EMBO reports

CBX4 contributes to HIV-1 latency by forming phase-separated nuclear bodies and SUMOylating EZH2

Liyang Wu, Ting Pan, Mo Zhou, Tao Chen, Shiyu Wu, Xi Lv, Jun Liu, Fei Yu, Yuanjun Guan, Bingfeng Liu, Wanying Zhang, Xiaohui Deng, Qianyu Chen, Anqi Liang, Yingtong Lin, Lilin Wang, Xiaoping Tang, Weiping Cai, Linghua Li, Xin He, Hui Zhang, and Xiancai Ma
DOI: 10.15252/embr.202153855

Corresponding author(s): Xiancai Ma (maxc6@mail.sysu.edu.cn) , Hui Zhang (zhangh92@mail.sysu.edu.cn)

Review Timeline:

Submission Date:	21st Aug 21
Editorial Decision:	4th Oct 21
Revision Received:	31st Jan 22
Editorial Decision:	1st Apr 22
Revision Received:	17th Apr 22
Accepted:	18th May 22

Editor: Martina Rembold

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.)

1st Editorial Decision **4th Oct 2021**

Dear Dr. Ma

Thank you for the submission of your research manuscript to our journal. We have now received the full set of referee reports that is copied below.

As you will see, while the referees acknowledge that the findings are potentially interesting, they all point out that significant revisions are required and that the data need to be strengthened before the study can be considered for publication here.

From these comments it is clear that publication of the manuscript in our journal cannot be considered at this stage. On the other hand, given the potential interest of your findings, I would like to give you the opportunity to address the concerns and would be willing to consider a revised manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Basically, all points raised by the referees need to be addressed. In particular, it will be essential to substantiate the proposed LLPS for endogenous CBX4 and to exclude that it results from an overexpression artefact. Also the proposed effect of phase separated CBX4 on EZH2 SUMOylation and HIV-1 repression as well as the relevance to HIV-1 latency need to be supported by further data. All missing control experiments, quantifications and repeats need to be provided.

Should you decide to embark on such a revision, acceptance of the 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 or rejection 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. I realize that the proposed experiments will likely take longer than 3 months and I can extend the revision duration to 5 months. If you decide to revise your manuscript for potential publication in EMBO Reports, I suggest to contact me so that we can discuss the revisions and their timeframe further.

***IMPORTANT NOTE: we perform an initial quality control of all revised manuscripts before re-review. Your manuscript will FAIL this control and the handling will be DELAYED if the following APPLIES:

1) A data availability section is missing.
2) Your manuscript contains error bars based on n=2. Please use scatter blots showing the individual datapoints in these cases. The use of statistical tests needs to be justified.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision.***

When submitting your revised manuscript, we will require:

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).
Please download our Figure Preparation Guidelines (figure preparation pdf) from our Author Guidelines pages https://www.embopress.org/page/journal/14693178/authorguide for more info on how to prepare your figures.

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 (). Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

5) 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
()

6) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc. See detailed instructions regarding expanded view here:

- Additional Tables/Datasets should be labeled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

7) Before submitting your revision, primary datasets (and computer code, where appropriate) produced in this study need to be deposited in an appropriate public database (see < https://www.embopress.org/page/journal/14693178/authorguide#dataavailability>).
- Cbx4 proteomics
- ATAC-seq

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 & Method) that follows the model below (see also < https://www.embopress.org/page/journal/14693178/authorguide#dataavailability>). Please note that the Data Availability Section is restricted to new primary data that are part of this study.

Data availability

The datasets (and computer code) 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. ***

8) Figure legends and data quantification:
The following points must be specified in each figure legend:
- the name of the statistical test used to generate error bars and P values,
- the number (n) of independent experiments (please specify technical or biological replicates) underlying each data point,
- the nature of the bars and error bars (s.d., s.e.m.)
- If the data are obtained from n {less than or equal to} 2, use scatter blots showing the individual data points.
Discussion of statistical methodology can be reported in the materials and methods section, but figure legends should contain a basic description of n, P and the test applied.
See also the guidelines for figure legend preparation:
https://www.embopress.org/page/journal/14693178/authorguide#figureformat
- Please also include scale bars in all microscopy images.

9) We would also encourage you to include the source data for figure panels that show essential data. Numerical data should be provided as individual .xls or .csv files (including a tab describing the data). For blots or microscopy, uncropped images should be submitted (using a zip archive if multiple images need to be supplied for one panel). Additional information on source data and instruction on how to label the files are available .

10) 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 .

11) As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. 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.

You are able to opt out of this by letting the editorial office know (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."

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

I look forward to seeing a revised form of your manuscript when it is ready. Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Please let me know if you have questions or comments regarding the revision.

Yours sincerely,

Martina Rembold, PhD
Senior Editor
EMBO reports

****

Referee #1:

Wu et al. have provided insight into Polycomb group (PcG) protein CBX4 mediated HIV-1 latency. The authors concluded CBX4 forms liquid droplet-like nuclear bodies on the HIV-1 LTR and CBX4 recruits EZH2 to CBX4 bodies and SUMOylates EZH2 utilizing its SUMO E3 ligase activity, which enhances the H3K27 methyltransferase activity of EZH2.
Overall, the manuscript is clearly written and the data are interesting. However, the data does not entirely support the conclusions. Some conclusions were overstated and premature. More data, especially in biochemical experiments, are necessary in order to validate the proposed mechanism. Therefore, I recommend the following suggestions for the authors to strengthen their manuscript.

◯Authors indicated that CBX4 forms phase-separated nuclear condensates which co-localize with HIV-1 proviruses and its activity is important for inactivation of HIV expression. However, CBX4 mediated phase-separated nuclear condensates were observed only in cells exogenously expressing CBX4. Authors does not show the ratio between endogenous and exogenous CBX4 protein level. Therefore, Authors cannot exclude the possibility CBX4 mediated phase-separated nuclear condensates is an artifact. Authors should show the endogenous CBX4 mediated phase-separated nuclear condensates.
In addition, SUMOlylated-EZH2 was observed only in the CBX4 over expressing cells.

◯Authors did not show KD efficiency of CBX4 at protein level. In order to accurately assess the function of CBX4, authors should perform western blot analysis using cells treated with shCBX4 and siCBX4.(Fig.1C-K, 5GH, 6, 7AB (7C-E, CBX4 mRNA should be analyzed)

◯For all the in vitro experiments using recombinant proteins or proteins purified from human cells, purity of proteins have to be analyzed by SDS-PAGE. Without quality checking, all in vitro assay are unreliable.

◯Figure 2A: Authors indicates CBX4 accumulation on HIV LTR. However, to be sure the ChIP signal of CBX4, CBX4 ChIP in CBX4 kd cells should be performed.

◯Figure3-E
To consolidate the conclusion, quantitative data on co-localisation (CBX4 and HIV) is necessary.

◯Figure4A
The proteins in CBX4 LLPS (EZH2, RING1B and SUMO4) should be able to internally diffused and diffuse across boundary. So, FRAP experiments for EZH2, RING1B and SUMO4 in CBX4 LLPS are required.

◯Figure4J-P
Does CBX4mut obtain the ability interact with EZH2? IP experiments for CBX4 and EZH2 like Fig EV2F can reveal this point. For direct evidence CBX4mut lacks the ability forming LLPS, droplet formation assay using recombinant CBX4mut (shown in Fig EV3G using wt CBX4) is needed.
In addition, CBX4mut possibly lacks the ability to recruit EZH2 on HIV LTR. So ChIP for EZH2 in cells expressing CBX4mut will support authors conclusion.

◯Figure6B-E
quantitative data is necessary.

◯Figure6E
Authors suggested CBX4-mediated EZH2 SUMOylation enhanced the methyltransferase activity of EZH2, resulting in elevated

H3K27me3. However, Fig 6 is not sufficient to draw this conclusion. The reasons are below.
-EZH2 mediated H3K27me3 methylation activity is quite low. In this experiment. As compared to lane1 with lane2 or lane3, there are no difference.
-Histone H3.1 is soluble? Reconstituted nucleosome or at least H3-H4 tetramer should be used in this assay.
-Moreover, to carry out such a complex system involving such a large number of proteins, it is necessary to use highly purified factors as recombinant proteins. Authors can not exclude the effect of contamination, therefore it is difficult to draw any conclusions from the authors' experiments.
-Efficiency of sumolylation of EZH2 in this experiment condition should be measured.
-Purity of each proteins should be analyzed.
-as control, experiments using EZH2 catalytic mutant and H3 K27R mutant are required.

◯Figure7A-B
Authors should access H3K27me3 and H2A119ub accumulation by ChIP in HIV infected primary CD4T cells with shCBX4.

◯Figure EV2F
Cell lysates for IP are treated with DNase and RNase? Interaction is dependent on DNA or RNA?

◯Figure EV4
Authors cannot conclude because there are a lot of non-specific bands. SUMOlylated CBX4 ladder is not clear.
For correct conclusion, the experiments using highly purified recombinant proteins are required.

Referee #2:

This study focusses on the function of CBX4 in transcriptional repression, specifically that of latent HIV-1. They report that CBX4 interacts with and SUMOylates EZH2 showing a connection between the PRC1 and PCR2 complexes. They claim this interaction occurs in a CBX4-driven biomolecular condensate and that this interaction is important for HIV-1 transcriptional repression, which maintains HIV-1 latency.

The interaction of CBX4 and EZH2 in promoting HIV-1 latency through a biomolecular condensate is interesting. The authors note that the condensate step could be a new target for intervention and therefore this study shows novelty and impact. However, the scientific support for the condensate mechanism and the importance of EZH2 SUMOylation for HIV-1 repression is insufficient.

A general weakness of the manuscript is the lack of replicates. Of many important experiments only single replicates are shown, making it impossible to evaluate the validity of the results. I have noted specific cases throughout my comments.

The evidence for CBX4 condensation is exclusively supported by over-expression studies. Since condensation can be driven by artificially high protein concentrations these studies are inconclusive without confirmation at physiological protein concentrations, through direct labeling or artificial expression at physiological levels. The single endogenous-level dataset (Fig 3B-D) is promising but insufficient on its own. Importantly, the major finding that EZH2 enters CBX4 condensates while over-expressed can be exclusively concentration-mediated and therefore an artifact.

Currently, several observations concerning CBX4 foci are inconsistent with a condensate model:

Fig4A: No replicates are shown for this data, at least 3+ biological replicates are necessary to gage the variability. The shown FRAP data is not consistent with LLPS. Almost no recovery is observed, ever after 10 minutes. This is significantly slower than previous reports of condensates. These kinetics are more consistent with aggregates, and the incomplete recovery indicates a large (50%) immobile fraction. These results argue against LLPS. Note that hardening of condensates can also be an effect of high protein concentrations in an artificial over-expression system.

Fig 4B The fused foci do not relax into a sphere. Again, directly arguing against LLPS. The authors refer to Alberti, Gladfelter and Mittag, but do not interpret the results consistent with the guidelines in this paper.

Fig 4B,C Both B and C show "mottled" pattern which is unexpected, because highly dynamic LLPS proteins would equally distribute in the dense phase. Do the authors have an interpretation of what is happening here?

Fig EV3H, does not show increasing droplet size with the highest concentration, completely counter to what the authors claim in the text. There might be more droplets in the higher concentration regime, but the authors only show a single image and do not perform quantification of droplet numbers or size. At least quantification of several images is necessary to support (or disqualify)

the claims made in the text.

Fig EV3I, J authors claim hundreds of nuclear puncta, but perform no quantification and show only single replicates

As a second major point the authors claim that the condensation of CBX4 with EZH2 is important for EZH2 SUMOylation and repression of HIV-1. However, the presented data does not support this point.

Fig 6B. While expression of CBX4 is shown to increase EZH2 SUMOylation, CBX4 mediated increase of EZH2 SUMOylation is NOT dependent on its capacity to form foci. Contrary to the claim of the authors in the text, in this panel there is no reduction in EZH2 SUMOylation (compare last two lanes). This would indicate that the ability to form foci is irrelevant for EZH2 SUMOylation.

Fig 6D. In this figure the authors claim to tie the increased SUMOylation of EZH2 to an increased level of H3K27Me3, which in turn could indicate more efficient repression of HIV-1. However, there is no clear difference between the two conditions and the bands are not quantified. As presented here, the data shows the opposite of what the authors claim in the text.

Fig 6E. In this figure the authors do show an increased level of H3K27Me3, however, here this is caused by the co-expression of CBX4. While this is presumably due to an increase of SUMOylation of EZH2, there are no bands of SUMO-EZH2 visible in this blot. Does CBX4 not induce EZH2 SUMOylation here? That would run counter to the model proposed by the authors.

Minor comments:

Fig 1 F-I Shows very similar data in different systems and therefore can be supplemental

Fig 2B, 2C, Unclear what amplicon is used here, please indicate the tested region.

Fig 2E, H3K27Ac does not change, why is this not further discussed/explained?

Fig 3E Only a single replicate is shown more replicates and quantification are needed to interpret these results.

Fig 4D concentration of hexanediol not mentioned

Fig EV3C-F Does not test what the authors claim. These experiments indicate that the observed foci are insensitive to osmotic stress, but do not inform about physiological conditions, since none of the tested regimes are physiological.

Referee #3:

CBX4 contributes to HIV-1 latency by forming phase-separated
2 nuclear bodies and SUMOylating EZH2

This work is based on the previous findings that Polycomb complexes PRC1 and PRC2 mediate HIV-1 latency by establishing and maintaining the heterochromatin environment through H3K27me3 at the viral promoter. The authors therefore screen for Polycomb specific factors and discover that CBX4 of PRC1 and EZH2 of PRC2 restrict HIV-1 transcription in TZM-bl cells. They then focus on CBX4 mainly, and investigate its relevance for HIV-1 transcription and latency in several latency models and HIV-1 patient derived CD4+ T cells.
While there are some interesting aspects of the manuscript, I think that many of the approaches need to be deepend and better clarified.

Although they talk about transcriptional latency mediated by CBX4 they start by measuring GFP fluorescence in FACS analysis and never measure the levels of viral transcripts in the cells. They perform ChIP and ATAC-Seq without providing adequate controls and explaining in details ATAC-seq. Of note, although apparently simple to perform , ATAC-seq analysis is rather complex and requires a better explanation (not just two sentences in the paragraph). As it is right now, leaves a reader with an idea that it is superfluous.

Moreover, the Authors suggest that CBX4 forms nuclear bodies with LLPS properties on the HIV-1 LTR, a claim based on ImmunoDNA FISH which shows no quantification or specific. .
The only somewhat solid data are in vitro SUMOylation assays, suggesting that CBX4 recruits EZH2 utilizing its SUMO E3

ligase activity. However, the fact that CBX4 sumoylates EZH2 is not really connected to HIV-1 latency. The authors state that CBX4 and EZH2. act synergistically to maintaining HIV-1 latency, but do not actually demonstrate synergism.

Major points:
• Introduction is not original with respect to the Authors previous publication ' Histone chaperone CAF-1 promotes HIV-1 latency by leading the formation of phase-separated suppressive nuclear bodies'
• 128-131. Language used to build the hypothesis is very trivial, and unclear
• Treatment controls are generally missing, including efficiencies of KD, cell viability, ChIP controls for cellular regions (good practice shows positive and negative regions).

More specific comments are listed bellow:

Figure 1. CBX4 contributes to HIV-1 latency. TZM-bl data are not convincing, 3 fold change in luc activity with respect to mock ie, leaky LTR-luc expression is not reactivation. Additional controls should be included, like, Tat transfection, PHA/TPA treatment. This system is not adequate to study latency and should be avoided. This is a system to study (at a very basal level) HIV-1 transcription.
Furthermore, in A) SiRNA KD efficiency for PRC1 and PRC2 complex is missing unless available online, access not provided to the Reviewer; although stated (Expanded View for this article is available online).
In B) there is an issue with data normalization. Shouldn't it be siNT in all reconstitution plasmid points at 1?
The reactivation of all latent models, except maybe for the JLat mix is very modest upon KD of CBX4, only few precents in all presented backgrounds.
171-172. Why did the authors opt for JQ1 and SAHA specifically? Is there a scientific rationale behind these compounds? They are just mentioned as 'other latency-reversing agents (LRAs) including HDAC inhibitor SAHA and BRD4 inhibitor JQ-1'

Figure 2. CBX4 contributes to H3K27me3 modification of HIV-1 promoter.
ChIP was performed first in TZM-bl cells, not an adequate model for HIV-1 chromatin binding as only a small portion of the viral genome is present. Moreover, enrichment of the main factor CBX4 is expressed over IgG. Could the authors show as percentage of input, and also add the specificity of the IP, validation for the CBX antibody
Validation ChIP in JLat clone for the CBX absence on LTR upon KD is missing, and should be included.
Authors should clarify ATAC-Seq analysis, specifically how the normalization was performed, if any, aside by normalizing to the total mapped reads

Figure 3. CBX4 recruits PcG proteins and forms nuclear bodies.
The authors overexpress CBX and perform Mass spec comparing it with control empty vector. They identify ~700 interacting proteins of CBX among which EZH2 and SUMO4. These proteins were then overexpressed in HEK293 with RFP tags. It is clear from the images that there is a colocalization between CBX4 and EZH2 as well as between CBX4 and SUMO4. However, the size of the colocalization bodies is significantly different, thus putting a question mark on the relevance of the overexpression(s). It also rises a possibility that many partners of CBX4 lead to the formation of heterogeneous CBX4 bodies, as can be inferred also from their various sizes. However, this further implies that it is very hard to understand which one of these bodies is responsible for HIV-1 silencing, and questions the ImmunoFiSH data which can be completely random, as there are way too many objects (bodies) and only one FISH signal. There are no control set objects and no quantifications either (Fig 3E and EV2A)

Figure 4. CBX4 bodies are phase-separated condensates B) 1,6 hexandiol co -staining for chromatin and viability should be included.
EV3 is referring to this figure: There is a major mislabelling of the concentration-dependent LLPS formation, should be the other way around.
Labelling is also random; the Authors should clarify construct labelling.

