Recurrent chromosomal translocations in sarcomas create a mega-complex that mislocalizes NuA4/TIP60 to Polycomb target loci

Deepthi Sudarshan[1], Nikita Avvakumov[1]*, Marie-Eve Lalonde[1]*, Nader Alerasool[2]^, Karine Jacquet[1]^, Amel Mameri[1]^, Justine Rousseau[3]^, Jean-Philippe Lambert[1,4], Eric Paquet[1], Samarth Thonta Setty[5], Jeremy Loehr[1], Anne-Claude Gingras[4], Benoit Coulombe[3], Mikko Taipale[2], Yannick Doyon[1]# and Jacques Côté[1,6]#

1- St-Patrick Research Group in Basic Oncology, Centre Hospitalier Universitaire de Québec-Université Laval Research Center, Laval University Cancer Research Center, Quebec City, QC, Canada; 2- Donnelly Centre for Cellular and Biomolecular Research, Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; 3- Institut de Recherches Cliniques de Montréal and Department of Biochemistry and Molecular Medicine, Université de Montréal, Montreal QC, Canada; 4- Centre for Systems Biology, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada; Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; 5- Computational Biology Laboratory, CHU de Québec - Université Laval Research Center, Québec City, QC, Canada.

* Equal contributions

^ Equal contributions

Corresponding authors: yannick.doyon@crchudequebec.ulaval.ca, jacques.cote@crhdq.ulaval.ca

6. Lead contact

Running Title: Molecular fusion of NuA4 and PRC2 complexes in sarcomas

ABSTRACT

Chromosomal translocations frequently promote carcinogenesis by producing gain-of-function fusion proteins. Recent studies have identified highly recurrent chromosomal translocations in patients with Endometrial Stromal Sarcomas (ESS) and Ossifying FibroMyxoid Tumors (OFMT) leading to an in-frame fusion of PHF1 (PCL1) to six different subunits of the NuA4/TIP60 complex. While NuA4/TIP60 is a co-activator that acetylates chromatin and loads the H2A.Z histone variant, PHF1 is part of the Polycomb repressive complex 2 (PRC2) linked to transcriptional repression of key developmental genes through methylation of histone H3 on lysine 27. In this study, we characterize the fusion protein produced by the *EPC1-PHF1* translocation. The chimeric protein assembles a mega-complex harboring both NuA4/TIP60 and PRC2 activities and leads to mislocalization of chromatin marks in the genome. These are linked to aberrant gene expression, in particular over an entire topologically-associated domain including part of the *HOXD* cluster. Furthermore, we show that JAZF1, implicated with PRC2 components in the most frequent translocations in ESS, is a potent transcription activator that physically associates with NuA4/TIP60. Altogether, these results indicate that most chromosomal translocations linked to these sarcomas employ the same molecular oncogenic mechanism through a physical merge of NuA4/TIP60 and PRC2 complexes leading to mislocalization of histone marks and aberrant polycomb target gene expression.

Highlights:

- *Recurrent oncogenic chromosomal translocations fuse Polycomb-like-1 (PHF1) with different subunits of the NuA4/TIP60 complex*
- *Translocation produces a mega-complex containing NuA4/TIP60 and PRC2 activities*
- *Translocation leads to mistargeting of histone marks throughout the genome and changes in gene expression*
- *Loss of H3K27me3 and gain of H4ac on HOXD cluster are restricted to a topologically-associated domain (TAD)*
- *A highly recurrent JAZF1-SUZ12 translocation uses the same mechanism merging NuA4/TIP60 with PRC2*

INTRODUCTION

ATP-dependent remodelers and histone modifiers are key regulators of the structure and function of chromatin, essential for genome expression and stability, cell proliferation, development and response to environmental cues. Post-translational modifications of specific residues on histones are also part of the epigenetic mechanisms ensuring gene expression memory during cell divisions (Zentner and Henikoff, 2013). Histone modifying enzymes are often part of large multisubunit protein complexes. They are highly conserved and composed of various combinations of subunits like readers, writers, and erasers of histone marks, histone chaperones, and chromatin remodelers. The combination of different modules ensures specific localization, histone target selection, enables epigenetic crosstalk and context-specific activity. The strictly coordinated activities of chromatin modifying complexes ensure the proper functioning of the cell (Lalonde et al., 2014). Disruption of chromatin marks and their regulators can lead to various pathologies including cancer (Shen and Laird, 2013).

Recurrent chromosomal translocations producing oncogenic fusion proteins are common in many cancers and particularly prevalent in hematopoietic malignancies and sarcomas. These fusions frequently involve chromatin and transcription regulators and, in many cases, are thought to be the primary drivers of cancer. In recent years, disruption of chromatin dynamics has emerged as a consistent oncogenic mechanism used by these fusion proteins (reviewed in (Brien et al., 2019)). Sarcomas are rare mesenchymal tissue cancers with distinct molecular profiles. 1/3rd of all sarcomas harbor chromosomal translocations with otherwise normal karyotypes, showing robust clustering of gene expression profiles compared to normal tissues. These chromosomal translocations are the driver oncogenes and the common

oncogenic mechanism employed is either transcriptional deregulation (EWS-FLI in Ewing's Sarcoma) or aberrant signaling (ALK fusions in Inflammatory Myofibroblastic Tumor) (Taylor et al., 2011). Of particular interest are the fusions involving proteins associated with chromatin regulators that impact gene expression programs. Recurrent translocations are found in Endometrial Stromal Sarcomas (ESS) and Ossifying FibroMyxoid Tumors (OFMT) that potentially fuse subunits of distinct chromatin modifying complexes with opposite functions in gene regulation, namely the NuA4/TIP60 histone acetyltransferase (HAT) complex and Polycomb Repressive complexes (Figure 1A).

Among the recurring chromosomal translocations that characterize endometrial stromal sarcoma, genes for five different subunits of the NuA4/TIP60 complex (EPC1/2, MBTD1, MEAF6, BRD8) are repeatedly found fused to genes for different PRC2 components (PHF1/SUZ12/EZHIP) (Ferreira et al., 2018; Hoang et al., 2018). Interestingly, other soft tissue sarcomas such as ossifying fibromyxoid tumors also harbor EPC1-PHF1, EP400-PHF1 and MEAF6-PHF1 fusions. OFMTs are rare cancers of uncertain cellular origin, and ~50-85% of cases show PHF1 translocations, with EP400-PHF1 being the most frequent (40%) (Schneider et al., 2016). JAZF1-PHF1 was reported in a case of Cardiac Ossifying sarcoma and fusions of NuA4/TIP60 subunits with the PRC1.1 component BCOR have also been reported in ESS, demonstrating that these types of fusion events represent a critical oncogenic mechanism in mesenchymal cancers (Schoolmeester et al., 2013) (Figure 1A).

NuA4/TIP60 is an evolutionarily conserved and multi-functional MYST-family HAT complex with 17 distinct subunits (Figure 1B)(Avvakumov and Cote, 2007; Doyon and Cote, 2004; Judes et al., 2015; Sheikh and Akhtar, 2019; Steunou et al., 2014; Voss and Thomas,

2018). The catalytic subunit KAT5/Tip60 acetylates histones H4 and H2A, as well as variants H2A.Z and H2A.X in the context of chromatin. In addition, the ATP-dependent remodeler subunit EP400 allows the exchange of canonical histone H2A in chromatin with variant histone H2A.Z (Billon and Cote, 2013; Pradhan et al., 2016). There are also multiple subunits with different reader domains that allow context-dependent site-specific activity of the complex and epigenetic crosstalk (Figure 1B)(Jacquet et al., 2016; Steunou et al., 2014). The 836aa EPC1 protein is a non-catalytic scaffolding subunit of the TIP60 complex. The conserved N-terminal EPcA domain (aa1-280) interacts with KAT5/Tip60, MEAF6, and ING3 thereby forming the human Piccolo NuA4 complex, enabling binding and acetylation of chromatin substrates (Boudreault et al., 2003; Chittuluru et al., 2011; Doyon et al., 2004; Huang and Tan, 2013; Lalonde et al., 2013; Selleck et al., 2005; Xu et al., 2016). The C-terminal part of EPC1 is thought to associate with the rest of the NuA4/TIP60 complex based on data from the yeast NuA4 complex (Auger et al., 2008; Boudreault et al., 2003; Setiaputra et al., 2018; Wang et al., 2018) (Figure 1B). The NuA4/TIP60 complex is an important transcriptional co-activator that can be recruited to gene regulatory elements by several transcription factors. Through acetylation of histones as well as non-histone substrates, it activates a multitude of gene expression programs, including: proliferation, stress response, apoptosis, differentiation and stem cell identity (Judes et al., 2015; Steunou et al., 2014; Voss and Thomas, 2018). The NuA4/TIP60 complex is also a key player in the response to DNA damage, assisting repair pathway choice as well as the repair process itself ((Cheng et al., 2018; Jacquet et al., 2016) and references therein). All these functions are essential for cellular homeostasis, hence NuA4/TIP60 subunits are often mutated or deregulated in cancers (Avvakumov and Cote, 2007; Judes et al., 2015;

Voss and Thomas, 2018) and Tip60/KAT5 itself is a haplo-insufficient tumor suppressor (Gorrini et al., 2007).

Polycomb Group proteins (PcG) are evolutionarily conserved proteins involved in development and transcriptional regulation, originally linked to *HOX* gene repression during Drosophila development (Schuettengruber et al., 2017). They are key regulators of mammalian development, differentiation, and cell fate decisions. PcG proteins assemble into multisubunit complexes, major ones being Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2). These complexes co-operate with each other to create repressive chromatin regions through histone modifications and chromatin organization. While PRC1 catalyzes H2AK119 mono-ubiquitination and chromatin compaction, PRC2 catalyzes H3K27 methylation. The PRC2 complex is composed of the core components EZH2, SUZ12, EED, and RBBP4/7. Through a read-write mechanism, it can deposit H3K27me3 over large chromatin regions such as the *HOX* clusters. Other associated factors function to stabilize the PRC2 complex on chromatin and modulate its activity, defining distinct variant complexes PRC2.1 and PRC2.2. PRC2.1 contains one of the Polycomb-like (PCL) proteins, PHF1, MTF2 or PHF19, as well as either PALI1/2 or EPOP (Figure 1C) (Laugesen et al., 2019; Loubiere et al., 2019). The 567aa PHF1 protein binds unmethylated CpG islands particularly at long linker DNA regions, stabilizes the binding of PRC2.1 complex on chromatin and increases its activity (Cao et al., 2008; Choi et al., 2017; Li et al., 2017; Sarma et al., 2008). PHF1 is also known to bind the H3K36me3 mark through its Tudor domain (Cai et al., 2013; Musselman et al., 2012), restricting the catalytic activity of EZH2 (Musselman et al., 2012) (Figure 1C).

Clinical studies frequently report new rearrangements in sarcomas, and many groups are working on the classification of ESS using recurrent fusions as molecular markers. However, there is currently little understanding about the molecular consequences of recurrent translocations and their primary role in tumorigenesis. To address this, we used biochemical and genomic approaches to study a recurrent translocation that fuses the EPC1 subunit of the NuA4/TIP60 complex to the PHF1 subunit of the PRC2.1 complex (Figure 1D). Low grade ESS, as well as OFMTs, harbor the EPC1-PHF1 fusion protein (Antonescu et al., 2014; Micci et al., 2006). We investigated the molecular impact of the EPC1-PHF1 fusion protein by generating isogenic cell lines, used affinity purification followed by mass spectrometry to identify interactors and analyzed chromatin occupancy of the fusion protein. Importantly, we analyzed the effect of the fusion protein on global chromatin dynamics while correlating differential gene expression.

RESULTS

To construct the fusion *EPC1-PHF1* gene (Figure 1D), we used a portion of the chimeric gene recovered by RT-PCR from total RNA isolated from surgically removed endometrial stromal sarcoma (a kind gift from Francesca Micci's group) (Micci et al., 2006). As necessary controls, we also subcloned full-length PHF1 and EPC1, as well as a portion of EPC1 corresponding to the fragment found in the fusion, referred to as EPC1(1-581). We confirmed the expression of all constructs and then tested whether the EPC1-PHF1 fusion protein can act like an oncogenic driver by performing Colony Formation Assay in HEK293T cells (Figure S1A-B). In contrast to full-

length and truncated EPC1 which strongly inhibit growth, we observed that expression of EPC1-PHF1 leads to a greater number of colonies compared to controls (Figure S1B), supporting the idea that this gene fusion may be a driver event, giving a growth advantage to the expressing cells.

EPC1-PHF1 forms a mega-complex merging NuA4/TIP60 and PRC2.1

To isolate the native complex(es) formed by the EPC1-PHF1 fusion protein, we generated isogenic K562 cell lines by targeted integration of C-terminally TAP-tagged cDNAs into the *AAVS1* safe harbor genomic locus (Dalvai et al., 2015). To circumvent issues related to protein overexpression, we used the moderately active *PGK1* promoter which was shown to achieve expression of NuA4/TIP60 and PRC2 subunits within 2.5-fold of the native levels (EP400, EPC1, MBTD1 and EZH2) (Dalvai et al., 2015). Our previous studies have demonstrated that this system enables purification of stable and stoichiometric protein complexes (Dalvai et al., 2015; Doyon and Cote, 2016; Jacquet et al., 2016). We chose the K562 cell line as it is an ENCODE tier-1 cell line and can grow in suspension culture to high cell densities, ideal for biochemical and genomic studies.

Nuclear extracts from K562 cell lines expressing the fusion protein, individual fusion partners or the empty tag (Mock, C-terminal 3xFlag-2xHA-2A-puromycin tag) were allowed to bind to anti-FLAG resin and the bound material was eluted with 3xFlag peptides. An alternate K562 cell line expressing EPC1-PHF1 with a C-terminal 3xFLAG-2xStrep tag was also used to improve yield and purity by binding the FLAG eluted fraction to Streptactin beads followed by elution with biotin. SDS-PAGE and silver-staining identified the components of the purified

complexes (Figure 2A). Mass spectrometry analysis identified all expected components of NuA4/TIP60 and PRC2 complexes (Figure 2B, S2A). We further confirmed the complex subunits by western blotting with NuA4/TIP60 and PRC2 specific antibodies (Figure 2C). The PHF1 fraction contained the core subunits EZH2, EED and SUZ12 of the PRC2 complex, as well as RBBP4 and the PRC2.1 specific PALI-1/LCOR. The EPC1(1-581) fraction contained all the subunits of NuA4/TIP60 complex except MBDT1, which was expected since it associates through an interaction with EPC1 C-terminus (aa644-672) (Jacquet et al., 2016; Zhang et al., 2020). Strikingly, the EPC1-PHF1 fraction contained subunits of both the TIP60 and PRC2.1 complexes. These results indicate that the EPC1-PHF1 fusion protein efficiently associates with both NuA4/TIP60 and PRC2.1. This was also supported by data in HEK293T cells (Figure S1C).

To determine if the fusion protein's association with the two complexes occur independently of each other or could occur simultaneously, we integrated 3xHA-tagged EPC1-PHF1 at the *AAVS1* locus in K562 cell lines in which TALEN/CRISPR was used to introduce a 3xFLAG tag at endogenous *EZH2* or *EP400* genes, respectively (Dalvai et al., 2015). This allowed us to purify endogenous PRC2 and NuA4/TIP60 complexes and verify if the fusion protein can physically bridge the two complexes. As shown in Figure 2D and 2E, it is clear that expression of the EPC1-PHF1 fusion leads to the association of the PRC2 complex with NuA4/TIP60 and vice versa, demonstrating the formation of a mega-complex.

We then performed *in vitro* histone acetyltransferase (HAT) and histone methyltransferase (HMT) assays with purified fractions on native chromatin as substrate. We observed robust HAT activity towards histone H4 and H2A in both EPC1(1-581) and EPC1-PHF1 complexes, as expected for NuA4/TIP60 (Figure 2F, S2B). Simultaneously, we detected HMT

activity towards histone H3 in PHF1 and EPC1-PHF1 complexes, as expected for PRC2 (Figure 2F, S2C). This demonstrates that the merged assembly maintains the enzymatic activities of both original complexes. Altogether, these results indicate that expression of the EPC1-PHF1 fusion protein in cells leads to the formation of a hybrid TIP60-PRC2 mega-complex that harbours opposite functions in term of chromatin modifications and impact on gene regulation (Figure 2G). It is important to note that we did not detect an effect of expressing the fusion protein or its partners on the bulk level of H3K27me3 (Figure S2D), in contrast to what has been previously proposed for the JAZF1-SUZ12 fusion (Ma et al., 2017).

The EPC1-PHF1 fusion complex is enriched at genomic loci containing bivalent chromatin

To determine the genomic locations where the EPC1-PHF1 mega-complex may act, we performed anti-FLAG chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-seq) using the isogenic cell lines. We analyzed binding regions (significant peaks) of EPC1-PHF1, EPC1(1-581) and PHF1 in comparison to the empty vector cell line (Figure 3A, S3A) and correlated them with chromatin regions containing H3K4me3 (usual NuA4/TIP60 targets), H3K27me3 (PRC2 targets) or a combination of both, i.e. bivalent chromatin. We found that the fusion protein is specifically enriched in regions of bivalent chromatin, which is likely facilitated by the presence of both H3K4me3 and H3K27me3 readers in the mega-complex (ING3 in NuA4/TIP60 and EED in PRC2) (Figure 3B).

Effect of EPC1-PHF1 on the chromatin and transcriptional landscape

Two major possible effects are expected from the presence of the fusion mega-complex at specific genomic loci: 1-appearance of chromatin acetylation where PRC2.1 is normally bound on silenced regions of the genome; 2-appearance of H3K27 methylation where TIP60 is normally bound on expressed regions of the genome. We also speculated that the fusion protein may mislocalize TIP60-mediated H4 acetylation to poised bivalent chromatin regions, pushing them towards transcriptional activation. To test this, we performed H4 Penta-acetyl (RRID:AB_310310) and H3K27me3 (RRID:AB_27932460) ChIP-sequencing in our isogenic cell lines. A first observation was that 28% of the genomic regions bound by EPC1-PHF1 show increased H4 acetylation but only 2.9% have increased H3K27me3, arguing for a dominant effect of Tip60 acetyltransferase vs PRC2 methyltransferase in the mega-complex. Furthermore, to understand the effect of EPC1-PHF1-induced changes in local histone modifications on transcription, we performed gene expression analysis. Upon integrative analysis, we found that EPC1-PHF1-induced changes in H4 acetylation correlate with changes in local gene expression (Figure 3C, S3B). In parallel, significant changes in H3K27me3 anti-correlate with changes in gene expression (Figure 3D, S3C).

Since the appearance of new H4 acetylation signals at EPC1-PHF1-bound regions was predominant over the appearance of new H3K27me3 signals, this suggests that subverting NuA4/TIP60 activity to new genomic loci by the fusion protein was the primary outcome. We found that the *HOX* gene clusters, in particular, demonstrated this effect in a striking fashion (Figures 4, S4 and S5). Representative ChIP-seq tracks of FLAG, H4 acetylation and H3K27me3 at the posterior *HOXD* genes in the isogenic cells are shown in Figure 4A (see also Figure S4A for *HOXC* genes). Binding of EPC1(1-581) alone is detected on two precise locations coinciding with

small pre-existing H4 acetylation islands at the end of the *EVX2* gene and between *HOXD12* and *HOXD13*. In contrast, PHF1 associates with the entire region, as expected from the H3K27me3 signal detected throughout the same region. Interestingly, while the H3K27me3 may be increased by the expression of exogenous PHF1, the two small H4 acetylation islands persist. Most strikingly, the EPC1-PHF1 cell line, while expressing much less fusion protein compared to EPC1(1-581) and PHF1 (Figure S4C), clearly demonstrates appearance and mislocalization of H4 acetylation over a relatively large region (~25kb highlighted in yellow). Importantly, we can also observe a decrease in the levels of H3K27me3 in this same region, compared to the isogenic cell lines and also compared internally to the neighboring region (starting from the *HOXD12* gene). This local increase in H4 acetylation correlates with specific increase in expression of the *EVX2* and *HOXD13* genes located within the region (Figure 4B, see also Figure S4B for *HOXC* genes). To validate this observation, we performed ChIP-qPCR with primers to four sites of *de novo* H4 acetylation (*HOXD13*-A, B, C, and D highlighted in Figure 4A). Correcting for nucleosome occupancy with total H2B signal, we could confirm the specific increase in H4 acetylation at these locations in the cell line expressing EPC1-PHF1 (Figure 4C).

Replacement of canonical H2A with H2A.Z in chromatin is another enzymatic activity of the NuA4/TIP60 complex through its EP400 subunit (Billon and Cote, 2013; Pradhan et al., 2016). ChIP-qPCR with a H2A.Z specific antibody revealed a similar mislocalization of this variant histone at the *HOXD* sites (Figure 4D). These data clearly validate our hypothesis that the EPC1-PHF1 fusion complex can localize at genomic regions normally occupied by PRC2, mislocalizing the activities of the TIP60 complex to deregulate transcription.

We were intrigued at the apparent reduction in the levels of H3K27me3 at the *HOXD* gene cluster in the EPC1-PHF1 expressing cells (Figure 4A-highlighted in yellow). We observed neither a global decrease in H3K27me3 levels (Figure S2D) nor destabilization of the PRC2 complex in our cell line (Figure 2). Moreover, we could detect productive histone methylation by the EPC1-PHF1 mega-complex (Figure 2F, Figure S2C). In a previous study (Musselman et al., 2012), we demonstrated that the Tudor domain of PHF1 binds to H3K36me3 and constrains PRC2 mediated H3K27me3 activity. We hypothesized that this could be part of the mechanism leading to a local decrease of H3K27me3. We checked for the levels of H3K27me3 and H3K36me3 at the *HOXD13*-C locus by ChIP-qPCR to confirm this possible cross-talk. A decrease in H3K27me3 was observed in EPC1-PHF1 cells compared to controls, validating our ChIP-seq results (Figure 4E). We also saw an increase in H3K36me3 at the *HOXD13*-C locus in EPC1-PHF1 cells (Figure 4F), likely linked to the increased transcription of the gene. This cross talk was not observed at the NuA4/TIP60-bound and highly transcribed gene *RPSA* (Figure S4D)(Jacquet et al., 2016) or at *HOXA9*, a repressed gene not occupied by the EPC1-PHF1 fusion complex (Figure S4E). We conclude that the mislocalization of H4 acetylation and the variant histone H2A.Z at the posterior *HOXD* gene locus in cells expressing EPC1-PHF1 leads to de-repression and productive transcription. Subsequent deposition of H3K36me3 – a histone mark linked to transcription elongation - blocks the spread of the repressive H3K27me3 mark, possibly in part through its recognition by the Tudor domain of PHF1 and direct effect on EZH2/PRC2 activity (Finogenova et al., 2020; Jani et al., 2019; Musselman et al., 2012; Schmitges et al., 2011).

