nature portfolio

Peer Review File

Virus-Human Chromatin Interactions Reorganise 3D Genome and Hijack KDM5B for Promoting Metastasis in Nasopharyngeal Carcinoma

Corresponding Author: Dr Wei Dai

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Chung et al has leveraged several technologies to examine the potential links between EBV positivity and epigenetic dysregulation in nasopharyngeal carcinoma, where EBV infection is prevalent. The authors have assembled an impressive amount of data from several cell lines, PDXs and primary samples. However, I the manuscript requires major improvement in data presentation, interpretation, and contextualization before this reviewer could provide any detailed critique of the content. Here are a few snippets of major shortcomings of the submitted manuscript and by no means reflects a point-to-point examination of the manuscript, which is unfortunately impossible given the current state of manuscript.

There are many inconsistencies in analysis. Many figures are not interpretable. Unites are missing, etc. This reviewer cannot evaluate the authors claim from the current figures. For examples here are a few items related to figures 1 and S1: It is unclear why active histone mark is equally enriched in A and B compartments of several cell lines in figure S1D. Where is NPC43neEBV in figures 1A and 1B? In tracks shown in Figure 1C, NPC53 is similar to NPC43. How the authors claimed these are regions with compartment switching? Could the authors explain the color-coding used for the heatmaps. What is data Xeno76, which is listed in a figure legend? The number of EBV+ lines/PDXs does not match across figure panels. As I mentioned earlier, providing a comprehensive list is beyond the scope and must be addressed by the authors before any submission.

The authors should faithfully state the conclusions of cited works. For instance, the authors' statement that "reduction in the number of chromatin loops could decrease topologically associating domains (TADs) and compartment formation [11]," has nothing to do with the data in reference (11).

This reviewer cannot fully examine scientific merit of the manuscript in its current form. There are many unsubstantiated statements. For instance, what is "Modified Hi-C experiment" (it is just a commercial kit!). What do the authors mean by "The EBV episomes bind the KDM5B promoter region with enhancer activation, forming cis-regulatory elements; this region gains a new chromatin loop and neoTAD, resulting in increased protein expression"? The sentence does not make any sense. What is "EBVIR-KDM5B-enhancer" signature, how was it defined? Again, providing a list of all the issues is beyond the scope of a review process.

It is impossible to understand the details of computational analysis given lack of details in Method section. There is no statistics. Again, using Figures 1 and S1 as an example, from the Methods is not possible to understand how A-to-B or B-to-A compartments are defined. Did the author define these loci per cell line/PDX or used a kind of consensus criteria? What threshold was used? Again, I cannot provide a point-to-point critique within a reasonable review.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career

Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
In this intriguing paper by Dai and colleagues, the investigators show how the EBV genome/DNA induces spatial reorganization of the human 3D genome through direct chromatin interactions, which affects chromatin loops, histone modifications and enhancer activation or inactivation at specific regions of the genome. More importantly, they highlight that through this mechanism, EBV hijacks KDM5B to promote aberrant gene expression that may promote cancer progression and metastasis. Lastly they use this data to support the notion that this could result in a better signature to predict metastases in NPC.
NPC is an EBV-driven cancer endemic in East and South-east Asia, and there has been a number of important finings in the last 10 years that have allowed a more streamlined approach to its management. Yet the fundamental/biological basis for carcinogenesis remains poorly understood, principally how a disease which is genetically bland, remains aggressive through epigenetic mechanisms. This publication certainly adds a lot of weight to a rapidly developing approach to looking at 3D chromatin conformation beyond just the histone code. Having said that, the absolute novelty of this paper is somewhat tempered by a similar paper (which they cite as reference 9) where a similar approach is shows in EBV-associated gastric cancer. In fact much of the analyses, figures and presentation format for Figures 1-3 follow the format of this Nature Genetics paper. This is not necessarily a criticism but perhaps accounts for why this paper has been bumped to Nat comms rather than Nature or Nature Genetics (although I have to say that the figures in the current paper are actually more refined than the Nature Genetics paper!). One advantage that the prior paper had which was not used here, was a very basic step-by-step explanation of the techniques used, why and what to expect as a result. Given that Nature Comms is a multi-D journal, this would be useful to the broad readership.

The major finding in my opinion here is the role of KDM5B (and perhaps Myc to a lesser extent), and the application of this to a well-curated dataset (which I assume has not been published before). There is a number of circumstantial evidence supporting its role in cancer progression in their models, but I feel that the authors should consider a functional analyses using knock-down or CRSPR to abrogate the function of KDM5B, to study the phenotype and whether there is a rescue in this instance. I would also like to see validation of the role of KDM5B across other EBV associated cancer types (lung, gastric, lymphoma). This would be important to address if KDM5B is context specific or more broadly associated with EBV-DNA interactions. Either result would be intriguing.

For the clinical section, it would also be important to address how KDM5B and EBV-DNA post treatment can be combined as a prognostic model to guide therapy and perhaps in the discussion even address whether KDM5B inhibitors may have an important role to play (again this depends on the functional analyses above).

The main figures and table are all appropriate and well formatted and I especially liked the word-cloud approach to showing functional terms in Figure 5D. The authors also need to make the clinical data/genomics data available which was not mentioned in the data availability statement.

Reviewer #4

(Remarks to the Author)
Chung et al. Cross-species Chromatin Interactions Reorganize the Human 3D Genome and Hijack KDM5B, a Risk Factor for Distant Metastasis, in Epstein–Barr Virus-associated Nasopharyngeal Carcinoma (NPC)

Chung et al. explore the role of Epstein-Barr virus (EBV) in the reorganization of the human 3D genome in nasopharyngeal carcinoma (NPC). It reveals that EBV–host chromatin interactions disrupt genome architecture, enhancing compartment switching and looping changes that trigger “enhancer infestation” and the activation of cancer-related genes. Specifically, the authors report that EBV hijacks KDM5B, a key protein associated with chromosomal stability, fostering an environment conducive to cancer cell proliferation, increased DNA damage response, and metastatic potential. The study introduces the EBVIR-enhancer-KDM5B signature as a predictive marker for distant metastasis and suggests its clinical relevance in risk assessment and prognosis for NPC patients.

Comments:

1. In Results Section 1, the authors should clearly explain how they define “active” vs. “inactive” compartments. Although partially explained in the Extended Data Figure 1 legend, it is unclear from the text how this analysis of H3K4me3 marks was performed. Also, the phenomenon of compartment switching in EBV-positive samples should be better explained.

2. The authors mention in lines 136-138: “The presence of H3K27me3 at only these select regions may form an epigenetic environment in cells that is favourable for EBV maintenance, and these cells may therefore be selected for during latent EBV infection”. I do not understand the basis of this statement, and this seems like mere speculation without any corroborating data. One suggestion would be to knock down the EZH2 enzyme (responsible for H3K27me3 methylation) in EBV-positive NPC cells and quantifying EBV copy number using qPCR at various time points post-knockdown to evaluate if reducing H3K27me3 affects EBV maintenance. Another option would be to use CUT&RUN or ChIP-qPCR to assess the binding of latency-associated EBV proteins (e.g., EBNA1) to H3K27me3-enriched regions in NPC cells.

3. More detail on the spatial transcriptomics pipeline is warranted. How was the data normalized across samples? How was the cellular deconvolution performed? Was single-cell data integrated to identify cell types?

4. It is unclear what information the bar plot in Figure 4b is trying to convey. I understand that the graph depicts EBVIR-enhancer-KDM5B pathway activity, but missing labels on the Y-Axis make it difficult to interpret.

5. Section 3 of Results, the authors' conclusions in this section rely heavily on speculative associations between EBV-induced enhancer activation, the KDM5B gene, and metastatic potential in NPC cells. Although they observe upregulation of KDM5B and associated enhancer activity in EBV-positive cells, there is limited experimental evidence to directly link these findings to functional outcomes such as increased cell proliferation, immune evasion, or metastatic behavior. To make these findings more credible, additional validation is needed, such as functional assays to confirm that the observed epigenetic changes directly influence these aggressive phenotypes in NPC cells. Experimental interventions like KDM5B inhibition or enhancer blockade in EBV-positive cells could help demonstrate causation rather than correlation, strengthening the biological relevance of the EBVIR-enhancer-KDM5B signature in NPC progression.

6. In Section 4 of Results, there is very little information in the section regarding the number of patients accrued in each centre. I understand that this information can be found in Table 1, but this should be included in the text as well. How did the authors assemble the genes in the EBVIR-enhancer-KDM5B signature? Although may be out of scope for this study, it will interesting to perform IHC for KDM5B on tissue microarrays assembled from patient samples from the multi-centre cohort to investigate if aberrant KDM5B protein expression could be used as a more simple risk factor for NPC metastasis in the future.

Minor comment: Some of the text on figures is to small to read.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Instead of improving the overall quality of the manuscript, the authors addressed only the specific examples I highlighted in my original review.

Nevertheless, the data presented in this manuscript could be valuable to the community. I recommend its publication as a resource paper, contingent on the availability of all raw data through a user-friendly and publicly accessible data portal.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
Thanks for the response. The functional and orthogonal analyses here certainly adds more weight to the data and validates the claims here. I am satisfied with the current version and support its publications. Congratulations to the authors for a fantastic study.

Reviewer #4

(Remarks to the Author)
I have reviewed the responses the authors have provided to my comments and I haver no further concerns regarding the publication of this study.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

We sincerely thank the reviewers' thoughtful comments and constructive feedback, which have significantly improved our manuscript. We have carefully addressed all concerns in the revised version and provided detailed point-by-point responses below. We believe these revisions have enhanced the quality and clarity of our work, and we hope the manuscript receives a favourable review and meets the publication standards.