Figure 5. CBX4 SUMOylates EZH2
Panel A) is not informative at all, as there is no any labelling included

Figure 6. LLPS-deficient CBX4 bodies are unable to suppress HIV-1 and SUMOylate EZH2
Panel A) Luciferase change is modest, again. Panel B) SUMO4 WB should be repeated, as signal is not well resolved, membrane seems burnt.
Figure 7. CBX4 depletion reactivates latent HIV-1 in cells from HIV-1-infected individuals.
The authors should show some quantification of viral mRNA, and not only use GFP as a measure for HIV-1 latency. It is only in FigC, D and E that they quantify HIV-1 mRNA through Vif (why Vif?)
Also as the cells were kept for prolonged time in culture, but no cell viability is shown. How did shCBX4 affect cell viability together with viral infection, as compared to the shControl?

Patient data are not convincing. As RNA levels are only doubled, and should be practically absent on ART, so doubling is only modest effect in this context. Raw dPCR data can clarify this. Or absolute quantification via ddPCR.

• Scientific and general English language is should be thoroughly corrected.

Minor points:
59. To what strategies do Authors refer saying 'and many others'
62-66. Sentence is completely unclear, both conceptually and semantically
68 . What do Authors mean by 'verified latency-reversing agents'
75-78. this sentence should be re-written. It is unclear what 'lies' refers to
79. what are epigenetic maintainers
95-96. preposition, for example in is missing
108. The tense of the sentence is incorrect
125. preposition is missing, for example into
128. 'less defined' should state not defined
133. 'many cellular models' which one, how many. This is a scientific article, and if demonstrated should be stated accordingly, quantified.

Point-by-point response to each comment:

Referee #1:

Wu et al. have provided insight into Polycomb group (PcG) protein CBX4 mediated HIV-1 latency. The authors concluded CBX4 forms liquid droplet-like nuclear bodies on the HIV-1 LTR and CBX4 recruits EZH2 to CBX4 bodies and SUMOylates EZH2 utilizing its SUMO E3 ligase activity, which enhances the H3K27 methyltransferase activity of EZH2.
Reply: We appreciate for reviewer's comprehensive summaries.

Overall, the manuscript is clearly written and the data are interesting. However, the data does not entirely support the conclusions. Some conclusions were overstated and premature. More data, especially in biochemical experiments, are necessary in order to validate the proposed mechanism. Therefore, I recommend the following suggestions for the authors to strengthen their manuscript.
Reply: We appreciate for reviewer's positive feedback and giving us the following insightful suggestions. We have provided more data to fully validate our proposed mechanisms.

◯*Authors indicated that CBX4 forms phase-separated nuclear condensates which co-localize with HIV-1 proviruses and its activity is important for inactivation of HIV expression. However, CBX4 mediated phase-separated nuclear condensates were observed only in cells exogenously expressing CBX4. Authors does not show the ratio between endogenous and exogenous CBX4 protein level. Therefore, Authors cannot exclude the possibility CBX4 mediated phase-separated nuclear condensates is an artifact. Authors should show the endogenous CBX4 mediated phase-separated nuclear condensates.*
In addition, SUMOlylated-EZH2 was observed only in the CBX4 over expressing cells.
Reply: We thank the reviewer for pointing out these deficiencies of our study. If we understand clearly, the reviewer gave us three kind suggestions in this comment. We would like to response to these suggestions one-by-one.

Firstly, the reviewer suggested us to show the ratio between endogenous and exogenous CBX4 protein levels. Thus, we conducted a new experiment to indicate that upon expressing equal amounts of GFP-CBX4 proteins (compared with the endogenous CBX4) in HEK293T cells, CF568-conjugated CBX4 proteins were co-localized with exogenous GFP-CBX4 proteins. The expression ratio of endogenous and exogenous CBX4 proteins has been evaluated by western blot, the result of which was 1.00 : 1.05. These results have been shown in **newly-added Appendix Figure S3F and G**. Besides, to exclude the possibility that CBX4 condensation might be caused by CBX4 overexpression, we knocked down the endogenous CBX4 by siRNA targeting 3'UTR of *CBX4* mRNA, followed by expressing equal amounts of GFP-CBX4 proteins. GFP-CBX4 proteins formed similar nuclear puncta as those in the overexpression system. The expression ratio of endogenous CBX4 and GFP-CBX4 was 1.00 : 1.04, which was evaluated by western blot. These results have been shown in **newly-added Appendix Figure S3D and E**.

Secondly, the reviewer suggested us to show the endogenous CBX4-mediated condensation. Actually, all the data showed in **Figure 3B and 3E, Figure EV2A and 2B, Appendix Figure S3A-C** were derived from endogenous CBX4 proteins, which showed that the endogenous CBX4 formed nuclear condensates. During the revision, we also tried to construct a GFP knock-in (before *CBX4* CDS) cell line which could be used to conduct live cell imaging in physiological conditions. However, we tried for three month. We were unable to knock in *GFP* ORF before *CBX4* CDS where contained extremely high percentages of "G" and "C" bases.

Thirdly, the reviewer also suggested us to show the endogenous EZH2 SUMOylation. As the SUMOylated proteins is only a small proportion (less than 5%) within total target proteins (Gareau and Lima, 2010, PMID: 21102611; Impens *et al.*, 2014, PMID: 25114211), it is very hard to observe SUMOylated proteins in total samples. Thus, we isolated primary $CD4^+$ T cells from three healthy donors and conducted immunoprecipitation (IP) assays to enrich EZH2 proteins. Besides, we compared the expression of EZH2 and SUMOylated EZH2 in both siNC and siCBX4 samples. We found that SUMOylated EZH2 was observed in IP samples. Upon knocking down endogenous CBX4, the SUMOylated EZH2 bands (higher than wildtype EZH2 bands) disappeared. These results have been shown in **newly-added Figure EV5D**.

◯*Authors did not show KD efficiency of CBX4 at protein level. In order to accurately assess the function of CBX4, authors should perform western blot analysis using cells treated with shCBX4 and siCBX4.(Fig.1C-K, 5GH, 6, 7AB (7C-E, CBX4 mRNA should be analyzed)*

Reply: We sincerely apologize for the lack of these important details. In our revised manuscript, we have added many essential KD or KO efficiencies data which contained both RT-qPCR and western blot data. The siRNA-mediated KD efficiencies for siRNA library (Figure 1A-C) have been shown in **newly-added Appendix Figure S1A**. The shRNA-mediated KD efficiencies for seven J-Lat cells (Figure 1F-I, Figure EV1A-K) have been shown in **newly-added Appendix Figure S1B, S1C, S1E, S1F, S1H and S1I**, containing both RT-qPCR data and western blot data. The siRNA-mediated KD efficiencies for TZM-bl ChIP assay and J-Lat 10.6 ChIP assay (Figure 2A-E, Appendix Figure S2E) have been shown in **newly-added Appendix Figure S2A, S2B, S2F and S2G**, containing both RT-qPCR and western blot data. The sgRNA-mediated KO efficiencies for ATAC-Seq (Figure 2H) have been shown in **newly-added Appendix Figure S2I**. The siRNA-mediated KD efficiencies for EZH2 SUMOylation assays (Figure 5G and 5H) have been shown in **newly-added Appendix Figure S5A and S5B**, containing both RT-qPCR and western blot data. The siRNA-mediated KD efficiencies in TZM-bl cells (Figure 6A) have been shown in **newly-added Appendix Figure S6C and S6D**, containing both RT-qPCR and western blot data. The shRNA-mediated KD efficiencies for primary $CD4^+$ T cells (Figure 7A-D) have been shown in **newly-added Appendix Figure S7B, S7C, S7F and S7G**, containing both RT-qPCR and western blot data. The siRNA-mediated KD efficiencies for clinical samples (Figure 7E-G) have been shown in **newly-added Appendix Figure S7I**.

◯*For all the in vitro experiments using recombinant proteins or proteins purified from human cells, purity of proteins have to be analyzed by SDS-PAGE. Without quality checking, all in vitro assay are unreliable.*

Reply: We are sorry for these omissions. The purities of *in vitro* purified GFP-CBX4 and GFP-CBX4mut proteins which were used for *in vitro* droplet formation assay (Figure EV4A-D) have been validated by both Coomassie blue staining and western blotting. These data have been shown in **newly-added Figure EV4E**. In our *in vitro* methyltransferase assay, we also used recombinant mononucleosomes which were composed of H3.1, H2A, H2B and H4. The purities of these mononucleosomes have been validated by Coomassie blue staining. The results have been shown in **newly-added Figure 6E and Appendix Figure S6F**.

◯*Figure 2A: Authors indicates CBX4 accumulation on HIV LTR. However, to be sure the ChIP signal of CBX4, CBX4 ChIP in CBX4 kd cells should be performed.*

Reply: We thank the reviewer for the kind suggestion. In our revised manuscript, we have added the CBX4 ChIP data upon CBX4 knockdown in both TZM-bl cells (**revised Figure 2A**) and J-Lat 10.6 cells (**newly-added Appendix Figure S2E**).

◯*Figure3-E*

To consolidate the conclusion, quantitative data on co-localisation (CBX4 and HIV) is necessary.

Reply: We thank the reviewer for this suggestion. We have added another two ImmunoFISH data in **newly-added Figure EV2A and EV2B**. Besides, we have measured the distances of nearest CBX4 bodies and HIV-1 provirus DNA with Imaris software. In each group, 15 cells were randomly imaged and measured the CBX4-HIV-1 distances. The quantification of distances was evaluated in both J-Lat 10.6 and 8.4 cells. These data have been shown in **newly-added Figure EV2C and EV2D**.

◯*Figure4A*

The proteins in CBX4 LLPS (EZH2, RING1B and SUMO4) should be able to internally diffused and diffuse across boundary. So, FRAP experiments for EZH2, RING1B and SUMO4 in CBX4 LLPS are required.

Reply: We thank the reviewer for the kind suggestions. In our revised manuscript, we have added more data on dual-color FRAP experiments. We co-overexpressed RFP-tagged EZH2, RING1B and SUMO4 with GFP-tagged CBX4 in HEK293T cells respectively. About 24 hours post transfection, we used strong 488 nm and 561 nm laser power to bleach CBX4 bodies. Live cell images were captured every 5 s. In each combination, six co-localization bodies were proceeded to FRAP. Three bodies (ROI 1, ROI 2 and ROI 3) were left untreated. Another three bodies (ROI 4, ROI 5 and ROI 6) were bleached with strong 488 nm and 561 nm laser power. We found that both RFP-tagged proteins and GFP-CBX4 proteins within bleached CBX4 bodies quickly recovered fluorescence intensities, while the fluorescence intensities of unbleached CBX4 bodies were almost unchanged. These results indicated that other proteins within CBX4 bodies (EZH2, RING1B and SUMO4) were also able to internally diffuse and diffuse across boundary. These results have been shown in **newly-added Figure EV3C-E**.

◯*Figure4J-P*

Does CBX4mut obtain the ability interact with EZH2? IP experiments for CBX4 and EZH2 like Fig EV2F can reveal this point. For direct evidence CBX4mut lacks the ability forming LLPS, droplet formation assay using recombinant CBX4mut (shown in Fig EV3G using wt CBX4) is needed.

In addition, CBX4mut possibly lacks the ability to recruit EZH2 on HIV LTR. So ChIP for EZH2 in cells expressing CBX4mut will support authors conclusion.

Reply: We thank the reviewer very much for giving us these insightful suggestions. The reviewer gave us three suggestions which could comprehensively elucidate the relationship between CBX4mut and EZH2.

Firstly, the CBX4mut was unable to interact with EZH2. We co-overexpressed Flag-tagged EZH2 with HA-tagged CBX4 and CBX4mut respectively. HA-tagged proteins were IP with anti-HA beads. We found that CBX4 was able to enrich EZH2, while CBX4mut was unable to IP with EZH2. This result has been shown in **newly-added Figure EV4F**.

Secondly, the CBX4mut was unable to form LLPS droplets *in vitro*. We purified GFP-CBX4mut proteins *in vitro*. About 25 μM of GFP-CBX4mut proteins were incubated with different droplet formation buffers which contained different NaCl concentrations. Besides, Different concentrations of GFP-CBX4mut proteins were incubated within 75 mM NaCl droplet formation buffer. We found that GFP-CBX4mut proteins were unable to form droplets at any NaCl concentration and at any protein concentration. These results have been shown in **newly-added Figure EV4C-E**.

Thirdly, the CBX4mut was unable to recruit EZH2 on HIV-1 LTR. We used siRNAs targeting 3'UTR of *CBX4* mRNA to knock down CBX4 in TZM-bl cells, followed by overexpressing with wildtype CBX4 and LLPS-deficient CBX4mut respectively. ChIP assays with antibodies against IgG and EZH2 were performed in these cells. We found that the enrichment of EZH2 on HIV-1 LTR could be rescued when re-overexpressing wildtype CBX4 in CBX4-KD cells. However, the CBX4mut was unable to recruit EZH2 on the HIV-1 LTR in CBX4-KD cells. This result has been shown in **newly-added Figure EV4G**.

◯*Figure6B-E*

quantitative data is necessary.

Reply: We apologize for not showing quantitative data on western blot in our original manuscript. In our revised manuscript, we have added many quantitative results for nearly all the western blot data including Figure 6B-E. In **newly-added Appendix Figure S3E and S3G**, we showed the expression ratio of endogenous CBX4 and exogenous GFP-CBX4. In **revised Figure 5G and 5H**, we showed the ratio of SUMO1-EZH2 and SUMO4-EZH2 respectively. In **revised Figure EV5A-D**, we showed the ratio of SUMO-EZH2 within different conditions. In **revised Figure 6B, 6D and 6E**, we showed the expression ratio of SUMO4-EZH2 and H3K27me3. All the quantifications were based on the signal intensities of target protein bands and analyzed by Image Studio.

◯*Figure6E*

Authors suggested CBX4-mediated EZH2 SUMOylation enhanced the methyltransferase activity of EZH2, resulting in elevated H3K27me3. However, Fig 6 is not sufficient to draw this conclusion. The reasons are below.
-EZH2 mediated H3K27me3 methylation activity is quite low. In this experiment. As compared to lane1 with lane2 or lane3, there are no difference.
-Histone H3.1 is soluble? Reconstituted nucleosome or at least H3-H4 tetramer should be used in this assay.
-Moreover, to carry out such a complex system involving such a large number of proteins, it is necessary to use highly purified factors as recombinant proteins. Authors can not exclude the effect of contamination, therefore it is difficult to draw any conclusions from the authors' experiments.
-Efficiency of sumolylation of EZH2 in this experiment condition should be measured.
-Purity of each proteins should be analyzed.
-as control, experiments using EZH2 catalytic mutant and H3 K27R mutant are required.
Reply: We sincerely thank the reviewer for the above insightful suggestions. And we are also very sorry for the deficiencies of our study. The reviewer gave us six suggestions if we understand clearly. Suggestion 1 (S1): EZH2-mediated H3K27me3 is quite low and should be re-confirmed. S2: Reconstituted soluble nucleosomes should be used to conduct this complex experiment. S3: The purities of recombinant proteins should be verified and avoid contamination. S4: SUMOylated EZH2 should be measured by western blot. S5: EZH2 catalytic mutant should be included as negative control. S6: H3 K27R mutant should be included as negative control. We would like to response these suggestions utilizing our new *in vitro* methyltransferase assays. (The first suggestion will be answered in the end.)

Firstly, we have re-conducted the experiments many times with purified recombinant mononucleosomes instead of Histone H3.1 (Suggestion 2). We also agreed with the reviewer that Histone H3.1 only was not suitable to conduct this experiment, although it was indeed soluble in our system. Thus, we have bought recombinant mononucleosomes (H3.1) (Cat. No.: 31467) and Lys27Met-mutated mononucleosomes (H3.1 K27M) (Cat. No.: 81264) from Active Motif company. Within each group, we incubated the reactions with equal amounts of these mononucleosomes.

Secondly, we have conducted Coomassie blue staining assay to evaluate the purities of recombinant proteins (Suggestion 3). Although recombinant mononucleosomes were bought from company and have been verified by the company, we still conducted Coomassie blue staining to ensure that the purities of these mononucleosomes were very high and no contaminated proteins were introduced. All the other proteins were expressed from cells. The purities of these proteins including EZH2 could only be verified by western blot.

Thirdly, we have measured the signals of SUMOylated EZH2 (Suggestion 4). In Group 4, 5 and 6, we introduced the SUMOylation system proteins including SUMO4, UBC9 and CBX4. All the SUMOylated EZH2 bands were able to be detected. Besides, we have shown the expression ratio of SUMOylated EZH2 below the figure.

Fourthly, we have introduced EZH2 catalytic mutant group (Suggestion 5). Apart from the other four groups which were conducted in our original manuscript, we also introduced a group which were overexpressed with equal amounts (compared with wildtype EZH2) of EZH2 catalytic mutant (Y731D). Y731D mutant has been found to

totally abort EZH2 methylation activity (Ernst *et al*, 2010, PMID: 20601953; Lavarone *et al*, 2019, PMID: 30976011). Indeed, in our study, we found that EZH2 Y731D mutant was unable to tri-methylate H3K27.