The EPC1-PHF1 mediated effect on histone acetylation/methylation at *HOXD* is restricted to a specific Topologically Associated Domain (TAD)

The *HOXD* gene cluster is involved in mammalian axial patterning, limb and genital development. Spatiotemporal regulation of the *HOXD* locus depends on long-acting multiple enhancer sequences located in "gene deserts" outside the *HOXD* cluster. Mammalian cells organize such regulatory landscapes in structural units called Topologically Associated Domains (TADs), which maintain genomic contacts even in the absence of transcription. These pre-formed 3D structures, delimited by CTCF and cohesin proteins, remain globally similar in cell types and are conserved from mouse to humans (reviewed in (Bompadre and Andrey, 2019; Szabo et al., 2019)). The *HOXD* locus is positioned between two large TADs (Rodriguez-Carballo et al., 2017)(Figure 5A). During limb development, the telomeric TAD is activated and controls the transcription of early *HOXD* genes (forearm), while later the centromeric TAD controls the posterior *HOXD* genes (digits)(Andrey et al., 2013). A strong boundary limits the posterior *HOX* genes from getting activated aberrantly (Lonfat and Duboule, 2015).

Chromatin modifications influence the dynamic mammalian TADs. The repressive H3K27me3 modification segregates chromatin into discrete sub-domains, and studies suggest that these 3D clusters require PcG proteins and H3K27me3 (Szabo et al., 2019). The *HOXD* cluster is known to harbor an H3K27me3 dense region, and the 1.8kb region between *HOXD11* and *HOXD12* (D11.12) may act as a putative mammalian Polycomb Responsive Element (PRE) (Woo et al., 2010). Recent studies have shown that this region acts as a nucleation hub for PRC2 to stably bind and spread H3K27me3 through intra- and interchromosomal interactions (Oksuz et al., 2018; Vieux-Rochas et al., 2015).

We observed that the Tip60/KAT5-mediated H4 acetylation and reduction of H3K27me3 at the *HOXD* locus in EPC1-PHF1 cells localized to the posterior *HOXD* genes. This led us to explore the possibility that a TAD boundary exists in this region. When we aligned the publicly available CTCF ChIP-seq data and Hi-C dataset in K562, we observed that the chromatin changes induced by the EPC1-PHF1 fusion protein are restricted to a large portion of the centromeric TAD at the *HOXD* locus (Figure 5B). Since Polycomb group protein-mediated long-range contacts are also at play at the *HOXD* locus, we looked for spreading of EPC1-PHF1-mediated H4 acetylation to other genomic loci that were shown to interact with the *HOXD* locus. Although existing Hi-C/5-C studies in K562 cells are not detailed enough to detect PcG-mediated TADs (Kundu et al., 2017) and many of the explored regions are not conserved from mouse to humans (Oksuz et al., 2018; Vieux-Rochas et al., 2015), we could detect the presence of H4 acetylation at the *HOXC* locus (H3K27me3 spreading site), *EVX1*, *HOXB13*, *CYP26B1* and *LHX2* (H3K27me3 nucleation sites) (Figure S5A-F). This observation suggests that the EPC1-PHF1 fusion protein could mislocalize activating histone marks through PcG-mediated chromatin contacts, activating an oncogenic transcriptional network. A similar mechanism was observed recently for an oncogenic EZH2 mutant that co-opted PcG-mediated long-range chromatin contacts to repress multiple tumor suppressors (Donaldson-Collier et al., 2019). Further work will be required to fully dissect the effect of the EPC1-PHF1 fusion complex on the structure and dynamics of TADs/polycomb-mediated chromatin contacts.

JAZF1 is a stochiometric interaction partner of the NuA4/TIP60 complex

Various subunits of the NuA4/TIP60 complex can be fused to different subunits from Polycomb Repressive complexes 1 and 2 in Endometrial Stromal Sarcoma (Figure 1A). However, the most frequently observed translocation is that of *JAZF1-SUZ12,* which occurs in more than 50% of all low grade ESS (LG-ESS (Ferreira et al., 2018)). There are other recurrent translocations involving the JAZF1 protein in sarcomas, namely *JAZF1-PHF1* and *JAZF1-BCORL1* (Allen et al., 2017; Micci et al., 2006; Schoolmeester et al., 2013). This connection led us to investigate whether JAZF1 interacts with the TIP60 complex.

The molecular function of the JAZF1 protein is largely unexplored. It is a transcription factor with three C2H2 zinc fingers, and studies have linked it to glucose production, lipid metabolism, ribosome biogenesis, metabolic disorders and pancreatic cancer (Kobiita et al., 2020; Liao et al., 2019; Nakajima et al., 2004; Zhou et al., 2020). The JAZF1 protein has a homolog in *S. cerevisiae* called Sfp1 (Figure 6A), an essential transcription factor involved in regulation of ribosomal protein (RP), ribosomal biogenesis(RiBi) and cell cycle genes, responding to nutrients and stress through regulation by TORC1 (Lempiainen et al., 2009; Marion et al., 2004). Crucially, Sfp1 and the yeast NuA4 complex cooperate functionally to regulate RP and RiBi genes (Figure 6B) (Loewith and Hall, 2011; Reid et al., 2000; Rossetto et al., 2014). High-throughput human interactome studies have also suggested a potential interaction of JAZF1 with TIP60 complex components (Hein et al., 2015; Huttlin et al., 2015).

Since *JAZF1* is not expressed in K562 cells (www.proteinatlas.org), we performed multiple cell compartment affinity purification coupled to tandem mass spectrometry (MCC-AP-MS/MS) in HEK293 as described in (Lavallee-Adam et al., 2013). This technique allowed us to isolate protein-protein interaction occurring on the chromatin fraction. By purifying JAZF1, we

recovered NuA4/TIP60 core components KAT5, EP400, RUVBL1/2, EPC1/2, TRRAP, BAF53a, BRD8, MBTD1, ING3, and DMAP1 with good reliability (Figure S6A). JAZF1 is highly expressed in the male and female reproductive tissues and may thus be involved in the tumorigenesis of LG-ESS of all fusion subtypes. To confirm this interaction, we ectopically expressed JAZF1-3xFLAG-2xSTREP in K562 cells from the *AAVS1* locus and tandem affinity-purified the protein from nuclear extracts. Silver staining of the purified fraction on gel, western blot and mass spectrometry analyses all confirm a very tight association of JAZF1 with the NuA4/TIP60 complex (Figure 6C, 6D and S6B), suggesting that JAZF1 can be a stable stoichiometric subunit. Accordingly, JAZF1 occupies the promoters of known TIP60-bound Ribosomal protein genes *RPSA* and *RPL36AL* as determined by ChIP-qPCR (Jacquet et al., 2016) (Figure 6E).

To analyze this tight interaction in the context of the *JAZF1-SUZ12* translocation, we performed co-immunoprecipitation experiments showing that the TIP60 complex still interacts strongly with the portion of JAZF1 (JAZF1(1-124)) that is fused to SUZ12 (Figure 6F, S6C). The same experiments showed that JAZF1 also still interacts with EPC1(1-581) and the EPC1-PHF1 mega-complex, but not with PHF1/PRC2.1. These conclusive data suggest that the pathogenesis of JAZF1 fusions could also involve the physical merge of NuA4/TIP60 with PRC2.

Conserved oncogenic mechanism of EPC1-PHF1 and JAZF1-SUZ12 fusion proteins

We have shown that the oncogenic EPC1-PHF1 fusion complex can lead to local transcription activation despite associating with a functional PRC2.1 complex. To further confirm the effect on transcription, we used a chemically inducible dCas9-based site-specific recruitment assay described in (Gao et al., 2016) in HEK293T cells (Figure 7A). In this system we could detect

specific transcriptional activation by the EPC1 protein, EPC1(1-581) and EPC1-PHF1, but not by PHF1 (Figure 7B). Interestingly, we found that the JAZF1 protein is a very robust transcriptional activator in this reporter system. This is also the case for the JAZF1-SUZ12 fusion protein, as well as for the truncated JAZF1 (1-124), but not SUZ12 by itself (Figure 7C).

To confirm that the JAZF1-SUZ12 fusion leads to the formation of a similar TIP60-PRC2 mega-complex, we repeated the purification of endogenous TIP60 and PRC2 subunits (as in Figure 2D-E), this time from cells expressing the JAZF1-SUZ12 fusion protein. Again, a merged TIP60-PRC2 mega-complex co-fractionated with TIP60 and PRC2 when the fusion protein was expressed in the cells (Figure 7D-E). Accordingly, the gene expression profiles of ESS with different fusions (JAZF1 and TIP60) cluster together (Dewaele et al., 2014; Micci et al., 2016; Przybyl et al., 2018). A recent interactome analysis of JAZF1-SUZ12 also validates our results (Piunti et al., 2019). Altogether, these data confirm that JAZF1 interacts with the TIP60 complex and that all of the currently detected fusion proteins in LG-ESS bridge the TIP60 complex to Polycomb repressive complexes.

To confirm that the JAZF1-SUZ12 fusion generates the same molecular events on chromatin as EPC1-PHF1 we looked again at the *HOXD* locus by ChIP-qPCR. In JAZF1-SUZ12 expressing cells, we again detected an increase in H4 acetylation (Figure 7F), a decrease in H3K27 methylation (Figure 7G, S7D) and an increase in transcriptional elongation-associated H3K36me3 (Figure 7H, S7C) in the region flanking the *HOXD13* gene. JAZF1-SUZ12 is thus able to mislocalize the TIP60 complex-associated histone mark through the JAZF1 protein, drawing parallels with the mislocalization mediated by EPC1-PHF1 (Figure 4).

The loss of H3K27me3 in these data can be explained by the loss of the ZnB domain of SUZ12 in the JAZF1-SUZ12 fusion protein, which has been shown to reduce the binding of JARID2 and EPOP subunits (Chen et al., 2018). SUZ12 acts as a platform to assemble the non-canonical subunits of the PRC2 complex, and a recent study shows that mutations in SUZ12 can modulate PRC2.1 vs PRC2.2 formation (Youmans et al., 2021). Furthermore, recent structural data suggest that the JARID2-containing PRC2.2 complex can partially override H3K36me3-mediated inhibition of EZH2 catalytic activity (Kasinath et al., 2021), whereas the PRC2.1 complex is known to be catalytically inhibited by the presence of H3K36me3 (Finogenova et al., 2020; Musselman et al., 2012). We hypothesize that the loss of JARID2 in JAZF1-SUZ12 skews it to associate with the PRC2.1 complex, which is catalytically inhibited by the presence of the H3K36me3 modification. Overall, our data indicate that the mislocalization of TIP60 activities could be the predominant oncogenic mechanism of the recurrent fusion proteins found in Endometrial Stromal Sarcomas and also other Sarcomas (Figure 7I).

DISCUSSION

Our study demonstrates that binding of the EPC1-PHF1 fusion protein increased histone acetylation and H2A.Z incorporation within regions regulated by PRC2, notably within the *HOXD* and *HOXC* gene clusters. This is usually correlated with increased gene expression, consistent with the notion that EPC1-PHF1 mistargets the HAT and histone exchange activities of NuA4/TIP60 to chromatin regions that are normally maintained in the silenced state (Figure 7D). Interestingly, mistargeting of histone acetylation at the *HOXD* cluster leads to loss of

H3K27 trimethylation over an entire specific topology-associated domain (TAD) without affecting the neighboring ones, albeit also bound by PRC2 complexes. In theory, since EPC1-PHF1 in integrated in a mega-complex with PRC2.1, it could act at most PRC2.1 binding sites. We speculate that the presence of small preexisting TIP60 binding sites within or in the vicinity of Polycomb/repressed regions, as we see at the *HOXD* cluster in the centromeric TAD, could make a big difference. The TIP60 moiety of the EPC1-PHF1 mega-complex could be efficiently recruited there (by DNA-bound factors and/or small regions of H3K4me3 in bivalent chromatin) while the PRC2 moiety would allow it to spread along the region like PRC2 normally does, leading to de novo acetylation. This would lead to increased transcription, coupled with H3K36 tri-methylation, which in turn inhibits the HMT activity of PRC2 (through recognition by PHF1 Tudor domain), leading to a decrease of H3K27me3 over the region (Figure 7D, middle panel). It is important to note that H3K36 methylation has also been described as a potent boundary mark to block H3K27me3 spreading (Streubel et al., 2018).

In addition, our study validates the protein JAZF1 as a potent transcription activator that stably associates with the NuA4/TIP60 complex. This information further integrates all of the reported translocations in Low-grade Endometrial Stromal sarcomas (LG-ESS) as producing a physical merge of TIP60 and PcG complexes, unifying the underlying molecular mechanism (Figure 7D, right panel). This is supported by the robust clustering of gene expression profiles of the LG-ESS with various fusion proteins when compared to High-grade ESS or tumors without fusion genes (Micci et al., 2016). Our findings further underline the importance of deregulated epigenetic modifiers in promoting oncogenesis, particularly in sarcomas (Nacev et al., 2020), in part through changes of the chromatin landscape over large genomic regions. The highly

recurrent physical merger of TIP60 and PcG complexes in these sarcomas, through several distinct subunits, indicates that this event is the driving force in the oncogenic mechanism. It is somewhat reminiscent of the MLL fusions proteins in pediatric acute myeloid leukemia, in which the H3K4 HMT MLL is recurrently fused with diverse components of the super-elongation complex, creating a large complex driving oncogenesis through altered transcription programs involving chromatin modifiers and readers as well as *HOXA* cluster misregulation (Mohan et al., 2010). A parallel can also be drawn with the discovery of various BRD4-linked Z4 complex proteins that are fused to NUT in NUT midline carcinoma (Shiota et al., 2018).

The chromatin modification changes associated with EPC1-PHF1 are delimited to a TAD (Figure 5). This is especially interesting given recent studies demonstrating the nucleation and spreading of H3K27me3 through Polycomb domain contacts (Oksuz et al., 2018; Vieux-Rochas et al., 2015). Interestingly, EPC1-PHF1 localizes to the *HOXD13*/*EVX2* nucleation region and mislocalizes H4 acetylation only in the centromeric TAD. We could also observe EPC1-PHF1 localization and H4 acetylation presence at "spreading" sites (Figure S5). A similar observation was made for the Ewing sarcoma fusion protein at the *HOXD* centromeric TAD leading to overexpression of *HOXD13* as seen here (Svoboda et al., 2014; von Heyking et al., 2016). A question that emerges from our observation is whether mislocalization of activating marks will lead to disruption of these H3K27me3-centric repressive domains and disrupt the organization of chromatin domains. Disruption of polycomb-mediated repression by fusion proteins to activate stem cell-related gene expression programs has emerged as a common theme recently (Brien et al., 2019). Unlike in Synovial sarcoma or NUT midline carcinoma, we do not observe extensive mislocalization/disruption in chromatin. This could be an artifact of our system of

study. Alternatively, the LG-ESS pathology is rare, milder, not rapidly progressing and with limited metastasis. LG-ESS may hence only deregulate few genes. In the BRD-NUT fusion cancers, the fusion of an acetylation reader protein (BRD4) and HAT interacting protein (NUT) creates a complex that can allosterically activate its activity, forming mega-domains of chromatin changes (Alekseyenko et al., 2017). Unlike BRD-NUT, in the EPC1-PHF1 fusion, the "Read-Write" allosteric activation ability of PRC2 is curtailed by the decrease in H3K27me3 and increase in H3K36me3. Intriguingly, a newly described subunit of PRC2 complexes, EZHIP (CXORF67), was also found fused to the TIP60 subunit MBTD1 in LG-ESS (Dewaele et al., 2014). EZHIP functions similarly to the oncogenic H3K27M mutant by binding to EZH2 and inhibiting its activity in trans (Hubner et al., 2019; Jain et al., 2019; Piunti et al., 2019). This mechanism seems similar to the trans-inhibition of EZH2 through PHF1 Tudor-H3K36me3 binding and recently described direct inhibitory effect of H3K36me3 on EZH2 catalytic activity (Finogenova et al., 2020; Jani et al., 2019). The molecular characterization of MBTD1-EZHIP will further reveal if parallels can be drawn with EPC1-PHF1.

While the *HOXA9* gene has been implicated in many cancers, *HOXD13* and *EVX2* are also deregulated in some cancers albeit more rarely (Luo et al., 2019). Posterior *HOX* genes are involved in the development of reproductive organs, are endometrial developmental markers and implicated in the differentiation of Endometrial stromal cells (Akbas and Taylor, 2004; Miyazaki et al., 2018). Since Endometrial Stromal sarcomas and OFMTs are very rare, we could not have timely access to patient tissues or patient tissue-derived cell lines to model EPC1-PHF1 fusion-mediated oncogenesis. Moreover, the cell of origin of soft tissue sarcomas is still not clear and differentiated endometrial cells are difficult to expand and clone. Our cellular model

system for the study is undifferentiated/immature and highly proliferative, therefore heterologous and does not truly represent the pathways and gene expression programs in ESS. Considering these limitations, we use caution in identifying ESS relevant gene expression changes from our study. Nevertheless, the data presented indicate that the mislocalization of TIP60 leads to histone H4 acetylation and H2A.Z incorporation in chromatin, favoring local gene expression that can subsequently inhibits PRC2 activity. This succession of molecular events is likely the predominant oncogenic mechanism of the recurrent fusion proteins found in Endometrial Stromal Sarcomas, Ossifying Fibromyxoid Tumors, and Cardiac Ossifying Sarcomas.

Since the mutational burden of translocation carrying cancers is relatively lesser than other cancers (clonal homogeneity), a true characteristic of sarcomas, targeted therapeutic approaches specific to the fusion protein are plausible (Brien et al., 2019; Taylor et al., 2011). According to our study, targeting TIP60 HAT activity with available small molecule inhibitors seems a viable approach (Brown et al., 2016). Alternatively, development of inhibitors designed to disrupt the interaction between JAZF1 and the TIP60 complex may also be a promising therapeutic approach.

Acknowledgments

We thank Céline Roques, Catherine Lachance, Valérie Côté and Philippe Cloutier for important technical support. We are very grateful to Prof. Francesca Micci for providing the cDNA obtained from patient sample that covered the EPC1-PHF1 fusion. We thank Compute Canada for the use of supercomputers, McGill Genome Center for sequencing/expression microarray and the CHUL proteomic platform. This work was supported by grants from the following: the

Canadian Institutes of Health Research (CIHR) to J.C. (FDN-143314); the Government of Québec, Ministry of Economy and Innovation to B.C.; the Natural Sciences and Engineering Research Council of Canada to Y.D. (RGPIN-2014-059680) and J.-P.L. (RGPIN-2017-06124); University of Toronto startup funds to M.T. D.S., M.-E.L. and K.J. were supported by PhD studentships from Fonds de la Recherche Québec-Santé (FRQS) and Nature/Technologie (FRQNT). N.A. was supported by a CIHR post-doctoral fellowship. B.C. holds the IRCM Bell-Bombardier Research Chair. A.-C.G. holds the Canada Research Chair in Functional Proteomics and the Lea Reichmann Chair in Cancer Proteomics. J.-P.L. and Y.D. are Junior 1 and Junior 2 FRQS scholars, respectively. J.C. holds the Canada Research Chair in Chromatin Biology and Molecular Epigenetics.

Author Contributions

D.S., N.A., M.-E.L., B.C., M.T., Y.D. and J.C. designed the experiments. D.S., N.A., M.-E.L., N.A., A.M., K.J., J.R., J.-P.L., J.L., and Y.D. performed the experiments. E.P., S.T.S. analyzed the genomic data. A.-C.G., M.T., B.C., Y.D. and J.C. supervised and secured funding. D.S. and J.C. wrote the manuscript with the help of co-authors.

Declaration of Interests: The authors declare no conflict of interest

FIGURE LEGENDS:

Figure 1. Recurrent translocations fusing NuA4/TIP60 components and Polycomb repressive complexes in soft tissue sarcomas.

(A) Table summarizing recurrent translocations found in soft-tissue sarcomas that fuse NuA4/TIP60 complex proteins and Polycomb group proteins.

(B) Schematic representation of the 17-subunit NuA4/TIP60 complex, illustrating the various subunits, activities (Lysine acetylation, H2A.Z exchange) and substrates recognized by reader domains.

(C) Schematic representation of the PRC2.1 complex illustrating the various subunits, activity (H3K27me1/2/3) and substrates recognized by reader domains.

(D) Schematic representation of the EPC1-PHF1 fusion protein. The numbers indicated are amino acids, protein domains retained in the fusion are indicated.

Figure 2. The EPC1-PHF1 fusion protein assembles a mega-complex merging NuA4/TIP60 and PRC2 complexes, and their activities.

(A) Silver stained SDS-PAGE showing affinity purified complexes. Purification was performed from K562 cell lines expressing the FLAG tag only (Mock), PHF1, EPC1 (1-581) or EPC1-PHF1 from the *AAVS1* locus. Labels on the right side are proteins that were identified based on western blotting and predicted molecular weights. The EPC1-PHF1 fraction clearly shows both NuA4/TIP60 as well as PRC2 subunits. See also Figure S1.

(B) Mass Spectrometry analysis of the affinity purified complexes. Table with total spectral counts of peptides present in the affinity purifications shown in (A). Peptides highlighted in green are NuA4/TIP60 complex components; peptides highlighted in purple are PRC2.1 complex subunits. See also Figure S2A.

(C) Western blots of selected NuA4/TIP60 and PRC2 complex subunits on the affinity purified complexes shown in (A). Antibodies specific to NuA4/TIP60 complex subunits (anti-TRRAP, EP400, KAT5, DMAP1, ING3) and PRC2 complex subunits (anti- EZH2 and SUZ12) are used.

(D, E) Affinity Purification of endogenous NuA4/TIP60 and PRC2 complexes confirming formation of a merged mega-complex by EPC1-PHF1. (D) K562 cells with endogenously tagged EP400 (C-terminal 3XFLAG-2XStrep) created by CRISPR/Cas9 gene targeting (Dalvai et al., 2015) were used to express EPC1-PHF1 (3xHA-2xStrep tag) or tag only (Mock) from the *AAVS1* locus. Upon immunoprecipitation of endogenous EP400 with anti-FLAG resin, endogenous SUZ12 is detected only in the EPC1-PHF1 expressing cells. (E) K562 cells with endogenously tagged EZH2 (N-terminal 3xFLAG-2xStrep) created by TALEN-gene targeting (Dalvai et al., 2015) were used to express EPC1-PHF1 (3xHA-2xStrep tag) or tag only (Mock) from the *AAVS1* locus. Upon immunoprecipitation of endogenous EZH2 with anti-FLAG resin, endogenous KAT5 and DMAP1 are detected in the EPC1-PHF1 expressing cells but not in the control.

