

学校代码: 10730

分类号:

密级:



兰州大学

博 士 学 位 论 文

(学 术 学 位)

小鼠多梳蛋白剪接体 CBX7C 通过相分离驱

论文题目 (中文) 动功能性 cPRC1 组装并调控细胞分化

Mouse polycomb protein isoform CBX7C drives

论文题目 (外文) functional cPRC1 assembly through phase
separation and regulates the cell differentiation

作 者 姓 名 关珊丽

学 科 专 业 生物学·细胞生物学

研 究 方 向 表观遗传学

教 育 类 型 学历教育

指 导 教 师 程博 教授

论 文 工 作 时 段 2017 年 9 月至 2024 年 5 月

答 辩 日 期 2024 年 5 月

校址: 甘肃省兰州市城关区天水南路 222 号

缩略词表

缩写	英文名称	中文名称
ATH	AT-Hook	AT 序列结合结构域
ATHL	AT-Hook like	类 AT 序列结合结构域
CBX	Chromobox homolog protein	染色质同源框蛋白
CD	Chromodomain	染色质结合结构域
CGI	CpG island	CpG 岛
cPRC1	canonical PRC1	经典多梳蛋白抑制复合体 1
EB	Embryoid body	类胚体
IDR	Intrinsically disordered protein region	内在无序区域
Hox	Homeotic gene	同源异型基因
HMGA	High mobility group A protein	高迁移率家族 A 蛋白
HMT	Histone methyl transferase	组蛋白甲基转移酶
HTM	Heterochromatic targeting module	异染色质靶向模块
LCR	Low-complexity region	低复杂度区域
LLPS	liquid-liquid phase separation	液液相分离
lncRNA	long non-coding RNA	长链非编码 RNA
mESC	Mouse embryonic stem cell	小鼠胚胎干细胞
ncPRC1	non-canonical PRC1	非经典多梳蛋白抑制复合体 1
PB	Polycomb body	多梳小体
PC	Polycomb repressor box	多梳抑制框结构域
PCGF4/6	Polycomb group RING finger 4/6	多梳环指蛋白 4/6
PCL	Polycomb-like proteins	多梳样蛋白
PcG	Polycomb group	多梳蛋白
PHC2	Polyhomeotic homolog 2	同源多聚物 2
PRC1/2	Polycomb group repressive complex 1/2	多梳蛋白抑制复合体 2

缩写	英文名称	中文名称
RING1A/B	RING finger protein 1A/B	环指蛋白 2
RNP	Ribonucleoprotein	核糖核蛋白
PRE	Polycomb response element	多梳蛋白响应元件
SGs	Stress granules	应激颗粒
TFs	Transcription factors	转录因子
TRE	Trithorax response element	三胸蛋白响应元件
TrxG	Trithorax group	三胸蛋白家族
TSS	Transcription start site	转录起始位点

小鼠多梳蛋白剪接体 CBX7C 通过相分离驱动功能性 cPRC1 组装并调控细胞分化

中文摘要

多细胞生物的生长发育需要通过分化获得形态和功能各异的细胞类群,并维持其正常的生物学特性。在特定的细胞类型中,特异性基因的表达对于细胞正常生物学功能的发挥至关重要,而其它细胞类型相关的基因则被稳定抑制。多梳蛋白家族 PcG (Polycomb group, PcG) 介导的表观遗传调控使其靶基因在分化过程中以细胞类型特异的表达模式被抑制,在发育和维持特定细胞状态中发挥重要作用,是调控细胞命运决定和发育过程中诸多重要细胞生物学过程的关键机制之一。PcG 蛋白家族中主要包括了两类多梳蛋白抑制复合体, PRC1 和 PRC2 (Polycomb group repressive complex 1/2), 它们彼此协同抑制基因表达并维持部分基因的沉默状态。PRC1 蛋白复合体识别并结合目标染色质主要依赖于含有染色质结合结构域(Chromodomain, CD)的染色质同源框蛋白(Chromobox homolog protein, CBX) 家族蛋白对组蛋白甲基化修饰标签的特异性识别并结合。CBX7 蛋白是哺乳动物胚胎干细胞(embryonic stem cells, ESC)中 PRC1 蛋白复合体的核心组分之一,介导对 H3K27me3 的识别和结合,进而在靶基因上装配 PRC1 并执行转录抑制调控。

本论文基于真核生物基因转录的选择性剪接现象,通过多种生理生化、细胞生物学及分子生物学方法探究了 CBX7 的三种不同蛋白亚型(CBX7A、CBX7B 和 CBX7C)的表达模式差异、参与复合体组装基因调控的分子机制、相分离性质及在小鼠 ESC (mESC) 分化过程中的独特生物学功能。

首先,研究发现 CBX7C 蛋白亚型主要在 mESC 和小鼠脾脏组织中特异性表达。与 CBX7A 和 CBX7B 相比,仅保留了介导染色质结合的 N 端 CD 结构域,而缺失了与 RING1B 互作并介导参与 cPRC1 复合体组装的 C 端 PC 结构域,并由于基因转录的选择性剪接而产生独特的 C 端区域。

其次, cPRC1 复合体组装的分子机制具有复杂性和多样性。本论文研究发现丧失了 RING1B 结合活性的 CBX7C 能够通过和 PHC2 之间的互作而参与完整的功能性 cPRC1 复合体在染色质上的组装,并作为表观遗传调控因子参与靶基因的转录抑制调控。利用独特的复合体组装机制研究实验系统(HTM 介导的异位

靶向系统),证实了CBX7能够作为“点火装置”招募并促进其它复合体组分在靶基因位点的组装起始。在复合体组装过程中,CBX7C能够作为CBX7A的竞争者与PHC2互作并替代其结合靶基因参与转录调控。

再次,在mESC中,表达水平相对较低的CBX7C对维持细胞正常生长具有重要作用。CBX7C蛋白的敲低能够导致严重的mESC生长缺陷。与维持mESC干性的CBX7A/B不同,CBX7C作为促进因子在mESC分化为类胚体(Embryoid body, EB)的过程中,其表达水平与PHC2协同显著增加,与EB的生长速率及状态呈正相关。

最后,研究发现含有IDRs基序的CBX7C能够驱动相分离的发生。CBX7C•PHC2凝集体的固有性质因其细胞剂量而异。在正常浓度下,形成了具有催化活性的Pc小体。而当浓度高于阈值时,发生了液-固相分离,形成致密的排斥染色质的水凝胶状结构。CBX7C的N端主要介导CBX7C•PHC2的二聚体形成,而C端主要在促进具有水凝胶性质的CBX7C•PHC2凝集体的形成中发挥作用。

此外,在mESC中共表达的CBX2和CBX7通过ATH或ATHL结构域而发生相互作用,并以此为分子基础共同结合至靶基因而发挥生物学功能。

综上所述,本论文揭示了CBX7C与PHC2的互作在PRC1在染色质上的组装及Pc小体形成过程中的独特作用和分子基础,强调了CBX7C在mESC分化中的重要作用,并开辟了一条通过调节其相分离性质来控制PRC1活性的新途径,揭示了含有CBX7C的cPRC1复合体执行表观抑制的新机制。研究结果不仅为进一步阐明相分离在转录调控中的功能提供了重要信息,而且为深入解析mESC中共存的不同CBX-PRC1复合体协同调控靶基因转录抑制的分子机制提供精确的分子生化证据。

关键词: PRC1, CBX 蛋白, CBX7C, 相分离; 胚胎干细胞, 表观遗传调控

Mouse polycomb protein isoform CBX7C drives functional cPRC1 assembly through phase separation and regulates the cell differentiation

Abstract

The growth and development of multicellular organism need to differentiate into cell groups with different shapes and functions, and maintain their normal biological characteristics. In specific cell types, the expression of specific genes is crucial for the normal biological function of cells, while other cell type related genes are stably suppressed. The epigenetic regulation mediated by the PcG allows its target genes to be appropriately repressed in cell type specific expression patterns during differentiation, playing an important role in development and maintaining specific cell states. It is one of the key mechanisms that regulate many important cellular biological processes in cell fate determination and development. The PcG protein family mainly includes two types of Polycomb group repressive complex 1/2, PRC1 and PRC2, which synergistically repress gene expression and maintain a state of partial gene silencing. The process of PRC1 protein complex recruiting to target chromatin mainly depends on the recognition and binding of specific histone modification by CBX proteins containing conservative CD domain. CBX7 is one of the core components of the PRC1 complex in mammalian ESCs, which mediates the recognition and binding of H3K27me3 to assemble PRC1 on target genes and perform transcriptional repression regulation.

Based on the Alternative splicing of eukaryotic gene transcription, we have explored the differential expression patterns of three different protein subtypes of CBX7 (CBX7A, CBX7B, and CBX7C), molecular mechanisms involved in cPRC1 complex assembly and gene regulation, phase separation properties, and unique biological functions in mouse embryonic stem cell (mESC) differentiation through various physiological, biochemical, cell biology, and molecular biology methods.

Firstly, the CBX7C protein subtype we focus on is mainly specifically expressed in mESCs and mouse spleen tissue. Compared with CBX7A and B, CBX7C is only

retained the N-terminal CD domain that mediates chromatin binding, while the C-terminal PC domain that interacts with RING1B and participates in cPRC1 complex assembly is missing. CBX7C acquires a unique C-terminal region due to Alternative splicing of gene transcription.

Secondly, the molecular mechanism of cPRC1 complex assembly is complex and diverse. CBX7C, which has lost the binding activity of RING1B, has been found that can participate in the assembly of the complete functional cPRC1 complex on chromatin through interaction with PHC2, and act as an epigenetic regulatory factor in transcriptional repression regulation of target genes. Through a unique complex assembly mechanism experimental system (HTM mediated ectopic targeting system), it has been confirmed that CBX7 can act as an “ignition device” to recruit and promote the assembly initiation of other complex components at target gene sites. During the assembly process of the cPRC1 complex, CBX7C, act as a competitor to CBX7A/B, interact with PHC2, replacing its binding target genes to participate in transcriptional regulation.

Again, in mESCs, CBX7C, with the relatively low expression level, plays an important role in maintaining normal cell growth. Knockdown of CBX7C protein can lead to severe cell growth defects. Unlike CBX7A/B, which maintains the stemness of mESC, CBX7C, as a promoting factor, is up-regulates significantly in synergy with PHC2 during the differentiation of mESC into embryoid bodies (EBs), and is positively correlated with the growth rate and status of EBs.

Finally, we found that CBX7C containing IDRs motifs can drive phase separation. The inherent properties of CBX7C•PHC2 condensates vary depending on their cellular dosage. At normal concentrations, Pc bodies with catalytic activity were formed. When the concentration is higher than the threshold value, liquid-solid separation occurs, forming a dense hydrogel like structure that excludes chromatin. The N-terminal of CBX7C mainly mediates the formation of CBX7C•PHC2 dimer, while the C-terminal mainly plays a role in promoting the formation of CBX7C•PHC2 condensates with hydrogel properties.

In addition, CBX2 and CBX7C co-existing in mESC interact with each other through the ATH or ATHL domain, and jointly bind to target genes and exert biological functions.

In summary, the study reveals the unique role and molecular basis of the

interaction between CBX7C and PHC2 in the assembly of cPRC1 on chromatin and the formation of Pc body, emphasizes the important role of CBX7C in mESC differentiation, and open a new avenue of controlling PRC1 activities via modulation of its phase separation properties. It reveals a new mechanism of epigenetic repression in CBX7C-cPRC1, providing momentous information for further defining the function of phase separation in transcriptional regulation, and providing precise molecular mechanisms for exploring the synergistic regulation of target gene transcription repression by different CBX-PRC1 complexes co-existing in mESCs.

Keywords: PRC1, CBX proteins, CBX7C, Phase Separation, ESCs, Epigenetic Regulation

目 录

第一章 前言.....	1
1.1 PcG 的研究现状及进展.....	1
1.2 PRC1 蛋白复合体简介.....	3
1.2.1 PRC1 的基本组成和功能.....	3
1.2.2 PRC1 蛋白复合体的组装.....	3
1.2.3 PRC1 和 PRC2 介导的组蛋白修饰及招募机制.....	6
1.3 CBX 蛋白的研究进展.....	8
1.3.1 CBX 蛋白的功能结构域研究.....	8
1.3.2 CBX 蛋白不同转录剪接体研究.....	12
1.3.3 CBX 蛋白在 ESC 分化中的时空交替及生理意义.....	16
1.3.4 CBX 蛋白的表达异常诱发机体发育缺陷及癌变的研究.....	20
1.4 蛋白质相分离在表观遗传调控中的生物学意义.....	21
1.4.1 相分离的生物学背景.....	21
1.4.2 相分离研究方法的进展.....	23
1.4.3 不同 CBX 蛋白及 PRC1 组分 PHC2 相分离的研究进展.....	25
1.5 本研究的立项依据与意义.....	27
1.5.1 本研究的立项依据.....	27
1.5.2 本研究的主要内容.....	28
1.5.3 本论文的技术路线.....	29
1.5.4 本研究的意义.....	30
第二章 材料与方法.....	31
2.1 实验材料.....	31
2.1.1 实验所用细胞、菌株及质粒载体.....	31
2.1.2 实验试剂.....	33
2.1.3 细胞培养试剂与耗材.....	35
2.1.4 实验仪器与耗材.....	35
2.2 实验方法.....	36
2.2.1 分子克隆.....	36
2.2.2 组织及细胞 RNA 提取、逆转录及 qPCR 检测.....	40
2.2.3 可诱导性表达 CBX7C-EGFP-Flag 细胞系构建.....	41
2.2.4 双分子荧光互补实验(BiFC).....	41
2.2.5 pull down 实验(GST-pull down 和 His-pull down).....	41
2.2.6 蛋白质免疫共沉淀实验(co-IP).....	43

2.2.7 免疫荧光.....	44
2.2.8 HTM 介导的细胞内异位靶向实验.....	45
2.2.9 原核蛋白诱导表达及纯化.....	45
2.2.10 纯化蛋白体外凝集实验分析.....	46
2.2.11 荧光漂白后恢复实验分析(FRAP).....	47
2.2.12 Cry2oligo 介导的光激活细胞内相分离实验	47
2.2.13 高分辨率活细胞成像系统(Delta Vision)实验.....	47
2.2.14 染色质免疫共沉淀(Ch IP)及 q PCR 实验分析	48
2.2.15 体外 EB 诱导实验及分析	51
2.2.16 双荧光素酶报告基因系统实验分析.....	52
2.2.17 数据统计与分析.....	52
第三章 实验结果	53
第一节 CBX7C 与 PHC2 蛋白互作促进功能性 cPRC1 组装并调节相分离特性	53
3.1 Cbx7 剪接体及其蛋白在 m ESC 及组织中的时空表达模式分析	53
3.2 丧失了与 RING1B 结合活性的 CBX7C 仍参与完整的 cPRC1 复合体组装	54
3.3 CBX7C-cPRC1 复合体识别靶基因并参与基因表达的表观遗传抑制调控	58
3.4 CBX7C 驱动功能性 cPRC1 复合体在染色质上的组装起始	62
3.5 CBX7 蛋白亚型具有不同的相分离能力	67
3.6 PHC2 与 CBX7A 或 CBX7C 的蛋白互作调节相分离特性.....	71
3.7 CBX7C 在 m ESC 分化过程中上调表达并参与调控 EB 的形成	84
第二节 CBX7 与 CBX2 蛋白互作并协同参与靶基因的表观遗传抑制调控	91
3.8 CBX7 与 CBX2 蛋白存在相互作用	91
3.9 ATH 或 ATHL 结构域介导了 CBX7 与 CBX2 蛋白的相互作用	93
3.10 ATH 或 ATHL 结构域的缺失降低了 CBX7 或 CBX2 蛋白在靶基因上的富集	96
小结.....	97
第四章 讨论与展望	98
4.1 CBX 蛋白亚型具有结构保守性与功能多样性	98
4.2 CBX7 的相分离影响 PRC1 功能及细胞生长状态	100
4.3 CBX7C 的蛋白表达及其相分离影响 m ESC 的生长状态及分化	101

第五章 结论.....	105
参考文献.....	107
附录.....	119

第一章 前言

表观遗传调控因子通过复杂而精密的分子调控网络在真核细胞的命运决定中发挥关键作用。多梳蛋白家族 (Polycomb group, PcG) 作为重要的基因表观遗传调控蛋白参与并维持目标基因的转录沉默, 在哺乳动物 ESC 的分化过程中参与构建及维持“细胞记忆”, 并产生可遗传的生物学表型^[1]。近年来, 关于 PcG 在表观遗传调控中的复杂分子机理、精细的调控网络及重要的生物学功能的深入探究逐渐成为表观遗传学研究的热点。近期多项研究表明 PcG 的一些核心组成亚基可以发生相分离, 在核内形成液滴状结构, 但是其参与转录调控的分子机制及其生物学意义尚不明确。

1.1 PcG 的研究现状及进展

染色质环境是维持不同细胞类型的基因特异性表达模式和细胞身份识别的关键影响因素, 而基因的表观遗传调控是发育过程中调节细胞命运选择的核心机制之一。不同的表观遗传调控蛋白可以维持或改变染色质的可及性和局部染色质的修饰模式^[2]。PcG 蛋白家族是一类进化保守的转录调控抑制因子, 与活性转录相关的三胸蛋白家族 (Trithorax group, TrxG) 作为发育进程的“总开关”, 在细胞分化和发育过程中起拮抗作用, 以协调关键基因的表达调控, 是大部分高等多细胞生物正常发育所必需的。

多梳抑制基因 (*Polycomb, Pc*) 是最初在黑腹果蝇中发现的第一个 PcG 家族基因^[3], 该基因的突变导致了同源异型基因 (Homeotic gene, *Hox*) 的异位表达而将前胚胎段转化为后胚胎段。随后的研究中, 通过基因筛查进一步发现了诸多相关基因, 这些基因的突变出现了与 *Pc* 突变而导致的功能丧失相似的表型, 并且发现除了正常定位于果蝇前足的性梳外, 果蝇后足也异常出现了性梳结构, 从而定义该家族蛋白为多梳蛋白家族 PcG^[3]。此后, *Hox* 基因表达调控的正向调节因子 TrxG 的基因 *Trithorax* 及其它编码 PcG 蛋白拮抗因子的基因从果蝇中分离出来, 进一步研究发现这些基因的突变导致胚胎片段通过拮抗 PcG 蛋白而转化为更前部的片段, 并与 *Hox* 功能丧失表型一致^[4-6], 因此, 将这些与 PcG 拮抗的蛋白定义为 TrxG 蛋白家族^[7]。近年来诸多研究表明, PcG 家族蛋白与 TrxG 家族蛋白不仅是 *Hox* 基因表达的表观遗传调控“守门人”, 它们还参与调节多个重要的细胞生物学过程, 包括 X 染色体失活、基因组印记、细胞周期控制、干细胞生

生物学和癌症发生等^[8-10]。蛋白的这种功能多样性是通过多种 PcG 和 TrxG 复合物在不同发育阶段以细胞特异性组装的方式实现的,包括通过组蛋白修饰或染色质重塑活动而修饰其靶基因处的染色质。尽管 PcG 和 TrxG 的靶基因在进化过程中是保守的,但这些复合物靶向染色质的方式在物种间可能存在显著的差异性。在果蝇中,多梳响应元件(Polycomb response elements, PRE)和三胸响应元件(Trithorax response elements, TRE)分别介导了 PcG 和 TrxG 复合物靶向染色质的过程,从而在整个发育过程中驱动染色质的沉默或激活状态的转换,发挥表观遗传调控的生物学功能^[11]。近年来研究表明,PcG 和 TrxG 复合物可以从改变局部染色质结构到调节高阶染色质组织和全基因组结构的多个层面调节其目标基因表达^[12]。自从 PcG 和 TrxG 蛋白家族的相关基因相继从果蝇中被分离克隆后,它们所介导染色质状态的表观遗传现象、分子机制和生物学功能的研究迅速成为研究热点。在过去十几年间,研究人员主要聚焦在其结构组成的多样性和生化互作关系、对染色质的募集机制、参与调节基因组功能的机制、在传递染色质状态遗传中的作用机制以及在正常和致病细胞生长过程中的意义,特别是胚胎干细胞生物学和癌症发生中的功能(图 1-1)。

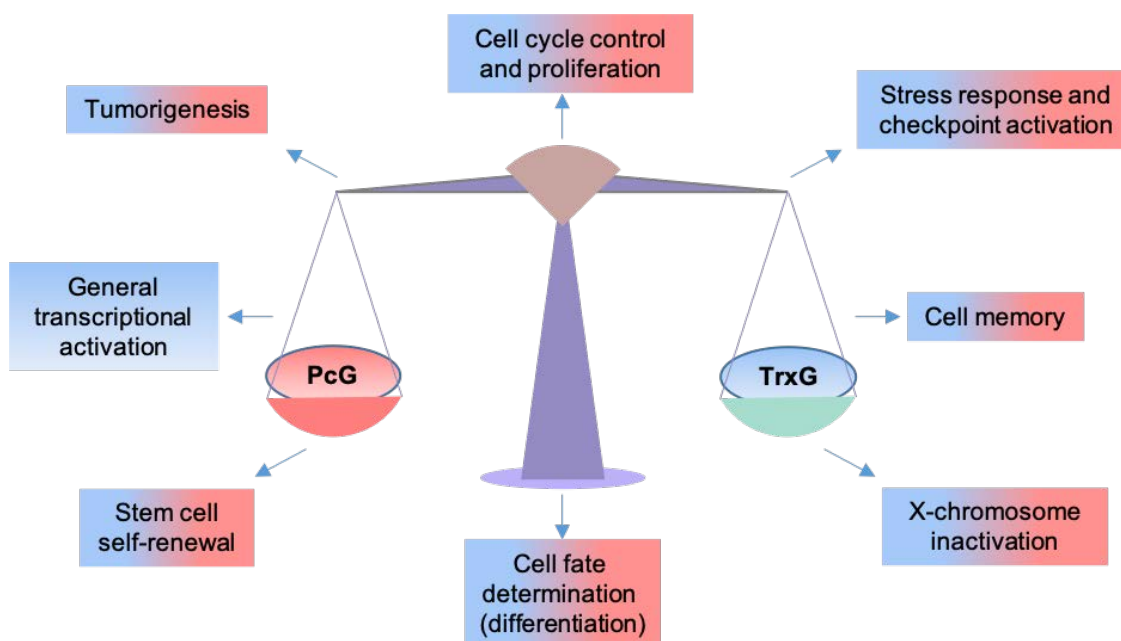


图 1-1 PcG 蛋白家族的生物学功能及其与 TrxG 蛋白家族的拮抗调控关系^[13]

Figure 1-1 The molecular function of PcG complexes and antagonistic action of TrxG ^[13]
PcG 蛋白家族通常与 TrxG 蛋白家族互为拮抗关系,参与许多发育相关基因的动态调节过程,在细胞命运决定和关键生物学过程调控中发挥重要作用。