Fifthly, we have introduced H3 K27R mutant (Suggestion 6). In Group 6, we used Lys27Met-mutated mononucleosomes (H3.1 K27M) as substrates instead of wildtype mononucleosomes (H3.1). We found that although EZH2 was co-overexpressed with EED, SUZ12, SUMO4, UBC9 and CBX4, K27M-mutated mononucleosomes were unable to be tri-methylated by EZH2.

Sixthly, we have shown distinct H3K27me3 signals upon conducting the above experiment as the reviewer suggested (Suggestion 1). In our new experiments, we have shown that H3K27 could be tri-methylated by EZH2, which was indicated by the intensities of H3K27me3 signals. Upon co-overexpressing with EED and SUZ12, the H3K27me3 signals were increased 2 fold. After further co-overexpressing with SUMO4, UBC9 and CBX4, the signals of H3K27me3 were enhanced much higher. The ratio of H3K27me3 was 6.38 fold compared with EZH2 only. The ratio of H3K27me3 within each group has been shown below the figure.

Overall, we have successfully conducted more complete *in vitro* methyltransferase assays according to the reviewer's insightful suggestions. These results have been shown in **revised Figure 6E and newly-added Appendix Figure S6F**. The corresponding main text, figure legends and methods have been revised accordingly as well.

◯*Figure7A-B*

Authors should access H3K27me3 and H2A119ub accumulation by ChIP in HIV infected primary CD4$^+$ T cells with shCBX4.

Reply: We thank the reviewer for the kind suggestion. We used wildtype HIV-1 viruses to infect PHA-activated primary CD4$^+$ T cells, followed by infecting with shCBX4 or shluc. On Day 6 post infection with HIV-1, ChIP assays with antibodies against H3K27me3 and H2AK119Ub were conducted in both shluc and shCBX4 groups. The results showed that upon KD CBX4 in primary CD4$^+$ T cells, the accumulation of both H3K27me3 and H2AK119Ub was significantly decreased. These results have been shown in **newly-added Figure 7B and 7C**.

◯*Figure EV2F*

Cell lysates for IP are treated with DNase and RNase? Interaction is dependent on DNA or RNA?

Reply: We are sorry for this omission. We haven't used DNase or RNase to treat IP samples in our original manuscript. To investigate whether the interaction of CBX4 and EZH2 was dependent on DNA or RNA, we re-performed this IP experiment. HA-tagged CBX4 and GFP were co-overexpressed with Flag-tagged EZH2, respectively. HA-tagged proteins were IP with anti-HA beads. In both groups, the IP reactions were supplemented with DNase/RNase or left untreated. Both total and IP samples were IB with anti-HA and anti-Flag antibodies. The results showed that upon DNase and RNase co-treatment in IP reactions, CBX4 still was able to interact with EZH2, which indicated that EZH2 was recruited by CBX4 in DNA- and RNA-independent manner. This result has been shown in **revised Figure EV2J**.

◯*Figure EV4*

Authors cannot conclude because there are a lot of non-specific bands. SUMOlylated CBX4 ladder is not clear.

For correct conclusion, the experiments using highly purified recombinant proteins are required.

Reply: We are sorry for not explaining these experiment clearly. Our experiments which conducted in Figure EV5 (former Figure EV4) were actually cell-based *in vivo* investigations. Thus, we haven't used *in vitro* purified recombinant proteins in these experiments. The reviewer also worried about that there seemed to be a lot of non-specific bands in both total and IP samples. Because a large number of cellular proteins would be SUMOylated upon overexpressing SUMO molecules, many SUMO-conjugated proteins would be IB in total samples, resulting in accumulated and laddered signals. In IP samples, we used beads to enrich EZH2 proteins. Those bands which were higher than wildtype EZH2 bands were actually SUMOylated EZH2 bands. SUMOylation often happens in chain reaction. SUMO molecules are able to conjugate to themselves and form SUMO chains like ubiquitin proteins do. Thus, a given candidate (such as EZH2) would be SUMOylated once, twice or thrice, resulting in the formation of 1×SUMO-EZH2, 2×SUMO-EZH2, 3×SUMO-EZH2. Those distinct bands which were not non-specific bands but SUMOylated EZH2 bands. We have shown the expression ratio of SUMO-EZH2 below each panel in **revised Figure EV5**.

****

Referee #2:

This study focusses on the function of CBX4 in transcriptional repression, specifically that of latent HIV-1. They report that CBX4 interacts with and SUMOylates EZH2 showing a connection between the PRC1 and PCR2 complexes. They claim this interaction occurs in a CBX4-driven biomolecular condensate and that this interaction is important for HIV-1 transcriptional repression, which maintains HIV-1 latency.

Reply: We thank the reviewer for the comprehensive summaries of our study.

The interaction of CBX4 and EZH2 in promoting HIV-1 latency through a biomolecular condensate is interesting. The authors note that the condensate step could be a new target for intervention and therefore this study shows novelty and impact. However, the scientific support for the condensate mechanism and the importance of EZH2 SUMOylation for HIV-1 repression is insufficient.

Reply: We thank the reviewer for the support of our study and pointing out the deficiencies of our work. We have re-conducted many experiments to improve the quality of previous data, and provided more new data to support the condensate mechanism and the importance of EZH2 SUMOylation on HIV-1 latency.

A general weakness of the manuscript is the lack of replicates. Of many important experiments only single replicates are shown, making it impossible to evaluate the validity of the results. I have noted specific cases throughout my comments.

Reply: We are very sorry for this severe weakness of our study. Replicates are very important to fully support our work, especially in important experiments. We thank the reviewer very much for pointing out these deficiencies in the following comments. We have corrected one by one and provided more experimental data and quantitative data.

The evidence for CBX4 condensation is exclusively supported by over-expression studies. Since condensation can be driven by artificially high protein concentrations these studies are inconclusive without confirmation at physiological protein concentrations, through direct labeling or artificial expression at physiological levels. The single endogenous-level dataset (Fig 3B-D) is promising but insufficient on its own. Importantly, the major finding that EZH2 enters CBX4 condensates while over-expressed can be exclusively concentration-mediated and therefore an artifact.

Reply: We thank the reviewer for pointing out these deficiencies of our study. We have provided more data to support CBX4 condensation in our revised manuscript.

Firstly, the reviewer suggested us to show more data on CBX4 condensation at physiological protein concentrations. We also agreed with the reviewer that the single endogenous-level dataset (**Figure 3B-E, Figure EV2A-D, Appendix Figure S3A-C**) were not sufficient. Thus, we conducted two experiments to mimic physiological CBX4 protein concentrations. We used siRNAs targeting 3'UTR of *CBX4* mRNA to knock down endogenous CBX4 in HEK293T cells. Then, we re-expressed equal amounts of GFP-tagged CBX4. The expression ratio of endogenous CBX4 and exogenous GFP-CBX4 was 1.00 : 1.04, which was evaluated by western blot. The result showed that equal amounts of GFP-CBX4 proteins formed similar nuclear condensates as endogenous CBX4 proteins. These results have been shown in **newly-added Appendix Figure S3D and S3E**. In another experiment, we overexpressed equal amounts of GFP-CBX4 proteins in wildtype HEK293T cells directly. Then, we used CF568-conjugated antibodies to capture total CBX4 proteins. We found that GFP-CBX4 proteins co-localized with CF568-conjugated CBX4 proteins (both endogenous and exogenous proteins), and formed similar nuclear puncta as endogenous CBX4 proteins which were shown in Figure 3B. The expression ratio of endogenous CBX4 and exogenous GFP-CBX4 was 1.00 : 1.05, which was evaluated by western blot. These results were shown in **newly-added Appendix Figure S3F and S3G**.

Secondly, the reviewer suggested us to show the co-localization of endogenous CBX4 and endogenous EZH2. Thus, we used rabbit anti-CBX4 antibody and mouse anti-EZH2 antibody to capture the endogenous CBX4 and EZH2 respectively. These cells were further incubated with AF488-conjugated anti-rabbit IgG antibody and CF568-conjugated anti-mouse IgG antibody. The co-localization was evaluated by line scan profile. The result showed that AF488-conjugated endogenous CBX4 co-localized with CF568-conjugated EZH2 and formed nuclear puncta. The result have been shown in **newly-added Figure EV2E**.

Currently, several observations concerning CBX4 foci are inconsistent with a condensate model:

Reply: We apologize for these deficiencies of our study. We have revised our manuscript according to the reviewer's following kind suggestions.

Fig4A: No replicates are shown for this data, at least 3+ biological replicates are necessary to gage the variability. The shown FRAP data is not consistent with LLPS. Almost no recovery is observed, ever after 10 minutes. This is significantly slower than previous reports of condensates. These kinetics are more consistent with aggregates, and the incomplete recovery indicates a large (50%) immobile fraction. These results argue against LLPS. Note that hardening of condensates can also be an effect of high protein concentrations in an artificial over-expression system.

Reply: We thank the reviewer for pointing out these deficiencies of our work. We have re-conducted FRAP experiments for GFP-CBX4 bodies and showed 3 biological replicates for each group. Six GFP-CBX4 bodies were circled and marked as region of interest 1 (ROI 1), ROI 2, ROI 3, ROI 4, ROI 5 and ROI 6. Three CBX4 bodies (ROI 4, ROI 5 and ROI 6) were bleached with strong 488 nm laser pulse. Another three unbleached CBX4 bodies (ROI 1, ROI 2 and ROI 3) were set as control. Images were captured every 4 s. Cells were imaged for 128 s in total. Relative fluorescence intensities of unbleached and bleached CBX4 bodies in each time point were calculated and represented as FRAP quantitation histogram. We found that bleached CBX4 bodies quickly recovered fluorescence intensities (50% within 30 s). These results, which have been shown in **revised Figure 4A**, indicated that CBX4 bodies were internally diffused and able to diffuse across boundary. Our new data indicated that CBX4 bodies are characterized of LLPS condensates rather than immobile aggregates.

Besides CBX4-only FRAP, we also conducted dual-color FRAP for CBX4 and its partners (EZH2, RING1B and SUMO4). We co-overexpressed RFP-tagged EZH2, RING1B and SUMO4 with GFP-tagged CBX4 in HEK293T cells respectively. About 24 hours post transfection, we used strong 488 nm and 561 nm laser power to bleach CBX4 bodies. We found that both RFP-tagged proteins and GFP-CBX4 proteins within bleached CBX4 bodies quickly recovered fluorescence intensities, while the fluorescence intensities of unbleached CBX4 bodies were almost unchanged. These results, which have been shown in **newly-added Figure EV3C-E**, indicated that CBX4 body partners were also internally diffused.

Fig 4B The fused foci do not relax into a sphere. Again, directly arguing against LLPS. The authors refer to Alberti, Gladfelter and Mittag, but do not interpret the results consistent with the guidelines in this paper.

Reply: We are sorry for the deficiencies of our data. The fused foci were actually spherical. The super-resolution SIM microscopy could easily remove transient weak signals during live cell imaging, resulting in the irregularity of CBX4 bodies. Thus, in our new experiment, we used Zeiss LSM 900 to conduct live cell imaging, the photobleaching and phototoxicity of which were weaker than SIM. The new data showed in **revised Figure 4B** indicated that two spherical CBX4 bodies gradually fused into a bigger spherical bodies.

Fig 4B,C Both B and C show "mottled" pattern which is unexpected, because highly dynamic LLPS proteins would equally distribute in the dense phase. Do the authors have an interpretation of what is happening here?

Reply: As we have mentioned in the above comment, our original data were captured with super-resolution SIM microscopy. Live cells actively moved during imaging.

Besides, it would cost us nearly 10 s to capture one image. The photobleaching and phototoxicity of SIM was also very high when we captured each image. The reconstitution algorithm of SIM also tended to remove weak signals. The final reconstructed images were indeed of high resolution while mottled protein condensates. The mottled pattern of our images might also be caused by the inexperience of our technicians. To this end, we have used high-speed confocal Zeiss LSM 900 to conduct all our live cell imaging experiments. The result of fusion experiment has been shown in **revised Figure 4B**. The result of fission experiment has been shown in **revised Figure 4C**, which showed that a spherical CBX4 body was able to split into two smaller spherical bodies. Both results showed that the LLPS proteins were equally distributed in the dense phase. We hope that the reviewer would be satisfied with our new data and our interpretation.

Fig EV3H, does not show increasing droplet size with the highest concentration, completely counter to what the authors claim in the text. There might be more droplets in the higher concentration regime, but the authors only show a single image and do not perform quantification of droplet numbers or size. At least quantification of several images is necessary to support (or disqualify) the claims made in the text.

Reply: We are very sorry for these deficiencies of our study. We have re-performed *in vitro* droplet formation assays in our revised manuscript. In our new experiment, we set more ladders of NaCl concentrations (500 mM NaCl, 250 mM NaCl, 125 mM NaCl, 62.5 mM NaCl, 31.25 mM NaCl, and 15.625 mM NaCl) and protein concentrations (1.5625 μM, 3.125 μM, 6.25 μM, 12.5 μM, 25 μM and 50 μM). Besides, we used Imaris software to count the numbers of droplets within each image. What's more, we also included *in vitro* droplet formation data of GFP-CBX4mut proteins besides those of GFP-CBX4 proteins. Our new data showed that CBX4 proteins were able to form more and bigger droplets along with the decrease of NaCl concentration. **CBX4 proteins also formed more and bigger droplets at higher protein concentrations.** However, GFP-CBX4mut proteins were unable to form droplets at any NaCl concentration and at any protein concentration. For each group, we conducted statistical analysis of droplets areas and numbers of GFP-CBX4/GFP-CBX4mut proteins. **Each statistical analysis was derived from three images.** These results have been shown in **revised and newly-added Figure EV4A-E**. The reviewer also suggested us to show more images for each group. However, to clearly show the numbers and sizes of droplets within each group, we captured large images and only showed one representative image. Showing three or more images would significantly decrease the visibilities of droplets within limited paper format, especially in higher NaCl concentrations and lower protein concentrations. We hope that the reviewer would allow us to show one representative image for each group. Our statistical analysis was indeed derived from three images.

Fig EV3I, J authors claim hundreds of nuclear puncta, but perform no quantification and show only single replicates.

Reply: We are sorry for these omissions. In our revised manuscript, we have shown two more images of GFP-CD-NLS-CBox. Besides, we also calculated the numbers of nuclear puncta based on 8 images utilizing Imaris software. The results showed that compared with GFP-CD-NLS, GFP-CD-NLS-CBox formed hundreds of nuclear puncta (average

262 bodies, ranging from 147 to 479). These data have been shown in **newly-added Appendix Figure S4A and S4B** (former Figure EV3I and EV3J).

As a second major point the authors claim that the condensation of CBX4 with EZH2 is important for EZH2 SUMOylation and repression of HIV-1. However, the presented data does not support this point.

Reply: We thank the reviewer for giving us the following insightful suggestions. We would like to response to these comments one-by-one.

Fig 6B. While expression of CBX4 is shown to increase EZH2 SUMOylation, CBX4 mediated increase of EZH2 SUMOylation is NOT dependent on its capacity to form foci. Contrary to the claim of the authors in the text, in this panel there is no reduction in EZH2 SUMOylation (compare last two lanes). This would indicate that the ability to form foci is irrelevant for EZH2 SUMOylation.

Reply: The SUMOylation of EZH2 was indeed decreased upon overexpressing CBX4mut based on intensities analysis. However, this SUMOylation experiment haven't been conducted good enough. The signal was not well-resolved. Thus, in our revised manuscript, we have re-performed this experiment. Our new data showed distinct SUMOylated EZH2 bands in SUMO4 overexpression groups. We found that the overexpression of CBX4 was able to enhance the SUMOylation of EZH2 (from 14.11 fold to 22.03 fold). While overexpression of CBX4mut did not enhance EZH2 SUMOylation (5.33 fold vs 4.70 fold). These results have been shown in **revised Figure 6B**. The ratio of SUMO4-EZH2 within each group has been shown below the figure, which was : 1.00 : 5.33 : 14.11 : 22.03 : 4.70.

The reviewer also concerned that CBX4 mediated increase of EZH2 SUMOylation might not dependent on its capacity to form foci. We have provided more data that CBX4mut was unable to enrich EZH2 and recruit EZH2 on the HIV-1 LTR (**newly-added Figure EV4F and EV4G**). Our imaging data indicated that CBX4mut was unable to colocalize with EZH2 (**Figure 4P)**. Thus, CBX4 utilized its CBox LLPS motif to recruit EZH2 to its condensate. Within CBX4 condensates, CBX4 utilized its SUMO E3 ligase activity to SUMOylate EZH2. Combined our SUMOylation assay data which showed in **revised Figure 6B**, we speculated that the condensation of CBX4 not only influenced the distribution of EZH2 but also altered the SUMOylation status of EZH2. Thus, we believe that CBX4-mediated EZH2 SUMOylation has strong positive correlation with its body formation capability.

Fig 6D. In this figure the authors claim to tie the increased SUMOylation of EZH2 to an increased level of H3K27Me3, which in turn could indicate more efficient repression of HIV-1. However, there is no clear difference between the two conditions and the bands are not quantified. As presented here, the data shows the opposite of what the authors claim in the text.