Reviewer #1, functional and computational epigenomics and cancer (Remarks to the Author):

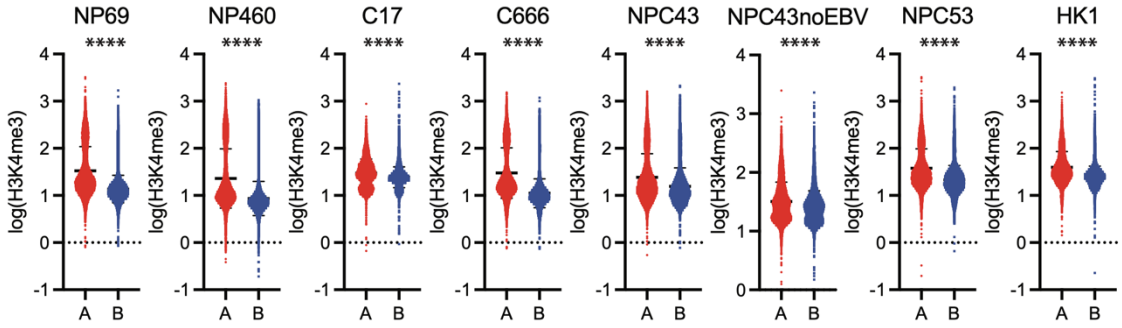
Chung et al has leveraged several technologies to examine the potential links between EBV positivity and epigenetic dysregulation in nasopharyngeal carcinoma, where EBV infection is prevalent. The authors have assembled an impressive amount of data from several cell lines, PDXs and primary samples. However, I the manuscript requires major improvement in data presentation, interpretation, and contextualization before this reviewer could provide any detailed critique of the content. Here are a few snippets of major shortcomings of the submitted manuscript and by no means reflects a point-to-point examination of the manuscript, which is unfortunately impossible given the current state of manuscript. Reviewer 1:

There are many inconsistencies in analysis. Many figures are not interpretable. Unites are missing, etc. This reviewer cannot evaluate the authors claim from the current figures. For examples here are a few items related to figures 1 and S1: It is unclear why active histone mark is equally enriched in A and B compartments of several cell lines in figure S1D.

1. It is unclear why active histone mark is equally enriched in A and B compartments of several cell lines in figure S1D.

Reply: We apologise for the confusion of the heatmap for the enrichment of H3K4me3 signals that may look similar between A and B compartments. In the revised **Extended Data Fig. 1d**, we illustrated the active H3K4me3 histone signals as a violin plot with statistics. The violin plots show that A compartments have more active H3K4me3 consistently across all the cell lines compared to B compartments *(p-value* <0.0001).

Extended Data Fig 1d

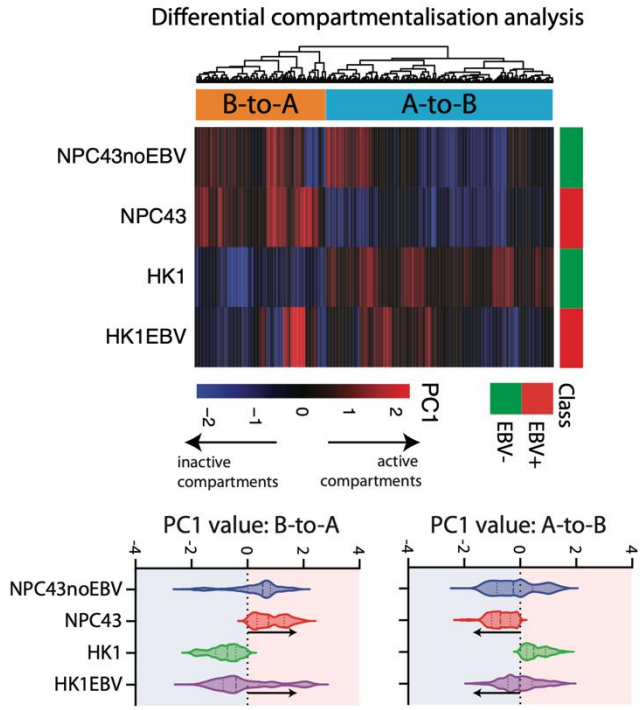


Extended Data Fig. 1d) Active histone modification of H3K4me3 at A/B compartments. *CUT&RUN-Seq was performed to investigate the distribution of the active histone mark H3K4me3 in different genomic compartments. The signals were normalised as RPKM (Reads Per Kilobase per Million mapped reads). Compartments were divided into 50kb window sizes and categorised as A (active) and B (inactive) based on the gene density within each compartment. The log10-transformed H3K4me3 signals were summarised by each 50kb window size. Representative profiles are shown for normal immortalised nasopharyngeal epithelial cells (NPE) - NP69 and NP460, EBV-positive nasopharyngeal carcinoma (EBV+NPC) - C17, C666, NPC43, HK1EBV, and EBV-negative NPC (EBV-NPC) - NPC43noEBV, NPC53, HK1. The violin plot showed that A compartments exhibit higher H3K4me3 signals compared to their individual B compartments, indicating that A compartments are more transcriptionally active than the B compartments. Mean and standard deviation were displayed in the plot, ****adjusted p-value ≤ 0.0001.*

2. Where is NPC43neEBV in figures 1A and 1B?

Reply: Thanks for pointing out this issue. In this study, NPC43 and NPC43noEBV were matched EBV-positive and EBV-negative cell lines derived from the same patients, while NPC43noEBV is not a cell line commonly used in the laboratory, therefore, we did not include it in the original **Figure 1**. To address this comment, we have added the heatmap using the same compartments shown in **Fig. 1a** for the paired EBV-positive and EBV-negative NPC43 and HK1 cell lines in the revised **Extended Data Fig. 2c**. We also further quantified the A and B compartments by the violin plots in the **Extended Data Fig. 2c**, which demonstrated a similar trend of B-to-A or A-to-B transition in EBV+ NPC cells when compared to EBV- NPC cells as shown in **Figure 1a**.

Extended Data Fig 2c



Extended Data Fig. 2c) Heatmap and violin plot of differential A/B compartments in two paired EBV- / EBV+ NPC cells. *Heatmap illustrates the differential A/B compartments identified from EBV+ NPC versus EBV- NPC cell lines to profile in NPC43noEBV - NPC43 and HK1 - HK1EBV cell lines. The in vitro EBV-infected HK1EBV has already shown the selective compartment changes from B to A. The NPC43noEBV has lost the EBV copies from NPC43, showing the loss of EBV episomes could partially reverse the change of A/B compartments.*

3. In tracks shown in Figure 1C, NPC53 is similar to NPC43. How the authors claimed these are regions with compartment switching?

Reply: Thanks for the comments. The overall B-to-A and A-to-B transitions of NPC53 are similar to EBV-negative NPC cells, as we demonstrated in the previous **Figure 1a-b**. Given the example we demonstrated in the previous **Figure 1c** is unclear, we have updated **Figure 1c** with another selected region for better illustration in the revised manuscript.

Figure 1c

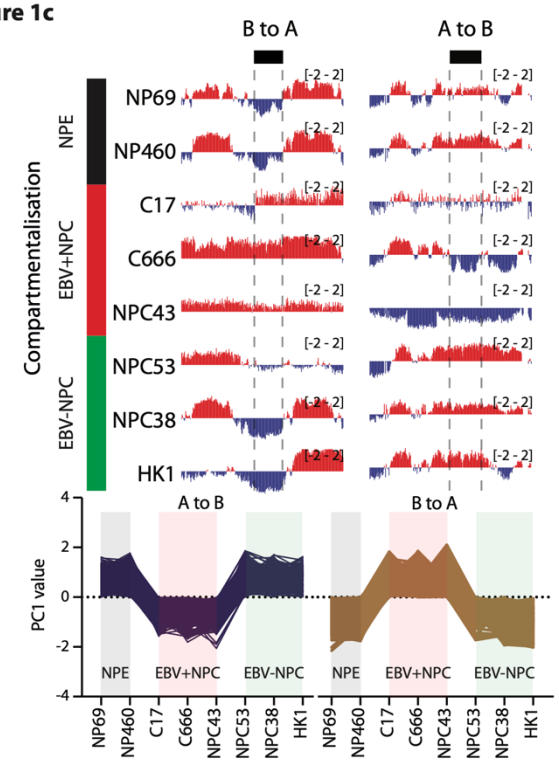


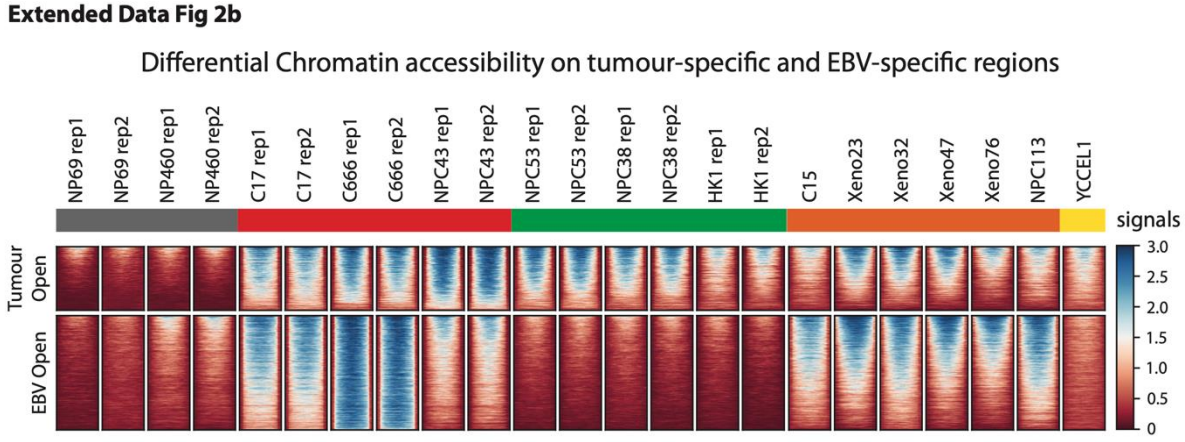
Figure 1c) Hierarchical clustering of EBV-associated compartment switching from B to A and A to B *in NPE, EBV+NPC and EBV-NPC cells visualised on the basis of PC1 value (a 50 kbp window size) through the IGV browser. A positive (negative) PC1 value indicates compartment A (compartment B). PC1 > 0 (active) in red colour and PC1 < 0 (inactive) in blue colour. The line plot summarised the genomic regions with B-to-A and A-to-B compartment switches across NPE, EBV+NPC and EBV-NPC cell lines.*

