Influenza A virus RNA Polymerase targets the chromatin of innate immune response genes

Jia Yi[1,3], Jessica Morel[4,5], Mickaël Costallat[1,2], Nathalie Lejal[5], Bernard Delmas[5], Christian Muchardt[1,2], Eric Batsché[1,2]*,

1 Epigenetics and RNA metabolism in human diseases. Institut de Biologie Paris-Seine. Université Sciences Sorbonne. 7-9 Quai Saint Bernard, 75005 Paris, France
2 CNRS UMR8256 - Biological Adaptation and Ageing, Paris, France
3 Doctoral School Life science complexity (ED515), University Sorbonne, Paris, France
4 Doctoral School Structure and dynamics of living systems (ED577), University Paris-Saclay
5 INRAE, UVSQ, VIM, Université Paris-Saclay,F-78352 Jouy-en-Josas, France
* corresponding author. Tel: +33 144273477; E-mail: eric.batsche@sorbonne-universite.fr

Abstract

The influenza A virus (IAV) profoundly affects host cell nuclear processes to accommodate for its transcription. In particular, the IAV RNA polymerase (FluPol) associates with host genes to snatch capped 5'-ends from nascent messenger RNAs, that are then used to prime viral transcription. The extend of FluPol-host gene interaction and its actual consequence on host transcription are still poorly characterized. We show here, using genome-wide approaches, that the PA, PB1, and PB2 subunits of FluPol establish contacts with both chromatin and RNA transcripts at promoters, but also within the coding region of genes. These interactions also reach into downstream regions where recruitment of FluPol subunits correlates with previously reported transcription termination defects induced by IAV, indicative of an implication of the FluPol in this viral strategy to limit host cell defense by interfering with transcription re-initiation. The latter is further suggested by a bias in the FluPol genomic targets, enriched in genes associated with anti-viral defense. Together, our observations suggest that FluPol-chromatin binding contributes to targeted dampening of host immune response pathways.

Introduction

Influenza A virus (IAV) is an RNA virus whose replication occurs in the nucleus of infected cells (Herz et al. 1981) and its own transcription is dependent of the cell RNA-polymerase II (RNAPII) activity (Amorim et al., 2007). This feature, atypical for RNA viruses arise from the activity of its RNA viral polymerase complex (FluPol) formed by the intricate association of the three subunits PA,PB1,and PB2. Indeed the FluPol requires RNA primers originating from the endonucleotidic cleavage of the 5'-end of capped cellular transcripts by the PA subunit to produce viral transcripts (Plotch et al., 1979; Dias et al., 2009; Reich et al., 2014). This process called ''cap snatching'' allows FluPol to produce viral mRNAs that can be translated into proteins from chimeric transcripts containing the 10-15 nucleotides downstream of the cap from host cell RNAs and the rest composed of viral sequences generated by copying the negative sense genomic vRNA(-).

Several observations suggest that FluPol steals capped RNAs directly at the promoters of genes. Firstly, FluPol interacts with RNAPII phosphorylated at the serine-5 (S5p) on the C-terminal domain (CTD) (Engelhardt et al., 2005; Martínez-Alonso et al., 2016; Lukarska et al., 2017), a phosphorylated form enriched at promoters (Komarnitsky et al., 2000; Boehm et al., 2003). Binding of FluPol to the RNAPII S5p-CTD activates the capping enzyme to proceed to the first step of the mRNA maturation (Ho and Shuman, 1999). Deep-sequencing of the 10-15 nucleotides, which prime the viral mRNAs indicates that about half of these heterogenous sequences matchs with the 5-ends of polyadenylated mRNAs (Sikora et al., 2014). Analysis by mNET-seq of viral RNAs co-immunoprecipitated with the RNAPII further showed a 1000-fold enrichment in S5p-RNAPII compared to the total pool of RNAPII, suggesting that FluPol generating the vRNA segments requires the association with the S5p-RNAPII (Bauer et al., 2018). FluPol was also found to interact with CHD1 (Marcos-Villar et al., 2016), a chromatin remodeler involved in transcriptional

elongation, binding to histone 3 trimethyled at lysine 4 (H3K4me3), an epigenetic hallmark of active promoters (Park et al., 2014). Finally, the interaction between FluPol and the nuclear RNA exosome complex was reported to favor the recruitment to promoters (Rialdi et al., 2017).

These data suggest that the cap-snatching occurs in the early phase of transcription elongation and provides an attractive explanation for the coupling of the vRNAs synthesis with the activity of the RNAPII, ensuring FluPol to find the capped-RNA primers.

However, the question remains unclear whether FluPol catches the 5'-end termini of host cells on promoters or also at other positions where S5p-RNAPII can be found. Indeed, S5p-RNAPII is also recruited to enhancers to produce the Capped eRNAs. Because active enhancers are more abundant than promoters (ENCODE Project Consortium et al., 2012; Shen et al., 2012; Zhu et al., 2013), these may be an alternative source of Capped RNAs for FluPol.

Initial approaches to study the "cap snatching" process have suggested that the virus preferentially targets the most abundant, highly transcribed; host mRNAs (Sikora et al., 2014). A more recent study based on an enrichment strategy by hybridization with vRNA likewise identified the protein-coding mRNAs as the main source of Capped extremities (Rialdi et al., 2017). However, in both studies, the limitations of the RNA sequencing may have prevented the detection of less abundant sites of RNA production, while the localization of the initiation sites may be biased by the very short human sequences snatched by the virus (10-15 nucleotides) that cannot always be mapped unambiguously on the human genome.

More recently, the IAV was shown to cause a global RNAPII termination defects, by inhibiting the termination complexes (Bauer et al., 2018; Zhao et al., 2018) which, in turn, disrupts the cohesin/CTCF-anchored chromatin loops (Heinz et al., 2018). These chromatin structures are extensively modified by the IAV infection, resulting in an opening of the topologically associating domains (TAD) to allow RNAPII long run-through events in intergenic regions.

We have investigated whether IAV transcription could also use other sources of capped RNAs, such as eRNAs from enhancers, for example. To this end, we aim at identifying additional chromatin loci where cap-snatching might occur by performing a series of chromatin immunoprecipitation (ChIP) using different antibodies recognizing the FluPol subunits PA, PB1 and PB2. We also generated viruses with the FluPol PA subunit labeled with a highly efficient V5 epitope to perform chromatin immunoprecipitations. The nascent transcripts associated to the chromatin at location recruiting FluPol were then analyzed by deep-sequencing. This unprecedented approach surprisingly revealed that FluPol remains linked to the nascent RNAs (probably through the RNAPII association) during phases of transcriptional elongation and termination. Together, these observations suggested that FluPol participates in inhibiting host gene transcriptional termination, a mechanism that facilitates evasion of the host-cell defense.

Results

High-throughput sequencing of chromatin-associated RNAs in IAV-infected cells revealed deep transcriptional perturbations

To directly examine the effect of influenza virus infection on host transcriptional activity, rather than its combined effect of transcription and RNA maturation, we carried out sequencing of RNA bound to the chromatin in human lung epithelial A549 cells infected at a high MOI (MOI = 5) with an H1N1 influenza virus (A/WSN/33). We observed that at 6 hours post infection (hpi), when focusing of the 3/4 most expressed genes, ~40 % were differentially regulated as evaluated by the DESeq2 package. (**Figure 1A and SupFig S1A**). This dysregulation was more extensive than those reported previously for this early times of infection (Bauer et al., 2018; Heinz et al., 2018; Zhao et al., 2018; Chung et al., 2018; Wang et al., 2020; Zhou et al., 2017).

The strong perturbation of the cellular transcriptome by IAV favors heterogeneous transcript distributions between samples (Bauer et al., 2018; Zhao et al., 2018), and the performance of statistical normalization may be affected (Robinson et al., 2010; Tarazona et al., 2011)(Conesa et al., 2016). Therefore, we used an alternative normalization method based on a set of reference genes. We reasoned that, after just 6 h of infection, pools of transcripts with long half-time would remain essentially unaffected. We first confirmed that mRNAs identified as long-lived in HeLa cells (Tani et al., 2012) preserved this characteristics in A549 cells, since there was a robust correlation between RNA life-span and their amount. Indeed, in both A549 chromatin-associated RNAs and total RNA extract, a higher proportion of shorter-lived RNA was among the weakest expressed genes, and inversely a higher proportion of longer-lived RNA was in the highest expressed genes (**SupFig S1B)**. Thus, considering that the 20% most expressed transcripts (> 300 rpkm) with long half-lives (> 15 h), we confirmed that approximately 40% of the 3/4 most expressed genes were differentially regulated (**Fig. 1B, SupFig S1C**), but with only a 75% overlap with the genes identified by DESeq2 pipeline (**Fig. 1C**).

Earlier studies have shown that IAV-dependent transcriptional termination defects allowed the production of ''downstream-of-gene'' transcripts (DoGs) from all active genes (Bauer et al., 2018; Zhao et al., 2018). This phenomenon, that may largely contribute to the bias we corrected, was well observed in chromatin-bound RNAs, with a general decrease in apparent exon coverage associated with an increase in the abundance of reads beyond the site of polyadenylation (**SupFig S1D** and **Fig. 1D**). The relevance of chromatin associated RNAs for the observation of DoGs was illustrated by their efficient vi-

