Nuclear RNA exosome targeting onto chromatin loci in the degradation of unstable transcripts affects transcriptional programs in liver cells.

Hiba SOUAIFAN 1,4, Mickael COSTALLAT 1, Laura SITKIEWICZ 3, Kylian GODEST 2, Florence CAMMAS 3, Carl MANN 2, Christian MUCHARDT 1*, Christophe RACHEZ 1*

1. Institut de Biologie Paris-Seine (IBPS), Sorbonne Université, CNRS UMR7238, 75252 Paris, France
2. Institut de Biologie Intégrative de la Cellule (I2BC), 1, Avenue de la Terrasse, 91190 Gif-sur-Yvette, France
3. Institute of Human Genetics, CNRS UMR9002 University of Montpellier, 34396 Montpellier, France
4. Ecole doctorale Complexité du Vivant, Sorbonne Université, Paris, France
* Co-corresponding authors.

Abstract

The regulation of RNA maturation and degradation processes is crucial for maintaining cellular homeostasis, with defective regulation potentially leading to cancer or cellular senescence. RNA homeostasis occurs co-transcriptionally and is influenced by chromatin-associated factors, highlighting the need for further exploration in this area. HP1 proteins, originally identified as markers of pericentromeric heterochromatin, play significant roles in co-transcriptional processes occuring on chromatin. This study investigates the role of HP1 proteins in the maturation and degradation of non-coding RNAs, specifically focusing on their interaction with the RNA exosome, a key machinery for RNA degradation.
The effects of depleting all three HP1 isoforms were investigated in mouse embryonic liver (BMEL) cell lines. Our results demonstrate that the RNA exosome interacts with HP1 proteins and is targeted to chromatin loci through HP1-dependent mechanisms. Furthermore, the absence of HP1 proteins alters various RNA species genome-wide, reactivating transposon-like repeat elements and stabilizing certain gene-associated transcripts thereby found in cytoplasm. In human cells, patients with highest levels of inflammation show reduced HP1 levels and less RNA degradation. This reduction leads to tumors with unique characteristics, being more rigid and exhibiting lower proliferation, which makes them less deadly than tumors with high HP1 levels.
These findings illustrate the role of HP1 proteins in regulating RNA homeostasis through their interaction with the RNA exosome, suggesting important implications for cancer biology.

Results

Several non-coding, unstable RNA species are found stabilized in the absence of HP1 proteins, in both chromatin and cytosolic fractions.

A series of studies have established links between HP1 proteins and the RNA degradation machineries. Independent approaches by Immunoprecipitation-mass spectrometry revealed the presence HP1 proteins in the interactomes of Exosome-associated factors such as ZC3H18 and MTR4, and vice-versa (Zaidan, 2017; Winczura, 2018; Silla, 2020). By String analysis of factors co-precipitating in a fraction containing both ZC3H18 and HP1 proteins (from Winczura, 2018) we could recapitulate a functional protein association network between HP1 proteins and the RNA Exosome complex and its nuclear cofactors (sup Fig1A). These clues prompted us to investigate the mechanisms linking HP1 with RNA stability/degradation on chromatin, in relationship with the nuclear RNA Exosome.
The implications of the three HP1 isoforms together in RNA processing and degradation were evaluated by comparing the transcriptomes of mouse embryonic liver cells (BMEL cells) invalidated for the expression of all three HP1 isoforms, triple knockout (TKO), compared to HP1-expressing cells (Het, heterozygote for HP1alpha and wild type for HP1beta and gamma). Six independent clones TKO1, -2, -3, and Het1, -32, -35, were compared. We looked at repressed/unstable RNA species therefore expected to be detected on chromatin only. Differences in RNA processing between these clones were assessed by analyzing RNA levels in both cytoplasmic and chromatin subcellular compartments. This way we expect to look at both mature RNA in the cytoplasm and chromatin-enriched RNA subject to ongoing co-transcriptional processing. To this aim, Het and TKO BMEL cells were fractionated into cytosol, nucleosol, and chromatin fractions (Fig. 1A). During the procedure, nucleosol extraction allowed us to clear the chromatin fraction of processed RNA species expected to be released from chromatin and transiting towards the cytoplasm. Fractionation quality was assessed by looking at several marker proteins and representative nuclear or cytoplasmic transcripts which appeared to be distributed as expected between the different fractions (Sup Figures 1 B-C-D). Transcriptome analysis of total RNA, and RNA from both chromatin and cytosolic fractions from the six independent clones was performed by illumina sequencing.
Previous reports pointed on the implications of HP1 proteins and of their homolog in S.pombe, SWI6, in the silencing of pericentromeric Satellite RNA, or in the repression of LTR repeat RNA (Keller, 2012; maison, 2016; Saksouk, 2020). We examined RNA levels of such repeat elements (LTR, LINE, SINE, MMSAT Satellite repeats) in . We focused on the stability of "self-expressed" repeats by counting RNA reads on extragenic repeat elements only. This way "gene-dependent" repeats embedded in coding gene transcripts thereby subject to maturation processes targeting their surrounding RNA were not considered. We also considered RNA levels at non-repeated genomic elements that are known sources of unstable non-coding transcripts, such as upstream antisense uaRNA at promoters (PROMPT RNA), and eRNA at active distal enhancers (enhD RNA).
Among all genomic elements that we considered and that we found expressed (i.e. more than two reads) in either Het or TKO or in both cell types, we found much higher proportions of these elements that were producing RNA in TKO cells compared to HP1-expressing cells. Especially, more than 50 percent of all expressed LTRs were found in TKO only (Figure 1B and 1C). The

number of all transcript-producing genomic elements analyzed so far was globally higher in TKO than in Het.
Looking at proportions of RNA showed that the mean RNA levels on repeat elements, on PROMPTs, and on distal enhancers (enhD) were increased in the chromatin fraction in TKO cells relative to Het cells (Figure 1D). In comparison, the levels of pre-mRNA at protein-coding genes were only minimally affected by HP1 depletion, The number of exonic reads counted on chromatin were minimally changed. At a global scale, coding gene transcripts were together not drastically changed both in terms of expression levels and Het-/TKO-specific expression (Figure 1C-E). We conclude that the absence of HP1 leads to an increase in both the number and the global expression level of genomic elements producing non-coding unstable RNA.
Some of the RNA species found to be increased in the chromatin fraction, especially LTR, LINE, and MMSAT transcripts, were also detected in higher proportions in the cytoplasm (Figure 1E). Transcripts initiated at LTR repeats were found to extend up to several kilobases. Interestingly, many of these LTR transcripts appeared to be spliced in the cytoplasm (RLTR17, RMER13A, figure 1B). This strongly suggests that LTR transcripts may be stable enough to follow RNA maturation processes that occur between chromatin and cytoplasm. Our results together suggest an increased stability of non-coding unstable RNA transcripts in the absence of HP1, possibly related to a defective RNA degradation on chromatin.

RNA Exosome and its nuclear cofactors are localized on chromatin, interact with HP1, and accumulate in subcellular soluble fractions in the absence of HP1.