4. Could the authors explain the color-coding used for the heatmaps.

Reply: We apologise for the confusion. We have carefully labelled all the figures with colour coding and added more relevant details in the revised figure legends. In **Figure 1a, Extended Data Fig 2a and Extended Data Fig 2c** of the heatmaps, we have labelled and described that the red and blue colour refers to active and inactive compartments. In **Figure 1d**, we have added the "Pearson correlation" in the heatmap legend. In **Figure 1c**, we have added "RPKM" (reads per kilobase of transcript per million mapped reads) as the normalised signal intensity in the heatmap legend. We also labelled the colour coding of the heatmap in **Extended Data Fig 3d.**

5. What is data Xeno76, which is listed in a figure legend? The number of EBV+ lines/PDXs does not match across figure panels. As I mentioned earlier, providing a comprehensive list is beyond the scope and must be addressed by the authors before any submission.

Reply: Thanks for the comment. Unfortunately, we did not obtain enough cells from Xeno76 to generate the Hi-C data. This is also the case for Xeno47. Therefore, we only have the Hi-C data for EBV+ NPC PDXs, including Xeno113, Xeno23, Xeno32 and C15 in **Figure 1a-b**. Nevertheless, we managed to get the ATAC-seq results for Xeno76 and Xeno47, which show similar 2D chromatin accessibility as other PDXs (**Extended Data Fig. 2b**). We listed out all PDXs that have been included for the Hi-C experiments in the method section in the revised manuscript **(see lines 484-495 on page 15).**



Extended Data Fig. 2b) Chromatin accessibility of EBVaGC at NPC tumour-specific and EBV-specific open regions. *The differential chromatin accessibility regions were identified from a previous study [1]. YCCEL1 showed open accessibility on tumour-specific open chromatin, similar to other EBV+ samples. However, in EBVaGC EBV did not affect EBV-specific open chromatin accessibility identified from EBV+ NPC samples, suggesting EBV influences the chromatin accessibility in a tissue-specific manner. The blue colour indicates the open chromatin accessibility of the regions.*

6. statement that "reduction in the number of chromatin loops could decrease topologically associating domains (TADs) and compartment formation [11]," has nothing to do with the data in reference (11).

Reply: Sorry for the confusion. We have corrected the reference and refined the statement as below.

"Given we have demonstrated that EBV episomes can reshape and unify genome organisation at the compartment level and studies showed the depletion of CTCF reduces chromatin loops, leading to vanish the topologically associating domains (TADs) [2, 3], we then examined whether EBV infection could induce the large-scale dysregulation of chromatin looping" (**see lines 107-111 on page 4**).

Both studies from Sanborn et al. [2] and Fudenberg et al. [3] demonstrated the relationship between CTCF protein, TADs and chromatin loops. As we observed the rearrangement of compartments associated with EBV infection, further investigation of genome-wide chromatin reorganisation at the chromatin loop level upon EBV infection is necessary.

7. What do the authors mean by "The EBV episomes bind the KDM5B promoter region with enhancer activation, forming cis-regulatory elements; this region gains a new chromatin loop and neoTAD, resulting in increased protein expression"? The sentence does not make any sense.

Reply: Thanks for your comment. We apologise for the confusion. We have revised these sentences as "this region harbours a newly formed topological-associated domain (neoTAD), reflecting altered chromatin architecture. Within this neoTAD, the promoter of KDM5B and its associated cis-regulatory elements exhibit increased chromatin accessibility, facilitating enhanced interactions with a local enhancer. Consistent with this open chromatin state, KDM5B expression was significantly upregulated" (**see lines 166-170 on page 6**). We also revised **Figure 3e** for a better illustration of this result, showing the gain of neoTAD formation in EBV+ cells compared to EBV- or normal control cells in the revised manuscript.

Figure 3e

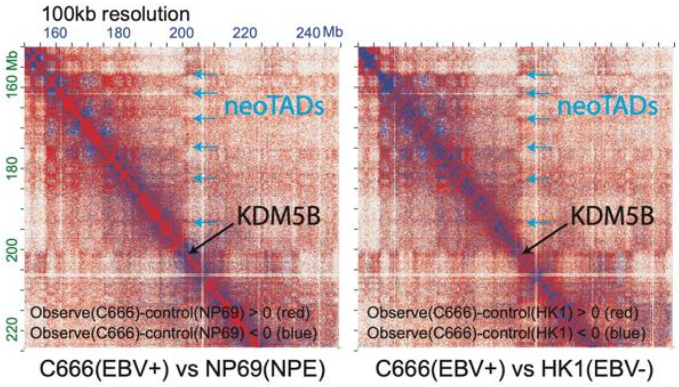


Figure 3e) neoTAD formation at the KDM5B gene locus. *Modified Hi-C data were used to visualise TAD formation, and data were comparing the EBV+ NPC to NPE and EBV- NPC cells using Juicebox.*

The plot was further normalised by sequencing coverage visualising at 100 kbp resolution. If the observed (C666) deduce control (HK1 or NP69) data is greater than zero (red), it indicates more chromatin contacts in C666, and vice versa. It showed multiple neoTADs were formed near the KDM5B locus in C666 (EBV+NPC), however, these neoTADs were absent in NPE and EBV- NPC cells. This suggests EBV episomes could facilitate the long-range of cis-regulatory chromatin interactions to, therefore, increase KDM5B gene regulation. Red indicates (observed)-(control) > 0, and blue indicates (observed)-(control) < 0.

8. What is "EBVIR-KDM5B-enhancer" signature, how was it defined?

Reply: Thanks for the comment. In the revised manuscript, we further clarified that the EBVIR-enhancer-KDM5B signature was defined as a gene set including KDM5B and KDM5B binding targets that were upregulated in EBV+ NPC with enhancer activation through EBV-host chromatin interactions (see **lines 162-164 on page 5**).

9. This reviewer cannot fully examine scientific merit of the manuscript in its current form. There are many unsubstantiated statements. For instance, what is "Modified Hi-C experiment" (it is just a commercial kit!).

Reply: Thanks for the comment. Unlike conventional Hi-C, which relies on sequence-specific restriction enzymes, our method uses a sequence-independent endonuclease to digest chromatin. This modification ensures more uniform coverage and reduces digestion bias. While this approach is based on a commercially available kit, we believe highlighting this distinction is valuable for readers. We have clarified this point in the method section (see **lines 495-496 on page 15**) to avoid confusion. To further address this comment, we have also replaced "modified Hi-C" with "Omni-C" throughout the revised manuscript.

10. It is impossible to understand the details of computational analysis given lack of details in Method section. There is no statistics. Again, using Figures 1 and S1 as an example, from the Methods is not possible to understand how A-to-B or B-to-A compartments are defined. Did the author define these loci per cell line/PDX or used a kind of consensus criteria? What threshold was used?

    Reply: We apologise for the lack of clarity in our Methods section. The A/B compartments were defined for each individual sample (cell lines/PDXs) based on the sign of the first

principal component (PC1) derived from Omni-C data analysis, where positive PC1 values correspond to compartment A and negative values to compartment B. An A-to-B transition was identified when a genomic region changed from a positive PC1 value in one condition to a negative value in another. No additional thresholds were applied beyond this sign change in PC1 values.

To address the reviewer's comments, we have added more detailed elaborations in the method section and figure legends.

The major changes are listed below,

In the method section,

1. We added the description about how we define active and inactive compartments, with specific cut-off values and statistical *p-value* in differential compartment analysis of Hi-C data (**lines 523-529 on page 16**).
2. We added detailed parameters of how the Aggregated peak analysis (APA) was performed, and how to interpret the APA score result (**lines 558-560 on page 17**).
3. We further clarify the sample used for each sequencing technique. For Omni-C (**lines 494-495 on page 15**), ATAC-seq (**lines 574-575 on page 17**), and CUT&RUN-seq (**lines 583-584 on page 18**).
4. We added detailed descriptions about how to analyse the spatial transcriptome data for 10X Visium, GeoMx Digital Spatial Profiler (DSP) and CosMx Spatial Molecular Imager (SMI) platforms. The parameters and cut-off value for differential gene expression analysis using GeoMs DSP and processing procedures including normalization, cell type identification, and neighbourhood network analysis using the CosMx platform were added in the revised manuscript (**lines 613-614, 616-626, 645-649, 663-677 on pages 19-21**).

In the figure legends,

1. The A/B compartments were defined by PC1 values derived from Omni-C data. In the revised figure legend for **Figure 1c**, we added "PC1 > 0 (active) in red colour and PC1 < 0 (inactive) in blue colour. The line plot summarised the genomic regions with B-to-A and A-to-B compartment switches across NPE, EBV+NPC and EBV-NPC cell lines." (**lines 844-846 on page 28**).
2. In **Figure 1e**, we revised the figure legend by adding the description that "a selective genomic region showed compartment A and B were clearly separated to each other without

EBV infection, while EBV infected cells have reorganised compartments with B-to-A transitions. The red colour indicates A (active) compartments and blue colour indicates B (inactive) compartments" (**lines 854-857 on page 28**).