(F) *In vitro* Histone Acetyltransferase (HAT) assay and Histone Methyltransferase (HMT) assay. HAT and HMT assays were performed with affinity purified complex fractions shown in (A). KAT5 purified through EPC1(1-581) as well as EPC1-PHF1 fusion show HAT activity towards H2A and H4 on chromatin (Upper panel, Lane 2 and 3). EZH2 purified through PHF1 as well as EPC1-PHF1 fusion show HMT activity towards H3 on chromatin (Upper panel, Lane 7 and 8).

Coomassie stained SDS-PAGE gel is the loading control for nucleosomal histones. See also Figures S2B-C.

(E) Schematic representation of the chimeric mega-complex assembled by the EPC1-PHF1 fusion protein retaining both enzymatic activities. The possible combined molecular activities and reader modules are indicated.

Figure 3. EPC1-PHF1 fusion complex is enriched at loci with bivalent chromatin and induces global changes in histone marks and gene expression.

(A) Venn diagram showing the number FLAG ChIP-seq peaks and overlap between EPC1(1-581), PHF1, and EPC1-PHF1 expressing K562 cell lines. See also Figure S3A.

(B) Percentage of FLAG ChIP-seq peaks overlapping bivalent chromatin: EPC1-PHF1 binds bivalent chromatin (genomic regions co-occupied by both H3K4me3 and H3K27me3 in K562 cells) at a higher percentage (45%) compared to EPC1(1-581) (30.2%) or PHF1 (8.8%).

(C and D) Integrative analysis of ChIP-seq and Microarray: (B) Genes with increased H4 acetylation show increased gene expression, EPC1-PHF1 expressing cells show higher change compared to controls; (C) Genes with decreased H3K27me3 occupancy show increased gene expression, EPC1-PHF1 expressing cells show higher change compared to controls. See also Figure S3B-C.

Figure 4. EPC1-PHF1 induces changes in the chromatin landscape at the *HOX* gene clusters.

(A) Representative ChIP-seq tracks at the *HOXD* gene cluster, highlighted region in yellow shows H4 acetylation mislocalization (especially at regions labelled *HOXD13*-A, B, C, D) and reduction

in H3K27me3 in EPC1-PHF1 expressing K562 cell line. The Y axis represents normalized reads per million total reads (rpm). See Figure S4A for similar effect at the *HOXC* cluster.

(B) Gene expression changes at the *HOXD* cluster in K562 cells expressing EPC1-PHF1 compared to cells expressing tag only (Mock). See Figure S4B for gene expression changes at the *HOXC* cluster.

(C) Mislocalization of H4 acetylation determined by ChIP-qPCR. EPC1-PHF1 expressing K562 cell line (Bar graph in maroon color) shows increased H4 acetylation compared to control cell lines at *HOXD13*-A, B, C, D regions, validating the results of ChIP-seq in (A). (Values are represented as ratio of %IP/input chromatin of H4ac and H2B) (n=2, error bars represent Range of the values). Intergenic (chr12: 65,815,182-65,815,318) and p21 promoter are negative controls.

(D) Mislocalization of variant histone H2A.Z determined by ChIP-qPCR. EPC1-PHF1 expressing K562 cell line (Bar graph in Maroon color) show increased H2A.Z occupancy compared to control cell lines at regions of H4 acetylation mislocalization (*HOXD13*-A, B, C, D) shown in (A). (Values are represented as ratio of %IP/input chromatin of H2A.Z and H2B) (n=2, error bars represent Range of the values). Intergenic (chr12: 65,815,182-65,815,318) and p21 promoter are negative controls.

(E, F) Decrease in H3K27me3 levels correlates with increase in H3K36me3: (E) ChIP-qPCR with H3K27me3 antibody shows decreased level at *HOXD13*-C region (gene body) only in EPC1-PHF1 expressing K562 cell line (Bar graph in Maroon color), validating our ChIP-seq results in (A). (F) ChIP-qPCR with H3K36me3 antibody shows increased level at *HOXD13*-C region (gene body) only in EPC1-PHF1 expressing K562 cell line (Bar graph in Maroon color). MyoD promoter is used as a negative control. (Values are represented as ratio of %IP/input chromatin of

H3K36me3 or H3K27me3 and H3) (n=2, error bars represent Range of the values). See also Figures S4D-E.

Figure 5. EPC1-PHF1 induced chromatin changes at the *HOXD* cluster are restricted to a Topologically Associated Domain (TAD).

(A) Schematic representation of the Topologically Associated Domains (TADs) and sub domains at the *HOXD* locus in mammalian cells.

(B) Alignment of Hi-C data in K562 cells (Rao et al., 2014) with CTCF ChIP-seq in K562 cells (ENCODE/Broad histone marks) and ChIP-seq from this study (Figure 4A). Highlighted regions indicate histone modification changes confined to a topological domain in EPC1-PHF1 expressing cell line compared to tag-only expressing cell line (Mock). See also Figure S5.

Figure 6: The JAZF1 protein stably associates with the NuA4/TIP60 complex.

(A) Protein sequence alignment: Human JAZF1 and yeast Sfp1 show homology in the zinc-finger domains. Black boxes show exact amino acid conservation, grey boxes show conservation of amino acids with similar polarity.

(B) Model depicting the relationship between SFP1 and NuA4 in *S. cerevisiae*.

Sfp1 is a transcriptional activator that responds to nutrients and stress, being regulated by TORC1 and Mrs6. In the nucleus, Sfp1 and the NuA4 complex co-operate to regulate the expression of Ribosomal Protein (RP) and Ribosomal Biogenesis (RiBi) genes.

(C) Silver stained SDS-PAGE showing tandem affinity-purified JAZF1. Purification was performed from K562 cell lines expressing the tag only (Mock), or JAZF1-3xFLAG-2xSTREP from the *AAVS1*

locus. Labels on the right side are proteins that were identified based on western blotting and predicted molecular weights. The JAZF1 fraction shows all subunits of the NuA4/TIP60 complex. See also Figure S6A-B for Mass Spectrometry analysis of JAZF1 purification from HEK293 and K562 cells, respectively.

(D) Western blots of selected NuA4/TIP60 subunits on the affinity purified fraction shown in (C). Antibodies specific to NuA4/TIP60 components (anti-KAT5, EP400, DMAP1, MBTD1, BRD8, MRG15) and anti-FLAG (JAZF1) are used.

(F) FLAG-JAZF1 ChIP-qPCR. Anti-FLAG ChIP-qPCR in HEK293 cells stably expressing FLAG tagged JAZF1. Enrichment of JAZF1 can be seen at ribosomal protein gene promoters (*RPSA* and *RPL36AL*) regulated by NuA4/TIP60. *CCNA2* is a positive control for JAZF1 (Values are represented as %IP/input chromatin, n=2, Error bars represent range of the values).

(E) Immunoprecipitation of JAZF1 with the EPC1-PHF1 fusion. Myc-tagged JAZF1 (1-124) construct was transiently expressed in FLAG tagged K562 cell lines with stable integration of Tag (Mock), EPC1(1-581), PHF1 or EPC1-PHF1 at the *AAVS1* locus. Immunoprecipitation was performed with Anti-FLAG-resin and analyzed by western blotting with Anti-Myc antibody. JAZF1 (1-124) corresponding to the portion in the JAZF1-SUZ12 fusion protein interacts with EPC1(1-581) and EPC1-PHF1. See Figure S6C for data with full length JAZF1 construct.

(F) FLAG-JAZF1 ChIP-qPCR in HEK293 cells. Anti-FLAG ChIP-qPCR in HEK293 cells stably expressing FLAG-tagged JAZF1. Enrichment of JAZF1 can be seen at ribosomal protein gene promoters (RPSA and RPL36AL) regulated by NuA4/TIP60. CCNA2 is a positive control for JAZF1. (Values are represented as % of input chromatin, n=2, Error bars represent range of the values).

Figure 7: The JAZF1-SUZ12 fusion protein has molecular impacts similar to EPC1-PHF1.

(A) Model for the dCas9-based inducible reporter gene activation assay in (B-C). dCas9 and the putative activator were fused to abscisic acid inducible heterodimerization domains ABI and PYL1, respectively. HEK293T cell line with TRE3G-EGFP reporter was used for recruiting dCas9 with gRNA targeting seven TetO repeats in the TRE3G promoter. Activation efficiency of the putative activator was tested by ABI-PYL1 dimerization induced with abscisic acid (ABA). ORFs which are activators induce the expression of EGFP, which is quantified by flow cytometry.

(B,C) Transcription activation assay quantified by Flow cytometry analysis. Panel B shows the % of cells transfected with EPC1, PHF1, EPC1-PHF1 or EPC1 (1-581) expressing high GFP upon ABA treatment. EPC1-PHF1 is a transcriptional activator similar to EPC1 and EPC1-N terminus, whereas PHF1 is not. The weaker EPC-PHF1-mediated activation compared to EPC1 may be linked to the usual lower expression level of the larger fusion. Panel C shows the % of cells expressing high GFP when transfected with JAZF1, SUZ12, JAZF1-SUZ12, JAZF1(1-129) and SUZ12(93-740) and treated with ABA. JAZF1 (Full length and N terminus) and JAZF1-SUZ12 act as transcriptional activators whereas SUZ12 does not. At least 25,000 cells were analyzed for each replicate. Error bars represent SEM of 5 independent repeats. DMSO is used as a negative control.

(D,E) Formation of a TIP60-PRC2 mega-complex by JAZF1-SUZ12. (D) K562 cells with endogenously tagged EZH2 (N-terminal 3xFLAG-2xStrep) created by TALEN-gene targeting (Dalvai et al., 2015) were used to express JAZF1-SUZ12 (3xHA-2xStrep tag) or tag only (Mock) from the *AAVS1* locus. Upon immunoprecipitation of EZH2 with anti-FLAG resin, endogenous

Tip60 can be detected only in the JAZF1-SUZ12 expressing cells. (E) K562 cells with endogenously tagged EP400 (C-terminal 3xFLAG-2xStrep) created by CRISPR-gene targeting (Dalvai et al., 2015) were used to express JAZF1-SUZ12 (3xHA-2xStrep tag) or tag only (Mock) from the *AAVS1* locus. Upon immunoprecipitation of EP400 with anti-FLAG resin, endogenous EZH2 can be detected only in the JAZF1-SUZ12 expressing cells.

(F) Mislocalization of H4 acetylation determined by ChIP-qPCR. JAZF1-SUZ12 expressing K562 cell line (Bar graph in Maroon color) shows increased H4 acetylation compared to control cell lines at *HOXD13*-A, B, C, D regions, similar to EPC1-PHF1 in Figure 4C (Values are represented as ratio of % of input chromatin of H4ac and H3) (n=2, error bars represent Range of the values). *RPSA* promoter is a positive control for H4 acetylation, *MYOD* promoter is a negative control. See also Figure S7A-B.

(G, H) Decrease in H3K27me3 levels at the *HOXD* locus correlates with increased H3K36me3 in JAZF1-SUZ12 expressing cells: (G) ChIP-qPCR with H3K27me3 antibody shows decreased level at *HOXD13*-C region (gene body) in JAZF1-SUZ12 expressing K562 cell line (Bar graph in Maroon color). *MyoD* promoter is a negative control. (H) ChIP-qPCR with H3K36me3 antibody shows increased level at *HOXD13*-C region (gene body) only in JAZF1-SUZ12 expressing K562 cell line (Bar graph in Maroon color). *MyoD* promoter is a negative control. (Values are represented as ratio of % of input chromatin of H3K27me3 or H3K36me3 and H3) (n=2, error bars represent Range of the values). See also Figure S7C-D.

(I) Model for the oncogenic mechanism of EPC1-PHF1 fusion protein in soft tissue sarcomas. 1) In normal cells, PRC2 complex occupies regions of poised bivalent chromatin that possess both activating H3K4me3 and repressive H3K27me3. 2) When EPC1-PHF1 fusion protein is

expressed, it assembles a mega-complex combining NuA4/TIP60 and PRC2 complexes. The mega-complex occupies bivalent regions (such as the *HOX* clusters in K562) and mislocalizes NuA4/TIP60 activities (H4, H2A acetylation, H2A.Z exchange). This tips the balance towards transcriptional activation. Levels of the transcription elongation associated histone mark H3K36me3 increases and inhibits the deposition of the repressive H3K27me3 histone mark. Thus, the changes in chromatin landscape potentiate the expression of oncogenes. 3) Sarcomas with fusions of JAZF1 and PcG proteins also employ a similar mechanism since JAZF1 strongly interact with the NuA4/TIP60 complex and mislocalize its activities.

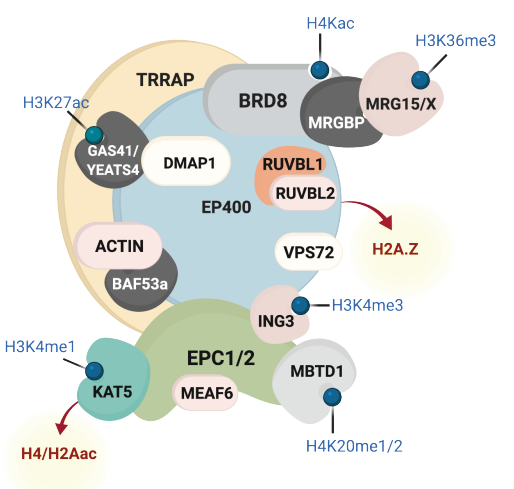
FIGURE 1

A

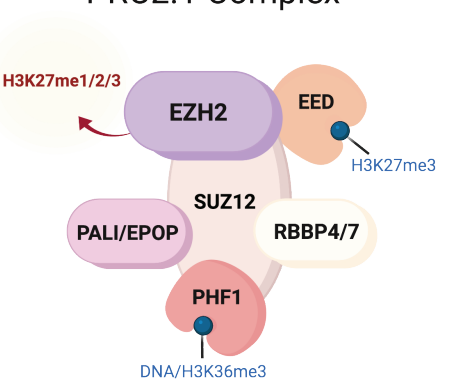
RECURRENT TIP60 & PcG TRANSLOCATIONS

FUSION PROTEIN	COMPLEXES INVOLVED	REFERENCE
Low Grade-Endometrial Stromal Sarcoma		
EPC1-PHF1	TIP60, PRC2.1	Micci F et al., 2006
EPC2-PHF1	TIP60, PRC2.1	Brunetti M et al., 2018
MEAF6-PHF1	TIP60, PRC2.1	Pangopoulos I et al., 2012
BRD8-PHF1	TIP60, PRC2.1	Micci F et al., 2017
MBTD1-PHF1	TIP60, PRC2.1	Han L et al., 2020
EPC1-SUZ12*	TIP60, PRC2	Dickson BC et al., 2018
MBTD1-EZHIP	TIP60, PRC2	Dewaele B et al., 2014
MEAF6-SUZ12	TIP60, PRC2	Makise N et al., 2019
EPC1-BCOR*	TIP60, PRC1.1	Dickson BC et al., 2018
ING3-BCOR**	TIP60, PRC1.1	Lin DI et al., 2020
JAZF1-PHF1	? , PRC2.1	Micci F et al., 2006
JAZF1-SUZ12	? , PRC2	Koontz JI et al., 2001
JAZF1-BCORL1	?, PRC1.1	Allen AJ., 2017
Ossifying Fibromyxoid Tumor		
EPC1-PHF1	TIP60, PRC2.1	Antonescu CR et al., 2014
EP400-PHF1	TIP60, PRC2.1	Endo M et al., 2013
MEAF6-PHF1	TIP60, PRC2.1	Antonescu CR et al., 2014
Cardiac Ossifying Sarcoma		
JAZF1-PHF1	?, PRC2.1	Schoolmeester JK et al., 2013

B TIP60 Complex



C PRC2.1 Complex



D SCHEMATIC OF THE EPC1-PHF1 FUSION PROTEIN

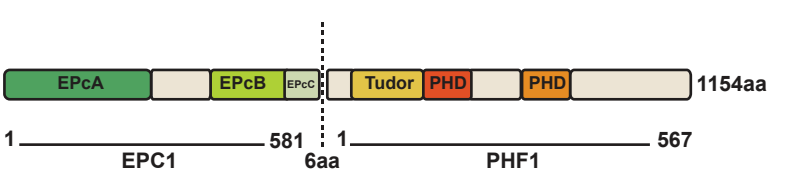
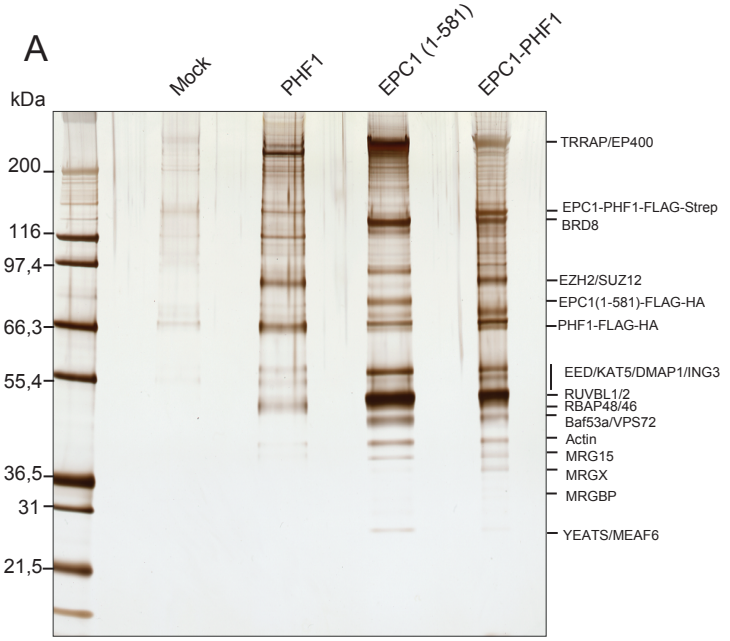


FIGURE 2

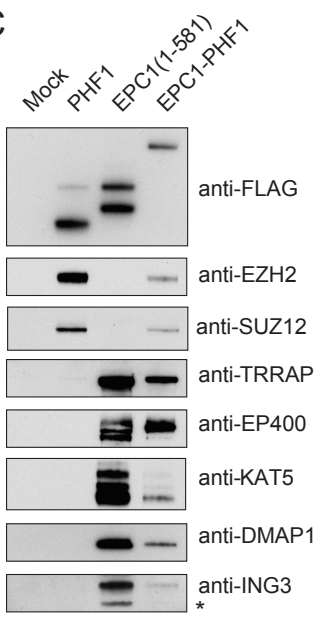
A



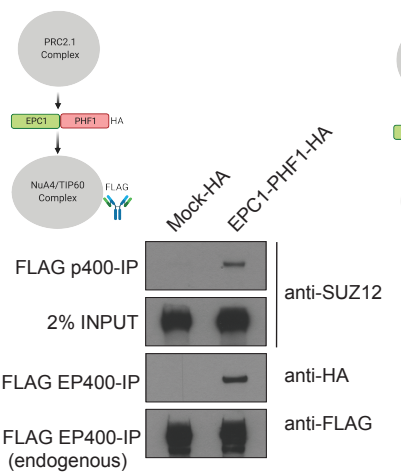
B

	PROTEIN	SPECTRAL COUNTS		
		PHF1	EPC1(1-581)	EPC1-PHF1
TIP60/NuA4 Complex	TRRAP		291	197
	EP400		216	138
	RUVBL2		101	94
	RUVBL1		125	92
	BRD8		57	54
	BAF53a		75	65
	DMAP1		47	37
	KAT5		41	31
	EPC1		54	43
	VPS72		26	18
	ING3		19	16
	MRG15		4	5
	MRGBP		1	2
	GAS41/YEATS4		3	2
PRC2.1 Complex	PHF1	25		32
	SUZ12	22		20
	EED	16		11
	RBBP4	13		13
	EZH2	9		14
	LCOR/PALI1	4		1