具有酶催化活性的 PcG 蛋白家族能够通过组蛋白的修饰来抑制特定基因位点的转录^[14],进而其充当转录阻遏物,介导目标基因的表观遗传调控。PcG 蛋

白家族主要分为两种功能协同的蛋白质复合物,包括多梳抑制复合物 PRC1 和 PRC2。其中,PRC2 催化靶基因位点的组蛋白 H3K27 的三甲基化修饰 (H3K27me₃),而 PRC1 能够识别并结合该位点。与此同时,PRC1 复合体的核心催化亚基 RING1B 相应的染色质位点组蛋白 H2AK119 的单泛素化修饰 (H2AK119ub),两者协同抑制目标基因的表达并维持目标基因沉默状态^[14, 15]。

1.2 PRC1 蛋白复合体简介

1.2.1 PRC1 的基本组成和功能

参与 PRC1 复合体构建的组分及其组合具有复杂性和多样性。在果蝇中,PRC1 复合体主要由 Pc (Polycomb)、Ph (Polyhomeotic)、Psc (Posterior sex combs) 以及 Sce (Sex combs extra 或称为 RING) 几种核心组分构建而成^[12]。果蝇 PRC1 复合体的各组分在生物进化过程中衍生出哺乳动物细胞中表达的多种同系物,包括 CBX 蛋白 (CBX2、4、和 6-8)、RING1 蛋白 (RING1A/B)、PHC 蛋白 (PHC1-3) 以及 PCGF 蛋白 (PCGF1-6) ^[16],这些组分以特定方式搭配组成哺乳动物中经典 PRC1 复合体 (canonical PRC1, cPRC1) (图 1-2)。其中,多梳同源多聚蛋白 (Polyhomeotic-like, Phc) 也具有一定的组蛋白识别及结合能力,在非经典的 PRC1 复合体 (non-canonical PRC1, ncPRC1) 在染色质位点上装配过程中作为 DNA 结合蛋白发挥功能^[17]。PRC1 复合体具有 E3 泛素连接酶活性,该活性单位由 RING1 及 PCGF 的 RING 结构域上的亲水性基团构成^[18, 19],因而具有催化组蛋白 H2AK119ub 修饰的酶学活性^[20]。研究表明,PRC1 复合体能够通过改变染色质凝缩状态以及染色质重塑进而抑制基因的表达^[21]。在果蝇及人类细胞中的发育相关基因位点上结合有大量的处于转录延伸暂停状态的 RNA 聚合酶 II (poised Pol II),并且这些基因的启动子区域通常积累一定水平的 H3K27me₃、H3K9ac 和 H3K14ac 等组蛋白修饰^[22, 23]。例如,RING1B 与 PCGF4 蛋白组合形成的异二聚体能够催化组蛋白形成 H2AK119ub 修饰^[24],进而导致 RNA 聚合酶 II 在靶基因的转录起始位点 (Transcriptional start site, TSS) 下游停滞,引发转录延伸的暂停^[25]。近年来研究发现,除上述 cPRC1 复合体外,哺乳动物细胞中还存在多种含有 RING 蛋白而缺失 CBX 蛋白的 ncPRC1 复合体亚型^[26, 27]。

1.2.2 PRC1 蛋白复合体的组装

在复杂的细胞内环境中,不同的蛋白通过分子间作用力等方式精准地组装成蛋白复合物后,才具有有效的功能并参与不同的细胞生物学过程。尽管 cPRC1 或

ncPRC1 蛋白复合物的不同分子组成已经被阐明，但是其各蛋白亚基不同的生物学功能及蛋白复合体组装的分子机制尚不完全清楚。

cPRC1 蛋白复合体对目标基因的靶向主要依赖于 CBX 家族蛋白识别并结合特异性组蛋白修饰的能力^[28-32]，而 RING1B 与 CBX 蛋白的相互作用成为调节 PRC1 蛋白复合体成员与核小体之间结合的接头机制（图 1-2）。RING1B 蛋白包含了 RING 结构域（N 端）和与 E3 泛素连接酶活性相关的 RAWUL 结构域（C 端）^[33]。研究表明，RAWUL 结构域不仅能够介导 RING1B 蛋白二聚体的形成，而且 RING1B 蛋白与 CBX 蛋白保守的 C 端 PC 结构域互作^[34, 35]。由于异型相互作用比 RING1B 同二聚作用更强，所以 RAWUL 结构域与 CBX7 蛋白的结合能够阻止 RING1B 蛋白二聚体的形成^[36, 37]，这也增加了含 RAWUL 结构域的 PRC1 蛋白复合体之间发生异源寡聚的可能性。

cPRC1 蛋白复合体的 E3 连接酶活性是由具有 RING 结构的蛋白异二聚体配对而获得的^[38, 39]。当与六类 PCGF 蛋白亚基之一配对时^[20, 24, 40]，PRC1 的 RING1A/B 亚基具有了 E3 连接酶活性。PCGF4 蛋白由三个不同区域组成，包括 N 端的 RING 结构域、中心结构域以及 C 端富含脯氨酸或丝氨酸的 UBL (ubiquitin-like domain) 结构域，其主要参与蛋白稳定性的调节^[41]。研究表明，PCGF4 自身没有泛素连接酶活性，但是其 RING 结构域与 RING1A/B 蛋白亚基互作并形成 PRC1 异二聚体后^[18, 19]，能够增强 RING1A/B 的稳定性及其催化 H2AK119 泛素化修饰的酶活性^[18, 42]。PCGF4 蛋白的 C 端 UBL 结构域也被称为 RAUL (RING finger-and WD40-associated ubiquitin-like) 结构域^[33]。RAUL 结构域构象折叠后参与了与 PHC1、PHC2 和 PHC3 的相互作用^[43, 44]。据文献报道，PCGF4 蛋白的 RAUL 结构域在维持 PRC1 复合体内的蛋白互作中发挥双重作用。首先，RAUL 结构域与 cPRC1 复合体亚基 PHC2 的 HD1 结构域互作，形成 β 发夹结构，涉及 PRC1 复合体的组装机制；其次，RAUL 结构域还参与了 PCGF4 蛋白低聚化^[45]。而 RAUL 结构域介导的这两种蛋白互作行为对 PCGF4 的生物学功能至关重要，如果蝇的体轴发育调控^[43]和致癌活性^[46-48]等。此外，序列比对表明在 PCGF 家族蛋白中，PCGF4 的 RAUL 结构域与 PCGF2 的序列同源性最高。与 PCGF4-PHC2 相比，尽管 PCGF2•PHC2 结合的亲和力略有降低，但 PCGF2 仍能够与 PHC2 蛋白互作，并且结构分析表明 PCGF4•PHC2 与 PCGF2•PHC2 具有相似的结合模式^[45]。

Ph 蛋白属于 cPRC1 蛋白复合体的核心亚基之一，它包含了 N 端的 HD1 结构域、C 端的 SAM 结构域以及邻近 HD1 结构域的 FSC 锌指结构域。研究表明，HD1 结构域介导了 PHC2 与 RING1B 和 PCGF2 的相互作用^[49]。Matteo Colombo 等人通过交联质谱 (crosslinking mass spectrometry, XL-MS) 分析发现，PHC2 的

FCS 锌指结构域能够与 RING1B 的 RING 结构域和 PCGF2 的 RAUL 结构域相互作用，并且 PHC1 的 SAM (sterile alpha motif) 结构域与 PCGF2 的 RAUL 结构域也存在相互作用^[50]。Ph 蛋白最显著的特征是 C 端含有一个保守的 SAM 结构域，该结构域能够组装成首尾相接的螺旋状聚合物^[51]。SAM 结构域存在于许多不同类型的蛋白质中，并且在许多情况下可以介导蛋白质的聚合^[52]。果蝇 Ph 蛋白的 SAM 结构域介导的蛋白完全聚合活性对其参与基因沉默至关重要^[53, 54]，并且促进了 PRC1 的整体酶促活性。在果蝇和哺乳动物细胞中，PRC1 蛋白复合体形成的斑点^[49, 55]，以及在果蝇细胞中形成大量的衍射受限团簇^[56]也与 SAM 结构域介导的蛋白质聚合相关。此外，SAM 结构域的突变会削弱 PcG 蛋白簇的形成，并减少 PcG 结合位点之间的长距离接触^[49, 56]，表明 Ph-SAM 在 PcG 蛋白和染色质的大规模组织以及基因调控中具有关键功能。

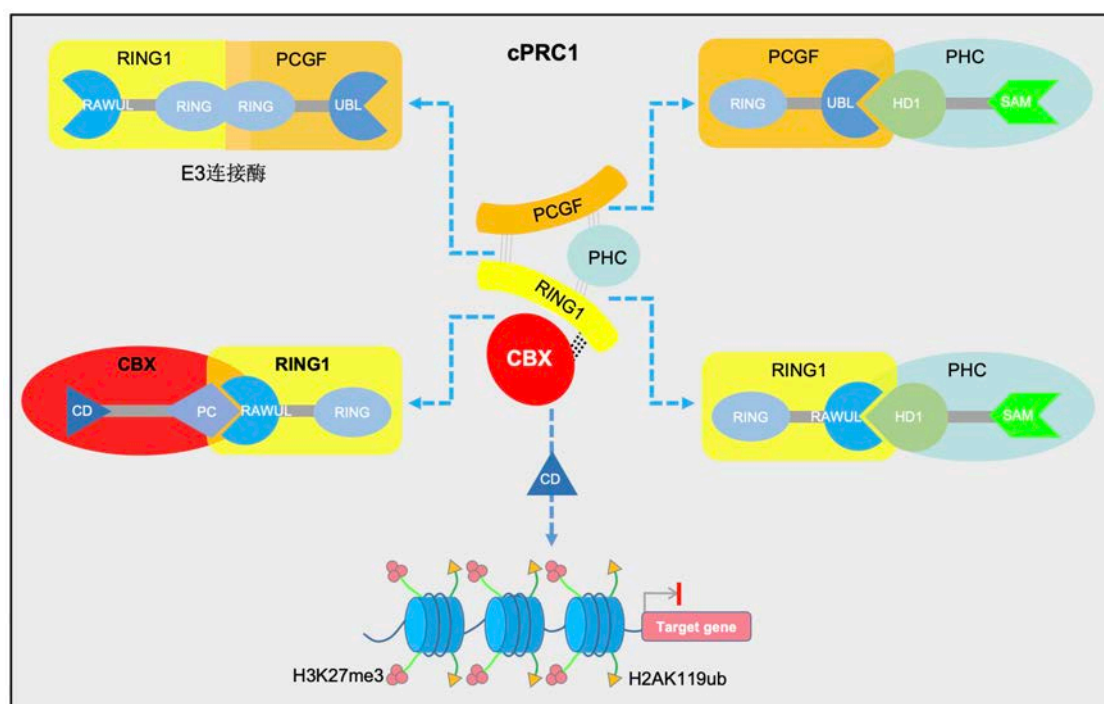


图 1-2 cPRC1 复合体的核心组分互作并识别催化靶位点的示意图

Figure 1-2 Schematic diagram of core component interactions and recognition and catalytic target sites in cPRC1 complex

cPRC1 蛋白复合体由 CBX、RING1、PHC 和 PCGF 通过相互作用组装而成，并通过 PRC2 催化的 H3K27me3 招募至靶基因，随后催化 H2AK119ub。

在上述 cPRC1 蛋白复合体的组装模型中，RING1B 作为桥梁分子与 CBX 蛋白相互作用，进而介导了 cPRC1 蛋白复合体各成员在染色质上的装配。由于基因转录过程中可变剪接的发生，产生了一种人源 CBX2 蛋白亚型 CBX2-2，它由于失去了 PC 结构域而无法与 RING1B 相互作用，也因此丧失了参与 cPRC1 复合体组装的能力^[57]。这一结果也符合 cPRC1 组装的经典模型。然而，在本论文

的研究中,作为研究对象的 CBX7 蛋白亚型 CBX7C 同样也缺失了与 RING1B 互作保守 PC 结构域,但是与 CBX2-2 不同的是,CBX7C 在细胞内能够通过 PHC2 的互作而保留了其驱动 cPRC1 组装的能力。这也说明了 cPRC1 蛋白复合体的组装机制不是单一的,不同 CBX 蛋白参与的复合体组装途径也有所差异。

1.2.3 PRC1 和 PRC2 介导的组蛋白修饰及招募机制

尽管 PcG 蛋白家族及其靶基因在进化过程中高度保守^[11],但 PcG 招募机制涉及的序列可能存在显著差异。在果蝇中,作为 DNA 调节元件的 PRE 能够将 PcG 因子招募至目标染色质,并在发育过程中介导和维持染色质沉默的表观遗传调控^[58-60]。在果蝇的 PcG 招募过程中,与特异的 PRE 结合的多种转录因子 TF 也发挥了重要的作用,如 Pho、Pho1、Trl/GAF、Spps、Dsp1 和 Cg^[61]。在经典的 PcG 招募模型中,TF 首先招募 PRC2 蛋白复合体,随后通过 Polycomb (Pc/CBX) 对 PRC2 催化积累的 H3K27me3 组蛋白修饰的识别及相互作用来招募 PRC1 蛋白复合体,完成 PcG 在目标基因的招募过程^[62] (图 1-3A)。此外,各种转录因子、PRC1 和 PRC2 之间也存在相互依赖的合作关系。研究表明 PRC1 可以通过正反馈调节而稳定维持 PhoRC 与 PRE 的结合,进而介导 PRE 上大量的 PRC2 结合^[63, 64]。此外,在某些特定基因组背景下,PRC1 和 PRC2 蛋白复合物也可以独立招募至目标染色质而发挥功能^[65]。PRC1 不依赖于 PRC2 的独立招募可能是由 PRC1 的辅因子 Scm 参与介导的,其 C 端的 SAM 结构域具有蛋白低聚能力,使 Scm 成为连接 PhoRC 和 PRC1 的分子桥^[66] (图 1-3B)。

在哺乳动物中,PcG 蛋白的全基因组图谱研究揭示了其招募与低甲基化 CGI (CpG island) 的高度相关性^[67, 68]。研究表明,在无 TF 结合的基因位点上进行人工富集 GC 后,可以异位招募 PcG 蛋白家族,因此在哺乳动物中 CGI 可能具有与果蝇 PRE 相似的生物学功能^[69, 70]。作为组蛋白 H3K4me3 和 H3K36me2 修饰的特异性去甲基化酶,KDM2B (FBXL10) 能够通过其 ZF-CxxC (zinc-finger CxxC) 结构域与 CGI 结合,协助 PRC1.1 招募至靶基因^[71]。然而,在 ESC 中,尽管 KDM2B 几乎能够与所有的 CGI 结合,但其仅参与了 30%左右 CGI 招募 PRC1/2 蛋白复合体的过程^[72, 73],这表明 PcG 复合物结合 CGI 的途径具有多样性。与 PcG 蛋白家族结合的 CGI 主要位于转录抑制基因的启动子区域,目标基因需被转录抑制而产生的分子信号介导了 PcG 蛋白家族招募至新的沉默基因的 CGI^[70]。在 Klose 等人提出的“染色质采样”模型中,PcG 蛋白与所有潜在结合位点具有较弱的相互作用,但其与 CGI 的稳定结合被活性转录或激活 TF 的结合所阻断^[74]。此外,在转录被抑制的区域仍有大量未结合 PcG 蛋白家族的 CGI,这可能是通过 PcG 蛋白家族与 TF (如 E2F6、MGA/MAX、REST、SNAIL、RUNX1、

JARID2、AEBP2 和 YY1 等)的相互作用,进而补充 PcG 的“染色质采样”模型的一种“补充机制”,并且这些 TFs 仅在特定情况下诱导 PcG 蛋白家族招募至目标基因位点^[12]。

长链非编码 RNA (long non-coding RNA, lncRNA) 也参与了 PcG 复合体招募至靶基因的过程(图 1-3C)。最初研究发现,在 ESCs 分化中 *Xist* 和它的反义抑制因子 *Tsix* 可转录产生不同的 lncRNA (包括 *XIST* 和 *TSIX*),二者随后协同参与 X 染色体失活的表观遗传调控,在此过程中,*XIST* 将 PcG 蛋白家族靶向至非活性 X 染色体进而发挥其转录抑制功能^[75]。此后,进一步研究发现其它 lncRNAs (如 *HOTAIR* 和 *KCNQ1*) 也介导了 PcG 蛋白家族招募至 *Hox* 或印记基因位点的过程^[76]。

调节 PcG 招募的第三个主要因素是其它组蛋白修饰复合物或 PcG 蛋白自身催化积累的染色质修饰标记(图 1-3D)。在经典的“层次结构招募”模型中,PRC2 催化的 H3K27me3 修饰增加了含有 Pc/CBX 的 PRC1 复合物对目标染色质的亲和力,介导了 cPRC1 复合体在目标基因位点的招募^[29, 31]。在 ncPRC1 复合体招募过程中不依赖于 PRC2,而是由 KDM2B 介导的,ncPRC1 催化积累的组蛋白修饰 H2AK119ub 能够促进 PRC2 与目标基因的结合^[77, 78]。此外,ncPRC1.3/1.5 复合体通过催化 H2AK119ub 修饰而起始了 ncPRC1 和 PRC2 向非活性 X 染色体的募集,表明 ncPRC1 和 H2AK119ub 在生理环境中具有启动 PcG 功能的作用^[79]。Kalb 和 Cooper 等人相继发现 H2AK119ub 介导 PRC2 的招募过程还可能取决于 AEBP2 和 JARID2 蛋白的参与,并且它们与 H2AK119ub 在体外存在直接的相互作用^[80, 81]。同时,SCM 或 EED 介导的 PRC1 和 PRC2 复合体之间的直接相互作用也进一步促进了二者与靶基因的共结合^[66, 82, 83]。通常认为组蛋白甲基化修饰 H3K36me3 与活性基因或 DNA 损伤位点相关^[84]。研究表明组蛋白修饰 H3K36me3 有助于 PcG 靶向目标基因位点^[84]。H3K36me3 能够与 PRC2 复合体相关因子 PCL (polycomb-like protein) 中所含有的 Tudor 结构域结合,并介导 PRC2 复合体侵入活性染色质区域以促进该区域的基因沉默^[85]。此外,H3K9me3 与 Polycomb 之间也建立了多重联系。研究发现 PRC2 复合体的亚基 EED 能够与 H3K9me3 结合^[86],而且 PRC2 复合体与 H3K9 甲基转移酶 GLP/EHMT1 和 G9A 之间也具有直接的相互作用,当解除这一互作关系后,PRC2 复合体与染色质的结合显著减少^[87]。

此外,其它相关因子(如 TF 和 ncRNA 等)与特异性染色质之间的相互作用将结合有 PRC2 复合体的染色质稳定在一个特定区域内,在该区域中 PRC2 复合体可以进一步催化累积 H3K27me3 并促进 cPRC1 的招募^[88]。同样,在 H3K27me3 缺失的情况下,多种招募相关的组分可能介导了 PRC1 复合体与染色质的特异性

结合^[89-91]。因此, PcG 复合物自身具有相对较低的 DNA 序列识别特异性, 但它们可以通过特异性 TF、新生 RNA、lncRNA 和染色质修饰而特异性地募集到特定区域或从特定区域解离来发挥功能^[12]。此外, 最新研究表明, PRC2 形成两个主要的复合体亚型 PRC2.1 和 PRC2.2, 其特异性辅助蛋白对多梳蛋白介导抑制具有重要贡献^[92]。PRC2.1 主要催化 Polycomb 靶基因的 H3K27me₃, 进而招募 CBX2/4-cPRC1, 但是其无法招募 CBX7-cPRC1。相反, 虽然 PRC2.2 不参与催化 H3K27me₃, 但其辅助蛋白 JARID2 对于 CBX7-cPRC1 的招募和随后的 Polycomb 靶基因的 3D 染色质相互作用至关重要^[92]。

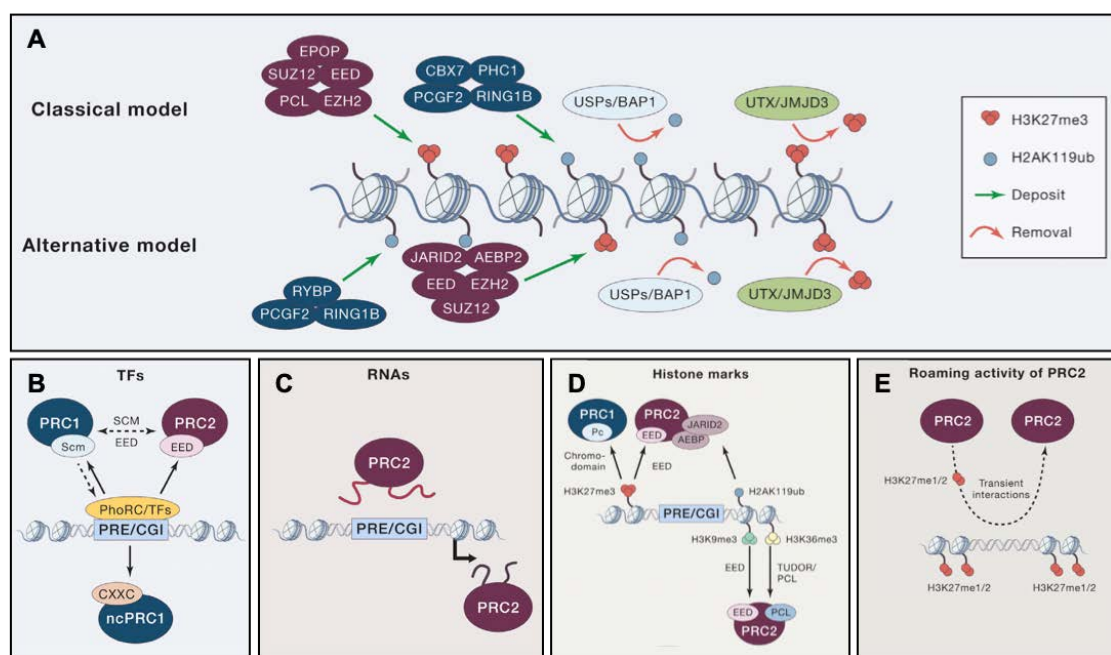


图 1-3 PRC1 和 PRC2 被招募至目标染色质的机制^[12]

Figure 1-3 Mechanism of recruitment of PRC1 and PRC2 to target chromatin^[12]

A: 经典招募模型: cPRC1 复合体的 CBX 蛋白亚基能够识别并结合 PRC2 复合体的靶向作用导致的 H3K27me₃, 并且 RING1 蛋白亚基进一步催化组蛋白 H2AK119ub。补充模型: ncPRC1 通过 FBXL10/KDM2B 蛋白亚基靶向未甲基化的 CpG, 并催化 PcG 靶位点的大多数 H2AK119ub 修饰, 同时有助于 PRC2 的募集; **B:** TFs 介导的招募; **C:** RNA 介导的招募; **D:** 染色质介导的招募; **E:** 在缺乏 PRE 序列和 PcG 招募组分时, PRC2 与染色质短暂结合并介导 H3K27 的低甲基化 (H3K27me₁/me₂)。