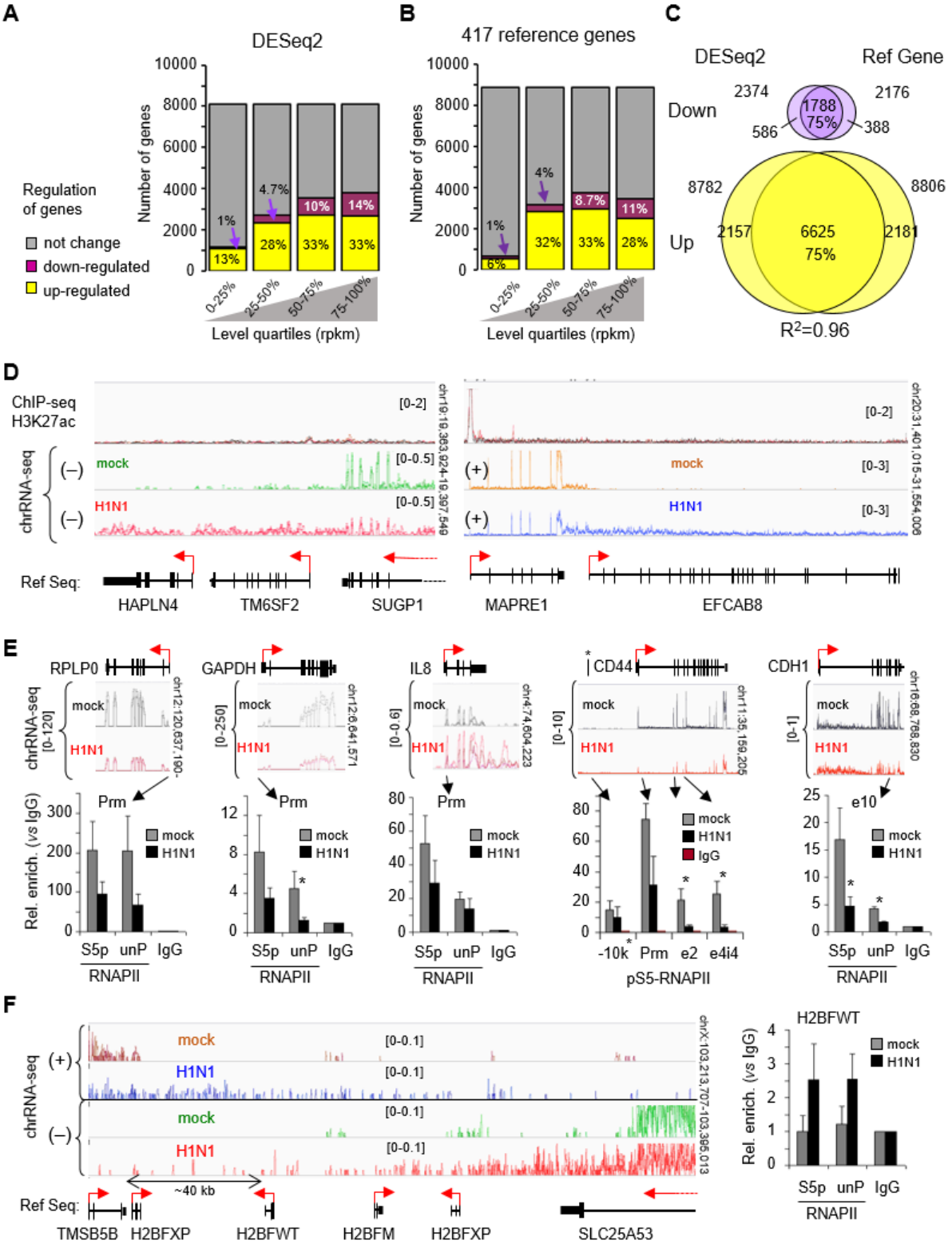


Figure 1 : RNAs associated to the chromatin illustrate cell transcription modifications induced by 6 h of infection by IAV H1N1 in A549 cells
A) Differential RNA levels between H1N1-infected cells versus non-infected cells (mock) using DESeq2 package. The genes were considered regulated for a fold change >1.5 and the adjusted p-value<0.05. Genes were ranked in quartiles as indicated. **B**) Differential RNA levels using the normalization based on the average variation of 424 reference genes expressing higher level (> 10 log2 cpm) of transcripts with long half-life (> 15 h), (**SupFig S1B**) from a paired t test (two-tails). Genes were ranked in quartiles from the mean of the two conditions. **C**) Venn-diagram for up and down regulated genes as evaluated by the two methods. R^2 is the Pearson's correlation coefficient comparing the two methods. **D**) Examples illustrating the variation in the distribution of reads in H1N1 infected cells from IGV visualization of indicated loci. The RNA-seq independent triplicate were overlayed in the same color as indicated, for each strand orientation. The top track shows the enrichment H3K27ac on the chromatin of infected or mock cells. The scale range of the tracks were indicated in brackets. The bottom track (Ref Seq) indicates the position of the genes and their orientation by the red arrows. **E, F**) ChIP assays with antibodies against S5-phosphorylated (S5p) or non-phosphorylated (unP) RNAPII of infected and non-infected cells. Immunoprecipitated DNA was quantified by qPCR with primers targeting indicated regions. Enrichment are expressed relatively to the signal obtained for ChIP with negative IgG control**.** Values are means (±deviation) of at least three independent experiments. Statistical significance of the differential levels between infected and non infected cells were evaluated using Student's t-test (two-tailed), with $P < 0.05$ (*).

sualization. Even on genes reported previously to not producing these RNAs (LY6E, APOL1, DEFB1 and IFI6) (Zhao et al., 2018), it was possible to observe transcripts produced by RNAPII terminal run-through (**SupFig S1E**). On the other hand, in agreement with Bauer et al. (Bauer et al., 2018), histone genes did not produce DoGs (**SupFig S1F**), due to the alternative maturation mechanism active at these genes and not relying on the canonical polyadenylation complex (Kolev and Steitz, 2005).

To gain further understanding of the mechanisms perturbing transcription, we carried out RNAPII ChIP-seq on the infected cells. This revealed that the apparent decrease in exon coverage in the RNA-seq data was also concomitant with an apparent decrease in RNAPII accumulation at the promoters of highly expressed genes not affected by the infection (GAPDH, RPLP0), at induced genes (IL8), and on the transcribed region of moderately expressed genes (CD44, CDH1) (**Fig. 1E**). Inversely, we noted an increased accumulation of RNAPII, possibly as a consequence of a defective termination at nearby genes, at loci corresponding to genes not expressed in A549 cells, such as H2BFWT, even after IAV infection (**Fig. 1F**). Re-analysis of publicly available ChIP-seq data from primary human monocyte-derived macrophages (MDMs) infected with IAV and harvested at 6 hpi (Heinz et al., 2018) revealed a very similar decreased RNAPII recruitment at expressed genes coupled to its increased accumulation on intergenic regions and at genes not expressed in these cells (**SupFig S1G).**

Quantification of chromatin-bound RNAs confirms that the IAV infection induced defects of transcriptional termination.

We next re-examined the DoGs in the light of their improved detection provided by the sequencing of chromatin-associated RNA. Visual examination in a genome browser showed that numerous genes produced DoGs in the absence of IAV infection, illustrating the intrinsic imperfections of the termination machinery (see examples **Fig. 1D, SupFig. S1D,E)**. As production of DoGs has been proposed to participate in the global reduction in host cell transcription, we examined the

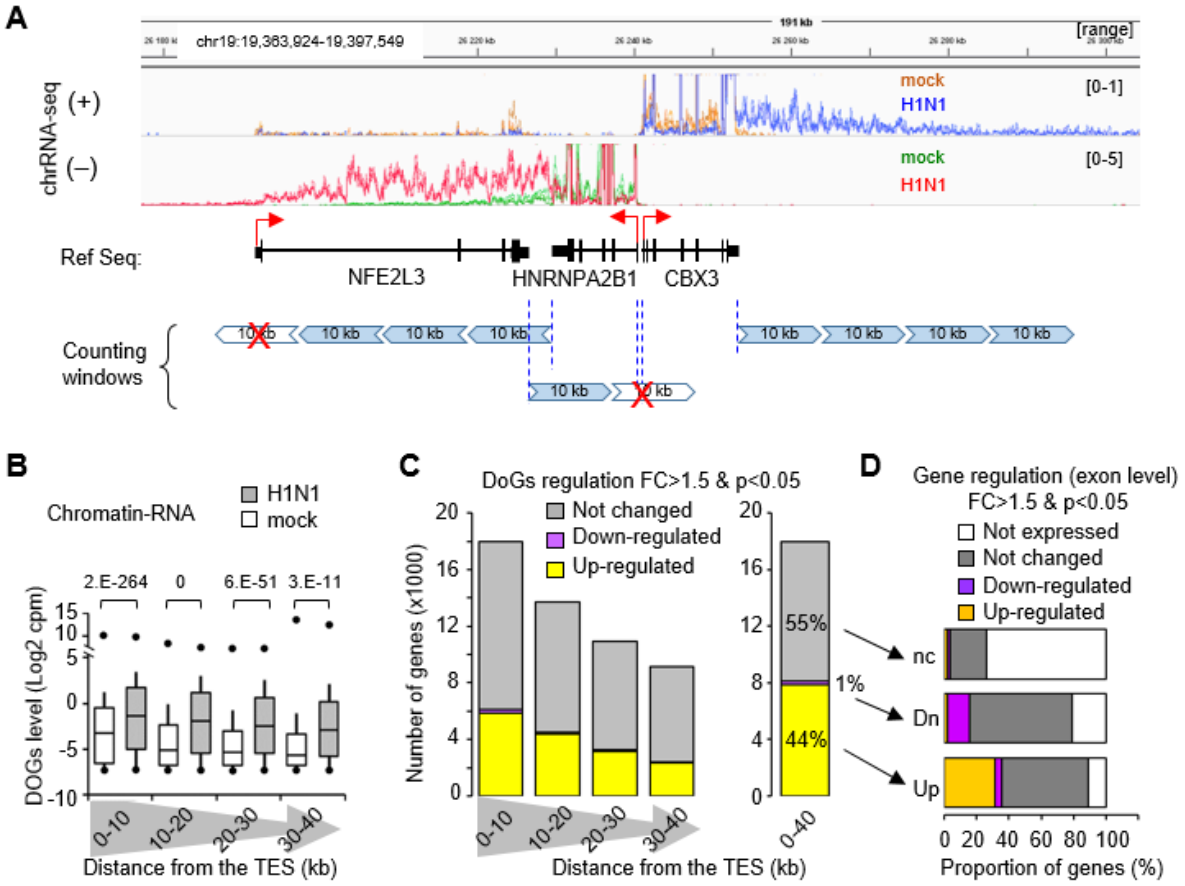


Figure 2 : chromatin RNAs were enriched in DoGs within the 3'end region of the genes. A) Strategy to quantify the 3'end transcripts downstream to the termination sites (TES). Oriented reads were counted in 10 kb windows downstream to the TES extending until 40 kb. The downstream-10 kb windows which overlapped with the transcription start site of another Ensembl gene in both orientation were not taken in account (red crosses). **B)** For each selected genes the indicated 10-kb windows from the TES were quantified for their DoGs levels evaluated in Log2(cpm). Data are displayed in box plot where the bottom line is the 1st centile, the box corresponds to the second and third quartile, the top line to the 9th centile, and the central line in the box is the median. Small circles indicate the minimum and the maximum. **C)** Counts of genes for which a 3'end downstream 10-kb windows exhibited differential transcription level. These windows were considered modified for a p-value<0.05 (paired t test, two tailed) and a fold change >1.5. The right panel is the same quantification for all pooled windows. The transcription downstream to TES covering the 40 kb was considered upregulated when the majority of 10-kb windows were upregulated (yellow), and the inverse for the down-regulation (magenta). The 40 kb region was considered unchanged (grey) in case of equality or when the 10 kb windows were themselves found unchanged, or when there was not a significant differential levels of DoGs on the four 10 kb-windows, as evaluated by a paired t-test (two-tailed, p<0.05). **D)** For each category of 3'end downstream region, the regulation of the genes (as evaluated in **Fig. 1B**) is indicated in proportion.

correspondence between gene expression and production of DoGs. In the aim to evaluate how the DoGs were modified in infected cells, oriented read density were measured in 10 kb windows downstream to the TES extending until 40 kb (**Figure 2A**). This DoGs quatification showed a robust overall upregulation of these RNA species even at sites located up to 40kb away from the TES of genes (**Fig. 2B**). We estimated that 44% of these downstream regions were upregulated by more than 1.5 times, while almost none were down-regulated (**Fig. 2C**). Among the upregulated DoGs, a majority (~55%) were related to genes unaffected in their expression in the presence of IAV (**Fig. 2D**), consistent with an interference between the biology of this virus and the host termination machinery. Yet, upregulated DoGs also originated from up-regulated genes (approximately