Because of the observed differential stability between chromatin-enriched and cytosolic RNA in the absence of HP1 proteins, we raised the hypothesis that the HP1 depletion may affect functionality of the nuclear Exosome complex. We checked differences in expression levels of components of the nuclear Exosome complexes (total extracts), we observed apparent identical levels for the tested proteins between representative Het and TKO clones (figure 2A). No drastic changes in expression levels we observed. The distribution among our three C, N, and X fractions, of components of the core Exosome, EXOSC3, EXOSC9, and cofactors associated to the NEXT complex, ZC3H18, or common to all nuclear complexes, MTR4, revealed that core components and nuclear cofactors are detected both in the nucleosol and associated with chromatin, co-fractionating with HP1 proteins (figure 2B). Since the exosome core components and nuclear cofactors analyzed are found both in the nucleosol and on chromatin, we checked whether these proteins associate with HP1. Indeed, by coimmunoprecipitation out of a nuclear fraction containing both soluble nuclear components and solubilized chromatin, HP1 proteins were found associated to EXOSC9, EXOSC10, MTR4, and ZC3H18 (Figure 2F-H, and Supplementary Figure 2A-B). From these results we raised the hypothesis that HP1 proteins may be chromatin associated factors involved in the accessibility of Exosome to chromatin. For this reason, we analyzed more carefully the distribution of the nuclear RNA exosome in the three C, N, and X fractions by following the core exosome component EXOSC9 in Het and TKO cells by Western blot analysis (Fig. 2D). Quantitation of the signal revealed a significant accumulation of this subunit in all soluble fractions while the protein does not globally disappear from the chromatin (Fig. 2E). Interestingly, accumulation of EXOSC9 in the soluble fractions coincides with a drastic increase in the cytosolic detection of DIS3L, the catalytic component of the SKI complex, the cytoplasmic Exosome. These results together suggest that the subtle decrease in

chromatin-associated exosome complexes may result in an increase in cytoplasmic localization of exosome components. Such changes in subcellular localization of these components may provide an explanation for the stabilization of chromatin-associated RNA species, and the concomitant increase in the degradation of cytoplasmic RNA. The observed differences suggest not a global loss of the nuclear RNA exosome activity but a mis-localization and re-distribution of Exosome complexes.

Association of nuclear RNA exosome and cofactors to chromatin loci is HP1-dependent.

Because we did not observe a global release of nuclear Exosome components from our chromatin fraction in HP1-depleted cells, we explored the hypothesis that more subtle mis-localization of the complex may be observed in the absence of HP1 at discrete loci. To this aim, we performed chromatin immunoprecipitation with high-throughput sequencing (ChIPseq) of components of the nuclear RNA exosome tested above in HP1-interaction assays, EXOSC10, ZC3H18, and MTR4, as well as of CTCF, and the active, phosphorylated form of RNA polymerase II (P-RNAPII, as a mix of both phospho-Ser2 and -Ser5 forms), as positive controls, and IgG as a negative control. In parallel, a genome-wide view of binding sites of the three HP1 isoforms in Het cells as well as regions enriched in the two repressive chromatin associated histone modifications, H3K9me3 and H3K27me3, was obtained by Cut&Tag.
Analysis of ChIPseq profiles in Het32 cells showed that the nuclear exosome components, MTR4, ZC3H18, and EXOSC10 are recruited on chromatin and enriched at specific genomic loci (Figure 3A). Most regions of exosome cofactor recruitment matched loci found to be enriched in all three HP1 proteins in Het cells (Figure 3A bottom lanes).Comparison of the two clones revealed however that the density of the profiles observed at these loci was clearly lost in TKO2 cells, while CTCF and P-RNAPII binding sites remained enriched in both clones. Intriguingly, the density profiles of P-RNAPII at these loci appeared to be clearly increased, despite the fact that chromatin-associated RNA density at most genes was unchanged (Figure 3A top lanes). These results are consistent with the fact that RNA reads counts on exons and introns appeared fairly similar between Het and TKO cells (Figure 1C-E). Together, ChIPseq profiles suggest that HP1 proteins specify regions where a tight, close association of exosome components with chromatin is achieved, rather than mere recruitment onto chromatin.

Genome-wide, regions of significant enrichment in MTR4, ZC3H18, EXOSC10, CTCF and P-RNAPII on chromatin were defined as ChIPseq peaks versus IgG by MACS2 analysis.
The number of peaks defined in Het32 cells for each nuclear exosome component was dramatically decreased in TKO2 cells, while the total number of CTCF peaks remained fairly unchanged and the number of P-RNAPII peaks was strongly increased in TKO2 (Figure 3B and 3C). CTCF, the CCCTC-binding factor, is a DNA-binding protein involved in the regulation of the three dimensional conformation of chromatin. The similar number of CTCF ChIPseq peaks obtained in both clones suggests that chromatin conformation in BMEL cells was not globally affected by the depletion of HP1 proteins. However, all the ChIPseq peaks combined from both cell lines for each factor showed a drastic decrease in intensity or disappeared in TKO2 cells (all peaks, Figure 3C). Our results together reflect a genome-wide loss of the targeting of these exosome components onto chromatin in the absence of HP1 proteins.

Since many ChIPseq peaks were located on the body of expressed genes (see: Tubb4b, Junb, Spin1 genes, figure 3A), we looked at ChIPseq densities by metagene analysis on a list of more than twenty thousand genes. Metagene profiles showed a preferential enrichment of MTR4, ZC3H18, EXOSC10 at the transcription start site (TSS) of the genes (Figure 3D). P-RNAPII showed a preferential recruitment at both TSS and beyond the gene body as expected from previous reports. Comparison of the ChIPseq profiles in Het32 and TKO2 revealed that our metagene profiles followed the same loss in absence of HP1 as for ChIPseq peaks. Chromatin enrichment at the TSS and along gene bodies was lost in TKO2 for MTR4, ZC3H18, EXOSC10, while P-RNAPII density profile was strongly increased all over gene bodies (Figure 3D, compare brown and dark blue profiles). These observations confirm a major role of HP1 proteins in targeting the exosome onto chromatin.

MTR4 ChIPseq peaks and chromatin accessibility at CTCF binding loci.

Association of the RNA exosome with a set of specific cofactors in the nucleus specifies various nuclear RNA exosome complexes (i.e. NEXT, PAXT, TRAMP,...) having slightly distinct activities. In an attempt to delineate distinct nuclear complexes associated to ZC3H18 and to MTR4, we analyzed the patterns of colocalization between ChIPseq peaks in Het32 cells. While ZC3H18 peaks colocalized with an array of other proteins tested in our ChIPseq assays (Fig. 4A), nearly all of the MTR4 chIPseq peaks detected in Het32 were colocalized with CTCF peaks (Fig. 4B, and 4C). Since MTR4 represents a common subunit to all nuclear exosome complexes, our results suggest a relationship between nuclear exosome complexes and the association of CTCF to chromatin. A comparison of CTCF ChIPseq peaks in both Het32 and TKO2 revealed that a majority of CTCF peaks (40913 peaks, Fig. 4D) were conserved between both Het32 and TKO2, these peaks are mostly already annotated in ENCODE database. However, two-thirds of our CTCF peaks were specific for either Het32 or TKO2 cells (29080 and 30213 peaks, respectively, Fig. 4D).
DIS3, a catalytic subunit of nuclear RNA exosome complexes, was previously found to affect chromatin architecture via of non-coding RNA at CTCF binding sites (Laffleur, 2021). To evaluate CTCF peaks in relationship with RNA exosome, the levels of chromatin-enriched RNA at CTCF peaks were quantified in our transcriptome, as in Fig. 1, on Het32- and TKO2-specific CTCF peaks. While CTCF peaks found in Het-only were not affected in terms of RNA detection, TKO-specific CTCF peaks were found to be slightly but significantly enriched in RNA in TKO versus Het (compare left and right panels in Fig. 4E, and 4F).
The cell-specific CTCF peaks, and peaks common to both cell types appear to have a tendency to cluster together defining domains of chromatin opening/closing between the two cell types (for example, see: TKO-only CTCF peaks in Fig. 4G). For this reason, Potential changes in chromatin accessibility between the two cell types were evaluated by ATACseq on CTCF peaks in all Het and TKO clones. Comparing the three categories of CTCF peaks revealed a gain in chromatin accessibility in TKO clones on TKO2-specific CTCF peaks only (Fig. 4H). Our results together suggest that RNA levels at CTCF binding sites may affect their detection as ChIPseq peaks, and local chromatin accessibility as well.

