THE EMBO JOURNAL

Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple negative breast cancer progression

Binbin Guo, Siqi Wu, Xun Zhu, Liyuan Zhang, Jieqiong Deng, Fang Li, Yirong Wang, Shenghua Zhang, Rui Wu, Jiachun Lu, Yifeng Zhou

Review timeline:

Submission date:	6th Apr 2019
Editorial Decision:	9th May 2019
Revision received:	26th Jul 2019
Editorial Decision:	1st Oct 2019
Revision received:	10th Oct 2019
Accepted:	15th Oct 2019

Editor: Daniel Klimmeck

Transaction Report:

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

1st Editorial Decision 9th May 2019

Thank you for the submission of your manuscript (EMBOJ-2019-102190) to The EMBO Journal. Your manuscript has been sent to three referees, and we have received reports from all of them, which I enclose below.

As you will see, the referees acknowledge the potential interest and novelty of your work, although they also express a number of issues that will have to be addressed before they can support publication of your manuscript in The EMBO Journal. While referee #1 is overall more positive and supportive of publication, referee #3 states that additional experiments are required to corroborate your findings on CIP2A competitive binding and the physiological relevance of the results (ref#2, pts. 2-5) This referee is also concerned about discrepancies with earlier literature (pt. 1) and potential indirect confounding effects (pt.8). Referee #3 points to unresolved selectivity of the effects and a number of controls required (ref#3, pts. 2,4). Further, the referees raise a number of issues related to accuracy and completeness of methods annotation, terminology and data illustration would need to be conclusively addressed to achieve the level of robustness needed for The EMBO Journal.

I judge the comments of the referees to be generally reasonable and given their overall interest, we are in principle happy to invite you to revise your manuscript experimentally to address the referees' comments.

REFEREE REPORTS:

Referee #1:

- General assessment
This manuscript presents a quite complete set of data illustrating how one long "non-coding" RNA

(lncRNA) encodes a regulatory micropeptide with crucial (patho)biological functions. This is an important point because the discovery that lncRNAs in fact encode peptides is recent and provides the scientific community with novel perspectives in the regulation of gene expression. The authors present the entire story from the screening of translationally regulated (by TGFbeta) lncRNAs to the description of the pathological function (in a triple-negative breast cancer setting) of one selected candidate including a solid biochemical and genetic characterization supporting the conclusions. Such candidate (named CIP2A-BP) is herein identified as a novel peptide regulating directly the activity of PP2A phosphatase by competition with its partner CIP2A. The authors provide many novel data expected to have a high fundamental impact on the EMBO's readership and more widely an impact on scientists working in the area of post-transcriptional (translational) regulation of gene expression in general and in the cancer field.
As such, I believe the data present sufficient scientific novelty and are technically sound to warrant publication in EMBO J without a long process of reviewing. I therefore suggest only a few minor revisions as listed below.

- Minor concerns
Title. Can the title incorporate the notion that the CIP2A-BP micropeptide emanates from translation of one lncRNA?

Figure 4. The data presented in this figure show an increase in the expression level of 4E-BP1 and in its binding to eIF4E next to TGFbeta treatment which likely explains the concomitant downregulation of LINC00665 lncRNA translation. Only hypophosphorylated 4E-BP1 is expected to inhibit translation. It would be therefore nice to test whether LINC00665 RNA translation is also sensitive to inhibition of 4E-BP1 phosphorylation for instance thanks to the use of mTOR inhibitors.

Figure 6. Through the regulation of PP2A activity, CIP2A-BP appears here as a modulator of the PI3K-AKT pathway. Because one downstream signaling cascade is mTOR/4E-BP1, one can anticipate that CIP2A-BP regulates also the phosphorylation status of 4E-BP1, and in turn its own translational expression. Is this the case? If so, what could be the best working model integrating all these parameters? These points could be discussed in the manuscript.

The manuscript contains typos, which should be corrected before publication. Ex:
- Page 3, line 18 "participates" should read "participate"
- Page 6 line 4 "ORF-GFPmut-GFPmut" should read "ORFmut-GFPmut"
- Page 10, line 9 "elongation" should read "initiation"
Etc...

Supplementary figures.
The legends to supplementary figures need careful proofreading as they contain many typos/errors. Some of them (far from being exhaustive) are listed below:
Supplementary Figure S2: lettering is wrong, A and B look inverted (see also reference to this figure in the main text).
Supplementary Figure S3 : What "respectively" stands for?
Supplementary Figure S4 : (A) What "respectively" stands for? The term "cohort" should read "cohorts"
Supplementary Figure S5: (A) Remove one "Lower".
Etc...