1.3 CBX 蛋白的研究进展

1.3.1 CBX 蛋白的功能结构域研究

在生物进化过程中, 哺乳动物细胞产生了五种与果蝇 Pc 蛋白同源的 CBX 蛋白, 包括 CBX2、4、6、7 和 8。在 NCBI 数据库中, 人源及鼠源基因组中至少含

有八种编码 CBX 蛋白的编码基因 (*Cbx1-8*)，其编码蛋白的 N 端都含有一个由 α 螺旋和 β 折叠构成染色质结合结构域 (Chromodomain, CD) [93]。CBX1、3 和 5 的结构和功能与果蝇异染色质蛋白 HP1 相似，因此分别被命名为 HP1 α 、 β 和 γ 。CBX2、4、6、7 和 8 蛋白的 C 端区域含有保守的多梳抑制框结构域 (Polycomb repressor box, PC)，主要通过 RING1B 蛋白的互作而介导 PRC1 复合体组装，进而发挥对靶基因的表观遗传调控功能[94]。在生物进化中，CBX 蛋白仅保留了高度保守的 N 端 CD 结构域和 C 端 PC 结构域。然而，通过蛋白质二级结构预测分析，CBX 蛋白的 CD 和 PC 结构域之间的区域具有典型的涉及蛋白互作的卷曲螺旋结构 coils[95,96]。ATH (AT-Hook) 结构域是具有 DNA 结合活性的卷曲螺旋结构之一，它能够与富含 AT 的 DNA 区域的小沟结构相互作用，进而促进蛋白与染色质或 DNA 之间的相互作用[97]。研究表明，在邻近于 CBX2 蛋白的 CD 结构域的 N 端区域含有一段相对保守的 ATH 结构域。通过序列分析比对，在 CBX4、6、7 和 8 中的 N 段区域也含有一个类 ATH 结构域 (AT-Hook like, ATHL) (图 1-4)。

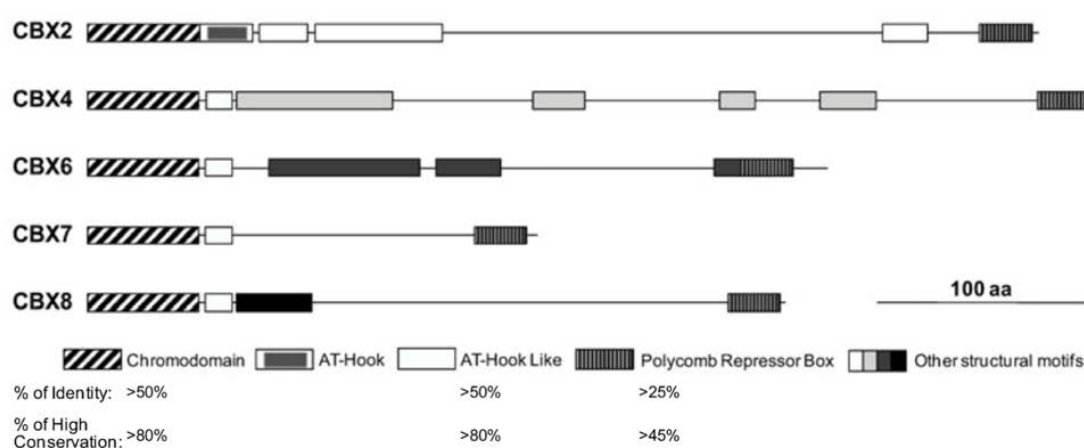


图 1-4 CBX 蛋白结构示意图及各结构域的同一致性和保守性分析[98]

Figure 1-4 Structural diagram of CBX protein and analyses of identity and conservation of each domain[98]

同一性：相同氨基酸 (aa) 的数量与总氨基酸数量的比率；高度保守性：高度保守性氨基酸的数量与总氨基酸数量的比率。

1.3.1.1 CD 结构域的结构及功能研究

研究表明，果蝇 Pc 蛋白与组蛋白修饰 H3K27me3 具有很强的亲和性，其 CD 结构域的第 36 位色氨酸残基 W 与第 40 位酪氨酸残基 Y 对 Pc 蛋白与 H3K27me3 的稳定结合密切相关[29]。此外，Pc 蛋白和组蛋白 H3 的互作与其疏水性氨基酸 (第 11 位缬氨酸 V 和第 50 位亮氨酸 L) 参与形成的疏水腔结构也密切相关[30]。

虽然 CD 结构域在生物进化过程中高度保守，但是随着蛋白及修饰的多样性

的增加,CD 结构域蛋白基序的关键氨基酸位点的改变可能导致了不同的 CBX 蛋白结合组蛋白甲基化修饰的种类及亲和性存在显著差异^[31, 99]。与果蝇 Pc 相比,尽管哺乳动物的 CBX 蛋白结合 H3K9me3 或 H3K27me3 的亲和性显著减弱,但是结合的组蛋白修饰标签的种类增加。结合结构解析数据,人源 CBX 蛋白表面则分布了较多的疏水基团,并且偏碱性,这也是导致 CBX 蛋白与组蛋白的亲和力下降的主要原因之一^[31]。比较不同 CBX 蛋白与组蛋白甲基化修饰之间的亲和性差异,CBX6 和 CBX8 与 H3K27me3 和 H3K9me3 的亲和性低于 CBX2、CBX4 和 CBX7^[32]。尽管 CBX6 和 CBX8 与 CBX2、CBX4 和 CBX7 具有相似的 CD 结构域,例如芳香笼及疏水基团等,但是由于 CBX6 和 CBX8 的 CD 结构域含有相对较高比例的精氨酸残基,改变了蛋白表面的电荷分布,进而导致了与组蛋白甲基化修饰的亲和性产生显著差异。根据多肽阵列数据分析,CD 结构域还可通过甲基化依赖的互作方式介导 CBX 蛋白与非组蛋白序列位点的结合^[30]。CBX 蛋白对不同的组蛋白甲基化修饰的亲和性也有差异。例如,尽管 CBX4 或 CBX7 均可识别并结合 H3K9me3 和 H3K27me3 修饰,但是 CBX4 或 CBX7 与 H3K9me3 的亲和力高于 H3K27me3^[32]。此外,CBX7 的 CD 结构域能够形成有助于组蛋白结合的“凹槽状”空间结构,但是该空间结构与蛋白肽链的结合不具特异性^[30]。CBX2 蛋白是与果蝇 Pc 蛋白同源性最高的 CBX 蛋白家族成员。与其他 CBX 蛋白相比,虽然 CBX2 优先靶向 H3K27me3,但是其亲和性却低于 Pc 蛋白^[32]。结构解析数据分析显示,CD 结构域的芳香笼结构包含了 CBX 蛋白与组蛋白甲基化修饰的结合区域^[30]。此外,在哺乳动物胚胎干细胞和成纤维细胞中,一些 CD 结构域之外的非保守区域也会影响不同的 CBX 蛋白结合到染色质的能力,这表明通过 CD 结构域介导的与 H3K27me3 修饰的结合并不是 CBX 蛋白特异性靶向至染色质的必要条件^[100]。上述研究表明,在复杂多样的生物进化过程中,尽管不同的 CBX 蛋白都具有特异性识别并结合目标染色质的功能,但是它们靶向的组蛋白甲基化修饰的种类及亲和性具有显著差异,因此使得不同的 CBX 蛋白具有其独特的下游靶基因群。

CBX 蛋白对组蛋白甲基化修饰的识别与结合介导了 PRC1 蛋白复合体靶向目标染色质的过程^[28-32]。之前的研究表明,哺乳动物的 CBX 蛋白能够靶向至近乎整个基因组中的组蛋白 H3K27me3 修饰^[101]。通过突变或缺失 CBX7 的 CD 结构域阻碍其与 H3K27me3 的结合会进一步影响 PRC1 复合体靶向至特异性基因位点的能力^[31, 32, 102, 103]。然而,除了 CBX4 蛋白外,其它 CBX 蛋白均能与 X 染色体失活相关的异染色质区域特异性结合,即使替换为 CBX4 的 CD 结构域也并未改变它们原有的靶向特异性^[32],这也说明 CD 结构域并不能独立完成 CBX 蛋白在靶基因位点上的招募。尽管 CBX 蛋白的 CD 结构域相较于其它区域具有较

高的序列保守性 (>80%), 但是其仍然需依赖于其它区域发挥其靶向特异性的功能^[104]。

除了 CBX2 外, 其余 CBX 蛋白 (CBX4、6、7 和 8) 的 CD 结构域都能与 RNA 在体外发生相互作用^[32]。此外, 一些具有明显的表达时空特异性的 lncRNA 也参与了 PcG 蛋白家族与染色质区域的特异性结合, 表明 lncRNA 与 PcG 蛋白家族的协同作用在表观遗传的时空特异性调控中可能发挥重要作用^[105]。例如, lncRNA (主要包括 *ANRIL* 和 *MOV10*) 通过与 CBX7 的 CD 结构域互作而被募集到编码细胞周期抑制、衰老以及应激性凋亡相关蛋白的 *Ink4b-Arf-Ink4a* 基因簇, 进而与 PcG 蛋白家族协同起始并维持其转录抑制^[106]。除此之外, CD 结构域自身被甲基化或去甲基化修饰的转换导致了 CBX4 与不同的 ncRNA 结合并分别靶向至亚核结构多梳小体 (Polycomb body, PB) (转录抑制状态) 或染色质间颗粒 (Interchromatin Granule, ICG) (转录激活状态)^[107], 进而引发转录因子 E2F1 介导的生长发育相关的调控基因在转录抑制状态和激活状态之间的转换^[107]。

1.3.1.2 PC 结构域的结构及功能

CBX 蛋白 C 端的 PC 结构域是其发挥转录抑制功能的关键区域^[94, 108-111], 该结构域在 Pc 蛋白及与其同源的 CBX 蛋白中高度保守。在哺乳动物中, PC 结构域是 CBX 蛋白区别于 HP1 蛋白 (CBX1、3 和 5) 的独特保守结构域。在果蝇中, PC 结构域的缺失导致机体出现 Pc 蛋白缺陷表型^[108], 并且 PC 结构域介导了 CBX 蛋白与 RING1 蛋白之间的直接相互作用^[111]。此外, 相关晶体结构研究表明, RYBP 和 CBX7 与 RING1B 同一区域互作, 二者竞争性地结合 RING1B 而形成不同的 PRC1 复合体亚型 (RYBP-ncPRC1 和 CBX7-cPRC1)^[34, 35]。PC 结构域的蛋白基序的保守性显著高于与其邻近的基序。然而, PC 结构域的部分基序在生物进化过程中也会发生细微变化, 进而影响 CBX 蛋白、RING1B 以及蛋白激酶等蛋白互作及其它特异性功能。例如, 某些疏水性氨基酸的变化引起蛋白物理性质的改变 (如 PC 结构域的第三位氨基酸在 CBX2 中由 Pc 中的缬氨酸 V 变为苯丙氨酸 F; CBX4 的第二、三位都由 Pc 中的缬氨酸 V 变为异亮氨酸 I 等), 因此可作为潜在蛋白翻译后修饰的新位点^[112]。

PC 结构域介导了 CBX 蛋白与 RING1B 的蛋白互作^[108]。尽管 PC 结构域的缺失或突变不会干扰 CBX 蛋白特异性识别并结合目标基因位点^[113], 但由于破坏了 PRC1 组装的桥梁 CBX•RING1B 而阻碍复合体的有效装配, 进而影响复合体对靶基因转录抑制调控活性^[94, 111]。然而, Divya Tej Sowpati 等人研究发现, 与其它 CBX 蛋白完全不同, CBX4 的亚细胞定位并未由于 CD 结构域的缺失而发生改变。相反, PC 结构域的缺失几乎完全改变了 CBX4 在细胞核内原有的分布

规律^[112]。基于此,不同的 CBX 蛋白的 PC 结构域在发挥生物学功能的过程中具有不同的作用。

1.3.1.3 ATH 或 ATHL 结构域的结构及功能

AT序列结合结构域(AT-Hook, ATH)或类AT序列结合结构域(AT-Hook like, ATHL)是结合富含AT并具凹陷结构的DNA序列的一类蛋白结构域^[114],最早在高迁移率家族A蛋白(High mobility group A proteins, HMGA)中被发现。CBX蛋白的ATH或ATHL位于其N端且邻近于CD结构域,由典型的GRP基序及侧翼序列构成,富含甘氨酸、精氨酸和脯氨酸^[115]。CBX蛋白的CD结构域和ATH结构域都能够结合含H3K27me3修饰的染色质核小体及富含AT的DNA卫星序列,二者共同介导了CBX2识别并结合着丝粒异染色质区域,进而发挥转录抑制调控功能^[116]。此外,CBX7的CD与ATHL结构域的融合表达蛋白在体外能够直接结合DNA,尤其是双链DNA。但是独立表达这两个结构域却使其丧失体外结合DNA活性,表明CD和ATHL结构域协同介导CBX7特异性识别并结合靶基因的过程^[117]。

此外,Pc蛋白或CBX蛋白均能通过ATH或者ATHL结构域中的GPR基序的介导,而结合TrxG蛋白复合体的组分CREB结合蛋白(CREB-binding protein, CBP)的组蛋白乙酰转移酶结构域(Histone acetyl transferase, HAT),进而通过抑制其组蛋白乙酰化活性而发挥拮抗转录激活的生物学功能^[118]。上述研究表明除CD和PC结构域外,相对保守的ATH或ATHL结构域在CBX蛋白以及PRC1复合体发挥表观遗传调控功能的过程中具有至关重要的作用。

1.3.2 CBX 蛋白不同转录剪接体研究

选择性剪接是真核生物基因转录过程中普遍存在的表达调控机制,能够使得单一基因产生多种独特的mRNA,细胞能够从有限数量的基因翻译更多的蛋白质多样性。选择性剪接可以通过多种方式实现mRNA多样性的增加,如外显子跳跃、互斥外显子选择、选择性剪接位点的使用(改变参与这种选择性剪接事件的内含子和外显子之间的边界,有助于转录物的多样性)以及内含子的保留,这些差异的产生也会影响mRNA的稳定性、定位或翻译过程^[119]。此外,一些剪接mRNA异构体可以通过改变阅读框而介导产生具有不同功能和定位的蛋白质异构体^[120]。选择性剪接增加了蛋白质组的复杂性^[121]。在正常生理或疾病状态的背景下,不断增加的细胞类型多样性以及不同蛋白质异构体之间的重要生物功能协调和替代事件的研究能够深入阐明组合调控的机理和生理意义,有助于相关疾病的治疗。

果蝇PRC1复合体的各组分在脊椎动物中衍生出多个同源物,这也使得在进化中PRC1蛋白复合体存在巨大的潜在组合多样性的空间。研究表明,细胞中存

在多种由不同类型蛋白亚基组成的PRC1复合物^[26]。PRC1不同组分的单一缺失能够导致小鼠出现具有不同表型的同源异型缺陷,这表明由不同种类的同源亚基组成的PRC1复合物可能通过控制一些非冗余靶基因的转录来影响个体发育^[122]。与这些由不同类型蛋白亚基组成的PRC1复合物的组成及功能研究相比,由基因转录的可变剪接而产生的不同蛋白亚型参与表观调控的研究相对较少。对于CBX蛋白而言,尽管已检测到同一基因产生的多种不同转录本及编码的蛋白产物,但是这些具有结构差异的蛋白是否具有相同转录调控的功能和机制,以及彼此在发挥功能时的拮抗或协同关系仍然有待深入探究。

研究表明,人源*Cbx2*基因位点产生了两种功能性转录本:*Cbx2-1*和*Cbx2-2*,并编码产生两种结构不同的蛋白产物CBX2-1和CBX2-2,二者都保留了保守的CD结构域(图1-5)。然而,由于转录时位于第4个剪接位点的后移,使得部分内含子转变为外显子,导致了*Cbx2-2*转录本缺失了后续的第五个外显子部分。其编码的蛋白产物CBX2-2缺乏PC结构域,进而丧失了与PRC1复合体的催化亚基RING1B互作的活性,也不参与PRC1蛋白复合体在染色质上的组装^[57]。但是与CBX2-1相比,脱离了PRC1复合体协同发挥功能的环境后,CBX2-2能够通过自身的聚合形成一个高分子复合物,并且仍具有抑制基因转录的调控功能^[57]。这一结果意味着即使缺失完整的功能性PRC1蛋白复合体,转录抑制过程也可能发生。

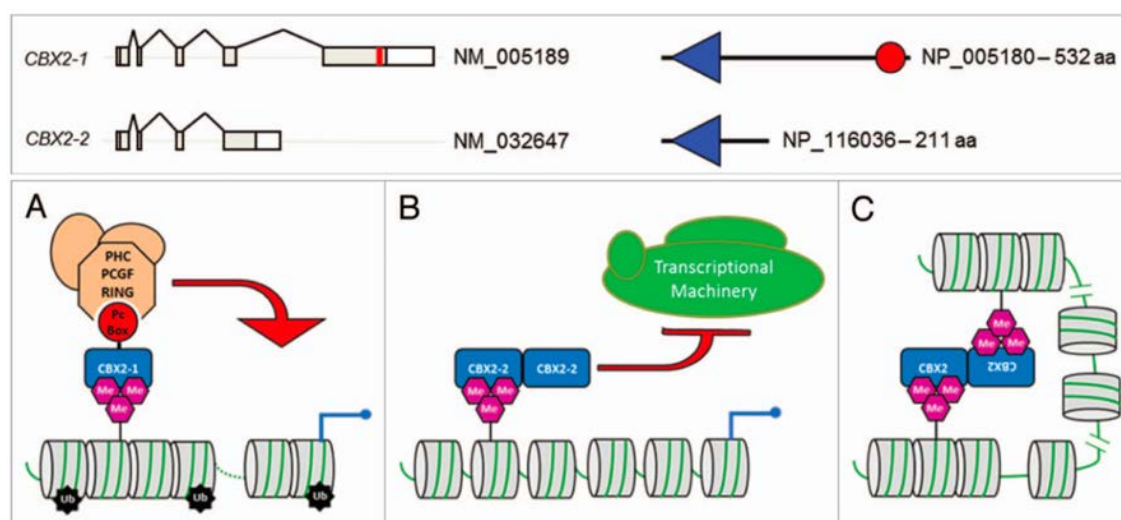


图1-5 人源*Cbx2*转录本与编码蛋白示意图, 以及其介导转录调控的模式图^[57]

Figure 1-5 Schematic diagram of human *Cbx2* and proteins, and CBX2-mediated repression^[57]

A: PRC1介导靶基因的转录抑制。CBX2-1通过其CD结构域特异性识别靶基因位点的H3K27me3修饰, 而其PC结构域招募参与基因转录抑制的其它PRC1组分; B: CBX2-2形成同源聚合物并阻止转录起始复合物在启动子处的结合; C: CBX2同源聚合物通过结合不同位点的H3K27me3修饰而介导染色质压缩。紫色六边形指示H3K27me3修饰。

与人类CBX2蛋白相似,斑马鱼CBX4、6和7蛋白也通过对转录物的选择性剪接而产生了不同的蛋白亚型(图1-6)。斑马鱼*Cbx4*基因通过选择性剪接多腺苷酸化事件可以产生两个转录物:*Cbx4a1*和*Cbx4a2*^[123]。*Cbx4a1*转录本包含5个外显子,编码产生含有477个氨基酸残基的CBX4a1蛋白质。而*Cbx4a2*包含6个外显子,编码包含有145个氨基酸的蛋白质CBX4a2。与CBX4a1相比,CBX4a2虽然保留了保守的CD结构域,但是缺失了PC结构域。斑马鱼*Cbx6a*基因的转录过程也有类似的特征。与*Cbx4a2*一样,*Cbx6a1*的转录物包含了6个外显子,其最后一个外显子仅由对应于终止密码子和未翻译序列的最后三个核苷酸TAA组成,它编码了包含有411个氨基酸的蛋白质CBX6a1,该蛋白质也缺失了PC结构域,但是保留了CD结构域和CBX6特有的CX6.1和CX6.2基序。体外的互作研究也进一步证实缺失PC结构域的蛋白亚型CBX4a2和CBX6a1丧失了与RING1B蛋白的直接相互作用^[123]。斑马鱼和人源的*Cbx7*基因产生了两种不同剪接形式的转录物*Cbx7a*和*Cbx7b*,它们编码的蛋白均包含有保守的CD和PC结构域。人源*Cbx7a*和*Cbx7b*均由6个外显子构成,而是*Cbx7b*是由第5个外显子延迟多聚腺苷酸化产生的。二者分别编码了由251个氨基酸构成的蛋白CBX7A和由158个氨基酸构成的蛋白CBX7B,并且这两种剪接体均包含了CD和PC结构域,且生物学功能相近。

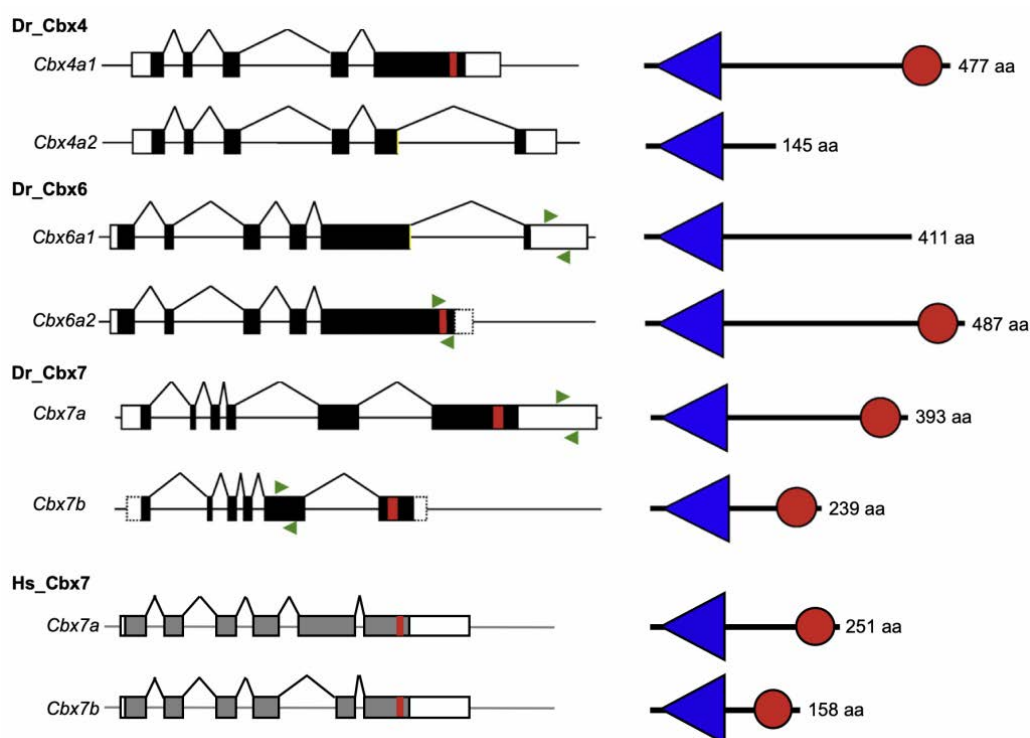


图 1-6 斑马鱼和人源 CBX 基因和蛋白质的示意图^[123]