Gain in chromatin accessibility at enhancers affects collagen genes in BMEL cells.

The gain and loss of CTCF ChIPseq peaks between Het and TKO cells prompted us to evaluate their impact on gene expression programs in these cells. To this aim, we used CTCF peaks as markers of new regulatory elements. As depicted in figure 5A, we first evaluated the relationship between CTCF peaks and their closest neighboring enhancer elements (within 5kb), we then evaluated the relationship between these enhancer elements and dysregulation of their closest genes (within 100kb). Out of the two hundred thousand distal enhancers (enhD) in the ENCODE database, four categories of enhD were defined based on their differential chromatin accessibility in ATACseq between Het and TKO cells. EnhD up, enhD down, unchanged enhD, and inactive enhD were found to have increased, decreased, unchanged, or no ATACseq signal, respectively. We expect these four categories of enhD to represent distal enhancers being activated in TKO, repressed in TKO, or unchanged or inactive in both cell types respectively.
Counting the proportions of the CTCF peaks found to be the closest to these enhD within a distance of 5kb revealed that TKO-only CTCF peaks were in highest proportions in the vicinity of activated enhD (orange category on enhD up, Fig. 5B), while Het-only CTCF peaks were mostly closest to enhD found to be repressed in TKO (blue category on enhD down, Fig. 5B). We then looked at the genes closest to enhD, and found to be significantly changed in their expression levels between Het and TKO by DESeq2 analysis. We observed that upregulated genes were found in highest proportions close to enhD up (genes up, orange bars, Fig.5C). Conversely, downregulated genes were found in highest proportions close to enhD down (genes down, blue bars, Fig.5C). Thus, similarly to the observation made on CTCF and enhD above, the genes closest to enhD revealed a tendency to be dysregulated in the same direction as their closest distal enhancers. Among the genes found to be upregulated by DESeq2 analysis, a number of collagen genes are significantly upregulated in TKO cells (Fig. 5D). A Gene Ontology (GO) analysis performed on the list of upregulated genes closest to activated enhancers (orange bar on top of enhD up, Fig.5C) identified Collagen-Containing Extracellular Matrix as the most significant GO cellular component corresponding to this category.
These results recapitulate a functional link between upregulated enhD and upregulation of collagen genes.

HP1 Deficiency and Reduced RNA Exosome Activity Converge to Deregulate Collagen Genes in Hepatocellular Carcinoma

To investigate the relevance of the cooperation between HP1 and the RNA exosome in hepatocellular carcinoma (HCC), we analyzed RNA sequencing data from 76 patients in Mongolia (Candia, 2020). The molecular subtypes of Mongolian HCC largely aligned with those found in patients from other regions, including Asia, Europe, and North America.
Our analysis revealed that HP1 gene expression (including CBX3, CBX1, and CBX5) were upregulated in the majority of tumors, with 71% of cases showing at least a twofold increase relative to adjacent healthy tissues (Fig. 6A). This finding is consistent with earlier reports indicating that elevated HP1 levels correlate with poor prognosis in HCC.
We next stratified patients based on their absolute HP1 expression levels, calculated as the summed expression of CBX3, CBX1, and CBX5 (Fig. 6A and 6B). This approach effectively divided the patients into two distinct subgroups in an unsupervised UMAP analysis (Fig. 6C).

Notably, patients with high total HP1 levels were enriched in populations previously associated with poor prognosis (Candia, 2020).

Comparative transcriptome analysis between high HP1 (Hi HP1) and low HP1 (Lo HP1) patients, defined by the first and third quartiles, respectively, revealed significant upregulation of alpha-fetoprotein and YAP1 in the Hi HP1 group. This upregulation correlated with increased expression of genes enriched in the KEGG pathway for cell cycle regulation. Altogether, this observation suggested that the Hi HP1 tumors were in average poorly differentiated, highly proliferative, and more aggressive (Fig. 6D, E).

In contrast, Lo HP1 patients exhibited activation of alpha-smooth muscle actin (alpha-SMA/ACTA2), vimentin, and multiple genes involved in collagen-containing extracellular matrix (ECM) formation, suggesting a link to liver fibrosis (Fig. 6F, G). This observation was particularly striking, as prior data mining from both human and mouse cell models with reduced RNA exosome activity also demonstrated upregulation of the same Gene Ontology (GO) cellular component pathway—collagen-containing ECM. These studies included Exosc3 inactivation in mouse embryonic stem cells (GSE100535), ZC3H18 depletion in HeLa cells (GSE212557). These findings suggested that genes involved in collagen ECM formation are particularly sensitive to reduced RNA exosome activity. Moreover, reduced HP1 activity and reduced RNA exosome activity both appeared to contribute to increased expression of these genes.

The exact mechanism underlying this specificity remains unclear. However, previous studies have shown that certain enhancer RNAs (eRNAs) and long noncoding RNAs (lncRNAs) are particularly susceptible to RNA exosome-mediated degradation and tend to accumulate upon its inactivation (Pefanis, 2015). Consistent with this, in our EXOSC3 depletion experiment in Wi38 cells, for which the sequencing depth was sufficient to detect eRNAs, we noted an accumulation of eRNAs and/or promoter RNAs upstream of several collagen genes, including COL5A1 and COL6A1. As These RNA species are suspected to play regulatory roles in gene activation, they may be inducing transcriptional activity by altering chromatin three-dimensional structure or by promoting elongation through neutralization of NELF.

An alternative explanation is that collagen genes are highly regulated at the mRNA stability level (as reviewed by Zhang and Stefanovic, 2016). Supporting this hypothesis, we observed an overall enhanced stabilization of otherwise labile mRNAs in the Lo HP1 patients, potentially due to reduced RNA turnover.

Overall, our findings suggest that absolute HP1 levels serve as valuable prognostic markers for grading HCC. This study thereby provides a more refined understanding of the diagnostic value of HP1 levels in this tumor type and highlights the shared molecular consequences of reduced HP1 and RNA exosome activity.

Figure legends

Figure 1: Chromatin-enriched RNA is increased in TKO cells and stabilized in the cytoplasm**.** **A.** Genotype and subcellular fractionation scheme of the six BMEL clones expressing (Het) or not (TKO) HP1 proteins used in this study. **B.** Genome views of RNA density profiles (merged triplicates) in chromatin (green) and cytosolic (red) fractions, at indicated genomic loci showing TKO-specific or TKO-increased RNA. **C, D, E.** RNA levels on repeat elements (LTR, LINE, SINE, MMSAT4), PROMPTS, and distal enhancers (enhD) are increased in TKO chromatin, but remain mostly unchanged on gene downstream regions (downstr.) and exons. **C.** Cumulative histograms of proportions of the indicated genomic regions expressing RNA (more than 2 reads) by comparison of the three Het clones with the three TKO clones. The five categories discriminate expression on chromatin in TKO only (brown), increased in TKO (salmon), in Het only (dark blue), increased in Het (light blue), unchanged (grey). **D, E.** Boxplots of mean RNA read counts in chromatin (D.) or in the cytosol (E.) on genomic regions as in C. in Het (blue) or TKO (orange). F, G, H. Histogram and boxplots as in figures C, D, E for proportions and counts on expressed exons (more than 10 reads).