Reply: We apologize for this confusing results. We agreed with the reviewer that this data was not consistent with what we claimed, especially without any quantitation data. Thus, in our revised manuscript, we have re-performed this experiment. Our original idea to conduct this experiment was based on the observation that CBX4-mediated EZH2 SUMOylation could significantly enhance H3K27me3 on the HIV-1 LTR. While the

CBX4mut neither enhanced the SUMOylation of EZH2 nor enhanced H3K27me3 on the HIV-1 LTR (**revised Figure 6B and 6C**). Thus, we speculated that CBX4 might influence EZH2-mediated H3K27me3 directly. To this end, we knocked down the endogenous EZH2 with siRNAs targeting 3'UTR of *EZH2* mRNA, followed by overexpressing EZH2, SUMO4, UBC9 and / or CBX4. In our **revised Figure 6D**, we found that the depletion of EZH2 significantly decreased H3K27me3. The re-expression of EZH2 could potentially rescue H3K27me3 to the basal level (from 0.13 fold to 1.39 fold). Upon co-overexpression of SUMO4 and UBC9 with EZH2, the H3K27me3 signals could be enhanced much higher (from 1.39 fold to 2.85 fold). The co-overexpression of SUMO4, UBC9 and CBX4 with EZH2 was able to further enhance the signals of H3K27me3 (from 2.85 fold to 4.75 fold). The ratio of SUMO4-EZH2 (1.00 : 0.00 : 1.26 : 2.93 : 11.07) as well as ratio of H3K27me3 (1.00 : 0.13 : 1.39 : 2.85 : 4.75) have been shown below the figure. Our data indicated that CBX4-mediated EZH2 SUMOylation has positive correlation with the H3K27me3 catalytic activity of EZH2, which resulted in more efficient repression of HIV-1.

Fig 6E. In this figure the authors do show an increased level of H3K27Me3, however, here this is caused by the co-expression of CBX4. While this is presumably due to an increase of SUMOylation of EZH2, there are no bands of SUMO-EZH2 visible in this blot. Does CBX4 not induce EZH2 SUMOylation here? That would run counter to the model proposed by the authors.

Reply: We sincerely apologize for omitting this important detail in our original manuscript. We have re-performed the *in vitro* methyltransferase assay and set more control groups to fully support our proposed model. Our experiment has been re-designed as below. In the first group, HEK293T cells were transfected with empty vector. In the second group, we overexpressed EZH2 in HEK293T cells. In the third group, we co-overexpressed EZH2 with EED and SUZ12, both of which were major subunits of PRC2. In the fourth and sixth groups, we co-overexpressed EZH2 with EED, SUZ12, SUMO4, UBC9 and CBX4. In the fifth group, we co-overexpressed EZH2 catalytic mutant (Y731D, named EZH2mut) with EED, SUZ12, SUMO4, UBC9 and CBX4. Forty-eight hours post transfection, we enriched EZH2 and EZH2mut proteins and incubated EZH2 or EZH2mut from each group with *in vitro* purified mononucleosomes and the cofactor S-adenosyl-L-methionine (SAM). Our data showed that the co-overexpression of EZH2, EED and SUZ12 could potentially enhance EZH2 activity on H3K27me3 modification compared with overexpression EZH2 only. Whereas, the co-overexpression of EZH2 with EED, SUZ12, SUMO4, UBC9 and CBX4 significantly enhanced EZH2-mediated H3K27me3 modification. However, Y731D-mutated EZH2mut was unable to catalyze H3K27me3 modifications on mononucleosomes. We also incubated EZH2 and CBX4-SUMOylated EZH2 with H3K27-mutated mononucleosomes (H3.1 K27M) and found that no H3K27me3 signals were observed within mononucleosomoes. These results have been shown in **revised Figure 6E**. **In Figure 6E, we also showed SUMO4-EZH2 bands in IP samples. The ratio of SUMO4-EZH2, which was 0.00 : 1.00 : 1.05 : 10.37 : 10.86 : 10.02, has been shown below the figure.**

Minor comments:

Fig 1 F-I Shows very similar data in different systems and therefore can be supplemental.

Reply: We thank the reviewer for this kind suggestion. We have moved these data to **newly-added Figure EV1A-D**. Besides, we have provided viral RNA quantification data for these 7 J-Lat cell lines, which have been shown in **newly-added Figure EV1E-K**.

Fig 2B, 2C, Unclear what amplicon is used here, please indicate the tested region.
Reply: We are sorry for this omission. All the ChIP signals which were shown in Figure 2B-G were measured on Position "B" (Nuc-free region). Thus, we have added this description in the figure legend of Figure 2B-G and represented as: **Only "B" position signals were shown and normalized to Input**. Besides, we have added essential descriptions in Figure 2B. We marked **"B" position signals** under the figure title.

Fig 2E, H3K27Ac does not change, why is this not further discussed/explained?
Reply: We thank the reviewer for pointing out this insufficiency. In the corresponding main text, we have added essential discussion for the change of H3K27me3, H2AK119Ub and H3K27Acetyl. After the results of Fig 2B, we add: "**As the main modifier of H3K27me3 was EZH2, this result suggested that CBX4 depletion might influence the distribution and function of EZH2 on the HIV-1 LTR.**" After the results of Fig 2D and E, we added: "**Because the H2AK119Ub modification was catalyzed by RING1B, the downregulation of H2AK119Ub upon knocking down CBX4 suggested that CBX4 might recruit RING1B to the HIV-1 LTR. The H3K27Acetyl was mainly deacetylated by HDAC1 and HDAC2. The stabilization of H3K27Acetyl upon CBX4 depletion indicated that CBX4 might not cross-talk with histone acetylation.**"

Fig 3E Only a single replicate is shown more replicates and quantification are needed to interpret these results.
Reply: We thank the reviewer for this kind suggestion. We have provided more immunoFISH data for naïve and activated J-Lat cells. These data have been shown in **newly-added Figure EV2A and B**. Besides, we have measured the distances of the nearest CBX4 bodies and HIV-1 proviruses DNA in both J-Lat 10.6 and J-Lat 8.4 cells utilizing Imaris software. In each group, 15 cells were randomly imaged and measured the CBX4-HIV-1 distances. These quantification data have been shown in **newly-added Figure EV2C and D**.

Fig 4D concentration of hexanediol not mentioned.
Reply: We apologize for this omission. We have marked the concentration of 1,6-Hexanediol (1,6-Hex) in Fig 4D, which was **1.5% (wt/vol)**. We also have shown this information of concentration within the figure legend of Figure 4 and the method of "Live cell imaging" as well.

Fig EV3C-F Does not test what the authors claim. These experiments indicate that the observed foci are insensitive to osmotic stress, but do not inform about physiological conditions, since none of the tested regimes are physiological.
Reply: We also agree with the reviewer's insightful comment. Our data which were shown in Figure 3 and 4 have already confirmed that CBX4 foci existed in physiological conditions. These crowding agents experiments were unnecessary. Thus, in our revised manuscript, these data and corresponding main text have been removed.

****

Referee #3:

CBX4 contributes to HIV-1 latency by forming phase-separated nuclear bodies and SUMOylating EZH2

This work is based on the previous findings that Polycomb complexes PRC1 and PRC2 mediate HIV-1 latency by establishing and maintaining the heterochromatin environment through H3K27me3 at the viral promoter. The authors therefore screen for Polycomb specific factors and discover that CBX4 of PRC1 and EZH2 of PRC2 restrict HIV-1 transcription in TZM-bl cells. They then focus on CBX4 mainly, and investigate its relevance for HIV-1 transcription and latency in several latency models and HIV-1 patient derived CD4$^+$ T cells.

Reply: We appreciate for reviewer's comprehensive summaries of our work.

While there are some interesting aspects of the manuscript, I think that many of the approaches need to be deepened and better clarified.

Reply: We appreciate for reviewer's positive feedback and pointing out the deficiencies of our study.

Although they talk about transcriptional latency mediated by CBX4 they start by measuring GFP fluorescence in FACS analysis and never measure the levels of viral transcripts in the cells. They perform ChIP and ATAC-Seq without providing adequate controls and explaining in details ATAC-seq. Of note, although apparently simple to perform , ATAC-seq analysis is rather complex and requires a better explanation (not just two sentences in the paragraph). As it is right now, leaves a reader with an idea that it is superfluous.

Reply: We thank the reviewer for pointing out these important details. **Firstly, we have measured the expression of viral RNAs to indicate HIV-1 reactivation.** We also agreed with the reviewer that measuring GFP-positive cells only was not sufficient to show CBX4-mediated HIV-1 reactivation. In our revised manuscript, we have used primer pairs (HIVTotRNA Forward Primer: 5'-CTGGCTAACTAGGGAACCCACTGCT-3' and HIVTotRNA Reverse Primer: 5'-GCTTCAGCAAGCCGAGTCCTGCGTC-3') which targeted conserved region from U5 to GLS (*Gag* leader sequence) to quantify intracellular HIV-1 RNAs. For reactivation in seven J-Lat cells (J-Lat 6.3, 8.4, 9.2, 10.6, 15.4, Mix and A2), we have provided more data of relative viral mRNA expression. These data have been shown in **newly-added Figure EV1E-K**. For reactivation in primary CD4$^+$ T cells, fold changes of HIV-1 mRNA expression at different time points within different groups were evaluated, which have been shown in **newly-added Appendix Figure S7A and S7E**. Our clinical samples data already have shown the intracellular HIV-1 RNAs. These data have been shown in **revised Figure 7E-G**.

Secondly, we have added essential controls for both ChIP and ATAC-Seq experiments. In Figure 2, we have conducted CBX4 ChIP in both siNC and siCBX4 TZM-bl cells (**revised Figure 2A**). We also have conducted CBX4 ChIP in J-Lat 10.6 cells. IgG and CBX4 signals in both siNC and siCBX4 groups were evaluated (**newly-**

added Appendix Figure S2E). In each ChIP assays, IgG groups were always included and treated as negative control. Besides the HIV-1 promoter regions, Cellular DNA and viral 5'LTR junction regions (G5 in TZM-bl cells and G5' in J-Lat 10.6 cells) were also quantified in both siNC and siCBX4 groups. The CBX4 KD efficiencies in TZM-bl cells and J-Lat 10.6 cells were validated by both RT-qPCR and western blot (**newly-added Appendix Figure S2A, S2B, S2F and S2G**). The CBX4 IP efficiencies were also validated in siNC and siCBX4 groups by western blot (**newly-added Appendix Figure S2D**). For ATAC-Seq data, we have shown the knockout efficiencies of sgCBX4 (**newly-added Appendix Figure S2I**). We also have measured cellular viabilities upon KO CBX4 in J-Lat 10.6 cells (**newly-added Appendix Figure S2J**).

Thirdly, we have added more descriptions and discussions for ATAC-Seq experiments within the main text and methods. The newly-added sentences were as below: The absence of heterochromatin modification and the presence of transcription machinery often indicated higher chromatin accessibility. Thus, we conducted ATAC-Seq in J-Lat 10.6 cells to directly probe the chromatin accessibility of the HIV-1 LTR upon CBX4 knockout (Buenrostro et al, 2015). Both the HIV-1 genome and the human genome were analyzed for accessibility. All the sequencing reads were aligned to both HIV-1 reference genome (K03455.1) and human reference genome (GRCh38) respectively. Only reads which contained both HIV-1 5'LTR and human integration junction DNA were sorted and normalized to the total mapped reads. The tag peaks were generated utilizing these reads and visualized by Igvtools. We found that the transposable tag density, which indicated the accessible region, was significantly increased on the HIV-1 LTR upon knocking out CBX4 (Fig 2H, and Appendix Fig S2I and J).

Moreover, the Authors suggest that CBX4 forms nuclear bodies with LLPS properties on the HIV-1 LTR, a claim based on ImmunoDNA FISH which shows no quantification or specific.

Reply: We are sorry for the lack of quantification of ImmunoFISH. In our revised manuscript, we have added more ImmunoFISH data for naïve and activated J-Lat cells. These data have been shown in **newly-added Figure EV2A and 2B**. Besides, we have measured the distances of the nearest CBX4 bodies and HIV-1 proviruses DNA in both J-Lat 10.6 and J-Lat 8.4 cells utilizing Imaris software. In each group, 15 cells were randomly imaged and measured the CBX4-HIV-1 distances. The results showed that CBX4 bodies tended to get close to or co-localize with HIV-1 proviruses in naïve J-Lat cells. Upon activation by TNFα, CBX4 bodies were away from HIV-1 proviruses, resulting in longer distances between the nearest CBX4 body and the HIV-1 provirus. These quantification data have been shown in **newly-added Figure EV2C and D**.

The only somewhat solid data are in vitro SUMOylation assays, suggesting that CBX4 recruits EZH2 utilizing its SUMO E3 ligase activity. However, the fact that CBX4 sumoylates EZH2 is not really connected to HIV-1 latency. The authors state that CBX4 and EZH2. act synergistically to maintaining HIV-1 latency, but do not actually demonstrate synergism.

Reply: We are sorry for the deficiencies of our study. The reviewer concerned about that CBX4-mediated EZH2 SUMOylation might not be connected to HIV-1 latency. Thus, in Figure 6, we have re-performed several experiments and added more new data to indicate

that CBX4 suppressed HIV-1 expression by SUMOylating EZH2. In **Figure 6A and newly-added Appendix Figure S6A and S6B**, we have shown that both the mutation of two SUMO-interacting motifs (SIMs), the deletion of CBox and the mutation of LLPS motif within CBox aborted CBX4-mediated HIV-1 re-suppression. These results indicated that both the LLPS property and the SUMO E3 ligase activity of CBX4 contributed to the suppression of HIV-1 expression. As we have found that CBX4 SUMOylated EZH2 in Figure 5, we conducted further ChIP assay to investigate whether CBX4-mediated EZH2 SUMOylation could influence EZH2 catalytic activity. Our ChIP assays with antibodies against H3K27me3 on the HIV-1 LTR showed that the co-overexpression of SUMOylation system (SUMO4, UBC9 and CBX4) with EZH2 was able to significantly enhance the H3K27me3 modifications on the HIV-1 LTR (**revised Figure 6C**). However, the co-overexpression of LLPS-deficient mutant CBX4mut with SUMO4, UBC9 and EZH2 was unable to further enhance H3K27me3 modifications. We also conducted the H3K27me3 rescue experiment (**revised Figure 6D**) and the *in vitro* methyltransferase assay (**revised Figure 6E**). All the results indicated that CBX4-mediated EZH2 SUMOylation enhanced the H3K27me3 catalytic activity of EZH2, resulting in elevated H3K27me3 on the HIV-1 LTR. Combined all the results showed in Figure 6, we do believe that CBX4 contributed to HIV-1 suppression and HIV-1 latency by SUMOylating EZH2 and enhancing H3K27me3 modifications on the HIV-1 LTR.

The reviewer also concerned about that our data might not fully indicate that CBX4 and EZH2 synergistically maintained HIV-1 latency. Previous reports have shown that PRC2 component EZH2 contributed to HIV-1 latency by catalyzing H3K27me3 modifications (Friedman *et al.*, 2011, PMID: 21715480). The established H3K27me3 was maintained by PRC1 components CBX4 and other CBX proteins (CBX2, CBX6, CBX7 and CBX8) (Khan *et al.*, 2018, PMID: 28246360; Guo *et al.*, 2021, PMID: 33494958). Thus, both PRC1 and PRC2 have been shown to contribute to HIV-1 latency. However, how PRC1 and PRC2 connected to each other was not defined. Our work presented here indicated that CBX4 bridged PRC1 and PRC2 by recruiting EZH2 and SUMOylating EZH2, resulting in elevated H3K27me3 on the HIV-1 LTR. The H3K27me3 modification acted as docking site for CBX4 and was further maintained by CBX4. Our data showed that CBX4 mediated positive feedback to PRC2 by SUMOylating EZH2 to enhance H3K27me3 activity of EZH2. The collaboration of CBX4 and EZH2 orchestrated both PRC1 and PRC2, which significantly maintained the H3K27me3-containing fHC on target genomic DNA including the HIV-1 promoter. Thus, we believed that CBX4 and EZH2 acted synergistically to maintain HIV-1 latency. **Our above interpretation has been further discussed within Discussion of the main text**.

Major points:

• *Introduction is not original with respect to the Authors previous publication ' Histone chaperone CAF-1 promotes HIV-1 latency by leading the formation of phase-separated suppressive nuclear bodies'.*

Reply: We sincerely apologize for the structural similarity of these two manuscript. In our revised manuscript, we have deleted many redundant descriptions. For many inaccurate phrases which were pointed out by the reviewer in the following comments, we have carefully checked and added essential descriptions. The part of the mechanisms of HIV-1 latency has been simplified. In the part of LLPS, we replaced many old

illustrations with more recent cases. We hope that the reviewer would be satisfied with our modifications.

• *128-131. Language used to build the hypothesis is very trivial, and unclear.*
Reply: We are sorry for the trivial and unclear hypothesis. We also agreed with the reviewer that these descriptions were too redundant and too subjective. In our revised manuscript, we have removed these sentences.

