Article

JmjC family of histone demethylases form nuclear condensates

Marta Vicioso-Mantis[1], Samuel Aguirre[1] and Marian A. Martínez-Balbás[1]#

[1] Department of Molecular Genomics. Instituto de Biología Molecular de Barcelona (IBMB), Consejo Superior de Investigaciones Científicas (CSIC), Barcelona 08028, Spain.

* Correspondence: mmbbmc@ibmb.csic.es; Tel: (+34) 93 403 4961; Fax: (+34) 93 403 4979

ABSTRACT: JmjC-family of lysine demethylases (JMJC-KDMs) plays an essential role in controlling gene expression and chromatin structure. In most cases, their function has been attributed to the demethylase activity. However, accumulating evidence demonstrates that these proteins play roles distinct from histone demethylation. This raises the possibility that they might share domains that contribute to their functional outcome. Here, we show that the JMJC-KDMs contain low complexity domains and intrinsically disordered regions (IDR), which in some cases reached 70% of the protein. Our data revealed that PHF2, KDM2A, and KDM4B cluster by phase separation. Moreover, our molecular analysis suggests that the PHF2 IDR contributes to transcription regulation. These data suggest that clustering via phase separation is a common feature that JMJC-KDMs utilize to facilitate their functional responses. Our study uncovers a novel potential function for the JMJC-KDM family that sheds light on the mechanisms to achieve the competent concentration of molecules in time and space within the cell nucleus.

Keywords: Epigenetic regulation/ Histone demethylases/ JmjC/ IDR/ Phase separation

1. INTRODUCTION

The regulated chromatin activity is the key event that allows multicellular organisms to support the process implicit in development, physiology, and homeostasis. It has become clear during the last decades that this control is mediated through histones and DNA modifications, the incorporation of histone variants, and the action of chromatin remodeling complexes [1]. One of the significant modifications, particularly relevant during development, is the histone lysine methylation, which encompasses the introduction of one (me1), two (me2), or three (me3) methyl groups to histones lysines or me1 and me2 to arginines [2]. Depending on their genomic localization and methylated residue, the different methylation states will imply transcriptional activation or repression [3]. Although histone methylation might impact chromatin structure, the central role of this modification is to be recognized by "reader" proteins that specify the downstream events. Early days, it was believed that histone methylation was irreversible. The discovery of the first histone lysine demethylase (KDM), KDM1A/LSD1, which removes methyl groups from histone H3 on lysine 4 (H3K4) [4] underscored the dynamic nature of histone methylation. Soon after that, the enzyme KDM2A/JHDM1A/FBXL11 was identified. It possesses a newly identified catalytic domain – the Jumonji-C (JmjC) domain [5]. KDM2A was the first member of an extended family of related demethylases, which coordinate iron to demethylate lysines using α–ketoglutarate (α-KG) as a cofactor, the JmjC domain-containing family of KDMs (from now on JMJC-KDM) [6, 7]. Since these initial findings, our understanding of how these enzymes function has progressed rapidly. Numerous studies over the last decade have highlighted their role in many genomic processes such as transcription control [8, 9], DNA replication [10, 11], and DNA repair [12-14]. Notably, histone demethylase (HDM) activity alteration is linked to a wide spectrum of human diseases, including mental disorders and cancer [15, 16]. In this way, KDMs have become new attractive therapeutic targets [17, 18].

The members of the JMJC-KDM family play an essential role in gene expression regulation, chromatin organization, and cellular homeostasis [19]. In most cases, this was attributed to their demethylase activity. However, accumulating evidence demonstrates that these proteins play roles distinct from histone demethylation [20, 21]. These data lead us to hypothesize that the major regulatory function of these proteins might rely not only on their catalytic capacity but other domains could be involved. Thus, whether JMJC-KDMs have any common domain that modulates their activity arises.

Recently, it has been shown that multiple components of the transcriptional machinery cluster via phase separation to efficiently accomplish their function during transcription [22-24]. These transcriptional hubs or condensates include RNA polymerase II (RNAPII) [22], cofactors (Mediator, Bromodomain containing 4 (Brd4)) [22, 25], transcription factors (Oct4, Smad) [26, 27], and even histone modifications themselves are involved in the process [28-30]. Phase separation is a dynamic process that allows the formation of biomolecular condensates by establishing weak, multivalent interactions. These commonly use electrostatic and hydrophobic residues [31, 32]. These interactions are often driven by intrinsically disordered regions (IDR) and low complexity domains [33, 34]. First analyses of the JMJC-KDMs indicated that a large part of the amino acids of these proteins is not organized into known domains. A closer analysis clarified that they belong to disordered regions. Thus, the possibility exists that JMJC-KDMs utilize these domains to cluster into biomolecular condensates to facilitate their function.

Here, we show that all JMJC-KDMs contain intrinsically disordered regions, which in some cases reach up to 70% of the protein. Our data revealed that plant homeodomain finger protein 2 (PHF2), KDM2A, and KDM4B could phase separate. Finally, our PHF2 molecular analysis suggests that clustering is a common feature that JMJC-KDMs might utilize in addition to their catalytic activity to promote their functional responses. Thus, our study uncovers a novel molecular property of JMJC-KDM family that shed light on the ways to achieve the competent concentration of molecules inside the cells.

2. RESULTS

2.1. JMJC-KDMs are intrinsically disordered proteins

In the last years, numerous works have shed light on the impact of IDRs on transcriptional regulation [35, 36]. Since a large part of JMJC-KDMs does not encompass any identified structure, we questioned whether JMJC-KDMs belong to the group of the intrinsically disordered proteins. To assess this, we analyzed the sequence of the JMJC-KDMs searching for disordered regions using the following algorithms: PONDR-VL3, IUPred, and VSL2 (see Material and Methods). Overall, the three algorithms agreed on the significant disorder score of these proteins (Figure 1A and Figure S1). In particular, PONDR-VL3 revealed a disorder score ranging from 0.61 (PHF2) to 0.40 (KDM4), considerably higher than the values obtained when analyzing the proteasome components PMS4A and PMSA7 and actin well-known structured proteins used as negative controls (Figure 1A and Figure S1). In addition, a compelling amino acids proportion (from 71.39% KDM6B to 32.67% KDM4C using PONDR-VL3) exists in disordered domains (Figure 1B). This value is higher than the obtained for the negative controls PSMA4 (23.37%) and similar to examples of disordered proteins such as MED1 (68.25%), RNAPII (36.35%) or HP1α (63.87%) [25] (Figure 1B).

Interestingly, 99% of identified IDRs on JMJC-KDMs contain amino acid tracks conserved within the family (see Material and Methods). These tracks were enriched in lysines (K) (9.7%) and prolines (P) (9.1%) as well as in hydrophobic amino acids (leucines, L, 10.3%) (Figure 1C), features that have been associated with the ability to phase separate [33].