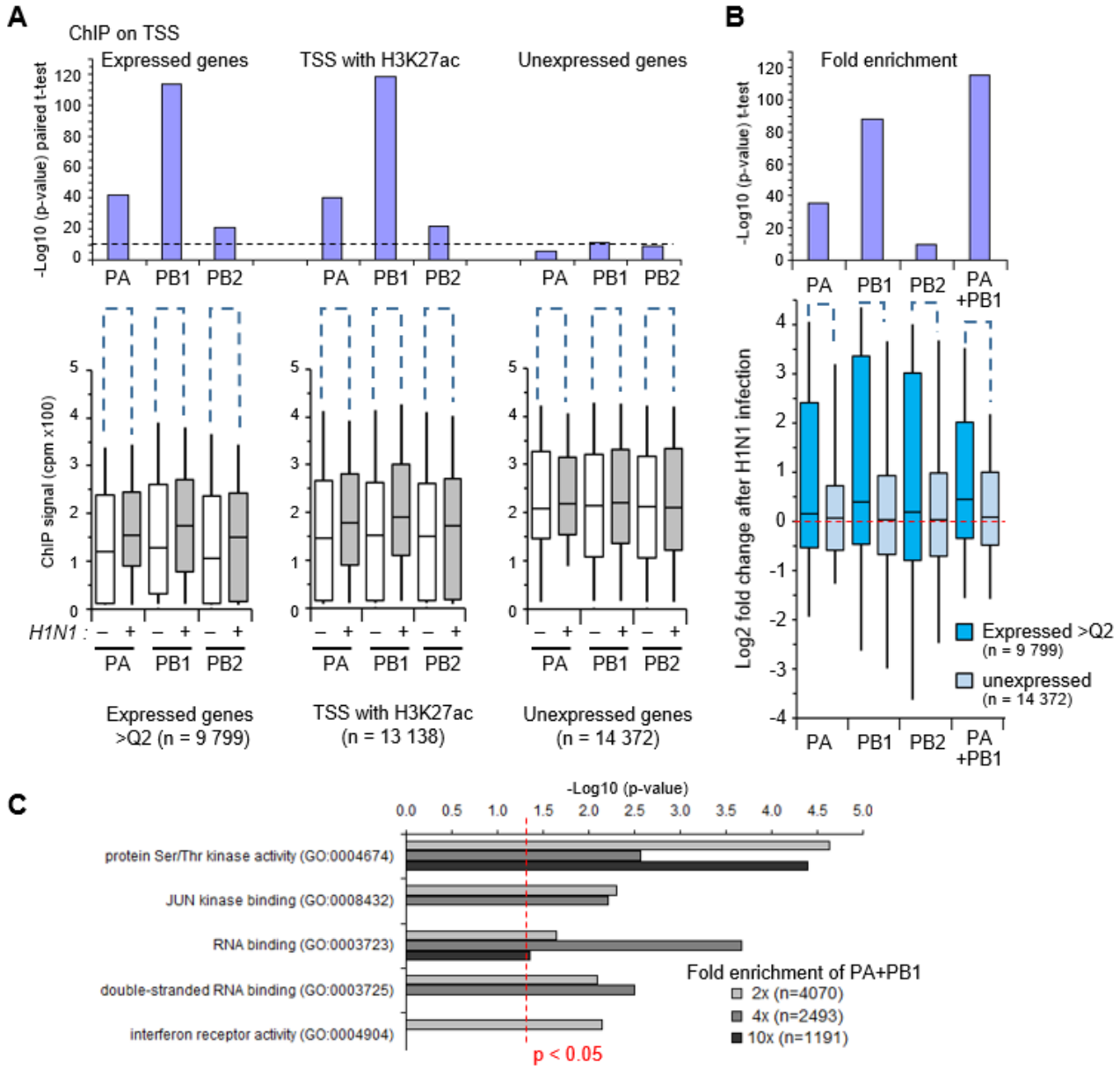


Figure 3 : Enrichment of FluPol subunits on the chromatin. Chromatin of non-infected (-) or infected (+) A549 cells were immunoprecipitated with indicated subunits of the FluPol and analysed by deep-sequencing. For each 500 bp-windows centered on annotated transcriptional start sites (TSS), the DNA enrichment were the means of replicates expressed in cpm (multiplied by 100 to facilitate the representation). Only TSS separated from other TSS by 5 kb were considered to avoid potential misleading with close neighbors or overlapped genes. **A**) Bottom box plots illustrate the distribution of the TSS from genes expressed more than the mediane (> Q2), of TSS with a H3K27ac signal greater than the mediane of the H3K27ac signals, or of TSS from non-expressed genes. The bottom and top lines of the box plots indicate the 1st and the 9th centiles. The upper graphs show the p-values evaluated on log2 normalized cpm by a paired t-test (two-tailed) assessing the enrichment induced by the H1N1 infection on each TSS in indicated categories. The dotted line indicate the p-value threshold based on the unexpressed genes. **B**) Fold change of enrichment for indicated TSS categories were the averages of TSS differences between log2 cpm from infected and non-infected cells. Statistical analysis has been carried out the log2 fold change using Student t-test (two-tailed) to compare TSS from unexpressed genes (light blue) and genes expressed more than mediane (blue). The dashed red line indicates the absence of change. The p-values were indicated above the box plots. In each case, the number (n) of tested TSS were indicated. PA+PB1 analysis were conducted by combining their ChIP-seq signals. **C**) Pathway analysis of genes on which PA+PB1 were enriched 2, 4, and 10 times on TSS between infected and mock cells. P-values of Gene Ontology (GO) molecular function 2021 were evaluated through Enrichr, and the threshold of $p<0.05$ is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category..

30%), suggesting that increased production of DoGs may also be caused by gene activation, and that production of DoGs does not *per se* prevent gene activation, at least at early times of infection. Finally, we noted that approximately 10% of the upregulated DoGs, as well as 70% of the non-significantly modified DoGs, that have been associated with unexpressed genes seemed produced by genes not expressed in the A549 cells. Manual verification indicated that these cases corresponded to genes entirely traversed by DoGs originating from neighboring genes.

Genome wide Analysis revealed that FluPol is recruited to gene promoters

We next examine the genome-wide recruitment of FluPol, using antibodies directed against each of its PA, PB1, and PB2 subunits. When tested on extracts from infected cells, these antibodies all resulted in a co-immunoprecipitation of phosphorylated RNAPII (**Sup Fig S3A**). confirming earlier observations documenting an interaction of the FluPol with the pS5-RNAPII.

Chromatin immunoprecipitation assays carried out with these antibodies. The relatively high background signal suggested that FluPol established only indirect contacts with the chromatin. However, examination of the ChIP signals on 500 bp-windows centered on promoter TSSs, showed a significantly increased signal upon viral infection, which was more obvious on expressed genes than on unexpressed genes, validating the specificity of these antibodies **(Fig. 3A)**. To avoid interference from neighboring genes, we focused on genes located at least 5 kb away from other genes. This was confirmed also by the increased signal observed at active promoters harboring H3K27ac marks, that was more important than at TSS of unexpressed genes (compare middle graph to the left one, **Fig 3A**). The differential TSS enrichment induced by the infection for each subunit was significantly greater on expressed than on unexpressed genes (above the red line, **Fig. 3B**). The PB2 antibody seemed to be the less efficient at discriminating the specific signal induced by the infection on expressed genes versus unexpressed genes (**Fig 3A, B**). Finally, the combined signal for the PA and PB1 subunits was more obviously enriched upon infection on expressed genes than on non-expressed genes, and more than the subunits taken individually (right bars in **Fig 3B**). In contrast, no enrichment upon H1N1 infection was detected in introns of genes, with no significant difference between expressed and silenced genes (**Suppl. Fig. S3B**). Together, these observations were consistent with a recruitment of the FluPol subunits to sites of transcriptional initiation.

We next examined whether the recruitment of FluPol was governed only by transcription, or whether its recruitment also involved site specificity. To that end, we carried out gene ontology analysis on the genes strongly recruiting PA+PB1 (2-, 4-, and 10-fold enrichment between infected and mock cells). This revealed a significant enrichment in genes associated with kinases involved in cell signaling, in RNA binding proteins especially of the double-stranded RNA recognition, as well as receptors to IFN (**Fig. 3C**). Furthermore, we observed a particular enrichment of the FluPol on promoters of OASL, OAS1, OAS2, DDX58/RIG-I, DDX60, ILF2, and IFNAR1/2, INFGR1/2 (**Sup. Table 1**), suggesting that FluPol specifically targets genes involved in the innate cell-defense and possibly interferes with their expression.

Chromatin-bound RNAs associated to the PA subunit revealed that FluPol is associated to the RNAPII in elongation.

As an alternative approach to study FluPol recruitment to the host genome, we next implemented an RNA-ChIP approach to explore the range of nascent host RNA species associating with the viral polymerase. As available antibodies were of insufficient quality for this more demanding approach, we generated a V5-tagged version of the PA subunit of FluPol. The resulting recombinant H1N1-derived IAV retained its ability to infect the A549 cells. Immunoprecipitation with an anti-V5 antibody under conditions used for RNA-ChIP (see Material &Methods) on cells infected with the V5-tagged virus, resulted in efficient immunoprecipitation of the tagged-PA (**Fig. 4A** and **SupFig. S4A**).

DNA ChIP experiments conducted with the V5 antibodies confirmed the presence of the engineered FluPol on promoters shown above to recruit the native FluPol, including moderately-to-strongly expressed genes (RPLP0, GAPDH, CD44), and at an induced gene (IL8), and only in infected cells (**Fig. 4B**). Interestingly, V5-PA was also detected with the coding region of CD44 and CDH1, at in regions covered by DoGs, but with no proper transcriptional activity, as illustrated by the H2BFWT/H2BW1 gene, silent in the A549 cells, but crossed by run-away RNAPII (**Fig. 4B**).

We then used the V5-tagged virus and anti-V5 antibody to carry out RNA ChIP-seq. This approach revealed that chromatin-RNAs associating with V5-PA distributed throughout the coding region of a large set of expressed genes (a total of 15299 genes, corresponding to 55% of the examined genes; see example of the OAS genes **Fig. 4C**).

Quantification confirmed that the immunoprecipitated RNAs were preferentially originating from the most expressed genes, although genes expressed at low levels were also contributing (**Fig 4D**).

Specificity of the immunoprecipitation was verified with non-immune IgG, yielding background signal not correlated with gene expression levels (**SupFig. S4B**). In addition, PA-bound RNAs covering the gene bodies concerned genes with an equivalent proportion of differentially expressed genes than what it was observed in the input (compare **Fig 4D to 1B**), suggesting that FluPol-RNA association could be a simple consequence of the highest transcriptional activity in these genes. Importantly, we also identified a small set of genes (1809 genes) at which V5-PA-bound RNAs was enriched only in matured mRNA sequences (only exonic sequences) (**Fig. 4E**). This set of genes was enriched in up-regulated