Figure 2: A. Protein levels in total extracts, and **B,** subcellular distribution in the three fractions (C, N, X) depicted in figure 1A of indicated exosome subunits and cofactors in Het and TKO cells by Western blot analysis and Ponceau staining of total proteins. **C, D, E.** Co-immunoprecipitation of HP1g by MTR4, ZC3H18, EXOSC10, or EXOSC9 in the nuclear compartment. **F. S**ubcellular distribution of EXOSC9 as in panel 2B in Het and TKO clones by quantification of the signal in Western blot analysis in supplementary Fig. 2C, graphed as a mean, and SD of all Het or TKO samples. Student's Test P-values are: *, 0.022; **, 0.003.

Figure 3: ChIPseq analysis reveals an HP1-dependent targeting of Nuclear RNA exosome and cofactors to chromatin loci. **A.** Genome views of chromatin RNA density profiles (top), ChIPseq profiles of indicated proteins, and HP1 Cut&Tag density profiles (bottom) on indicated loci. **B.** Number of ChIPseq peaks versus IgG defined by MACS2 analysis for the indicated proteins in Het32 (blue) and TKO2 (brown). **C.** Heatmaps of ChIPseq densities centered on peaks depicted in B, from both Het32 and TKO2 combined in a single list (all peaks). **D.** Metagene anaysis of ChIPseq profiles of the indicated proteins on 23808 gene bodies defined as TSS to TES intervals, in Het32 (dark blue), in TKO2 (dark red), compared to IgG profiles in Het32 and TKO2 (salmon and light blue profiles, respectively). **E, F.** Cut&Tag profiles at TSS for HP1 proteins (HP1, Fig. 3E), and for H3K9me3 and H3K27me3 histone marks (F).

Figure 4: CTCF is detected at most MTR4 peaks and marks chromatin loci gaining accessibility in TKO cells. **A, B.** Number of ZC3H18 (blue bars, A) and MTR4 peaks (purple bars, B) colocalized with the other indicated ChIPseq peaks, compared to the lists of both respective peaks with randomized genomic locations (gray bars). **C,** Heatmaps of CTCF, MTR4, and IgG ChIPseq densities centered on MTR4 peaks detected in Het32. **D.** Venn diagram of all CTCF ChIPseq peaks detected in both Het32 and TKO2 cells, compared to CTCF peaks in ENCODE liver database. Three distinct categories of CTCF peaks were defined : 29080 Het32-only peaks, 30213 TKO-only peaks, and 40913 peaks common to both cell types. **E, F.** RNA levels on Het32-only and TKO2-only CTCF peaks. **E.** Cumulative histograms of proportions of indicated CTCF peak loci expressing RNA as in figure 1C. **F.** Boxplots of mean RNA read counts in chromatin on indicated CTCF peaks, as in figure 1D, in Het (blue) or TKO (orange). **G.** Genome view in IGV, as in figure 3A, of chromatin-enriched RNA, CTCF and MTR4 densities at two regions found to be enriched in TKO-only CTCF peaks, or peaks common to both cell types. **H.** Profiles of CTCF ChIPseq density in Het32 (dark blue), in TKO2 (dark red), compared to IgG profiles in Het32 and TKO2 (salmon and light blue profiles, respectively), centered on the CTCF peaks in the three categories defined in figure 4D (top panels), and their corresponding ATACseq profiles of chromatin accessibility, in all six Het (blue) and TKO (red) clones depicted in figure 1A.

Figure 5: Changes in chromatin accessibility at enhD enhancers affect collagen gene expression in BMEL cells. **A,** Scheme depicting the two types of relationships between enhD and thier closestCTCF, or enhD and their closest genes, analyzed in Fig. 5B and 5C, respectively. **B.** Cumulative histogram of proportions of the different categories of CTCF peaks closest to enhD. **C.** Histogram of proportions of the different categories of genes up- or downregulated by more than 1.5 fold, and closest to enhD. **D.** Log2 fold change in expression of collagen genes by DESeq2 analysis in TKO versus Het cells. **E.** GO term analysis by enrichR of upregulated genes closest to upregulated enhD. F. Genome view of the Col6a1 and Col6a2 locus showing TKO-only CTCF peaks (orange arrows).

Figure 6: HP1 expression affects collagen genes in hepatocellular carcinoma (HCC) patient samples. **A,** Histogram of tumor/non-tumor ratio for the total HP1 expression in 76 HCC patient transcriptomes. **B,** Total HP1 expression levels in two groups of high HP1 (Hi HP1) and Low HP1 (Lo HP1) patients defined by the first and third total HP1 expression quartiles, respectively. **C,** Unsupervised UMAP analysis of patient samples. **D, E,** Histograms of the number of Hi HP1 and Lo HP1 patients clustered according to four Molecular subclasses (MO) sorted by increasing severity of the disease from MO1 to MO4 (D), or showing survival criteria (E). **F, G,** Histograms of the most significant indicated Gene Ontology (GO) terms and their adjusted p values for upregulated genes in Lo HP1 patient samples.