3. **Figure 1f**, we added the description that "P2LL denotes peak-to-lower-left of the interaction matrices. The higher P2LL score indicates stronger chromatin interactions." (**line 863 on page 28**).
4. In **Figure 2b**, we added the description for the statistical analysis. The added sentence is "Fisher's Exact Test has been performed to test if EBVIRs associated with subsets of compartment switches, * denotes *p-value* ≤ 0.05" (**lines 880-882 on page 31**).
5. In **Figure 2c**, we added the description that "CUT&RUN-seq signals were normalised in Reads Per Kilobase per Million mapped reads (RPKM)" (**lines 891-892 on page 31**).
6. In **Figure 3a**, we added "the normalised KDM5B expression is examined and linked to high expression of EBV latent genes" (**lines 905-906 on page 33**).
7. In **Figure 3b**, we added the description that "a total of 154 genes close to EBVIRs with enhancer H3K27ac signals and upregulated in EBV-high cancer cell cluster derived from integrated scRNA-seq data were identified in the Venn Diagram. KDM5B and its binding targets in overlapping 154 genes, named as EBVIR-enhancer-KDM5B signature" (**lines 911-914 on page 33**).
8. In **Figure 3c**, we added "Transcription factor (TF) analysis using Enrichr for 154 genes showed KDM5B is the top one TF for regulating these genes." (**lines 916-917 on page 33**).
9. In **Figure 3e**, we added the description that "Omni-C data were used to visualise TAD formation, and data were comparing the EBV+ NPC to NPE and EBV- NPC cells using Juicebox. The plot was further normalised by sequencing coverage visualising at 100 kbp resolution. If the observed (C666) deduce control (HK1 or NP69) data is greater than zero (red), it indicates more chromatin contacts in C666, and vice versa. It showed multiple neoTADs were formed near the KDM5B locus in C666 (EBV+NPC), however, these neoTADs were absent in NPE and EBV- NPC cells. Red indicates (observed)-(control) > 0, and blue indicates (observed)-(control) < 0" (**lines 924-931 on page 33**).
10. In **Figure 3f**, we added the description that "the sum of normalised expression level in EBVIR-enhancer-KDM5B signature has been calculated in EBV-high and EBV-low cells. The expression levels in the EBV-high and EBV-low clusters were statistically evaluated

by an unpaired t-test, ****$p\text{-}value \leq 0.001$. It showed the EBVIR-enhancer-KDM5B signature transcriptionally upregulated in EBV-high cluster of cells than EBV-low cluster" (**lines 932-936 on page 33**).

11. In **Figure 3g**, we added the description that "The expression of KDM5B was estimated for each circle using the normalised data. The average expression levels of genes in the EBVIR-enhancer-KDM5B signature were calculated for each circle, visualised by image and quantified by the violin plots. *$p\text{-}value \leq 0.05$. It showed that the EBV-high cluster of cells was associated with upregulation of EBVIR-enhancer-KDM5B signature than the EBV-low cluster at spatial level" (**lines 939-944 on page 34**).
12. In **Figure 4a**, we added the description that "neighbourhood analysis using Jaccard distance stratified cells into three niches in this selected region of interest" and the colour code for high and low expression of the genes were added in the figure legend (**lines 952-954 on page 36**).
13. In **Figure 4b**, we added "the GSVA score for EBVIR-enhancer-KDM5B pathway activity was calculated for each cluster of cells" (**lines 956-957 on page 36**).
14. In **Figure 4c**, we added "RPMS1 and LMP1 were used as indicators for EBV-high cancer cells as discovered by scRNA-seq data. Expression levels of the genes upregulated in epithelial clusters 1-3 were selected in the plot" (**lines 959-960 on page 36**).
15. In **Figure 4d**, we added the description that "the selective hallmarks with statistical enrichment in epithelial clusters 1-3 were visualised in the plot" (**lines 962-963 on page 36**).
16. In **Figure 4f**, we added the description that "the proportion of cells in each epithelial cell cluster was examined for MET and non-MET patients. Epi-1/-2/-3 clusters account for about 25% of the total epithelial cells in MET patients" (**lines 972-974 on page 36**).
17. In **Figure 5a**, we added the description that "genes were ranked based on differential expression analysis using the bulk RNA-seq data from MET patients and non-MET patients in the HK and GZ cohorts to calculate the enrichment score of EBVIR-enhancer-KDM5B signature" (**lines 979-981 on page 38**).

Reviewer #2, ECR (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3, expertise in EBV-associated nasopharyngeal carcinoma and single cell sequencing (Remarks to the Author):

In this intriguing paper by Dai and colleagues, the investigators show how the EBV genome/DNA induces spatial reorganization of the human 3D genome through direct chromatin interactions, which affects chromatin loops, histone modifications and enhancer activation or inactivation at specific regions of the genome. More importantly, they highlight that through this mechanism, EBV hijacks KDM5B to promote aberrant gene expression that may promote cancer progression and metastasis. Lastly they use this data to support the notion that this could result in a better signature to predict metastases in NPC.

NPC is an EBV-driven cancer endemic in East and South-east Asia, and there has been a number of important finings in the last 10 years that have allowed a more streamlined approach to its management. Yet the fundamental/biological basis for carcinogenesis remains poorly understood, principally how a disease which is genetically bland, remains aggressive through epigenetic mechanisms. This publication certainly adds a lot of weight to a rapidly developing approach to looking at 3D chromatin conformation beyond just the histone code. Having said that, the absolute novelty of this paper is somewhat tempered by a similar paper (which they cite as reference 9) where a similar approach is shows in EBV-associated gastric cancer. In fact much of the analyses, figures and presentation format for Figures 1-3 follow the format of this Nature Genetics paper. This is not necessarily a criticism but perhaps accounts for why this paper has been bumped to Nat comms rather than Nature or Nature Genetics (although I have to say that the figures in the current paper are actually more refined than the Nature Genetics paper!). One advantage that the prior paper had which was not used here, was a very basic step-by-step explanation of the techniques used, why and what to expect as a result. Given that Nature Comms is a multi-D journal, this would be useful to the broad readership.

Reply: We are deeply grateful for the reviewer's positive feedback on our study. In the revised manuscript, we added more details about the methods and techniques used for this study. We hope this would be helpful in facilitating understanding of these techniques for the broad readers of Nature Communications.

1. The major finding in my opinion here is the role of KDM5B (and perhaps Myc to a lesser extent), and the application of this to a well-curated dataset (which I assume has not been published before). There is a number of circumstantial evidence supporting its role in cancer progression in their models, but I feel that the authors should consider a functional analysis using knock-down or CRSPR to abrogate the function of KDM5B, to study the phenotype and whether there is a rescue in this instance. I would also like to see validation of the role of KDM5B across other EBV associated cancer types (lung, gastric, lymphoma). This would be important to address if KDM5B is context specific or more broadly associated with EBV-DNA interactions. Either result would be intriguing.

Reply: Thanks for the constructive comments and suggestions. As suggested, we inhibited the expression of KDM5B through knockdown (KD) in the EBV+ NPC cells and observed a reduction of cancer cell growth (**Figure 6a-b**). Similarly, treatment of the KDM5B inhibitor (GSK-467) has demonstrated a strong suppression of cancer cell proliferation in EBV+ NPC and EBV+ gastric cancer cells (**Figure 6c, Extended Data Fig. 10a**). As we discovered, the cancer cells with a strong EBVIR-enhancer-KDM5B signature were highly proliferative and associated with metastasis in the clinical samples, we further investigated the cancer cell growth upon a genome-wide loss-of-function (LOF) analysis using the CRISPR screening approach. Targeting the EBVIR-enhancer-KDM5B signature, a noticeable effect on suppression of cancer cell growth was observed in this experiment (NES = -1.36, FDR q-value = 0.039) (**Figure 6d**). This result further confirmed that not only KDM5B alone but EBVIR-enhancer-KDM5B targets together also strongly influence tumour cell growth. The bulk RNA-seq data analysis for KDM5B-KD further revealed that KDM5B regulated the genes in this EBVIR-enhancer-KDM5B signature, controlling for viral life cycle, regulation of viral entry into host cells, and oxidative stress-induced cell death (**Figure 6e**). To investigate the role of KDM5B in regulating the EBV life cycle, we examined the KDM5B binding signals on the EBV genome. As previously reported, EBNA1 is important for binding to the EBV-oriP region for persistent stable EBV infection. Interestingly, the co-localisation of

KDM5B, EBNA1 bindings and H3K27ac at the EBV-oriP region indicates that KDM5B has a potential role to regulate EBV latency and maintain EBV genome and it is likely involved in EBV-host chromatin interactions (**Figure 6f**). Consistent with this finding, the KDM5B-KD cells showed an obvious reduction in cellular EBV copies (**Figure 6g**). This result was further confirmed by the clinical specimens, where cellular EBV copies were positively correlated with KDM5B pathway activity ($r = 0.633$) in the tumour tissues (**Figure 6h**). Collectively, these results showed that EBV hijacked KDM5B to promote cancer cell growth and maintain its survival in a cellular context.

To explore the role of KDM5B on metastasis, the differentially expressed genes (DEGs) analysis to compare KDM5B-KD and KDM5B-WT showed that expression of VCAM1 and VEGFA, which have been identified from our SMI data, were suppressed by KDM5B-KD (**Figure 4c, Extended Data Fig. 10b**). VCAM1, encoding Vascular Cell Adhesion Molecule 1, is known to play a significant role in various cancers, promoting tumour cell invasion, metastasis, and cellular immune response. VEGFA, encoding Vascular Endothelial Growth Factor A, has a crucial role in metastasis by promoting tumour angiogenesis and inhibiting T-cell functions. Functional enrichment analysis further revealed that the other genes repressed by KDM5B-KD were highly involved in the metastasis-related pathways, such as hedgehog signalling, epithelial-mesenchymal transition (EMT), and hypoxia (**Extended Data Fig. 10c**).