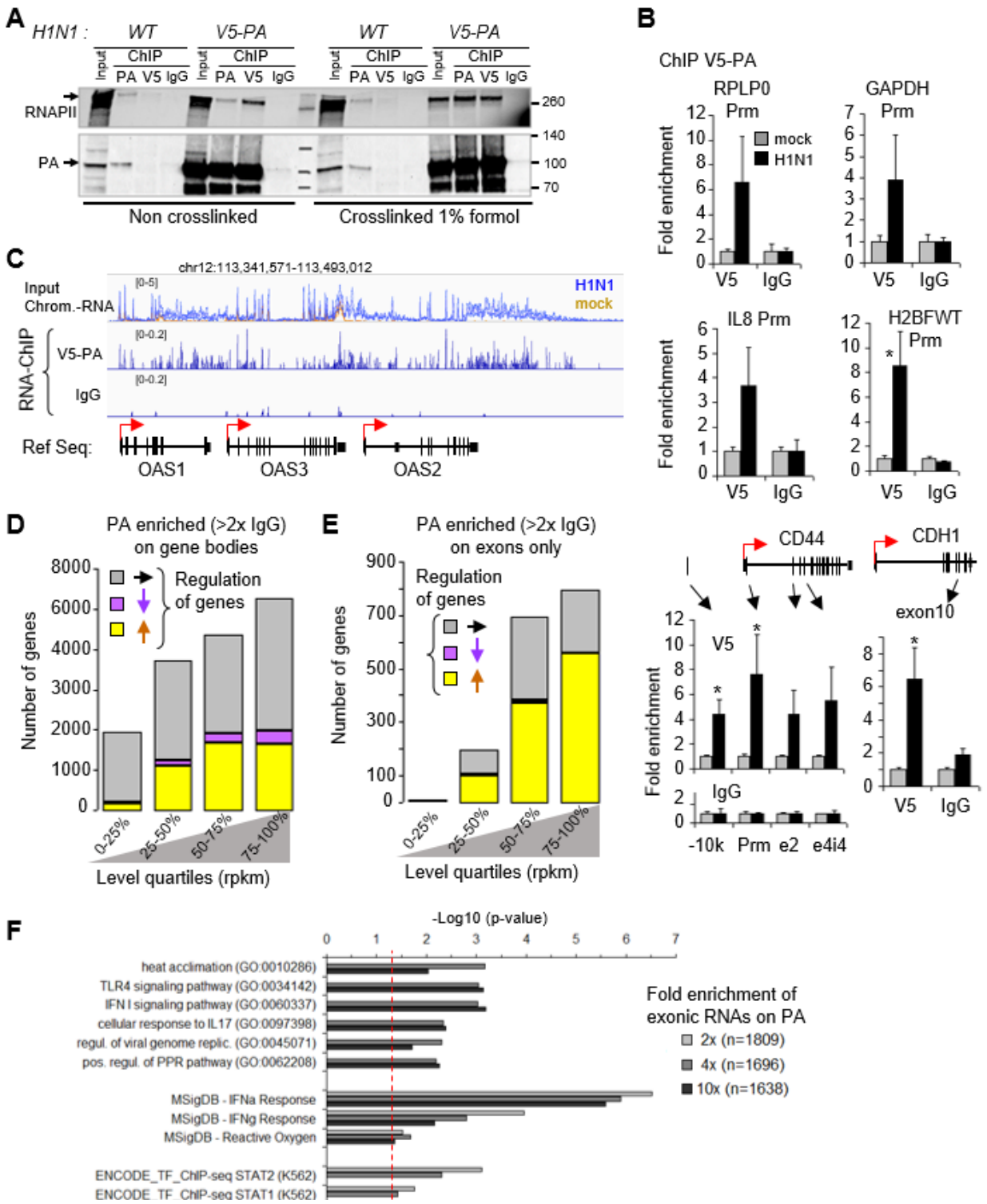


Figure 4 : FluPol subunits on chromatin is linked to nascent RNAs from host-cells A) A549 cells infected cells by WT IAV or V5-PA engineered IAV were crosslinked by 1% formaldehyde or not. Chromatin fraction were subjected to immunoprecipitation with PA, V5 antibodies or non immun IgG. Eluates and 10% of extracts were resolved by western blot. The lower part of the membranes was revealed with the PA antibodies and the upper part with phospho-RNAPII antibodies. **B**) ChIP assays with V5 and IgG antibodies of infected and non-infected cells. Immunoprecipitated DNA was quantified by qPCR with primers targeting indicated regions. Enrichment are expressed relatively to the signal obtained for ChIP in uninfected cells. Values are means (±dev.) of three independent experiments. Statistical significance of the differential levels were evaluated using Student's t-test (two-tailed), with $P < 0.05$ (*) **C**) Example of PA enriched RNAs bound to the chromatin with IGV. The RNA-seq independent triplicate were overlayed in the same color as indicated, for the (+) strand. The scale range of the tracks were indicated in brackets. The bottom track (Ref Seq) indicates the position of the genes and their orientation by the red arrows. **D**) RNAs from ChIP by V5-PA or by non-immun IgG in infected cells were analysed by deep-sequencing. RNAs covering the gene bodies enriched with PA were evaluated versus IgG and have been considered significant when the average of the changes subtracted with the triplicate deviation is superior of two times to the IgG. The number of enriched genes that were upregulated (yellow), downregulated (magenta) or unmodified (gray) were plotted in function of gene expression as determined **Fig. 1B** and ranked in quartiles. **E**) The same analysis as in **C**) was conducted for RNAs covering the exons only. **F**) Pathway analysis of genes on which PA were enriched 2, 4, and 10 times on exonic RNAs between infected and mock cells. P-values of Gene Ontology (GO) biological function 2021, Molecular Signatures Database 2020 (MSigDB), and ENCODE ChIP-seq 2015 were evaluated through Enrichr, and the threshold of $p<0.05$ is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category.

genes) at which V5-PA-bound RNAs was enriched only in matured mRNA sequences (only exonic sequences) (**Fig. 4E**). This set of genes was enriched in up-regulated genes, an enrichment not observed with non-immune IgG (**SupFig. S4C**). Most of the PA-bound exonic RNAs corresponded to upregulated genes in all quartiles of expression level (compare **Fig 4E to 1B**), suggesting a particular gene targeting of FluPol.

Gene ontology analysis on this subset of genes indicated a significant enrichment in genes involved in host-cell defense, associated with Toll-like receptor (TLR) and IFN pathways (**Fig. 4F**). This enrichment was not observed when analyzing the entirety of the 8824 upregulated genes. Thus, this approach confirmed the IFNα/γ pathway as a particular target of the FluPol recruitment, and suggested that binding of FluPol to mature mRNA may participate in the viral strategy to dampen the host-cell defense machinery. Finally, we noted an enrichment of the V5-PA subunit on the transcripts bound to the chromatin matched genes recruiting FluPol at their promoters, as documented by our DNA-ChIP assay, including the OAS genes, DDX58/60, IFIT2/3/M1/M2, and ISG15/20 genes (**Sup. Table 2**). This suggests that cap-snatching at promoters requires or causes a meta-stable association between the FluPol and the messenger RNAs, lasting at least through some of the pre-mRNA maturation steps.

Chromatin-bound RNAs associated to the PA subunit revealed the FluPol is associated to the RNAPII to termination sites and downstream regions.

We next used the V5-PA RNA ChIP-seq to examine in more details the association of the FluPol with with DoGs. The distribution of RNAs bound to the V5-PA subunit suggested an enrichment at the 3'end of genes (TES) and at regions covered by DoGs (**Fig. 4C**). This was verified quantitatively by counting reads in the regions extending 40kb downstream of the TES of all genes, in bins of 10 kb (Left panel, **Fig. 5A)**. When consolidating reads over the entirety of the 40kb-regions, the median density of V5-PA-associated RNAs was 2-fold higher than that observed with the non-immune IgG (right panel, **Fig. 5A**). Among V5-PA-associated DoGs, there were 3 to 4 times more upregulated than unmodified DoGs, and virtually none diminished (compare brown and black bars, **Fig. 5B**). At note, DoGs identified as upregulated in the V5-PA RNA ChIP-seq data out numbered those identified with the RNA-seq data (**Fig. 3C**), and V5-PA-bound DoGs were greatly enriched at DoGs upregulated in the RNA-seq (47%-55%, vs 6%-12% in unmodified DoGs and 0-7% in downregulated DoGs). This strongly suggested that the FluPol is an active player in the mechanism causing the production of DoGs in infected cells and may participate in the viral complex involved in the inhibition termination processes.

To further document the presence of the FluPol in the DoG regions, we reassess the DNA ChIP-seq data with the anti-FluPol subunits, to quantify enrichment of individual subunits at TESs and at downstream regions. In order, to discriminate between TESs and potentially overlapping TSSs, we selected genes separated from each other by at least 10 kb. A significant enrichment in downstream regions of expressed genes versus unexpressed genes was mainly observed for PA and PB1 (**Fig 5C**). The recruitment of these FluPol subunits was in average strongest at regions proximal to the TES (compare quantifications at 500bp, 1kb, 10kb and 40 kb away from the TES), suggesting that coding regions and proximal terminal downstream region are preferential sites for FluPol binding to nascent transcripts.

Gene ontology analysis on genes at which the FluPol is enriched on the 3'-ends revealed an enrichment in genes involved in the cell-defense (pattern recognition receptors (PRR) and NF-kB pathways), in signaling pathways, in oxidative stress associated to cell death, and in the degradation of mRNA 3'end (**Fig 5D**). Genes combining an enrichment in FluPol on their TSS, their exons, and on their TES were few, yet again significantly enriched in genes involved in the cell defense (PLSCR1; IRF5; DDX60; SOD2, CXCL3; SOD2; RELA).

Discussion

We have shown that infection of cells with IAV causes a significant redistribution of transcripts into the chromatin compartment (**Fig. 1C**) which is a consequence of the inhibition of the transcriptional termination process. This phenomenon is caused by a significant increase in DoGs that can be easily observed by analysis of chromatin-associated RNAs in ChIP experiments (**Fig. 2**). This inhibition of gene termination results in a general decrease of the RNAPII at promoters and in gene bodies, and an increase in intergenic regions (**Fig 1D-F, S1D-S1G**).

By using the presence of the 3 subunits of the viral RNA polymerase in the chromatin of infected cells, we showed that FluPol is not only associated with promoters where it is thought to snatch the caps of RNAPII-initiated RNAs. In fact, our observations suggest that FluPol remains associated with RNAPII during elongation, at transcriptional termination sites and also in downstream regions where RNAPII continues to run in intergenic regions (**Fig 3, 4, 5**). FluPol was not only bound to the DNA of transcriptionally active genes but remains associated with host cells nascent transcripts (**Fig 4C-E, 5AB**), with a concentration to DoGs. This is the first time to our knowledge that FluPol is observed along the gene bodies and beyond. The association with DoGs suggests that FluPol could, in addition to NS1, also participate in the process of transcriptional termination inhibition.

Analysis of the regulatory pathways defined by the genes for which FluPol is most strongly enriched indicates that FluPol particularly targets genes of the anti-viral innate response. The participation of FluPol in the inhibition of transcriptional termination particularly on cellular defense genes would add a function of FluPol, in addition to cap-snatching, to participate in the viral escape strategy.