Referee #2:

In this manuscript the authors found that the lnc RNA, LINC00665, can be translated. They found that it can be translated in a cap-dependent manner and produce a small polypeptide termed CIP2A-BP. Its production is inhibited by TGFb, which induces EMT. The authors showed that low expression of CIP2A-BP is associated with poor survival of triple negative breast cancer patients. Its overexpression inhibited migration and invasion, whereas its ablation increased migration, invasion and metastasis. However, there are several concerns that should be addressed:

1. It has been shown that TGF-β promotes PP2A activation to dephosphorylate S6K and AKT

(Genes Dev, 2000. 14(24): p. 3093-101), which contradicts the results shown here.
2. PP2A-B56 α and γ have been shown to associate with dimerized CIP2A to stabilize CIP2A. However, the authors indicated that CIP2A-BP decreases CIP2A - B56 binding without affecting CIP2A stability. This should be further addressed.
3. Since the activity of 4EBP1 is determined by its phosphorylation, the authors should show the phosphorylation status of 4EBP1 in the immunoblots.
4. 4EBP1 should affect the translation of multiple mRNAs and should have a much more global effect.
5. Fig. 4 -The authors should show that endogenous 4EBP1 mRNA is elevated by TGFb and SMAD4.
6. Fig. 3 -The authors should quantify metastasis.
7. Fig. 4H- total eIF4E should be shown.
8. The interaction of CIP2A with other proteins could affect the phenotype. For example, the interaction with Filmin A could affect migration. The authors should exclude the possibility that the effect on migration, invasion, and metastasis is through Filmin A.
9. The authors should show the interaction of endogenous CIP2A with endogenous CIP2A-BP.
10. Fig. 6A- The authors should include TGFb treatment.
11. Fig. 6B- The effect of CIP2A-BP on PP2A activity is modest. Is that sufficient to inhibit Akt activity to the extent that CIP2A-BP affect metastasis?
12. Fig. 6F- The authors should show controls without MK2206 treatment.
13. The authors should determine the mechanisms by which Akt affects MMP2, MMP9, SNAIL and E-cadherin (protein stability, mRNA translation, transcription).
14. The authors should quantify all the panels in Fig. 7.
15. The authors should explain how the CIP2A-BP peptide enters the mammary gland tumor cells in vivo. It is also not clear when the peptide was injected in the mammary fat pad of MMTV-PyMT mice.

Referee #3:

This manuscript investigates the role of long non-coding RNA (which codes) in triple negative breast cancer and undercovers that it encodes a micro-peptide which functions to regulate CIP2A. This is an interesting manuscript and with changes, that I hope the authors should already have in the majority of cases, should be acceptable for publication.

Major points, which need adding to this manuscript.

1) More detailed materials and methods section (in supplemental information) should be included which outline how the data was generated needs to be included. For examples how was the ribosome profiling data analysed, currently unclear.

2) A: Analysis of linc00665 across the whole sucrose gradient +/- TGFb to show where the message migrates. B: Under control conditions puromycin release experiments should be conducted to show that ribosomes are actively translocating on this RNA. Puromycin should release the RNA into the lighter fractions if translated.

3) The ribosome profiling data should show which ORF is being translated I don't understand why they have mutated the different ORFs?

4) Why is 4E-BP1 being selective for repressing translation of this RNA, or is it, could they clarify. They need to do more mRNAs to see if this is selective (which can be coupled to experiments outlined in point 2A).

5) Statistics on the metastasis figures need including.

6) Figure 7 is confusing to me, why do they think injecting peptide is going to do anything? Is there any evidence that this peptide moves from cell to cell or is in circulation? Of course, small peptides might do this and in fact have been shown to but with this one is there any evidence?

7) If this micro-peptide is important I would expect it to be conserved, is it? Of course, could you have a look in humans but you should look more broadly.

Minor points

1) Introduction "Long chain noncoding RNAs" should be replaced with "Long non-coding RNAs"

2) Western blot in Fig 4L should be repeated.

3) They should comment of the role that this non-coding RNA that has been published Cong et al 2019 CDDis "Long non-coding RNA linc00665 promotes lung adenocarcinoma progression and functions as ceRNA to regulate AKR1B10-ERK signaling by sponging miR-98" and comments on the different in mechanism of action.