We have tested the KDM5B inhibitor and enhancer inhibitor (JQ1) on EBV+NPC (C666 and NPC43) and EBVaGC (YCCEL1) cells. The data demonstrated that both inhibitors could suppress cancer cell proliferation in both EBV+NPC and EBVaGC cells (**Figure 6i**), suggesting the general mechanisms for promoting cancer cell growth through enhancer activation and KDM5B upregulation in both EBV+ cancer types.

We have added all these results as new **Figure 6 and Extend Data Figure 10**, and the relevant contents have been added in the new result section "**KDM5B is one of the key host factors to promote cancer cell growth and maintain EBV DNA copies"** in the revised manuscript (**lines 260-306 on page 8-10**). We also added a short discussion about this result to address that "despite the tissue specificity of the epigenomic profile, EBV docking to the human genome may utilise a similar mechanism leading to enhancer infestation to promote tumour growth in both EBVaGC and EBV+ NPC, which is supported by the findings in this study that inhibition of KDM5B and

enhancer blockade could suppress cancer cell growth in both EBV+ cancer types" (**see lines 337-341 on page 11**).

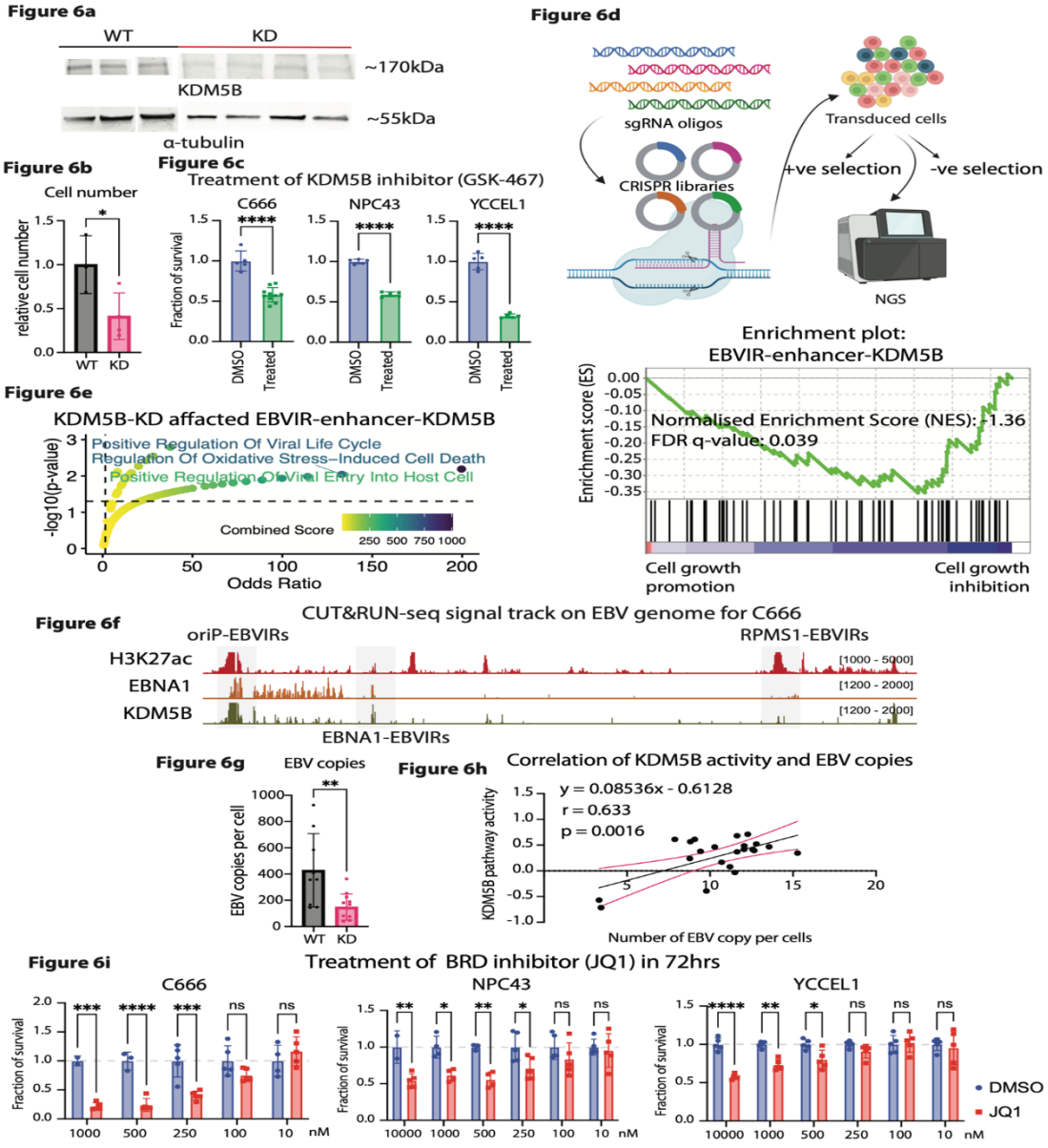
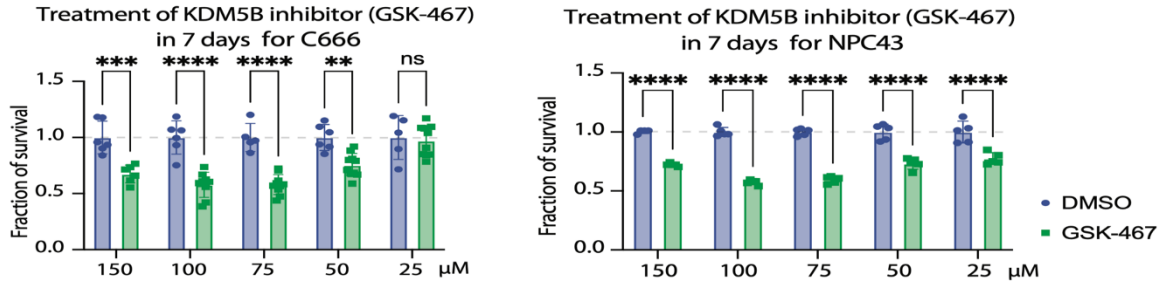


Figure 6. EBV hijacks host factor - KDM5B to maintain its genome copies for host cell survival. ***A) Generating KDM5B wild-type (WT) and knockdown (KD) cells in C666.*** *A western blot result for KDM5B at ~170kDa and α- tubulin control at ~55kDa demonstrate the success of establishing KDM5B-KD clones.* ***B) Quantification of cells in KDM5B-WT and -KD cells.*** *An equal number of cells were cultured for 7 days. A cell count chamber was used to calculate the number of cells in WT and KD cells. The number of cells was normalised to WT. An unpaired t-test was used to test the significance, showing KD cells have*

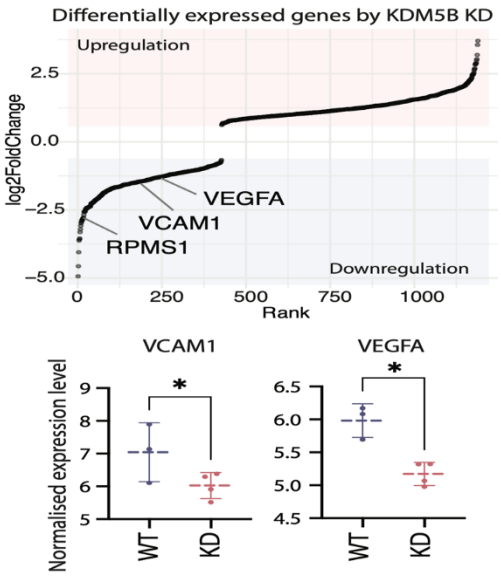
approximately 60 % cell growth reduction with p-value ≤ 0.05. The error bar indicates standard deviation (SD) of the data. ***C) MTT assay results for 7-day treatment of KDM5B inhibitor (GSK-467) in C666, NPC43, and YCCEL1 cells.*** *The KDM5B inhibitor results were normalised to matched DMSO control. The error bar indicates standard deviation (SD) of the data. A two-way ANNOVA test was performed with Bonferroni correction, ****adjusted p-value ≤ 0.0001. A concentration of 75 μM of KDM5B inhibitor treatment showed significant suppression of cancer cell growth in EBV+NPC and EBVaGC cells.* ***D) CRISPR screening revealed that the EBVIR-enhancer-KDM5B signature suppresses tumour cell growth.*** *A genome-wide loss-of-function (LOF) analysis employing the CRISPR screening approach was conducted. CRISPR libraries were transduced into the cells, and cells were subsequently selected by puromycin. The sgRNA sequences were amplified and sequenced by the Illumina platform. The LOF of a gene conferring a growth advantage, characterised by a higher abundance of sgRNAs, was quantified based on the sequencing. Conversely, the LOF of a gene lacking a growth advantage was determined by sequencing with significant reduction of the sequencing signals. The GSEA result showed significant negative enrichment of EBVIR-enhancer-KDM5B signature, NES = -1.36, FDR q-value = 0.039, suggesting LOF of EBVIR-enhancer-KDM5B signature strongly inhibited cancer cell growth.* ***E) KDM5B regulates genes in response to the EBV life cycle and cell death.*** *Bulk RNA-seq has been conducted for KDM5B-WT with three biological replicate and KDM5B-KD with four biological replicate clones. Pre-ranked genes based on log2 fold change of differentially expressed genes were subjected to GSEA. The downregulated genes in the EBVIR-enhancer-KDM5B signature were extracted for GO – BP analysis. A scatter plot shows the odds ratio against -$\log_{10}$(p-value). The combined score ($\log_{10}$(p-value) from Fisher's exact test × the Z score for deviation from the expected rank) is indicated by colour. It showed these genes are involved in the regulation of the viral life cycle, regulation of oxidative stress-induced cell death, and regulation of viral entry into host cells, p-value ≤ 0.05.* ***F) Profile of EBNA1, KDM5B binding, and H3K27ac on EBV genome visualised in IGV browser.*** *The RPKM normalised EBNA1 and KDM5B binding signals generated by CUT&RUN-seq show the involvement of EBNA1 and KDM5B at EBV-oriP-EBVIR with H3K27ac signals.* ***G) Quantification of EBV copy per cell in KDM5B-WT and -KD cells.*** *The ct values from qPCR obtained were used to calculate average EBV copies based on the standard calibration curve established with an equal amount of DNA input. qPCR was performed in triplicates for each clone. An unpaired t-test was used to test the significance, showing a significant reduction of EBV copies in KD cells with p-value ≤ 0.01.* ***H) Pearson correlation plots between the KDM5B pathway and cellular EBV copies in the patient biopsies.*** *The DNA samples collected from the patient biopsies were extracted, and the cellular average EBV copies were measured by the known amount of EBV copy standards and adjusted by the tumor purity estimated from the matched RNA-seq data. The KDM5B pathway activity was derived by GSVA score using the RNA-seq data. The Pearson correlation tests showed a strong positive correlation of KDM5B pathway activity*