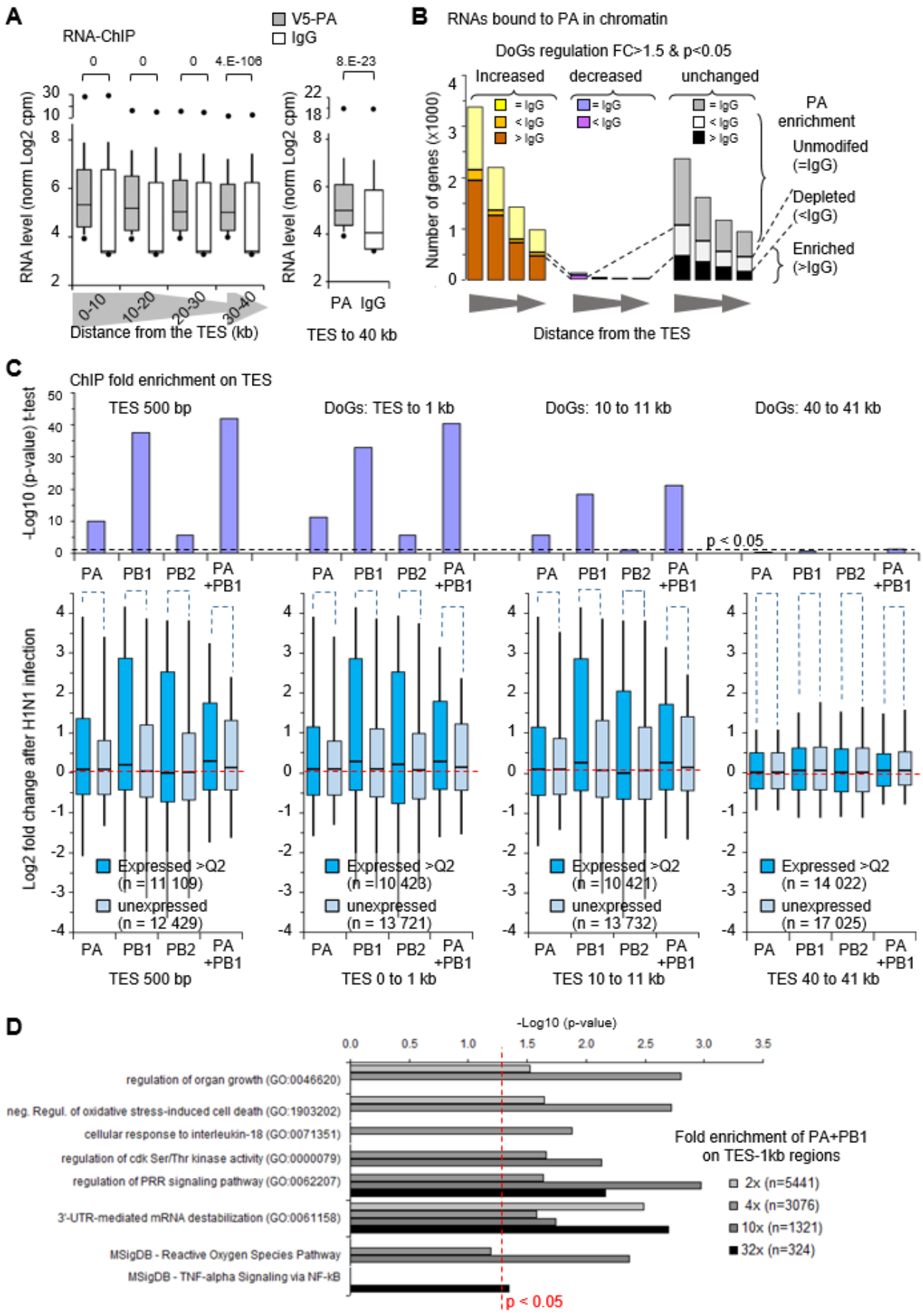

A
RNA-ChIP
V5-PA
IgG
0
0
0
4.E-106
8.E-23
RNA level (norm Log2 cpm)
0-10
10-20
20-30
30-40
Distance from the TES (kb)
PA
IgG
TES to 40 kb
B
RNAs bound to PA in chromatin
DoGs regulation FC>1.5 & p<0.05
Increased
decreased
unchanged
= IgG
< IgG
> IgG
Number of genes (x1000)
PA enrichment
Unmodifed (=IgG)
Depleted (<IgG)
Enriched (>IgG)
Distance from the TES
C
ChIP fold enrichment on TES
-Log10 (p-value) t-test
TES 500 bp
DoGs: TES to 1 kb
DoGs: 10 to 11 kb
DoGs: 40 to 41 kb
p < 0.05
PA
PB1
PB2
PA +PB1
Log2 fold change after H1N1 infection
Expressed >Q2 (n = 11 109)
unexpressed (n = 12 429)
Expressed >Q2 (n = 10 423)
unexpressed (n = 13 721)
Expressed >Q2 (n = 10 421)
unexpressed (n = 13 732)
Expressed >Q2 (n = 14 022)
unexpressed (n = 17 025)
TES 500 bp
TES 0 to 1 kb
TES 10 to 11 kb
TES 40 to 41 kb
D
-Log10 (p-value)
regulation of organ growth (GO:0046620)
neg. Regul. of oxidative stress-induced cell death (GO:1903202)
cellular response to interleukin-18 (GO:0071351)
regulation of cdk Ser/Thr kinase activity (GO:0000079)
regulation of PRR signaling pathway (GO:0062207)
3'-UTR-mediated mRNA destabilization (GO:0061158)
MSigDB - Reactive Oxygen Species Pathway
MSigDB - TNF-alpha Signaling via NF-kB
Fold enrichment of PA+PB1 on TES-1kb regions
2x (n=5441)
4x (n=3076)
10x (n=1321)
32x (n=324)
p < 0.05

Figure 5 : Enrichment of FluPol subunits on chromatin downstream to the 3'ends of genes. A) RNA bound to the chromatin of infected A549 cells and immunoprecipitated with indicated antibodies were analysed by deep-sequencing. For each selected genes the indicated 10-kb windows from the TES were quantified for their DoGs levels evaluated in Log2(cpm). Data are displayed in box plot as in **Fig 2B**. The right panel is the same quantification for all pooled windows. **B)** Counts of DoG regions enriched with V5-PA and their differential transcription level upon infection. These regions were considered enriched by PA (>IgG, in brown and black) when the average of the ChIP V5-PA signals minus the deviation is greater than twice the level of the control IgG. The regions were considered depleted in PA (<IgG, orange, violet, white) when the PA signal plus the deviation was weaker than twice the IgG signal. DoG levels were considered modified for a p-value<0.05 (paired t test, two tailed) and a fold change >1.5. **C**) Fold change of enrichment for categories of indicated terminal region relative to TES. These were the averages differences between log2 cpm from infected and non-infected cells. Statistical analysis has been carried out the log2 fold change using Student t-test (two-tailed) to compare DoGs regions from unexpressed genes (light blue) and genes expressed more than mediane (blue). The dashed red line indicates the absence of change. The p-values were indicated above the box plots. The dashed line indicate the limit p-value=0.05. In each case, the number (n) of tested TSS were indicated. PA+PB1 analysis were conducted by combining their ChIP-seq signals. **D**) Pathway analysis of genes on which PA+PB1 were enriched 2, 4, 10 and 32 times on 1 kb windows downstream to the TES between infected and mock cells. P-values of Gene Ontology (GO) biological process 2021 and MSigDB were evaluated through Enrichr, and the threshold of p<0.05 is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category.

Transcriptional changes during viral infection confound the analysis of differential gene expression.

It was known the IAV-induced terminal defects allowed the production of ''downstream-of-gene'' transcripts (DoGs) (Bauer et al., 2018; Zhao et al., 2018). These reads beyond the poly-A termination signal are very generally transient RNA species, and are too rare to be observed at high levels by RNA-seq, but not with data from IAV-infected cells. The DoGs and the subsequent pervasive transcription become so high that they negatively affect the power of mRNA quantification by RNA-seq. This lies on the modification the distribution of reads among the transcriptome which significantly affects the quantification of gene expression (Robinson et al., 2010; Tarazona et al., 2011). The change in transcript distribution in the chromatin fraction partially defeated the assessment of differential gene expression by algorithms, such as DeSeq2, validated to measure mRNA expression changes. The method we proposed using reference genes assumed to be constant, in our case here based on the lifetime of mRNAs while the biological triggering phenomenon (viral infection) was in a much shorter time frame, demonstrated that there is an assessment uncertainty with DeSeq2 for 25% of genes predicted to be differentially expressed (**Fig 1C**). Our results highlighted the importance of the choice of normalization methods in the case of biological phenomena, such as here the viral infection by IAV, which disrupt the distribution of RNAs in the genome, in addition to the expression level of genes. Chromatin RNA enrichment allowed the observation of many transcriptional changes, quite in agreement with those obtained by more complicated techniques such as NET-Seq (Bauer et al., 2018). Part of these differences could lie on the RNAPII passage through silent genes, for which the promoters are inactive (**Fig. 1D, 1F, S1D**).

The distribution of FluPol chromatin association indicates that the site of Cap snatching is not necessarily the TSS.

ChIP experiments with several antibodies directed against FluPol subunits allowed us to show that FluPol localizes **5C**). To our knowledge this is the first time that FluPol has been localized on a large scale across the genome.

Although we noted a certain inefficiency in ChIP of antibodies directed against the PA, PB1 subunits and even worse with the anti-PB2 antibody, statistical processing of the enrichment data allowed us to show a specific targeting effect of FluPol on transcriptionally active versus silent genes. These results are also confirmed by the detection of cellular RNAs associated with ChIP by an anti-V5-tag antibody recognizing a virus modified to express a V5-tagged version of PA (**Fig 5A B**). Interactions between FluPol and the Ser5 phosphorylated form of RNAPII (Engelhardt et al., 2005; Martínez-Alonso et al., 2016; Lukarska et al., 2017), as well as interaction with the chromatin reader CHD1 that recognizes the promoter-specific H3K4me3 marks (Marcos-Villar et al., 2016) would suggest that the process of cap-snatching must occur when nascent RNAs exit from RNAPII paused at promoters (Walker and Fodor, 2019). We have no good reason to question this hypothetical model. However, studying the localization of FluPol on chromatin greatly expands the windows of opportunity where this reaction might occur.

PA-associated RNAs are more often found up-regulated and more significantly involved in the cellular response when they were exonic RNAs than when considering RNAs spanning the entire gene body (**Fig 4DE**). This suggests that spliced RNAs would be associated with PA, in the same time with DoGs at the 3'-end of genes. We propose a model where FluPol is recruited to TSSs but would remain attached to RNAPII during elongation. Due to the coupling between transcription and RNA splicing, maturing RNAs arrive at the 3' end with RNAPII and FluPol. At the TES level, termination inhibition causes transcription to continue after the poly-adenylation site of the genes where FluPol still remains associated with the elongating complex carrying spliced RNAs (**Figure 6**).

The association of FluPol with RNAPII in elongation could be part of the viral strategy of escape from cellular defenses.

What might be the advantage of FluPol remaining attached to the transcriptional complex throughout gene

transcription? On the one hand, FluPol recruited to the TSS preferentially on genes involved in antiviral defense (**Fig 3C**), could also participate in termination inhibition of these same genes (**Fig 5D**). This could be an alternative pathway to the cap-snatching process by which FluPol would inhibit the expression of defense genes. On the other hand, by remaining associated with the elongating transcriptional complex, FluPol may ensure that its own transcripts are in close proximity to the factors necessary for the splicing of genomic segments 7 and 8 (which are partially spliced to generate M2 and NS2 viral proteins, respectively), or for nuclear export (Ferhadian et al., 2018). The increase in DoGs found very often associated with PA also suggests that FluPol may also play a role in the stability of these usually very unstable RNAs.

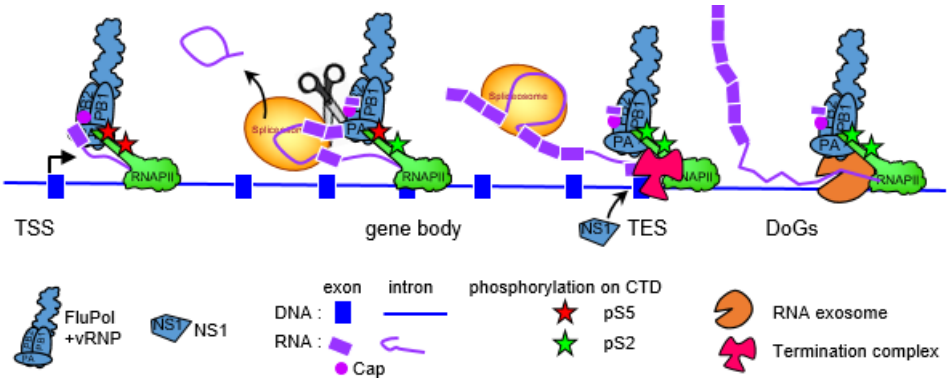


Figure 6 Working model on the functional association of FluPol with RNAPII. The FluPol is reccruited first at the level of TSS and PB2 can bind the cap of the nascent RNA. During elongation the FluPol may stay linked to the RNAPII either through its interaction with the S5p which remain present all along the elongation process or either the link by the nascent transcript. The splicing of the nascent transcript occur during the elongation, producing intermediate of spliced forms. At the 3'end of the gene of the termination process is inhibited by NS1. At this position the FluPol could participate to the termination inhibition and also to the inhibition of the RNA exosome promoting the stabilization of DoGs. The cutting reaction of the nascent RNA by the PA subunit (scissors) could occur at anytime of these steps. The nascent RNA remains linked to the RNAPII in elongation until the 3'end region when the RNAPII continues after the TES because of the termination inhibition. This proposed model could explain why the FluPol is found enriched to DoGs and exonic RNAs on the chromatin.