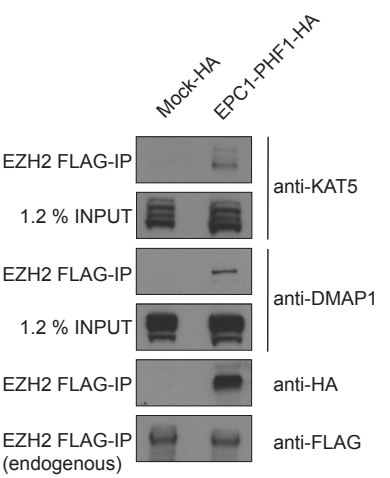
C



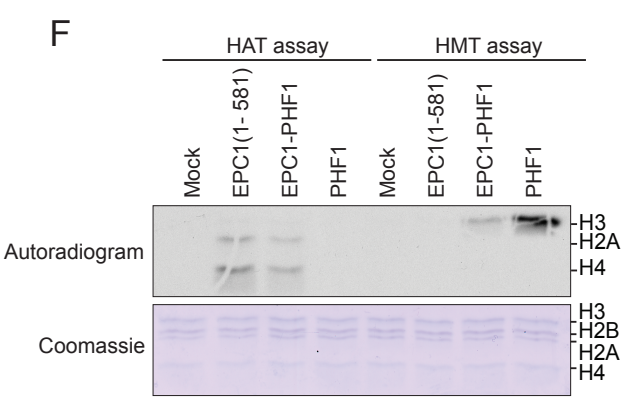
D



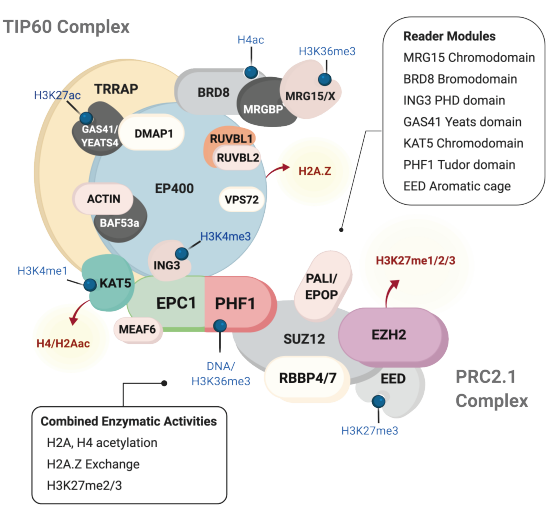
E



F



G



EPC1-PHF1 MEGA COMPLEX

FIGURE 3

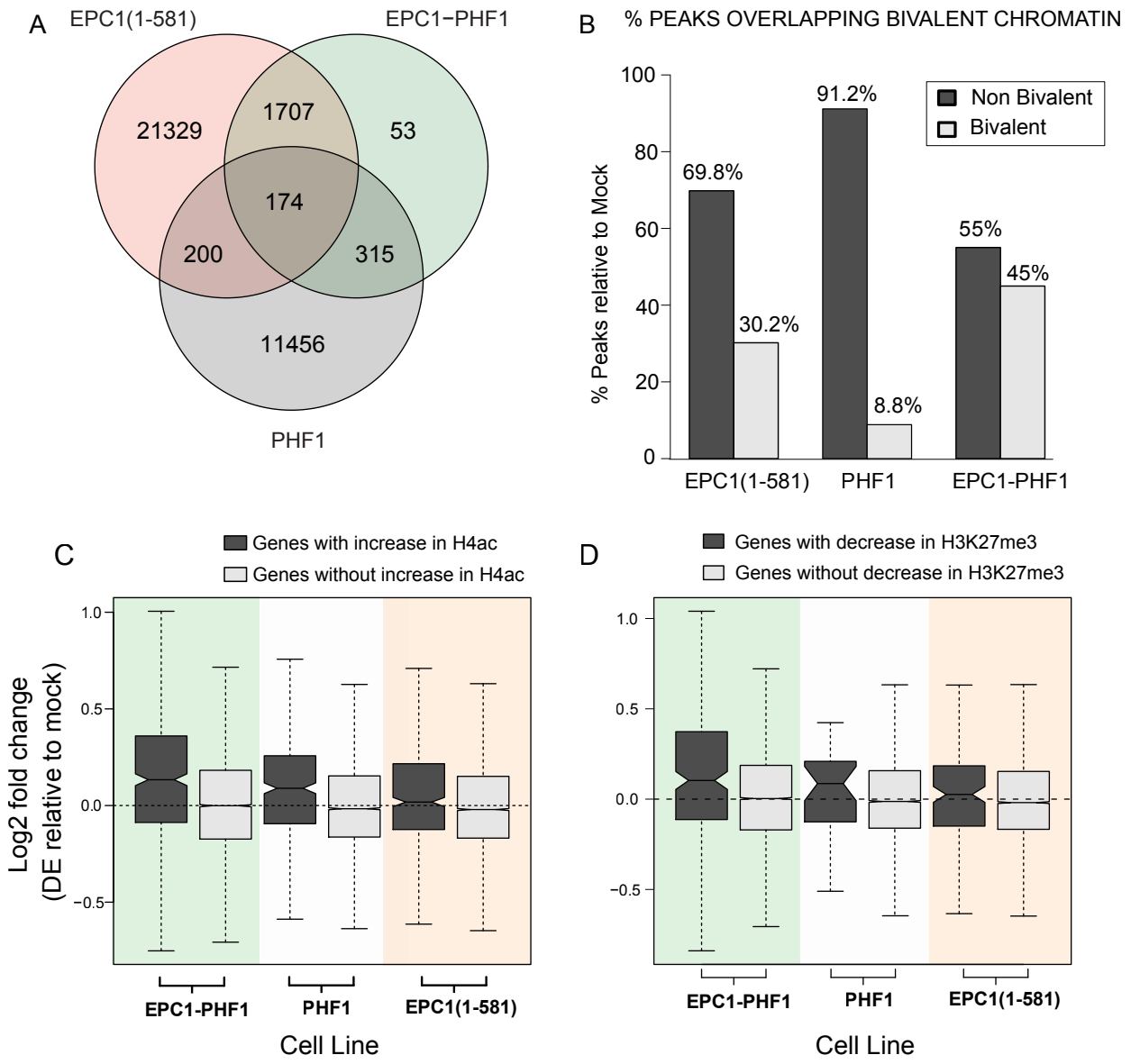

A
EPC1(1-581)
EPC1−PHF1
PHF1
21329
1707
53
174
200
315
11456
B
% PEAKS OVERLAPPING BIVALENT CHROMATIN
Non Bivalent
Bivalent
% Peaks relative to Mock
100
80
60
40
20
0
69.8%
30.2%
91.2%
8.8%
55%
45%
EPC1(1-581)
PHF1
EPC1-PHF1
C
Genes with increase in H4ac
Genes without increase in H4ac
Log2 fold change
(DE relative to mock)
1.0
0.5
0.0
−0.5
EPC1-PHF1
PHF1
EPC1(1-581)
Cell Line
D
Genes with decrease in H3K27me3
Genes without decrease in H3K27me3
1.0
0.5
0.0
−0.5
EPC1-PHF1
PHF1
EPC1(1-581)
Cell Line

FIGURE 4

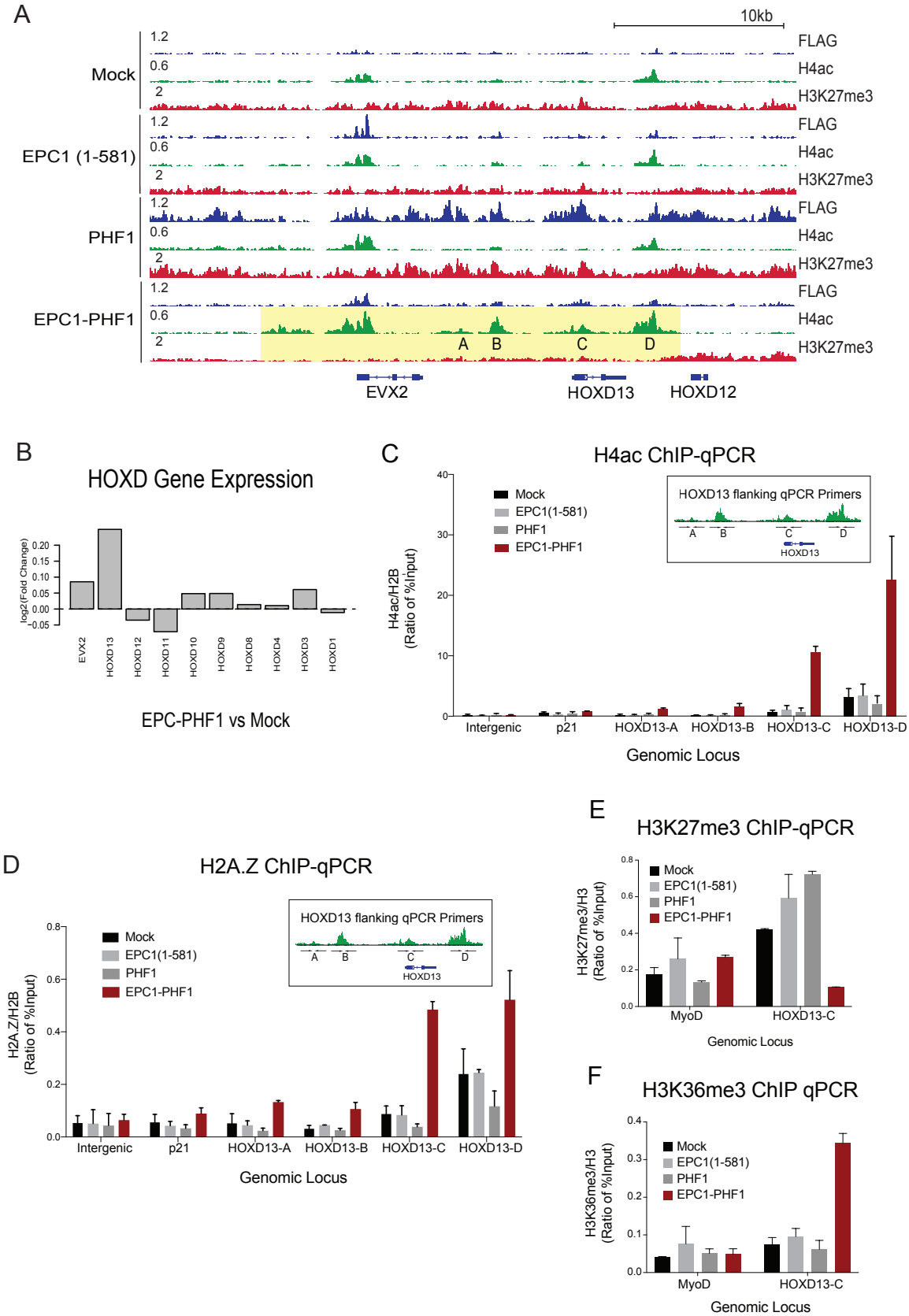

A
10kb
Mock
EPC1 (1-581)
PHF1
EPC1-PHF1
FLAG
H4ac
H3K27me3
1.2
0.6
2
A
B
C
D
EVX2
HOXD13
HOXD12
B
HOXD Gene Expression
log2(Fold Change)
EVX2
HOXD13
HOXD12
HOXD11
HOXD10
HOXD9
HOXD8
HOXD4
HOXD3
HOXD1
EPC-PHF1 vs Mock
C
H4ac ChIP-qPCR
H4ac/H2B (Ratio of %Input)
Mock
EPC1(1-581)
PHF1
EPC1-PHF1
HOXD13 flanking qPCR Primers
Intergenic
p21
HOXD13-A
HOXD13-B
HOXD13-C
HOXD13-D
Genomic Locus
D
H2A.Z ChIP-qPCR
H2A.Z/H2B (Ratio of %Input)
E
H3K27me3 ChIP-qPCR
H3K27me3/H3 (Ratio of %Input)
MyoD
F
H3K36me3 ChIP qPCR
H3K36me3/H3 (Ratio of %Input)

FIGURE 5

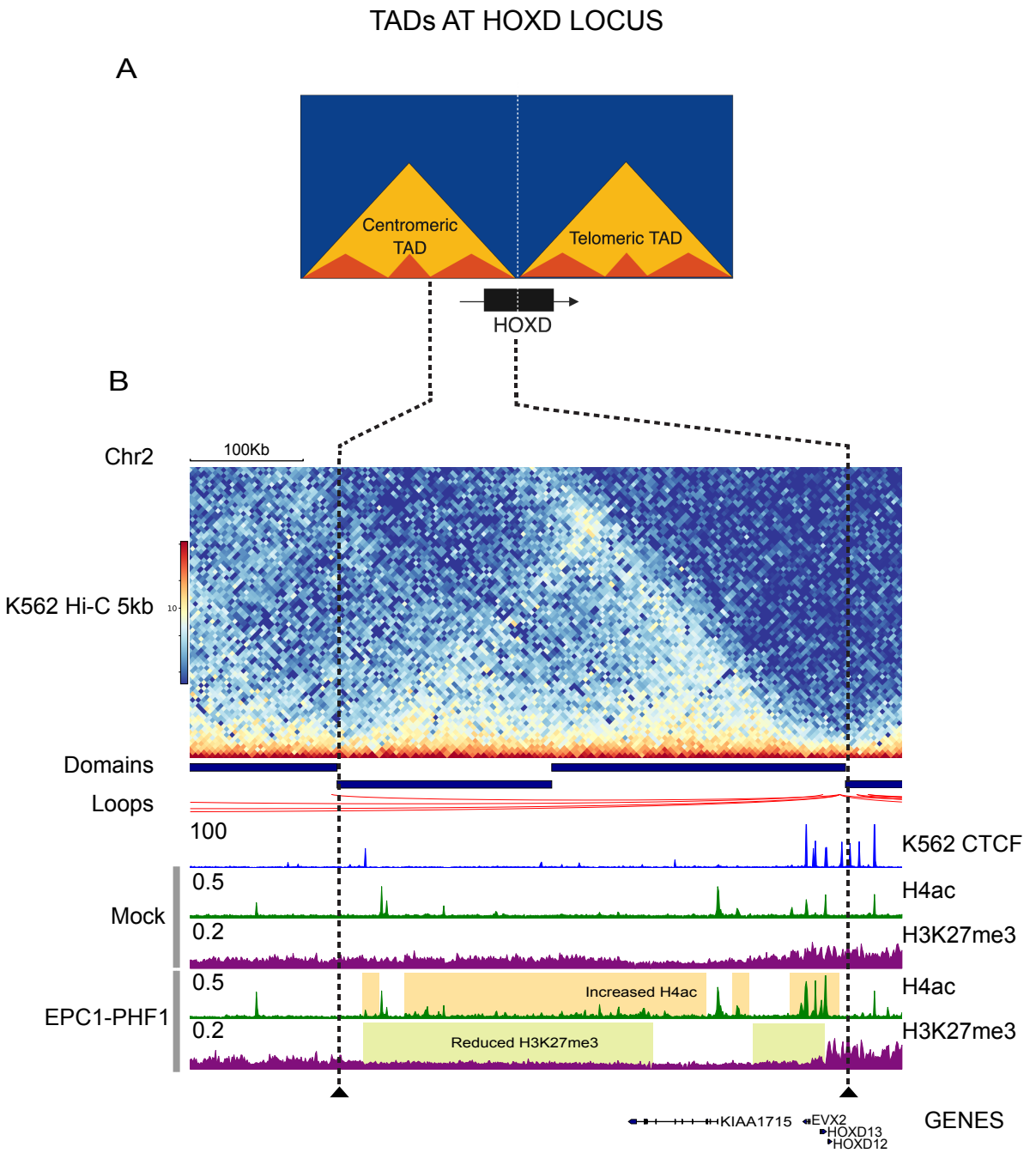
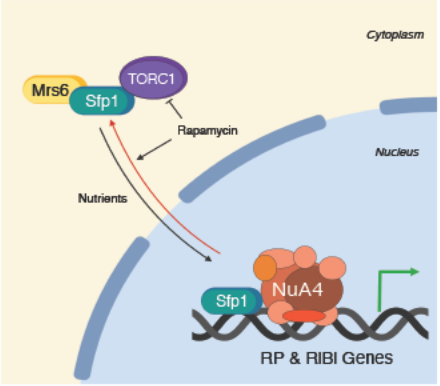

TADs AT HOXD LOCUS
A
Centromeric TAD
Telomeric TAD
HOXD
B
Chr2
100Kb
K562 Hi-C 5kb
10
Domains
Loops
100
K562 CTCF
0.5
H4ac
Mock
0.2
H3K27me3
0.5
Increased H4ac
H4ac
EPC1-PHF1
0.2
Reduced H3K27me3
H3K27me3
KIAA1715
EVX2
HOXD13
HOXD12
GENES

FIGURE 6

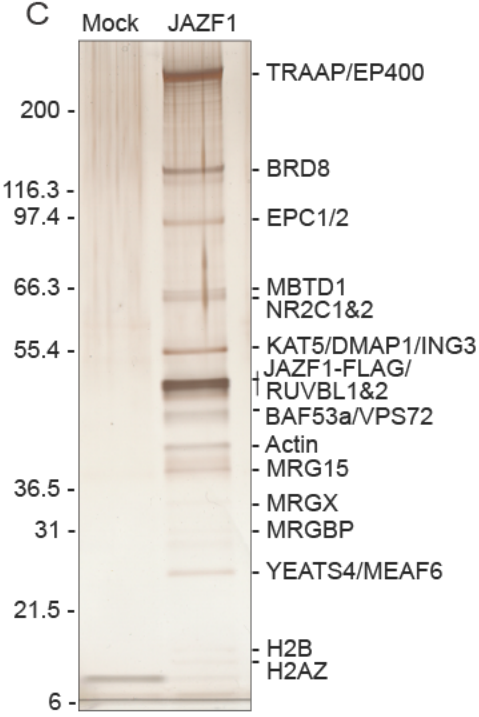
A

SFP1	198	QYCKDYSCCGLSLEGLHDLLRHYEEAHISTSP	Zinc Finger (12 - 37)
JAZF1	12	NTCR-FGGCGLHFETLADLIEHIEINHIDTDP	
SFP1	596	KPEKCPVIGCEKTYKNQNGLKYHRLHGHQNQK	Zinc Finger (173 - 198)
JAZF1	171	KPEACPVPGCKKRYKNVNGIKYHAKNGHRTQI	
SFP1	655	KPYRCEVCGKRYKNLNGLKYH	Zinc Finger (208 - 230) (degenerate)
JAZF1	204	KPFKCR-CGKSYKTAQGLRHH	

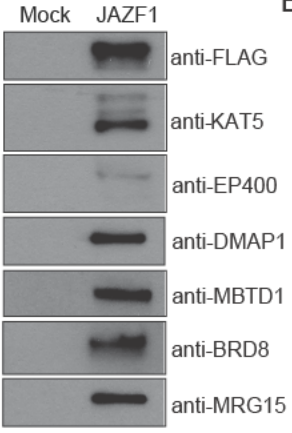
B



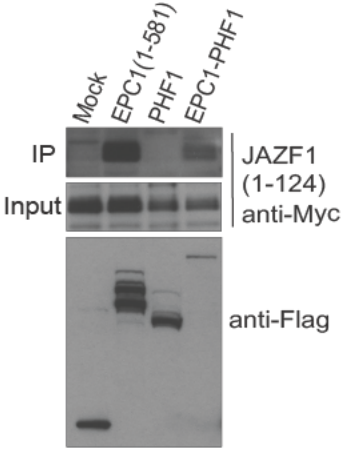
C



D



E



F

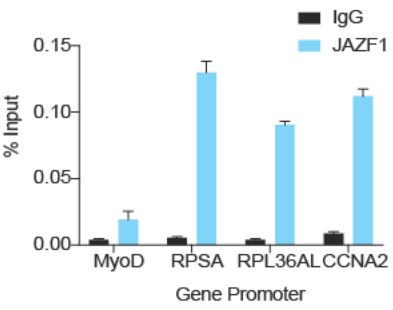
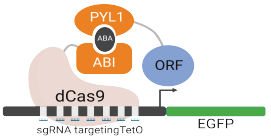
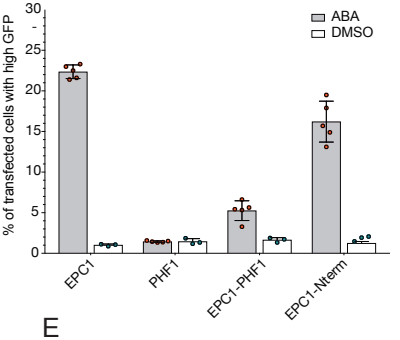


Figure 7

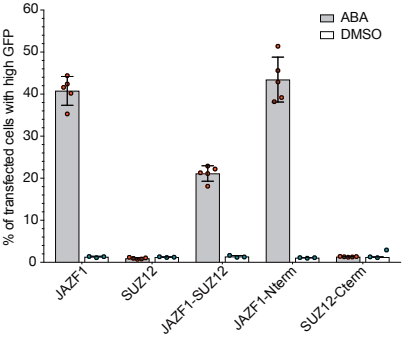
A



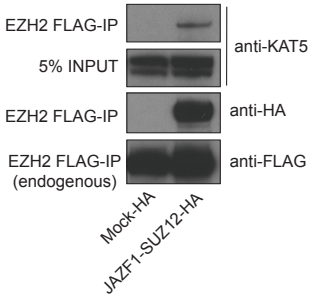
B



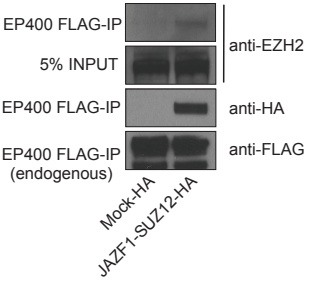
C



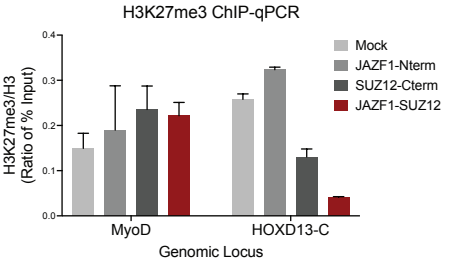
D



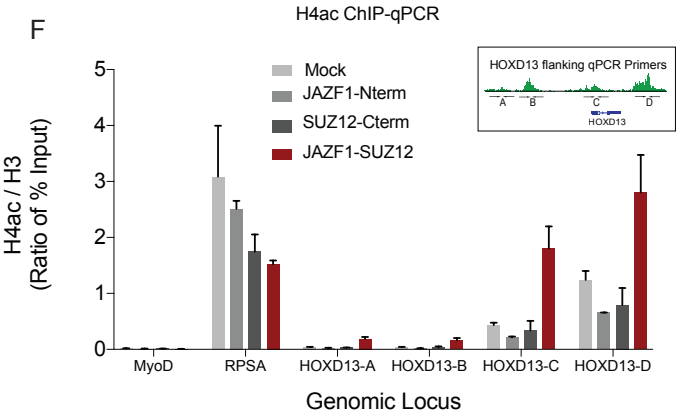
E



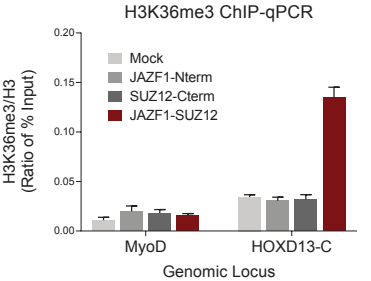
G



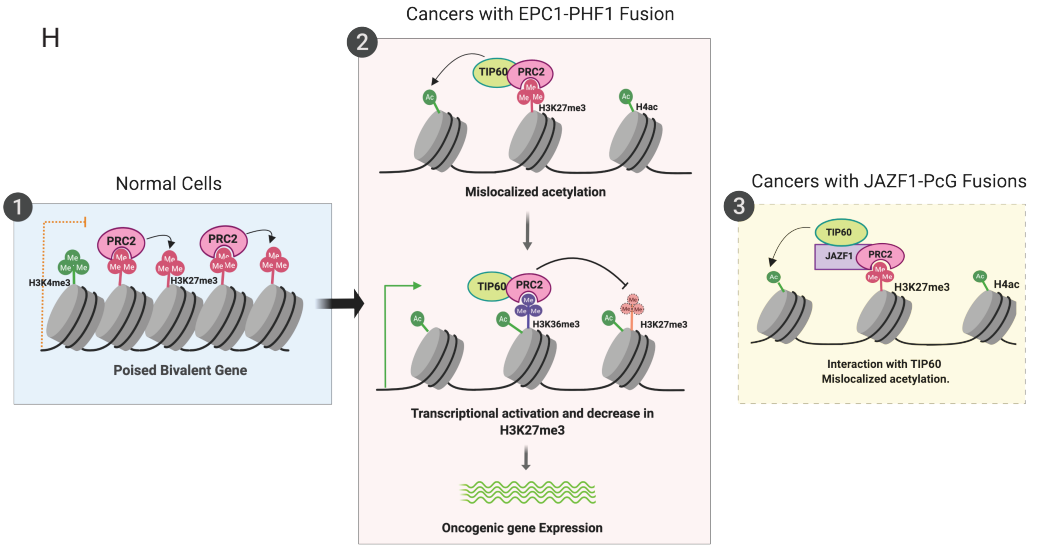
F



H



H



STAR Methods:

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jacques Côté (Jacques.Cote@crhdq.ulaval.ca)

EXPERIMENTAL MODEL AND SUBJECT DETAILS:

Cell culture:

K562 cells were obtained from the ATCC and maintained at 37°C under 5% CO2 in RPMI medium supplemented with 10% fetal bovine serum and GlutaMAX. 25 mM HEPES-NaOH (pH 7.4) was added during culture in Spinner flasks.