• *Treatment controls are generally missing, including efficiencies of KD, cell viability, ChIP controls for cellular regions (good practice shows positive and negative regions).*
Reply: We thank the reviewer for pointing out the above deficiencies of our study. Throughout our revised manuscript, we have added many essential treatment control data. We have shown the KD efficiencies of siRNA library in **newly-added Appendix Figure S1A**. We also have shown KD efficiencies of shCBX4 in 7 J-Lat cell lines (J-Lat 6.3, 8.4, 9.2, 10.6, 15.4, Mix and A2) utilizing both RT-qPCR and western blot (**newly-added Appendix Figure S1B, S1C, S1E, S1F, S1H and S1I**). We also have shown KD efficiencies of siCBX4 in ChIP assays (**newly-added Appendix Figure S2A, S2B, S2D, S2F and S2G**). We also have shown KO efficiencies of sgCBX4 in ATAC-Seq assay (**newly-added Appendix Figure S2I**). We also have shown KD efficiencies of siCBX4 in SUMOylation assays (**newly-added Appendix Figure S5A and S5B**). We also have shown KD efficiencies of siCBX4 in TZM-bl cells (**newly-added Appendix Figure S6C and S6D**). The shCBX4-mediated KD efficiencies in primary $CD4^+$ T cells have been shown in **newly-added Appendix Figure S7B, S7C, S7F and S7G**. The siCBX4-mediated KD efficiencies in clinical samples were shown in **newly-added Appendix Figure S7I**. Cellular viabilities upon KD or KO CBX4 also have been shown in corresponding appendix figures (**newly-added Appendix Figure S1D, S1G, S1J, S2C, S2H, S2J, S5A, S5B, S6E, S7D, S7H and S7J-L**). For ChIP assays, in any targeted protein ChIP experiment, we always included IgG negative controls (**revised Figure 2A-G, Appendix Figure S2E, Figure EV2K, Figure EV4G, Figure 6C, Figure 7B and 7C**). We also have shown CBX4 IP efficiencies by western blot in both siNC and siCBX4 samples (**newly-added Appendix Figure S2D**). Besides the HIV-1 promoter regions, cellular DNA and viral 5'LTR junction regions (G5 in TZM-bl cells and G5' in J-Lat 10.6 cells) were also quantified in both siNC and siCBX4 groups (**revised Figure 2A and Appendix Figure S2E**).

More specific comments are listed bellow:
Figure 1. CBX4 contributes to HIV-1 latency. TZM-bl data are not convincing, 3 fold change in luc activity with respect to mock ie, leaky LTR-luc expression is not reactivation. Additional controls should be included, like, Tat transfection, PHA/TPA treatment. This system is not adequate to study latency and should be avoided. This is a system to study (at a very basal level) HIV-1 transcription.
Reply: We thank the reviewer for pointing out the deficiencies of our study and giving us constructive suggestions. In our revised manuscript, we have added another two panels of data which showed that the co-treatment of Tat transfection or TNFα stimulation with siRNAs were able to significantly enhance the expression of LTR-driven luciferase. Especially for CBX4, co-treating TZM-bl cells with siCBX4 and Tat could induce 6 fold

of luciferase expression compared with co-treating with siNC and Tat. Co-treating TZM-bl cells with siCBX4 and TNFα could induce 4.7 fold of luciferase expression compared with co-treating with siNC and TNFα. These data have been shown in **newly-added Figure 1A-C**. We also agreed with the reviewer that the TZM-bl cell line was not an adequate model to study HIV-1 latency. Thus, we only used these data to indicate that CBX4 might suppress the HIV-1 promoter activity. Our major data showed in **Figure 1F-I and Figure EV1** indicated that CBX4 contributed to HIV-1 latency as well, mainly through restricting the HIV-1 promoter activity.

Furthermore, in A) SiRNA KD efficiency for PRC1 and PRC2 complex is missing unless available online, access not provided to the Reviewer; although stated (Expanded View for this article is available online).
Reply: We apologize for these omissions. In our revised manuscript, we have provided all the siRNA-mediated KD efficiencies for the siRNA library (Figure 1A-C) utilizing RT-qPCR assay. These data have been shown in **newly-added Appendix Figure S1A**.

In B) there is an issue with data normalization. Shouldn't it be siNT in all reconstitution plasmid points at 1?
Reply: We thank the reviewer for pointing out this improper data normalization. In our revised manuscript, we have normalized all the data in siCBX4 groups to those in siNC groups in all reconstitution plasmid points, which means siNC groups in all reconstitution plasmid points have been at 1.

The reactivation of all latent models, except maybe for the JLat mix is very modest upon KD of CBX4, only few precents in all presented backgrounds.
Reply: The reactivation in these J-lat models was indeed not very high. The reason was all the monoclonal J-Lat cell lines including J-Lat 6.3, 8.4, 9.2, 10.6, 15.4 and A2 integrated into different cellular genomic positions. The reactivation potentials of these cell lines were significantly influenced by their integration sites. So the overall reactivation of these cell lines was very different from each other. Although the reactivation upon KD CBX4 was modest, it was significantly higher compared with those in shluc groups. The fold change of GFP-positive cells (indicating latency-reactivated cells) upon KD CBX4 by shCBX4 was higher when comparing with those in shluc groups. Besides, we also provided data of fold change of HIV-1 mRNA expression upon KD CBX4. The relative viral mRNA expression was more sensitive than the percentages of GFP-positive cells. These data have been shown in **newly-added Figure EV1E-K**.

171-172. Why did the authors opt for JQ1 and SAHA specifically? Is there a scientific rationale behind these compounds? They are just mentioned as 'other latency-reversing agents (LRAs) including HDAC inhibitor SAHA and BRD4 inhibitor JQ-1'.
Reply: We are sorry for not explaining that clearly. We chose SAHA and JQ-1 based on three reasons. Firstly, SAHA and JQ-1 have been widely used to serve as traditional LRAs ((Ait-Ammar *et al.*, 2020, PMID: 32038533; Spivak & Planelles, 2018, PMID: 29099677). Secondly, SAHA targets HDAC and JQ-1 targets BRD4. Both compounds have no influence on the activity of PcG proteins including CBX4. Thirdly, the enhancement of HIV-1 latency reactivation upon co-treating other LRAs with CBX4 KD

reflected that CBX4 could be a potential target to develop novel LRAs. We have added essential descriptions after Line 172, which was : "**Both SAHA and JQ-1 have been widely used to reactivate latent HIV-1 and have no influence on the activity of PcG proteins (Ait-Ammar et al., 2020; Spivak & Planelles, 2018).**"

Figure 2. CBX4 contributes to H3K27me3 modification of HIV-1 promoter.
ChIP was performed first in TZM-bl cells, not an adequate model for HIV-1 chromatin binding as only a small portion of the viral genome is present. Moreover, enrichment of the main factor CBX4 is expressed over IgG. Could the authors show as percentage of input, and also add the specificity of the IP, validation for the CBX antibody.
Reply: We thank the reviewer for these kind suggestions. If we understand clearly, the reviewer gave us three insightful suggestions to further improve our study.

Firstly, other models besides the TZM-bl cell line should be used to conduct ChIP assays. We agreed with the reviewer that the TZM-bl cell line, which is an HIV-1 expression model, is not an adequate model to study chromatin alterations and HIV-1 latency. While the purpose of our ChIP assays in TZM-bl cells was to study epigenetic modifications upon KD CBX4 in a very simple model. We tried to verify that CBX4 contributed to HIV-1 suppression in TZM-bl cells, mainly by influencing the epigenetic modifications on HIV-1 LTR (**Figure 2A-E**). Then, we conducted further ChIP assays in a true HIV-1 latency model J-Lat 10.6. We found that CBX4 contributed to HIV-1 latency as well. CBX4 could enrich on the HIV-1 LTR and maintain suppressive epigenetic modifications including H3K27me3 on the latent HIV-1 LTR (**Figure 2F and 2G**). We have added more essential data, especially CBX4 ChIP data (**newly-added Appendix Figure S2E**), to verify that CBX4 promoted HIV-1 latency.

Secondly, the enrichment of main factor CBX4 should be normalized to Input and represented as percentages of Input. We are sorry for this improper data presentation. All the ChIP assay data including **Figure 2A-G, 6C, 7B and 7C, Appendix Figure S2E, Figure EV2K and EV4G** have been carefully checked and represented as the percentages of Input (% Input).

Thirdly, CBX4 antibody validation data should be included to indicate the IP specificity. We have added several data to evaluate the siRNA KD efficiencies and CBX4 IP efficiencies in ChIP assays. The efficiencies of siCBX4-mediated CBX4 KD in both TZM-bl cells and J-Lat 10.6 cells have been evaluated by RT-qPCR and western blot (**newly-added Appendix Figure S2A, S2B, S2F and S2G**). The IP efficiencies of ChIP-grade CBX4 antibodies have been validated by western blot (**newly-added Appendix Figure S2D**).

Validation ChIP in JLat clone for the CBX absence on LTR upon KD is missing, and should be included.
Reply: We thank the reviewer for the kind suggestions. Firstly, we have added the CBX4 ChIP data upon CBX4 KD in TZM-bl cells (**revised Figure 2A**). Secondly, we also performed CBX4 ChIP assays in both siNC and siCBX4 J-Lat 10.6 cells. This data has been shown in **newly-added Appendix Figure S2E**.

Authors should clarify ATAC-Seq analysis, specifically how the normalization was performed, if any, aside by normalizing to the total mapped reads.

Reply: We are sorry for not describing ATAC-Seq analysis clearly. **In our revised manuscript, we have added more descriptions and discussions for ATAC-Seq experiments within the main text and methods.** The newly-added sentences were as below: "The absence of heterochromatin modification and the presence of transcription machinery often indicated higher chromatin accessibility. Thus, we conducted ATAC-Seq in J-Lat 10.6 cells to directly probe the chromatin accessibility of the HIV-1 LTR upon CBX4 knockout (Buenrostro et al, 2015). Both the HIV-1 genome and the human genome were analyzed for accessibility. All the sequencing reads were aligned to both HIV-1 reference genome (K03455.1) and human reference genome (GRCh38) respectively. Only reads which contained both HIV-1 5'LTR and human integration junction DNA were sorted and normalized to the total mapped reads. The tag peaks were generated utilizing these reads and visualized by Igvtools. We found that the transposable tag density, which indicated the accessible region, was significantly increased on the HIV-1 LTR upon knocking out CBX4 (Fig 2H, and Appendix Fig S2I and J)." Besides, we have modified Figure 2H by adding "Normalized to total mapped reads" in the title of vertical axis.

Figure 3. CBX4 recruits PcG proteins and forms nuclear bodies.
The authors overexpress CBX and perform Mass spec comparing it with control empty vector. They identify ~700 interacting proteins of CBX among which EZH2 and SUMO4. These proteins were then overexpressed in HEK293 with RFP tags. It is clear from the images that there is a colocalization between CBX4 and EZH2 as well as between CBX4 and SUMO4. However, the size of the colocalization bodies is significantly different, thus putting a question mark on the relevance of the overexpression(s). It also rises a possibility that many partners of CBX4 lead to the formation of heterogeneous CBX4 bodies, as can be inferred also from their various sizes. However, this further implies that it is very hard to understand which one of these bodies is responsible for HIV-1 silencing, and questions the ImmunoFiSH data which can be completely random, as there are way too many objects (bodies) and only one FISH signal. There are no control set objects and no quantifications either (Fig 3E and EV2A).
Reply: We thank the reviewer for the above insightful comment. Co-localization bodies were indeed different from each other. However, we believe that **CBX4 bodies themselves were heterogeneous in physiological conditions**. Firstly, our data showed in Figure 3B-D and Appendix Figure S3A-C have indicated that the diameters of endogenous CBX4 bodies were highly heterogeneous (average 558 nm, ranging from 93 nm to 1432 nm). Actually, when we examining endogenous CBX4 bodies under microscope, the sizes, the distributions and the numbers of CBX4 bodies within single cells were very different from each other. Maybe the variabilities were caused by different co-partners as the reviewer speculated. Investigating the determinant of variable CBX4 bodies was out of the scope of our study.

Secondly, we have conducted two experiments to mimic physiological CBX4 protein concentrations. We used siRNAs targeting 3'UTR of *CBX4* mRNA to knock down endogenous CBX4. Then, we re-expressed equal amounts of GFP-tagged CBX4. The expression ratio of endogenous CBX4 and exogenous GFP-CBX4 was 1.00 : 1.04, which was evaluated by western blot. The result showed that equal amounts of GFP-CBX4 proteins formed similar nuclear condensates as endogenous CBX4 proteins. These results

have been shown in **newly-added Appendix Figure S3D and S3E**. In another experiment, we overexpressed equal amounts of GFP-CBX4 proteins in wildtype HEK293T cells directly. Then, we used CF568-conjugated antibodies to capture total CBX4 proteins. We found that GFP-CBX4 proteins co-localized with CF568-conjugated CBX4 proteins (both endogenous and exogenous proteins), and formed similar nuclear puncta as endogenous CBX4 proteins which were shown in Figure 3B. The expression ratio of endogenous CBX4 and exogenous GFP-CBX4 was 1.00 : 1.05, which was evaluated by western blot. These results were shown in **newly-added Appendix Figure S3F and S3G**.

Thirdly, We also have conducted endogenous co-localization assay. We used rabbit anti-CBX4 antibody and mouse anti-EZH2 antibody to capture the endogenous CBX4 and EZH2 respectively. These cells were further incubated with AF488-conjugated anti-rabbit IgG antibody and CF568-conjugated anti-mouse IgG antibody. The co-localization was evaluated by line scan profile. The result showed that AF488-conjugated endogenous CBX4 co-localized with CF568-conjugated EZH2 and formed nuclear puncta. The result has been shown in **newly-added Figure EV2E**. The distribution and size of endogenous co-localized bodies were also similar to those of exogenous co-localized bodies (Figure 3F).

Fourthly, although CBX4 bodies were heterogeneous, these nuclear puncta tended to get close to or co-localize with HIV-1 proviruses in naïve J-Lat cells. In our revised manuscript, we have added more ImmunoFISH data for naïve and activated J-Lat cells. These data have been shown in **newly-added Figure EV2A and 2B**. Besides, we have measured the distances of the nearest CBX4 bodies and HIV-1 proviruses DNA in both J-Lat 10.6 and J-Lat 8.4 cells utilizing Imaris software. In each group, 15 cells were randomly imaged and measured the CBX4-HIV-1 distances. The results showed that CBX4 bodies tended to get close to or co-localize with HIV-1 proviruses in naïve J-Lat cells. Upon activation by TNFα, CBX4 bodies were away from HIV-1 proviruses, resulting in longer distances between the nearest CBX4 body and the HIV-1 provirus. These quantification data have been shown in **newly-added Figure EV2C and D**. The mechanisms why heterogenous CBX4 bodies were spatiotemporally close to HIV-1 proviruses in latent cells were complex. The 3D genomics and the epigenetic memory theory might explain the phenomenon, which was out of our study scope. Although the sizes, the distributions and the numbers of CBX4 bodies were variable in different cells, these bodies in live cells were dynamic and moved quickly (could internally diffuse, fuse and split). It was indeed very hard to figure out which body was responsible for HIV-1 silencing as the reviewer remarked. But our CBX4-HIV-1 distance quantification data reminded us that considerable amounts of CBX4 bodies were possibly and statistically close to and co-localized with HIV-1 proviruses in latent cells.

We hope that the reviewer would be satisfied with our above interpretation.

Figure 4. CBX4 bodies are phase-separated condensates
1,6 hexandiol co -staining for chromatin and viability should be included.

Reply: We thank the reviewer for the kind reminding. We have performed new disruption experiment with 1.5% 1,6-Hexanediol (1,6-Hex). We used Hoechst to dye DNA (also indicate chromatin) in live cells. Cells were imaged every 5 seconds and imaged for 2 minutes. The numbers of CBX4 bodies before 1,6-Hex treatment (0 min)

and after 1,6-Hex treatment (2 min) were calculated. These data have been shown in **revised Figure 4D** and **newly-added Figure EV3F**. Besides, we also measured the percentages of amine-reactive fluorescent dye non-permeant cells before 1,6-Hex treatment (0 min) and after 1,6-Hex treatment (2 min) to show cellular viabilities. The viability data have been shown in **newly-added Figure EV3G**.

EV3 is referring to this figure: There is a major mislabeling of the concentration-dependent LLPS formation, should be the other way around.
Reply: We are very sorry for the mislabeling. We have carefully checked the correspondence of the main text and figures and corrected any mislabeling. The NaCl concentration-dependent CBX4 proteins LLPS formation should be linked with **newly-added Figure EV4A**. The protein concentration-dependent LLPS formation should be linked with **newly-added Figure EV4B**. The LLPS formation abilities of GFP-CBX4mut proteins within different NaCl concentrations and different protein concentrations should be linked with **newly-added Figure EV4C and EV4D**.

Labelling is also random; the Authors should clarify construct labelling.
Reply: We are sorry for these mistakes. We have carefully checked our manuscript again and again. All the random labeling has been corrected to make sure that the order of each figure was consistent with the order of the experimental logic flow of our study.

Figure 5. CBX4 SUMOylates EZH2
Panel A) is not informative at all, as there is no any labelling included.
Reply: We are sorry for this omission. We used Figure 5A to show the overall result of CBX4 SUMO-MS experiment. The identified 1,928 SUMOylated candidates formed a large STRING network, which could be clustered into 12 major sub-clusters utilizing *k*-means clustering. Each color in Figure 5A indicated a different interconnected sub-cluster. We have **revised Figure 5A** by adding an essential title named "CBX4-SUMOylated proteins (12 sub-clusters)". Besides, we have added more descriptions within the main text, figure legend and methods of mass spectrometry (MS).

Figure 6. LLPS-deficient CBX4 bodies are unable to suppress HIV-1 and SUMOylate EZH2
Panel A) Luciferase change is modest, again. Panel B) SUMO4 WB should be repeated, as signal is not well resolved, membrane seems burnt.
Reply: We apologize for the deficiencies of our study and thank the reviewer for pointing out the above important problems.