1st Revision - authors' response 26th Jul 2019

Referee #1:

- General assessment

This manuscript presents a quite complete set of data illustrating how one long "non-coding" RNA (lncRNA) encodes a regulatory micropeptide with crucial (patho)biological functions. This is an important point because the discovery that lncRNAs in fact encode peptides is recent and provides the scientific community with novel perspectives in the regulation of gene expression. The authors present the entire story from the screening of translationally regulated (by TGFbeta) lncRNAs to the description of the pathological function (in a triple-negative breast cancer setting) of one selected candidate including a solid biochemical and genetic characterization supporting the conclusions. Such candidate (named CIP2A-BP) is herein identified as a novel peptide regulating directly the activity of PP2A phosphatase by competition with its partner CIP2A. The authors provide many novel data expected to have a high fundamental impact on the EMBO's readership and more widely an impact on scientists working in the area of post-transcriptional (translational) regulation of gene expression in general and in the cancer field.

As such, I believe the data present sufficient scientific novelty and are technically sound to warrant publication in EMBO J without a long process of reviewing. I therefore suggest only a few minor revisions as listed below.

-Minor concerns

-Title. Can the title incorporate the notion that the CIP2A-BP micropeptide emanates from translation of one lncRNA?

Response : Thanks for the comment. We changed the title to "A micropeptide CIP2A-BP encoded by LINC00665 inhibits triple negative breast cancer progression".

-Figure 4. The data presented in this figure show an increase in the expression level of 4E-BP1 and in its binding to eIF4E next to TGFbeta treatment which likely explains the concomitant downregulation of LINC00665 lncRNA translation. Only hypophosphorylated 4E-BP1 is expected to inhibit translation. It would be therefore nice to test whether LINC00665 RNA translation is also sensitive to inhibition of 4E-BP1 phosphorylation for instance thanks to the use of mTOR inhibitors.

Response : Thanks for the comment and helpful suggestion. We have carefully thought and addressed your question with additional experiments. p-mTOR promotes phosphorylation of 4E-BP1, reduced binding of 4E-BP1 to eIF4E, which ultimately promotes translation. We determined the effect of mTOR inhibitor in triple negative breast cancer cells. Without TGF-β stimulation, mTOR inhibitor treatment significantly reduced 4E-BP1 phosphorylation and translation of micropeptide CIP2A-BP. With TGF-β stimulation, although mTOR inhibitor treatment reduced 4E-BP1 phosphorylation, mTOR inhibitor treatment had no

effect on translation of micropeptide CIP2A-BP. We have included this data in the manuscript (shown in Supplemental Figure 7D). Our data suggest that translation of micropeptide CIP2A-BP from LINC00665 is regulated by 4E-BP1 phosphorylation status, additional signaling pathways are involved in regulating translation of micropeptide CIP2A-BP under TGF-β.

-Figure 6. Through the regulation of PP2A activity, CIP2A-BP appears here as a modulator of the PI3K-AKT pathway. Because one downstream signaling cascade is mTOR/4E-BP1, one can anticipate that CIP2A-BP regulates also the phosphorylation status of 4E-BP1, and in turn its own translational expression. Is this the case? If so, what could be the best working model integrating all these parameters? These points could be discussed in the manuscript.

Response : Thanks for the comments and helpful suggestion. As explained in the previous comment, our additional data suggest that without TGF-β stimulation, mTOR inhibition reduced translation of micropeptide CIP2A-BP. But such effect was overcome by TGF-β stimulation. Normally, mTOR/4E-BP1 acts as downstream signaling cascade of PI3K-AKT pathway, this suggests that micropeptide CIP2A-BP could inhibit PI3K-AKT and its downstream mTOR/4E-BP1 signal pathway through CIP2A/PP2A pathway, thereby inhibit its own translation. This might be a negative feedback regulation mechanism evolved for homeostasis. However, in the microenvironment of advanced tumors, the presence of high levels of TGF-beta could lead to the activation of TGF-β/Smad signaling pathway, which promotes the expression of 4E-BP1 and inhibits the expression of CIP2A-BP. Under this scenario, even if the mTOR/4E-BP1 signaling pathway is activated, it could not increase translation of micropeptide CIP2A-BP. Based on

these data, we hypothesize that in triple negative breast cancer cells, TGF-β/Smad signaling pathway dominates over mTOR/4E-BP1 signaling pathway in regulation of the translation of micropeptide CIP2A-BP. Therefore, in triple negative breast cancer cells, micropeptide CIP2A-BP does not regulate its own translation and expression. We have included this discussion in the manuscript.