HEK293 cells were obtained from the ATCC and maintained at 37°C under 5% CO2 in DMEM medium supplemented with 10% fetal bovine serum.

METHOD DETAILS

Construction of Recombinant DNA:

RT-PCR product from Total RNA isolated from surgically removed endometrial stromal sarcoma tumor was a kind gift from Dr. Fransesca Micci (Micci et al., 2006). The cDNA containing the entire truncation of EPC1 and a fused portion of PHF1 up to base pair (bp) 139 , the included portion of PHF1 contained a BglI restriction site, which was then utilized to reconstruct the full length fusion gene using a clone of PHF1 obtained from Open Biosystems Products (Hunstville, AL). This construct was then subcloned into AAVS1- PGK Puro_2XHA-3XFlag (Jacquet et al., 2016) AAVS1_Puro_PGK1_3XFlag_Twin_Strep (Addgene plasmid #68375)(Dalvai et al., 2015) or pREV-CMV-3xFLAG (Avvakumov et al., 2012). We also sub-cloned full length PHF1 and EPC1, as well as a portion of EPC1 that included bp 1-1743, corresponding to the fragment of this gene found in the fusion. All constructs were verified by sequencing, and their expression was tested by transient transfection.

Human cDNA of *JAZF1* was purchased from GE Healthcare. Open reading frame (ORF) of *JAZF1* was subcloned into the pMZI vector (Zeghouf et al., 2004) containing a tandem affinity purification tag at its C terminus (Rigaut et al., 1999). ORF of *JAZF1*, *was* subcloned in either the p3X-CMV-Flag-14 (Sigma-Aldrich) and/or the pcDNA3Myc6 vector (a kind gift from Dr. Sylvain Meloche) as previously described (Julien et al., 2003).

Cell Line Generation:

EPC1-PHF1 and controls were targeted to the *AAVS1* safe harbour locus in K562 cells, using ZFN (Hockemeyer et al., 2009) as described previously in (Dalvai et al., 2015)

Briefly, Lipofectamine 2000 was used to transfect 1.25 million cells with 1 µg ZFN expression vector and 4 µg donor constructs according to manufacturer's instructions. Clonal selection was performed with 0.25 µg/ml puromycin for 10 days in methylcellulose-based semi-solid RPMI medium after 48hr of transfection.

Purification of Complexes:

Native complexes were purified as described in detail before in (Dalvai et al., 2015; Doyon and Cote, 2016). K562 cells expressing 3XFlag-2XHA or 3xFlag-2xStrep tag at the C-terminus of EPC1(1-581), PHF1, EPC1-PHF1 fusion, JAZF1 or mock controls, were expanded in spinner flasks (3 L cultures at 0.6–1.0 million cells/ml density). Nuclear cell extracts prepared as described in (Abmayr et al., 2006) and ultracentrifuged at 100,000 G for 1h. Extracts were precleared with 300 µl Sepharose CL-6B (Sigma). 250µl of anti-FLAG M2 affinity resin (Sigma) was added to the extract and incubated for 2 hr at 4°C. The resin was washed consecutively with buffer containing 300mM salt and 150mM salt. Complexes were eluted in two fractions with 200 µg/ml of 3xFLAG peptide (Sigma) and 150mM salt for 1hr in 4°C. These fractions were used for experiments or pooled to perform a second round of affinity purification with Streptactin resin (3xFlag-2xStrep Cell lines). Eluted fractions were mixed with 125µl Strep-Tactin Sepharose (IBA) affinity matrix for 2 hr at 4°C, and the beads were washed with buffer containing 150mM salt. Complexes were eluted in two fractions with buffer containing 4mM D-Biotin, 150mM salt. The purified complexes were loaded on NuPAGE 4–12% Bis-Tris gels (Invitrogen) and visualized by silver staining. Fractions were then analyzed by mass spectrometry.

Sample Digestion for Mass spectrometry Analysis

Purified complexes were run on preparative gels (Bolt 12% Bis-Tris, Invitrogen) over 1 cm, to stack all proteins into 1 band, stained with Sypro Ruby (Bio-Rad) and excised under UV lamp. The gel slice was reduced, alkylated and digested based on the original protocol (Wilm et al., 1996). Briefly, proteins were reduced with 10 mM DTT at 56°C for 15 minutes and alkylated with 100 mM iodoacetamide for 15 minutes. Trypsin digestion was performed using modified porcine trypsin (Sequencing grade, Sigma) at 37°C for 16 h. Digestion products were extracted sequentially using 25 mM ammonium bicarbonate, then 50% acetonitrile with 5% formic acid and finally 100% acetonitrile. The recovered extracts were pooled, vacuum centrifuge dried and stored at -80°C until MS analysis.

Proteins identification by mass spectrometry – TripleTOF 5600

Each sample (5 µL) was directly loaded at 400 nL/min onto an equilibrated HPLC column. The peptides were eluted from the column over a 90 min gradient generated by a NanoLC-Ultra 1D plus (Eksigent, Dublin CA) nano-pump and analysed on a TripleTOF 5600 instrument (AB SCIEX, Concord, Ontario, Canada). The gradient was delivered at 200 nL/min starting from 2%

acetonitrile with 0.1% formic acid to 35% acetonitrile with 0.1% formic acid over 90 min followed by a 15 min clean-up at 80% acetonitrile with 0.1% formic acid, and a 15 min equilibration period back to 2% acetonitrile with 0.1% formic acid, for a total of 120 min. To minimize carryover between each sample, the analytical column was washed for 3 h by running an alternating sawtooth gradient from 35% acetonitrile with 0.1% formic acid to 80% acetonitrile with 0.1% formic acid, holding each gradient concentration for 5 min. Analytical column and instrument performance were verified after each sample by loading 30 fmol bovine serum albumin (BSA) tryptic peptide standard (Michrom Bioresources Inc. Fremont, CA) with 60 fmol α-casein tryptic digest and running a short 30 min gradient. TOF MS calibration was performed on BSA reference ions before running the next sample to adjust for mass drift and verify peak intensity. The instrument method was set to data dependent acquisition (DDA) mode, which consisted of one 250 milliseconds (ms) MS1 TOF survey scan from 400–1300 Da followed by 20 100 ms MS2 candidate ion scans from 100–2000 Da in high sensitivity mode. Only ions with a charge of 2+ to 4+ that exceeded a threshold of 200 cps were selected for MS2, and former precursors were excluded for 10 s after one occurrence.

Proteins identification by mass spectrometry – Orbitrap Fusion

Peptide samples were separated by online reversed-phase (RP) nanoscale capillary liquid chromatography (nanoLC) and analyzed by electrospray mass spectrometry (ESI MS/MS). The experiments were performed with a Dionex UltiMate 3000 nanoRSLC chromatography system (Thermo Fisher Scientific) connected to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source. Peptides were trapped at 20 µl/min in loading solvent (2% acetonitrile, 0.05% TFA) on a 5mm x 300 µm C18 pepmap cartridge pre-column (Thermo Fisher Scientific) during 5 minutes. Then, the pre-column was switch online with a self-made 50 cm x 75 um internal diameter separation column packed with ReproSil-Pur C18-AQ 3-µm resin (Dr. Maisch HPLC) and the peptides were eluted with a linear gradient from 5-40% solvent B (A: 0,1% formic acid, B: 80% acetonitrile, 0.1% formic acid) in 90 minutes, at 300 nL/min. Mass spectra were acquired using a data dependent acquisition mode using Thermo XCalibur software version 3.0.63. Full scan mass spectra (350 to 1800m/z) were acquired in the orbitrap using an AGC target of 4e5, a maximum injection time of 50 ms and a resolution of 120 000. Internal calibration using lock mass on the m/z 445.12003 siloxane ion was used. Each MS scan was followed by acquisition of fragmentation spectra of the most intense ions for a total cycle time of 3 seconds (top speed mode). The selected ions were isolated using the quadrupole analyzer in a window of 1.6 m/z and fragmented by Higher energy Collision-induced Dissociation (HCD) with 35% of collision energy. The resulting fragments were detected by the linear ion trap in rapid scan rate with an AGC target of 1E4 and a maximum injection time of 50ms. Dynamic exclusion of previously fragmented peptides was set for a period of 20 sec and a tolerance of 10 ppm.

Multiple cell compartment affinity purification coupled to tandem mass spectrometry (MCC-AP-MS/MS)

Inducible EcR-293 stable cell line containing pMZI empty vector as control, pMZI-JAZF1 were generated as previously described (Jeronimo et al., 2007). Protein expression was induced with 3 µM ponasterone A (Life Technologies) and cells were harvested after 48h. Protein-protein interactions of JAZF1 were identified using two biological replicate runs for each bait of the multiple cell compartment affinity purification coupled to tandem mass spectrometry (MCC-AP-MS/MS) procedure, as recently described (Lavallee-Adam et al., 2013), (Cloutier et al., 2014)). The eluates were precipitated with trichloroacetic acid (TCA) and stored at -80 °C until liquid chromatography (LC) MS/MS analysis.

Affinity Purification followed by Immunoblotting

Figure 2D, 2E and 6G: K562 cells with endogenously tagged (3XFLAG-2XStrep) EP400 or EZH2 (from (Dalvai et al., 2015)) were used to stably express EPC1-PHF1 or JAZF1-SUZ12 3X-HA tagged constructs from the *AAVS1* safe harbor locus. Selected clones were expanded to get 6-8 million cells. Cells were collected and washed twice in 1X PBS, lysed for 30 min in 2X volume Lysis Buffer (450mM NaCl, 10% Glycerol, 50mM Tris-HCl pH8, 1%Triton-X 100, 2mM MgCl2, 0.1mMZnCl2, 2mM EDTA, 1mM DTT, Protease inhibitors), same volume of lysis buffer without salt was added to get a final salt concentration of 225mM. Fractions were centrifuged to prepare whole cell extracts. The extracts were incubated with FlagM2 agarose resin (Sigma) for 4hr at 4°C. The resin was centrifuged, washed in Lysis buffer with 225mM salt and Eluted with 3X Flag peptide (Sigma). The eluted fraction was loaded on 4-15% gradient gels with Input and immunoblotted with anti-HA, anti-KAT5, anti-DMAP1, anti-SUZ12 or anti-EZH2 antibody as appropriate.

Figure 6F, S6: EPC1-PHF1-Flag tagged K562 Cells and controls were grown in 100mm dishes, transiently transfected with JAZF1-full length or JAZF1 (1-281) Myc tagged constructs. Cells were collected after 48 hours and affinity purified as described above. The eluted fraction was loaded on 4-15% gradient gels with Input and immunoblotted with anti-Myc antibody.

Chromatin Immunoprecipitation

FLAG ChIPs were performed according to the protocol described in (Jacquet et al., 2016). 1mg of cross-linked chromatin from K562 cells was incubated with 10µg of anti-FLAG antibody (Sigma, M2) pre-bound on 300 µl of Dynabeads Protein G (Invitrogen) overnight at 4°C. The beads were washed extensively and eluted in 0.1% SDS, 0.1M NaHCO3. Crosslink was reversed with 0.2M NaCl and incubation overnight at 65°C. Samples were treated with RNase and Proteinase K for 2h and DNA was recovered by phenol chloroform and ethanol precipitation.

Histone Mark ChIPs were performed according to the protocol described in (Lalonde et al., 2013). Briefly 200 µg of chromatin was immunoprecipitated with 1-3ug of histone mark antibodies, overnight at 4 and incubated with 100ul of Dynabeads Protein-A 4h at 4°C. The beads were washed extensively and eluted in 0.1% SDS, 0.1M NaHCO3. Crosslink was reversed with 0.2M NaCl and incubation overnight at 65C. Samples were treated with RNase and Proteinase K for 2h and DNA recovered by phenol chloroform and ethanol precipitation.
Libraries for sequencing were prepared as described (Jacquet et al., 2016). Samples were sequenced by 50 bp single reads on HiSeq 2000 platform (Illumina).
Quantitative real-time PCRs were performed on a LightCycler 480 (Roche) with SYBR Green I (Roche) to confirm the specific enrichment at defined loci. The error bars represent standard errors based on two independent experiments. Quantitative real-time PCR primers used are listed in Supplementary Table.

HAT and HMT Assays

Fractions of purified complexes were assayed for enzymatic activity on short oligonucleosomes and free histones isolated from HeLa S3 cells as described previously (Doyon et al., 2004; Musselman et al., 2012; Utley et al., 1996).

Histone Acetyltransferase (HAT) Assay: *In vitro* histone acetylation assay 0.5µg of the indicated substrates in a 15µl reaction containing 50mM Tris-HCl pH8.0, 10% glycerol, 1mM EDTA, 1mM DTT, 1mM PMSF, 10mM sodium butyrate and 0.25 µCi/µl (4.9 Ci/mmol) of [3H]-labeled Acetyl-CoA (Perkin Elmer) or 0.15mM unlabeled Acetyl-CoA (Sigma). Samples were either spotted on P81 membrane (GE Healthcare) for count or analyzed on 18% SDS-PAGE. For radioactive gel assays, Coomassie staining was followed by treatment with Enhance® (Perkin Elmer) and fluorography, or by western blot analysis for non-radioactive assays. In results from liquid counting, error bars represent standard deviation of independent reactions.

Histone Methyltransferase (HMT) Assay : *In vitro* histone acetylation assay 0.5µg of the indicated substrates in a 15µl reaction containing 20mM Tris-HCl pH8.0, 5% glycerol, 0.1mM EDTA, 1mM DTT, 1mM PMSF, 100mM KCl and 1 mCi of [3H]AdoMet (S-adenosyl-L-[methyl-3H] methionine, 80 Ci/mmol; SAM) (Perkin Elmer) or 0.15mM unlabeled S- adenosyl methionine for 45 min at 30°C. Samples were either spotted on P81 membrane (GE Healthcare) for Scintillation count or analyzed on 15% SDS-PAGE. For radioactive gel assays, Coomassie staining was followed by treatment with Enhance® (Perkin Elmer) and fluorography, or by western blot analysis for non-radioactive assays. In results from liquid counting, error bars represent standard deviation of independent reactions.

Recruitment-activator assay

The reporter HEK293T cell line with a TRE3G-EGFP reporter was a kind gift from Stanley Qi (Stanford University). The reporter cell line was transduced with a lentivirus expressing ABI-dCas9 followed by two rounds of selection under 6µg/ml blasticidin. These cells were then transduced with a lentivirus expressing EBFP2 and a gRNA targeting seven tetO repeats in the TRE3G promoter (gRNA sequence GTACGTTCTCTATCACTGATA). Single-cell derived clonal cell lines were generated and a clone showing robust EGFP induction by a strong transcriptional activator VPR was selected for downstream assays. 96-well plates were seeded with $3x10^4$ cells per well one day prior to transfection. 150 ng of each construct was transfected using polyethylemine (PEI). Transfected cells were induced 24 hours after transfection by treatment with 100 µM abscisic acid. 48 hours after induction, cells were dissociated and resuspended in flow buffer using a liquid handing robot and analyzed by LSRFortessa (BD). Flow cytometry data was analyzed using FlowJo by gating for positive gRNA (EBFP2), then further for construct (TagRFP) expression. At least 25,000 cells were analyzed for each replicate.

Microarray:

RNA samples were extracted using TRIzol reagent (Invitrogen), following manufacturer's instructions. Duplicate RNA samples from EPC1-PHF1 expressing and control K562 cell lines were compared. We performed gene expression microarray experiments using the Human Illumina HumanHT-12_V4 platform.

Data Analysis:

ChIP-seq

Reads from ChIP-seq experiments were obtained from the Genome Innovation Center at McGill University. All the single-end reads were mapped on the human genome build hg18 using bowtie version 0.12.8. The read alignments in sam format were then converted to bam files using samtools (Langmead et al., 2009). Samtools was then successively utilized to remove duplicated reads, sort, and index bam files. Peak calling was performed using MACS version 1.4.0 by comparing EPC (1-581), PHF1 and EPC-PHF1 ChIP-seq signals to the control signal (Zhang et al., 2008). We identified intersection between two sets of peaks in hg18 coordinates using bedtools "intersect" function (Quinlan and Hall, 2010). We identified significant changes in H4ac and H3K27me3 ChIP-seq read counts at EPC-PHF1 bound regions using DESeq (Anders and Huber, 2010). Custom analyses specifically tailored to answer questions presented in this study were implemented in R version 2.14.0. For all the analyses using bivalent chromatin regions in K562 cells, we defined bivalent chromatin regions as the overlap between H3K4me3 and H3K27me3 significant ChIP-seq peaks (adjusted Bonferroni P-value < 0.05) identified in K562 cells by the ENCODE project (Consortium, 2012) and available from the UCSC genome

browser website (Kent et al., 2002). We used the UCSC hg18 gene definition for all the analyses that necessitate transcription start and end coordinates in genomic context.

Microarray

We first log2 transformed and quantile normalized the data before using them for the analyses presented in the paper. All the data were pre-processed using the lumi Bioconductor package (Du et al., 2008).

Mass Spectrometry

Mass spectrometry data were stored, searched and analyzed using the ProHits laboratory information management system (LIMS) platform (Liu et al., 2016). Within ProHits, AB SCIEXWIFF files were first converted to an MGF formatusing WIFF2MGF converter and to an mzML format using ProteoWizard (v3.0.4468) (Kessner et al., 2008) and the AB SCIEX MS Data Converter (V1.3 beta). The mzML files were searched using Mascot (v2.3.02). The spectra were searched with the RefSeq database (version 57, January 30th, 2013) acquired from NCBI against a total of 72,482 human and adenovirus sequences supplemented with common contaminants from the Max Planck Institute (http://141.61.102.106:8080/share.cgi?ssid=0f2gfuB) and the Global Proteome Machine (GPM; http://www.thegpm.org/crap/index.html). The database parameters were set to search for tryptic cleavages, allowing up to two missed cleavage sites per peptide with a mass tolerance of 40 ppm for precursors with charges of 2+ to 4+ and a tolerance of +/- 0.15 amu for fragment ions. Deamidated asparagine and glutamine and oxidized methionine were allowed as variable modifications.

Multiple cell compartment affinity purification coupled to tandem mass spectrometry (MCC-AP-MS/MS)

Protein were solubilized, digested with trypsin and analysed by tandem mass spectrometry as previously described (Cloutier et al., 2014). Protein database searching was performed with Mascot 2.2 (Matrix Science) against the human NCBInr protein database. Reliability assessment of protein-protein interactions identified with MCC-AP-MS/MS was performed using our previously published software package Decontaminator (Lavallee-Adam et al., 2011). pMZI empty vector controls were used to train Decontaminator's Bayesian inference algorithm. Decontaminator assigned a False Discovery Rate (FDR) to each protein-protein interaction in our dataset using a leave-one-out procedure. Interactions with a FDR below 20% in both replicate runs of MCC-AP-MS/MS and interactions observed in only one replicate run with a FDR < 10% were reported.

Hi-C data alignment

Alignment of the ChIP sequencing data with previously published Hi-C data in K562 cells (Rao et al., 2014) was performed by converting the ChIP sequencing alignment from hg18 to hg19 using CrossMap Tool (Zhao et al., 2014). Hi-C data and ChIP sequencing data was then visualized using PyGenome Tracks tool V2.1 (Ramirez et al., 2018). The CTCF ChIP-sequencing data used is from GEO: GSM733719 (Broad Institute /ENCODE group). The Hi-C data (Heat map, domains and loops) was downloaded from the datasets provided at Chorogenome (http://chorogenome.ie-freiburg.mpg.de/data_sources.html)

DATA AND SOFTWARE AVAILABILITY

The ChIP-seq data and expression analysis reported in this study are available at the GEO repository under accession number GSE162544.

All Mass Spectrometry files generated as part of this study were deposited at MassIVE (http://massive.ucsd.edu). The MassIVE ID is MSV000083618 and MSV000086476. The MassIVE FTP download links are ftp://massive.ucsd.edu/MSV000083618 and ftp://massive.ucsd.edu/MSV000086476. The password for download prior to final acceptance is “fusion”.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1 (Related to Figure 1-2). EPC1-PHF1 fusion protein associates with TIP60 as well as PRC2 complexes and enhances cell proliferation.

(A) Indicated FLAG tagged constructs were transiently transfected in HEK293T and affinity purified by FLAG-resin after co-transfection with HA tagged KAT5 and ING3 and analyzed by western blotting. The left panel shows the normalized expression of transiently transfected FLAG tagged constructs. The right panel is the FLAG purification showing interaction with KAT5 and ING3.

(B) Colony formation assay in HEK293T cells. EPC1-PHF1 transfected cells show higher colony counts compared to controls. Upper panel shows quantification (n=2, error bars represent range of the values).

(C) FLAG affinity purification from stably transfected HEK293T cell lines (over-expression). The left panel shows relative expression levels of the indicated cell lines. The right panel is the western blot analysis of FLAG affinity purification showing interaction of EPC1-PHF1 fusion protein with endogenous TIP60 complex subunits (TRRAP, EP400, BRD8, KAT5, ING3, Baf53a, DMAP1) and EZH2.

Supplementary Figure 2 (Related to Figure 2). HAT and HMT activity of the mega-complex and the effect of EPC1-PHF1 on global H3K27me3 levels.

(A) Mass Spectrometry analysis of the affinity purified EPC1-PHF1 complex. Table with total spectral counts of peptides enriched in the affinity purification. Peptides highlighted in light-

grey are NuA4/TIP60 complex components, peptides highlighted in darker-grey are PRC2.1 complex subunits. Replicate of Figure 2A-B using the EPC1-PHF1-3FLAG-HA-2A-puro cell line.