For Figure 6A, we have added co-treatment data as we conducted in Figure 1A-C. The siCBX4-mediated enhancement of luciferase expression in TZM-bl cells was indeed modest while statistically significant (**Figure 6A**). The co-treatment of Tat transfection or TNFα stimulation with siCBX4 could induce much higher amounts of luciferase (from 2.87 fold to 6.77 fold for Tat co-transfection, from 2.87 fold to 4.76 fold). While the co-treatment of Tat transfection or TNFα stimulation did not influence wildtype CBX4-mediated rescue. These data have been shown in **newly-added Appendix Figure S6A and S6B**.

For Figure 6B, we have repeated this experiment which showed the dysfunction of CBX4mut on EZH2 SUMOylation. The **revised Figure 6B** has shown distinct bands clearly. The expression ratio of SUMO4-EZH2 also has been marked below the panel.

Figure 7. CBX4 depletion reactivates latent HIV-1 in cells from HIV-1-infected individuals.
The authors should show some quantification of viral mRNA, and not only use GFP as a measure for HIV-1 latency. It is only in FigC, D and E that they quantify HIV-1 mRNA through Vif (why Vif?).
Reply: We thank the reviewer for the kind suggestions. We also agreed with the reviewer that quantification of viral mRNAs would be more sensitive to indicate the reactivation of HIV-1 latency, while the percentage of GFP-positive cells only showed how many cells were reactivated but could not truly reflect the intensity of reactivation within a population of single cells. We have used primer pairs (HIVTotRNA Forward Primer: 5'-CTGGCTAACTAGGGAACCCACTGCT-3' and HIVTotRNA Reverse Primer: 5'-GCTTCAGCAAGCCGAGTCCTGCGTC-3') which targeted region from U5 to GLS (*Gag* leader sequence) to quantify intracellular HIV-1 RNAs. For reactivation in seven J-Lat cells (J-Lat 6.3, 8.4, 9.2, 10.6, 15.4, Mix and A2), we have provided more data of relative viral mRNA expression. These data have been shown in **newly-added Figure EV1E-K**. For reactivation in primary $CD4^+$ T cells, fold changes of HIV-1 mRNA expression at different time points within different groups were evaluated, which have been shown in **newly-added Appendix Figure S7A and S7E**. Our clinical samples data already have shown the intracellular HIV-1 RNAs. However, in our original manuscript, the targeted region was *Vif* mRNA which could easily acquire many different mutations during natural infection and latency formation. We agreed that *Vif* region was not a good choice, which might underestimate the true reactivation. Thus, we have re-performed these experiments in clinical samples utilizing the above primer pairs which targeted a more conserved region from U5 to GLS. We also used Lonza 4D-Nucleofector System to transfect siRNAs instead of lentiviruses-based shRNAs. Based on our experience, the knockdown efficiencies of siCBX4 by 4D-Nucleofector System were much higher than those of shRNAs in unactivated primary $CD4^+$ T cells. These data have been shown in **revised Figure 7E-G**.

Also as the cells were kept for prolonged time in culture, but no cell viability is shown. How did shCBX4 affect cell viability together with viral infection, as compared to the shControl?
Reply: We are sorry for the lack of these cellular viabilities data. During revision, we have evaluated cellular viabilities by measuring the percentages of amine-reactive fluorescent dye non-permeant cells with Zombie Violet Fixable Viability Dye (BioLegend). In the experiment of CBX4-mediated delay of HIV-1 latency, cellular viabilities of HIV-1-infected primary $CD4^+$ T cells on Day 0, 3, 6, 9, 12 and 15 were evaluated in both shluc and shCBX4 groups (**newly-added Appendix Figure S7D**). In the experiment of CBX4-mediated activation of HIV-1 latency, cellular viabilities in Mock group, αCD3/αCD28/IL-2 group, shluc group, shCBX4 group, shluc+JQ-1 group and shCBX4+JQ-1 group were evaluated at the same time (**newly-added Appendix Figure S7H**). In the study of patient samples, cellular viabilities of Mock group,

αCD3/αCD28/IL-2 group, siNC group and siCBX4 group were evaluated for all the three clinical samples (n**ewly-added Appendix Figure S7J-L**). The results showed that cellular viabilities in CBX4 KD groups (shCBX4 or siCBX4) were similar to those in control groups (shluc or siNC). CBX4 knockdown in primary $CD4^+$ T cells unlikely affected cellular viabilities.

Patient data are not convincing. As RNA levels are only doubled, and should be practically absent on ART, so doubling is only modest effect in this context. Raw dPCR data can clarify this. Or absolute quantification via ddPCR.

Reply: We are sorry for having not presented these data clearly. **Actually, the intracellular HIV-1 RNAs levels were presented as 10^3 (1000) copies per million $CD4^+$ T cells.** So the vertical axis-labeled copy numbers within each group should be multiplied by 1000. The "doubling" should be "2000 fold". For example, in Patient 1 samples, the average copy number of HIV-1 RNAs in Mock group was 15.06×10^3 copies per million $CD4^+$ T cells. The average copy number in αCD3/αCD28/IL-2 group was 93.71×10^3 copies per million cells. The average copy number in siNC group was 16.03×10^3 copies per million cells. The average copy number in siCBX4 group was 67.97×10^3 copies per million cells. So the copy number in αCD3/αCD28/IL-2 group was 6222 (6.22×10^3) fold of that in Mock group. The copy number in siCBX4 group was 4240 (4.24×10^3) times more than that in siNC group. As we have mentioned in the previous comments, we have **revised Figure 7E-G** by re-conducting clinical sample experiments. We used siRNAs targeting CBX4 in primary $CD4^+$ T cells utilizing Lonza 4D-Nucleofector System directly. The KD efficiencies and corresponding cellular viabilities have been shown in **newly-added Appendix Figure S7I-L**. The reviewer also kindly suggested us to use dPCR or ddPCR to quantify HIV-1 RNAs. Our data showed in revised Figure 7E-G were sufficient to prove that CBX4 depletion could reactivate latent HIV-1 from HIV-1-infected clinical samples. We hope the reviewer would allow us to keep Figure 7E-G.

Secondly, the reviewer also concerned that RNA levels in untreated samples should be practically absent on ART. However, based our experience and the communication with clinicians in the department of infectious diseases, ART can efficiently suppress the virus budding from reservoir cells to the plasma (HIV seronegative), but cannot totally suppress the transcription of latent HIV-1 within cells, resulting in the transcribing of intracellular HIV-1 RNAs. Our primer pairs (HIVTotRNA Forward Primer: 5'-CTGGCTAACTAGGGAACCCACTGCT-3' and HIVTotRNA Reverse Primer: 5'-GCTTCAGCAAGCCGAGTCCTGCGTC-3') amplified regions from U5 to GLS (*Gag* leader sequence), which could potentially represent the early and total transcripts of HIV-1. These primers would be more sensitive to detect residue viral RNAs. Thus, we were able to detect viral RNAs in untreated samples.

• *Scientific and general English language should be thoroughly corrected.*

Reply: We apologize for making many scientific and general language mistakes. We also thank the reviewer for pointing out the following omissions. We have read our revised manuscript again and again, and tried our best to correct all the language mistakes.

Minor points:

59. To what strategies do Authors refer saying 'and many others'
Reply: We apologize for this inadequate description. We have replaced "and many others" with actual strategies which included "**stem cell transplantation, CRISPR/Cas9 gene editing, immune checkpoint blockade, and control antigen-driven proliferation**".

62-66. Sentence is completely unclear, both conceptually and semantically
Reply: We are sorry for these unclear and redundant descriptions. We have replaced the original description with a more concise and comprehensive sentence, which is: "**Multiple molecular mechanisms involve in HIV-1 latency.**"

68 . What do Authors mean by 'verified latency-reversing agents'
Reply: Our original purpose was to indicated that some of the LRAs have been verified for their effectiveness on latency reversing. However, we realized that we have not used this word properly. Thus, in our revised manuscript, we have replaced this word with "**widely-used**".

75-78. this sentence should be re-written. It is unclear what 'lies' refers to
Reply: We are sorry for this confusing word. We have re-written this sentence as below: "**Simultaneously, many transcription suppressors including LSF, YY1, CTIP2, DSIF, NELF and TRIM28 enrich on the HIV-1 long terminal repeat (LTR) (Coull *et al*, 2000; He & Margolis, 2002; Ma *et al*, 2019; Marban *et al*, 2007; Ping & Rana, 2001). The 5' end of HIV-1 LTR contains the HIV-1 promoter and drives viral gene transcription.**"

79. what are epigenetic maintainers
Reply: We are sorry for this confusing word. Our original purpose was to refer to those proteins which maintained H3K9me3, H3K27me3 and H4K20me1 modifications. Obviously, the word "maintainers" have been wrongly used. Thus, in our revised manuscript, we replaced "epigenetic modifiers and maintainers" with a simple phrase "**epigenetic proteins**".

95-96. preposition, for example in is missing
Reply: We thank the reviewer for pointing out this omission. We have replaced "regulate" with "**participate in**".

108. The tense of the sentence is incorrect
Reply: Thanks for pointing out this mistake. We have re-written this sentence as below: "**Viral processes also have been found to be associated with LLPS.**"

125. preposition is missing, for example into
Reply: Thanks for pointing out this mistake. We have added "**into**" after the word "coalesces".

128. 'less defined' should state not defined
Reply: Sorry for this mistake. We have changed "less" with "**not**".

133. 'many cellular models' which one, how many. This is a scientific article, and if demonstrated should be stated accordingly, quantified.
Reply: We are sorry for this improper phrase. We have replaced "many cellular models" with "**seven cellular models (J-Lat 6.3, 8.4, 9.2 ,10.6, 15.4, Mix and A2)**".

Again, we sincerely thank all the editors and reviewers for pointing out these important details and giving us many constructive suggestions. We have corrected all the mistakes which the reviewers have reminded. We also have carefully read our revised manuscript many times to correct any typos and language issues. We hope that these changes are satisfactory.

Dear Prof. Ma

Thank you for the submission of your revised manuscript to EMBO reports. I apologize for the delay in handling your manuscript, but we have only now received the full set of referee reports that is copied below. In addition, I have discussed the reports further with the referees.

As you will see, the referees acknowledge that the revised manuscript has been strengthened. However, I also note that neither referee 2 nor referee 3 are fully satisfied with the revision. Many concerns relate to data presentation and discussion and can be addressed by clarification and textual changes. However, I also note that referee 2 raises concerns that some of the experiments have been repeated with different outcomes, such as the FRAP experiments in Figure 4A. In agreement with the referee, I would like to give you the chance to address these concerns. Please clarify why the results differed from the first version and please also include additional orthogonal experimental approaches to support the new findings.

From the editorial side, there are also a few things that need your attention:

- Appendix: please add page numbers to the file and to the table of contents.

- Appendix Figure legends: Please define whether the triplicate experiments (S1, S2, S5, S6, S7) or n=8 (S4) refer to technical or biological/independent replicates.
- Figure S4B: Please define the central band, boxes and whiskers of the boxplot.

- Please change the heading 'Declaration of interests' to 'Disclosure Statement and Competing Interests'. See also https://www.embopress.org/competing-interests for more information.

- Author contributions: Please discriminate the 2 "LWs"; one should be LiyW and the other LilW.

- Author checklist: Please complete the section on "Ethics, human participants and written consent" since it applies to your study.

- I attach to this email a related manuscript file with comments by our data editors. Please address all comments and upload a revised file with tracked changes with your final manuscript submission. I have also taken the liberty to make some changes to the Abstract to make it more accessible to our general readership.

We look forward to seeing a final version of your manuscript as soon as possible.

Yours sincerely,

Martina Rembold, PhD
Senior Editor
EMBO reports

****

Referee #1:

The authors responded adequately to my previous comments. I highly recommend the publication of the revised manuscript.

Referee #2:

The authors have performed considerable work to address my previous comments. However, on several occasions they did not address unexpected experimental results by orthogonal methods or biological reasoning, but simply repeated the exact same experimental protocol and reported different, sometimes opposite, results that fit the hypotheses. Examples are FRAP in Figure 4A where the recovery was reduced from >10 minutes to 30 seconds, Figure 4B where the change to a lower resolution microscope resulted in spherical droplets and the increased SUMOylation of EZH2 in figure 6B. Importantly, the authors often do not explain why the new results differ from the initially reported ones. I can only interpret this to mean that either the initial results were inaccurate, or the newly reported results are. I cannot reasonably distinguish between these two possibilities and can therefore only advise against publication of the manuscript in this form. Additional orthogonal data are needed to confirm the new results.

Referee #3:

In the new version of the manuscript the authors have performed and added a number of additional controls. Despite the fact that these new results have been added, the manuscript does not result to well curated. Listed below are some of the issues that could be further improved:

Starting with Line 79, the authors list different transcription repressors, but do not mention HIV-1 silencing by PML, mediated also by G9a. This HMT is mentioned in line 89, but in a different context. This is a bit surprising because PML involvement in HIV-1 silencing could further be elaborated in lines 104-105 where they talk about LLPS, as PML NBs were one of the first exemplary NBs shown to be formed through phase separation. Moreover, as they are also related to sumoylation, this could be further cited.

There are several recurring mistakes in the wording. For example, they say that CBX depletion mediated HIV-1 reactivation was enhanced much higher (line 173, 182, 189).

In line 199 the say that they inspected many modifications - this sounds like lame language, and should be corrected to be more scientifically appropriate.

Line 201: CBX4 depletion significantly decreases the enrichment - what exactly does that mean? CBX4 depletion results in significant occupancy ???

Lines 213,214, 215 would be more appropriate for discussion.

Line 220 : The absence of heterochromatin modification and the presence of transcription
machinery often indicates (rather that indicated) higher chromatin accessibility.

I continue to miss the point of ATAC-Seq - why have they done ATAC -Seq to look then only at the HIV-1 5' integration site (+/- 1kb)?

Line 264 they describe FISH in 10.6 JLats and mention that upon TNFa activation CBX4 bodies are observed at larger distance from HIV-1. This is again in line with what observed previously for PML NBs and HIV-1 (10.1016/j.chom.2013.05.016).

Line 298: In recent years (in is missing)

Line 504: patient data are clearly showing that there is a significant degree of variability between them, so it is a bit difficult to conclude that CBX4 depletion causes considerable reactivation of HIV1 in all patients. I would be more careful here as CBX4 is probably one of the possible mechanism that contributes to HIV-1 silencing and as such could have variable (and patient dependent) effects.

Line 529: Recent years - remove it, or add in.

General remark for the discussion : The authors comment that CBX4 contributes to HIV-1 latency by forming LLPS nuclear bodies and SUMOylating EZH2. However, they do not even hypothesize how is CBX4 recruited to the viral LTR promoter - what is causing the accumulation of CBX4 and deposition of H3K27me3. This could be further proposed in the discussion, with a goal of further dissection of the process.

Point-by-point response to each comment:

Referee #1:

The authors responded adequately to my previous comments. I highly recommend the publication of the revised manuscript.
Reply: We thank the reviewer for being satisfied with our revision and highly recommending the publication of our work.

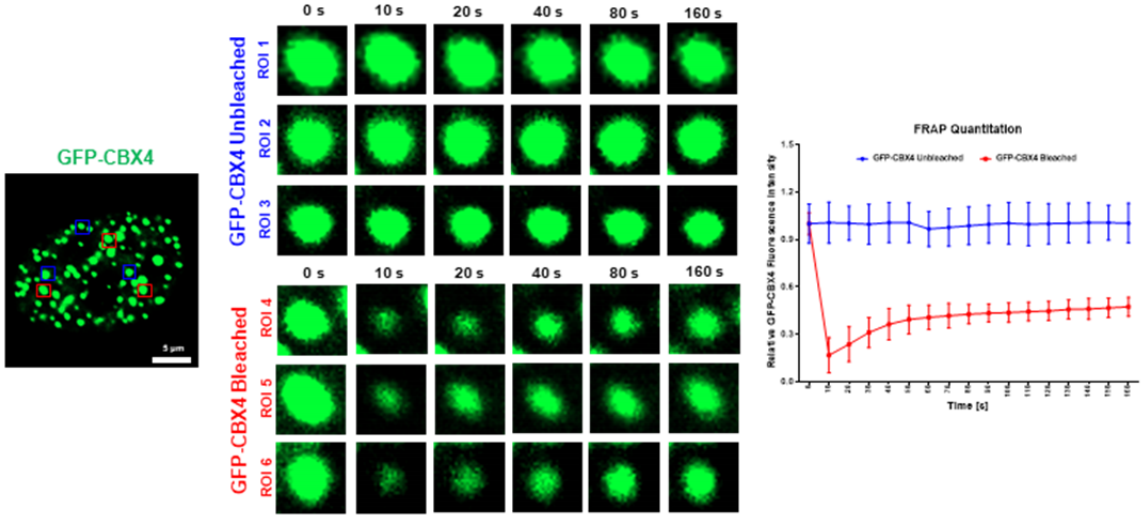
****

Referee #2:

The authors have performed considerable work to address my previous comments. However, on several occasions they did not address unexpected experimental results by orthogonal methods or biological reasoning, but simply repeated the exact same experimental protocol and reported different, sometimes opposite, results that fit the hypotheses. Examples are FRAP in Figure 4A where the recovery was reduced from >10 minutes to 30 seconds, Figure 4B where the change to a lower resolution microscope resulted in spherical droplets and the increased SUMOylation of EZH2 in figure 6B. Importantly, the authors often do not explain why the new results differ from the initially reported ones. I can only interpret this to mean that either the initial results were inaccurate, or the newly reported results are. I cannot reasonably distinguish between these two possibilities and can therefore only advise against publication of the manuscript in this form. Additional orthogonal data are needed to confirm the new results.
Reply: We thank the reviewer for pointing out the above deficiencies of our previous manuscript and responses. More importantly, we sincerely apologize for having not fully elucidated the new results and the improvement of experimental methods. During the first revision, both Reviewer 1 and Reviewer 3 suggested us to re-perform LLPS- and SUMOylation-related experiments, as they thought that the qualities of these figures should be improved. However, we sincerely apologize for having not explained the changes which we made for these experiments, which also have been kindly reminded by Reviewer 2 (you) during the first revision. We hope the reviewer would allow us to further explain these experiments and corresponding results. During the second revision, we also have performed some new experiments, collected more data which could represent the improvement of the previous experiments, and added significantly more details of these experiments in the main text, figure legends and methods. The reviewer has raised several major and important issues in the first revision, three of which were further concerned in this second revision. We would like to response to these suggestions according to our newly revised manuscript.