Recently, we have shown that RNA DoGs in intergenic regions are enriched in repeat regions that promote cellular senescence via the PRR pathway (Mullani et al., 2021). This increase in DoGs is caused by a decrease in the expression of certain exosome RNA subunits. Because the intergenic regions crossed by DoGs are rich in repeated sequences (SINE, LINE, LTR), this pervasive transcription can result in the formation of double-stranded RNAs recognized by intracellular PRRs (Chow et al., 2018). In this regard, it is interesting to recall that FluPol has been shown to be associated with the nuclear exosome RNA which plays a positive role in viral multiplication (Rialdi et al., 2017). Loss of function of the nuclear RNAexosome makes viral multiplication less efficient possibly by decreasing the recruitment of FluPol to the TSS. It can be noticed in the results of this study that depletion of the RNA exosome, in addition to the decrease in FluPol at the TSS (5'end), instead caused an increase in FluPol in the gene terminal region (3' region). This confirms our data on the presence of FluPol on termination sites of gene. It is therefore conceivable that FluPol may have an inhibitory action on the activity of the exosome RNA which therefore promotes the increased presence of of DoGs (and intergenic RNAs) usually targeted by rapid degradation. The robust increase of DoGs during viral infection combined with the transcriptional repression of genes involved in the PRR pathway could allow the saturation of the double-stranded RNA detection system and thus leave the way clear for viral processes, such as the formation of cellular:viral RNA hybrids, essential for the viral transcription.

Material and Methods

H1N1 virus construction and preparation

Amplification of WSN H1N1 // Construction of V5-tag-PA virus WSN H1N1

Antibodies :

Antibodies were purchased from Diagenode for H3K27ac (C15410196) ; H3K4me1 (C15410037) from Abcam for RNAPII (8WG16) , RNAPII pS5 (ab5131) for Western blot, RNAPII pS5 (CTD 4H8) and for ChIP, histone H3 (Ab1791), from ThermoFisher for PA (PA5-31315), PB1 (PA5-34914) and PB2 (PA5-32220); and from Bethyl for anti-V5 epitope (A190-120A).

Cell culture, H1N1 infection

A549 cells (CCL-185) were purchased from ATCC. Cells were maintained in Dulbecco's modified Eagle's medium (Gibco) supplemented with 7% (v/v) fetal bovine serum (Thermo Scientific) and 100 U/ml penicillin-streptomycin (Gibco). A549 pellet of cells, infected (MOI=5, h.p.i=6h) with V5-tag-PA (virusV5) or normal virus WSN H1N1. Cells have been not fixed or fixed by 0.5%, 1%, 2% formaldehyde in PBS solution during 10 min at room temperature prior to harvest. Cells were pelleted and frozen at -80°C prior performing the extracts.

Chromatin immunoprecipitation (ChIP and RNA-ChIP) and RNA-chromatin extraction

ChIP assays were performed as previously described (Ameyar-Zazoua et al., 2012; Mauger et al., 2015). Cells were crosslinked during 10 min at room temperature in PBS with 1% (v/v) formaldehyde, followed by a 5 min wash in PBS containing 125 mM Glycine before harvest.

Resuspended cells (10-20E6 in 1.5 mL µtube) were incubated on ice 5 min in 1 mL of chilled buffer A (0.25% TRITON, 10 mM TRIS pH8, 10 mM EDTA, 0.5 mM EGTA), 30 min in 1 mL of buffer B (250 mM NaCl, 50 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA), and then resuspended in 100 to 200 µL of buffer C (1% SDS, 10 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA) at room temperature. All buffers were extemporately supplemented by 0.5 U/µL RNAsin (Promega), COmplete protease and PhosSTOP phosphatase inhibitors (Roche). Cell suspensions in 0.6 mL µtube were sonicated in water bath at 4°C during 12 min (15 sec. ON, 15 sec. OFF) with a Bioruptor apparatus (Diagenode) setted on high power and then clarified by 10 min centrifigation at 10 000 rpm, 4°C. Shearing of the DNA was checked after reversing the crosslinking on agarose gel electrophoresis to be around 300-500 bp. Sheared chromatin was quantified by optical density (260 nm) and diluted 10-fold in IP buffer to final concentrations: 1% TRITON, 0.1% NaDeoxycholate, 0.1% SDS, 150 mM NaCl, 10 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA, 1 U/µL RNasin, 1X protease and pshotase inhibitors). For ChIP-PCR assays 30 µg of chromatin was incubated with 2 µg of antibodies. For ChIP-seq or RNA-ChIP-seq assays, 200 µg of chromatin was incubated with 10 µg of antibodies at 4°C during 16 h. 20 µL of saturated magnetic beads coupled to protein G (Dynabeads) per µg of antibody were used to recover the immuno-complexes. After 2 h of incubation the bound complexes were washed extensively 10 min at room temperature on a wheel in the following wash buffers: WBI (1% TRITON, 0.1% NaDOC, 150 mM NaCl, 10 mM TRIS pH8), WBII (1% TRITON, 1% NaDOC, 150 mM KCl, 10 mM TRIS pH8), WBIII (0.5% TRITON, 0.1% NaDOC, 250 mM NaCl, 10 mM TRIS pH8), WBIV (0.5% Igepal CA630 (Sigma), 0.5% NaDOC, 250 mM LiCl, 10 mM TRIS pH8, 1 mM EDTA), WBV (0.1% Igepal, 150 mM NaCl, 20 mM TRIS pH8, 1 mM EDTA), WBVI (0.001% Igepal, 10 mM TRIS pH8). 20-100 µg of sheared chromatin used as input and ChIP beads were then boiled 10 min in 100 µL H2O containing 10% (V/W) chelex resin (BioRad), followed by Proteinase K (0.2 mg/mL)-digestion for 30 min at 55°C, and then finally incubated 10 min at 100°C. 0.5 µL was used for quantitative real-time PCR assays. For RNA-chromatin extraction, supernatant were purified by acidic phenol/chloroform method and ethanol-precipitation with 5 µg of Glycoblue (Ambion) as carrier. After resuspension these input, as well as the eluate from RNA-ChIP were treated by 5 U of TURBO-DNase (Ambion) 1 h a 37°C prior purification by acidic phenol/chloroform method and ethanol-precipitation with 5µg of Glycoblue (Ambion) as carrier. These samples were then checked for the absence of DNA contamination by qPCR with 1 µL of eluate and primers targeting genomic repetitive elements. Another round of DNAse digestion and purification was performed in case of qPCR signal < 35 Ct.

Library preparation and high throughput sequencing

1 to 5 nanograms of immunoprecipitated DNA were used to build DNA libraries. Libraries were prepared following the instruction from NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB #E7645L) and indexes (NEB #7335S; #E7500S; #E7710S). No sized selections were applied. The quality and size of the libraries were assessed using the 2100 Bioanalyzer (Agilent). The amount of the DNA used for library preparation and the amount of DNA libraries were measured using Qubit (Invitrogen). Indexed libraries were pooled according to the Illumina calculator (https://support.illumina.com/help/pooling-calculator/pooling-calculator.htm) and deep sequenced by the Illumina platform provided by Novogene (HiSeq-SE50). RNA from IgG ChIP triplicate were pooled to ensure sufficient material equivalent to individual V5-PA ChIP. Libraries from ChIPed RNA and input chromatin-RNA were performed by the GenomEast platform, a member of the 'France Genomique' consortium (ANR-10-INBS-0009) (IGBMC, Illkirch, France). RNA-Seq libraries were generated from 100 to 300 ng of RNA using TruSeq Stranded Total RNA Library Prep Gold kit and TruSeq RNA Single Indexes kits A and B (Illumina, San Diego, CA), according to manufacturer's instructions. Briefly, ribosomal RNA was removed by Ribo-Zero beads. After chemical fragmentation depleted RNA were reverse-transcribed with random primers. After second strand synthesis by DNA Polymerase I and RNase H and blunting, the cDNA were ligated to adapters and amplified by PCR for 12 cycles. Residual primers were removed by AMPure XP beads and the libraries were sequenced on Illumina Hiseq 4000 sequencer as Paired-End 100 base reads. Image analysis and base calling were performed using RTA 2.7.7 and bcl2fastq 2.17.1.14.

Data availability

ChIP-seq, RNA-ChIP-seq, RNA-seq data have been deposited in the NCBI Gene Expression Omnibus database under GEO accession number GSExxxxxx.

Protein extraction for immunoprecipitation.

Pellet of frozen cells corresponding to 10E6 cells were thawed in ice and resuspended in 500 µL of Buffer A with protease inhibitors (Roche). 150 µL of suspension were supplemented with 180 µL of buffer B and sonicated 5 min at 4°C by BioRuptor (high power, 15 sec. ON / 15 sec. OFF). Clarified supernatants were quantified by DO (260 nm). 20 µg of protein extract were incubated with 1 µg of indicated antibodies during 16 h at 4°C on a wheel and then 2 h with 30 µL of anti-rabbit Dynabeads. Beads were washed 5 times in wash buffer WBV, resuspended

in laemmli buffer with 100 mM DTT and boiled 20 min and supernatant was analyzed by western blot.

Western blot

Denaturated proteins were resolved by SDS–PAGE (4–12% Criterion XT Bis-Tris Protein Gel; Bio-Rad), and transferred to nitrocellulose membrane (Bio-Rad). Blocked membrane were incubated with 1/1000 diluted first antibody as indicated in figures in PBS with 0.1% Tween-20 and 5% (V/W) non-fat milk, washed 10 min 3 times in PBS-tween(0.1%), incubated 1 h at room temperature with 1/3000 anti-Mouse StarBright blue700 (BioRad) or 1/2000True blot HRP anti-rabbit antibodies. After 4x10 min washes in PBS-tween(0.1%), the membranes were revealed revealed by chemiluminescence (for HRP), and quantified using Chemidoc MP imaging system (Bio-Rad).

Real-time quantitative PCR:

1 µL of ChIP eluate was used for quantitative real-time PCR (qPCR) in 10 µl reactions with Brillant III Ultra Fast SYBR-Green Mix (Agilent) using a Stratagene MX3005p system. The analysis of qPCR was performed using the MxPro software. The sequences of primers used for PCR are in **SupTab. S3**. Statiscal analysis and graphs were produced with Microsoft Excel.