The manuscript contains typos, which should be corrected before publication. Ex:

- Page 3, line 18 "participates" should read "participate"
- Page 6 line 4 "ORF-GFPmut-GFPmut" should read "ORFmut-GFPmut"
- Page 10, line 9 "elongation" should read "initiation"

Etc...

Supplementary figures.

The legends to supplementary figures need careful proof-reading as they contain many typos/errors. Some of them (far from being exhaustive) are listed below:

Supplementary Figure S2: lettering is wrong, A and B look inverted (see also reference to this figure in the main text).

Supplementary Figure S3: What "respectively" stands for?

Supplementary Figure S4: (A) What "respectively" stands for? The term "cohort" should read "cohorts"

Supplementary Figure S5: (A) Remove one "Lower".

Etc...

Response : Thanks for the comments and helpful suggestion. We have performed the proof-reading and corrected all the typos/errors in the manuscript.

Supplementary Figure S3: "respectively" means "MCF-10A" and "MDA-MB-231"

cells. Supplementary Figure S4: "respectively" means "Suzhou" and "Guangzhou" cohorts.

Referee #2:

In this manuscript the authors found that the lnc RNA, LINC00665, can be translated. They found that it can be translated in a cap-dependent manner and produce a small polypeptide termed CIP2A-BP. Its production is inhibited by TGFb, which induces EMT. The authors showed that low expression of CIP2A-BP is associated with poor survival of triple negative breast cancer patients. Its overexpression inhibited migration and invasion, whereas its ablation increased migration, invasion and metastasis. However, there are several concerns that should be addressed:

-1. It has been shown that TGF-β promotes PP2A activation to dephosphorylate S6K and AKT (Genes Dev, 2000. 14(24): p. 3093-101), which contradicts the results shown here.

Response：Sorry for the confusion. In the cited paper (Genes Dev, 2000. 14(24): p. 3093-101), TGF-β inhibits S6K and AKT phosphorylation by acting on the Bα subunit of PP2A, leading to G1 phase arrest of epithelial cells, and inhibiting cell cycle progression. In our study, we found that TGF-β inhibits the activity of PP2A through downregulating translation of micropeptide CIP2A-BP, activating the PI3K/AKT pathway, and promoting tumor cell metastasis and invasion. We speculate following two mechanisms for this discrepancy: (1) the mammalian PP2A holoenzyme is a heterotrimer composed of structural subunit A, regulatory

subunit B and catalytic subunit C. Subunits A and C have two isomers (α and β) each, while regulatory B subunit has four families of isomers (B' or PPP2R2; B'' or PPP2R5; B' or PPP2R3; B' or PPP2R6), each containing several isomers. The activity and targets of PP2A holoenzyme are mainly determined by regulatory subunit B. Therefore, the PP2A holoenzyme could have different activities and targets based on different regulatory subunits B involved. The difference between our study and the cited study (Genes Dev, 2000. 14(24): p. 3093-101) could be due to different subunits involved in PP2A; (2) TGF-β can act both as a tumor suppressor gene and an oncogene during malignant transformation and tumor progression. On one hand, TGF-β can inhibit cell proliferation and induce apoptosis; on the other hand, TGF-β can promote tumor metastasis and angiogenesis. Such a dual role of TGF-β in tumorigenesis and tumor progression could explain the discrepancy between our study and the cited study. We cited this paper and discussed it in the manuscript.

-2. PP2A-B56 α and γ have been shown to associate with dimerized CIP2A to stabilize CIP2A. However, the authors indicated that CIP2A-BP decreases CIP2A - B56 binding without affecting CIP2A stability. This should be further addressed.

Response : Thanks for the comment and helpful suggestion. We determined the protein level of CIP2A in CIP2A-BP knockout and overexpression cells when translation was inhibited by cycloheximide D. Our results showed that the speed of CIP2A protein degradation was not affected by CIP2A-BP level (Supplemental Figure 8B). Our results support the hypothesis that CIP2A-BP interferes the interaction between PP2A-B56 and CIP2A but does not affect the stability of CIP2A.