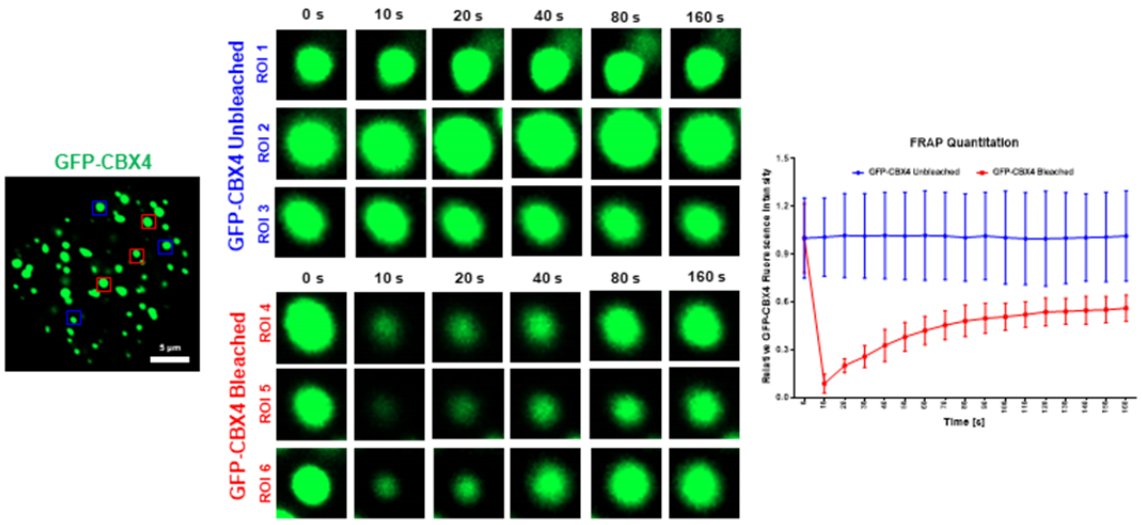
Firstly, the reviewer concerned about that FRAP results showed in Figure 4A were different from the results presented in the previous version. We apologize for not fully explaining the differences between these results. Our previous experiments showed that CBX4 bodies recovered slowly upon bleaching (50% within 10 min). While our new experiments showed that CBX4 bodies recovered quickly (50% within 30 s). The experimental procedure was indeed improved according to the reviewer's kind suggestion

that FRAP should be conducted in or close to physiological protein concentrations. Thus, in our new FRAP experiment, we expressed equal or lower amounts of GFP-CBX4 in HEK293T cells compared with the amounts of endogenous CBX4 proteins. In this condition, bleached CBX4 bodies were able to recover fluorescence intensities quickly, within almost 30 s (**revised Figure 4A**). While the FRAP data of previous Figure 4A were collected from cells overexpressing higher amounts of GFP-CBX4 proteins than physiological conditions, which resulted in slower recovery, over 10 min. We agree with the reviewer that high protein concentrations in our artificial overexpressing system could result in hardening of CBX4 condensates. We do think that extremely high amounts of CBX4 proteins would turn LLPS CBX4 bodies to solid aggregates, resulting in large immobile fraction. Indeed, our preliminary experiments have confirmed this hypothesis. When we overexpressed extremely high amounts of CBX4 proteins (2 μg GFP-CBX4-expressing plasmids in 35 mm confocal dish), over half of each CBX4 body was immobilized. These harden condensates recovered fluorescence intensities extremely slowly, resulting in less than 50% recovery within 3 min (**Author response imaging 1 and previous Figure 4A**). When we expressed a little lower amount of CBX4 proteins in HEK293T cells (1 μg GFP-CBX4-expressing plasmids in 35 mm confocal dish), bleached CBX4 bodies could recover 50% fluorescence intensities within 3 min (**Author response imaging 2**). Only when we expressed equal or lower amounts of GFP-CBX4 proteins (less than 500 ng GFP-CBX4-expressing plasmids in 35 mm confocal dish) compared with the amounts of endogenous CBX4 proteins, could bleached CBX4 bodies acquire almost full recovery (**Author response imaging 3 and current Figure 4A**). Besides, to further confirm that CBX4 bodies were indeed internally diffused in simulated physiological conditions, we conducted new FRAP experiments which bleached only half of one CBX4 body. These GFP-CBX4-expressing cells were expressed with equal or lower amounts of GFP-CBX4 proteins compared with the amounts of endogenous CBX4 proteins. The results showed that the bleaching not only quenched the fluorescence intensities of bleached parts, but also impaired the intensities of unbleached parts (**newly added Appendix Fig S4A**). The fluorescence intensities of both parts quickly recovered, within 60 s.

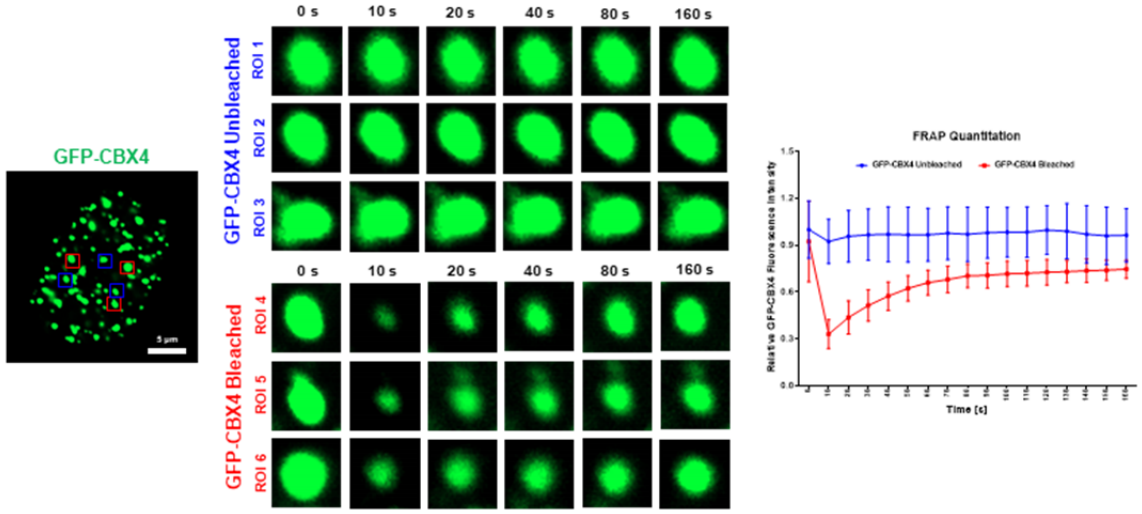


Author response imaging 1: HEK293T cells within 35 mm confocal dish were transfected with 2 μg GFP-CBX4-expressing plasmids. About 24 hours post transfection, three CBX4 bodies (ROI

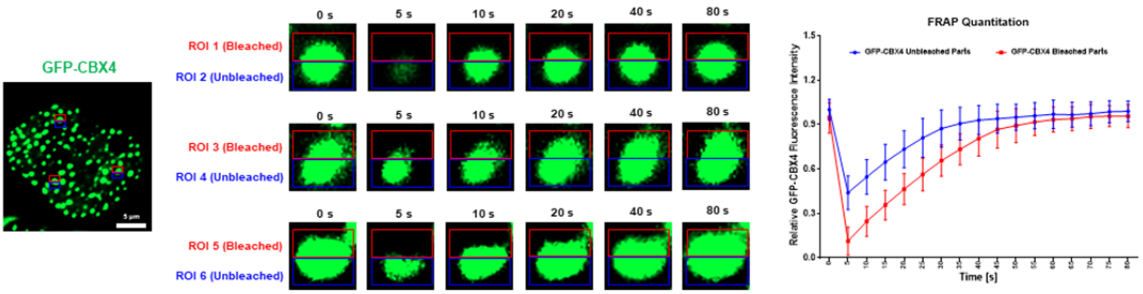
4, ROI 5 and ROI 6) were bleached with strong 488 nm laser pulse. Another three unbleached CBX4 bodies (ROI 1, ROI 2 and ROI 3) were set as control. The relative fluorescence intensities were measured every 5 s.



Author response imaging 2: HEK293T cells within 35 mm confocal dish were transfected with 1 μg GFP-CBX4-expressing plasmids. FRAP experiments were conducted as above.



Author response imaging 3: HEK293T cells within 35 mm confocal dish were transfected with 500 ng GFP-CBX4-expressing plasmids. FRAP experiments were conducted as above.



Newly added Appendix Fig S4A: HEK293T cells within 35 mm confocal dish were transfected with 250 ng GFP-CBX4-expressing plasmids. Three GFP-CBX4 bodies were split into two parts. Only half parts of each body were bleached and marked as ROI 1, ROI 3 and ROI 5. Another half parts were unbleached and marked as ROI2, ROI 4 and ROI 6. The relative fluorescence intensities were measured every 5 s

In our first revision, we also provided dual-color FRAP data for CBX4 and its partners (EZH2, RING1B and SUMO4). We found that both RFP-tagged proteins and GFP-CBX4 proteins within bleached CBX4 bodies quickly recovered fluorescence intensities (**current Figure EV3C-E**). Overall, our above results and analysis indicated that CBX4 bodies were characterized of both liquid phase and solid phase. In overexpression system, CBX4 bodies harbored large amounts of immobilized condensates, which resulted in the insensitivity of fluorescence recovery (slower and less recovery). While in physiological conditions or close to physiological conditions, CBX4 bodies were LLPS condensates which harbored large amounts of mobilized proteins. Further elucidating the function and characteristic of solid CBX4 bodies was beyond the scope of our study. We hope the reviewer would allow us to focus on the LLPS of CBX4 bodies.

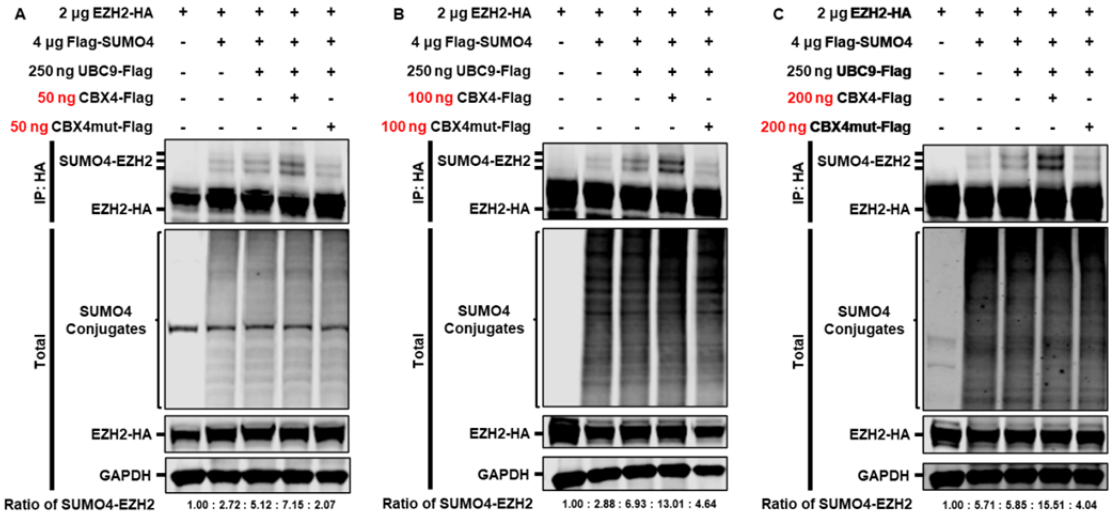
Secondly, the reviewer concerned about why we changed to a lower resolution microscope which resulted in more spherical droplets in Figure 4B-D. Our previous live cell imaging was conducted utilizing super-resolution SIM microscopy. While our current revised live cell images were acquired with Zeiss LSM 900 equipped with Airyscan 2 system. During our revision, we have consulted with technicians from both Nikon company (Dr. Jinli Lu, Dr. Pu Wang and Dr. Yan Duan) and Zeiss company (Dr. Yi Zhang). We also had close conversation with technicians from imaging platforms (Miss Yuanjun Guan from Sun Yat-sen University, and Mr Zepeng Guo from Guangdong Provincial People's Hospital). Actually, the lateral resolutions of both microscopy are not significantly different. The resolution of SIM is 115 nm. While the resolution of LSM 900 is 120 nm. The reason why we revised our live cell imaging utilizing LSM 900 refers to the higher photobleaching and phototoxicity of SIM. In our preliminary experiments, we have found that GFP fluorescence intensities within live cells could easily bleach during long-term imaging procedure (2 min to 10 min) upon capturing with SIM. While the photobleaching and phototoxicity of LSM 900 were weaker than SIM. Another reason that we chose LSM 900 refers to the imaging speed. Although SIM can present images with higher resolution, the imaging speed of SIM is very slow. To acquire a complete 3D-SIM image, fifteen images (5 phases and 3 angles) will be sequentially captured in each focal plane to acquire a complete image. That means if we want to capture a high-resolution SIM image, fifteen times more times will be needed. For example, if we set the exposure time with 300 ms, then we will need 4.5 s to acquire one SIM image. While we only need 300 ms for one LSM 900 image. For fixed cell imaging, it will be better to use SIM because of higher resolution. However, for live cell imaging, protein molecules within live cells are moving quickly during imaging. If CBX4 bodies move more quickly than the imaging speed, then the reconstructed SIM images will present mottled pattern. Our preliminary experiments have shown that LSM 900 imaging will be quicker than SIM and avoid the presence of mottled pattern. Another small reason that we utilized LSM 900 to conduct FRAP experiments is our SIM imaging platform has not been equipped with FRAP module. While LSM 900 are equipped with FRAP. Thus,

we think that the imaging platform differences may influence the results of different image pattern. For fixed cells, we still used high-resolution SIM. While for live cells, we prefer to use LSM 900. We hope the reviewer would allow us to use different imaging platforms in different experiments.

Thirdly, the reviewer concerned about that current data of SUMOylation assays were a little different from previous data, especially data showed in Figure 6B, 6D and 6E. We sincerely apologize that we have not clearly explained the differences of these experiments. The reviewer concerned about that there is no reduction in EZH2 SUMOylation compared both CBX4-OE and CBX4mut-OE groups in Figure 6D. However, our previous data showed in Figure 6D indeed showed that the overexpression of CBX4mut reduced the SUMOylation of EZH2 compared with wildtype CBX4. We apologize that we did not show the exact ratios of SUMOylated EZH2 within different groups. Besides, the SUMOylation signals of previous data were not well-resolved. Both Reviewer 1 and Reviewer 3 also suggested us to re-perform this experiment and conduct quantitative analysis. During our revision, we have performed this SUMOylation assay for many times. Based on our previous experience, the amounts of SUMO E3 ligase have significant influence on the SUMOylation percentages of substrates (Xiancai Ma *et al.*, 2019, PMID: 30652970). Thus, we have conducted many SUMOylation assays in cells transfected with different amounts of CBX4- or CBX4mut-expressing plasmids (50 ng, 100 ng, 200 ng and 250 ng). Our preliminary data showed that the percentages of SUMOylated EZH2 would be higher upon transfecting more CBX4 plasmids (**Author response imaging 4 and current Figure 6B**). While the percentages of SUMOylated EZH2 were almost unchanged upon transfecting different amounts of CBX4mut plasmids. The ratios of SUMO4-EZH2 within each group have been shown below each figure. For 50 ng CBX4/CBX4mut experiment, the ratio was 1.00 : 2.72 : 5.12 : 7.15 : 2.07 (**Author response imaging 4A**). For 100 ng CBX4/CBX4mut experiment, the ratio was 1.00 : 2.88 : 6.93 : 13.01 : 4.64 (**Author response imaging 4B**). For 200 ng CBX4/CBX4mut experiment, the ratio was 1.00 : 5.71 : 5.85 : 15.51 : 4.04 (**Author response imaging 4C**). For 250 ng CBX4/CBX4mut experiment, the ratio was 1.00 : 5.33 : 14.11 : 22.03 : 4.70 (**current Figure 6B**). The reason why we have not transfected much higher CBX4- and CBX4mut-expressing plasmids is that we still wish to mimic or be close to physiological conditions instead of high overexpression conditions. All of our preliminary data, previous data showed in Figure 6B and current data showed in Figure 6B indicate that CBX4-mediated EZH2 SUMOylation has strong positive correlation with its body formation capability. To facilitate other researchers to repeat our work and conjugate to their own work, we have listed the detailed transfection strategies and protocols within figure legends and methods.