Bioinformatic analysis

The human genome of reference considered for this study was hg19 homo sapiens primary assembly from Ensembl. After reads alignment, the SAM files were then converted to BAM files and sorted by coordinate and indexed using samtools (v1.7) (Danecek et al., 2021). For ChIP-seq data from this study of from those in MDM cells (GSE103477) (Heinz et al. 2018), reads were mapped to hg 19 using bowtie2 (v2.3.4) (Langmead and Salzberg, 2012) (parameters: -N 0 -k 1 –very-sensitive-local). We then selected reads with a MAPQ equal or higher than 30 corresponding to uniquely mapped reads for further analysis. After indexing MACS2 (v.2.1.1) (Zhang et al., 2008) was used to call peaks (parameters: -p 0.05). For RNA-chromatin and RNA-ChIP-seq data, mapping was carried out with STAR (v2.6.0b) (Dobin et al., 2013) (parameters: --outFilterMismatchNmax 1 --outSAMmultNmax 1 --outMultimapperOrder Random --outFilterMultimapNmax 30). Data observations were performed with the Integrative Genomics Viewer software (Robinson et al., 2011) using the CPM normalized BigWig files produced from the BAM files by Deeptools suite (parameter: --normalizeUsing CPM) (Ramírez et al., 2016)

Evaluation of gene expression from chromatin-RNA (our study) or from total RNA extract of A549 cells mock or infected by normal H1N1 at 12 hpi (GSE103604) (Zhao et al. 2018) were performed by using featureCounts (v1.28.1) (Liao et al., 2014) from the Rsubreads package (parameter: -s 2). These read counts were analysed using the DESeq2 (Love et al., 2014) package in order to test for the differential gene expression (DGE). The normalization, the dispersion estimation and the statistical analysis were performed with DESeq2 using the default parameters. Raw *P*-values were adjusted for multiple testing according to the Benjamini and Hochberg (BH) procedure and genes with an adjusted *P*-value lower than 0.05 and a fold change greater than 50% were considered differentially expressed. The raw count from featureCount were also used to produce normalized gene levels using reference genes as follow : counts covering mRNAs were converted in rpkm by considering the library total count of each sample and the length of each gene as defined in Ensembl. The average level of genes with half-life greater than 15 h as defined in (Tani et al., 2012) and that were more expressed than 75% of all expressed genes (corresponding 417 genes with level higher than 300 rpkm for chromatin-RNA, and 281 genes with more than 17 rpkm in total RNA extract from Zhao et al. 2018) was used to calculate a corrective index applied on the gene levels in infected cells. The maximun mRNA level between mock and infected cells were used to calculate the correlation between the level quartiles and the life-span quartile (**SupFig S1B**) and to draw the MA-plot (**SupFig S1C**). Genes with a *P*-value from a paired t test (n=3) lower than 0.05 and a fold change greater than 50% were considered differentially expressed.

Signal evaluation of ChIP-seq, RNA-ChIPseq and RNA-seq on defined windows were performed as followed. Multiples regions based on the Ensembl annotations (hg19, version 87) were generated using the bedtools suite (v2.27.1) (Quinlan and Hall, 2010), 500 bp around the center of the intron that is closest to the center of the gene, 500 bp around the TSS, 500 bp around the TES, from TES to 1 kb downstream gene, from TES to 10 kb downstream gene and 40 kb to 41 kb downstream gene. In the case of the downstream region of genes, only regions retained were those not overlapping and distant by at least 10 kb with with any other gene from the TES to the end of the counting region. For TSS regions, windows retained were those distant by at least 10 kb from both extremities of other genes. Consecutive 10 kb windows were generated after TES of genes with a maximum of 40 kb as long as there was no overlap with downstream genes oriented in the same direction. All previously generated counting tracks were converted to the SAF format and then used as annotation file in featureCounts (v1.6.1) (parameter: -s 2 for RNA-ChIP-seq or -s 0 for ChIP-seq) . Regions with a *P*-value from a paired t test (n=3) lower than 0.05 and a fold change greater than 50% were considered differentially enriched between infected and mock cells. In the case of V5-PA RNA-ChIP the relative enrichment in infected cells was evaluated in comparison with non-immun IgG RNA-ChIP as explained in the legend of **Fig. 4D**.

Statiscal analysis and graphs were produced with Microsoft Excel. Pathway analysis were performed with Enrichr (Kuleshov et al., 2016).

Acknowledgements

The authors thank our colleagues C. Rachez and S. Thouroude for helpful discussions. A special thanks also to Catherine Bodin for technical assistance and to Edith Ollivier for administrative assistance.

Fundings

J.Y. is part of the Pasteur - Paris University (PPU) International PhD Program; European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie [665807]. Ministry of Research (MENRT; to J.B.); Centre National de Recherche Scientifique (CNRS; to E.B. and C.M.); Agence Nationale de la Recherche [ANR-11-BSV8-0013] ANR FluChromatin; REVIVE-Investissement d'Avenir (to J.Y., M.C., C.M.). Funding for open access charge: CNRS recurrent funding;
Conflict of interest statement. None declared.

References

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.-C., et al. (2012). Argonaute proteins couple chromatin silencing to alternative splicing. Nat. Struct. Mol. Biol. *19*, 998–1004.

Amorim, M.-J., Read, E.K., Dalton, R.M., Medcalf, L., and Digard, P. (2007). Nuclear export of influenza A virus mRNAs requires ongoing RNA polymerase II activity. Traffic Cph. Den. *8*, 1–11.

Bauer, D.L.V., Tellier, M., Martínez-Alonso, M., Nojima, T., Proudfoot, N.J., Murphy, S., and Fodor, E. (2018). Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. Cell Rep. *23*, 2119-2129.e3.

Boehm, A.K., Saunders, A., Werner, J., and Lis, J.T. (2003). Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. Mol. Cell. Biol. *23*, 7628–7637.

Chow, K.T., Gale, M., and Loo, Y.-M. (2018). RIG-I and Other RNA Sensors in Antiviral Immunity. Annu. Rev. Immunol. *36*, 667–694.

Chung, M., Cho, S.Y., and Lee, Y.S. (2018). Construction of a Transcriptome-Driven Network at the Early Stage of Infection with Influenza A H1N1 in Human Lung Alveolar Epithelial Cells. Biomol. Ther. *26*, 290–297.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., et al. (2016). A survey of best practices for RNA-seq data analysis. Genome Biol. *17*, 13.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008.

Dias, A., Bouvier, D., Crépin, T., McCarthy, A.A., Hart, D.J., Baudin, F., Cusack, S., and Ruigrok, R.W.H. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature *458*, 914–918.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinforma. Oxf. Engl. *29*, 15–21.

ENCODE Project Consortium, Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Engelhardt, O.G., Smith, M., and Fodor, E. (2005). Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. J. Virol. *79*, 5812–5818.

Ferhadian, D., Contrant, M., Printz-Schweigert, A., Smyth, R.P., Paillart, J.-C., and Marquet, R. (2018). Structural and Functional Motifs in Influenza Virus RNAs. Front. Microbiol. *9*, 559.

Heinz, S., Texari, L., Hayes, M.G.B., Urbanowski, M., Chang, M.W., Givarkes, N., Rialdi, A., White, K.M., Albrecht, R.A., Pache, L., et al. (2018). Transcription Elongation Can Affect Genome 3D Structure. Cell *174*, 1522-1536.e22.

Ho, C.K., and Shuman, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. Mol. Cell *3*, 405–411.

Kolev, N.G., and Steitz, J.A. (2005). Symplekin and multiple other polyadenylation factors participate in 3′-end maturation of histone mRNAs. Genes Dev. *19*, 2583–2592.

Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452–2460.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90-97.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinforma. Oxf. Engl. *30*, 923–930.

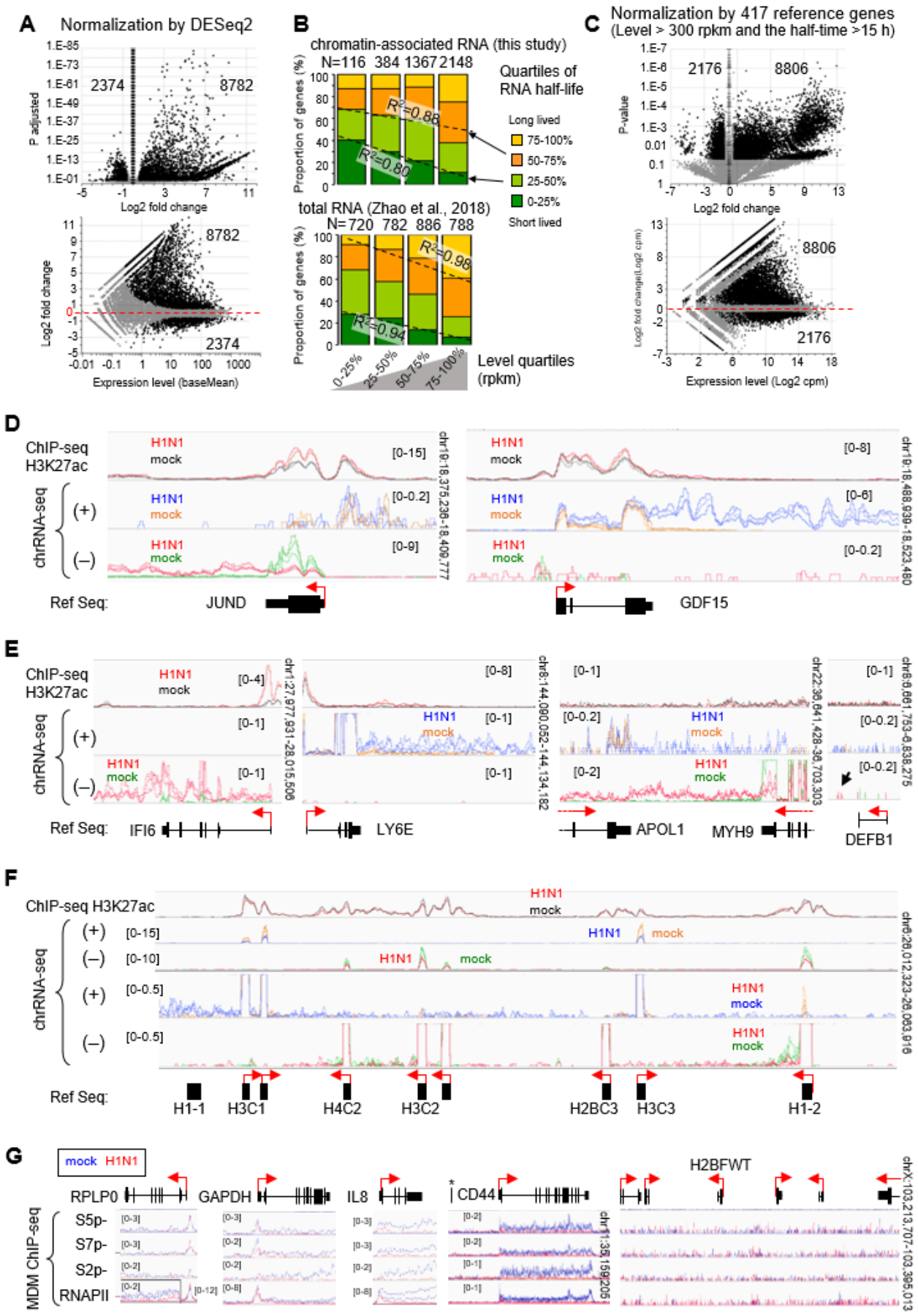
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lukarska, M., Fournier, G., Pflug, A., Resa-Infante, P., Reich, S., Naffakh, N., and Cusack, S. (2017). Structural basis of an essential interaction between influenza polymerase and Pol II CTD. Nature *541*, 117–121.