-3. Since the activity of 4E-BP1 is determined by its phosphorylation, the authors should show the phosphorylation status of 4E-BP1 in the immunoblots.

Response : Thanks for the comment and helpful suggestion. We re-performed these experiments and added 4E-BP1 phosphorylation result in Figure 4 and Supplemental Figure 7D.

-4. 4E-BP1 should affect the translation of multiple mRNAs and should have a much more global effect.

Response : Thanks for the comment and helpful suggestion. We re-analyzed the original gene expression dataset (GSE59817) by bioinformatics and identified additional 429 protein-coding genes, whose translation but not transcription was reduced by TGF-β treatment (logFC<-1, FDR<0.05). Because TGF-β treatment upregulates the expression of 4E-BP1, we hypothesized that 4E-BP1 could regulate the translation of these genes. When we knocked-out 4E-BP1, we showed that the protein levels of CXADR, CABLES1, and SOCS3 were significantly upregulated (Supplemental Figure 7, A-C). Our data suggest that 4E-BP1 does have a global effect, influence the translation of multiple mRNAs.

-5. Fig. 4 -The authors should show that endogenous 4E-BP1 mRNA is elevated by TGFb and SMAD4.

Response : Thanks for the comment and helpful suggestion. We showed that 4E-BP1 mRNA level was increased by TGF-β treatment in triple-negative breast cancer cells. 4E-BP1 mRNA level was further increased by Smad4 overexpression but completely abolished by Smad4 knockdown. Smad4 expression level had no

effect on 4E-BP1 mRNA level without TGF-β stimulation (Supplemental Figure 6, A and B). Our data suggest that TGF-β stimulates 4E-BP1 transcription through Smad4.

-6. Fig. 3 -The authors should quantify metastasis.

Response : Thanks for the comment and helpful suggestion. We have quantified the number of metastatic nodules in the lung for Figure 3 and Figure 7.

-7. Fig. 4H- total eIF4E should be shown.

Response : Thanks for the comment and helpful suggestion. We re-performed these experiments and added total eIF4E in Figure 4H now.

-8. The interaction of CIP2A with other proteins could affect the phenotype. For example, the interaction with Filmin A could affect migration. The authors should exclude the possibility that the effect on migration, invasion, and metastasis is through Filmin A.

Response : Thanks for the comment and helpful suggestion. First, we used immunoprecipitation and western blot analysis to validate the high-confidence micropeptide CIP2A-BP binding candidate proteins detected by mass spectrometry. We showed that only CIP2A but not other proteins (FLNA, ANXA2, MYL6, CLTC and PFKP) bound to micropeptide CIP2A-BP (Figure 5B). Second, we did not detect interaction between the indicated proteins (FLNA, ANXA2, MYL6, CLTC and PFKP) and CIP2A using CIP2A antibody in immunoprecipitation and western blot analysis. Our data suggest that CIP2A-BP regulated cell migration, invasion and metastasis through CIP2A but not others.

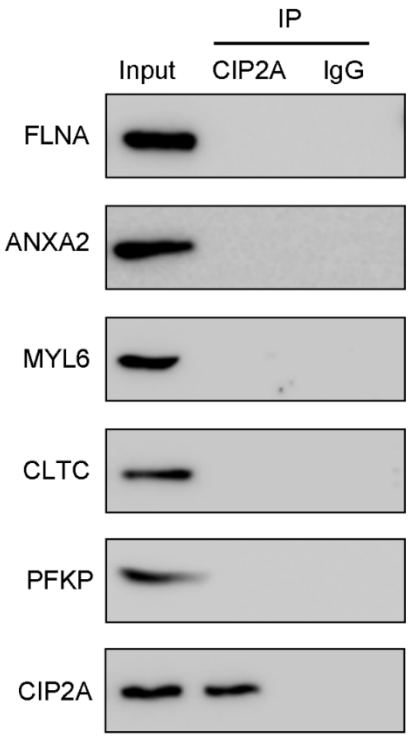

Figure Legends

Co-immunoprecipitation assays showed that there was no interaction between CIP2A and the indicated proteins.

-9. The authors should show the interaction of endogenous CIP2A with endogenous CIP2A-BP.

Response：Thanks for the comment and sorry for the confusion. First, we detected interaction between his-tagged CIP2A-BP and endogenous CIP2A by immunoprecipitation and western blot analysis using anti-His antibody (Figure 5B). Second, we detected interaction between CIP2A and endogenous CIP2A-BP by immunoprecipitation and Western blot using CIP2A antibody (Figure 5D). All these results have been incorporated in the revised manuscript.