to EBV copies with r = 0.633, y = 0.08536x - 0.6128, and p-value = 0.0016. ***I) MTT assay results for 72 hours of treatment of BRD inhibitor (JQ1) in C666, NPC43, and YCCEL1 cells.*** *Multiple concentrations of BRD inhibitors have been tested. The concentration-matched DMSO controls were included. The BRD inhibitor results were normalised to their matched DMSO control. The error bar indicates standard deviation (SD) of the data. A two-way ANNOVA test was performed with Bonferroni correction, *adjusted* p-value < 0.05, **adjusted* p-value < 0.01, ***adjusted* p-value < 0.001 *and* ****adjusted* p-value < 0.0001.*

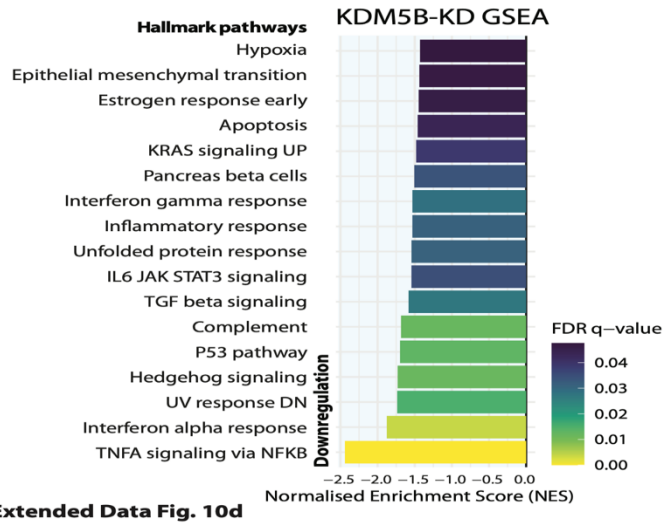
Extended Data Fig. 10a



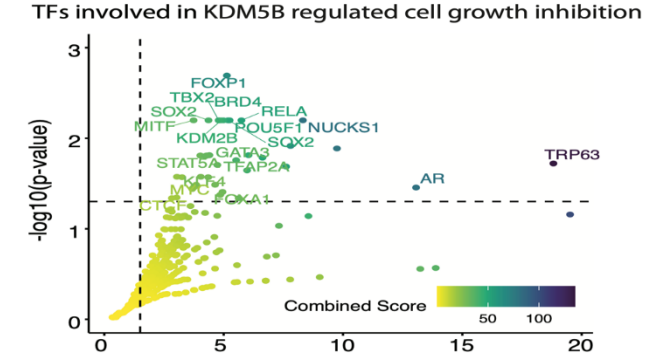
Extended Data Fig. 10b



Extended Data Fig. 10c



Extended Data Fig. 10d



Extended Data Fig. 10 KDM5B inhibition and KDM5B knockdown regulated TFs and hallmark pathways. *A) MTT assay results for 7-day treatment of KDM5B inhibitor (GSK-467) in C666, NPC43, and YCCEL1 cells. Multiple concentrations of KDM5B inhibitors have been tested. The concentration-matched DMSO controls were included. The KDM5B inhibitor results were normalised to DMSO control. The error bar indicates standard deviation (SD) of the data. A two-way ANNOVA test was performed with Bonferroni correction, **adjusted p-value < 0.01, ***adjusted p-value < 0.001 and ****adjusted p-value < 0.0001.* ***B) The KDM5B-KD experiment identified differentially expressed genes related to metastasis.*** *VCAM1 and VEGFA showed significant downregulation after KDM5B knockdown with L2FC = -1.44 and -1.29, respectively, with adjusted p-value ≤ 0.05. The RPMS1 expression has downregulation with log2 fold-change (L2FC) = -2.68, adjusted p-value = 0.063. The box plot with the error bar represents the normalised expression level in the KDM5B-WT and -KD cells from RNA-seq data.* ***C) Functional terms enrichment for KDM5B-KD regulated genes from the molecular signature database (MSigDB).*** *GSEA was conducted based on the ranked L2FC of KDM5B-KD genes using MSigDB Hallmark 2020. Terms with significant enrichment for downregulated genes and FDR q-value ≤ 0.05 were plotted.* ***D) Putative transcription factors involvement in KDM5B regulating genes to suppress cell growth.*** *A CRISPR screening result identified from C666 with L2FC ≤ -2 and adjusted p-value ≤ 0.01 indicates the loss of function of genes shown significant inhibition on cancer cell growth compared to non-target gRNA control. Bulk RNA-seq identified differentially downregulated genes by KDM5B-KD in C666 with L2FC ≤ -0.05 and adjusted p-value ≤ 0.1. The overlapping 36 genes from the CRISPR screening and knockdown experiments were subjected to TF analysis using ChEA 2022. The combined score was calculated (log_{10} (p-value) from Fisher's exact test × the Z score for deviation from the expected rank).*

2. For the clinical section, it would also be important to address how KDM5B and EBV-DNA post treatment can be combined as a prognostic model to guide therapy and perhaps in the discussion even address whether KDM5B inhibitors may have an important role to play (again this depends on the functional analyses above).

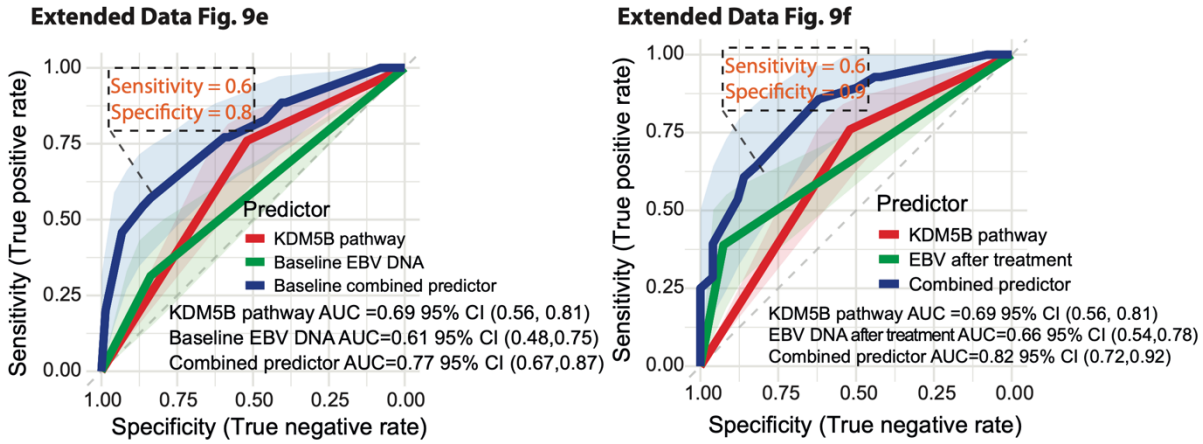
Reply: Thanks for this constructive comment. As suggested, we have further addressed the importance of combined KDM5B activity and EBV-DNA levels as a predictor for patient stratification to guide therapy. We have added the statements in the result section in the revised manuscript (**see lines 252-258 on page 8**) as below.

"Furthermore, KDM5B activity can be used in combination with the overall stage and the plasma EBV DNA level at baseline (AUC = 0.77, 95% CI: 0.67-0.87) or after treatment (AUC = 0.82, 95% CI: 0.72-0.92), as a promising predictor of risk (**Extended Data Fig 9e, Supplementary Table 4-**

5). The late-stage patients (stage IV) with high KDM5B pathway activities and post-treatment EBV DNA levels have a much higher risk for metastasis (OR = 5.8, 95% CI 2.2-15.4, *p-value* = 0.00004, specificity = 0.9 and sensitivity = 0.6, **Extended Data Fig. 9f**)."

To further address the clinical application of the findings, we added a discussion in the revised manuscript as below.

"The specificity to identify high-risk patients is as high as 0.9 when KDM5B activities combine with overall stage and EBV DNA copy number post-treatment. Our functional data further suggest that inhibition of KDM5B and its targets dramatically suppress tumour cell growth, indicating the potential of KMD5B inhibitor as the maintenance therapy for this subset of patients. Alternatively, JQ1 treatment targeting enhancer activation could be another promising treatment option to prevent metastasis and improve treatment outcomes in these high-risk patients" (**see lines 406-412 on page 13)**.