Figure 1

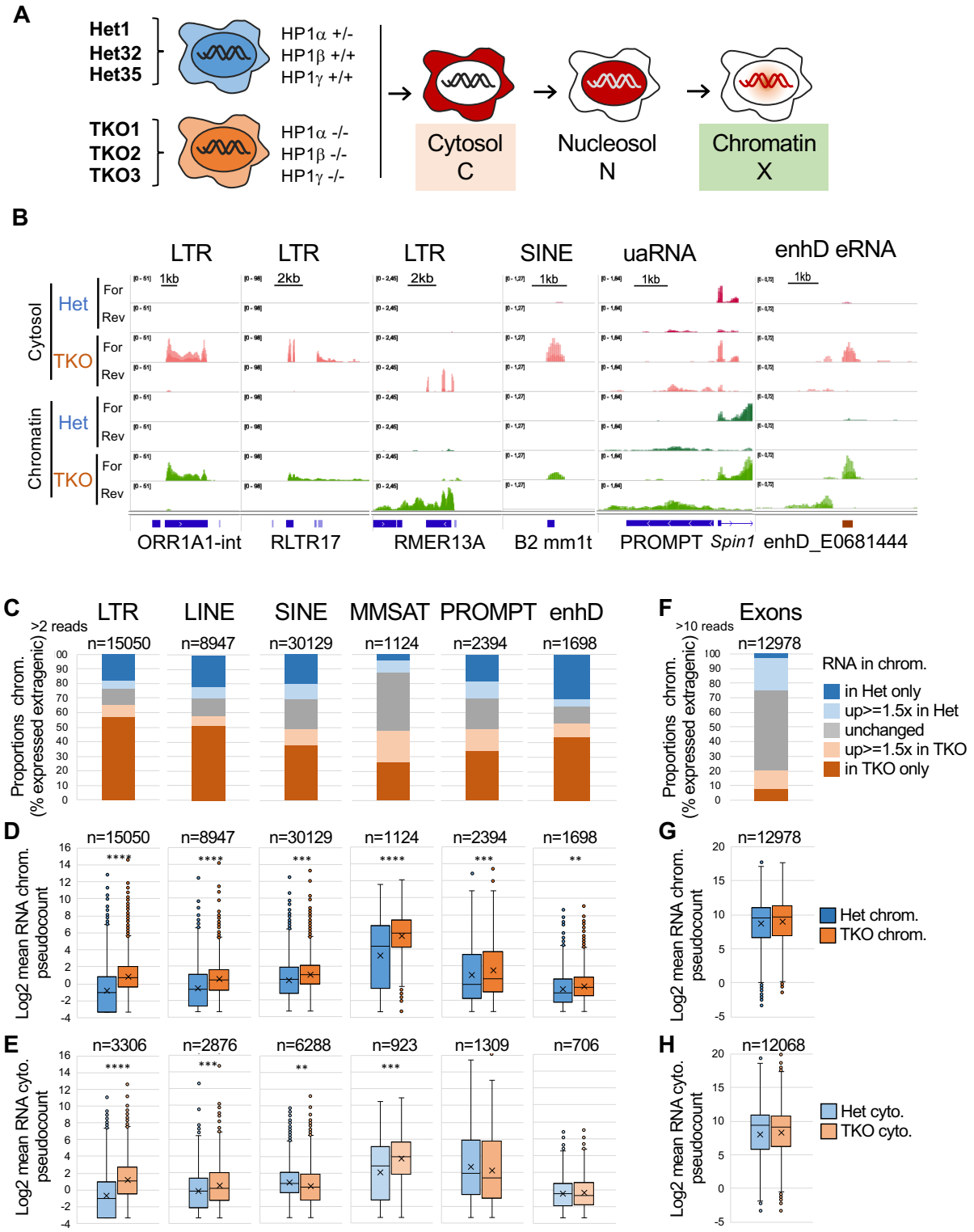

A
Het1
Het32
Het35
HP1α +/-
HP1β +/+
HP1γ +/+
TKO1
TKO2
TKO3
HP1α -/-
HP1β -/-
HP1γ -/-
Cytosol
C
Nucleosol
N
Chromatin
X
B
LTR
LTR
LTR
SINE
uaRNA
enhD eRNA
Cytosol
Chromatin
Het
TKO
For
Rev
1kb
2kb
ORR1A1-int
RLTR17
RMER13A
B2 mm1t
PROMPT
Spin1
enhD_E0681444
C
LTR
LINE
SINE
MMSAT
PROMPT
enhD
>2 reads
n=15050
n=8947
n=30129
n=1124
n=2394
n=1698
Proportions chrom. (% expressed extragenic)
F
Exons
>10 reads
n=12978
RNA in chrom.
in Het only
up>=1.5x in Het
unchanged
up>=1.5x in TKO
in TKO only
D
Log2 mean RNA chrom. pseudocount
G
Het chrom.
TKO chrom.
E
n=3306
n=2876
n=6288
n=923
n=1309
n=706
Log2 mean RNA cyto. pseudocount
H
n=12068
Het cyto.
TKO cyto.

Figure 2

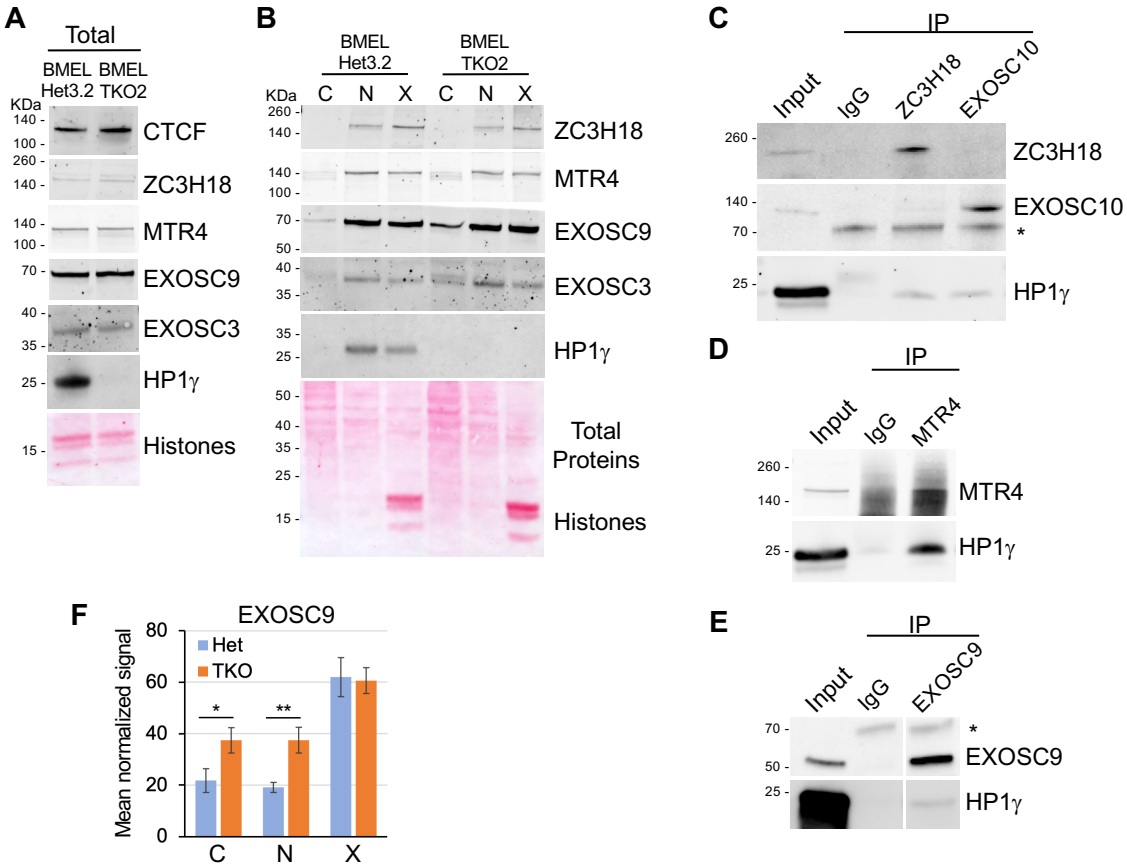

A
Total
BMEL Het3.2
BMEL TKO2
KDa
CTCF
ZC3H18
MTR4
EXOSC9
EXOSC3
HP1γ
Histones
B
BMEL Het3.2
BMEL TKO2
C N X C N X
ZC3H18
MTR4
EXOSC9
EXOSC3
HP1γ
Total Proteins
Histones
C
IP
Input
IgG
ZC3H18
EXOSC10
ZC3H18
EXOSC10
HP1γ
D
IP
Input
IgG
MTR4
MTR4
HP1γ
E
IP
Input
IgG
EXOSC9
EXOSC9
HP1γ
F
EXOSC9
Mean normalized signal
Het
TKO
C
N
X