Finally, we would like to further explain the new results showed in revised Figure 6D and 6E. Our previous data showed in Figure 6D were indeed not well-performed. As the reviewer commented in the first revision, we did not show the percentages of SUMOylated EZH2 or the contribution of CBX4 on H3K27me3 modification. Both data should be included to fully support the claim that CBX4 and increased SUMOylated EZH2 tied to the increased level of H3K27me3. Thus, in our revised Figure 6D, we have re-performed this experiment by including CBX4-OE group. Besides, we also showed the bands of SUMOylated EZH2. We also conducted statistical analysis of SUMO4-EZH2 and H3K27me3. The detailed transfection strategies have been shown in the

corresponding figure legends and methods. The experiments to output previous Figure 6E have not been well-performed either. Reviewer 1 also kindly suggested us to re-perform this experiment by adding the following modifications: using highly purified recombinant mononucleosomes instead of Histone H3.1 only, measuring the signals and percentages of SUMOylated EZH2, introducing EZH2 catalytic mutant group, and introducing H3 K27R mutant group. Reviewer 2 (you) also kindly suggested us to show the signals and percentages of SUMOylated EZH2. Based on all the above suggestions, we re-performed this *in vitro* methyltransferase assay and set more control groups. Detailed *in vivo* transfection and *in vitro* incubation strategies have been shown in corresponding figure legends and methods. As the experimental strategies have been improved to consider more control conditions, we think that our new results showed in Figure 6B, 6D and 6E were more accurate than previous immature results. We hope the reviewer would allow us to add these improved modifications and be satisfied with the above responses.



Author response imaging 4: HeLa cells within 6 cm dishes were transfected with 2 μg of HA-tagged EZH2, 4 μg of Flag-tagged SUMO molecules, 250 ng of E2 UBC9, and/or different amounts of E3 CBX4/CBX4mut (ranging from 50 ng to 200ng). The ratios of SUMO4-EZH2 were shown below each panel.

****

Referee #3:

In the new version of the manuscript the authors have performed and added a number of additional controls. Despite the fact that these new results have been added, the manuscript does not result to well curated. Listed below are some of the issues that could be further improved:

Reply: We thank the reviewer for the support of our study and pointing our many important issues of our revised manuscript. In the new version, we have carefully addressed each comment and revised our manuscript accordingly. Besides, we also have corrected many more mistakes which we made in our previous manuscript.

Starting with Line 79, the authors list different transcription repressors, but do not mention HIV-1 silencing by PML, mediated also by G9a. This HMT is mentioned in line

89, but in a different context. This is a bit surprising because PML involvement in HIV-1 silencing could further be elaborated in lines 104-105 where they talk about LLPS, as PML NBs were one of the first exemplary NBs shown to be formed through phase separation. Moreover, as they are also related to sumoylation, this could be further cited.
Reply: We sincerely apologize for these omissions. We agree with the reviewer that we should introduce PML bodies-mediated HIV-1 latency. Thus, in Line 80 to 83, we added PML as the transcription suppressors, which was modified as below: "**Simultaneously, many transcription suppressors including LSF, YY1, CTIP2, DSIF, NELF, PML and TRIM28 enrich on the HIV-1 long terminal repeat (LTR) (Coull *et al*, 2000; He & Margolis, 2002; Lusic *et al*, 2013; Ma *et al*, 2019; Marban *et al*, 2007; Ott & Verdin, 2013; Ping & Rana, 2001).**"
In Line 88 to 95, we have mentioned PML-mediated HIV-1 silencing, which was modified as: "**Deacetylated H3K9 is methylated by histone methyltransferases SUV39H1, G9a and GLP, which shape the suppressive marks H3K9me2 and H3K9me3 on the HIV-1 LTR (Chéné *et al*, 2007; Ding *et al*, 2013; Imai *et al*, 2010). Particularly, the transcription suppressor PML protein can orchestrate the G9a methyltransferase to bind to the latent HIV-1 promoter and shape the H3K9me2-containing facultative heterochromatin on the proviral DNA (Lusic *et al.*, 2013; Ott & Verdin, 2013). H3K9me2 and H3K9me3 are further maintained by heterochromatin protein 1α (HP1α), HP1β and HP1γ (Chéné *et al.*, 2007).**"
When introducing LLPS in Line 114 to 118, we added the following discussion: "**The PML nuclear body, which is also one of the HIV-1 latency contributors, is one of the first exemplary nuclear bodies shown to be formed through phase separation (Banani *et al.*, 2016; Lusic *et al.*, 2013). Further biochemical analysis indicates that PML forms LLPS nuclear bodies through SUMOylation-mediated SUMO-SIM multivalent interactions (Banani *et al.*, 2016; Corpet *et al*, 2020).**"

There are several recurring mistakes in the wording. For example, they say that CBX depletion mediated HIV-1 reactivation was enhanced much higher (line 173, 182, 189).
Reply: We are sorry for these lame languages. We have modified "CBX4 depletion-mediated HIV-1 reactivation was enhanced much higher" with a simple and professional phrase "**The HIV-1 reactivation was more significant**" in Line 173. In Line 182, the phrase "The reactivation was enhanced much higher" was replaced with "**The reactivation was more significant**". In Line 189, the phrase "The reactivation was further enhanced much higher" was replaced with "**The reactivation was more significant**". In Line 190, the phrase "CBX4 depletion-mediated HIV-1 latency activation in these cell lines" was replaced with "**The latent HIV-1 reactivation which mediated by knocking down CBX4 in these cell lines**".

In line 199 the say that they inspected many modifications - this sounds like lame language, and should be corrected to be more scientifically appropriate.
Reply: We are very sorry for this lame description. In our revised manuscript, we have changed this sentence with a more scientifically appropriate one, which is: "**we evaluated the alterations of histone modifications upon CBX4 knockdown**".

Line 201: CBX4 depletion significantly decreases the enrichment - what exactly does that mean? CBX4 depletion results in significant occupancy ???
Reply: Sorry for this confusing description. We have modified the original sentence with a more appropriate one, which is: "**The knockdown of CBX4 resulted in the loss of CBX4 proteins on the HIV-1 LTR**".

Lines 213,214, 215 would be more appropriate for discussion.
Reply: We thank the reviewer for pointing out the deficiency of the discussion in Line 213, 214 and 215. In our revised manuscript, we have added more descriptions and cited more appropriate papers. The new discussion is modified as below: "**The H2AK119Ub modification, which was catalyzed by RING1B, also has been found to contribute to HIV-1 latency (Khan *et al.*, 2018; Yoon *et al*, 2014). The downregulation of H2AK119Ub on the HIV-1 LTR upon knocking down CBX4 indicated that CBX4 might recruit RING1B and establish the suppressive H2AK119Ub modification on the HIV-1 LTR. While the active mark H3K27Acetyl was mainly deacetylated by HDAC1 and HDAC2 on the HIV-1 LTR (Marban *et al.*, 2007). The stabilization of H3K27Acetyl on the HIV-1 LTR upon CBX4 depletion indicated that CBX4 might not cross-talk with histone acetylation and corresponding deacetylases.**"

Line 220: The absence of heterochromatin modification and the presence of transcription machinery often indicates (rather that indicated) higher chromatin accessibility.
Reply: We apologize for this mistake. We have changed "indicated" with "**indicate**" in our revised manuscript. We think that "indicate" would be more proper than "indicates", as the subject contains two mechanisms which are "the absence of heterochromatin modification" and "the presence of transcription machinery".

I continue to miss the point of ATAC-Seq - why have they done ATAC -Seq to look then only at the HIV-1 5' integration site (+/- 1kb)?
Reply: We are sorry for having not explained this issue clearly in our previous revision. The reason why we conducted ATAC-Seq to probe the chromatin accessibility of HIV-1 LTR was to indicate that the HIV-1 LTR was accessible for RNAP II upon CBX4 knockout. Our ChIP-qPCR assays have confirmed that the depletion of CBX4 was able to remove heterochromatin modifications and recruit transcription machinery on the HIV-1 LTR which contains the HIV-1 promoter. Many previous reports also showed that the accessibility of the HIV-1 LTR was significantly influenced by the specific integration sites, the specific host factors and the presence of LRAs (Conrad et al, 2017, PMID: 28844864; Dupont et al, 2021, PMID: 33811831; Einkauf et al, 2022, PMID: 35026153; Jefferys et al, 2021, PMID: 33635929). By probing the chromatin accessibility of the whole HIV-1 genome, the alteration of chromatin accessibility of the HIV-1 LTR was more significant and more important to evaluate the contribution of host factors and the potency of LRAs. In our previous reports on HIV-1 latency contributors, we also have found that TRIM28 and CAF-1 were able to restrict the chromatin accessibility of the HIV-1 LTR, which significantly suppressed HIV-1 expression (Xiancai Ma *et al.*, 2019, PMID: 30652970; Xiancai Ma *et al.*, 2021, PMID: 33739466). Indeed, our ATAC-Seq data could be further annotated to fully elucidate the contribution of CBX4 on other host genes besides HIV-1 (PRJNA797956). However, we hope the reviewer would allow us to

focus on the contribution of CBX4 on the chromatin accessibility of HIV-1 LTR only, which is also the major region targeted by CBX4. To make the logic flow more reasonable, we have added the following sentences to bridge ChIP-qPCR results and ATAC-Seq results: "**The absence of heterochromatin modification and the presence of transcription machinery often indicate higher chromatin accessibility. The integration sites, the removement of specific host factors and the presence of LRAs have been found to alter the chromatin accessibility of the HIV-1 LTR which contains the HIV-1 promoter (Conrad et al, 2017; Dupont et al, 2021; Einkauf et al, 2022; Jefferys et al, 2021). Thus, we conducted ATAC-Seq……**"

Line 264 they describe FISH in 10.6 JLats and mention that upon TNFa activation CBX4 bodies are observed at larger distance from HIV-1. This is again in line with what observed previously for PML NBs and HIV-1 (10.1016/j.chom.2013.05.016).
Reply: We apologize for the lack of proper discussion and citations. In our revised manuscript, we have specifically mentioned the similarities of CBX4 bodies and PML bodies, which is modified as below in Line 273 to 283: "**We next investigated the positions of CBX4 bodies and HIV-1 proviruses in different stimulation conditions utilizing immunolabeling-based fluorescence in situ hybridization (ImmunoFISH) assay (Lusic *et al.*, 2013; Marini *et al*, 2015). We found that the HIV-1 genomic DNA was close to or co-localized with CBX4 bodies in naïve J-Lat 10.6 and 8.4 cells (Figs 3E, EV2A and EV2B). Upon TNFα stimulation, the HIV-1 genomic DNA was away from CBX4 bodies (Figs 3E and EV2A-D). These results are in line with what observed previously for PML nuclear bodies and latent HIV-1 proviruses (Lusic *et al.*, 2013; Ott & Verdin, 2013). The latent HIV-1 provirus resides in close proximity to PML bodies. While TPA treatment results in the progressive displacement of the HIV-1 provirus from the PML bodies.**"

Line 298: In recent years (in is missing)
Reply: We apologize for this omission. We have added "**in**" before "recent years" in our revised manuscript.

Line 504: patient data are clearly showing that there is a significant degree of variability between them, so it is a bit difficult to conclude that CBX4 depletion causes considerable reactivation of HIV-1 in all patients. I would be more careful here as CBX4 is probably one of the possible mechanisms that contributes to HIV-1 silencing and as such could have variable (and patient dependent) effects.
Reply: We thank the reviewer for pointing out this deficiency. We also agree with the reviewer that the LLPS of CBX4 should be only one of the possible mechanisms that contributes to HIV-1 latency, which results in the variability of HIV-1 reactivation in different patient samples. Our patient samples data also indeed show significant degree of variability between them. Thus, we have modified the original descriptions as below: "**We found that the depletion of CBX4 was able to reactivate considerable amount of HIV-1 RNAs in Patient 1 samples, while the absence of CBX4 only induced slight increase of intracellular HIV-1 RNAs in samples from Patient 2 and Patient 3 (Figs 7E-G, and Appendix Fig S7I-L). Other possible mechanisms besides CBX4 might**

also contribute to HIV-1 silencing in different patient samples, which could result in patient-dependent variabilities of HIV-1 reactivation."

Line 529: Recent years - remove it, or add in.
Reply: Apologize for this mistake. We have added "**in**" before "recent years".

General remark for the discussion : The authors comment that CBX4 contributes to HIV-1 latency by forming LLPS nuclear bodies and SUMOylating EZH2. However, they do not even hypothesize how is CBX4 recruited to the viral LTR promoter - what is causing the accumulation of CBX4 and deposition of H3K27me3. This could be further proposed in the discussion, with a goal of further dissection of the process.
Reply: We thank the reviewer for this kind suggestion. We also agree with the reviewer that we should discuss more on how CBX4 is recruited to the HIV-1 promoter. Thus, in Line 582 to 604, we hypothesize four possible mechanisms, which is shown as below: "**Previous reports have shown that PRC2 subunit EZH2-mediated H3K27me3 can act as docking sites for CBX4 of PRC1 (Guo *et al.*, 2021). Thus, the H3K27me3 modification may serve as the initial signal for the accumulation of CBX4 on the HIV-1 LTR. Our data show that CBX4 can mediate positive feedback to PRC2 by SUMOylating EZH2 to enhance H3K27me3 activity of EZH2. The collaboration of CBX4 and EZH2 can orchestrate both PRC1 and PRC2, which significantly maintains the H3K27me3-containing fHC on target genomic DNA including the HIV-1 promoter. Other proteins may recruit CBX4 on the HV-1 LTR directly. YY1 can recruit PRC1 to its DNA binding site by physically interacting with YAF2 and CBX proteins (Basu *et al*, 2013). While YY1 has been found to bind to the HIV-1 LTR directly and contribute to HIV-1 latency (Bernhard *et al*, 2013; Coull *et al.*, 2000). Interestingly, SUV39H1 proteins, which catalyze H3K9me3 modifications on the HIV-1 LTR, can also methylate CBX4 and recruit PRC1 to cellular gene promoters, which represses the expression of target genes (Sewalt *et al*, 2002; Yang *et al*, 2011). We hypothesize that CBX4-containing PRC1 suppressive complex may be recruited by SUV39H1 to the latent HIV-1 promoter as well. CpG islands (CGIs) have been found to act as polycomb response elements (PREs) and be recognized by both PRC1 and PRC2 (Ku *et al*, 2008). Another report shows that KDM2B can recognize non-methylated DNA in CGIs and recruit PRC1 to these regions (Farcas *et al*, 2012). We hypothesize that CBX4 and corresponding PRC1 may also be recruited by KDM2B to the HIV-1 LTR and the transcription start site which harbor two CGIs (Blazkova *et al.*, 2009; Kauder *et al.*, 2009). Future work should focus on the exact mechanisms of how CBX4 is recruited to the HIV-1 promoter.**"

Again, we sincerely thank all the editors and reviewers for pointing out the above important details and giving us many constructive suggestions. We have addressed all these important issues which the reviewers have suggested in our revised manuscript. In our new revision, we also have carefully read our revised manuscript many times to correct any typos and language issues. We hope that these changes are satisfactory.

Prof. Xiancai Ma
Institute of Human Virology, Sun Yat-sen University
No. 74, Zhongshan Road 2
Guangzhou, Guangdong 510080
China

Dear Prof. Ma,

As I informed you, we have meanwhile received the report from referee #2, who is very positive about the final revisions and supports publication. Thank you for sending the modified manuscript text and Appendix and for introducing the few minor changes that were needed. I have uploaded these files for you.

I am now 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]. 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." Please note that the author checklist will still be published even if you opt out of the transparent process.

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,

Martina Rembold, PhD
Senior Editor
EMBO reports

****

THINGS TO DO NOW:

Once your article has been received by Wiley for production, the corresponding author will receive an email from Wiley's Author Services system which will ask them to log in and will present them with the appropriate license for completion.

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-2021-53855V3 and be addressed to emboreports@wiley.com.

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

EMBO Press Author Checklist

Corresponding Author Name: Hui Zhang and Xiancai Ma
Journal Submitted to: EMBO reports
Manuscript Number: EMBOR-2021-53855V3

USEFUL LINKS FOR COMPLETING THIS FORM

The EMBO Journal - Author Guidelines
EMBO Reports - Author Guidelines
Molecular Systems Biology - Author Guidelines
EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January 2022)

This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your manuscript.

Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for 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.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots 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. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship 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 χ2 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.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Yes	Data Availability Section

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Yes	Materials and Methods

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Yes	Materials and Methods

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Yes	Meterials and Methods
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Yes	Meterials and Methods
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Yes	Meterials and Methods

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Not Applicable	

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	
Microbes: provide species and strain, unique accession number if available, and source.	Yes	Materials and Methods

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Not Applicable	

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Not Applicable	

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Not Applicable	
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Materials and Methods, Figure Legends
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Yes	Materials and Methods, Figure Legends
Include a statement about **blinding** even if no blinding was done.	Yes	Materials and Methods, Figure Legends
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Yes	Materials and Methods, Figure Legends
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Materials and Methods, Figure Legends

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Yes	Materials and Methods, Figure Legends
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Materials and Methods, Figure Legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Yes	Materials and Methods
Studies involving **human participants**: 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.	Yes	Materials and Methods
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	
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.	Not Applicable	
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.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Data Availability Section
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Not Applicable	
If publicly available data were reused, provide the respective **data citations in the reference list**.	Not Applicable	