Figure 1-6 Schematic representation of zebrafish and human CBX genes and proteins^[123]

在左侧基因结构示意图中,白色部分表示未翻译区域,而红色部分指示编码PC结构域的区域。右侧蛋白质结构示意图中,蓝色三角形指示CD结构域,红色圆圈指示PC结构域。

CBX家族成员CBX7参与细胞增殖调控^[124]。然而，CBX7在细胞增殖中的功能具有多样性。CBX7最早因其参与细胞寿命调控而闻名，并在脑、肺、肝、骨骼肌、肾和心脏中检测到其表达^[124]。之后，CBX7被认定为癌基因并检测到在前列腺癌、淋巴瘤和胃癌中表达水平显著上调^[125-127]。然而，CBX7在一些特定类型的癌症中也可作为抑癌因子参与转化和细胞周期相关基因（如Wnt/ β -catenin和Cyclin E1）的转录抑制，包括乳腺癌、癌症胰腺癌、癌症肺癌、癌症甲状腺癌、癌症结肠癌、癌症膀胱癌和脑癌等^[124-126, 128-132]。据此，CBX7可能在特定组织和环境中发挥其独特的生物学作用，具有时空特异性。然而，这种时空特异性对细胞增殖的潜在作用机制尚不清楚。

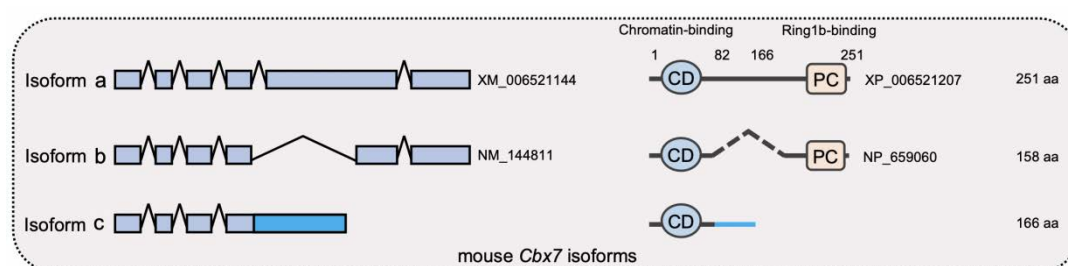


图 1-7 小鼠中 *Cbx7* 三种转录本及其编码的蛋白质产物示意图

Figure 1-7 The schematic diagram of the three mouse *Cbx7* isoforms and the protein products
鼠源 *Cbx7* 基因产生三个转录物：*Cbx7a*、*b* and *c*，分别编码了由 251aa 构成的 CBX7A 和由 158aa 构成的 CBX7B，二者均含有 CD 和 PC 结构域。由于选择性剪接，CBX7C 仅保留了 CD 结构域，并获得不同于 CBX7A 和 CBX7B 的 C 端区域。

目前，在NCBI数据库中包含有*Cbx7*基因产生的三种人源及三种鼠源的转录本（<http://genome.ucsc.edu>, mouse genome NCBI37/mm9），但是三种鼠源的CBX7蛋白亚型并未得到验证。近年来，研究报道了新鉴定的鼠源CBX7的两种蛋白亚型：p36CBX7（本论文中的CBX7A）和p22CBX7（本论文中的CBX7B）^[133]。与人源CBX7A同源的鼠源p36CBX7定位于细胞核，但其过表达不影响细胞增殖。然而，与人源CBX7B同源p22CBX7则主要定位于细胞质，其过表达抑制了细胞增殖^[133]。在本论文的研究中，由于不同的多聚腺苷酸化位点及可变剪接方式，鼠源*Cbx7*基因可产生三个转录物：*Cbx7A*、*Cbx7B* 和 *Cbx7C*（图1-7）。与上述人源*Cbx7*相似，*Cbx7A*和*Cbx7B*均由6个外显子构成，分别编码了由251个氨基酸残基构成的蛋白CBX7A和由158个氨基酸残基构成的蛋白CBX7B，并且这两种剪接体均包含了CD和PC结构域，生物学功能相近。与*Cbx7A*和*Cbx7B*相比，鼠源的第三种转录物*Cbx7C*仅含有4个外显子，前3个外显子与*Cbx7A*和*Cbx7B*相同，第4个外显子转录时改变了原有的选择性剪接事件中内含子和外显子之间的边界，将原来的第4个内含子的前端转变为外显子进行转录。其编码的蛋白与斑马鱼中表达

的CBX4a2、CBX6a1以及人源CBX2-2相似，仅保留了N端的CD结构域。鼠源CBX7C与人源CBX2-2相比，结构相似而分子机制却不相近。本论文的研究结果中显示，与CBX2-2相比，缺失PC结构域的CBX7C蛋白虽然也丧失了与RING1B的互作，但是它保持了与另一组分PHC2互作的能力，能够参与功能性PRC1复合体的组装。另外，CBX7C-cPRC1也可靶向至已知的CBX7A/B-cPRC1的靶基因。同时转录组数据分析结果显示鼠源CBX7C也具有转录抑制功能。本论文的研究从不同角度出發，进一步深入探究了其独特的生物学特性及功能。

此外，PRC1的另一核心组分Ph的编码基因*Ph*基因家族的转录也存在选择性剪接多腺苷酸化事件的发生^[123]。在进化过程中，除了保守的基因序列和合成酶外，*Phc2*基因转录调控方式也是保守的。人源*Phc2*基因通过启动子转换而产生了两种不同的转录物*Phc2a*和*Phc2b*，从而编码两种分子量为90 kDa和36 kDa的PHC2蛋白亚型。短亚型PHC2b由长亚型PHC2a的C端的最后323个氨基酸组成，并且仍含有保守的FCS基序和SAM结构域。在斑马鱼中，*Phc2*基因以相同的转录剪接形式也产生了两种转录物*Phc2a*和*Phc2b*，并编码两种大小不同的蛋白亚型PHC2a和PHC2b。短亚型PHC2b与长亚型PHC2a的C端的氨基酸组成一致。二者在斑马鱼的发育过程中表现出不同的表达模式^[134]。PHC2a主要是在胚胎的体节期表达，而PHC2b则在斑马鱼的整个发育阶段表达。此外，沿着胚胎发育中的不同体节，这两种蛋白亚型表现出不同时空表达模式。PHC2a最初是在第7-8体节阶段表达；PHC2b则出现在第1个体节中。随着体节的形成，PHC2a和PHC2b的表达在每个体节中平行地向尾芽移动位置。在单个体节中，PHC2a的表达在后部边界较高，在前部区域较低。与之相反，PHC2b的表达在相对应体节的前部边界较高，在后部区域较低。

与之类似，在本论文的研究结果分析中，CBX7C与CBX7A/B在mESC和小鼠各组织中表现出不同的时空表达差异，CBX7C主要在mESC和小鼠脾脏组织中表达，这意味着CBX7C在机体发育的不同阶段具有特殊的生物学功能。

1.3.3 CBX 蛋白在 ESC 分化中的时空交替及生理意义

果蝇中 PcG 蛋白家族的突变会导致机体的同源异型转化和其它相关发育缺陷，而在小鼠中大多数 PRC2 或 PRC1 复合体亚基的缺失会直接导致原肠胚形成阶段的胚胎死亡^[10]。近年来，越来越多的表观遗传学证据表明，PcG 蛋白家族的缺失不仅能够导致早期多能干细胞生存能力的缺陷^[135]，而且还影响细胞命运决定及维持。在多能干细胞^[136]以及分化程度更高的祖细胞和终末分化细胞中^[137]，PRC1 和 PRC2 复合体在全基因组中的定位也进一步证实了 PcG 蛋白家族在细胞属性（cell identity）决定以及在发育中正常分化程序的维持中具有关键作用。在

胚胎和成体干细胞的基因组中,大量的 PcG 蛋白家族结合区被 H3K27me3 抑制标记和 H3K4me3 激活标记修饰^[138]。PRC2 和 MLL2 复合体协同作用催化产生上述两种转录状态的组蛋白甲基化修饰,进一步促进细胞命运决定和发育调控的相关基因的启动子设定为稳定的二价状态。这种“二价启动子”的转录水平很低,其激活或抑制状态的选择取决于机体产生的发育信号^[139]。

在不同谱系细胞中,启动子区域和广泛结合 PcG 蛋白家族的抑制基因组会有所不同^[12]。在哺乳动物的多能干细胞中,PcG 蛋白家族参与了几乎所有的谱系特异性基因的表现遗传抑制调控,而与干细胞增殖及干性维持相关的基因处于持续转录激活状态。在诱导胚胎干细胞分化过程中,干性维持相关的基因表达沉默,而谱系特异性基因逐渐转去抑制转变为激活状态。在此过程中,PcG 蛋白仍然对细胞命运决定相关的基因具有一定程度的转录抑制调控,以保证多能干细胞能够顺利分化为特定的细胞类型,并防止细胞发生异常的去分化或反分化。PRC1 和 PRC2 复合体协同介导的这种精细调节和全基因组再分配过程主要依赖于 PcG 蛋白家族各组分的组装和拆卸,以实现多能性和谱系定型之间的适当平衡(图 1-8)。

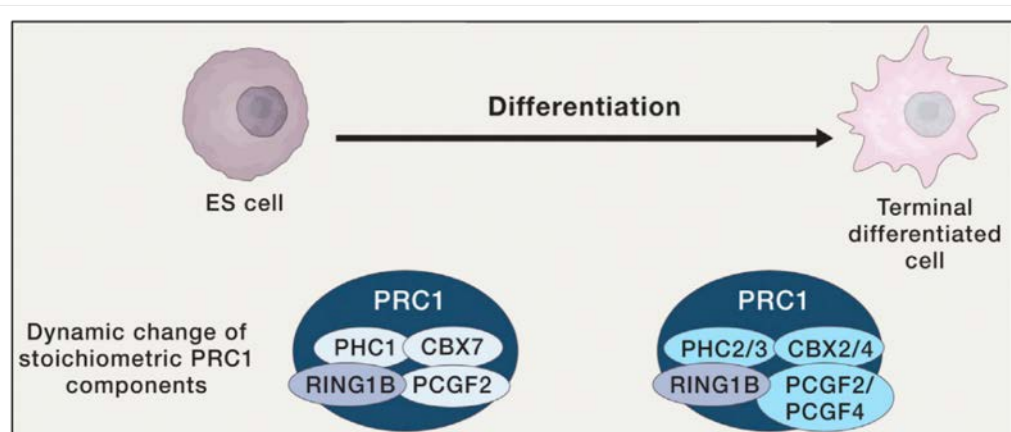


图 1-8 胚胎干细胞分化过程中 PRC1 复合体的动态组装^[12]

Figure 1-8 Dynamic Assembly of PRC1 complex during ESCs Differentiation^[12]

PRC1核心组分不同同源蛋白(如PCGF或CBX蛋白)在ESCs增殖和分化过程中分别参与 PRC1复合体组装中,进而动态调节PRC1功能。

在果蝇和哺乳动物中,CBX 蛋白是从机体早期发育到成年阶段过程中基因表达的表观遗传调控的关键蛋白。Polycomb 基因在脊椎动物中编码的同源物 CBX2、4、6、7 和 8 在进化中产生了不同的基序组合,这也使得这些同源蛋白具有了各自特异性的结构与功能。在机体发育过程中,它们表现出来不同的时空表达模式。在 ESC 分化为神经或脂肪细胞的过程中,CBX 蛋白亚细胞定位的变化会显著影响 PRC1 复合体在全基因组染色质上分布^[140]。CBX7 蛋白是哺乳动物 ESC 中 PRC1 复合体的主要 CBX 家族成员。CBX7 蛋白的表达能够抑制 ESC

的分化并促进其自我更新^[141, 142], 在 ESC 的干性维持和发育相关基因的转录沉默中发挥关键作用。在 ESC 的自我更新中, CBX7 蛋白介导了 *Cbx2*、*Cbx4* 和 *Cbx8* 的转录抑制。随着 ESCs 分化的发生, CBX7 蛋白被其它旁系家族成员取代, 导致 *Cbx7* 基因的表达下调, 而 *Cbx2*、*Cbx4* 和 *Cbx8* 的转录去抑制引起对应的编码蛋白表达逐渐上调^[136]。如果 CBX7 在机体发育过程中未被其它 CBX 家族成员取代, 会导致造血干细胞的过度增殖, 从而引起白血病的发生^[102]。另一方面, *Cbx7* 基因敲除后促进了 ESC 的进一步分化, 导致外胚层发育相关的基因转录去抑制。相反, CBX2、CBX4 和 CBX8 蛋白的高表达促进了 ESC 分化的发生^[102]。例如, 在 ESC 中过表达的 CBX4 蛋白能够识别并结合至神经祖细胞特异性表达的基因位点, 并抑制与干细胞干性维持相关的基因表达, 从而促进其分化的发生^[143]。在 ESC 分化为不同类型细胞的过程中, CBX7 特异性结合的基因位点逐渐被其它 CBX 蛋白所取代。在向成纤维细胞分化的过程中, CBX7 在 ESC 中靶向的 1589 个基因位点部分被 CBX2 取代^[136]。PRC1 复合体在 ESC 和分化细胞 K562 (白血病相关细胞) 的靶基因群也具有高的相似度, 但是仍有少量基因位点在细胞分化前未结合 CBX7, 而分化后却是 CBX2 或 CBX8 的下游调控基因^[144]。综合而言, 哺乳动物中不同 CBX 蛋白的大部分下游靶基因群是相同的, 但是也存在某种 CBX 蛋白所特有的基因结合位点, 并且不同 *Cbx* 基因的表达水平差异也会影响同一时空 PRC1 复合体的下游靶基因群的变化^[101]。

随着近年来研究的不断深入, 揭示了 CBX7 不是 ESC 中唯一的 CBX 家族蛋白成员, CBX2 与 CBX6 也在 ESC 中具有较高水平的表达^[145, 146] (图 1-9A)。与其它 CBX 蛋白相比, CBX2 具有较强的压缩染色质能力^[147]。丧失了压缩染色质活性的 CBX2 突变体导致 CBX2-PRC1 无法靶向至 H3K27me3 标记的基因位点, 进而引起其在 mESC 中的靶基因去抑制。此外, 在胚胎发生过程中, 压缩染色质活性的丧失还会抑制 CBX2-PRC1 的功能, 表达 CBX2 突变体的转基因小鼠形成的胚胎会在轴向骨骼形成中表现出前后转换, 这与 CBX2^{-/-}小鼠表现为相同的经典 PcG 缺陷表型^[148, 149]。另外, 研究表明 CBX6 通过保持多能性和分化状态之间的平衡而在维持 ESC 干性方面发挥重要作用, 它的表达异常使得 Wnt 和 MAPK 信号通路都受到了干扰, 导致 ESCs 的增殖和分化之间的平衡崩溃, 最终引发了 ESC 的分化^[146]。

在另一方面, 除了 CBX6 之外, mESC 中存在的 CBX2-PRC1 和 CBX7-PRC1 复合物均能发生凝聚, 形成多梳小体 PB^[150]。PRC1 复合体也是以此来介导启动子区域的远距离互作进而压缩染色质^[21, 150, 151]。在这一调控模型中, CBX-PRC1 复合物整合遗传信息和表观遗传学修饰, 通过液液相分离 (liquid-liquid phase separation, LLPS) 而影响基因组的三维结构 (图 1-9B)。CBX2-PRC1 作为上述

模型中的支架，在 mESC 中发生相分离形成凝集体，进而驱动 CBX2-PRC1 复合体的凝集并形成 PB^[152]。CBX7-PRC1 在该模型中充当衔接子的作用，它能够与 H3K27me3 标记的染色质相互作用^[117]。CBX2-PRC1 和 CBX7-PRC1 之间通过 PHC 蛋白的聚合将二者共定位至 CBX2-PRC1 形成的凝集体 PB 中，与此同时将 CBX7-PRC1 结合的 H3K27me3 标记的染色质也带入了凝集体 PB 中^[150]。我们在研究 CBX7 蛋白在 mESC 中的表观遗传调控机制及功能时发现 CBX7 与 CBX2 能够共靶向至相同靶基因位点，并且二者通过 ATHL-CBX7 及 ATH-CBX2 发生了直接相互作用。这一结果为上述模型提供了更加精准地分子生化证据。

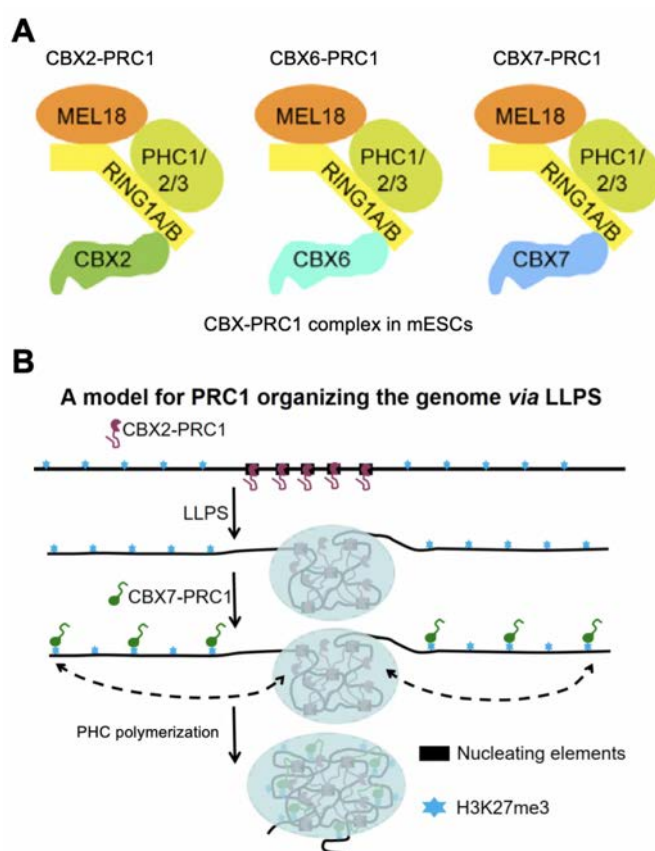


图 1-9 PRC1 复合体整合遗传信息及表观遗传修饰，通过相分离影响基因组三维结构^[150]

Figure 1-9 PRC1 complexes integrate genetic information and epigenetic modifications to organize the genome through LLPS^[150]

A: mESC中，CBX-PRC1示意图；B: PRC1通过相分离压缩染色质的串联模型：CBX2-PRC1作为支架，CBX7-PRC1作为衔接子，靶基因的H3K27me3修饰作为下游因子，通过CBX7-PRC1与H3K27me3的互作以及PHC介导的CBX2-PRC1与CBX7-PRC1的聚合，含有H3K27me3修饰的染色质被募集并压缩在含有CBX2-PRC1和CBX7-PRC1凝集体中。

为了深入探究不同 CBX 蛋白在分化中发生的时空转换的分子机制，Ana O' Loghlen 等人将 CBX7 与其它 CBX 蛋白的表达模式进行了系统比较。结果显示 *Cbx7* 基因在 ESC 中正常表达，并在细胞分化过程中急剧下调，但是其它 CBX 蛋

白的表达没有显著变化,而是在分化后逐渐增加。进一步通过功能筛选发现 miR-125 和 miR-181 家族是在 ESC 分化过程中参与 *Cbx7* 基因转录调控的关键调节因子。研究表明 miR-125b 和 miR-181a/b 在 ESCs 分化过程中被诱导产生,并通过对 *Cbx7* 基因的 3' UTR 区域的特异性靶向而抑制其转录的进行^[142]。

在本论文研究中,我们发现在 ESC 诱导形成 EB 的过程中, *Cbx7C* 转录本和其编码的蛋白水平急剧增加,这一变化与 *Cbx7A* 和 *Cbx7B* 的表达模式是相反的,这意味着 *Cbx7C* 可能逃逸了 miR-125 和 miR-181 家族的抑制。与此同时,我们还发现 CBX7C 在 mESC 中靶基因位点上的富集在一定程度上排斥了 CBX7A 与 CBX7B 的结合。同时,在亚细胞共定位的检测结果显示相同的“替代效应”,即 CBX7C-PRC1 凝集体中 CBX7A 和 CBX7B 被排斥在外,而 CBX7C 同样具有结合 PRC1 靶基因并抑制其转录的功能。同时,我们发现 CBX7A 和 CBX7B 的急剧下调并未引起其靶基因表达水平的显著变化。我们推测可能存在某种转换机制,即在 ESCs 中 CBX7A 与 CBX7B 是作为主导来调控基因表达;随着分化的进行,增加的 CBX7C 可能逐渐替代了被 miR-125 和 miR-181 家族抑制的 CBX7A 和 CBX7B,进而参与对靶基因的抑制和状态维持。

1.3.4 CBX 蛋白的表达异常诱发机体发育缺陷及癌变的研究

CBX 蛋白是重要的表观遗传调节蛋白,广泛参与多种生物学过程,如胚胎发育、干细胞的干性维持以及细胞的增殖和凋亡调控等^[153, 154]。近年来越来越多的证据表明,在肿瘤发生过程中, CBX 蛋白的表达紊乱以及功能障碍通常会导致发育过程的阻断或异位活化,对肿瘤细胞的干性^[155]、细胞周期^[156]、放化疗敏感性^[157, 158]以及肿瘤代谢^[159]等产生显著影响,进而促进癌症的发生和发展。CBX 蛋白调节肿瘤发生机制的研究为抗癌策略提供了有效的新靶点。

临床肿瘤样本的免疫组化和表达水平等多项分析证实,在多种类型的肿瘤细胞中 CBX 蛋白表达发生异常变化,并且这些变化与肿瘤的恶性表型和预后密切相关。转录和蛋白表达水平的生物信息学分析也进一步证明 CBX 可作为癌症预后的生物标记物^[160, 161]。与正常组织相比,部分 CBX 蛋白的表达在癌症组织中呈现异常上调^[162]。而作为肿瘤抑制物, CBX7 蛋白在大多数肿瘤组织中表达水平较低^[129]。此外, CBX 蛋白的表达水平的变化趋势与肿瘤的恶性表型密切相关。研究表明, CBX7 的表达水平与宫颈癌等的肿瘤阶段和恶性等级呈负相关,并且与其血管入侵有关^[163]。而 CBX2 与 CBX4 蛋白在肝癌中的表达水平与肿瘤大小呈正相关^[156],其高表达导致肿瘤复发率显著升高^[164]。另一方面, CBX 蛋白家族成员的差异表达与癌症患者的总生存率(OS)、无复发生存率(RFS)、无瘤生存率(DFS)、无进展生存率(PFS)、疾病相关存活率(DSS)、进展后生存率(PPS)