Extended Data Fig. 9e) The receiver operating characteristic (ROC) curves for the KDM5B pathway activity, plasma EBV DNA copy at diagnosis and combination of the KDM5B pathway activity, plasma EBV DNA at diagnosis and Stage. *The combination of the KDM5B pathway activity and plasma EBV DNA at diagnosis and Stage has an AUC of 0.77 (CI: 0.67 – 0.87) higher than the KDM5B pathway activity (AUC = 0.69, CI: 0.56 – 0.81) and plasma EBV DNA at diagnosis (AUC = 0.61, CI: 0.48 – 0.75) alone.*

Extended Data Fig. 9f) The ROC curves for the KDM5B pathway activity, plasma EBV DNA copy after treatment and combination of the KDM5B pathway activity, *plasma* ***EBV DNA at diagnosis, plasma EBV DNA after treatment and Stage.*** *The combination of the KDM5B pathway, and plasma EBV DNA at diagnosis, plasma EBV DNA after treatment and Stage has an AUC of 0.82 (CI: 0.72 – 0.92) higher than*

the KDM5B 1290 pathway (AUC = 0.69, CI: 0.56 – 0.81), and plasma EBV DNA after treatment (AUC = 0.66, 1291 CI: 0.54 – 0.78) alone.

3. The main figures and table are all appropriate and well formatted and I especially liked the word-cloud approach to showing functional terms in Figure 5D. The authors also need to make the clinical data/genomics data available which was not mentioned in the data availability statement.

Reply: We deeply appreciate the kind comments from the reviewer. We have made the data publicly available as described in "Data availability" in the revised manuscript (**see line 423-431 on page 13**).

Reviewer #4, expertise in head and neck cancer and spatial transcriptomics (Remarks to the Author):

Chung et al. Cross-species Chromatin Interactions Reorganize the Human 3D Genome and Hijack KDM5B, a Risk Factor for Distant Metastasis, in Epstein–Barr Virus-associated Nasopharyngeal Carcinoma (NPC)

Chung et al. explore the role of Epstein-Barr virus (EBV) in the reorganization of the human 3D genome in nasopharyngeal carcinoma (NPC). It reveals that EBV–host chromatin interactions disrupt genome architecture, enhancing compartment switching and looping changes that trigger "enhancer infestation" and the activation of cancer-related genes. Specifically, the authors report that EBV hijacks KDM5B, a key protein associated with chromosomal stability, fostering an environment conducive to cancer cell proliferation, increased DNA damage response, and metastatic potential. The study introduces the EBVIR-enhancer-KDM5B signature as a predictive marker for distant metastasis and suggests its clinical relevance in risk assessment and prognosis for NPC patients.

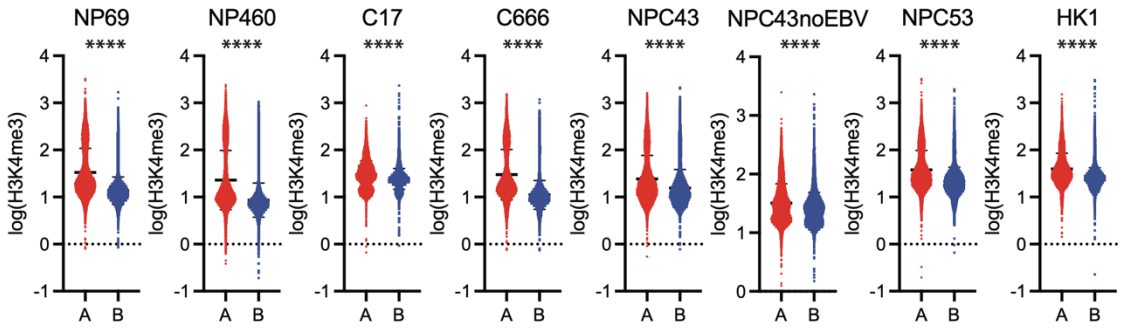
Comments:

1. In Results Section 1, the authors should clearly explain how they define "active" vs. "inactive" compartments. Although partially explained in the Extended Data Figure 1 legend, it is unclear

from the text how this analysis of H3K4me3 marks was performed. Also, the phenomenon of compartment switching in EBV-positive samples should be better explained.

Reply: Thanks for the comment. We have defined the active (A) compartments and inactive (B) compartments by active histone mark - H3K4me3. The A compartments are enriched for higher active histone modifications than the B compartments. The H3K4me3 signals were obtained through CUT&RUN sequencing and normalised as RPKM (Reads Per Kilobase per Million mapped reads). We made a comparison of these normalized signals for the A and B compartments. The normalised signals were summarized in **Extended Data Fig. 1d**, showing the active A compartments have statistically higher levels of H3K4me3 than inactive B compartments across all the cell lines. To further clarify this, we have added more details in the method section (**see lines 523-529 on page 16**) and in the figure legend for extended Data Fig. 1d (**see lines 1088-1092 on pages 44-45**) of the revised manuscript.

Extended Data Fig 1d



Extended Data Fig. 1d) Active histone modification of H3K4me3 at A/B compartments. *CUT&RUN-Seq was performed to investigate the distribution of the active histone mark H3K4me3 in different genomic compartments. The signals were normalised as RPKM (Reads Per Kilobase per Million mapped reads). Compartments were divided into 50kb window sizes and categorised as A (active) and B (inactive) based on the gene density within each compartment. The log10-transformed H3K4me3 signals were summarised by each 50kb window size. Representative profiles are shown for normal immortalised nasopharyngeal epithelial cells (NPE) - NP69 and NP460, EBV-positive nasopharyngeal carcinoma (EBV+NPC) - C17, C666, NPC43, HK1EBV, and EBV-negative NPC (EBV-NPC) - NPC43noEBV, NPC53, HK1. The violin plot showed that A compartments exhibit higher H3K4me3 signals compared to their individual B compartments, indicating that A compartments are more transcriptionally active than the B compartments. Mean and standard deviation were displayed in the plot, ****adjusted p-value ≤ 0.0001.*

2. The authors mention in lines 136-138: "The presence of H3K27me3 at only these select regions may form an epigenetic environment in cells that is favourable for EBV maintenance, and these cells may therefore be selected for during latent EBV infection". I do not understand the basis of this statement, and this seems like mere speculation without any corroborating data. One suggestion would be to knock down the EZH2 enzyme (responsible for H3K27me3 methylation) in EBV-positive NPC cells and quantifying EBV copy number using qPCR at various time points post-knockdown to evaluate if reducing H3K27me3 affects EBV maintenance. Another option would be to use CUT&RUN or ChIP-qPCR to assess the binding of latency-associated EBV proteins (e.g., EBNA1) to H3K27me3-enriched regions in NPC cells.

Reply: Thanks for the comment and suggestion. We agree with the reviewer that this statement is speculative without further evidence, and it would be good to have functional data, such as EZH2 knockdown, to further support this statement. However, as our focus of this manuscript is enhancer activation associated with H3K27ac rather than H3K27me3, we removed this statement in the revised manuscript.

3. More detail on the spatial transcriptomics pipeline is warranted. How was the data normalized across samples? How was the cellular deconvolution performed? Was single-cell data integrated to identify cell types?

Reply: Thanks for the suggestion. We have added more details for NanoString GeoMx (segment level), CosMx (single-cell level) and 10X visium spatial transcriptome data analysis in the method section in the revised manuscript.

The GeoMx data were normalised using the Q3 method, which utilises the top 25% of expressers to normalize across the region of interest (ROI). The CosMx data were normalised using the Seurat "LogNormalize" default setting and adjusting for library size factors to ensure that cell-specific total transcript abundance and distribution of counts did influence downstream data analysis, and 10X visum data were normalised using SCTransform based on regularised negative binomial regression (**see lines 646-649 on page 20**).

For cellular deconvolution using 10X Visium data, the gene markers and parameters from the integrated scRNA-seq data in a previous study [1] were adopted to deconvolute the spatial data using FindTransferAnchors and TransferData with the SCT normalised data. The cell-type

probabilities for each spot were estimated by the 'anchor'-based integration method [4] based on this well-annotated scRNA-seq data (see **lines 613-625 on page 19)**.

We indeed used integrated single-cell RNA sequencing datasets derived from three studies [5-7] as the reference to estimate the probabilities of the cell types in the spatial transcriptome data based on 10X Visium platform. To further clarify this point, we added the statement that "the cell-type probability of the spatial transcriptome data is estimated by the 'anchor'-based integration method [4] based on the integrated scRNA-seq data from three studies, which were downloaded from CNGBdb database - CNP0000428 [5], GEO database - GSE150825 [6] and GSE162025 [7]" (**see lines 616-619 on page 19**)

4. It is unclear what information the bar plot in Figure 4b is trying to convey. I understand that the graph depicts EBVIR-enhancer-KDM5B pathway activity, but missing labels on the Y-Axis make it difficult to interpret.

Reply: Thanks for pointing out this issue and we apologise for the missing label. It is indeed the pathway activity estimated by GSAV using the EBVIR-enhancer-KDM5B signature, as mentioned by the reviewer. We have added the label for the y-axis in **Figure 4b** in the revised manuscript.

5. they observe upregulation of KDM5B and associated enhancer activity in EBV-positive cells, there is limited experimental evidence to directly link these findings to functional outcomes such as increased cell proliferation, immune evasion, or metastatic behavior. Experimental interventions like KDM5B inhibition or enhancer blockade in EBV-positive cells could help demonstrate causation rather than correlation, strengthening the biological relevance of the EBVIR-enhancer-KDM5B signature in NPC progression.

Reply: Thanks for this constructive suggestion. To address the function of KDM5B and enhancer activation in EBV+ cells, we have conducted experiments to examine the functional impact of KDM5B inhibition and enhancer blockade in EBV+ cells. **Please see our detailed reply in point 1 for reviewer 3.** In brief, the results showed that KDM5B inhibition and enhancer blockade could suppress cancer cell growth in vitro. The RNASeq data analysis after KDM5B knockdown indicated that the genes repressed by KDM5B-KD were highly involved in the metastasis-related pathways, such as hedgehog signalling, epithelial-mesenchymal transition (EMT), and hypoxia. More specifically, two metastasis-related genes, *VCAM1* and *VEGFA* found to be associated with the EBVIR-enhancer-KDM5B signature by NanoString SMI data in the clinical samples, were

confirmed to be significantly downregulated by KDM5B-KD. These results have been added as **Figure 6a-i** and **Extended Fig. 10b** in the revised manuscript.