Figure 3

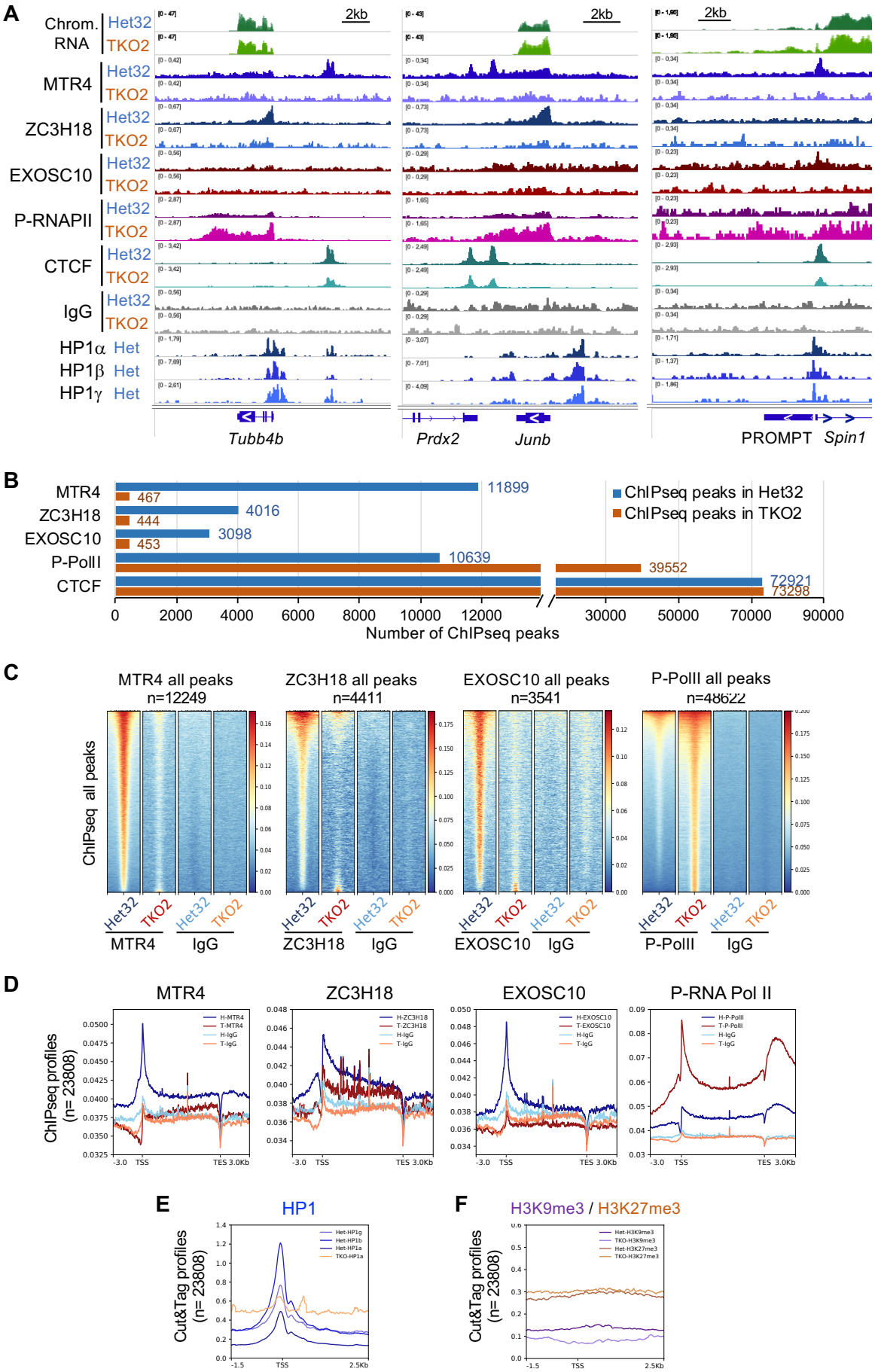

A
Chrom. RNA
MTR4
ZC3H18
EXOSC10
P-RNAPII
CTCF
IgG
Het32
TKO2
HP1α Het
HP1β Het
HP1γ Het
2kb
Tubb4b
Prdx2
Junb
PROMPT
Spin1
B
MTR4
ZC3H18
EXOSC10
P-PolII
CTCF
11899
467
4016
444
3098
453
10639
39552
72921
73298
ChIPseq peaks in Het32
ChIPseq peaks in TKO2
Number of ChIPseq peaks
C
MTR4 all peaks n=12249
ZC3H18 all peaks n=4411
EXOSC10 all peaks n=3541
P-PolII all peaks n=48622
ChIPseq all peaks
Het32
TKO2
MTR4
IgG
ZC3H18
EXOSC10
P-PolII
D
MTR4
ZC3H18
EXOSC10
P-RNA Pol II
ChIPseq profiles (n= 23808)
TSS
TES 3.0Kb
E
HP1
Cut&Tag profiles (n= 23808)
TSS
2.5Kb
F
H3K9me3 / H3K27me3
Cut&Tag profiles (n= 23808)
TSS
2.5Kb

Figure 4

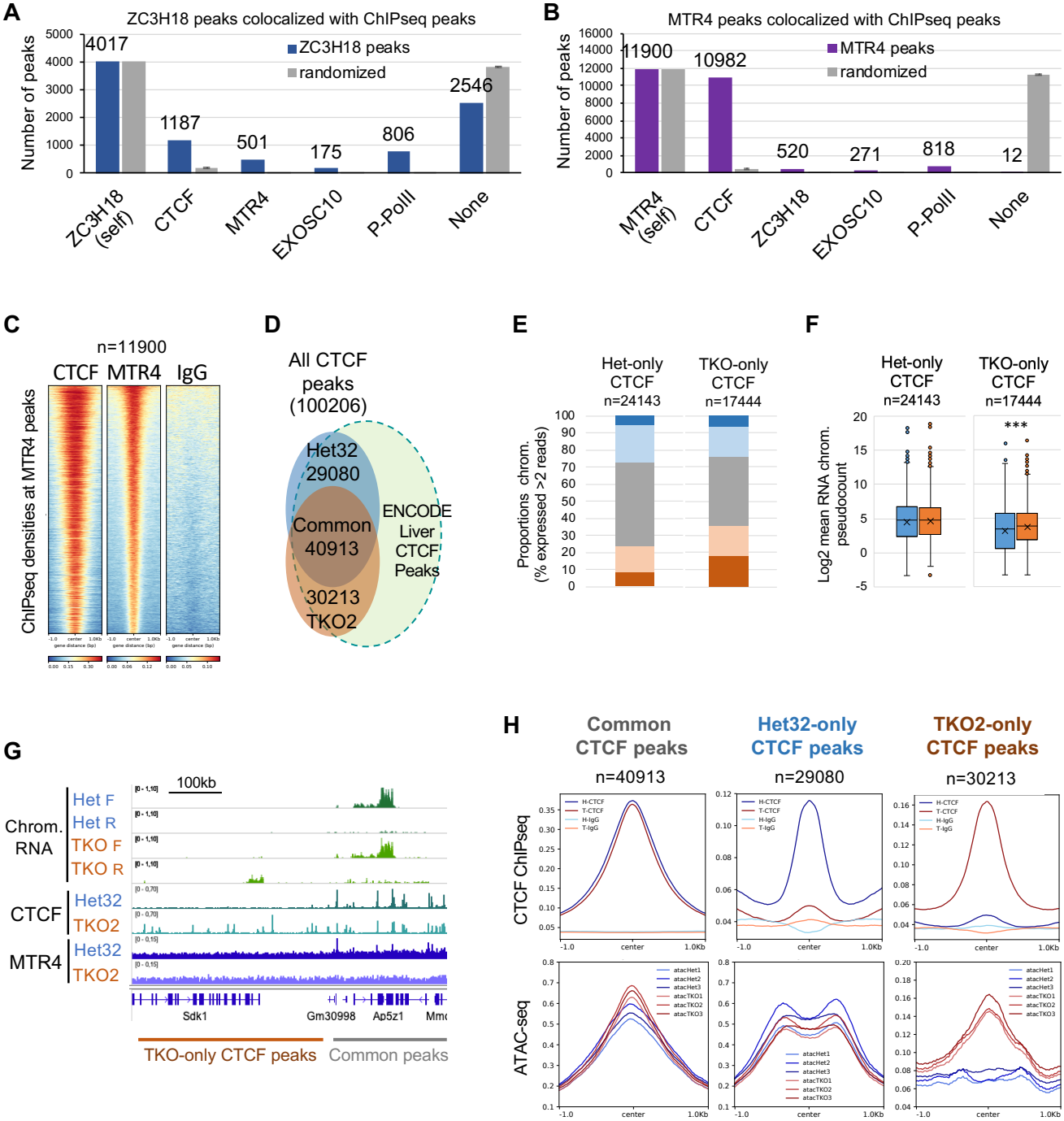
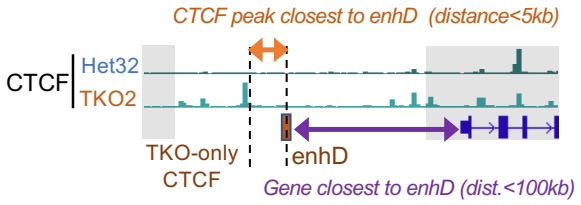
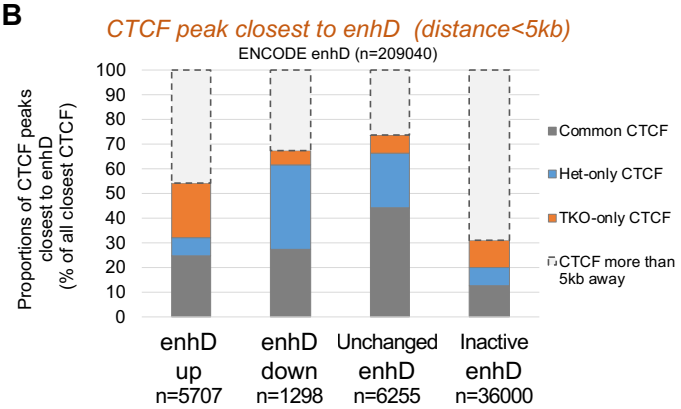