A

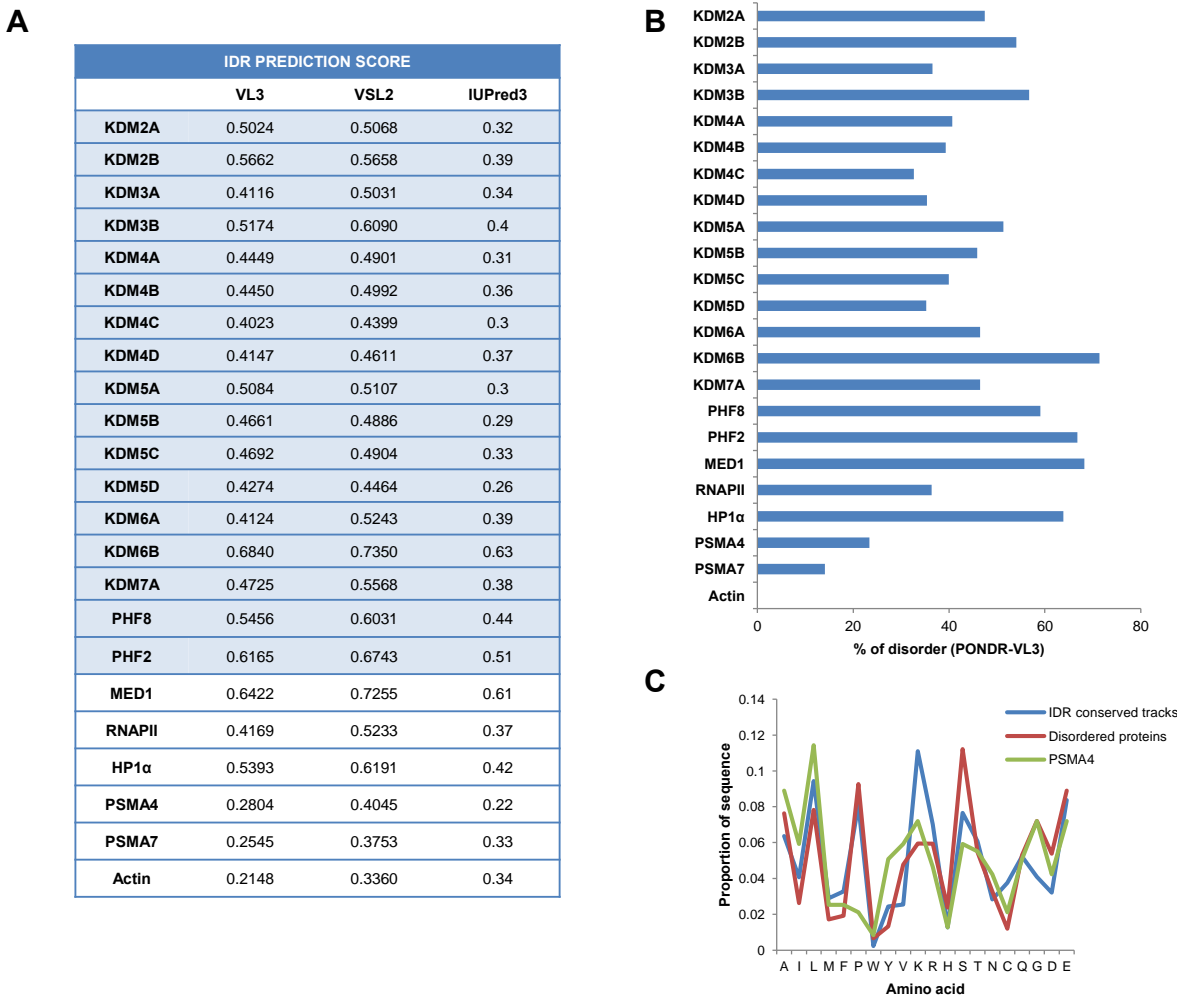| IDR PREDICTION SCORE			
	VL3	VSL2	IUPred3
KDM2A	0.5024	0.5068	0.32
KDM2B	0.5662	0.5658	0.39
KDM3A	0.4116	0.5031	0.34
KDM3B	0.5174	0.6090	0.4
KDM4A	0.4449	0.4901	0.31
KDM4B	0.4450	0.4992	0.36
KDM4C	0.4023	0.4399	0.3
KDM4D	0.4147	0.4611	0.37
KDM5A	0.5084	0.5107	0.3
KDM5B	0.4661	0.4886	0.29
KDM5C	0.4692	0.4904	0.33
KDM5D	0.4274	0.4464	0.26
KDM6A	0.4124	0.5243	0.39
KDM6B	0.6840	0.7350	0.63
KDM7A	0.4725	0.5568	0.38
PHF8	0.5456	0.6031	0.44
PHF2	0.6165	0.6743	0.51
MED1	0.6422	0.7255	0.61
RNAPII	0.4169	0.5233	0.37
HP1α	0.5393	0.6191	0.42
PSMA4	0.2804	0.4045	0.22
PSMA7	0.2545	0.3753	0.33
Actin	0.2148	0.3360	0.34

Figure 1

Figure 1. JMJC-KDMs are intrinsically **disordered proteins.** (A) Table representing the IDR prediction score of JMJC-KDMs using PONDR-VL3, VSL2, and IUPred3 algorithms. (B) Disorder prediction of human JMJC-KDMs using PONDR-VSL3 algorithm (length of disordered segments > 50 amino acids). (C) The amino acid composition of the IDR conserved tracks between KDMs families and disordered proteins. The ordered protein PSMA4 was also included as a negative control (see Methods).

2.2. JMJC -KDMs are prone to phase separate

IDRs drive condensate formation via phase separation, so we used the PSPredictor and catGRANULE algorithms (see Material and Methods) to predict JMJC-KDMs phase separation propensity. Both tools returned high scores (from 0.99 PHF2 to 0.01 KDM4A for PSPredictor) (Figure 2A), in some cases similar to proteins known to be involved in phase separation (e.g., MED1, 0.99, RNAPII, 0.64 or HP1α, 0.96) and consistently higher than the proteasome proteins PSMA4 (0.001), PSMA7 (0.079) or actin (0.014) (Figure 2A). Altogether these data point to JMJC-KDMs as disordered proteins with the potential to undergo phase separation.

To further analyze the ability of JMJC-KDMS to cluster together via phase separation, we chose three KDMs among those with high (PHF2), low (KDM4B), and medium (KDM2A) scores according to both predictors and that belong to different families (Figure 2A, B). Our analysis did not include KDM6 family members that show very high scores because it has recently been demonstrated that their two members undergo phase separation [9, 37]. The amino acid composition has also been shown to play an essential role in the IDR-mediated formation of transcriptional condensates. We thus analyzed the amino acid composition of the selected KDMs and found an enrichment of lysines and serines for PHF2; glutamate, serine, and lysine for KDM4B; glutamate, arginine, and leucine for KDM2A (Figure 3A) compared to the average in the mouse disorders proteins [38,

39]. These residues are associated with the ability to phase separate [33]. We used MED1 and PMSA4 proteins as positive and negative controls, respectively.

Moreover, disordered regions frequently concur with low-complexity domains that contribute to phase separation [25, 31, 33, 40]. Then, we check whether the KDMs are enriched in low-complexity domains using the SEG algorithm (see Material and Methods). As shown in Figure 3B, between 2.6% (KDM4B) and 13.5% (PHF2) of KDM sequences were predicted to contain low-complexity segments. Furthermore, it has been previously described that hydrophobic interactions might be crucial for biomolecular condensate formation [23, 41]. Then, we analyze the hydrophobicity of the selected JMJC-KDMs, finding the presence of hydrophobic tracks (Figure 3C). Thus, our data suggest that the amino acid composition of JMJC-KDMs primes them to be involved in the phase separation process.