和无远处转移生存率（DMFS）密切相关，它们作为癌症预后标志物具有很大的应用潜力。

CBX 蛋白参与调控癌症发生和发展的机制是复杂多样的。作为表观遗传学调控因子，CBX 蛋白通过经典的表观遗传学修饰调节染色质状态和靶基因的表达，如组蛋白甲基化/乙酰化^[165, 166]和 DNA/RNA 甲基化。CBX 蛋白还可通过信号转导^[167]和代谢重编程^[159]等调节机制促进或抑制这些肿瘤相关的生物学过程。尽管目前临床癌症治疗中已经开始使用针对组蛋白修饰酶的小分子抑制剂，但这些抑制剂并不能特异性地“抹去或写入”表观遗传修饰标记，这可能会导致严重的治疗后果。而 CBX 作为关键的表观遗传调控蛋白，其在肿瘤发生发展过程中的功能和机制的研究有助于新的癌症诊疗策略的设计，针对于 CBX 蛋白的小分子抑制剂等药物的研发和应用在癌症治疗具有广阔前景。

1.4 蛋白质相分离在表观遗传调控中的生物学意义

诸多研究表明生物分子凝聚体的形成在机体细胞生命活动中发挥重要的生物学作用。其中，相分离是细胞中形成许多凝聚物的分子基础^[168-170]。在物理学研究中相分离是生物分子形成凝集体的必经过程。而在生物学研究中，相分离的发生在时间和空间上受到严格的调控，进而产生具有生物学功能的凝集体。生物分子凝聚体表现出的粘性液体和类似凝胶甚至固态的物理特性对其功能也至关重要^[171]。生物分子发生相分离而形成的凝集体并不是溶剂系统中的简单聚合物，而是存在于复杂溶剂（即细胞质或细胞核）中的高度复杂的结构^[172]。对细胞中发生相分离而形成复杂凝集体的机制研究，以及驱动相分离发生的关键蛋白质的作用机理的探索，对揭示相分离在生命活动中的重要功能及潜在的生物医药及临床治疗应用至关重要。

1.4.1 相分离的生物学背景

真核细胞中的凝集体是由生物分子发生相分离凝集而形成的微米级结构，能够发挥浓缩蛋白质和核酸的生物性功能，但是不具有典型的周围膜结构。研究表明这些凝集体参与多种生物学过程，包括 DNA 损伤应答、核糖体生成、RNA 代谢以及信号转导等^[171]。近年来研究发现，由多价大分子互作驱动发生的液液相分离（liquid-liquid phase separation, LLPS）是生物分子凝集体形成的重要机制之一，这一物理框架能够解释相分离形成的生物凝集体结构的组成、组装、物理和生化特性。

早在 1899 年动物学和遗传学家 Edmund Beecher Wilson 提出，作为细胞的

主要部分之一,细胞质是一种类似于由不同化学悬浮液滴组成的液态混合物^[173]。Tony Hyman 等人在 2009 年首次证实了 P granules 表现出类似液体的行为,并且其胞内的定位是由生物分子快速溶解和凝集过程所驱动的。自此科研人员逐渐意识到 LLPS 可能是各种无膜细胞器形成的分子基础^[174]。此后越来越多的研究表明,当多价蛋白互作时,随着蛋白质浓度的增加,会发生从小复合物到大聚合物凝集的快速转变,在宏观上表现为 LLPS^[175]。而 LLPS 的发生在体外反应中决定了 RNA 颗粒的组成及结构^[176],并且这种体内与体外相结合的研究相分离的方法加速了对相分离的探索和认知。LLPS 参与多种复杂且精细的生物学过程,如适应性和先天免疫信号传导、应激颗粒组装、异染色质形成、转录调控、miRNA 诱导的沉默复合体(miRISC)组装以及自噬等^[177-182]。

在经典的热力学理论研究中,溶解在溶液中的分子浓度达到溶解度极限时就会引发相分离的形成,该浓度也被称为其相分离的阈值浓度。而在细胞中,具有不同理化性质(如浓度、电荷、亲水及疏水性等)的生物分子通过快速分子运动形成不同的相,进而保持胞内分子理化平衡。形成凝集体的生物分子通常具有多价性,它能够调控分子内或分子间的互作,并且多价分子自发凝集而产生寡聚或多聚复合物^[175]。同时,由于熵驱动效应,分子的多价性降低了分子溶解度,进而导致相分离的发生^[175]。细胞内生物分子凝集体的形成主要通过多价高分子耦合组装及相分离的发生而介导的。这些生物分子主要包括了具有多模块互作结构域或含有弱粘性序列的无序区域的蛋白质。而具有多个与其他蛋白质等生物分子互作序列的 RNA 和 DNA 核酸分子也同样具有发生相分离的能力^[183]。在研究中,精细的分子实验数据能够进一步确定具有多模块互作结构域的蛋白质的互作模块数量及模块间亲和力,通过调控这两个相分离发生的关键参数,即增加模块数和增强亲和力,能够在较低浓度下实现蛋白质分子发生相分离^[175, 183]。

第二类能够形成凝集体的生物分子是含有连续且占比较高的内在无序区域(Intrinsically disordered protein regions, IDR)的蛋白质,它们能够在生理条件下发生相分离^[184]。缺乏稳定三级结构的 IDR 主要通过其重复基序而介导多价弱粘性分子互作。含 IDR 的蛋白驱动相分离而形成的凝集体中可能包含了 RNA 等多种生物分子,如 P-body、应激颗粒(stress granules, SGs)、生殖颗粒(germ granules)和一些核结构等。此类蛋白能够在体外多种溶液环境中自发地发生相分离^[185]。其包含的 IDR 通常由特定的氨基酸残基构成低复杂性的蛋白基序,主要包括丝氨酸、甘氨酸、天冬氨酸、谷氨酰胺、酪氨酸和苯丙氨酸。由这些氨基酸残基组成的多个重复基序在蛋白质靶向 RNA 颗粒^[186]或有丝分裂纺锤体^[187]并发生相分离的过程中发挥关键作用。此外,研究表明芳香类氨基酸残基在 IDR 驱动的相分离中也具有重要作用,它们能够通过和精氨酸残基在分子内或分子间的互作来

促进蛋白质相分离的发生^[185]。相反,蛋白 IDR 中含有的芳香类氨基酸残基的突变会导致这些蛋白质的相分离能力降低^[187, 188]。IDR 的电荷分布也是影响其相分离能力的关键因素之一。当电荷分布均匀且蛋白基序的净电荷趋于 0 时,蛋白发生相分离会受到阻碍。而当带同向电荷的氨基酸残基聚集时,则促进了蛋白相分离的发生^[184, 185, 189, 190]。蛋白 IDR 发生互作及驱动相分离的能力均取决于其侧链及骨架的氨基酸残基组成和蛋白整体序列模式^[191]。与具有多模块互作结构域的蛋白质相比, IDR 参与互作的基序尚不明确,但其多效性在促进生物分子发生相分离形成凝集体的过程中发挥核心作用。对蛋白 IDR 的精准预测和深入研究将成为未来生命科学研究的重要领域之一。

当蛋白质或核酸等生物分子发生 LLPS 时,它们通常在溶液中可凝集成若干个大小不一的类似液滴的致密相^[168]。在这个过程中,发生了由最初的大分子与水的结合转变为后来的大分子与大分子以及水与水的结合^[192]。除了生物分子自身的理化性质外,驱动相分离发生的主要因素还包括生物分子类型及周围环境条件,例如,溶液盐离子类型及浓度、环境温度、其它共存溶质及 pH 等^[185, 193, 194]。LLPS 可以在细胞中产生不同物理状态的凝集体。在液态的凝集体中,距离较近的分子间排列是规则的,而随着距离的增加,分子排列方向会发生变化,呈现出无序状态^[170]。液态凝集体具有典型的液体特征,它在溶液中能够形成球状液滴,大液滴能够分散成若干个小液滴,而小液滴也能聚合成为一个大液滴,该过程是由液体表面张力所决定的^[195]。构成液态凝集体致密相的分子通常可以在致密相内或致密相与稀释的周围环境之间交替,且具有分子流动性^[196]。通过 LLPS 形成液态致密相的过程通常是可逆的。然而,这种液态的凝集体通过不可逆转变而产生淀粉样纤维结构的水凝胶状态,并且在高盐或变性剂环境中才能重新溶解至周围环境溶液中^[170, 176, 197]。蛋白可以利用这种由液态向近固态的转变来实现其不同生理状态下的功能^[198]。

1.4.2 相分离研究方法的进展

随着对 LLPS 的理化机制及生物学功能的深入研究,研究者研发出诸多不同策略的相分离研究方法(图 1-10)。

首先,通过预测算法识别目的蛋白中是否含有无序区域,进而预测其是否具有发生相分离的能力(图 1-10A)。例如,最初由 Robert B. Russell 等人构建的 PONDR (Predictor of Natural Disordered Regions) 数据库能够预测目的蛋白的无序区,并能够明确其氨基酸残基位置及长度^[199]。此后,由 IUPred 数据库和 ANCHOR 数据库结合生成的 Iupred2a (Prediction of Intrinsically Unstructured Proteins) 数据库,该数据库不仅能够预测蛋白无序区,还能够分析推测其中参与

互作结合的无序区^[200]。在 2012 年，由 Julian Gough 等人构建的 D²P² (Database of Disordered Protein Predictions) 数据库能够对全基因组编码的蛋白质进行无序区预测^[201]。除了对目的蛋白无序区域的预测外，几个相分离数据库还提供了关于 LLPS 的实验验证的数据信息，整合了关于这些分子体系的生物物理驱动力、生物学功能及调控研究等的信息。

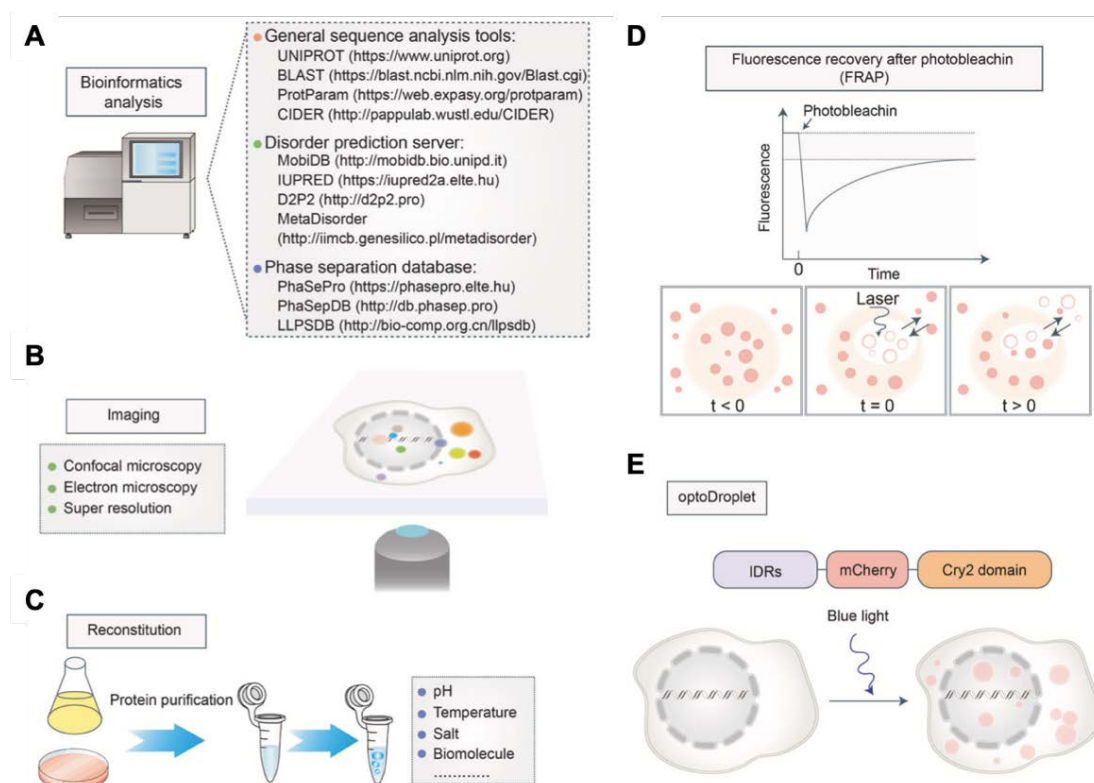


图 1-10 鉴定及研究 LLPS 的方法^[202]

Figure 1-10 The methods to identify or study LLPS^[202]

A: 用于研究LLPS的选定生物信息学工具或数据库; B: 显微镜是观察生物分子凝集物结构和组成的重要工具; C: 体外重构相分离; D: 通过FRAP实验检测生物分子凝集物的流动性; E: OptoDroplet系统可以利用蓝光激发调节蛋白多价性，进而促进或逆转细胞内生物分子凝集物的形成。

显微镜是可视化生物分子凝集体结构和组成的重要工具（图 1-10B）。共聚焦显微镜和超高分辨率成像技术的研发及应用为生物分子凝集物的研究提供了更精细的数据信息。例如，利用超高分辨率成像技术揭示了 SARS-CoV-2 病毒中由核衣壳蛋白-RNA-膜蛋白发生相分离形成凝集体的核心外壳结构^[201]。然而，这两种显微技术的局限性在于用于观察的生物分子凝集物必须带有荧光标签或被荧光抗体等标记。与此相比，电子显微镜可实现对无标签的生物分子凝集物进行精确地可视化观察^[203]。目前，多种显微技术与相分离体外重建分析的结合已被广泛应用于 LLPS 的研究领域。显微技术可用于从大肠杆菌、酵母或昆虫细胞中

纯化的蛋白质在体外环境中发生相分离的研究(图 1-10C)。由于 LLPS 对 pH 值、温度、RNA、盐浓度和翻译后修饰等因素敏感,因此显微技术与体外重建分析能够帮助我们进一步确定目的蛋白在体外发生相分离的特定条件^[204, 205]。

生物分子在凝集状态时的理化性质和动力学特征是其生物学功能的决定因素。例如,呈水凝胶状态的核衣壳蛋白形成的生物分子凝集体能够促进病毒核衣壳的组装,而液体状态的核衣壳蛋白凝集体则能够促进病毒基因在宿主细胞中的整合^[206]。生物分子凝集体的理化性质和动力学特征可以使用基于显微成像技术的方法进行研究,包括光漂白后荧光恢复 (Fluorescence Recovery After Photobleaching, FRAP)、荧光相关光谱 (Fluorescence Correlation Spectroscopy, FCS) 和光漂白荧光损失 (Fluorescence Loss in Photobleaching, FLIP)^[207]。其中,FRAP 是使用最广泛的相分离研究技术,其通过激光在选定区域进行光漂白后,凝集体的荧光会随着时间的推移而逐渐恢复(图 1-10D)。通常,当漂白处凝集体的荧光恢复的时间越短,则证明其流动性越高。

虽然相分离的体外研究有助于研究生物分子凝集物的多种理化性质及其影响因素,但这些生物分子凝集物在体内的组成成分、分子浓度和及生物学功能尚不清楚。目前,已开发的一个基于光遗传学的 OptoDroplet 系统能够通过蓝光激发而调节蛋白的多价性,以促进或逆转体内生物分子凝析体的形成^[208](图 1-7E)。CRY2 是存在于拟南芥中的蓝光受体,在蓝光激活后能够形成寡聚物^[209]。将 CRY2 与含有 IDR 的目的蛋白融合表达能够显著加速凝集过程。总体而言,这些新方法的研发对深入探究 LLPS 的作用机制及其生物学功能发挥了至关重要的作用。

1.4.3 不同 CBX 蛋白及 PRC1 组分 PHC2 相分离的研究进展

近年来,研究表明生物大分子发生相分离的现象广泛存在于具有细胞中,涉及多种生物学过程,相分离形成失稳与一些神经/肌肉退行性疾病等的发生密切相关。生物大分子(如蛋白质、DNA 和 RNA)的浓度、理化性质以及环境因素等密切影响相分离的发生发展,如环境温度、pH、盐离子类型及浓度^[205]。近年来相分离在转录调控领域中的分子机制及生物学功能研究成为热点,如相分离可以迅速感知周围环境变化作出快速应答^[193, 198, 210]。此外,相分离形成的凝集体参与了基因组的空間重塑,结合在某些基因位点的蛋白质经由相分离使位置较远的多个基因位点靠近,同时产生空间位阻而排斥不含结合位点的无关基因区域。这一可变结构促进并维持了超级转录因子在染色质结构相对疏松的区域的集中分布,为染色质的折叠创造了空间支点,从而有助于激活基因的表达或者抑制位于被浓缩的异染色质区域的基因转录^[211]。例如,转录因子、辅活化因子和 RNA 聚

合酶 II 协同组装为转录凝集体以激活转录^[212-214]。异染色质蛋白质通过相分离聚集形成凝聚体,进而参与异染色质折叠压缩^[215, 216]。

多梳抑制复合物 PRC 对染色质的修饰及构象调控为将基因组划分为功能不同的亚区和调节潜在基因的活性提供了重要方式。PRC2 的酶活性及其染色质的募集或排斥都是通过与辅因子和 DNA 元件、组蛋白及 RNA 等相互作用进行精确调控的。PRC2 介导的 H3K27me3 被不同类别的效应物识别,例如经典的 PRC1 和含有 BAH 模块的蛋白质(尤其是人类中的 BAHCC1)^[217]。这些效应子通过不同的机制介导基因沉默,包括涉及相分离发生的染色质压缩和组蛋白脱乙酰化^[152, 218]。近年来已报道的研究关注到 PcG-CBX2 的相分离是兼性异染色质形成的基础^[152, 218]。CBX2 可以独立于 H3K27me3 和 CBX-PRC1 在染色质上成核。CBX2 和 DNA 之间的相互作用对于 CBX-PRC1 在染色质上的成核和凝聚体的组装至关重要。由于相分离凝集体能提高生化反应的效率和特异性,从而加速了 CBX2 对目标基因的搜索过程^[150, 168]。此外,研究人员在深入发掘艾滋病致病病毒 HIV-1 的新型潜伏感染建立和维持因子时发现:CBX4 蛋白利用自身 LLPS 能力,不仅可以招募 PRC1 催化亚基 RING1B,还可以招募 PRC2 催化亚基 EZH2 至 CBX4 驱动形成的凝集体中;另一方面具有 SUMO E3 连接酶活性的 CBX4 蛋白有助于增强 EZH2 的 H3K27me3 催化能力,进而促进 PRC1 和 PRC2 对 HIV-1 潜伏感染的协同维持作用。而 CBX4 驱动的相分离的逆转可能是重新激活潜伏的 HIV-1 的潜在干预措施^[219]。尽管目前相分离在 CBX 蛋白参与基因转录调控中的功能研究取得了一定的进展,但揭示在活细胞中 CBX 蛋白参与相分离发生的潜在机制仍然是一个挑战,LLPS 如何组织核酸代谢的生化反应的分子机制尚不清楚。此外,PRC1 的核心组分 PHC2 具有 4 个保守的结构域。其中 HD1 介导了 PHC2 与 PCGF2 或 RING1B 的互作,而 SAM 结构域能够引发 PHC2 的自聚^[49]。SAM 域包含 ML 和 EH 结合面,这两个部分能够以首尾相接的形式介导蛋白自聚^[51]。并且 EH 结合面的突变(L307R)或 ML 结合面的突变(L293R/H298R)都将导致 SAM 结构域介导的蛋白自聚的能力的丧失^[51]。在前期研究 CBX7 不同剪接体蛋白功能的过程中,我们发现 CBX7A/C 与 PHC2 形成的互作复合物在细胞核内形成明显的相分离现象,并且 CBX7A/C 自身能够在核内形成 foci。在前期研究中,我们应用已开发的一些能够预测靶蛋白进行相分离的工具和数据库(PONDER, Predictor of Natural Disordered Regions)对 CBX7A 和 CBX7C 进行分析,发现 CBX7A 和 CBX7C 确实包含了参与相分离的无序区域 IDR。与 CBX7A 相比,CBX7C-PRC1 复合体更易发生相分离形成高度聚集且动态变化的凝集体颗粒。与此同时,在基于光遗传学的 OptoDroplet 系统中,在拟南芥蓝光受体 CRY2 的介导下,通过蓝光调节多价态以促进或逆转体内生物分子凝聚物的形成。而在此

过程中,与 CRY2 及荧光标签 mCherry 融合表达的 CBX7C 表现出更加显著的对蓝光的敏感性,其迅速形成凝集体。然而 CBX7C 发生相分离的生理条件及生物学意义尚不清楚。

1.5 本研究的立项依据与意义

1.5.1 本研究的立项依据

表观遗传调控因子如何通过复杂而精细的调控网络在细胞命运决定中发挥作用是近年来重要的科学问题。多梳蛋白家族 PcG 作为重要的表观遗传调控蛋白参与目标基因的转录抑制调控,在干细胞分化过程中参与形成并维持“细胞记忆”,产生可遗传的表型^[1]。在生物进化过程中,PRC1 的不同亚基在脊椎动物中都衍生出多个同源物,以及在基因转录过程中的选择性剪接多腺苷酸化事件引起单个基因产生不同的异构体,为 PRC1 复合体潜在的组合多样性提供了巨大的空间。由此引发的功能多样性的产生有待深入探究。近期多项研究表明 PcG 的一些核心组成亚基可发生相分离,在核内形成液滴状结构,但是其参与转录调控的分子机制及其存在的生物学意义尚不明晰,有待进一步研究。

首先,在 cPRC1 蛋白复合体组装的经典模型中,RING1B 与 CBX 蛋白的互作是复合体各成员在染色质上完成装配的关键。由于基因转录过程中的可变剪接而失去了 PC 结构域的人源 CBX2 的蛋白亚型 CBX2-2 因无法与 RING1B 互作,而丧失了参与 cPRC1 复合体组装的能力^[57],通过自身的自聚保留了对靶基因的转录抑制功能。然而,在本论文的研究中,我们发现 cPRC1 蛋白复合体的组装机制不是单一的,不同 CBX 蛋白参与的复合体组装途径也有所差异。据此,本研究着重探究由于相同可变剪接方式产生的鼠源 CBX7 的蛋白亚型 CBX7C 与其余亚型(CBX7A 和 CBX7B)及 CBX2-2 介导靶基因转录沉默的不同分子机制。

其次,在本论文探究 CBX7 不同蛋白亚型参与 cPRC1 复合体组装的分子机制的过程中,我们发现 CBX7 具有与 CBX2 类似的驱动相分离的能力。因此我们通过蛋白序列分析、体内及体外相分离理化性质及相关结构域的鉴定,进一步深入研究了 CBX7 驱动相分离的能力及分子机制,比较了不同 CBX7 蛋白亚型在相分离形成中的差异。

再次,CBX7A/B 在维持 mESC 的干性及发育相关基因的沉默中具有关键作用。尽管与 CBX7A/B 相比,CBX7C 的表达量较低,但同样的转录可变剪接形式在斑马鱼及脊椎动物进化过程中中具有较高的保守性,这也意味着这种缺失 PC 结构域结构的蛋白亚型可能具有不可或缺的生物意义。因此,我们进一步探究

了 *Cbx7C* 的基因沉默对 mESC 的生长稳态的影响, 以及在 mESC 诱导分化为类胚体 (Embryoid body, EB) 的过程中, *Cbx7* 的不同转录本及对应的编码蛋白的表达水平的变化。