(B) Liquid HAT assay: Quantification of the Histone Acetyltransferase (HAT) assay in Figure 2F measuring the incorporation of ^{3}H-Acetyl moiety on chromatin substrates by scintillation counting (Counts Per Minute/CPM, error bars are based on technical replicates). Reaction with purified EPC1-PHF1 complex (Figure 2A) shows elevated CPM, indicating HAT activity. EPC1 (1-581) is the positive control.

(C) Liquid HMT assay: Quantification of the Histone Methyltransferase (HMT) assay in Figure 2F measuring incorporation of ^{3}H-Methyl moeity on chromatin substrates by scintillation counting (Counts Per Minute/CPM, error bars are based on technical replicates). Reaction with purified EPC1-PHF1 complex (Figure 2A) shows elevated CPM, indicating HMT activity. PHF1 fraction (PRC2) is the positive control.

(D) Immuno-blotting of acid extracted histones from EPC1-PHF1 and control cell lines (K562) with H3K27me3 antibody. H4 antibody is used as loading control for histones. Global level of H3K27me3 does not change in K562 with the expression of the EPC1-PHF1 fusion protein.

Supplementary Figure 3 (Related to Figure 3): Size of EPC1-PHF1-bound regions in the genome and changes in chromatin marks induced by EPC1-PHF1 correlate with effects on transcription levels.

(A) Average sizes (bp) of FLAG ChIP-seq peaks of EPC1-PHF1, EPC1(1-581) and PHF1 (determined in relation to Mock) shown as box plot. EPC1-PHF1 peaks show similarity to PHF1 rather than EPC1(1-581).

(B, C) Integrative analysis of ChIP-seq and Microarray: (B) Genes with decreased H4 acetylation show decreased gene expression; (C) Increased H3K27me3 occupancy and gene expression are anti-correlated.

Supplementary Figure 4 (Related to Figure 4): EPC1-PHF1 induces changes in chromatin landscape at the *HOXC* gene cluster.

(A) Representative ChIP-seq tracks at the *HOXC* cluster, highlighted region shows H4 acetylation mislocalization and reduction in H3K27me3 in EPC1-PHF1 expressing K562 cell line. The Y axis represents normalized reads per million total reads (rpm).

(B) Gene expression changes at the *HOXC* cluster in K562 cells expressing EPC1-PHF1 compared to cells expressing tag only (Mock).

(C) Relative protein expression levels of EPC1(1-581), PHF1 and EPC1-PHF1 K562 cell lines used for ChIP-seq (in Figure 3, 4, 5). EPC1-PHF1 is less expressed than other constructs. Western blotting from whole cell extracts was done with anti-FLAG and Tubulin is used as loading control.

(D,E) H3K27me3 and H3K36me3 levels at control gene regions: (D) ChIP-qPCR at a H3K36me3 positive locus, *RPSA* gene body; (E) ChIP-qPCR at a H3K27me3 positive region, *HOXA9*. (Values are represented as ratio of %IP/input chromatin of H3K36me3 or H3K27me3 and H3) (n=2, error bars represent Range of the values).

Supplementary Figure 5 (Related to Figure 5): Spreading of the EPC1-PHF1 complex and its associated histone marks.

(A) Schematic representation of a possible spreading of the EPC1-PHF1 complex and its associated histone marks, based on the reported mechanism of H3K27me3 spreading in cis and far cis through long range chromatin contacts (Oksuz et al., 2018).

(B) Representative ChIP-seq tracks at a spreading site - *HOXC* gene cluster, highlighted regions show changes in chromatin marks compared to control.

C) Representative ChIP-seq tracks at a nucleation site - *EVX1* at *HOXA* gene cluster, highlighted regions show changes in chromatin marks compared to control.

(D) Representative ChIP-seq tracks at a nucleation site - *HOXB* gene cluster, highlighted regions show changes in chromatin marks compared to control.

(E) Representative ChIP-seq tracks at a spreading site - *CYP26B1*, proximal to *HOXD* gene cluster, highlighted regions show changes in chromatin marks compared to control.

(F) Representative ChIP-seq tracks at a nucleation site - *LHX2* (Lim Homeobox 2), highlighted regions show changes in chromatin marks compared to control.

Supplementary Figure 6 (Related to Figure 6): **JAZF1 interacts with the NuA4/TIP60 complex.**

(A) JAZF1 interacts with the TIP60 complex. Table of proteins (numbers represent total spectral counts) co-purified with JAZF1 in the chromatin fractions from HEK293 cells upon multiple cell compartment affinity purification coupled to tandem mass spectrometry analysis (MCC-AP-MS/MS) (Data from two biological replicates; Interactions with an FDR below 20% in both replicate runs and interactions observed in only one replicate run with an FDR < 10% are reported).

(B) Mass Spectrometry analysis of the JAZF1-3xFLAG-2xStrep affinity purified from K562 cells. Table with total spectral counts of peptides present in the affinity purifications shown in Figure 6C-D. Peptides highlighted in purple are NuA4/TIP60 complex components, peptides highlighted in red are other known JAZF1 interacting proteins, in green are histones.

(C) Immunoprecipitation of JAZF1 with the EPC1-PHF1 fusion. Myc-tagged JAZF1 construct was transiently expressed in FLAG-tagged K562 cell lines expressing Tag (Mock), EPC1(1-581), PHF1 or EPC1-PHF1) from the *AAVS1* locus. Immunoprecipitation was performed with Anti-FLAG-resin and analyzed by western blotting with Anti-Myc antibody. JAZF1 interacts with EPC1(1-581) and EPC1-PHF1. See Figure 6E for data with JAZF1(1-124) construct.

Supplementary Figure 7 (Related to Figure 7): The JAZF1-SUZ12 fusion mislocalises histone marks.

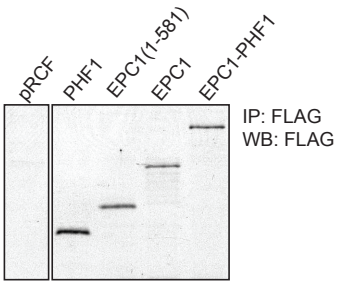
(A) Schematic representation of the JAZF1-SUZ12 fusion protein. The numbers indicated are amino acids, protein domains retained in the fusion are indicated.

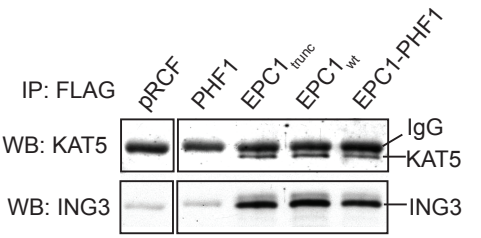
(B) Protein expression level of JAZF1-SUZ12, JAZF1-Nterm, SUZ12-C-term and empty tag (3xFLAG-2xStrep) from the *AAVS1* locus in K562 cells. Anti-FLAG western blotting on whole cell extracts and GAPDH is used as loading control.

(C, D) H3K27me3 and H3K36me3 levels in JAZF1-SUZ12 expressing cells at control loci. (C) ChIP-qPCR at a H3K36me3 positive locus, *RPSA* gene body; (D) ChIP-qPCR at a H3K27me3 positive region, *HOXA9*. (Values are represented as ratio of % of input chromatin of H3K36me3 or H3K27me3 and H3) (n=2, error bars represent Range of the values).

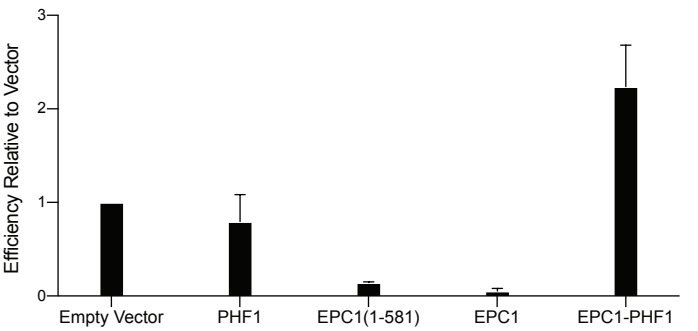
SUPPLEMENTARY FIGURE 1 (RELATED TO FIGURE 1)

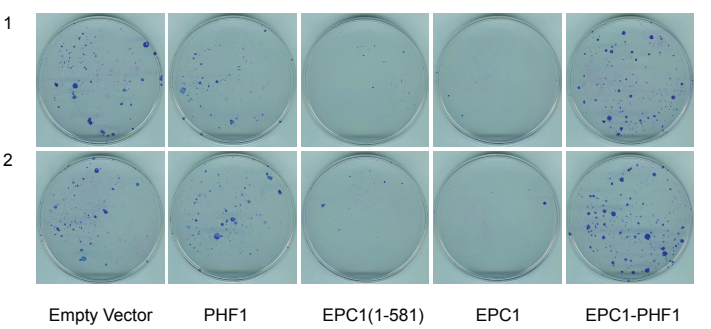
A



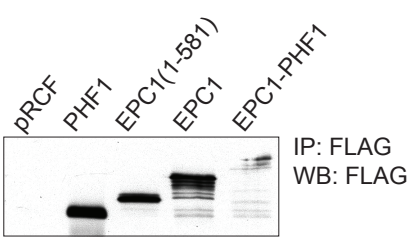


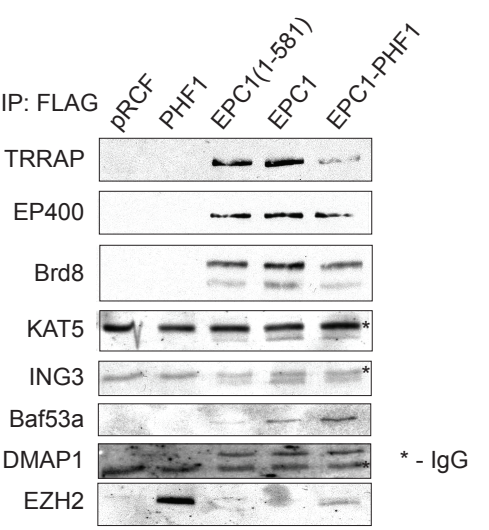
B





C



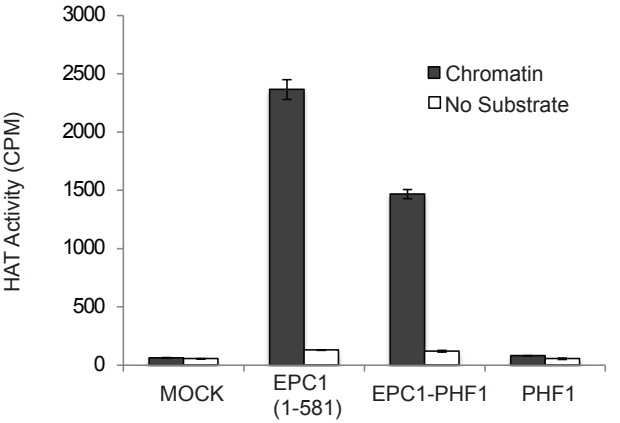


SUPPLEMENTARY FIGURE 2 (RELATED TO FIGURE 2)

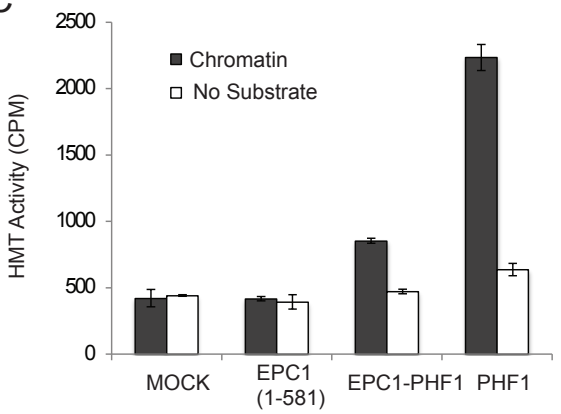
A

	Protein	Spectral Counts
TIP60	TRRAP	57
	RUVBL1	38
	EP400	28
	RUVBL2	25
	EPC1	9
	DMAP1	9
	BRD8	8
	BAF53a	8
	ACTB	7
	ING3	5
	KAT5	5
	VPS72	4
	GAS41/ YEATS4	4
	MRG15	3
PRC2	PHF1	4
	RBBP4	3
	EZH2	3

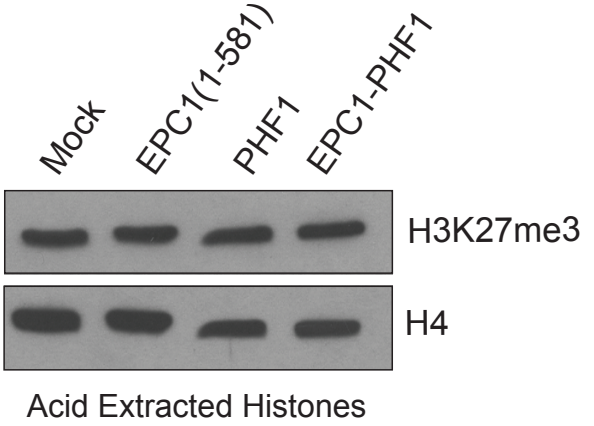
B



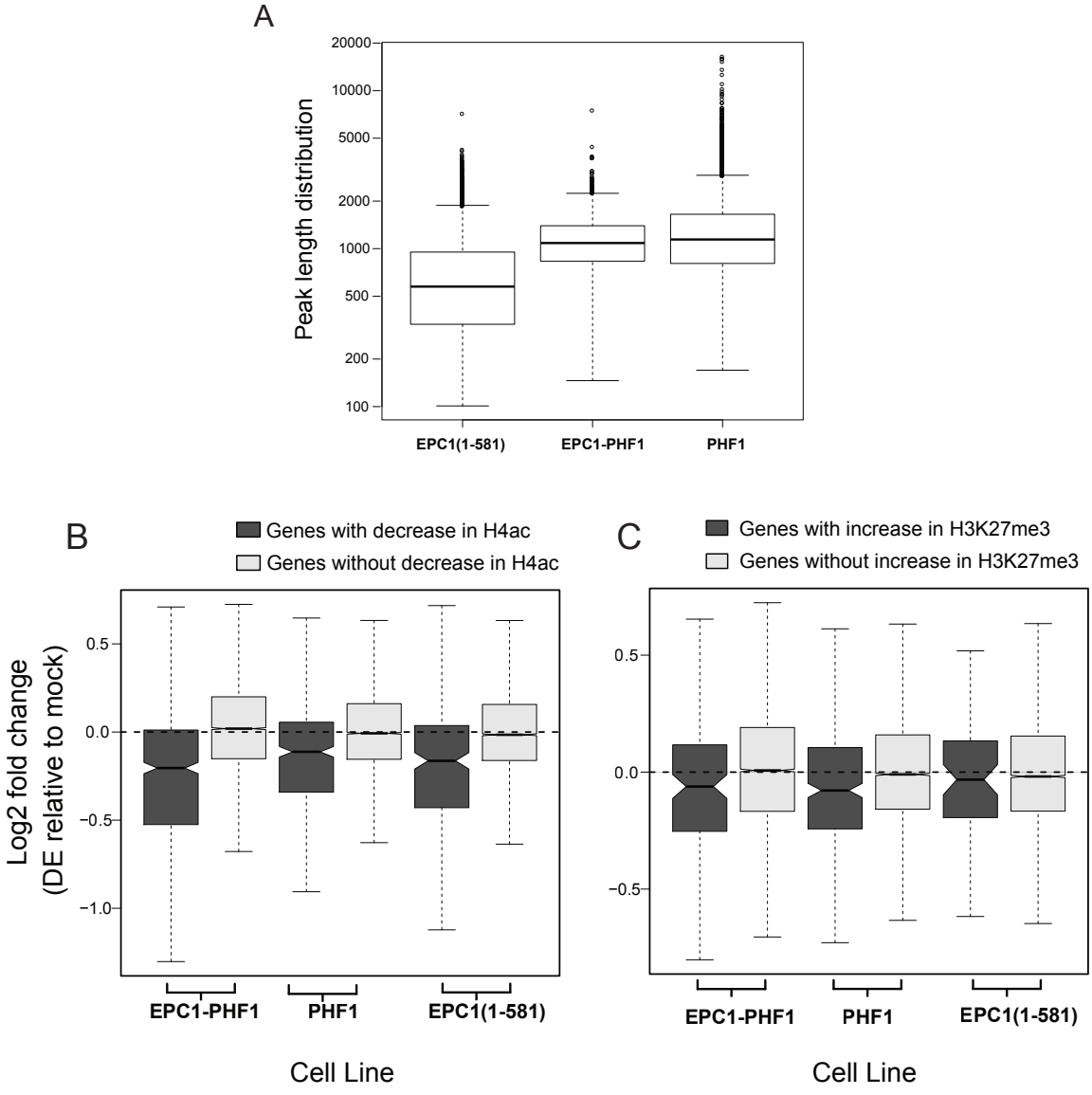
C



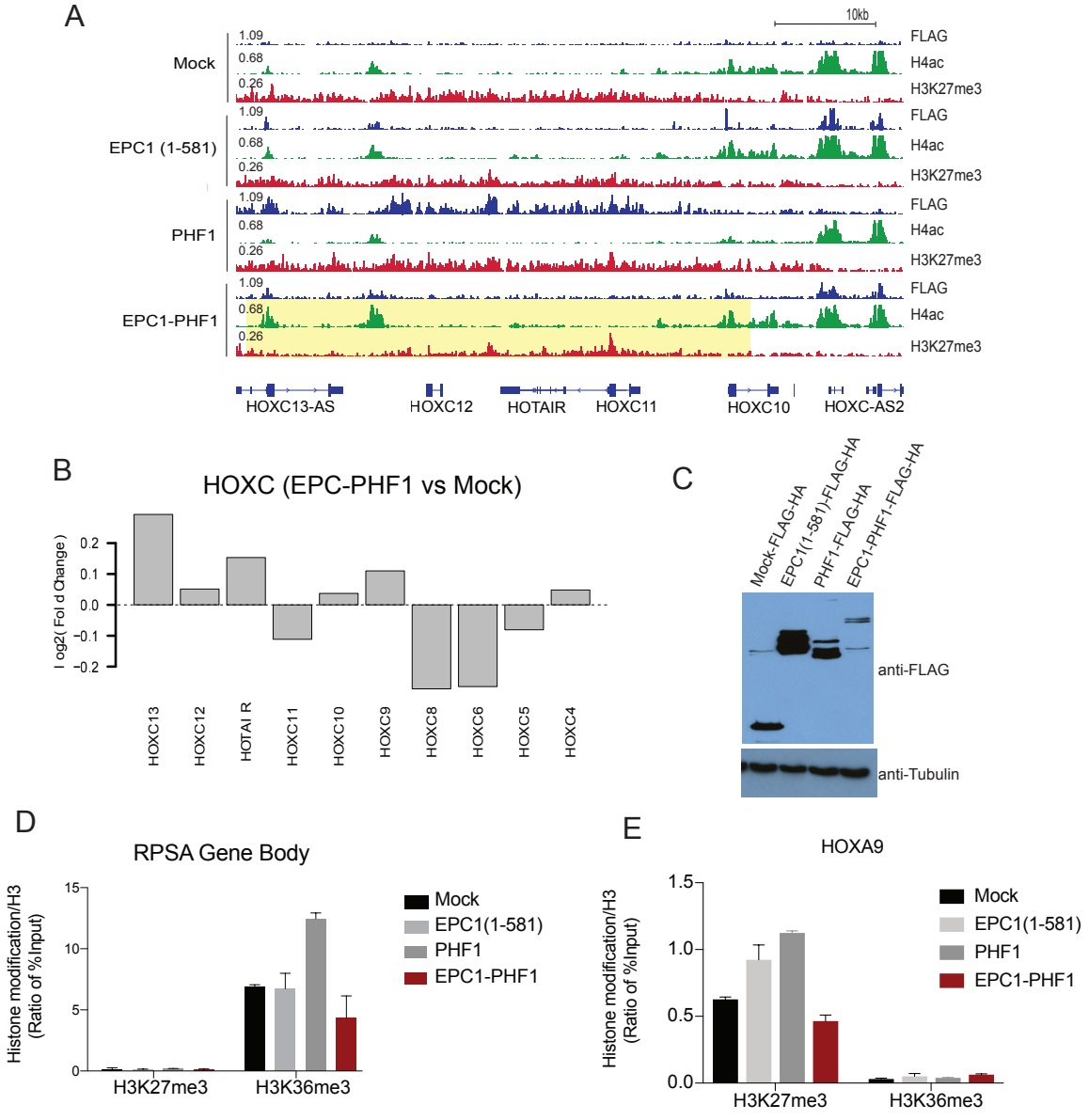
D



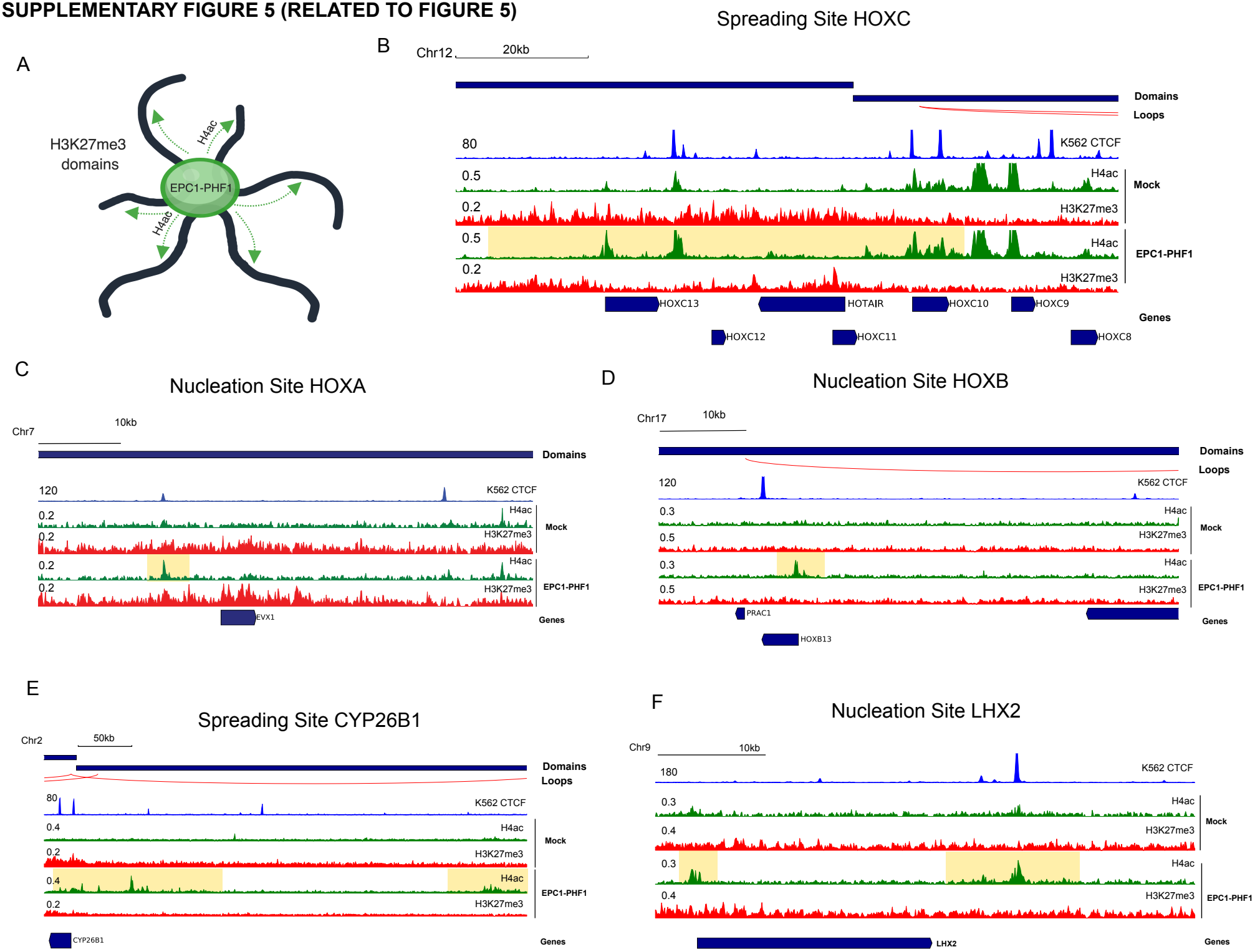
SUPPLEMENTARY FIGURE 3 (RELATED TO FIGURE 3)