Figure 2

A

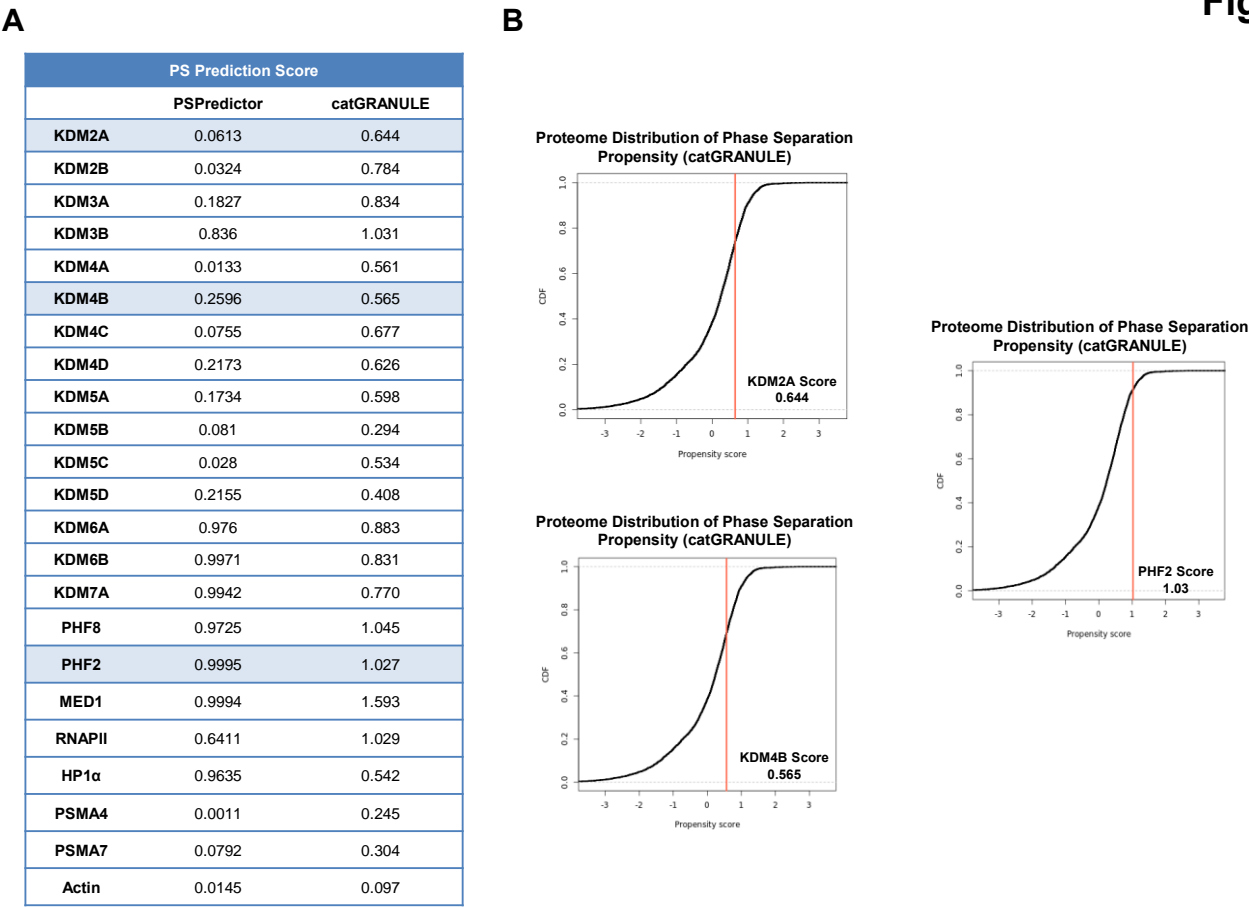| PS Prediction Score		
	PSPredictor	catGRANULE
KDM2A	0.0613	0.644
KDM2B	0.0324	0.784
KDM3A	0.1827	0.834
KDM3B	0.836	1.031
KDM4A	0.0133	0.561
KDM4B	0.2596	0.565
KDM4C	0.0755	0.677
KDM4D	0.2173	0.626
KDM5A	0.1734	0.598
KDM5B	0.081	0.294
KDM5C	0.028	0.534
KDM5D	0.2155	0.408
KDM6A	0.976	0.883
KDM6B	0.9971	0.831
KDM7A	0.9942	0.770
PHF8	0.9725	1.045
PHF2	0.9995	1.027
MED1	0.9994	1.593
RNAPII	0.6411	1.029
HP1α	0.9635	0.542
PSMA4	0.0011	0.245
PSMA7	0.0792	0.304
Actin	0.0145	0.097

B

Figure 2. JMJC -KDMs are prone to phase-separate. (A) Predictions of JMJC-KDMs putative capacity to phase separate, determined with the PSPredictor (on the left) and the catGRANULE (on the right) algorithms. (B) Distribution of phase separation propensity of KDM2A, KDM4B, and PHF2 determined by the catGRANULE algorithm.
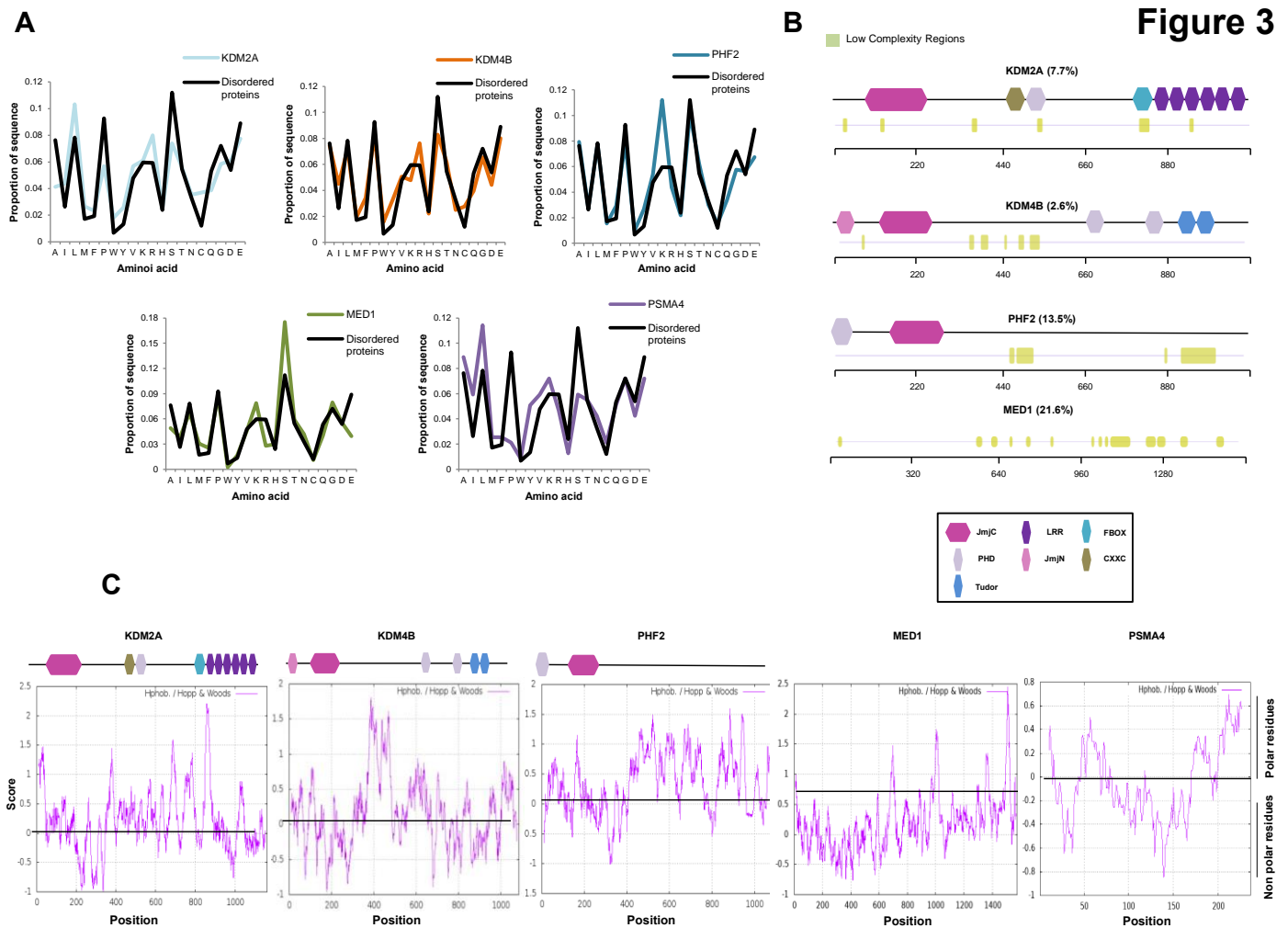
Figure 3. JMJC-KDMs contain low-complexity domains. (A) Amino acid composition of KDM2A, KDM4B, PHF2, MED1, PMSA4, and disordered proteins defined by the presence of a 50 residues fragment whose IUPRED median score is at least 0.55 and that it is not found in Pfam. (B) Analysis of the presence of low-complexity domains in KDM2A, KDM4B, PHF2 and MED1 using the SEG algorithm. The percentage of low complexity regions is indicated. Low complexity regions are depicted in green. A schematic representation of the KDM domains is shown on top. (C) The hydrobicity profile of KDM2A, KDM4B, PHF2, MED1, and PMSA4 was determined using the ExPASy website with the Hopp and Woods scale and a sliding window of 21. A scheme of the KDMs domains is depicted on top of each panel.