最后, 在 mESC 中存在 CBX2-PRC1、CBX6-PRC1 及 CBX7-PRC1 三种复合体。研究表明 CBX2-PRC1 与 CBX7-PRC1 能够共存于 CBX2 驱动形成的多梳小体 PB 中^[150]。本研究前期结果也发现 CBX2 与 CBX7 在 mESC 中能够共靶向至相同靶基因位点上。基于此, 我们从分子及生化水平上探究了 CBX2 与 CBX7 的互作关系, 以及介导二者互作的关键结构域。

1.5.2 本研究的主要内容

在本论文中, 我们讨论了鼠源 CBX7 的三种不同蛋白亚型的结构, 相分离特质及参与基因调控的分子机制和生物学功能。结合多种生化及分子生物学方法及显微成像分析, 我们发现了缺失 PC 结构域的 CBX7C 仍然能够作为 cPRC1 的功能亚基而参与基因转录抑制调控, 并在 mESC 分化过程中表达上调, 可能替代了在该过程中下调的 CBX7A 和 CBX7B。与 CBXA 相比, CBX7C 在体内和体外具有更强的驱动相分离的能力。此外, 在 mESC 中, CBX7 与 CBX2 通过其 ATHL 和 ATH 结构域能够发生直接互作。本论文的主要研究内容总结如下:

1. 分析 mESC 及正常组织中不同 *Cbx7* 剪接体及其蛋白的时空表达模式差异;
2. 解析不同 CBX7 蛋白亚型与 cPRC1 其它组分的互作关系及招募和组装机制;
3. 比较 CBX7 不同蛋白亚型驱动相分离的能力差异;
4. 探究 CBX7C 识别 PRC1 靶基因的分子机制及介导转录抑制调控的生物学功能;
5. 探索 CBX7C 参与 mESC 中基因表观调控过程中的生物学功能及对 mESC 的干性维持及分化的影响;
6. 解析 CBX 家族蛋白间的互作关系。

1.5.3 本论文的技术路线

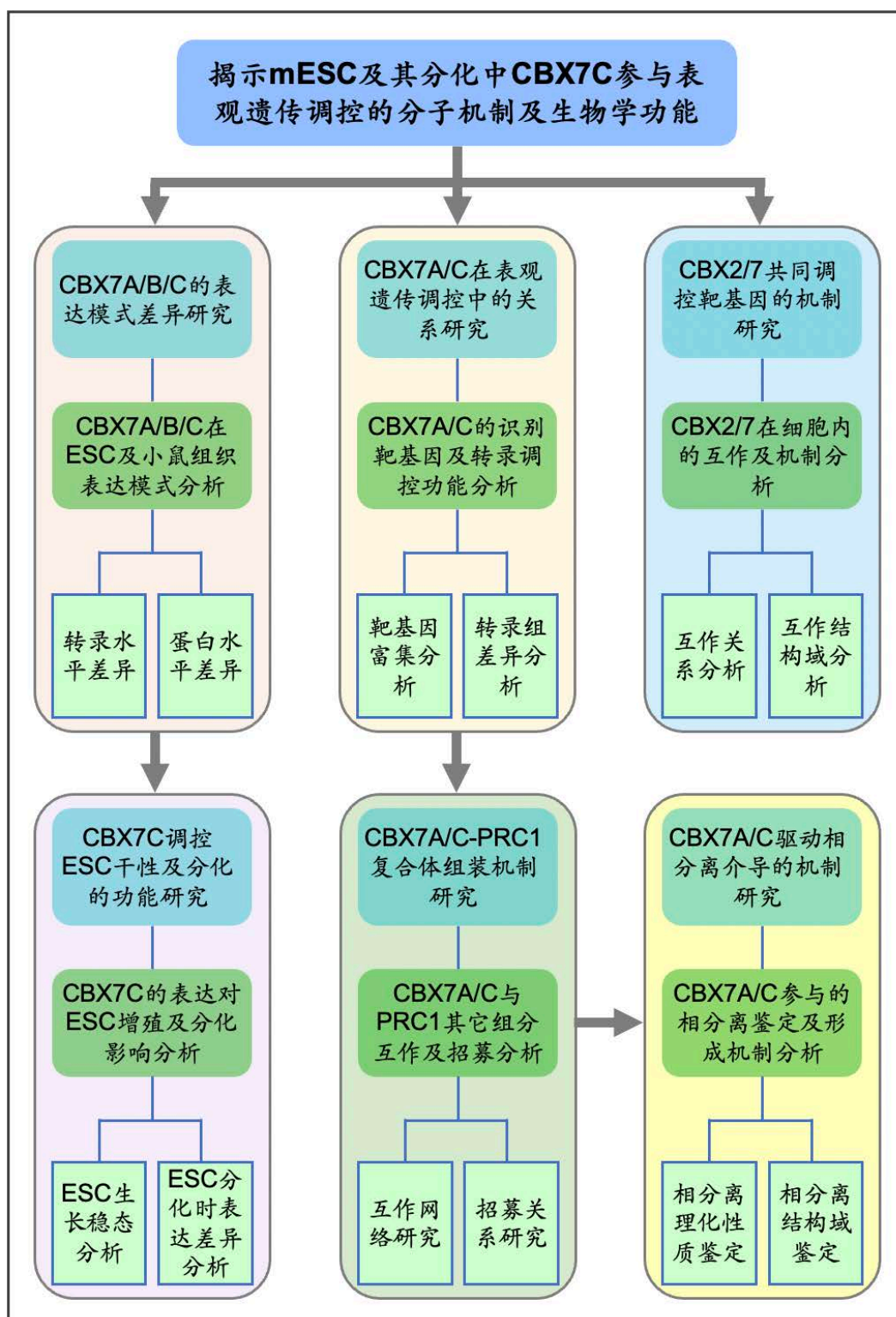


图 1-11 本论文研究路线图

Figure 1-11 Roadmap of this thesis

1.5.4 本研究的意义

多梳蛋白家族 PcG（包含 PRC1 和 PRC2）介导的表观遗传调控使其靶基因在分化过程中以细胞类型特异的表达模式被抑制，在发育和维持特定细胞状态中发挥重要作用，是调控细胞命运决定和发育过程中诸多重要细胞生物学过程的关键机制之一。CBX7 蛋白是哺乳动物 ESC 中 PRC1 蛋白复合体的核心组分之一，介导对 H3K27me3 的识别和结合进而在靶基因上装配 PRC1 并执行转录抑制调控。

基于基因转录的选择性剪接，我们首次报道了一种在 mESC 及脾脏组织中特异性表达的 *Cbx7* 剪接异构体 *Cbx7C*，并证明了其编码的蛋白 CBX7C 在缺失了关键的结构域后仍具有对靶基因转录的表观抑制调控功能。CBX7C 通过与 PHC2 的互作促进了 PRC1 的组装。在 mESC 分化为 EB 的过程中，CBX7C 与 PHC2 的表达协同上调，并且 CBX7C 的表达促进了 EB 的形成。此外，我们还发现表达水平较低的 CBX7C 与 PHC2 在互作后发生相分离并有效地形成了功能性多梳复合体。而当细胞中表达较高计量的 CBX7C 与 PHC2 后促进了大量多梳复合体的凝集，导致形成水凝胶状的凝集体的形成，并排斥其中的染色质成分。综上所述，我们的工作揭示了 CBX7C 与 PHC2 的互作在 PRC1 在染色质上的组装及 Pc 小体形成过程中的独特作用和分子基础，并开辟了一条通过调节其相分离性质来控制 PRC1 活性的新途径。

第二章 材料与方法

2.1 实验材料

2.1.1 实验所用细胞、菌株及质粒载体

本研究所用的细胞、菌株及质粒载体如下表所示（表 2-1，2-2，2-3）：

表 2-1 本研究所用的实验细胞

细胞株种类	细胞株名称
人胚肾细胞	HEK293T（本实验室保存）
小鼠胚胎干细胞	PGK12.1（本实验室保存）
宫颈癌细胞	HeLa（本实验室保存）

表 2-2 本研究所用的实验菌株

菌株（用途）	制造商
<i>E. coli</i> Stabl3（质粒扩增）	本实验室菌株
<i>E. coli</i> DH5 α （质粒扩增）	生工生物（B58413-0020）
<i>E. coli</i> Rosseta（蛋白表达）	本实验室菌株

表 2-3 本研究所用的实验载体

载体种类	载体名称
双分子荧光互补载体 (原始载体为本实验室保存)	pBiFC-VN-CBX7A
	pBiFC-VN-CBX7A (1-82)
	pBiFC-VN-CBX7A (83-251)
	pBiFC-VN-CBX7A (Δ CD)
	pBiFC-VN-CBX7A (Δ ATHL)
	pBiFC-VN-CBX7A (Δ CD+ATHL)
	pBiFC-VN-CBX7B
	pBiFC-VN-CBX7C
	pBiFC-VN-CBX7C (1-110)
	pBiFC-VN-CBX7C (111-166)
	pBiFC-VN-CBX7C (Δ CD+ATHL)
	pBiFC-VN-CBX2
	pBiFC-VN-CBX4

载体种类	载体名称
原核表达载体 (原始载体为本实验室保存)	pBiFC-VN-CBX6
	pBiFC-VN-CBX8
	pBiFC-VN-PCGF2
	pBiFC-VN-RING1B
	pBiFC-YC-RING1B
	pBiFC-YC-PHC2
	pBiFC-YC-PHC2(Δ SAM)
	pBiFC-YC-PCGF2
	PGEX-6P-1
	PGEX-6P-1-CBX7A
真核可诱导性表达载体 (原始载体为本实验室保存)	PGEX-6P-1-CBX7A (1-82)
	PGEX-6P-1-CBX7A (83-251)
	PGEX-6P-1-CBX7C
	PGEX-6P-1-CBX7C (1-51)
	PGEX-6P-1-CBX7C (1-110)
	PGEX-6P-1-CBX7C (52-166)
	PGEX-6P-1-CBX7C (111-166)
	PGEX-6P-1-EGFP
	PGEX-6P-1-EGFP-CBX7A
	PGEX-6P-1-EGFP-CBX7C
真核表达载体 (原始载体为本实验室保存)	pET28a-RING1B
	pET28a-PHC2
	pcDNA3.1-TRE-pCMV-CBX7C-EGFP-Flag
异源位靶向表达载体 (原始载体为本实验室保存)	pcDNA3.1-pCMV-CBX7A-EGFP-Flag
	pcDNA3.1-pCMV-CBX7C-EGFP-Flag
	pBiFC-mCherry-PHC2
optoDroplet	pcDNA3.1-pCMV-mCherry-HTM
	pcDNA3.1-pCMV-mCherry-HTM-CBX7A
	pcDNA3.1-pCMV-mCherry-HTM-CBX7C
	pcDNA3.1-pCMV-mCherry-HTM-PHC2
	pcDNA3.1-pCMV-mCherry-Cry2oligo

载体种类	载体名称
(原始载体为本实验室保存)	pcDNA3.1-pCMV-CBX7A-mCherry-Cry2oligo pcDNA3.1-pCMV-CBX7C-mCherry-Cry2oligo
慢病毒包装载体	pMD2.G
(原始载体为本实验室保存)	pPAX2
敲除载体	pLko.1-shRNA
(原始载体由南晶老师惠赠)	
荧光蛋白报告载体	pcDNA3.1- 5×UAS-CYP26A1-EGFP-Flag
(原始载体为本实验室保存)	GAL4 DBD GAL4 DBD-CBX7A GAL4 DBD-CBX7C
双荧光素酶报告载体	GAL4 DBD-IRES-Renilla Luciferase
(原始载体为本实验室保存)	GAL4 DBD-CBX7A-IRES-Renilla Luciferase GAL4 DBD-CBX7C-IRES-Renilla Luciferase GAL4 DBD-PHC2-IRES-Renilla Luciferase

2.1.2 实验试剂

本研究所用的实验抗体和试剂如下表所示（表 2-4，2-5）：

表 2-4 本研究所用的实验抗体

试剂名称	制造商	货号
一抗 CBX7C	本实验室制备	———
Tubulin	博奥龙	B1052
GST	Abmart	M20007
His	本实验室制备	———
Flag	Abmart	M20008
RING1B	CST	5694
H3K27me3	PTM BIO	PTM622
H2AK119ub	CST	8240
H3K9me3	Abcam	Ab8898
H3K4me3	PTM BIO	PTM613
PCGF2	Abclonal	Ab17327
PCGF4	CST	5856S

试剂名称	制造商	货号
CBX7A+B	Abcam	Ab21873
二抗		
HRP-山羊抗小鼠 IgG	北京博奥龙	BF03001
HRP-山羊抗兔 IgG	北京博奥龙	BF03008
山羊抗小鼠 IgG (Alexa Fluor® 594)	Abcam	Ab150116
山羊抗小鼠 IgG (Alexa Fluor® 488)	Abcam	Ab150113
山羊抗小鼠 IgG (Alexa Fluor® 594)	Abcam	Ab150080
山羊抗兔 IgG (Alexa Fluor® 488)	Abcam	Ab150077

表 2-5 本研究所用的实验试剂

试剂名称	制造商	货号
质粒 DNA 小量抽提试剂盒	生工	B518191-0100
高纯度质粒小提中量试剂盒	天根	DP107-02
PCR 纯化试剂盒	生工	B518141-0100
DNA 胶回收试剂盒	生工	B518131-0100
总 RNA 提取试剂盒	Yeasten	19211ES60
反转录试剂盒	Yeasten	11141ES10
HB-infusion 无缝连接试剂盒	汉恒生物	HB-3847385
双荧光素酶报告系统试剂盒	Promega	E1910
2×SYBR GREEN qPCR MASTER MIX	BIOMAKE	B21702
2×Taq master mix	Vazyme	P112-02
蛋白酶 K	生工	04-N-1-5
DNase	TRANSGENE	J10326
限制性内切酶、T4-DNA 连接酶	NEB/Vazyme	略
100 bp DNA marker	擎科生物	DL2000-01A
1 Kb DNA marker	Takara	3426A
Protein marker	全式金	DM111-02
96 well qPCR 板	百赛生物	BS30030962
高敏型 ECL 发光液	Vazyme	E4112-01
PVDF 膜	GE	10600023
Protein A/G IP beads	苏州海狸	222020-20

试剂名称	制造商	货号
Protease Inhibitor Cock tail	Bimake	B14012
37%HCHO	生工	A501912-0500
蛋白胨	生工	A505247
酵母提取物	生工	A100850
BeaverBeads™ GSH	Beaver-Bio	70601
BeaverBeads™ IDA-Co	Beaver-Bio	70502
BSA	索莱宝	A8020
AgaroseN	Invitrogen	920008
ExFect Transfection Reagent	Vazyme	T101-01
DMEM（高糖）	Gibco	12800017

2.1.3 细胞培养试剂与耗材

本研究所用的细胞培养试剂及耗材如下表所示（表 2-6）：

表 2-6 细胞培养试剂及耗材

细胞培养耗材	制造	货号
100 mm 细胞培养皿	Corning	L30196
6 孔、12 孔、24 孔细胞培养板	JETBIOFIL	TCP001-006/01
新生牛血清	Biological Industries	04-102-1A
胎牛血清	GEMINI	A87F82H
胰酶消化液	Biosharp	BL501A
β -巯基乙醇	Sigma	RNBJB366
DNA/RNA 转染试剂	Vazyme	T101/201-01
明胶	生工	A85522-50G
L-谷氨酰胺	生工	A6600224

2.1.4 实验仪器与耗材

本研究所用的实验仪器与耗材如下表所示（表 2-7）：

表 2-7 实验仪器与耗材

仪器名称	型号	生产厂家
-80℃冰箱	BCD-649WE	海尔

仪器名称	型号	生产厂家
二氧化碳培养箱	INC153H9	MEMMERT
二级生物洁净安全柜	BHL-1300IIIH2	江苏泰安
倒置显微镜	Motic-AE2000	麦克奥迪
研究级倒置荧光显微镜	IX71	OLYMPUS
激光共聚焦显微镜	A1R+Ti2-E	Nikon
气浴震荡培养箱	SHZ-82	常州国华
普通 PCR 仪	2720	Life technologies
实时定量 PCR 仪	ABI-Q5	Life technologies
液氮罐	30 L	上海绿地
高速 4℃离心机	TGL-16M	上海卢湘仪
低速离心机	TCL-151P	上海卢湘仪
恒温水浴锅	HH-2	常州朗越
LDZX 型立式压力蒸汽灭菌器	LDZZX-50KBS	上海申安
凝胶成像分析仪	WD-9413C	上海予申
电泳仪	EPS-300	上海天能
垂直电泳槽	VE-180	上海天能
水平电泳槽	VE-90	上海天能
转移电泳槽	VE-186	上海天能
制冰机	TIVIS-30	雪科
漩涡混合器	XH-2	常州朗越
分光光度计	UV-235	龙尼柯（上海）
干式恒温器	DTH-100	常州朗越
电热鼓风干燥箱	DHG-9030A	上海一恒科学
单管式化学发光仪	PromegaGlomax20/20	Promega
化学发光成像系统	Fx6	Viber

2.2 实验方法

2.2.1 分子克隆

1. **BiFC 实验载体构建：**分别以 pBiFC-CBX7A/B/C-F/R 为引物，以 cDNA

为模板克隆 CBX7A/B/C 片段,得到的产物用 Xba I 和 BamH I 进行酶切反应并回收得到插入片段,利用 Xba I 和 BamH I 对载体 pBiFC-VN 进行酶切反应并回收得到载体片段,将回收的插入片段和载体片段进行连接得到载体 pBiFC-VN-CBX7A/B/C; 分别以 pBiFC-CBX7A (1-82)/(83-251)/(Δ CD)/(Δ CD+ATHL)-F/R、pBiFC-CBX7C (1-110)/(111-166)/(Δ CD+ATHL)-F/R 和 pBiFC-PHC2 (Δ SAM)-F/R 为引物,以载体 pBiFC-VN-CBX7A/C 和 BiFC-YC-PHC2 为模板克隆 CBX7A (1-82)/(83-251)/(Δ CD)/(Δ CD+ATHL)、CBX7C (1-110)/(111-166)/(Δ CD+ATHL) 和 PHC2(Δ SAM)片段,得到的产物用 Xba I 和 BamH I 进行酶切反应并回收得到插入片段,利用 Xba I 和 BamH I 对载体 pBiFC-VN 和 pBiFC-YC 进行酶切反应并回收得到载体片段,将回收的插入片段和载体片段进行连接得到载体 pBiFC-VN-CBX7A (1-82)/(83-251)/(Δ CD)/(Δ CD+ATHL)、pBiFC-VN-CBX7C (1-110)/(111-166)/(Δ CD+ATHL)和 pBiFC-YC-PHC2 (Δ SAM); 分别以 pBiFC-CBX7A (Δ ATHL)-1F/R 和 pBiFC-CBX7C (Δ ATHL)-2F/R 为引物,以载体 pBiFC-VN-CBX7A 为模板克隆 CBX7A (Δ ATHL)-1 和 CBX7A (Δ ATHL)-2 片段,以 pBiFC-CBX7A (Δ ATHL)-1F 和 pBiFC-CBX7C (Δ ATHL)-2R 为引物,以片段 CBX7A (Δ ATHL)-1 和 CBX7A (Δ ATHL)-2 为共同模板,通过搭桥 PCR 得到 CBX7A (Δ ATHL)片段,得到的产物用 Xba I 和 BamH I 进行酶切反应并回收得到插入片段,利用 Xba I 和 BamH I 对载体 pBiFC-VN 进行酶切反应并回收得到载体片段,将回收的插入片段和载体片段进行连接得到载体 pBiFC-VN-CBX7A (Δ ATHL)。

2. 原核表达载体构建: 分别以 PGEX-6p-1-CBX7A/(1-82)/(83-251)-F/R、PGEX-6p-1-CBX7C/(1-51)/(1-110)/(52-166)/(111-166)-F/R 和 PGEX-6p-1-EGFP-F/R 为引物,以载体 pBiFC-VN-CBX7A/C 和 pcDNA3.1-EGFP-Flag 为模板克隆 CBX7A/(1-82)/(83-251)、CBX7C/(1-51)/(1-110)/(52-166)/(111-166)和 EGFP 片段,得到的产物用 BamH I 和 Xho I 进行酶切反应并回收得到插入片段,利用 BamH I 和 Xho I 对载体 PGEX-6p-1 进行酶切反应并回收得到载体片段,将回收的插入片段和载体片段进行连接得到载体 PGEX-6p-1-CBX7A/(1-82)/(83-251)、PGEX-6p-1-CBX7C/(1-51)/(1-110)/(52-166)/(111-166) 和 PGEX-6p-1-EGFP; 分别以 PGEX-6p-1-EGFP-protein-F/R、PGEX-6p-1-EGFP-CBX7A-F/R 和 PGEX-6p-1-EGFP-CBX7C-F/R 为引物,以载体 pBiFC-VN-CBX7A/C 和 pcDNA3.1-EGFP-Flag 为模板克隆 EGFP 和 CBX7A/C 片段,分别以 PGEX-6p-1-EGFP-protein-F 和 PGEX-6p-1-EGFP-CBX7A/C-R 为引物,以片段 EGFP 和 CBX7A/C 为共同模板,通过搭桥 PCR 得到 EGFP-CBX7A/C 片段,得到的产物用 BamH I 和 Xho I 进行酶切反应并回收得到插入片段,利用 BamH I 和 Xho I 对载体 PGEX-6p-1 进行酶切反应并回收得到载体片段,将回收的插入片段和载体片段进行连接得到载体

PGEX-6p-1-EGFP-CBX7A/C; 分别以 pET28a-RING1B-F/R 和 pET28a-PHC2-F/R 为引物, 以载体 pBiFC-YC-RING1B 和 pBiFC-YC-PHC2 为模板克隆 RING1B 和 PHC2 片段, 得到的产物用 Nde I 和 Xho I 进行酶切反应并回收得到插入片段, 利用 Nde I 和 Xho I 对载体 pET28a 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pET28a-RING1B 和 pET28a-PHC2。

3. **真核可诱导性表达载体构建:** 以 pcDNA3.1-TRE-pCMV-CBX7C-EGFP-Flag-F/R 为引物, 以载体 pBiFC-VN-CBX7C 为模板克隆 CBX7C 片段, 得到的产物用 EcoR I 和 Cla I 进行酶切反应并回收得到插入片段, 利用 EcoR I 和 Cla I 对载体 pcDNA3.1-TRE-pCMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-TRE-pCMV-CBX7C-EGFP-Flag。

4. **真核表达载体构建:** 以 pcDNA3.1-pCMV-CBX7C-EGFP-Flag-F/R 为引物, 以载体 pBiFC-VN-CBX7C 为模板克隆 CBX7C 片段, 得到的产物用 Hind III 和 Cla I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 Cla I 对载体 pcDNA3.1-CMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-pCMV-CBX7C-EGFP-Flag; 以 pBiFC-EGFP-PHC2-F/R 为引物, 以载体 pTripZ 为模板克隆 mCherry 片段, 得到的产物用 Hind III 和 BamH I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 BamH I 对载体 pBiFC-YC-PHC2 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pBiFC-mCherry-PHC2。