However, the transcription of the immunosuppressive genes, such as CD47 and Galectin-9, and MHC-II genes were not altered upon KDM5B-KD. The elevation of these genes may likely be directly associated with enhancer activation, but independent from KDM5B function. To address this point, we have added a short discussion about this result in the revised manuscript. We comment that "transcription of these immunosuppressive genes and MHC-II genes were not altered upon KDM5B-KD. The elevation of these genes may likely be directly associated with enhancer activation, but independent from KDM5B regulation" (see **lines 374-377 on page 12**).

6. In Section 4 of Results, there is very little information in the section regarding the number of patients accrued in each centre. I understand that this information can be found in Table 1, but this should be included in the text as well.

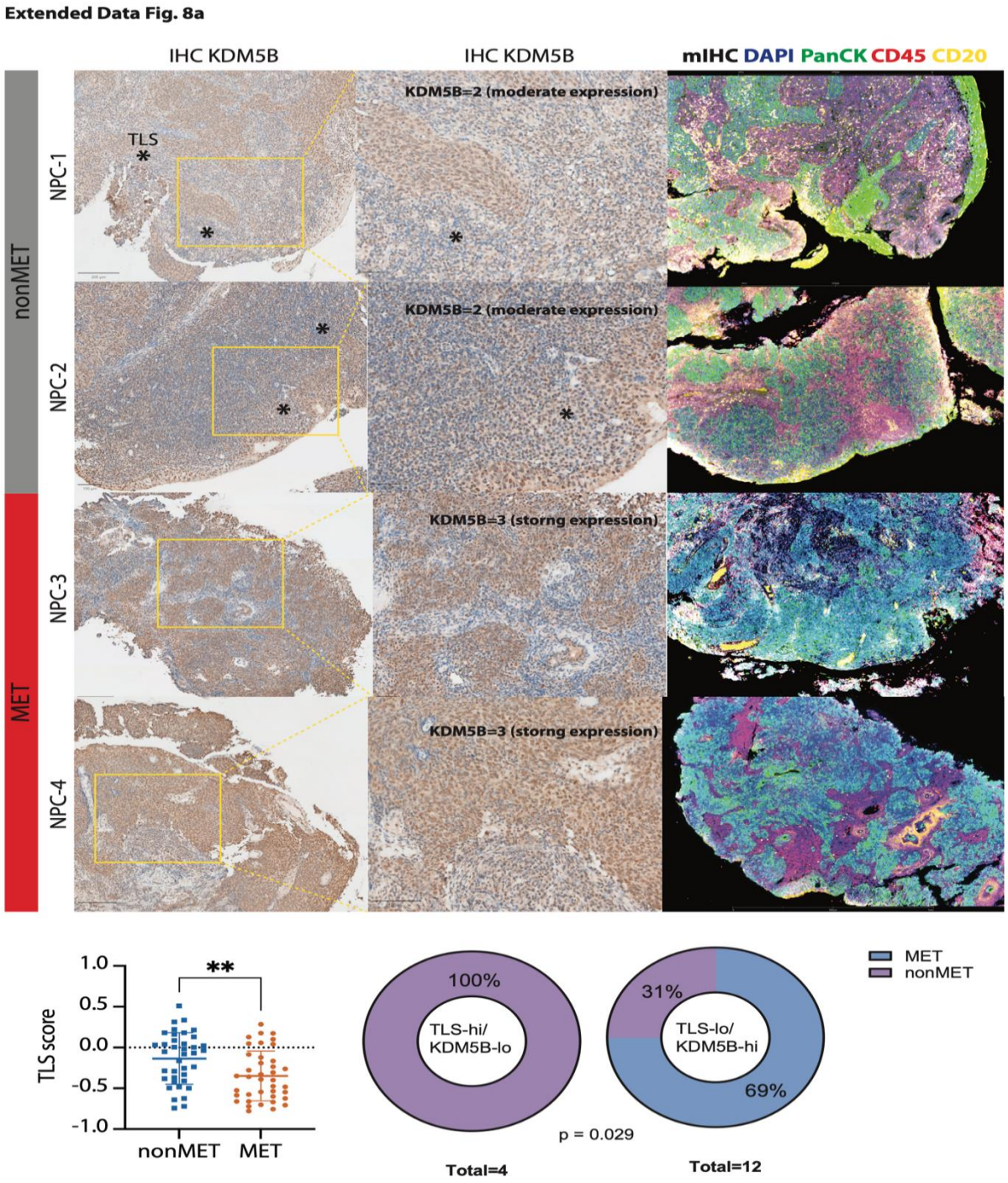
Reply: Thanks for your comment. As suggested, we have added this information in the main text in the revised manuscript (**see lines 218-223 on page 7**).

7. How did the authors assemble the genes in the EBVIR-enhancer-KDM5B signature? Although may be out of scope for this study, it will be interesting to perform IHC for KDM5B on tissue microarrays assembled from patient samples from the multi-centre cohort to investigate if aberrant KDM5B protein expression could be used as a simpler risk factor for NPC metastasis in the future.

Reply: We identify the EBVIR-enhancer-KDM5B by integrating Hi-C data, CUT&RUN-seq data, and scRNA-seq (**Figure 3b**). To further clarify the definition of the "EBVIR-enhancer-KDM5B" signature, we added the statement that "this signature was defined as a gene set including KDM5B and KDM5B binding targets that were upregulated in EBV+ NPC with enhancer activation through EBV-host chromatin interactions" (see **lines 162-164 on page 5**).

Thanks for the constructive comments about KDM5B IHC staining. As suggested, we have conducted IHC staining to quantify the levels of KDM5B. In addition, we used the multiplex IHC staining to quantify the tertiary lymphoid structures (TLS), given that we observed a high association of TLS signature with the development of distant metastasis in our multi-centre multiomics study. The results showed that the patients with a low number of TLS and high

expression of KDM5B are more likely to have metastasis after conventional treatment (P=0.029). This result has been added as **Extended Data Fig. 8a** in the revised manuscript. The result further highlighted the importance of cancer cell features together with the host tumour microenvironment for promoting metastasis in NPC.



Extended Data Fig. 8 High level of KDM5B expression with low level of tertiary lymphoid structures

(TLS) related to NPC metastasis. A) The high expression levels of KDM5B and low levels of TLS were related to NPC metastasis. *The IHC staining of KDM5B and multiplex IHC (mIHC) staining results of PanCK, CD20, CD45 in MET and nonMET patients. The intensity of KDM5B immunohistochemical staining was scored as 0: negative, 1: weak, 2: moderate, and 3: strong expression. A score of 3 was categorised as KDM5B-hi and a score of less than 3 as KDM5B-lo. As shown in the figure, NPC-1 and NPC-2 were scored 2, while NPC-3 and NPC-4 were scored 3. Tertiary lymphoid structures (TLS), marked * asterisk in the figure, is defined as an aggregate consisting of more than 100 CD20+CD45+PanCK- cells, surrounded by CD20-CD45+PanCK- cells, and located inside or adjacent to tumour nests based on multiplex immunofluorescent staining (DAPI: blue, PanCK: green, CD45: red, CD20: yellow). 77 NPC cases from HK cohort have been evaluated. TLS of every case were calculated from the whole-slide image and normalised by the tissue area. The expression of KDM5B and were evaluated by the HKU pathologist. TLS score was estimated by GSVA method using the gene signature derived from a previous study [8]. Cases with normalised TLS more than 3 were classified as TLS-hi and less than 3 as TLS-lo. Unpaired t-test has been performed for TLS score in MET and nonMET patients, **adjusted p-value ≤ 0.01. The box plot showed MET patients has significantly less TLS compared to nonMET patients. There is a statistically significant difference between high expression levels of KDM5B (KDM5B-hi) with low TLS (TLS-lo) and low expression levels of KDM5B (KDM5B-lo) with high level of TLS (TLS-hi), p-value ≤ 0.029. About 69% of MET patient samples contain high expression levels of KDM5B and low levels of TLS. However, it cannot be found in low expression levels of KDM5B with high TLS samples.*

8. Minor comment: Some of the text on figures is too small to read.

Reply: Thanks for the comment. We have enlarged the overall size of the text and highlighted the important information in the revised manuscript. We hope the readability of the text on these figures improves.

Reference

1. Ka-Yue Chow, L., et al., *Epigenomic landscape study reveals molecular subtypes and EBV-associated regulatory epigenome reprogramming in nasopharyngeal carcinoma.* EBioMedicine, 2022. **86**: p. 104357.
2. Sanborn, A.L., et al., *Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. E6456-65.
3. Fudenberg, G., et al., *Formation of Chromosomal Domains by Loop Extrusion.* Cell Rep, 2016. **15**(9): p. 2038-49.
4. Stuart, T., et al., *Comprehensive Integration of Single-Cell Data.* Cell, 2019. **177**(7): p. 1888-1902 e21.
5. Chen, Y.P., et al., *Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma.* Cell Res, 2020. **30**(11): p. 1024-1042.
6. Gong, L., et al., *Comprehensive single-cell sequencing reveals the stromal dynamics and tumor-specific characteristics in the microenvironment of nasopharyngeal carcinoma.* Nat Commun, 2021. **12**(1): p. 1540.
7. Liu, Y., et al., *Tumour heterogeneity and intercellular networks of nasopharyngeal carcinoma at single cell resolution.* Nat Commun, 2021. **12**(1): p. 741.
8. Liu, Y., et al., *Single-cell and spatial transcriptome analyses reveal tertiary lymphoid structures linked to tumour progression and immunotherapy response in nasopharyngeal carcinoma.* Nat Commun, 2024. **15**(1): p. 7713.