2.3. KDM2A, KDM4B, and PHF2 KDMS cluster in vitro and inside the cell

Considering the potential of the analyzed KDMs to phase separate and their role in transcriptional control, we further proceed to study the capacity of these proteins to form biomolecular condensates. We fused KDM2A, KDM4B to EGFP, and PHF2 to mCherry fluorescence tags (Figure S2A), we expressed them in HEK293T cells (Figure S2B), and we performed *in vitro* droplet assays using nuclear extracts. Our data show that the selected KDMs formed droplets that do not appear when we overexpressed EGFP or mCherry alone, used as negative controls, suggesting that the droplets can be attributed to KDMS proteins and not to the fluorescence tags (Figure 4A). The number of droplets was similar to those obtained for MED15 and JMJD3, two proteins that experience LLPS [9, 42], used as positive controls. Furthermore, the obtained droplets showed features [roundness, convexity, and aspect ratio] characteristic of a liquid-like nature (Figure 4B).

Next, we analyzed the ability of fluorescence KDMs to form puncta in fixed cells, observing that when these KDMs were overexpressed in HEK293T cells, condensates could be detected as nuclear puncta (Figure 4C). It has been previously demonstrated that the aliphatic alcohol 1,6-hexanediol disrupts the hydrophobic interactions that maintain the phase-separated droplets. Thus, we tested the sensitivity of the KDM's condensates to

1,6-hexanediol. The data demonstrated that the treatment reduced the number and size of KDMs puncta (Figure 4D) without affecting the total amount of protein (Figure S2C). We also analyzed the formation of puncta by the endogenous PHF2. Immunofluorescence experiments in HEK293T and NIH3T3 cells showed that the endogenous protein could form nuclear condensates (Figure S2D), ruling out that the observed puncta were due to an overexpression artifact.

These data suggest that JMJC-KDMs condensate and that these condensates are a separated phase inside the cell.
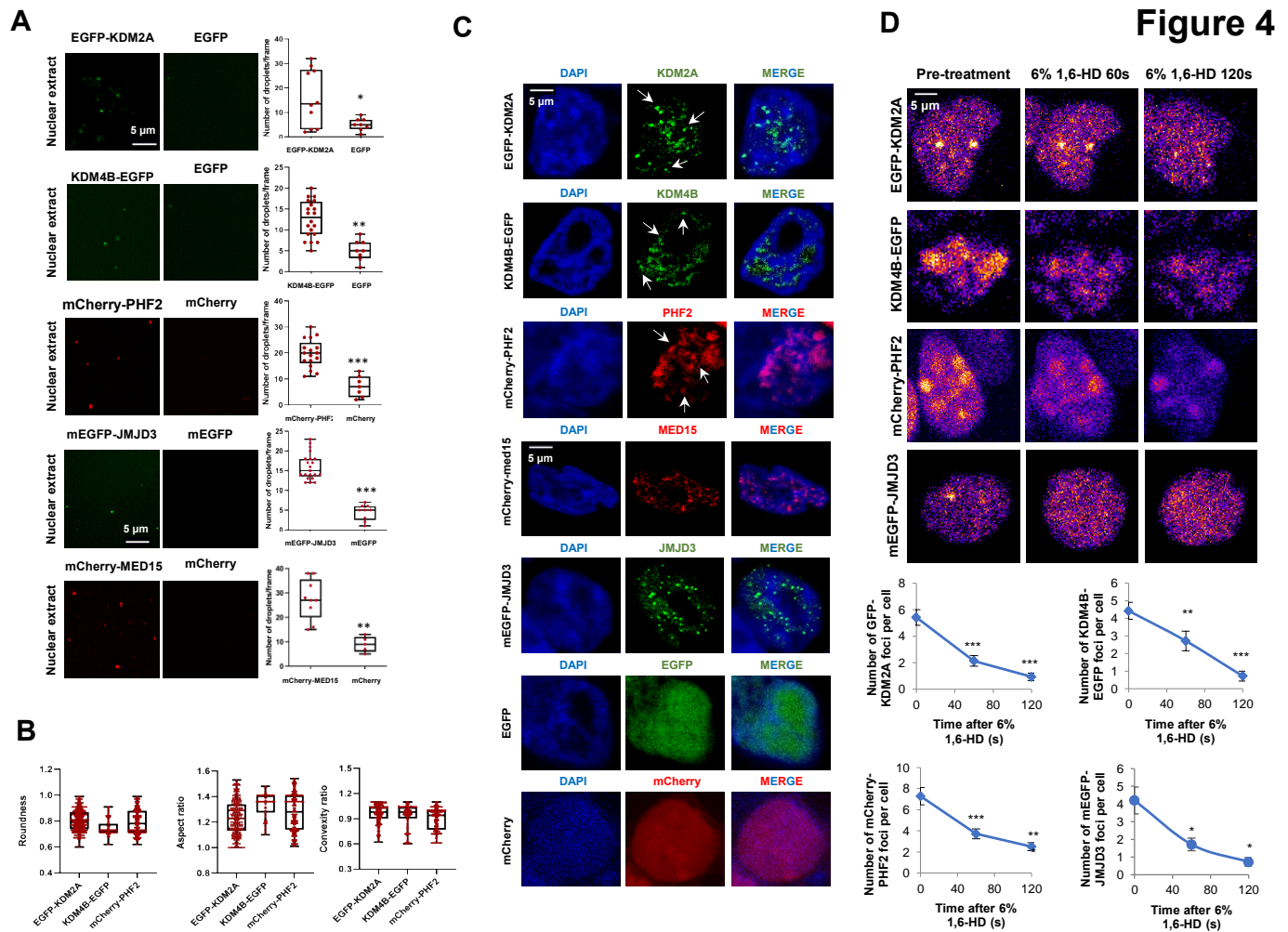
Figure 4. JMJC-KDMs undergo LLPS *in vitro* and inside the cell. (A) EGFP-KDM2A, KDM4B-EGFP, mCherry-PHF2, EGFP-JMJD3, mCherry-MED15, EGFP, and mCherry proteins were analyzed using droplet-formation assays in nuclear extracts at room temperature in the presence of 150 mM NaCl. Quantification of the droplets is displayed on the bottom. Data are the mean ± SEM. ***p<0.001 (Student's *t*-test). Droplets in 5 fields from three biologically independent experiments were quantified. Scale bar, 5 μm. (B) Boxplots represent the roundness, convexity, and aspect ratio features of the fluorescent-tagged KDMs droplets. Data are the mean ± SEM. Quantifications correspond to droplets in 5 fields from three biologically independent experiments for each protein. (C) Confocal microscopy images of HEK293T cells transfected with EGFP-KDM2A, KDM4B-EGFP, mCherry-PHF2, EGFP-JMJD3, mCherry-MED15, EGFP, or mCherry and visualized with the fluorescence tag and DAPI to stain the DNA. Images are representative of 3 biologically independent experiments. Scale bar, 5 μm. (D) HEK293T cells were transfected with 0.05ug plasmid encoding EGFP-KDM2A, KDM4B-EGFP, mCherry-PHF2, or EGFP-JMJD3; they were treated with 6% 1,6-Hexanediol for 5 minutes and imaged at 60 and 120 s. Nuclei were visualized with DAPI (blue). Quantification of the nuclear puncta displayed per cell along

the treatment is shown on the right. Data are the mean ± SEM. **p<0.01; ***p<0.001 (Student's *t*-test). *n*=30 transfected cells were quantified; Images are representative of three biologically independent experiments. Scale bar, 5 μm.

2.4. KDM2A, KDM4B, and PHF2 condensates correlate with transcription

The data above indicate that JMJC-KDMs undergo LLPS in the nucleus; then, we hypothesized that clustering propensity could contribute to the transcriptional regulation mediated by JMJC-KDMs. To test this idea, we analyzed the localization of KDM condensates with the repressive H3K9me3 and active H3K36me3 histone marks by immunofluorescence assays. These histone marks were used because they form nuclear foci and are landmaks of transcriptional repression and activation respectively. KDM4B-EGFP and mCherry-PHF2 nuclear condensates excluded the repressive mark H3K9me3 (Figure 5A), consistent with their role in transcription activation. Moreover, according to its role as a transcriptional repressor, EGF-KDM2A puncta were excluded from condensates marked by the active mark H3K36me3 (Figure 5A).