5. **异位靶向表达载体构建:** 以 pcDNA3.1-CMV-mCherry-HTM-1F/R 和 pcDNA3.1-CMV-mCherry-HTM-2F/R 为引物, 以载体 pTripZ 和 cDNA 为模板克隆 mCherry 和 HTM 片段, 以 pcDNA3.1-CMV-mCherry-HTM-1F 和 pcDNA3.1-CMV-mCherry-HTM-2R 为引物, 以片段 mCherry 和 HTM 为共同模板, 通过搭桥 PCR 得到 mCherry-HTM 片段, 得到的产物用 Hind III 和 Xba I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 Xba I 对载体 pcDNA3.1-CMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-pCMV-mCherry-HTM; 以 pcDNA3.1-CMV-mCherry-HTM-protein-F/R 和 pcDNA3.1-pCMV-mCherry-HTM-CBX7A/C-F/R/PHC2-F/R 为引物, 以载体 pcDNA3.1-pCMV-mCherry-HTM、pBiFC-VN-CBX7A/C 和 pBiFC-YC-PHC2 为模板克隆 mCherry-HTM、CBX7A、CBX7C 和 PHC2 片段, 以 pcDNA3.1-pCMV-mCherry-HTM-protein-F 和 pcDNA3.1-pCMV-mCherry-HTM-CBX7A/C/PHC2-R 为引物, 以片段 mCherry-HTM 和 CBX7A/CBX7C/PHC2 为共同模板, 通过搭桥 PCR 得到 mCherry-HTM-CBX7A/C/PHC2 片段, 得到的产物用 Hind III

和 Xba I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 Xba I 对载体 pcDNA3.1-CMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-CMV-mCherry-HTM-CBX7A/C/PHC2。

6. optoDroplet 表达载体构建: 以 pcDNA3.1-pCMV-mCherry-Cry2oligo-1F/R 和 pcDNA3.1-pCMV-mCherry-Cry2oligo-2F/R 为引物, 以载体 pTripZ 和拟南芥 cDNA 为模板克隆 mCherry 和 Cry2oligo 片段, 以 pcDNA3.1-pCMV-mCherry-Cry2oligo-1F 和 pcDNA3.1-pCMV-mCherry-Cry2oligo-2R 为引物, 以片段 mCherry 和 Cry2oligo 为共同模板, 通过搭桥 PCR 得到 mCherry-Cry2oligo 片段, 得到的产物用 Hind III 和 Xba I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 Xba I 对载体 pcDNA3.1-pCMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-pCMV-mCherry-Cry2oligo; 以 pcDNA3.1-pCMV-CBX7A/C-mCherry-Cry2oligo-F/R 和 pcDNA3.1-pCMV-protein-mCherry-Cry2oligo-F/R 为引物, 以载体 pBiFC-VN-CBX7A/C 和 pcDNA3.1-CMV-mCherry-Cry2oligo 为模板克隆 CBX7A/C 和 mCherry-Cry2oligo 片段, 以 pcDNA3.1-pCMV-CBX7A/C-mCherry-Cry2oligo-F 和 pcDNA3.1-pCMV-protein-mCherry-Cry2oligo-R 为引物, 以片段 CBX7A/CBX7C 和 mCherry-Cry2oligo 为共同模板, 通过搭桥 PCR 得到 CBX7A-mCherry-Cry2oligo 和 CBX7A-mCherry-Cry2oligo 片段, 得到的产物用 Hind III 和 Xba I 进行酶切反应并回收得到插入片段, 利用 Hind III 和 Xba I 对载体 pcDNA3.1-CMV-CBX7A-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段进行连接得到载体 pcDNA3.1-CMV-CBX7A/C-mCherry-Cry2oligo。

7. 荧光蛋白报告载体构建: 以 pcDNA3.1-5×UAS-CYP26A1-EGFP-Flag-1F/R 和 pcDNA3.1-5×UAS-CYP26A1-EGFP-Flag-2F/R 为引物, 以载体 pcDNA3.1-5×UAS-5×GFP sg 和 cDNA 为模板克隆 5×UAS 和 CYP26A1 片段, 以 pcDNA3.1-5×UAS-CYP26A1-EGFP-Flag-1F 和 pcDNA3.1-5×UAS-CYP26A1-EGFP-Flag-2R 为引物, 以片段 5×UAS 和 CYP26A1 为共同模板, 通过搭桥 PCR 得到 5×UAS-CYP26A1 片段, 回收后得到插入片段, 利用 Xba I 和 Age I 对载体 pcDNA3.1-TRE-CMV-EGFP-Flag 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段通过同源重组进行连接得到载体 pcDNA3.1-5×UAS-CYP26A1-EGFP-Flag; 以 GAL4 DBD-CBX7A/C-F/R 为引物, 以载体 pBiFC-VN-CBX7A/C 为模板克隆 CBX7A/C 片段, 回收后得到插入片段, 利用 Xba I 和 Bgl II 对载体 GAL4 DBD 进行酶切反应并回收得到载体片段, 将回收的插入片段和载体片段通过同源重组进行连接得到载体 GAL4 DBD-CBX7A/C。

8. 双荧光素酶报告系统载体构建：以 GAL4 DBD-IRES-Renilla Luciferase-1F/R 和 GAL4 DBD-IRES-Renilla Luciferase-2F/R 为引物，以载体 pTripZ 和 GAL4 DBD 为模板克隆 IRES 和 Renilla Luciferase 片段，以 GAL4 DBD-IRES-Renilla Luciferase-1F 和 GAL4 DBD-IRES-Renilla Luciferase-2R 为引物，以片段 IRES 和 Renilla Luciferase 为共同模板，通过搭桥 PCR 得到 IRES-Renilla Luciferase 片段，回收后得到插入片段，利用 EcoR I 和 Bgl II 对载体 GAL4 DBD 进行酶切反应并回收得到载体片段，将回收的插入片段和载体片段通过同源重组进行连接得到载体 GAL4 DBD-IRES-Renilla Luciferase；以 GAL4 DBD-CBX7A/C-IRES-Renilla Luciferase-F/R 为引物，以载体 pBiFC-VN-CBX7A/C 为模板克隆 CBX7A/C 片段，回收后得到插入片段，利用 EcoR I 和 Bgl II 对载体 GAL4 DBD 进行酶切反应并回收得到载体片段，将回收的插入片段和载体片段通过同源重组进行连接得到载体 GAL4 DBD-CBX7A/C-IRES-Renilla Luciferase。

2.2.2 组织及细胞 RNA 提取、逆转录及 qPCR 检测

1. RNA 提取

- (1) 试剂及耗材：总 RNA 提取试剂盒、 β -巯基乙醇，RNase free 枪头和 EP 管等；
- (2) 细胞或组织：PGK12.1，小鼠睾丸、子宫、脾脏、肾脏肝脏和心脏组织；
- (3) RNA 提取方法参照总 RNA 提取试剂盒说明书中步骤。

2. 逆转录

- (1) gDNA 消化。在 RNase free 的离心管中预混以下溶液（表 2-8），42℃放置 2 min 后置于室温 5 min：

表 2-8 gDNA 消化体系配制

成分	用量
Total RNA	2 μ g
5 $\times$ gDNA Digester Mix	3 μ L
RNase free H ₂ O	至 15 μ L

- (2) 逆转录反应体系配制。向上述管中加入以下试剂（表 2-9），并轻轻混匀：

表 2-9 逆转录反应体系配制

成分	用量
第一步的反应液	15 μ L
4 $\times$ HafairIII SuperMix plus	5 μ L

- (3) 将上述逆转录反应体系依次置于 25℃/5 min 和 55℃/15 min，最终将其置于 85℃/5 min 以终止反应；

(4) 所得 cDNA 可置于-80℃进行保存。

3. qPCR 检测

- (1) 将上述得到的 cDNA 稀释 10 倍后作为 qPCR 检测模板；
- (2) 配制 10 μ L 反应体系，以三步法进行 qPCR 反应，反应结束后使用 $2^{-\Delta\Delta CT}$ 法计算目的基因的相对表达量。

2.2.3 可诱导性表达 CBX7C-EGFP-Flag 细胞系构建

电转法构建 CBX7C-EGFP-Flag 稳定表达的 mESC 细胞系：对质粒 pcDNA3.1-TRE-CMV-CBX7C-EGFP-Flag 进行扩增提取，需使用去除内毒素质粒提取试剂盒。将细胞消化下来离心后弃掉上清，用无血清的 DNEM 进行重悬，按比例加入 5 $\times$ 电转适配液，以及适量的质粒（按照 6-8 μ g 质粒/ 1×10^7 个细胞）。将 200 μ L 细胞-DNA 混合液加入到电转杯（2 mm）中并置于电转槽中按照 450 V-0.5 sec 的参数进行操作。电转后将细胞平铺到培养皿中，于 24 h 后进行换液，继续培养至细胞覆盖率 70%-80%左右，以 1 mg/mL Neomycin 的筛选浓度进行筛选。荧光显微镜下进行荧光鉴定。

2.2.4 双分子荧光互补实验（BiFC）

1. 试剂及材料：
- (1) 试剂：ExFect Transfection Reagent、DMEM 和 BiFC 相关质粒；
- (2) 细胞：HEK293T 或 HeLa。
2. 配制转染体系（表 2-10），室温静置 15 min 后滴加至细胞培养基中：

表 2-10 转染体系配制（24 孔板）

成分	用量
DMEM	100 μ L
质粒	共 1 μ g
ExFect Transfection Reagent	2 μ L

3. 转染 36 h 后可置于倒置荧光显微镜或激光共聚焦观察（提前将细胞接种与盖玻片上），以阴性对照组为参照，分析目的蛋白之间是否具有互作。

2.2.5 pull down 实验（GST-pull down 和 His-pull down）

1. 试剂及材料：
- (1) 试剂：LB 液体和固体培养基、PBS、抗生素（Amp⁺和 Kan⁺）、IPTG、BeaverBeads™ GSH 或 BeaverBeads™ IDA-Co，实验相关质粒，以及表 2-11 和表 2-12 中的各缓冲液：

表 2-11 GST-pull down 体系缓冲液配制

缓冲液	成分	终浓度
Binding buffer (pH 7.4)	NaCl	140 mM
	KCl	2.7 mM
	Na ₂ HPO ₄	10 mM
	KH ₂ PO ₄	1.8 mM
	PMSF	1 mM
Washing buffer (pH 7.4)	NaCl	140 mM
	KCl	2.7 mM
	Na ₂ HPO ₄	10 mM
	KH ₂ PO ₄	1.8 mM
	DTT	2 mM
	Triton X-100	1.5%
	PMSF	1 mM

表 2-12 His-pull down 体系缓冲液配制

缓冲液	成分	终浓度
Binding buffer (pH 7.4)	Phosphate buffer	20 mM
	NaCl	500 mM
	Imidazole	20 mM
	PMSF	1 mM
Washing buffer (pH 7.4)	Phosphate buffer	20 mM
	NaCl	500 mM
	Imidazole	40 mM
	Reduced glutathione	5 mM
	Triton X-100	1.5%
	PMSF	1 mM

(2) 菌种: *E. coli* Rosseta

2. 蛋白原核表达:

(1) 将质粒转化至 *E. coli* Rosseta 并在 LB 固体培养基上生长, 挑取单菌落接种于 LB 液体培养基中生长, 待细胞对数生长期 ($OD \approx 0.6-0.7$) 时加入 IPTG (终浓度为 0.5 mM) 后在 28°C 条件下 180 rpm/8 h 诱导蛋白表达;

(2) 在 4°C 条件下, 5000 rpm 离心 10 min, 弃上清, 收集菌体, 以 2 倍于菌体体

积的预冷 PBS 重悬后 5000 rpm 离心 5 min，弃上清，再次收集菌体；

(3) 用适当体积的预冷 Binding buffer 重悬菌体，在冰浴条件下超声破碎，4℃条件下 12000 rpm 离心 10 min，收集上清；

3. 蛋白纯化：

(1) 取适量含有 BeaverBeads™ GSH（与 GST 融合蛋白结合）或 BeaverBeads™ IDA-Co（与 His 融合蛋白结合）的溶液，并在磁力架上收集 beads 至离心管底部，弃上清，加入 10 倍于 beads 体积的预冷 PBS 重悬后在磁力架上收集 beads，弃上清，重复 3 次；

(2) 将上述蛋白裂解液与处理后的 BeaverBeads™ GSH 或 BeaverBeads™ IDA-Co 在 4℃条件下旋转孵育 1 h；

(3) 在磁力架上收集 beads，弃上清，加入 10 倍于 beads 体积的预冷 Binding buffer 重悬后在磁力架上收集 beads，弃上清，重复 5 次。

4. 体外蛋白结合及检测：

(1) 将结合有 GST 融合蛋白的 BeaverBeads™ GSH 或结合有 His 融合蛋白的 BeaverBeads™ IDA-Co 与适量体积的 His 融合蛋白或 GST 融合蛋白裂解液在 4℃条件下旋转孵育 2 h；

(2) 在磁力架上收集 beads，弃上清，加入 10 倍于 beads 体积的预冷 Washing buffer 重悬后在磁力架上收集 beads，弃上清，重复 5 次；

(3) 在磁力架上收集 beads，弃上清，加入适量体积的 Loading buffer 重悬 beads，置于 100℃金属浴孵育 10 min 后放入-80℃进行保存；

(4) 通过 Western blot 进行检测。

2.2.6 蛋白质免疫共沉淀实验（co-IP）

1. 试剂及材料：

(1) 试剂：PBS、Protein A/G IP beads、抗体以及表 2-13 中的各缓冲液：

表 2-13 co-IP 体系缓冲液配制

缓冲液	成分	终浓度
Cell lysis buffer	Tris-HCl	20 mM
	NaCl	300 mM
	EDTA	0.5 mM
	Na ₃ VO ₄	1 mM
	NaF	50 mM
	NP-40	1%

缓冲液	成分	终浓度
	Triton X-100	1%
	PMSF	1 mM
	Protease Inhibitor Cock tall	1 Tab
Block buffer	BSA of PBS	4 mg/mL

(2) 细胞: mESC: iCBX7C-EGFP-Flag, HEK293T: iCBX7C-EGFP-Flag

2. 细胞处理及与抗体孵育:

(1) 收集细胞, 用适量体积的预冷 PBS 重悬细胞, 4°C条件下, 800 rpm 离心 5 min, 弃上清, 再次收集细胞;

(2) 取适量体积的预冷 Cell lysiss buffer 重悬细胞, 在 4°C条件下旋转裂解 2 h;

(3) 在 4°C条件下, 12000 rpm 离心 10 min, 收集上清, 取 5%蛋白裂解液作为 Input 样品待测;

(4) 取适量体积的抗体加入上述蛋白裂解液, 在 4°C条件下旋转孵育过夜。

3. Protein A/G IP beads 封闭:

(1) 取适量含有 Protein A/G IP beads 的溶液, 并在磁力架上收集 beads 至离心管底部, 弃上清, 加入 10 倍于 beads 体积的预冷 PBS 重悬后在磁力架上收集 beads, 弃上清, 重复 3 次;

(2) 以 10 倍于 beads 体积的含有 4 mg/mL BSA 的 PBS 溶液重悬 beads, 在 4°C条件下旋转封闭过夜。

4. 免疫共沉淀及检测:

(1) 在磁力架上收集封闭过夜的 beads, 弃上清, 加入 10 倍于 beads 体积的预冷 PBS 重悬后在磁力架上收集 beads, 弃上清, 重复 3 次;

(2) 用上述抗原-抗体混合液重悬处理后的 beads, 在 4°C条件下旋转孵育 4 h;

(3) 在磁力架上收集 beads, 弃上清, 加入适量体积的 Loading buffer 重悬 beads, 置于 100°C金属浴孵育 10 min 后放入-80°C进行保存;

(4) 通过 Western blot 进行检测。

2.2.7 免疫荧光

1. 试剂及材料:

(1) 试剂: PBS、一抗、荧光二抗以及表 2-14 中的各缓冲液:

表 2-14 IF 体系缓冲液配制

缓冲液	成分	终浓度
细胞固定液	Formaldehydel of PBS	4%
细胞通透液	Triton X-100	0.5%

缓冲液	成分	终浓度
细胞封闭液	BSA of PBS	3%
TBST	Tris-HCl (pH 7.6)	50 mM
	NaCl	150 mM
	Tween-20	0.1%

(2) 细胞：生长在圆形盖玻片上的细胞

2. 细胞固定：弃去培养基，加入预冷的 PBS 洗涤 3 次后弃去洗涤液，加入适量体积的含 4% 多聚甲醛的 PBS 溶液，室温静置固定 15 min；
3. 细胞通透：弃去固定液，用预冷的 PBS 洗涤 3 次后弃去洗涤液，加入适量体积的含 0.5% Triton X-100 的 PBS 溶液，室温静置孵育 15 min；
4. 细胞封闭：弃去通透液，用预冷的 PBS 洗涤 3 次后弃去洗涤液，加入适量体积的含 3% BSA 的 PBS 溶液，室温静置封闭 60 min；
5. 免疫反应：

- (1) 弃去封闭液，用预冷的 PBS 洗涤 3 次后弃去洗涤液。取 50 μ L 相应的一抗稀释液（用 TBST 以 1:200 稀释）滴加在铺有封口膜的盖玻片上。取出培养皿中的盖玻片置于抗液体滴上，使细胞浸润与抗体稀释液中，将载玻片置于湿盒中，于 4°C 条件下静置孵育过夜；
- (2) 将盖玻片放回培养皿中，用预冷的 TBST 洗涤 10 min，并重复 3 次，以相同的方法将盖玻片置于荧光二抗液滴上，于室温静置孵育 1 h；
- (3) 将盖玻片放回培养皿中，用预冷的 TBST 洗涤 10 min，并重复 3 次，保存于 4°C；
- (4) 将盖玻片置于载玻片上用封片剂封片后，即可用于共聚焦成像分析。

2.2.8 HTM 介导的细胞内异位靶向实验

1. 试剂及材料：

- (1) 试剂：ExFect Transfection Reagent、DMEM 和相关质粒；
- (2) 细胞：mESC 或 HeLa。
2. 根据表 2-8 配制转染体系，室温静置 15 min 后滴加至细胞培养基中；
3. 转染 36 h 后可置于倒置荧光显微镜或激光共聚焦观察（提前将细胞接种与盖玻片上）；或进一步进行免疫荧光检测分析。

2.2.9 原核蛋白诱导表达及纯化

1. 试剂及材料：

- (1) 试剂：LB 液体和固体培养基、PBS、Amp⁺ 抗生素、IPTG、BeaverBeads™ GSH，

实验相关质粒以及表 2-9 中所示缓冲液;

(2) 菌种: *E. coli* Rosseta

2. 原核蛋白诱导表达

(1) 将表达载体转化至原核表达菌株 Rosetta, 挑取单克隆接种于适量体积的含 Amp 抗生素的 LB 液体培养基中培养;

(2) 培养至细胞对数生长期(OD_{600} 为 0.6-0.7)时加入 IPTG(终浓度为 0.5 mM), 低温低速诱导 8 h。

3. 原核表达蛋白纯化

(1) 将所得菌液在 4°C 条件下, 5000 rpm 离心 10 min, 弃上清, 收集菌体, 以 2 倍于菌体体积的预冷 PBS 重悬后 5000 rpm 离心 5 min, 弃上清, 再次收集菌体;

(2) 用适当体积的预冷 Binding buffer 重悬菌体, 在冰浴条件下超声破碎, 4°C 条件下 12000 rpm 离心 10 min, 收集上清;

(3) 参照 BeaverBeads™ GSH 的蛋白纯化使用手册对 GST 融合蛋白进行蛋白亲和纯化。纯化蛋白用 BCA 法进行蛋白定量后, 离心管分装, 于 -80°C 保存备用。

2.2.10 纯化蛋白体外凝集实验分析

5. 试剂及材料:

(1) 试剂: 1, 6-己二醇, PBS;

(2) 材料: 纯化的 GST 融合蛋白。

6. 各样品浓度校正: 参照试剂盒说明书, 应用 BCA 方法测量纯化蛋白的浓度, 然后用 PBS 将各样品稀释校正为 12 μ M;

3. 纯化蛋白在离心管内凝集沉淀: 将上述蛋白样品在室温下静置孵育 20 min 后以 12000 rpm 离心 10 min, 在紫外线下观察离心管底部是否有蛋白凝集沉淀。

4. (4) 纯化蛋白凝集颗粒的显微观察: 参考 Ann Boija 等人报道的实验方法^[213], 使用双面胶带将两个盖玻片粘合在一起, 形成一个狭窄的腔室。将上述离心的蛋白质样品再次充分悬浮, 并将少量的蛋白质样品注入腔室中, 以便于在倒置荧光显微镜下以 488 nm 激发光激发成像。

5. 1, 6-己二醇处理:

(1) 将高浓度的 1, 6-己二醇溶液添加至纯化蛋白溶液中, 使其终浓度为 10%或 50%, 室温孵育 2 min 后以 12000 rpm 离心 10 min, 在紫外线下观察离心管底部是否有蛋白凝集沉淀。

(2) 将上述离心的蛋白质样品再次充分悬浮, 并将其注入盖玻片腔室中, 在倒置

荧光显微镜下成像。使用 ImageJ 软件分析每个样本的不同视野中的荧光颗粒数，并以直方图显示数据。

2.2.11 荧光漂白后恢复实验分析 (FRAP)

1. 试剂及材料:

(1) 试剂: ExFect Transfection Reagent、DMEM 和相关质粒;

(2) 细胞: HeLa。

2. 将 HeLa 细胞接种到 35mm 的共聚焦专用培养皿中培养，根据表 2-8 配制转染体系，室温静置 15 min 后滴加至细胞培养基中:

3. 转染 36 h 后使用激光扫描共聚焦显微镜在 100 倍镜下观察并获取细胞图像。在光漂白之前在 488 nm 的激发波长下采集一次图像。此后，在选定区域中用强度为 2% 的 488 nm 激光进行光漂白，随后立即以适当的时间间隔采集图像。同时，应用 NIS Elements 软件实时采集漂白区域的荧光值。计算漂白后的荧光强度与漂白前的比值，以获得分析颗粒动力学特征的荧光恢复曲线分析。

2.2.12 Cry2oligo 介导的光激活细胞内相分离实验

1. 试剂及材料:

(1) 试剂: ExFect Transfection Reagent、DMEM 和相关质粒;

(2) 细胞: HeLa。

2. 将 HeLa 细胞接种到 35mm 的共聚焦专用培养皿中培养，根据表 2-8 配制转染体系，室温静置 15 min 后滴加至细胞培养基中:

3. 转染 36 h 后，将表达与 mCherry-Cry2oligo 融合的蛋白细胞激光共聚焦观察；同时打开 488 nm 和 543 nm 的激光通道，设置 488 nm 通道的激光强度为 0.1%，分辨率为 1024×1024；设置连续激发扫描程序：488 nm 激光激发前用 560 nm 激光通道采集蓝光激发前的细胞图像，此后，488 nm 激光通道持续激发选定区域，在此过程中，每隔 30 s 利用 560 nm 激光通道进行图像采集，观察在蓝光激发后，胞内的荧光信号聚集的情况。