A
ZC3H18 peaks colocalized with ChIPseq peaks
Number of peaks
ZC3H18 peaks
randomized
4017
1187
501
175
806
2546
ZC3H18 (self)
CTCF
MTR4
EXOSC10
P-PolII
None
B
MTR4 peaks colocalized with ChIPseq peaks
Number of peaks
MTR4 peaks
randomized
11900
10982
520
271
818
12
MTR4 (self)
CTCF
ZC3H18
EXOSC10
P-PolII
None
C
n=11900
CTCF MTR4 IgG
ChIPseq densities at MTR4 peaks
D
All CTCF peaks (100206)
Het32 29080
Common 40913
30213 TKO2
ENCODE Liver CTCF Peaks
E
Het-only CTCF n=24143
TKO-only CTCF n=17444
Proportions chrom. (% expressed >2 reads)
F
Het-only CTCF n=24143
TKO-only CTCF n=17444
Log2 mean RNA chrom. pseudocount
***
G
100kb
Chrom. RNA
Het F
Het R
TKO F
TKO R
CTCF
Het32
TKO2
MTR4
Het32
TKO2
Sdk1
Gm30998
Ap5z1
TKO-only CTCF peaks
Common peaks
H
Common CTCF peaks n=40913
Het32-only CTCF peaks n=29080
TKO2-only CTCF peaks n=30213
CTCF ChIPseq
ATAC-seq
H-CTCF
T-CTCF
H-IgG
T-IgG
atacHet1
atacHet2
atacHet3
atacTKO1
atacTKO2
atacTKO3
-1.0
center
1.0Kb

Figure 5

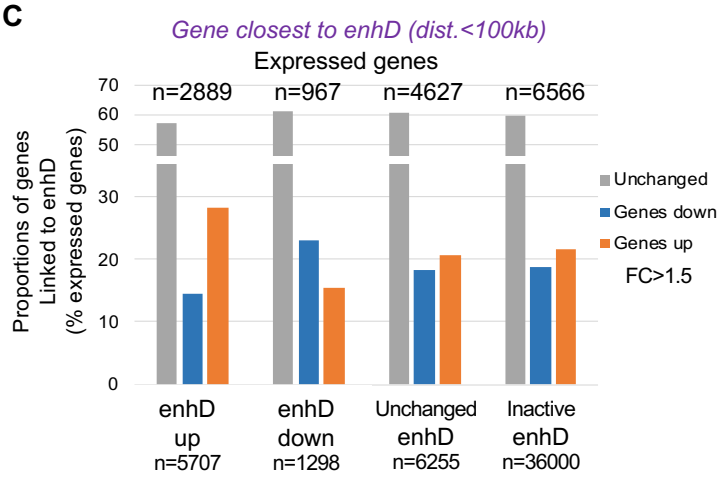
A



B



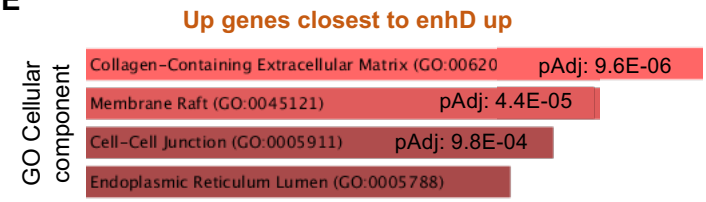
C



D Collagen Transcriptome (DESeq2)

Gene name	BaseMean	log2FC	adj-pVal	
Col1a1	26	2.25	4.98E-02	*
Col2a1	62	5.86	2.54E-05	***
Col4a1	4191	0.31	9.29E-01	
Col4a2	2231	0.16	9.67E-01	
Col4a4	34	1.28	7.31E-01	
Col4a5	1575	-0.80	8.08E-01	
Col5a1	939	0.77	9.10E-01	
Col6a1	38	3.94	1.02E-02	*
Col6a2	45	7.51	9.37E-05	****
Col6a3	43	4.22	1.91E-06	****
Col7a1	153	1.62	7.91E-02	
Col9a3	112	0.42	8.16E-01	
Col11a2	311	-0.06	9.92E-01	
Col12a1	91	-1.61	3.88E-01	
Col15a1	144	-1.22	3.15E-01	
Col16a1	57	4.01	2.66E-04	****
Col17a1	14	1.34	6.46E-01	
Col18a1	2373	0.38	7.94E-01	
Col20a1	300	0.68	8.62E-01	
Col24a1	36	0.62	8.92E-01	
Col27a1	198	2.28	3.91E-02	*