To further analyze the contribution of the IDRs to transcription regulation, we selected the PHF2 protein (which has been previously studied by our lab) to construct a mutant protein that lacks part of its IDR. In particular, we depleted a strikingly lysine-enriched region (amino acids 487-806) (Figure S3A) that is widely conserved among mammals (Figure S3B). We named this mutant PHF2 ΔCharged (see Material and Methods). Next, we tested its ability to form nuclear puncta. When compared to the PHF2 WT protein, the mutant could not phase separate (Figure 5B). Thus, we concluded that the lack of the charged region impacted the phase separation potential of PHF2. Interestingly, the mutation did not affect the ability to demethylate H3K9me2 (Figure 5B). Next, we decided to study whether the lack of phase separation of the PHF2 mutant translated into a functional impairment, affecting the transcription regulation. We measured the expression levels of some PHF2-dependent genes (*E2f3* and *Pcna*, [43]) in NIH3T3 cells that overexpress either PHF2 WT or PHF2 ΔCharged (Figure 5C, left). While the WT protein increased the mRNA level of these genes, the mutant acted as a dominant-negative protein hindering their expression (Figure 5C, right). The housekeeping gene *Rps23*, which was used as a negative control, was not affected by the deletion of the charged track (Figure 5C). Thus, the capacity of PHF2 to activate transcription seems to be related to the presence of the charged domain in its IDR.

Altogether these results support a model in which JMJC-KDM might use a common mechanism that might regulate their outcome by forming IDR-driven condensates. Thus, our study uncovers a novel molecular property of JMJC-KDM family that shed light on the ways to achieve the competent concentration of molecules inside the cells.

3. DISCUSSION

The requirement of different molecules to perform a function in the correct concentration, time, and space inside the cells is a central question. The cells solve this problem continuously in a dynamic way. With the evolution, some organelles have developed the formation of a membrane to facilitate biological functions such as mitochondria, nuclei, Golgi apparatus, etc. Other exists in a membraneless fashion (Cajal bodies, nucleoli, or signaling bodies). Recently, the membraneless condensates have received much attention; however, how they are formed and their function is unclear. In recent years, it has been shown that many of these condensates have liquid-like properties and are formed by phase separation [31]. Phase separation plays a myriad of biological functions, from signal transduction, filament polymerization, and chromatin structure to transcriptional control [44]. Moreover, phase separation is related to many diseases, from cancer to neurodegenerative disorders [45-47].

Recently, the contribution of phase separation to transcription regulation has become an intense area of research [38]; however, the molecular mechanisms through which this control is exerted are not fully understood. Our study reveals a common feature of the chromatin-acting enzymes, the JMJC-KDMs. We show that they are prone to phase

separate and that representative examples form nuclear condensates that are important for their functional outcome.

Bearing this in mind, we can question the role of the JMJC-KDMs condensates. We will discuss some relevant aspects that could be affected by the phase separation of these regulatory proteins, such as the chromatin structure and the machinery involved in transcription.

3.1. JMJC-KMD condensates may affect their functional substrate, the chromatin structure.

Although our results for PHF2 suggest that phase separation might not affect its catalytic activity, an interesting possibility is that phase separation might affect the catalytic activity of some JMJC-KDM enzymes, as recently demonstrated for the JMJC-KDM KDM6A [37]. That, in turn, might have a profound effect on chromatin structure and/or phase separation itself. Indeed, it has been described that histone modifications can affect chromatin phase separation. For instance, in the presence of BRD4 protein, highly acetylated chromatin forms a new phase-separated state similar to the chromatin subcompartments [28]. The droplets formed by BRD4 have a repelling effect on chromatin, which leads to the remodeling of chromatin structure [48]. These data suggest that changes in KDM's activity mediated by phase separation might affect chromatin structure.

On the other hand, it is well known that the chromatin structure also affects the phase separation process [28, 49]. Phase separation droplets tend to be formed in the chromatin structure with open and low-density areas [50]. The surrounding chromatin has a certain mechanical resistance to the formation of droplets [51, 52]. Then, changes in the methylation status might impact chromatin conformation and its functional outcome.

3.2. JMJC-KMD condensates may affect transcriptional machinery function

RNAPII is a key molecule in eukaryotic gene transcription. It forms phase-separated condensates through the interaction of its C-terminal domain (CTD) with many components of the transcriptional machinery [53]. In particular, Mediator and Brd4 factors and coactivators stabilize the pre-initiation complex at gene promoters [54]. Other key components of the transcriptional machinery, the transcription factors, have also been shown to contribute to transcription through a phase separation process [26, 27]. The current model proposes that phase-separated condensates regulate gene transcription [23, 55, 56]. At promoter, an essential condensate formed by Mediator, transcription factors, coactivators, and non-phosphorylated RNAPII allows initiation. Downstream of the initiation site, phosphorylated RNAPII, nascent RNA, elongation factors, RNA processing factors, and specific elongation coactivator form a condensate that allows elongation [22, 56]. In addition to the promoter, it has been demonstrated that enhancers and enhancer-clusters participate in gene regulation through phase separation [57]. JMJD-KDMs are essential cofactors for transcription. They contribute to transcriptional control at promoter and enhancer sites and during elongation. Thus, KDM's ability to drive or form part of nuclear condensates might be critical to transcriptional hub formation. KDM's condensates might concentrate transcription factors, cofactors, and the general elongation factors at promoters or enhancers in a compartment to make transcription kinetically more efficient, as proposed for other factors [22, 26, 56]. If this were the case, JMJC-KDM biomolecular condensates would work as transcriptional hubs that compartmentalize coactivators and transcriptional machinery to favor the transcription process, as recently proposed [9, 37].

In addition, other possible scenarios can be envisioned.

-JMJC-KDM condensates may physically insulate the transcriptional machinery from its regulators to prevent inactivation. In this way, KDM's condensates maintain the transcriptional process ongoing

-They could stabilize complexes formed by multiple components, which generally are involved in transcription initiation and elongation.

-Condensate formation might facilitate contacts between distal genomic regions, that would favor transcription burst frequency or intensity, as it has been suggested for KDM6A and KDM6B [9, 37].

-JMJC-KDMs condensation might serve as a mechanism for accessing the compacted and closed chromatin, as demonstrated for KL4 during reprogramming [58]; the deformable properties of many IDRs might facilitate the access of these enzymes into the compacted facultative or constitutive heterochromatin to favor nucleosome unwrapping. This may be particularly important during development or reprogramming, where LLPS and KDMs have been shown to play a relevant role.

Phase separation is linked to the occurrence of many pathologies, such as cancer, neurodegeneration, or infections [45-47]. Recently, the histone demethylase KDM6A tumor suppressor capacity has been shown to be dependent on the ability of its IDR to phase separate [37]. Interestingly, many mutations on JMJC-KDMs that lead to diseases occur in their IDRs. In particular, PHF2 nonsense mutations in exons 12, 16, and 18 have been associated with different cancer types [59]. In all cases, PHF2 IDR has been compromised. A more detailed understanding is required of the direct or indirect involvement of the IDR from JMJC-KDM on pathologies and the possible condensing proteins mechanism in this process, and their potential as therapeutic targets.