SUPPLEMENTARY FIGURE 4 (RELATED TO FIGURE 4)



SUPPLEMENTARY FIGURE 5 (RELATED TO FIGURE 5)



SUPPLEMENTARY FIGURE 6 (RELATED TO FIGURE 6)

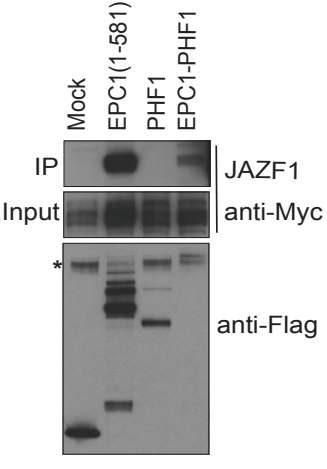
A

PROTEIN	SPECTRAL COUNTS
TRRAP	207
EP400	127
RUVBL2	48
RUVBL1	44
KAT5	17
EPC2	16
DMAP1	16
MBTD1	17
EPC1	11
JAZF1	10
BAF53a	24
ING3	3
BRD8	2

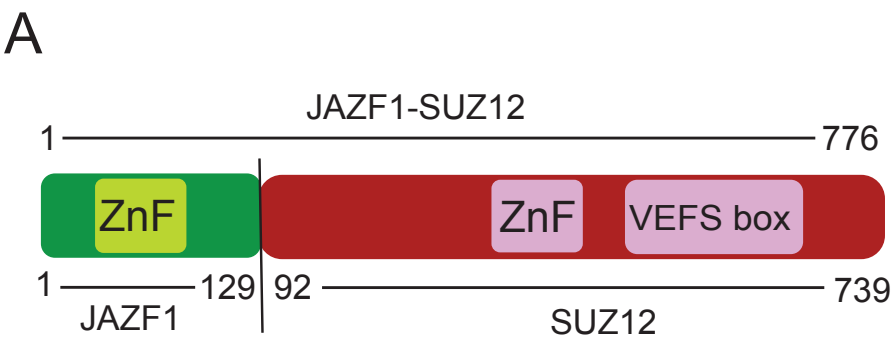
B

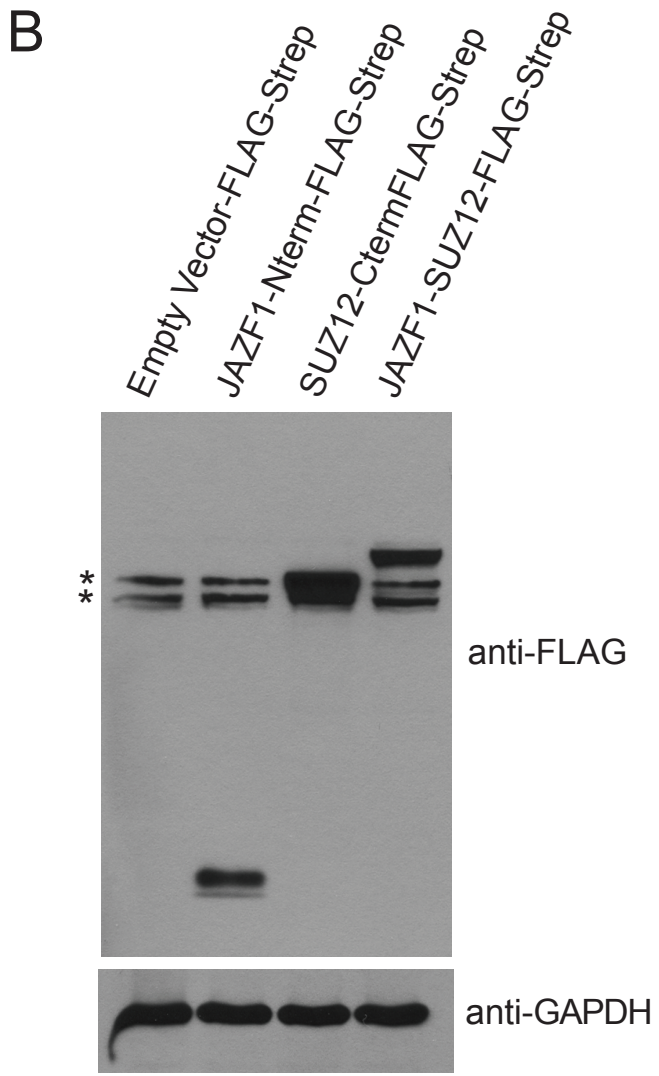
	PROTEIN	SPECTRAL COUNTS
TIP60 COMPLEX	TRRAP	1071
	EP400	595
	RUVBL1	352
	RUVBL2	346
	EPC2	231
	BRD8	186
	EPC1	155
	KAT5	143
	ING3	118
	DMAP1	117
	ACTB	107
	BAF53a	94
	MBTD1	94
	MRG15	70
	MRGX	56
	VPS72	54
	YEATS4	51
	MEAF6	43
	MRGBP	32
	JAZF1	33
	NR2C2	34
	NR2C1	15
	H2B	09
	H2A.Z	05

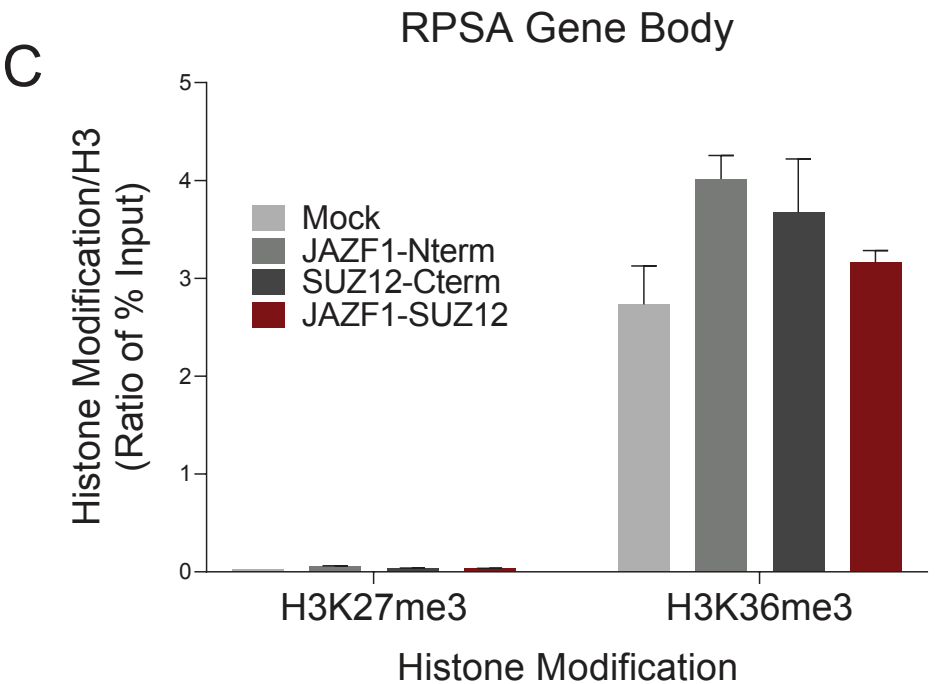
C

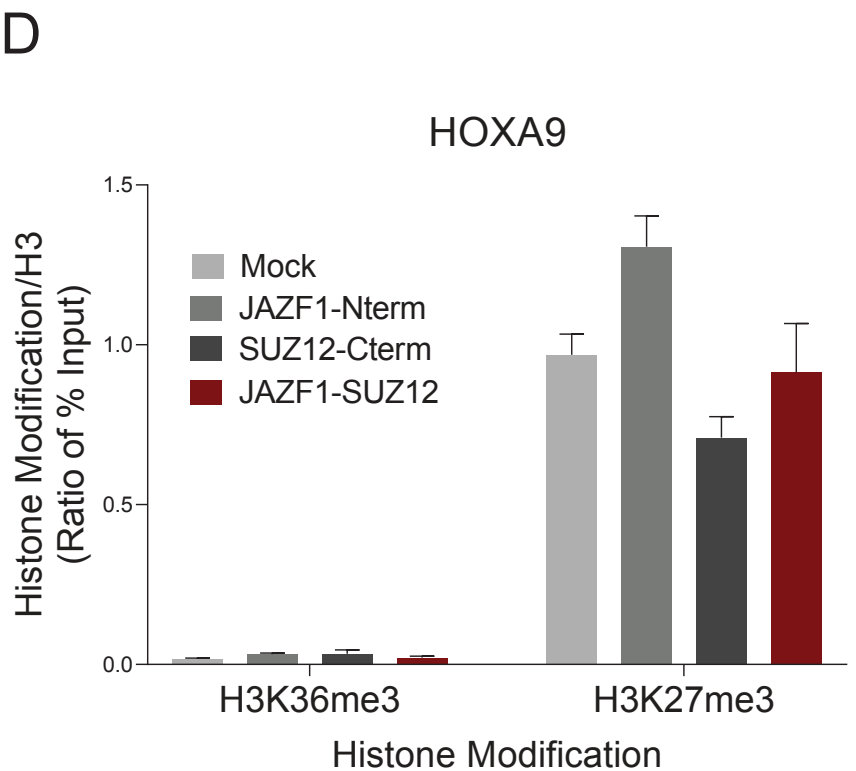


SUPPLEMENTARY FIGURE 7 (RELATED TO FIGURE 7)









REFERENCES

Abmayr, S.M., Yao, T., Parmely, T., and Workman, J.L. (2006). Preparation of nuclear and cytoplasmic extracts from mammalian cells. Current protocols in molecular biology *Chapter 12*, Unit 12 11.

Akbas, G.E., and Taylor, H.S. (2004). HOXC and HOXD gene expression in human endometrium: lack of redundancy with HOXA paralogs. Biology of reproduction *70*, 39-45.

Alekseyenko, A.A., Walsh, E.M., Zee, B.M., Pakozdi, T., Hsi, P., Lemieux, M.E., Dal Cin, P., Ince, T.A., Kharchenko, P.V., Kuroda, M.I.*, et al.* (2017). Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma. Proceedings of the National Academy of Sciences of the United States of America *114*, E4184-E4192.

Allen, A.J., Ali, S.M., Gowen, K., Elvin, J.A., and Pejovic, T. (2017). A recurrent endometrial stromal sarcoma harbors the novel fusion JAZF1-BCORL1. Gynecologic oncology reports *20*, 51-53.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology *11*, R106.

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science *340*, 1234167.

Antonescu, C.R., Sung, Y.S., Chen, C.L., Zhang, L., Chen, H.W., Singer, S., Agaram, N.P., Sboner, A., and Fletcher, C.D. (2014). Novel ZC3H7B-BCOR, MEAF6-PHF1, and EPC1-PHF1 fusions in ossifying fibromyxoid tumors--molecular characterization shows genetic overlap with endometrial stromal sarcoma. Genes, chromosomes & cancer *53*, 183-193.

Auger, A., Galarneau, L., Altaf, M., Nourani, A., Doyon, Y., Utley, R.T., Cronier, D., Allard, S., and Cote, J. (2008). Eaf1 is the platform for NuA4 molecular assembly that evolutionarily links chromatin acetylation to ATP-dependent exchange of histone H2A variants. Molecular and cellular biology *28*, 2257-2270.

Avvakumov, N., and Cote, J. (2007). The MYST family of histone acetyltransferases and their intimate links to cancer. Oncogene *26*, 5395-5407.

Avvakumov, N., Lalonde, M.E., Saksouk, N., Paquet, E., Glass, K.C., Landry, A.J., Doyon, Y., Cayrou, C., Robitaille, G.A., Richard, D.E.*, et al.* (2012). Conserved molecular interactions within the HBO1 acetyltransferase complexes regulate cell proliferation. Molecular and cellular biology *32*, 689-703.

Billon, P., and Cote, J. (2013). Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. Biochimica et biophysica acta *1819*, 290-302.

Bompadre, O., and Andrey, G. (2019). Chromatin topology in development and disease. Curr Opin Genet Dev *55*, 32-38.

Boudreault, A.A., Cronier, D., Selleck, W., Lacoste, N., Utley, R.T., Allard, S., Savard, J., Lane, W.S., Tan, S., and Cote, J. (2003). Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. Genes & development *17*, 1415-1428.

Brien, G.L., Stegmaier, K., and Armstrong, S.A. (2019). Targeting chromatin complexes in fusion protein-driven malignancies. Nature reviews. Cancer *19*, 255-269.

Brown, J.A., Bourke, E., Eriksson, L.A., and Kerin, M.J. (2016). Targeting cancer using KAT inhibitors to mimic lethal knockouts. Biochemical Society transactions *44*, 979-986.

Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D.J., Allis, C.D.*, et al.* (2013). An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Molecular cell *49*, 571-582.

Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2008). Role of hPHF1 in H3K27 methylation and Hox gene silencing. Molecular and cellular biology *28*, 1862-1872.
Chen, S., Jiao, L., Shubbar, M., Yang, X., and Liu, X. (2018). Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. Molecular cell *69*, 840-852 e845.
Cheng, X., Jobin-Robitaille, O., Billon, P., Buisson, R., Niu, H., Lacoste, N., Abshiru, N., Cote, V., Thibault, P., Kron, S.J.*, et al.* (2018). Phospho-dependent recruitment of the yeast NuA4 acetyltransferase complex by MRX at DNA breaks regulates RPA dynamics during resection. Proceedings of the National Academy of Sciences of the United States of America *115*, 10028-10033.
Chittuluru, J.R., Chaban, Y., Monnet-Saksouk, J., Carrozza, M.J., Sapountzi, V., Selleck, W., Huang, J., Utley, R.T., Cramet, M., Allard, S.*, et al.* (2011). Structure and nucleosome interaction of the yeast NuA4 and Piccolo-NuA4 histone acetyltransferase complexes. Nature structural & molecular biology *18*, 1196-1203.
Choi, J., Bachmann, A.L., Tauscher, K., Benda, C., Fierz, B., and Muller, J. (2017). DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. Nature structural & molecular biology *24*, 1039-1047.
Cloutier, P., Lavallee-Adam, M., Faubert, D., Blanchette, M., and Coulombe, B. (2014). Methylation of the DNA/RNA-binding protein Kin17 by METTL22 affects its association with chromatin. Journal of proteomics *100*, 115-124.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
Dalvai, M., Loehr, J., Jacquet, K., Huard, C.C., Roques, C., Herst, P., Cote, J., and Doyon, Y. (2015). A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. Cell reports *13*, 621-633.
Dewaele, B., Przybyl, J., Quattrone, A., Finalet Ferreiro, J., Vanspauwen, V., Geerdens, E., Gianfelici, V., Kalender, Z., Wozniak, A., Moerman, P.*, et al.* (2014). Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma. International journal of cancer *134*, 1112-1122.
Donaldson-Collier, M.C., Sungalee, S., Zufferey, M., Tavernari, D., Katanayeva, N., Battistello, E., Mina, M., Douglass, K.M., Rey, T., Raynaud, F.*, et al.* (2019). EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains. Nature genetics *51*, 517-528.
Doyon, Y., and Cote, J. (2004). The highly conserved and multifunctional NuA4 HAT complex. Curr Opin Genet Dev *14*, 147-154.
Doyon, Y., and Cote, J. (2016). Preparation and Analysis of Native Chromatin-Modifying Complexes. Methods in enzymology *573*, 303-318.
Doyon, Y., Selleck, W., Lane, W.S., Tan, S., and Cote, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. Molecular and cellular biology *24*, 1884-1896.
Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: a pipeline for processing Illumina microarray. Bioinformatics *24*, 1547-1548.
Ferreira, J., Felix, A., Lennerz, J.K., and Oliva, E. (2018). Recent advances in the histological and molecular classification of endometrial stromal neoplasms. Virchows Archiv : an international journal of pathology *473*, 665-678.
Finogenova, K., Bonnet, J., Poepsel, S., Schafer, I.B., Finkl, K., Schmid, K., Litz, C., Strauss, M., Benda, C., and Muller, J. (2020). Structural basis for PRC2 decoding of active histone methylation marks H3K36me2/3. eLife *9*.

Gao, Y., Xiong, X., Wong, S., Charles, E.J., Lim, W.A., and Qi, L.S. (2016). Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. Nature methods *13*, 1043-1049.
Gorrini, C., Squatrito, M., Luise, C., Syed, N., Perna, D., Wark, L., Martinato, F., Sardella, D., Verrecchia, A., Bennett, S.*, et al.* (2007). Tip60 is a haplo-insufficient tumour suppressor required for an oncogene-induced DNA damage response. Nature *448*, 1063-1067.
Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F.*, et al.* (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712-723.
Hoang, L., Chiang, S., and Lee, C.H. (2018). Endometrial stromal sarcomas and related neoplasms: new developments and diagnostic considerations. Pathology *50*, 162-177.
Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B.*, et al.* (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nature biotechnology *27*, 851-857.
Huang, J., and Tan, S. (2013). Piccolo NuA4-catalyzed acetylation of nucleosomal histones: critical roles of an Esa1 Tudor/chromo barrel loop and an Epl1 enhancer of polycomb A (EPcA) basic region. Molecular and cellular biology *33*, 159-169.
Hubner, J.M., Muller, T., Papageorgiou, D.N., Mauermann, M., Krijgsveld, J., Russell, R.B., Ellison, D.W., Pfister, S.M., Pajtler, K.W., and Kool, M. (2019). EZHIP / CXorf67 mimics K27M mutated oncohistones and functions as an intrinsic inhibitor of PRC2 function in aggressive posterior fossa ependymoma. Neuro-oncology.
Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K.*, et al.* (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell *162*, 425-440.
Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J.P., Roques, C., Pandita, R.K., Paquet, E., Herst, P., Gingras, A.C., Pandita, T.K.*, et al.* (2016). The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 through Direct H4K20me Binding and H2AK15 Acetylation. Molecular cell *62*, 409-421.
Jain, S.U., Do, T.J., Lund, P.J., Rashoff, A.Q., Diehl, K.L., Cieslik, M., Bajic, A., Juretic, N., Deshmukh, S., Venneti, S.*, et al.* (2019). PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism. Nature communications *10*, 2146.
Jani, K.S., Jain, S.U., Ge, E.J., Diehl, K.L., Lundgren, S.M., Muller, M.M., Lewis, P.W., and Muir, T.W. (2019). Histone H3 tail binds a unique sensing pocket in EZH2 to activate the PRC2 methyltransferase. Proceedings of the National Academy of Sciences of the United States of America *116*, 8295-8300.
Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J.*, et al.* (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Molecular cell *27*, 262-274.
Judes, G., Rifai, K., Ngollo, M., Daures, M., Bignon, Y.J., Penault-Llorca, F., and Bernard-Gallon, D. (2015). A bivalent role of TIP60 histone acetyl transferase in human cancer. Epigenomics *7*, 1351-1363.
Julien, C., Coulombe, P., and Meloche, S. (2003). Nuclear export of ERK3 by a CRM1-dependent mechanism regulates its inhibitory action on cell cycle progression. The Journal of biological chemistry *278*, 42615-42624.
Kasinath, V., Beck, C., Sauer, P., Poepsel, S., Kosmatka, J., Faini, M., Toso, D., Aebersold, R., and Nogales, E. (2021). JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications. Science *371*.
Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome research *12*, 996-1006.

Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics *24*, 2534-2536.
Kobiita, A., Godbersen, S., Araldi, E., Ghoshdastider, U., Schmid, M.W., Spinas, G., Moch, H., and Stoffel, M. (2020). The Diabetes Gene JAZF1 Is Essential for the Homeostatic Control of Ribosome Biogenesis and Function in Metabolic Stress. Cell reports *32*, 107846.
Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J.T., Sadreyev, R.I., Dekker, J., and Kingston, R.E. (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. Molecular cell *65*, 432-446 e435.
Lalonde, M.E., Avvakumov, N., Glass, K.C., Joncas, F.H., Saksouk, N., Holliday, M., Paquet, E., Yan, K., Tong, Q., Klein, B.J.*, et al.* (2013). Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. Genes & development *27*, 2009-2024.
Lalonde, M.E., Cheng, X., and Cote, J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. Genes & development *28*, 1029-1041.
Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.
Laugesen, A., Hojfeldt, J.W., and Helin, K. (2019). Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. Molecular cell *74*, 8-18.
Lavallee-Adam, M., Cloutier, P., Coulombe, B., and Blanchette, M. (2011). Modeling contaminants in AP-MS/MS experiments. Journal of proteome research *10*, 886-895.
Lavallee-Adam, M., Rousseau, J., Domecq, C., Bouchard, A., Forget, D., Faubert, D., Blanchette, M., and Coulombe, B. (2013). Discovery of cell compartment specific protein-protein interactions using affinity purification combined with tandem mass spectrometry. Journal of proteome research *12*, 272-281.
Lempiainen, H., Uotila, A., Urban, J., Dohnal, I., Ammerer, G., Loewith, R., and Shore, D. (2009). Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. Molecular cell *33*, 704-716.
Li, H., Liefke, R., Jiang, J., Kurland, J.V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D.J., Bulyk, M.L.*, et al.* (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. Nature *549*, 287-291.
Liao, Z.Z., Wang, Y.D., Qi, X.Y., and Xiao, X.H. (2019). JAZF1, a relevant metabolic regulator in type 2 diabetes. Diabetes/metabolism research and reviews *35*, e3148.
Liu, G., Knight, J.D., Zhang, J.P., Tsou, C.C., Wang, J., Lambert, J.P., Larsen, B., Tyers, M., Raught, B., Bandeira, N.*, et al.* (2016). Data Independent Acquisition analysis in ProHits 4.0. Journal of proteomics *149*, 64-68.
Loewith, R., and Hall, M.N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics *189*, 1177-1201.
Lonfat, N., and Duboule, D. (2015). Structure, function and evolution of topologically associating domains (TADs) at HOX loci. FEBS letters *589*, 2869-2876.
Loubiere, V., Martinez, A.M., and Cavalli, G. (2019). Cell Fate and Developmental Regulation Dynamics by Polycomb Proteins and 3D Genome Architecture. BioEssays : news and reviews in molecular, cellular and developmental biology *41*, e1800222.
Luo, Z., Rhie, S.K., and Farnham, P.J. (2019). The Enigmatic HOX Genes: Can We Crack Their Code? Cancers *11*.
Ma, X., Wang, J., Wang, J., Ma, C.X., Gao, X., Patriub, V., and Sklar, J.L. (2017). The JAZF1-SUZ12 fusion protein disrupts PRC2 complexes and impairs chromatin repression during human endometrial stromal tumorogenesis. Oncotarget *8*, 4062-4078.
Marion, R.M., Regev, A., Segal, E., Barash, Y., Koller, D., Friedman, N., and O'Shea, E.K. (2004). Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. Proceedings of the National Academy of Sciences of the United States of America *101*, 14315-14322.

Micci, F., Gorunova, L., Agostini, A., Johannessen, L.E., Brunetti, M., Davidson, B., Heim, S., and Panagopoulos, I. (2016). Cytogenetic and molecular profile of endometrial stromal sarcoma. Genes, chromosomes & cancer *55*, 834-846.
Micci, F., Panagopoulos, I., Bjerkehagen, B., and Heim, S. (2006). Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma. Cancer research *66*, 107-112.
Miyazaki, K., Dyson, M.T., Coon, V.J., Furukawa, Y., Yilmaz, B.D., Maruyama, T., and Bulun, S.E. (2018). Generation of Progesterone-Responsive Endometrial Stromal Fibroblasts from Human Induced Pluripotent Stem Cells: Role of the WNT/CTNNB1 Pathway. Stem cell reports *11*, 1136-1155.
Mohan, M., Lin, C., Guest, E., and Shilatifard, A. (2010). Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. Nature reviews. Cancer *10*, 721-728.
Musselman, C.A., Avvakumov, N., Watanabe, R., Abraham, C.G., Lalonde, M.E., Hong, Z., Allen, C., Roy, S., Nunez, J.K., Nickoloff, J.*, et al.* (2012). Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nature structural & molecular biology *19*, 1266-1272.
Nacev, B.A., Jones, K.B., Intlekofer, A.M., Yu, J.S.E., Allis, C.D., Tap, W.D., Ladanyi, M., and Nielsen, T.O. (2020). The epigenomics of sarcoma. Nature reviews. Cancer.
Nakajima, T., Fujino, S., Nakanishi, G., Kim, Y.S., and Jetten, A.M. (2004). TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. Nucleic acids research *32*, 4194-4204.
Oksuz, O., Narendra, V., Lee, C.H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P.P., Garcia, B.A.*, et al.* (2018). Capturing the Onset of PRC2-Mediated Repressive Domain Formation. Molecular cell *70*, 1149-1162 e1145.
Piunti, A., Smith, E.R., Morgan, M.A.J., Ugarenko, M., Khaltyan, N., Helmin, K.A., Ryan, C.A., Murray, D.C., Rickels, R.A., Yilmaz, B.D.*, et al.* (2019). CATACOMB: An endogenous inducible gene that antagonizes H3K27 methylation activity of Polycomb repressive complex 2 via an H3K27M-like mechanism. Science advances *5*, eaax2887.
Pradhan, S.K., Su, T., Yen, L., Jacquet, K., Huang, C., Cote, J., Kurdistani, S.K., and Carey, M.F. (2016). EP400 Deposits H3.3 into Promoters and Enhancers during Gene Activation. Molecular cell *61*, 27-38.
Przybyl, J., Kidzinski, L., Hastie, T., Debiec-Rychter, M., Nusse, R., and van de Rijn, M. (2018). Gene expression profiling of low-grade endometrial stromal sarcoma indicates fusion protein-mediated activation of the Wnt signaling pathway. Gynecologic oncology *149*, 388-393.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Ramirez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Gruning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nature communications *9*, 189.
Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.
Reid, J.L., Iyer, V.R., Brown, P.O., and Struhl, K. (2000). Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. Molecular cell *6*, 1297-1307.
Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature biotechnology *17*, 1030-1032.
Rodriguez-Carballo, E., Lopez-Delisle, L., Zhan, Y., Fabre, P.J., Beccari, L., El-Idrissi, I., Huynh, T.H.N., Ozadam, H., Dekker, J., and Duboule, D. (2017). The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. Genes & development *31*, 2264-2281.

Rossetto, D., Cramet, M., Wang, A.Y., Steunou, A.L., Lacoste, N., Schulze, J.M., Cote, V., Monnet-Saksouk, J., Piquet, S., Nourani, A.*, et al.* (2014). Eaf5/7/3 form a functionally independent NuA4 submodule linked to RNA polymerase II-coupled nucleosome recycling. The EMBO journal *33*, 1397-1415.
Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008). Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Molecular and cellular biology *28*, 2718-2731.
Schmitges, F.W., Prusty, A.B., Faty, M., Stutzer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S.*, et al.* (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Molecular cell *42*, 330-341.
Schneider, N., Fisher, C., and Thway, K. (2016). Ossifying fibromyxoid tumor: morphology, genetics, and differential diagnosis. Annals of diagnostic pathology *20*, 52-58.
Schoolmeester, J.K., Sukov, W.R., Maleszewski, J.J., Bedroske, P.P., Folpe, A.L., and Hodge, J.C. (2013). JAZF1 rearrangement in a mesenchymal tumor of nonendometrial stromal origin: report of an unusual ossifying sarcoma of the heart demonstrating JAZF1/PHF1 fusion. The American journal of surgical pathology *37*, 938-942.
Schuettengruber, B., Bourbon, H.M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell *171*, 34-57.
Selleck, W., Fortin, I., Sermwittayawong, D., Cote, J., and Tan, S. (2005). The Saccharomyces cerevisiae Piccolo NuA4 histone acetyltransferase complex requires the Enhancer of Polycomb A domain and chromodomain to acetylate nucleosomes. Molecular and cellular biology *25*, 5535-5542.
Setiaputra, D., Ahmad, S., Dalwadi, U., Steunou, A.L., Lu, S., Ross, J.D., Dong, M.Q., Cote, J., and Yip, C.K. (2018). Molecular Architecture of the Essential Yeast Histone Acetyltransferase Complex NuA4 Redefines Its Multimodularity. Molecular and cellular biology *38*.
Sheikh, B.N., and Akhtar, A. (2019). The many lives of KATs - detectors, integrators and modulators of the cellular environment. Nature reviews. Genetics *20*, 7-23.
Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38-55.
Shiota, H., Elya, J.E., Alekseyenko, A.A., Chou, P.M., Gorman, S.A., Barbash, O., Becht, K., Danga, K., Kuroda, M.I., Nardi, V.*, et al.* (2018). "Z4" Complex Member Fusions in NUT Carcinoma: Implications for a Novel Oncogenic Mechanism. Molecular cancer research : MCR *16*, 1826-1833.
Steunou, A.-L., Rossetto, D., and Côté, J. (2014). Regulating Chromatin by Histone Acetylation. In Fundamentals of Chromatin, J.L. Workman, and S.M. Abmayr, eds. (Springer New York), pp. 147-212.
Streubel, G., Watson, A., Jammula, S.G., Scelfo, A., Fitzpatrick, D.J., Oliviero, G., McCole, R., Conway, E., Glancy, E., Negri, G.L.*, et al.* (2018). The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. Molecular cell *70*, 371-379 e375.
Svoboda, L.K., Harris, A., Bailey, N.J., Schwentner, R., Tomazou, E., von Levetzow, C., Magnuson, B., Ljungman, M., Kovar, H., and Lawlor, E.R. (2014). Overexpression of HOX genes is prevalent in Ewing sarcoma and is associated with altered epigenetic regulation of developmental transcription programs. Epigenetics *9*, 1613-1625.
Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. Science advances *5*, eaaw1668.
Taylor, B.S., Barretina, J., Maki, R.G., Antonescu, C.R., Singer, S., and Ladanyi, M. (2011). Advances in sarcoma genomics and new therapeutic targets. Nature reviews. Cancer *11*, 541-557.

Utley, R.T., Owen-Hughes, T.A., Juan, L.J., Cote, J., Adams, C.C., and Workman, J.L. (1996). In vitro analysis of transcription factor binding to nucleosomes and nucleosome disruption/displacement. Methods in enzymology *274*, 276-291.
Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. Proceedings of the National Academy of Sciences of the United States of America *112*, 4672-4677.
von Heyking, K., Roth, L., Ertl, M., Schmidt, O., Calzada-Wack, J., Neff, F., Lawlor, E.R., Burdach, S., and Richter, G.H. (2016). The posterior HOXD locus: Its contribution to phenotype and malignancy of Ewing sarcoma. Oncotarget *7*, 41767-41780.
Voss, A.K., and Thomas, T. (2018). Histone Lysine and Genomic Targets of Histone Acetyltransferases in Mammals. BioEssays : news and reviews in molecular, cellular and developmental biology *40*, e1800078.
Wang, X., Ahmad, S., Zhang, Z., Cote, J., and Cai, G. (2018). Architecture of the Saccharomyces cerevisiae NuA4/TIP60 complex. Nature communications *9*, 1147.
Wilm, M., Neubauer, G., and Mann, M. (1996). Parent ion scans of unseparated peptide mixtures. Analytical chemistry *68*, 527-533.
Woo, C.J., Kharchenko, P.V., Daheron, L., Park, P.J., and Kingston, R.E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell *140*, 99-110.
Xu, P., Li, C., Chen, Z., Jiang, S., Fan, S., Wang, J., Dai, J., Zhu, P., and Chen, Z. (2016). The NuA4 Core Complex Acetylates Nucleosomal Histone H4 through a Double Recognition Mechanism. Molecular cell *63*, 965-975.
Youmans, D.T., Gooding, A.R., Dowell, R.D., and Cech, T.R. (2021). Competition between PRC2.1 and 2.2 subcomplexes regulates PRC2 chromatin occupancy in human stem cells. Molecular cell *81*, 488-501 e489.
Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J.F. (2004). Sequential Peptide Affinity (SPA) system for the identification of mammalian and bacterial protein complexes. Journal of proteome research *3*, 463-468.
Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology *20*, 259-266.
Zhang, H., Devoucoux, M., Song, X., Li, L., Ayaz, G., Cheng, H., Tempel, W., Dong, C., Loppnau, P., Cote, J.*, et al.* (2020). Structural Basis for EPC1-Mediated Recruitment of MBTD1 into the NuA4/TIP60 Acetyltransferase Complex. Cell reports *30*, 3996-4002 e3994.
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 biology *9*, R137.
Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.P., and Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics *30*, 1006-1007.
Zhou, M., Xu, X., Wang, H., Yang, G., Yang, M., Zhao, X., Guo, H., Song, J., Zheng, H., Zhu, Z.*, et al.* (2020). Effect of central JAZF1 on glucose production is regulated by the PI3K-Akt-AMPK pathway. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *34*, 7058-7074.

SUPPLEMENTARY METHODS

Acid Extraction of Histones

EPC1-PHF1-Flag Tagged K562 and control cells were grown to 10^7 cell density in 100mm dishes, harvested, washed twice in 1XPBS. Cells were pelleted and resuspended in 1ml lysis buffer (PBS containing 0.5% Triton X 100 (v/v), 2 mM phenylmethylsulfonyl fluoride (PMSF), 0.02% (w/v) NaN3, 5mM sodium Butyrate) for 10 min, nuclei were spun down, resuspended in half volume of lysis buffer and centrifuged. Pellet was resuspended in 250 µl 0.2N HCl, incubated overnight at 4C, centrifuged. Supernatant containing acid extracted histones was neutralized with 1/10 volume of NaOH. Histones were quantified by Bradford assay, resolved on 15% SDS-PAGE and Immunoblotted with anti-H3K27me3 and anti-H4 antibodies.
(https://www.abcam.com/protocols/histone-extraction-protocol-for-western-blot**)**

Colony formation assay: HEK293T were transfected using the calcium phosphate method and then grown under selection with 300ug/ml hygromycin B for two weeks. Colonies that formed were fixed with methanol and colored with a solution of Giemsa stain for easier visualization. Plates were scanned using an Epson Perfection 4870 scanner; colonies in the images were counted using Alpha Imager software.

ChIP-qPCR PRIMERS USED IN THE STUDY		
Primer Sequence 5' to 3'	Name	Source
TGCCTTTCCTAACCAGTTCAGC	HOXA9	Dalvai M et al., 2015
CGGGCAGAAGACGCACATCCCG	HOXA9	Dalvai M et al., 2015
CGCCTGAGCAAAGTAAATGAG	MyoD	This Study
GATATAGCGGATGGCGTTG	MyoD	This Study
CCATGCCTGAGACCTTTTTC	RPL36AL	Dalvai M et al., 2015
TGCTCAGAATCCTGGGTAGG	RPL36AL	Dalvai M et al., 2015
AAGGGGCCTGAGTTCTTCTC	Intergenic chr12	Jacquet K et al., 2016
CACAGCTTCCATCTGGCTCT	Intergenic chr12	Jacquet K et al., 2016
TGCTTGGGCAGCAGGCTGTG	p21	This Study
CCTCACCTGAAAACAGGCAGCCC	p21	This Study
GCAGGTAGCACGGAAAAGAC	RPSA	Jacquet K et al., 2016

AGAAAGCGGGCTAACATCCT	RPSA	Jacquet K et al., 2016
AGA GAC AAT GGA CTC TGC GG	HOXD13-A	This Study
TGA GGG GTT GTC CTT GGA A	HOXD13-A	This Study
GAA AAG GGA CCG ACA CGA	HOXD13-B	This Study
CAC AGG CAC AAC TCC CAC TC	HOXD13-B	This Study
GCA GTC TGC ACA GAG CAG TT	HOXD13-C	This Study
GGC CAT TTC CCT GAG TG	HOXD13-C	This Study
GGA GAG CTC CTC AGA AAC CC	HOXD13-D	This Study
CCC TTT GTC TGA CCG TGA GT	HOXD13-D	This Study
TTACGGGACTTGGGTTGGGTC	RPSA-coding	This Study
ACCTCTTACCCCACTTTCCACT	RPSA-coding	This Study
TCCCTTTACCCGTCTCGTCT	CCNA2	This Study
CGGTACTGAAGTCCGGGAAC	CCNA2	This Study

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Histone H2B (Rabbit Polyclonal)	Abcam	Cat# ab1790, RRID:AB_302612
Histone H2A.Z (Rabbit Polyclonal)	Abcam	Cat# ab4174, RRID:AB_304345
IgG (Rabbit)	Millipore	Cat# PP64, RRID:AB_97852
Histone H3 (Rabbit Polyclonal)	Abcam	Cat# ab1791, RRID:AB_302613
Histone H4 hyperacetylated (Penta) (Rabbit Polyclonal)	Millipore	Cat# 06-946, RRID:AB_310310
H3K27me2/me3 (Mouse monoclonal)	Active Motif	Cat# 39535, RRID:AB_2793246
H3K27me3 (Rabbit Polyclonal)	Active Motif	Active Motif Cat# 39155, RRID:AB_2561020
Histone H3K36me3 (Rabbit Polyclonal)	Abcam	Cat# ab9050, RRID:AB_306966
c-Myc (Mouse Monoclonal)	Sigma-Aldrich	Cat# M5546, RRID:AB_260581
EZH2 (Goat Polyclonal)	Santa Cruz Biotechnology	Cat# sc-17270, RRID:AB_640300
SUZ12 (Rabbit Polyclonal)	Cell Signaling	Cat# 3737, RRID:AB_2196850
FLAG M2-Peroxidase (HRP) (Mouse Monoclonal)	Sigma-Aldrich	Cat# A8592, RRID:AB_439702
TRRAP (Goat Polyclonal)	Santa Cruz Biotechnology	Cat# sc-5405, RRID:AB_2209666
p400 (Rabbit Polyclonal)	Abcam	Cat# ab5201, RRID:AB_304780
KAT5 (Goat Polyclonal)	Santa Cruz Biotechnology	Cat# sc-5727, RRID:AB_2128817
DMAP1 (Rabbit Polyclonal)	Thermo Fisher Scientific	Cat# PA1-886, RRID:AB_2292767
ING3 (Goat Polyclonal)	Abcam	Cat# ab3713, RRID:AB_2127784
BAF53A (Rabbit Polyclonal)	Abcam	Cat# ab3882, RRID:AB_304125
KAT5 (Rabbit Polyclonal)	Epitomics	Cat# S2475

BRD8 (Rabbit Polyclonal)	Bethyl Laboratories	Cat# A300-219A, RRID:AB_263340
Biological Samples		
RT-PCR fragment from Patient Tumor	Micci F et al., 2006	
Chemicals, Peptides, and Recombinant Proteins		
Protein-G-Dynabeads	Invitrogen	Cat#10004D
Protein-A-Dynabeads	Invitrogen	Cat#10002D
Anti-FLAG M2 Affinity Beads	Sigma-Aldrich	Cat# A2220, RRID:AB_10063035
Streptactin Sepharose Beads	Iba/Neuromics	Cat#171201-002
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668027
Ponasterone A	Invitrogen	Cat#H10101
[3H]-labeled Acetyl-CoA	Perkin Elmer	
Acetyl-CoA	Sigma-Aldrich	
[3H]AdoMet (S-adenosyl-L-[methyl-3H] methionine	Perkin Elmer	
Critical Commercial Assays		
SF Cell Line 4D-NucleofectorTM X Kit L	Lonza	Cat #: V4XC-2012
Deposited Data		
ChIP seq		
Microarray		
Mass Spectrometry	MassIVE	MSV000083618 MSV000086476
Experimental Models: Cell Lines		
K-562	ATCC	CCL-243
HEK-293	ATCC	CRL-1573
HEK293T/17	ATCC	CRL-11268
K562-Mock-HA-FLAG	This study	
K562-EPC1(1-581)-HA-FLAG	This study	
K562-PHF1-HA-FLAG	This study	
K562-EPC1-PHF1-HA-FLAG	This study	
K562-EPC1-PHF1-3X FLAG-2XStrep	This study	
HEK393- N term Flag	This study	
HEK393- Flag-JAZF1	This study	
EcR-293	Jeronimo et al., 2004	
EcR-293-pMZI -JAZF1-TAP	This study	

HEK293T-pRCF	This study	
HEK293T-pRCF-EPC1	This study	
HEK293T-pRCF-EPC1(1-581)	This study	
HEK293T-pRCF-EPC1-PHF1	This study	
HEK293T-pRCF-PHF1	This study	
K562-3X FLAG-2XStrep-EP400 (CRISPR)	Dalvai M et al., 2015	
K562-EZH2-3X FLAG-2XStrep (TALEN)	Dalvai M et al., 2015	
Oligonucleotides		
ChIP qPCR Primers	This Study (see Table)	
Recombinant DNA		
PHF1 ORF	Open Biosystems	
JAZF1 ORF	GE Healthcare	
AAVS1- PGK Puro_2XHA-3XFlag	Jacquet K et al., 2016	
AAVS1_Puro_PGK1_3XFlag_Twin_Strep	Dalvai M et al., 2015	RRID:Addgene_68375
p3X-CMV-Flag-14	Sigma-Aldrich	
pREV-CMV-3xFLAG	Avvakumov N et al., 2012	
pMZI_TAP	Zeghouf M et al., 2004 Rigaut et al., 1999	
pcDNA3_6XMyc	Julien C et al., 2003	
AAVS1_Puro_PGK1_3XFlag_Twin_Strep_EPC1-PHF1	This study	
AAVS1- PGK Puro_2XHA-3XFlag-PHF1	This study	
AAVS1- PGK Puro_2XHA-3XFlag-EPC1(1-581)	This study	
AAVS1- PGK Puro_2XHA-3XFlag-EPC1-PHF1	This study	
pREV-CMV-3xFLAG-EPC1	This study	
pREV-CMV-3xFLAG-EPC1(1-581)	This study	
pREV-CMV-3xFLAG-PHF1	This study	
pREV-CMV-3xFLAG-EPC-PHF1	This study	
pcDNA3_6XMyc_JAZF1	This study	
pMZI_TAP_JAZF1	This study	
pCMV-3Tag 1A_Flag_JAZF1	This study	
AAVS1- PGK Puro_2XHA-3XFlag-JAZF1	This study	
Software and Algorithms		
Prism	GraphPad Software	RRID:SCR_002798
SAMtools	Langmead B et al., 2009	RRID:SCR_005611

MACS	Zhang Y et al., 2008	RRID:SCR_013291
BEDTools	Quinlan and Hall, 2010	RRID:SCR_006646
DESeq(R package)	Anders and Huber, 2010	RRID:SCR_000154
lumi (Bioconductor package)	Du P et al., 2008	RRID:SCR_012781
pyGenometracks	F Ramírez et al., 2018	
ProHits Platform	Liu G et al. 2016	
ProteoWizard ((v3.0.4468)	Kessner et al., 2008	RRID:SCR_012056
AB SCIEX MS Data Converter	SciEX	
Mascot	Matrix Science	RRID:SCR_014322
Decontaminator	Lavallée-Adam M et al., 2011	
CrossMap	Zhao H et al., 2014	RRID:SCR_001173