-10. Fig. 6A- The authors should include TGFb treatment.

Response : Thanks for the comment and helpful suggestion. We have included TGF-β treatment data (Supplemental Figure 9A).

-11. Fig. 6B- The effect of CIP2A-BP on PP2A activity is modest. Is that sufficient to inhibit Akt activity to the extent that CIP2A-BP affect metastasis?

Response : Thanks for the thoughtful comment. First, our results indicated that CIP2A-BP significantly upregulated PP2A activity and AKT phosphorylation (fold change>2). Second, the level of AKT phosphorylation change also significantly upregulated AKT/NFκB signaling pathway and increased protein levels of MMP2, MMP9, and Snail. Finally, both in vitro and in vivo results indicated that these changes significantly affected invasion and metastasis of TNBC. We concluded that the effect of CIP2A-BP on PP2A is sufficient to inhibit AKT activity and further affects cell metastasis.

-12. Fig. 6F- The authors should show controls without MK2206 treatment.

Response : Thanks for the comment and helpful suggestion. We have included controls without MK2206 treatment (Supplemental Figure 11A).

-13. The authors should determine the mechanisms by which Akt affects MMP2, MMP9, SNAIL and E-cadherin (protein stability, mRNA translation, transcription).

Response : Thanks for the comment and helpful suggestion. Previous studies indicate that activation of PI3K/AKT/NF-kB signaling pathway leads to NF-kB releasing from NF-kB/IkBa complex to form p50/p65 dimer in the cytoplasm. Subsequently, p50/p65 dimer is transported to nucleus and p65 acts as a

transcription factor to regulate downstream target gene expression. In the current study, we found that p65 acted as a transcription factor for MMP2, MMP9 and SNAIL transcription by CHIP, dual luciferase reporter assays and qPCR. We also confirmed that transcription of MMP2, MMP9 and SNAIL was increased, after activation of PI3K/AKT/NF-kB signaling pathway (Supplemental Figure 10). Our data suggest that the PI3K/AKT/NF-kB signaling pathway affects transcription of MMP2, MMP9 and SNAIL. Finally, we showed that SNAIL affected the transcription of downstream gene E-cadherin using CHIP, dual luciferase reporter assays and qPCR experiments (Supplemental Figure 10, A and E).

-14. The authors should quantify all the panels in Fig. 7.

Response：Thanks for the comment and helpful suggestion. We have quantified all panels in Figure 7.

-15. The authors should explain how the CIP2A-BP peptide enters the mammary gland tumor cells in vivo. It is also not clear when the peptide was injected in the mammary fat pad of MMTV-PyMT mice.

Response：Thanks for the comment and sorry for the confusion. In the current study, the micropeptide CIP2A-BP was synthesized with a cell-penetrating peptide at the C-terminus, which improved the efficiency of cell entrance of CIP2A-BP. Using immunohistochemistry, we detected CIP2A-BP in mice breast tumor tissue after injection of micropeptide CIP2A-BP into mice breast fat pad (Supplemental Figure 11, B and C). In the MMTV-PyMT mice model experiment, we first injected the micropeptide CIP2A-BP through the mammary fat pad when MMTV-PyMT mice were 8 weeks old, then injected CIP2A-BP once every week for five

weeks through the mammary fat pad. We have included these details in the Materials and Methods section.

-Referee #3:

-This manuscript investigates the role of long non-coding RNA (which codes) in triple negative breast cancer and undercovers that it encodes a micro-peptide, which functions to regulate CIP2A. This is an interesting manuscript and with changes, that I hope the authors should already have in the majority of cases, should be acceptable for publication.

-Major points, which need adding to this manuscript.

-1) More detailed materials and methods section (in supplemental information) should be included which outlines how the data was generated needs to be included. For examples how was the ribosome profiling data analysed, currently unclear.