Our study uncovers a novel molecular property of JMJC-KDM family that shed light on the ways to achieve the competent concentration of molecules inside the cells. Thus, JMJC-KDM's condensate formation might be a general and key physical platform for regulation chromatin structure and activity.
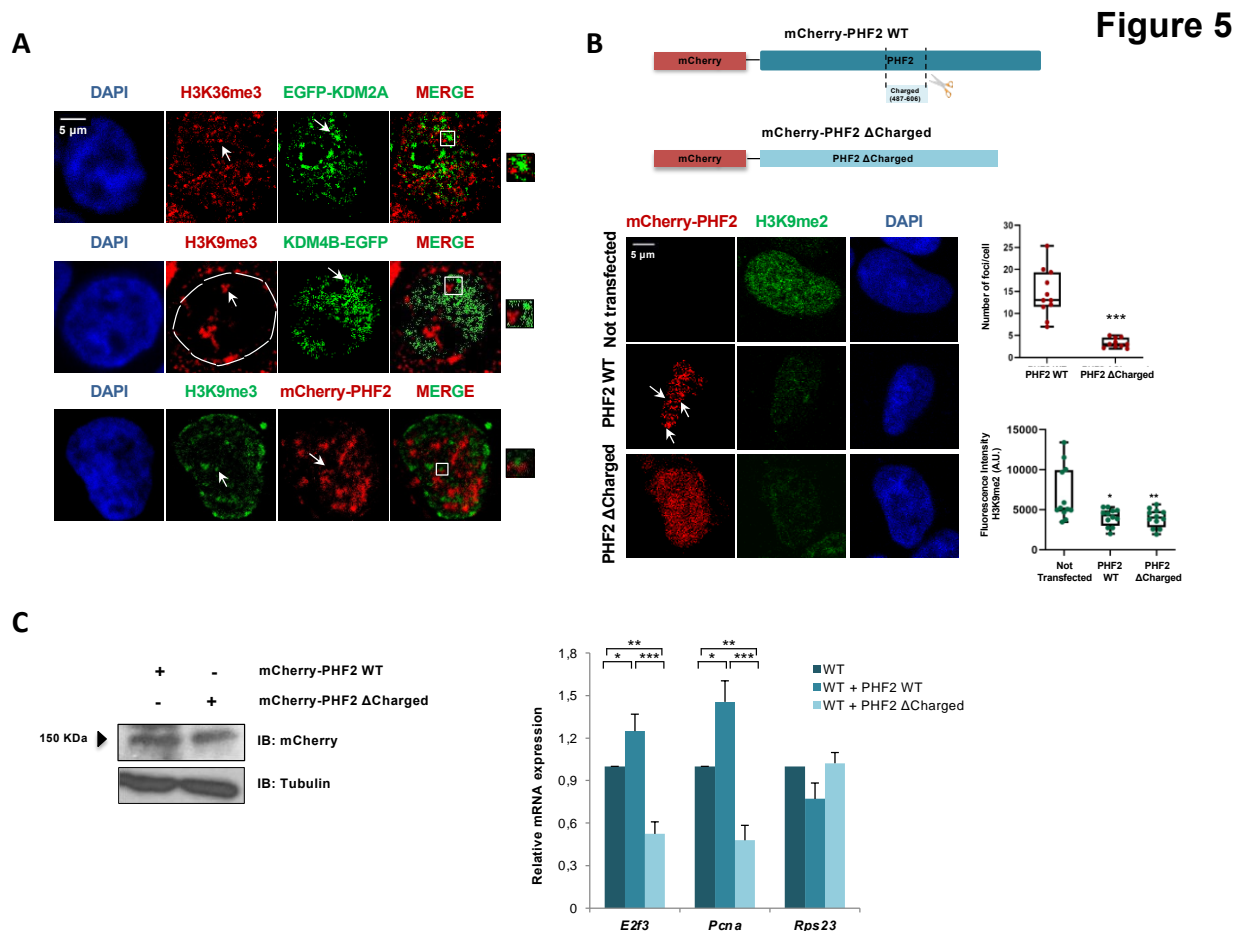

Figure 5. KDM2A, KDM4B, and PHF2 condensates correlate with transcription. (A) HEK293T cells were transfected with 0.05μg of EGF-KDM2A, KDM4B-EGFP, and mCherry-PHF2, and the localization of the active (H3K36me3) or the repressive (H3K9me3) or the active (H3K36me3) histone marks were analyzed using immunofluorescence staining with anti-H3K36me3 or anti-H3K9me3 or anti-H3K36me3. The JMJC-KDMs localization was determined following the

fluorescence signal. Nuclei were visualized with DAPI (blue). Scale bar, 5 μm. The images are representative of three independent experiments with similar results. (B) mCherry-PHF2 WT and mCherry-PHF2 ΔCharged expression vectors (top panel) were transfected into HEK293T cells. The formation of nuclear puncta was analyzed by the fluorescence signal. The levels of H3K3me2 were determined by immunofluoresce using anti H3K9me2 antiboy. The boxplots on the right shows the number of foci per cell and the H3K9me2 fluorescence intensity. *n*=30 transfected cells for condition. Data show the mean ± SEM. For the number of foci per cell, statistic comparisons were done between cells overexpressing PHF2 WT and cells overexpressing PHF2 ΔCharged (***p<0.001, Student's *t*-test). For fluorescence intensity, comparisons were performed between Not Transfected cells and cells overexpressing PHF2 WT (*p<0.05, Student's *t*-test), and between Not Transfected cells and cells overexpressing PHF2 ΔCharged (**p<0.01, Student's *t*-test). Images are representative of 3 biologically independent experiments. Scale bar, 5 μm. (C) PHF2 WT and PHF2 ΔCharged were transfected. 24h later, total protein extracts were prepared, and the expression levels of PHF2 WT and the mutant were determined by immunoblot using anti mCherry antibody. The image shown is representative of two independent experiments (left). The expression levels of *E2f3* and *Pcna* genes were quantified by qPCR in NIH3T3 cells expressing the empty vector (WT, dark blue), expressing PHF2 WT (WT + PHF2 WT, blue) or PHF2 ΔCharged (WT + PHF2 ΔCharged, light blue) (right). As a negative control, the mRNA levels of the housekeeping gene *Rps23* were measured. Error bars represent SEM. Statistic comparisons were done between WT cells (dark blue) and cells overexpressing PHF2 WT (WT + PHF2 WT, blue) (*p<0.05, Student's *t*-test); between WT cells (dark blue) and cells overexpressing PHF2 ΔCharged (WT + PHF2 ΔCharged , light blue) (**p<0.01, Student's *t*-test); and between cells overexpressing PHF2 WT (WT + PHF2 WT, blue) and cells overexpressing PHF2 ΔCharged (WT + PHF2 ΔCharged , light blue) (***p<0.01, Student's *t*-test). Results are representative of three biologically independent experiments.

4. MATERIALS AND METHODS

4.1. Sequence analysis and predictions

To calculate the Protein disorder estimations three prediction algorithms were used, PONDR-VL3 [60] IUPred [61] and PONDR-VSL2 [62]. The predictors provide a score between 0 and 1 for each amino acid where above 0.5 lies within a disordered region longer than 50 AA. The phase separation property of each protein was determined using the PSPredictor and catGRANULE [63] predictors. Low-complexity domains were determined using the SEG algorithm with the MobiDB database [64]. The Prot Pi Protein Tool website was used to analyze amino acid composition. The presence of a 50 residues fragment whose IUPRED median score was at least 0.55 and that was not found in Pfam was used as criteria to define disorder proteins, so that functional domains were avoided. The hydrophobicity was estimated with the ExPASy website [65] using the Hopp and Woods scale [66] and a sliding window of 21.

4.2. Amino acid composition of conserved IDRs within KDM families

The IDR conservation between the members of each JMJC-KDM family was assessed using the Basic Local Alignment Search Tool (BLAST ®) [67]. Then, to calculate the amino acid composition of these conserved IDRs, the number of each specific amino acid in the conserved sequences was divided by the total number of amino acids in the conserved sequences. This way, the amino acid proportion within a KDM family was calculated for each amino acid (e.g., proportion of K in KDM4 conserved tracks = number of K in KDM4 conserved tracks / number of total amino acids in KDM4 conserved tracks). Finally, total amino acid composition of the conserved IDRs was calculated as the mean between the amino acid proportions for the conserved IDRs in each family (e.g., total K content in IDR conserved tracks = mean of the proportion of K in KDM2 conserved tracks, KDM3 conserved tracks, KDM4 conserved tracks, KDM5 conserved tracks, KDM6 conserved tracks and KDM7 conserved tracks).

4.3. Cell culture and cell treatments

Human HEK293T and mouse NIH3T3 cells were grown in DMEM supplemented with 10% of fetal bovine serum (Gibco) and 1% of Penicillin/Streptomycin [68].