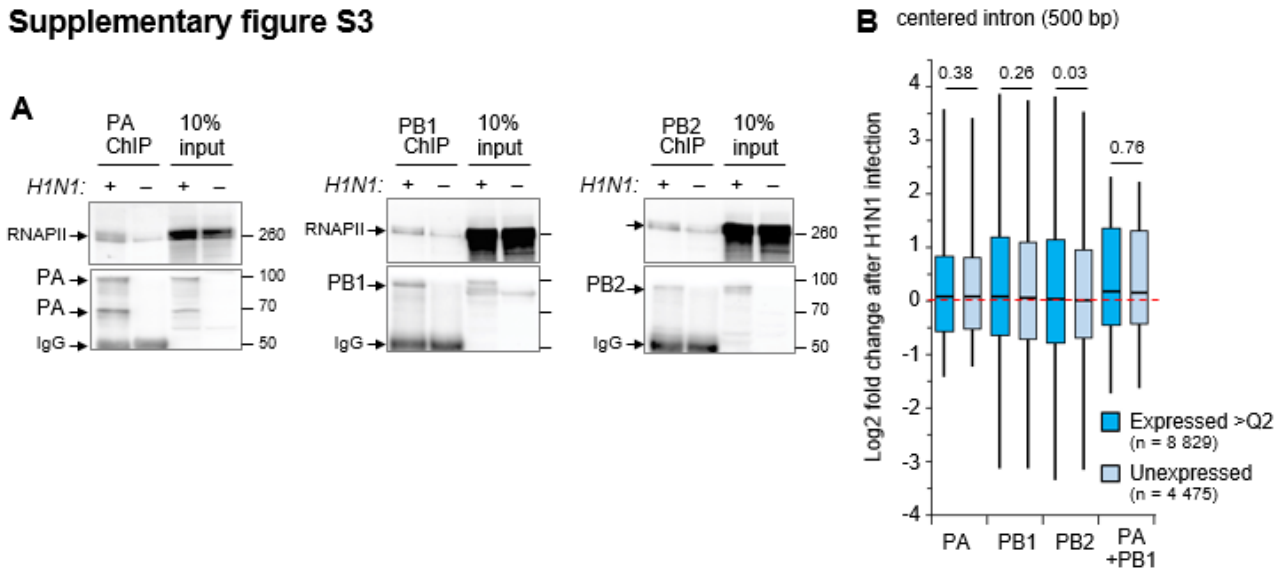
Marcos-Villar, L., Pazo, A., and Nieto, A. (2016). Influenza Virus and Chromatin: Role of the CHD1 Chromatin Remodeler in the Virus Life Cycle. J. Virol. *90*, 3694–3707.

Martínez-Alonso, M., Hengrung, N., and Fodor, E. (2016). RNA-Free and Ribonucleoprotein-Associated

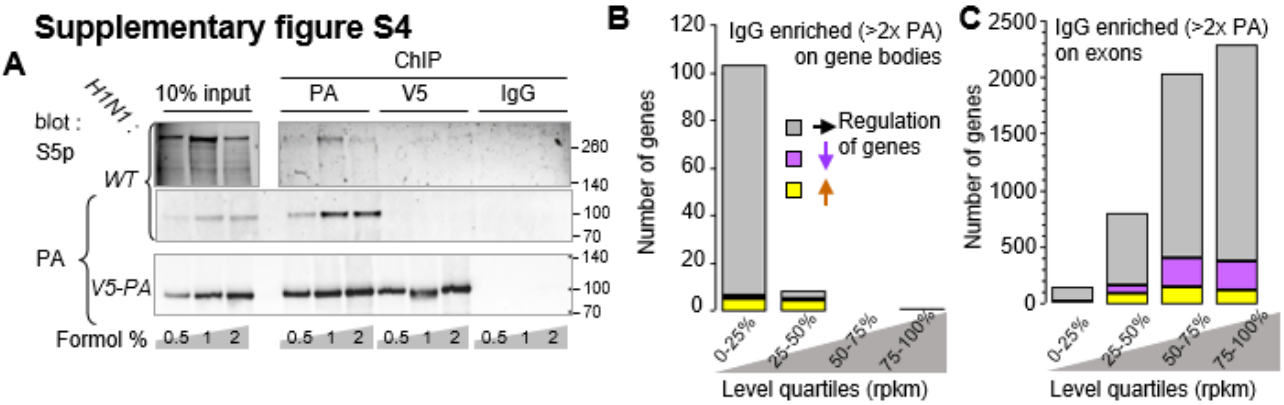
Influenza Virus Polymerases Directly Bind the Serine-5-Phosphorylated Carboxyl-Terminal Domain of Host RNA Polymerase II. J. Virol. *90*, 6014–6021.
Mauger, O., Klinck, R., Chabot, B., Muchardt, C., Allemand, E., and Batsché, E. (2015). Alternative splicing regulates the expression of G9A and SUV39H2 methyltransferases, and dramatically changes SUV39H2 functions. Nucleic Acids Res. *43*, 1869–1882.
Mullani, N., Porozhan, Y., Mangelinck, A., Rachez, C., Costallat, M., Batsché, E., Goodhardt, M., Cenci, G., Mann, C., and Muchardt, C. (2021). Reduced RNA turnover as a driver of cellular senescence. Life Sci. Alliance *4*.
Park, D., Shivram, H., and Iyer, V.R. (2014). Chd1 co-localizes with early transcription elongation factors independently of H3K36 methylation and releases stalled RNA polymerase II at introns. Epigenetics Chromatin *7*, 32.
Plotch, S.J., Bouloy, M., and Krug, R.M. (1979). Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. Proc. Natl. Acad. Sci. U. S. A. *76*, 1618–1622.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma. Oxf. Engl. *26*, 841–842.
Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160-165.
Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R.W.H., et al. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature *516*, 361–366.
Rialdi, A., Hultquist, J., Jimenez-Morales, D., Peralta, Z., Campisi, L., Fenouil, R., Moshkina, N., Wang, Z.Z., Laffleur, B., Kaake, R.M., et al. (2017). The RNA Exosome Syncs IAV-RNAPII Transcription to Promote Viral Ribogenesis and Infectivity. Cell *169*, 679-692.e14.
Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.
Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.
Sikora, D., Rocheleau, L., Brown, E.G., and Pelchat, M. (2014). Deep sequencing reveals the eight facets of the influenza A/HongKong/1/1968 (H3N2) virus cap-snatching process. Sci. Rep. *4*, 6181.
Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. *22*, 947–956.
Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: A matter of depth. Genome Res. *21*, 2213–2223.
Walker, A.P., and Fodor, E. (2019). Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. Trends Microbiol. *27*, 398–407.
Wang, C., Forst, C.V., Chou, T., Geber, A., Wang, M., Hamou, W., Smith, M., Sebra, R., Zhang, B., Zhou, B., et al. (2020). Cell-to-Cell Variation in Defective Virus Expression and Effects on Host Responses during Influenza Virus Infection. MBio *11*.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.
Zhao, N., Sebastiano, V., Moshkina, N., Mena, N., Hultquist, J., Jimenez-Morales, D., Ma, Y., Rialdi, A., Albrecht, R., Fenouil, R., et al. (2018). Influenza virus infection causes global RNAPII termination defects. Nat. Struct. Mol. Biol. *25*, 885–893.
Zhou, B., Li, J., Liang, X., Yang, Z., and Jiang, Z. (2017). Transcriptome profiling of influenza A virus-infected lung epithelial (A549) cells with lariciresinol-4-β-D-glucopyranoside treatment. PLOS ONE *12*, e0173058.
Zhu, J., Adli, M., Zou, J.Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., et al. (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. Cell *152*, 642–654.


A
Normalization by DESeq2
P adjusted
2374
8782
Log2 fold change
Expression level (baseMean)
B
chromatin-associated RNA (this study)
N=116 384 1367 2148
Proportion of genes (%)
Quartiles of RNA half-life
Long lived
75-100%
50-75%
25-50%
0-25%
Short lived
total RNA (Zhao et al., 2018)
N=720 782 886 788
Level quartiles (rpkm)
C
Normalization by 417 reference genes
(Level > 300 rpkm and the half-time >15 h)
P-value
2176
8806
Log2 fold change
Expression level (Log2 cpm)
D
ChIP-seq H3K27ac
chrRNA-seq
H1N1
mock
Ref Seq:
JUND
GDF15
E
IFI6
LY6E
APOL1
MYH9
DEFB1
F
ChIP-seq H3K27ac
H1-1
H3C1
H4C2
H3C2
H2BC3
H3C3
H1-2
G
mock H1N1
RPLP0
GAPDH
IL8
CD44
H2BFWT
MDM ChIP-seq
S5p-
S7p-
S2p-
RNAPII

Supplementary figure S1 : Differential gene expression at 6 h in H1N1-infected A549 cells versus non-infected cells
.**A**) Differential RNA levels between H1N1-infected cells versus non-infected cells using DESeq2 package. Variations have been considered significant (black dot) for a mean fold change (FC)>1.5 and the adjusted p-value <0.05. The number of black dots were indicated. **B**) The top graph indicates the proportion of genes ranked by their level of RNA associated to the chromatin in A549 cells and their half-life as evaluated in HeLa cells (Tani et al., 2012). Among the 4678 genes investigated for their RNA life span, the number of genes expressed in each quartile in A549 cells are indicated on the top of graph. Pearson's coefficient of determination (R^2) were calculated between minimal RNA level of each quartile (rpkm) and the percentage of genes with the half-time indicated by the dotted lines and arrow. The bottom graph is the same analysis performed with the mean of RNA levels from total RNA extract in uninfected and 12-hpi infected A549 cells (Zhao et al. 2018). **C**) Differential levels of RNA associated to the chromatin using the normalization based on the average variation of 424 reference genes expressing higher level (> 10 log2 cpm) of transcripts with long half-life (> 15 h). Variations have been considered significant (black dot) for a mean fold change is >1.5 and the p-value <0.05 calculated by a paired T-test (two-tailed) on the log2(cpm). **D, E,F**) Examples illustrating the variation in the distribution of reads in H1N1 infected cells from IGV visualization of indicated loci. The RNA-seq independent triplicate were overlayed in the same color as indicated, for each strand orientation. The top track shows the enrichment H3K27ac on the chromatin of infected or mock cells. The scale range of the tracks were indicated in brackets. **G**) Distribution of indicated phosphorylated (on Serine5, 7 and 2) and whole RNAPII on genes analysed **Fig 1E**. ChIP-seq assays with antibodies targeting the indicated RNAPII were conducted in infected (red line) or uninfected MDM cells (blue line) were used from GSE103477 datasets (Heinz et al., 2018). The range scale are indicated in brackets



Supplementary figure S3 : immunoprecipitation efficiencies of Antibodies directed against the Flu Pol subunits. A) Infected (+) or uninfected (-) cells were extracted and subjected to immunoprecipitation with the indicated antibodies used in ChIP assay. Eluates and 10% of the extracts (input) were resolved by western blot. The lower part of the membranes (50 to 120 kDa) was revealed with the corresponding antibodies and the upper part (120 kDa-top) with RNAPII antibodies. **B**) The same analysis, as in **fig. 3B**, has been conducted for the averages of intronic coverage. For each gene the more central intron has been used to defined a centered 500 bp windows of counting. Only windows separated by 10 kb from TSS were considered.



Supplementary figure S4 : immunoprecipitation efficiencies of Antibodies directed against the Flu Pol subunits. A) A549 cells infected cells by WT IAV or V5-PA engineered IAV were crosslinked by indicated % of formaldehyde. Chromatin fraction were subjected to immunoprecipitation with PA, V5 antibodies or non immun IgG. Eluates and 10% of extracts were resolved by western blot. The lower part of the membranes was revealed with the PA antibodies and the upper part with phospho-RNAPII antibodies. **B, C**) The same analysis as described for **figure 4D, E** was carried out but this time evaluating the signal enrichment for IgG control compared to V5-PA