Response：Thanks for the comment and sorry for the confusion. We have included detailed materials and methods section in the “Supplemental Materials and Methods”. For example, the ribosome profiling and RNA sequencing data of TGF-β-treated MCF10A cells (GSE59817) were obtained from GEO database. For ribosome profiling data, the raw reads were preprocessed by cutadapt software, then the filtered reads were aligned to Genome Reference Consortium Human Build 37 (GRCh37) using Tophat2 algorithm. After alignment to the human genome, the gene expression levels were calculated by Cufflinks, and the

differentially expressed genes were calculated by Cuffdiff. Finally, differentially expressed genes were defined as those genes with FDR<0.05. For RNA sequencing data, the reads were aligned to Genome Reference Consortium Human Build 37 (GRCh37) using Tophat2 algorithm. After alignment to the human genome, the gene expression levels were calculated by Cufflinks, and the differentially expressed genes were calculated by Cuffdiff. Finally, differentially expressed genes were defined as those genes with FDR<0.05.

-2) A: Analysis of linc00665 across the whole sucrose gradient +/- TGFb to show where the message migrates. B: Under control conditions puromycin release experiments should be conducted to show that ribosomes are actively translocating on this RNA. Puromycin should release the RNA into the lighter fractions if translated.

Response：Thanks for the comment and helpful suggestion. First, we analyzed the distribution of LINC00665 in cells with or without TGF-b treatment by sucrose gradient and qPCR. We showed that after TGF-β treatment significantly reduced the distribution of LINC00665 in 40-80s and polysome fractions (Supplemental Figure 1C). Second puromycin treatment reduced distribution of LINC00665 in the polysome group, but increased distribution in the 40-80S group (Figure 1D). Our results suggest that TGF-β inhibited translation of LINC00665 through initiation of ribosome binding.

-3) The ribosome profiling data should show which ORF is being translated I don't understand why they have mutated the different ORFs?

Response：Thanks for the comment and sorry for the confusion. Although the ribosome profiling data indicated that only ORF1 was translated, sequence analysis of LINC00665 suggested the presence of four potential ORFs (Figure 1A). Therefore, we used in vitro experiments to confirm that only ORF1 could translate into a micropeptide (Figure 1D). In Figure 1E, GFPmut was used as negative control, GFPmut and ORFmut-GFPmut were used as negative controls to confirm that GFP signal was derived from CIP2A-BP-GFP fusion protein. In Figure 1H, ORFmut-His was used to further confirm the necessity of LINC00665 ORF1 start codon for translation. All these mutated ORFs were used to confirm that LINC00665 ORF1 was translated from its own start codon.

4) Why is 4E-BP1 being selective for repressing translation of this RNA, or is it, could they clarify. They need to do more mRNAs to see if this is selective (which can be coupled to experiments outlined in point 2A).

Response：Thanks for the comment and helpful suggestion. First, we found that the level of micropeptide CIP2A-BP was significantly increased after 4E-BP1 knockdown. In 4E-BP1 knockdown cells, TGF-β treatment did not affect the level of CIP2A-BP, suggesting that 4E-BP1 is essential for the translation of micropeptide CIP2A-BP. Second, by analyzing the original data of this study (GSE59817), we identified 429 genes whose translation but not transcription was decreased significantly by TGF-β treatment (logFC<-1, FDR<0.05). We selected three genes (CXADR, CABLES1, and SOCS3) and determined their mRNA distribution by sucrose gradient after TGF-β treatment. Our results showed that the fraction of CXADR, CABLES1 and SOCS3 mRNA in 40-80S and polysome groups decreased significantly after TGF-β treatment (Supplemental Figure 7, A-

C). Previous studies showed that TGF-β could significantly upregulate the expression level of 4E-BP1, therefore our data suggest that 4E-BP1 could inhibit the translation of multiple genes.

-5) Statistics on the metastasis figures need including.

Response：Thanks for the comment and helpful suggestion. We have included statistics on Figure 3 and Figure 7.

-6) Figure 7 is confusing to me, why do they think injecting peptide is going to do anything? Is there any evidence that this peptide moves from cell to cell or is in circulation. Of course, small peptides might do this and in fact have been shown to but with this one is there any evidence?

Response：Thanks for the comment and sorry for the confusion. In the current study, the micropeptide CIP2A-BP was synthesized with a cell-penetrating peptide at the C-terminus, which improved the efficiency of cell entrance of CIP2A-BP. Using immunohistochemistry, we detected CIP2A-BP in mice breast tumor tissue after injection of micropeptide CIP2A-BP into mice breast fat pad (Supplemental Figure 11B). Simultaneously, we detected the micropeptide in the serum of mice injected via tail vein (Supplemental Figure 11C). Our results suggest that micropeptides could enter circulation.

-7) If this micro-peptide is important I would expect it to be conserved, is it? Of course, could you have a look in humans but you should look more broadly.