4.4. Plasmids

pCDNA-FLAG-EGFP-KDM2A and pCDNA-KDM4B-EGFP were kindly provided by Dr Till Bartke (IFE Helmholtz Munich, Neuherberg, Germany) and Dr Thomas Jenuwein (Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany), respectively. To construct mCherry-PHF2, PHF2 was extracted from the p3xFLAG-PHF2 vector (kindly provided by Dr Jiemin Wong, East China Normal University, Shangai, China) through digestion with BamHI restriction enzyme and inserted into BamHI digested pCMV-mCherry-C1 plasmid. p3xFLAG-PHF2 ΔCharged vector was obtained from p3xFLAG-PHF2 through a PCR amplification with specific primer pairs (Table S1) so that the deleted region was excluded. mCherry-PHF2 ΔCharged plasmid was obtained by replacing the PHF2 from mCherry-PHF2 vector with the ΔCharged mutant version, which was extracted from the p3xFLAG-PHF2 ΔCharged digesting with BamHI. pCMV-mEGFP-JMJD3 was cloned as described in [9]. mCherry-MED15 plasmid was kindly provided by Dr. R. Young.

4.5. Antibodies

Antibodies used were anti: DAPI (ThermoFisher, D1306), β-TUBULIN (Millipore, MAB3408), H3K9me3 (Abcam, ab8898), H3K9me2 (Abcam, Abcam, 1220), H3K36me3 (Abcam, ab9050), PHF2 (Cell Signaling, D45A2), mCherry (ThermoFisher Scientific, MA5-32977) and GFP (Roche, 11814460001).

4.6. RNA extraction and qPCR

Total RNA was prepared using TRIZOL reagent (Invitrogen), following the manufacturer instructions. 200-1000ng of RNA were used for reverse transcription using High Capacity cDNA reverse transcription kit (Invitrogen). qPCR was performed with SYBR Green (Roche) in a QuantStudio™ 5 Real-Time PCR System (ThermoFisher Scientific, A34322) with specific primer pairs (Table S1).

4.7. Western Blot

Immunoblotting was performed following standard method. An ECL kit (Amersham) was used to visualize the results.

4.8. Droplet assays in nuclear extracts

$20x10^6$ HEK293T cells were transfected using 5μg of vector encoding KDMs cDNA fused to mEGFP or mCherry as previously described [69]. Nuclear extracts were prepared at 4mg/ml concentration. They were used for a droplet formation assay by 1:1 diluting them with droplet-buffer (10% glycerol, 20 mM HEPES). The final droplet buffer conditions were 20 mM HEPES, 150 mM NaCl, 15% glycerol, 3.75 mM EGTA, 2.5 mM MgCl2, 1.25 mM CaCl2. The reactions were incubated for 30 minutes and loaded onto a glass-bottom 384-well plate (Cellvis P384-1.5H-N) 5 minutes. Then, they were imaged on an Automated Inverted Microscope Leica Thunder 3D Live Cell using a 63x water immersion objective (NA=1.2).

4.9. Droplets liquid-like features quantification

A quantification of the shape descriptors "aspect ratio" and "roundness" of droplets was performed with the Fiji plugin "Analyze particles". The Fiji macro "Calculate Convexity and Solidarity" was used to quantify droplets "convexity" [32, 70]. Each droplet was individualized as an object by fixing the threshold. Results shown in the graph correspond to n=75 droplets.

4.10. 1. ,6-Hexanediol treatment for live imaging cells

$20x10^6$ HEK293T cells growing on glass dishes coated with 5 μg/ml of poly-D-lysine were transfected with 0.05μg of EGFP-KDMs and mCherry-PHF2 vectors. They were imaged on a 37ºC heated stage of a Zeiss LSM780 Confocal/Multiphoton using Zen software to determine a baseline using the X detector and a 40x water objective. After the fifth

acquisition, 1,6-Hexanediol (#240117, Sigma) was added to the culture medium at a final concentration of 6%, and images were again taken for 5 minutes. Raw images were analyzed using Fiji software for subsequent quantifications. Representative images of puncta disassembly at 60 and 120 seconds are presented.

4.11. Indirect immunofluorescence

Immunofluorescence assays were performed as previously described [71]. Basically, cells were fixed with 4% paraformaldehyde (20 minutes) and permeabilized with PBS-Triton X-100 (0.5%). They were blocked in 0.5% BSA (in PBS with 0.1% Triton X-100) for 1 hour at room temperature. Thus, they were incubated overnight at 4°C with primary antibodies. Next, cells were incubated with Alexa-conjugated secondary IgG antibodies (Jackson ImmunoResearch). DAPI (Sigma) was added for 2 hours at room temperature. The intrinsic fluorescence of EGFP and mCherry was imaged without using primary/secondary antibodies. Images were obtained using a Leica SP5 confocal microscope utilizing LAS-AF software.

4.12. Focus calling (Immunofluorescence, 1,6-Hexanediol treatment)

Foci were called using the "Object Counter 3D" plugin in Fiji. For each image, the "threshold" was set in the way that each focus could be seen as an individual object.
$$\Sigma$$

4.13. Statistical analysis

Quantitative data was expressed as mean and standard error mean (SEM) (for immunofluorescence counting and RNA transcription experiments). At least two to three biologically independent experiments were carried out for each experiment. The significance of differences was determined using the Student's t-test (*$p<0.05$; **$p<0.01$, *** $p<0.001$).

AUTHOR CONTRIBUTION: M.V. performed experiments, contributed to experimental design, analyzed the data and revised the paper; S.A. conducted experiments and revised the manuscript; M.A.M.B. conceived the project, analyzed the data and wrote the paper with editing support from all authors.

DATA AVAILABILITY: All relevant data supporting the key finding of this work are available in Supplementary Information and within the article.

DECLARATION OF COMPETING INTERESTS: The authors declare no competing interests.

REFERENCES

[1] Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016;17:487-500.

[2] Greer EL, Shi Y. Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet. 2012;13:343-57.

[3] Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21:381-95.

[4] Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell. 2004;119:941-53.

[5] Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, et al. Histone demethylation by a family of JmjC domain-containing proteins. Nature. 2006;439:811-6.

[6] Johansson C, Tumber A, Che K, Cain P, Nowak R, Gileadi C, et al. The roles of Jumonji-type oxygenases in human disease. Epigenomics. 2014;6:89-120.

[7] Black JC, Van Rechem C, Whetstine JR. Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol Cell. 2012;48:491-507.

[8] Turberfield AH, Kondo T, Nakayama M, Koseki Y, King HW, Koseki H, et al. KDM2 proteins constrain transcription from CpG island gene promoters independently of their histone demethylase activity. Nucleic Acids Res. 2019;47:9005-23.

[9] Vicioso-Mantis M, Fueyo R, Navarro C, Cruz-Molina S, van Ijcken WFJ, Rebollo E, et al. JMJD3 intrinsically disordered region links the 3D-genome structure to TGFbeta-dependent transcription activation. Nat Commun. 2022;13:3263.

[10] Black JC, Allen A, Van Rechem C, Forbes E, Longworth M, Tschop K, et al. Conserved antagonism between JMJD2A/KDM4A and HP1gamma during cell cycle progression. Mol Cell. 2010;40:736-48.

[11] Feng H, Lu J, Song X, Thongkum A, Zhang F, Lou L, et al. CK2 kinase-mediated PHF8 phosphorylation controls TopBP1 stability to regulate DNA replication. Nucleic Acids Res. 2020;48:10940-52.

[12] Khoury-Haddad H, Nadar-Ponniah PT, Awwad S, Ayoub N. The emerging role of lysine demethylases in DNA damage response: dissecting the recruitment mode of KDM4D/JMJD2D to DNA damage sites. Cell Cycle. 2015;14:950-8.

[13] Khoury-Haddad H, Guttmann-Raviv N, Ipenberg I, Huggins D, Jeyasekharan AD, Ayoub N. PARP1-dependent recruitment of KDM4D histone demethylase to DNA damage sites promotes double-strand break repair. Proc Natl Acad Sci U S A. 2014;111:E728-37.