E



F

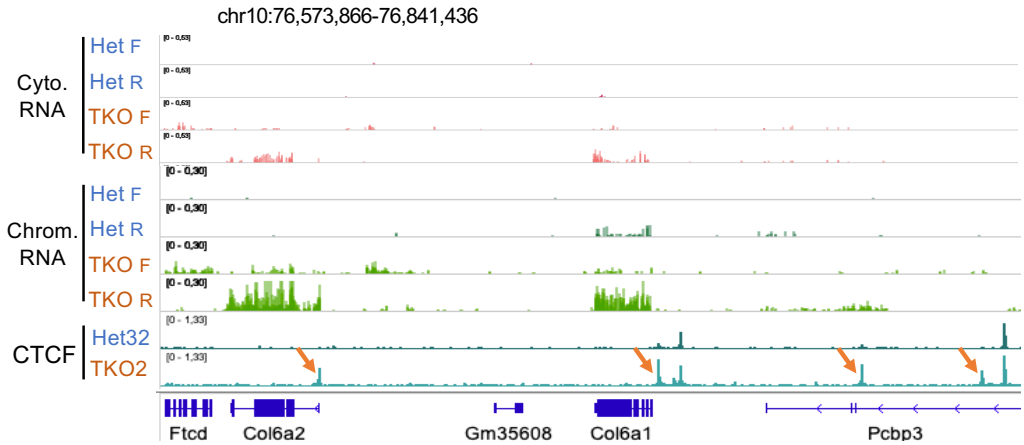
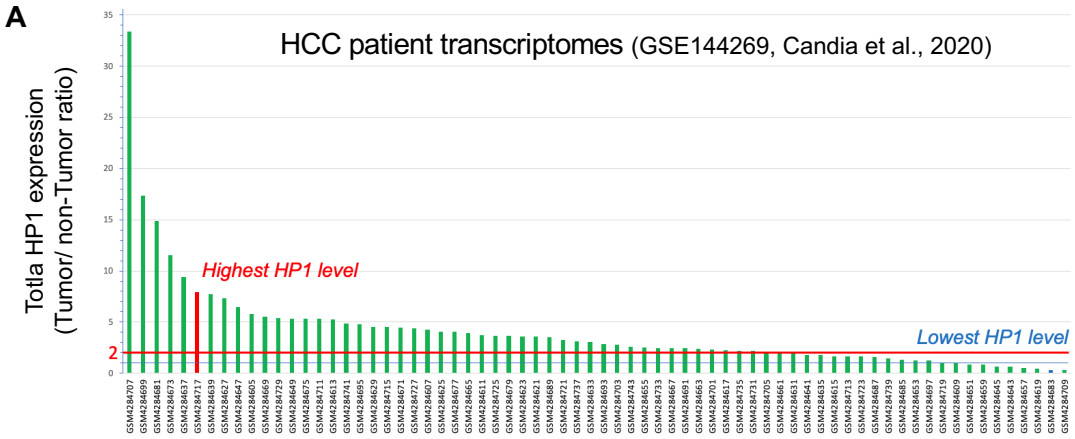
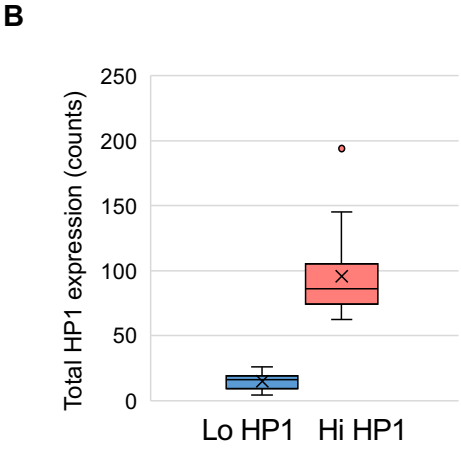
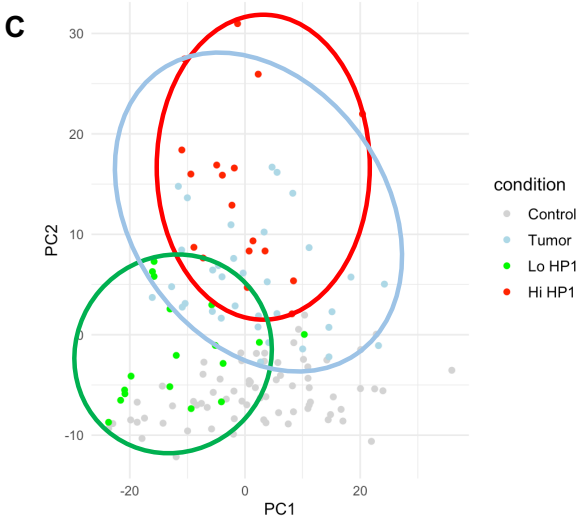
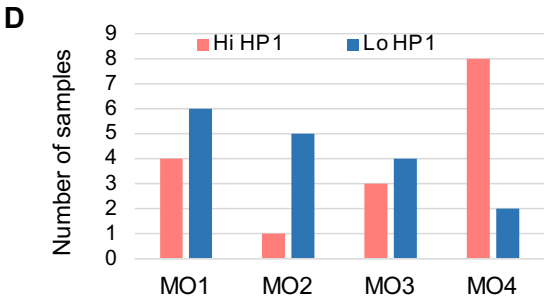


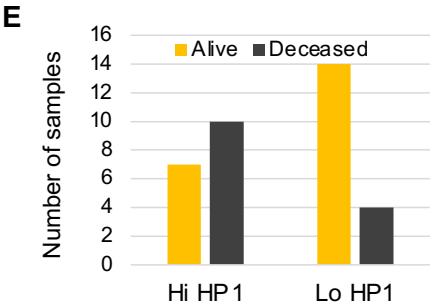
Figure 6

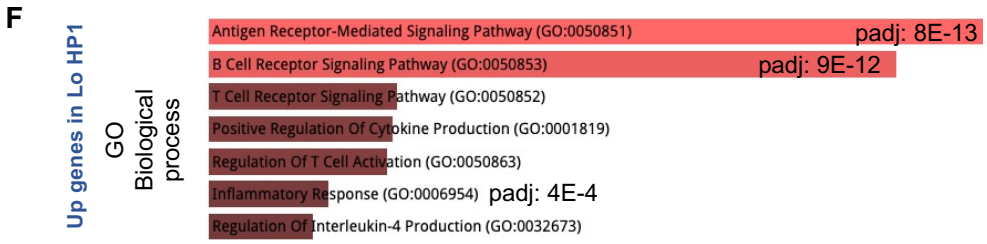

A
HCC patient transcriptomes (GSE144269, Candia et al., 2020)
Totla HP1 expression
(Tumor/ non-Tumor ratio)
Highest HP1 level
Lowest HP1 level
2

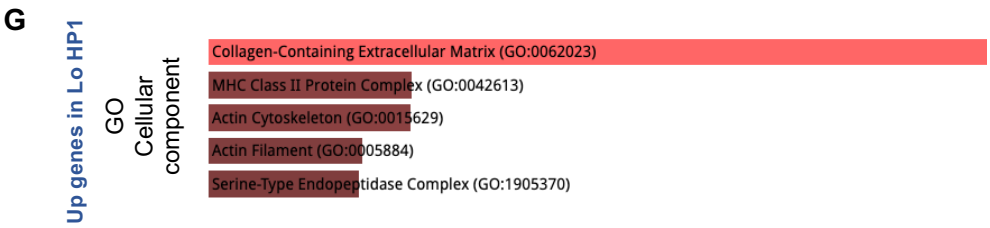

B
Total HP1 expression (counts)
250
200
150
100
50
0
Lo HP1
Hi HP1


C
PC2
PC1
condition
Control
Tumor
Lo HP1
Hi HP1


D
Number of samples
Hi HP1
Lo HP1
MO1
MO2
MO3
MO4


E
Number of samples
Alive
Deceased
Hi HP1
Lo HP1


F
Up genes in Lo HP1
GO
Biological
process
Antigen Receptor-Mediated Signaling Pathway (GO:0050851)
padj: 8E-13
B Cell Receptor Signaling Pathway (GO:0050853)
padj: 9E-12
T Cell Receptor Signaling Pathway (GO:0050852)
Positive Regulation Of Cytokine Production (GO:0001819)
Regulation Of T Cell Activation (GO:0050863)
Inflammatory Response (GO:0006954)
padj: 4E-4
Regulation Of Interleukin-4 Production (GO:0032673)


G
Up genes in Lo HP1
GO
Cellular
component
Collagen-Containing Extracellular Matrix (GO:0062023)
MHC Class II Protein Complex (GO:0042613)
Actin Cytoskeleton (GO:0015629)
Actin Filament (GO:0005884)
Serine-Type Endopeptidase Complex (GO:1905370)