Response：Thanks for the comment. This micropeptide is also present in rhesus monkeys.

-Minor points

-1) Introduction "Long chain noncoding RNAs" should be replaced with "Long non-coding RNAs"

Response：Thanks for the comment and helpful suggestion. We have changed "Long chain noncoding RNAs" to "Long non-coding RNAs" in the manuscript.

-2) Western blot in Fig 4L should be repeated.

Response：Thanks for the comment and helpful suggestion. We have repeated the WB experiment shown in figure 4 L in the revised manuscript.

-3) They should comment of the role that this non-coding RNA that has been published Cong et al 2019 CDD is "Long non-coding RNA linc00665 promotes lung adenocarcinoma progression and functions as ceRNA to regulate AKR1B10-ERK signaling by sponging miR-98" and comments on the different in mechanism of action.

Response：Thanks for the comment. Cong et al. showed that LINC00665 transcript acts as an oncogene in lung adenocarcinoma progression, while our study suggested that the micropeptide encoded by LINC00665 acts as a tumor suppressor in triple negative breast cancer. This might be due to tissue specificity of this lncRNA. Future studies are needed to investigate whether this lncRNA could also be translated in other types of tumor.

2nd Editorial Decision 1st Oct 2019

Thank you for submitting your revised manuscript for consideration by The EMBO Journal. Please accept my sincere apologies for the unusual delay in getting back to you. Your revised study was sent back to one of the referees for re-evaluation. As you will see the referee finds that his-her concerns have been sufficiently addressed and is now broadly in favour of publication. Please note that we have editorially considered your response to the other referees and concluded that they have been addressed satisfactorily.

Thus, we are pleased to inform you that your manuscript has been accepted in principle for publication in The EMBO Journal, pending some minor issues related to formatting and data representation as listed below, which need to be adjusted at re-submission.

REFEREE REPORTS:

Referee #3:

The authors have addressed my comments.

2nd Revision - authors' response 10th Oct 2019

The authors performed all the requested changes.

3rd Editorial Decision 15th Oct 2019

Thank you for submitting the revised version of your manuscript. I have now evaluated your amended manuscript and concluded that the remaining minor concerns have been sufficiently addressed.

Thus, I am pleased to inform you that your manuscript has been accepted for publication in the EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Yifeng Zhou
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2019-102190R

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.

USEFUL LINKS FOR COMPLETING THIS FORM

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

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-tun

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

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?	Sample size was chosen based on the need for statistical power. Reported in Materials and Methods, "Statistical analysis" subsection.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Reported in Materials and Methods, "Animal breeding and treatments" subsection.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	All patients in this study met the following inclusion criteria: the resected nodules were identified by pathological examination, no anti-cancer treatments were given before surgery.
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.	Yes. Reported in Materials and Methods, "Animal breeding and treatments" subsection.
For animal studies, include a statement about randomization even if no randomization was used.	For experiments involving mice the allocation of animals to different experimental treatments was randomized.
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.	We have included both Suzhou cohort (training data set) and Guangzhou cohort (validation data set) to minimize the effects of subjective bias.
4.b. For animal studies, include a statement about blinding even if no blinding was done	The investigators were blinded to the group allocation.
5. For every figure, are statistical tests justified as appropriate?	Yes.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes.

Is there an estimate of variation within each group of data?	Yes.
Is the variance similar between the groups that are being statistically compared?	Yes.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Reported in Appendix table S3.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Reported in Materials and Methods, "Cell culture and treatments" subsection. These cell lines were all characterized by DNA finger printing analysis and passaged less than 6 months in this study. Cells were tested for mycoplasma prior to use for experiments.

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	Reported in Materials and Methods, "Animal breeding and treatments" subsection.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	Reported in Materials and Methods, "Study approval" subsection.
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.	We confirm compliance with the ARRIVE guidelines.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	Reported in Materials and Methods, "Study approval" subsection.
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.	The study was approved by the Ethical Committee of the Soochow University and Guangzhou Medical University, and conducted according to Declaration of Helsinki principles.
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Yes, we have followed the REMARK guideline.

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The ribosome profiling and RNA sequencing data of TGF-β-treated MCF10A cells were obtained from Gene Expression Omnibus DateSets, and accession ID is GSE59817.We provided references to our datasets in the Data Availability section at the end of our manuscript.
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).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as 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.	NA

G- Dual use research of concern

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