[14] Alonso-de Vega I, Paz-Cabrera MC, Rother MB, Wiegant WW, Checa-Rodriguez C, Hernandez-Fernaud JR, et al. PHF2 regulates homology-directed DNA repair by controlling the resection of DNA double strand breaks. Nucleic Acids Res. 2020;48:4915-27.

[15] D'Oto A, Tian QW, Davidoff AM, Yang J. Histone demethylases and their roles in cancer epigenetics. J Med Oncol Ther. 2016;1:34-40.

[16] Swahari V, West AE. Histone demethylases in neuronal differentiation, plasticity, and disease. Curr Opin Neurobiol. 2019;59:9-15.

[17] Natoli G, Testa G, De Santa F. The future therapeutic potential of histone demethylases: A critical analysis. Curr Opin Drug Discov Devel. 2009;12:607-15.

[18] Hojfeldt JW, Agger K, Helin K. Histone lysine demethylases as targets for anticancer therapy. Nat Rev Drug Discov. 2013;12:917-30.

[19] Accari SL, Fisher PR. Emerging Roles of JmjC Domain-Containing Proteins. Int Rev Cell Mol Biol. 2015;319:165-220.

[20] Dimitrova E, Turberfield AH, Klose RJ. Histone demethylases in chromatin biology and beyond. EMBO Rep. 2015;16:1620-39.

[21] Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Exp Mol Med. 2017;49:e324.

[22] Cho WK, Spille JH, Hecht M, Lee C, Li C, Grube V, et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science. 2018;361:412-5.

[23] Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A Phase Separation Model for Transcriptional Control. Cell. 2017;169:13-23.

[24] Peng L, Li EM, Xu LY. From start to end: Phase separation and transcriptional regulation. Biochim Biophys Acta Gene Regul Mech. 2020;1863:194641.

[25] Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science. 2018;361.

[26] Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science. 2018;361.

[27] Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell. 2018;175:1842-55 e16.

[28] Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, Henry L, et al. Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell. 2019;179:470-84 e21.

[29] Gallego LD, Schneider M, Mittal C, Romanauska A, Gudino Carrillo RM, Schubert T, et al. Phase separation directs ubiquitination of gene-body nucleosomes. Nature. 2020;579:592-7.

[30] Fuchs G, Oren M. Writing and reading H2B monoubiquitylation. Biochim Biophys Acta. 2014;1839:694-701.

[31] Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell. 2015;57:936-47.

[32] Hyman AA, Weber CA, Julicher F. Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol. 2014;30:39-58.

[33] Pak CW, Kosno M, Holehouse AS, Padrick SB, Mittal A, Ali R, et al. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. Mol Cell. 2016;63:72-85.

[34] Fuxreiter M. Fuzziness in Protein Interactions-A Historical Perspective. J Mol Biol. 2018;430:2278-87.

[35] Dunker AK, Bondos SE, Huang F, Oldfield CJ. Intrinsically disordered proteins and multicellular organisms. Semin Cell Dev Biol. 2015;37:44-55.

[36] Wallmann A, Kesten C. Common Functions of Disordered Proteins across Evolutionary Distant Organisms. Int J Mol Sci. 2020;21.

[37] Shi B, Li W, Song Y, Wang Z, Ju R, Ulman A, et al. UTX condensation underlies its tumour-suppressive activity. Nature. 2021;597:726-31.

[38] Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18:285-98.

[39] Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, et al. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. Nat Struct Mol Biol. 2019;26:637-48.

[40] Wheeler RJ, Hyman AA. Controlling compartmentalization by non-membrane-bound organelles. Philos Trans R Soc Lond B Biol Sci. 2018;373.

[41] Mir M, Bickmore W, Furlong EEM, Narlikar G. Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? Development. 2019;146.

[42] Shi Y, Chen J, Zeng WJ, Li M, Zhao W, Zhang XD, et al. Formation of nuclear condensates by the Mediator complex subunit Med15 in mammalian cells. BMC Biol. 2021;19:245.

[43] Pappa S, Padilla N, Iacobucci S, Vicioso M, Alvarez de la Campa E, Navarro C, et al. PHF2 histone demethylase prevents DNA damage and genome instability by controlling cell cycle progression of neural progenitors. Proc Natl Acad Sci U S A. 2019;116:19464-73.

[44] Niewidok B, Igaev M, Pereira da Graca A, Strassner A, Lenzen C, Richter CP, et al. Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. J Cell Biol. 2018;217:1303-18.

[45] Alberti S, Dormann D. Liquid-Liquid Phase Separation in Disease. Annu Rev Genet. 2019;53:171-94.

[46] Taniue K, Akimitsu N. Aberrant phase separation and cancer. FEBS J. 2021.

[47] Zbinden A, Perez-Berlanga M, De Rossi P, Polymenidou M. Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. Dev Cell. 2020;55:45-68.

[48] Di Pierro M, Zhang B, Aiden EL, Wolynes PG, Onuchic JN. Transferable model for chromosome architecture. Proc Natl Acad Sci U S A. 2016;113:12168-73.

[49] Frank L, Rippe K. Repetitive RNAs as Regulators of Chromatin-Associated Subcompartment Formation by Phase Separation. J Mol Biol. 2020;432:4270-86.

[50] Wang N, Liu C. Implications of liquid-liquid phase separation in plant chromatin organization and transcriptional control. Curr Opin Genet Dev. 2019;55:59-65.

[51] Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017;547:241-5.

[52] Watson M, Stott K. Disordered domains in chromatin-binding proteins. Essays Biochem. 2019;63:147-56.

[53] Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, et al. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell. 2013;155:1049-60.

[54] Cooper DG, Fassler JS. Med15: Glutamine-Rich Mediator Subunit with Potential for Plasticity. Trends Biochem Sci. 2019;44:737-51.

[55] Cramer P. Organization and regulation of gene transcription. Nature. 2019;573:45-54.

[56] Guo YE, Manteiga JC, Henninger JE, Sabari BR, Dall'Agnese A, Hannett NM, et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature. 2019;572:543-8.

[57] Shrinivas K, Sabari BR, Coffey EL, Klein IA, Boija A, Zamudio AV, et al. Enhancer Features that Drive Formation of Transcriptional Condensates. Mol Cell. 2019;75:549-61 e7.

[58] Sharma R, Choi KJ, Quan MD, Sharma S, Sankaran B, Park H, et al. Liquid condensation of reprogramming factor KLF4 with DNA provides a mechanism for chromatin organization. Nat Commun. 2021;12:5579.

[59] Lee JH, Yoo NJ, Kim MS, Lee SH. Histone Demethylase Gene PHF2 Is Mutated in Gastric and Colorectal Cancers. Pathol Oncol Res. 2017;23:471-6.

[60] Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. J Bioinform Comput Biol. 2005;3:35-60.

[61] Dosztanyi Z, Csizmok V, Tompa P, Simon I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J Mol Biol. 2005;347:827-39.

[62] Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics. 2006;7:208.

[63] Bolognesi B, Lorenzo Gotor N, Dhar R, Cirillo D, Baldrighi M, Tartaglia GG, et al. A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. Cell Rep. 2016;16:222-31.

[64] Piovesan D, Tabaro F, Paladin L, Necci M, Micetic I, Camilloni C, et al. MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins. Nucleic Acids Res. 2018;46:D471-D6.

[65] Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31:3784-8.

[66] Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A. 1981;78:3824-8.

[67] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389-402.

[68] Blanco-Garcia N, Asensio-Juan E, de la Cruz X, Martinez-Balbas MA. Autoacetylation regulates P/CAF nuclear localization. J Biol Chem. 2009;284:1343-52.

[69] Zamudio AV, Dall'Agnese A, Henninger JE, Manteiga JC, Afeyan LK, Hannett NM, et al. Mediator Condensates Localize Signaling Factors to Key Cell Identity Genes. Mol Cell. 2019;76:753-66 e6.

[70] Weber SC, Brangwynne CP. Getting RNA and protein in phase. Cell. 2012;149:1188-91.

[71] Estaras C, Akizu N, Garcia A, Beltran S, de la Cruz X, Martinez-Balbas MA. Genome-wide analysis reveals that Smad3 and JMJD3 HDM co-activate the neural developmental program. Development. 2012;139:2681-91.