2.2.13 高分辨率活细胞成像系统 (DeltaVision) 实验

1. 试剂及材料:

(1) 试剂: ExFect Transfection Reagent、DMEM 和相关质粒;

(2) 细胞: HeLa。

2. 将 HeLa 细胞接种到 35mm 的共聚焦专用培养皿中培养，根据表 2-8 配制转染体系，室温静置 15 min 后滴加至细胞培养基中:

3. 转染 18 h 后, 将细胞置于成像系统中用于活细胞培养的环境控制箱, 使细胞在 37°C, 5% CO₂ 的环境中正常生长;
4. 选取多个视野, 在 60 倍镜下每 3 小时采集一次图像, 观察 CBX7C-PHC2 互作凝集体的形成及其与染色质位置关系变化的过程。

2.2.14 染色质免疫共沉淀 (ChIP) 及 qPCR 实验分析

1. 试剂及材料:

- (1) 试剂: PBS、Protein A/G IP beads、salmon sperm DNA、37% Formaldehyde、抗体、RNase、蛋白酶 K、苯酚/异戊醇混合液、Glycogen 以及表 2-15 中的各缓冲液:

表 2-15 ChIP 体系缓冲液配制

缓冲液	成分	终浓度
Lysis buffer 1	HEPES-KOH (pH 7.5)	50 mM
	NaCl	140 mM
	EDTA	1 mM
	Glycerol	10%
	NP-40	0.5%
	Triton X-100	0.25%
	PMSF	1 mM
	Protease Inhibitor Cock tall	1 Tab
Lysis buffer 2	Tris-HCl (pH 8.0)	10 mM
	NaCl	200 mM
	EDTA	1.5 mM
	PMSF	1 mM
	Protease Inhibitor Cock tall	1 Tab
Lysis buffer 3	Tris-HCl (pH 8.0)	10 mM
	NaCl	300 mM
	EDTA	1.5 mM
	Na-Deoxycholate	0.1%
	Na-Laurylsarcosine	0.5%
	PMSF	1 mM
	Protease Inhibitor Cock tall	1 Tab
Low salt wash buffer	Tris-HCl (pH 8.0)	20 mM

缓冲液	成分	终浓度
	NaCl	150 mM
	EDTA	1 mM
	SDS	0.1%
	Triton X-100	1%
High salt wash buffer	Tris-HCl (pH 8.0)	20 mM
	NaCl	350 mM
	EDTA	1 mM
	SDS	0.1%
	Triton X-100	1%
LiCl wash buffer	Tris-HCl (pH 8.0)	10 mM
	LiCl	250 mM
	EDTA	1 mM
	NP-40	0.5%
	Na-Deoxycholate	0.5%
Block buffer	BSA of PBS	0.5%
Elution Buffer	Tris-HCl (pH 8.0)	50 mM
	EDTA	10 mM
	SDS	1%
TE Buffer	Tris-HCl (pH 8.0)	10 mM
	EDTA	1 mM

(2) 细胞: mESC: iCBX7C-EGFP-Flag。

2. 细胞收集并交联:

- (1) 准备足量的细胞由于实验检测 (通常每个反应需 1×10^8 细胞左右), 弃去细胞培养皿中的培养基后用预冷的 PBS 洗涤 3 次, 弃去洗涤液, 加入含有 1% 的 Formaldehyde 的培养基, 于室温下静置交联 10 min;
- (2) 向上述培养基中加入适量体积的 2.5 M 甘氨酸 (终浓度为 0.125 M) 并混匀, 于室温下静置 5 min 而终止交联;
- (3) 弃去培养基, 用预冷的 PBS 洗涤 3 次, 弃去洗涤液, 加入适量体积的预冷 PBS 后用细胞刮刀收集细胞于 50 mL 离心管, 在 4°C 条件下, 5000 rpm 离心 5 min, 弃去上清, 收集细胞;

3. 细胞裂解及超声破碎:

- (1) 取适量 Lysis buffer 1 重悬上述交联后的细胞, 在 4°C 条件下旋转孵育裂解 10

min;

- (2) 在 4℃条件下, 5000 *rpm* 离心 5 min, 弃去上清, 收集细胞裂解沉淀, 加入与 Lysis Buffer 1 等量的 Lysis Buffer 2 充分重悬细胞裂解沉淀, 在 4℃条件下旋转孵育裂解 10 min;
- (3) 在 4℃条件下, 5000 *rpm* 离心 5 min, 弃去上清, 收集细胞裂解沉淀, 加入适量体积的 Lysis Buffer 3 充分重悬细胞裂解沉淀;
- (4) 将上述细胞裂解液置于冰水浴中, 进行超声处理 (70 W, 3 sec ON/7 sec OFF, 45 min);
- (5) 在超声后的细胞裂解液中加入 TritonX-100(终浓度为 1%), 并在 4℃条件下, 13000 *rpm* 离心 10 min, 收集上清;
- (6) 对上述裂解液进行 OD₂₆₀ 检测, 用 Lysis Buffer 3 稀释校正每个反应样品的 OD₂₆₀ 的数值为 20-25/mL, 并取出 5%的细胞裂解液作为后续检测的 Input 样品, 保存于-80℃条件下待处理。

4. Protein A/G IP beads 封闭:

- (1) 取适量含有 Protein A/G IP beads 的溶液, 并在磁力架上收集 beads 至离心管底部, 弃上清, 加入 10 倍于 beads 体积的预冷 PBS 重悬后在磁力架上收集 beads, 弃上清, 重复 3 次;
- (2) 以适量体积的 Block buffer 重悬 beads, 并加入适量体积的浓度为 0.5 mg/mL 的 salmon sperm DNA, 在 4℃条件下旋转封闭过夜;
- (3) 在磁力架上收集 beads 至离心管底部, 弃上清, 用适量体积的含有 1% TritonX-100 的 Lysis Buffer 3 重悬 beads, 重复 3 次;

5. 免疫沉淀反应:

- (1) 将处理后的 beads 加入上述细胞裂解液中, 并同时加入适量体积的抗体, 三者构成的免疫反应体系在 4℃条件下旋转孵育过夜;
- (2) 在磁力架上收集 beads 至离心管底部, 弃上清, 依次用 Low salt wash buffer、High salt wash buffer、LiCl wash buffer 和 TE Buffer 在 4℃条件下旋转洗涤 beads 20 min, 每种缓冲液重复洗涤 3 次;
- (3) 在磁力架上收集 beads 至离心管底部, 弃去洗涤液后加入适量体积的 Elution Buffer 并充分重悬 beads。

6. 蛋白质-染色质解交联:

- (1) 将上述免疫反应样品置于 65℃水浴锅中孵育 30 min, 使免疫反应中的蛋白质-染色质交联物与 beads 解离并融于溶液中, 在室温条件下 12000 *rpm* 离心 2 min, 收集上清;
- (2) 将上述获得的上清溶液置于 65℃条件下静置孵育过夜, 使得蛋白质与 DNA

解交联；同时 Elution Buffer 将前述的 Input 样品补足至于其它样品同体积后置于同样条件下进行解交联；

- (3) 向解交联后的样品中加入等体积的 TE Buffer 稀释 Elution Buffer 中的 SDS 浓度；
- (4) 加入适量体积 RNAase（终浓度为 0.2 mg/mL）并混匀后，在 37°C 条件下孵育 1 h；
- (5) 加入适量体积蛋白酶 K（终浓度为 0.2 μg/mL）并混匀后，在 55°C 条件下孵育 2 h。

7. DNA 回收：

- (1) 向上述样品中加入等体积的苯酚/异戊醇混合液（P:C:IA=25:24:1），充分混匀后在常温条件下 12000 rpm 离心 5 min，取上层溶液；
 - (2) 向上述样品中加入 5 M NaCl（终浓度为 200 mM）和 20 μg/μL 的 Glycogen（终浓度为 0.075 μg/μL）溶液；
 - (3) 加入 2 倍于反应物体积的无水乙醇后置于 -80°C 条件下 10 min；
 - (4) 在 4°C 条件下 12000 rpm 离心 10 min，弃上清，DNA 与 Glycogen 沉淀与离心管底部；
 - (5) 加入 80% 乙醇，于 4°C 条件下 12000 rpm 离心 5 min，重复 2 次；
 - (6) 尽量去尽残留的乙醇后，置于通风处加速液体挥发；
 - (7) 用适量的 ddH₂O 充分重悬沉淀后于 -80°C 保存待用。
8. qPCR 检测：将 Input 样品进行 10 倍梯度稀释后作为模板，并获得标准曲线，基于此分析目的蛋白在靶基因上的富集度。

2.2.15 体外 EB 诱导实验及分析

1. 试剂及材料：

- (1) 试剂：PBS、去 LIF 的分化培养基；
 - (2) 细胞：mESC。
2. 从培养皿中消化正常培养的 mESC 并获得细胞悬液；
 3. 用 PBS 洗涤细胞，800 rpm 离心 3 min，弃去洗涤液，重复 2 次；
 4. 用去 LIF 的分化培养基重复重悬细胞，获得单细胞悬液；
 5. 对上述单细胞悬液进行细胞计数，将其稀释至约 2×10^5 个细胞/mL 的浓度（5000 个细胞/滴）。
 6. 悬滴法诱导 EB：将上述 mESC 细胞悬液以 25 μL/滴的体积逐滴置于培养皿盖上，然后翻转并将其盖在含有 10 ml PBS 的培养皿上，随后放入培养箱中待其分化；

7. 24-36 h 后,可观察到液滴中形成的白色颗粒状拟胚体,用微量移液管小心地将液滴转移至加有去 LIF 的分化培养基的培养皿中进行进一步培养;
8. 收集不同时间点的 EB,进行 RNA 和蛋白提取,用于后续的 qPCR 和 Western blot 检测。

2.2.16 双荧光素酶报告基因系统实验分析

1. 试剂及材料:

- (1) 试剂: PBS、双荧光素酶报告系统试剂盒、ExFect Transfection Reagent、DMEM 和相关质粒;
- (2) 细胞: HEK293T: GAL4-TK-Firefly Luciferase;
2. 将细胞接种至 24 孔板中,12 h 后细胞覆盖率达到约 80%-90%,根据表 2-8 配制转染体系,室温静置 15 min 后滴加至细胞培养基中;
3. 将 HeLa 细胞接种到 35mm 的共聚焦专用培养皿中培养,根据表 2-8 配制转染体系,室温静置 15 min 后滴加至细胞培养基中;
4. 细胞处理:
 - (1) 弃去培养基,加入 PBS 洗涤细胞,重复 3 次;
 - (2) 参照双荧光素酶报告系统试剂盒说明书,向上述细胞加入适量体积的 $1 \times$ PLB,在室温条件下置于水平摇床裂解 15 min;
 - (3) 收集培养皿中的细胞裂解物至离心管中,在 4°C 条件下 12000 rpm 离心 10 min,收集上清细胞裂解液。
5. 化学发光检测:参照说明书的推荐比例,取 $10 \mu\text{L}$ 细胞裂解液与 $50 \mu\text{L}$ LAR II buffer 在 1.5 mL 离心管中充分混匀后置于单管式化学发光仪中,首先快速读取 Firefly Luciferase 活性值,其后加入 $50 \mu\text{L}$ Stop&Glo[®] Buffer 后,继续放入单管式化学发光仪中读取作为内参的 Renilla Luciferase 活性值。

2.2.17 数据统计与分析

通过 Excel 和 Origin 2021 完成数据的分析及作图。使用 Photoshop CS4 进行图片整合处理。SPSS 用于分析显著性差异,主要应用 Student's *t* test 和 Duncan's multiple range test,通过 *t* 检验的显著性差异用星号表示 (ns, 不显著, $p > 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$),通过 Duncan 检验的显著性差异用不同小写字母表示。

第三章 实验结果

第一节 CBX7C 与 PHC2 蛋白互作促进功能性 cPRC1 组装并调节相分离特性

3.1 *Cbx7* 剪接体及其蛋白在 mESC 及组织中的时空表达模式分析

3.1.1 小鼠 *Cbx7* 的选择性剪接方式及其蛋白的结构域信息分析

基于真核细胞内发生的编码基因的选择性剪接多腺苷酸化事件, 小鼠 *Cbx7* 基因产生了三种转录物 (<http://genome.ucsc.edu>, mouse genome NCBI37/mm9), 我们将其分别命名为 *Cbx7A*、*Cbx7B* 和 *Cbx7C* (图 1-7)。据序列信息分析, *Cbx7A* 和 *Cbx7B* 的转录本均包含 6 个外显子。在转录过程中, 由于第 5 个外显子区域发生了内含子保留 (Intron Retention) 而转录产生了 *Cbx7A*。编码蛋白 CBX7A 和 CBX7B 都含有保守的 N 端 CD 结构域和 C 端 PC 结构域。CD 是 PRC1 与染色质结合的关键结构域。PC 结构域介导了 CBX 蛋白与 PRC1 复合体酶学核心组分 RING1B 的结合, 其在 PRC1 复合体的组装过程中发挥关键作用^[35, 94]。与 CBX7A 和 CBX7B 相比, CBX7C 仅保留了保守的 N 端 CD 结构域, 而缺失了 C 端的 PC 结构域, 并由于基因转录过程中在第 4 个外显子区域发生了内含子保留而产生独特的 C 端区域。目前, 有研究发现 CBX7A/B 参与调控发育相关等多种生物功能基因的转录, 但 CBX7C 剪切体的分子作用尚不清楚, 本论文进一步对 CBX7C 的表达模式、参与 PRC1 组装的分子机制及其转录调控活性进行了研究。

3.1.2 CBX7C 是一种在 mESC 及小鼠脾脏中特异性表达的 CBX7 蛋白亚型

为了深入研究不同 *Cbx7* 剪接体的生物学功能的异同, 我们首先在 mESC 和小鼠各种组织中检测了 *Cbx7A*、*B* 和 *C* 剪接体的表达模式。RT-qPCR 结果显示, 尽管 *Cbx7C* 的转录丰度远低于 *Cbx7A* 和 *Cbx7B*, 但是在 mESC 和小鼠脾脏组织中均表达三种 *Cbx7* 转录本 (图 3-1A 和 B)。随后, 为了探究 *Cbx7C* 转录本在蛋白水平上的表达模式, 我们利用原核表达的 His-SUMO-CBX7C (83-166 aa) (区别于 CBX7A 和 CBX7B 的区域) 融合蛋白制备抗原, 将其注射至新西兰大白兔体内进行免疫反应, 并通过亲和纯化获得了特异性识别 CBX7C 的多克隆抗体 (图 3-1C)。WB 实验分析显示, *Cbx7C* 转录本编码产生了分子量约为 20 kDa 的蛋白质产物 (图 3-1D), 表明 *Cbx7C* 确实能够作为编码基因在 mESC 和小鼠脾脏组

织中产生特异性表达的蛋白。综上所述, *Cbx7* 转录产生的不同转录本及其编码蛋白在 mESC 及小鼠组织中表现出不同的表达模式, 作为 CBX7 的蛋白亚型之一, CBX7C 在 mESC 和小鼠脾脏组织中特异性表达。

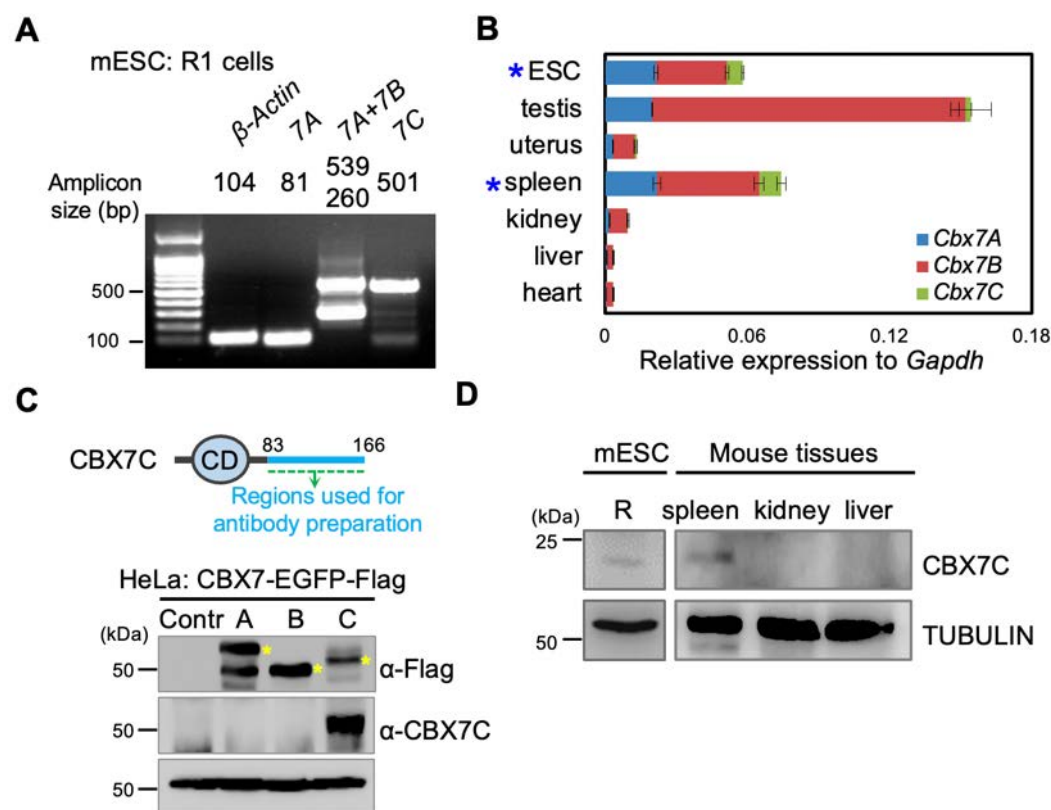


图 3-1 *Cbx7* 转录本及其编码的蛋白在小鼠 ESCs 及组织中的表达水平检测

Figure 3-1 The expression detection of *Cbx7* and encoded proteins in mESCs and tissues

A: PCR 检测 mESC(R1)中 *Cbx7* 各转录本; B: RT-qPCR 检测 mESC 及小鼠不同组织中 *Cbx7* 各剪接体的转录水平; C: CBX7C 抗体的特异性验证。CBX7C 多克隆抗体的特异性识别区域示意图(上)。在瞬时表达 CBX7-EGFP-Flag 的 HeLa 细胞中, 通过 WB 验证 CBX7C 抗体的特异性(下)。“*”指示目的蛋白条带; D: WB 检测 CBX7 各蛋白亚型在 mESC(R1)及所示的小鼠组织中的蛋白表达。CBX7C 蛋白的理论分子量约为 20 kDa。

小结

在本论文研究中, 我们关注了一种未报道的 *Cbx7* 剪接异构体 *Cbx7C* 的功能特性。该异构体编码的蛋白 CBX7C 蛋白亚型缺乏 C 端 PC 结构域, 在 mESC 和小鼠脾脏组织中特异性表达。

3.2 丧失了与 RING1B 结合活性的 CBX7C 仍参与完整的 cPRC1 复合体组装

3.2.1 缺失 PC 结构域的 CBX7C 丧失了与 RING1B 的直接相互作用

研究表明 cPRC1 蛋白复合体靶向目标基因的过程主要依赖于其核心组分之一 CBX 家族蛋白识别并结合被特异性修饰的组蛋白标签,其中 RING1B 与 CBX 蛋白的相互作用是调节 cPRC1 蛋白复合体成员与核小体之间结合的接头机制。作为 cPRC1 复合体的重要组成部分之一, CBX7 通过表观遗传调控目的基因的表达,并且在细胞分化时与其它 CBX 蛋白交替表达^[136]。CBX7A 和 CBX7B 蛋白 C 端保守的 PC 结构域能够与 RING1B 蛋白 C 端 RAWUL 结构域直接相互作用,进而促进 PRC1 复合体组装。基于已知的蛋白结构域互作信息, PC 结构域的缺失可能会阻断 CBX7C 与 RING1B 的直接相互作用,并进一步改变 CBX7C 与其它 cPRC1 复合体组分的互作模式。

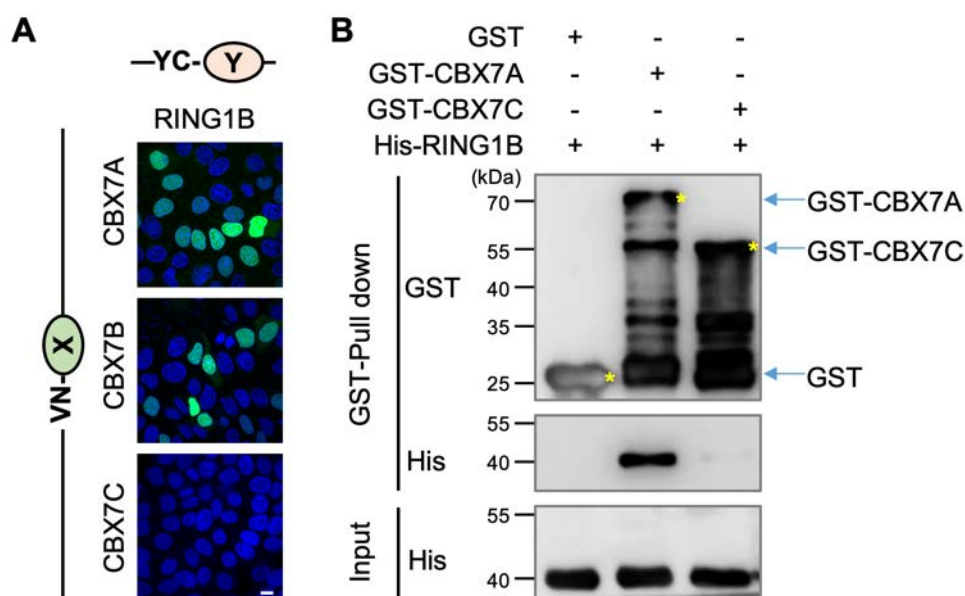


图 3-2 CBX7 各蛋白亚型与 RING1B 的相互作用检测

Figure 3-2 The interaction detection between CBX7 protein subtypes and RING1B

A: 通过 BiFC 及共聚焦成像分析检测 VN-CBX7A、B 和 C 与 YC-RING1B 在 HEK293T 细胞中的互作。Hoechst 信号指示细胞核。比例尺为 10 μm ; B: GST-pull down 及 WB 验证原核表达蛋白 GST-CBX7A 或 C 与 His-RING1B 的互作。“*”指示目的蛋白条带。

为了验证这一假设,我们首先在 HEK293T 细胞中通过 BiFC 实验及共聚焦成像检测了 CBX7 的不同蛋白亚型与 RING1B 的互作。结果证实 CBX7A 和 CBX7B 能够与 RING1B 蛋白互作,而 CBX7C 确实失去了与 RING1B 互作的能力(图 3-2A)。此外,我们进一步通过 GST-pull down 实验检测了 CBX7 与 RING1B 是否存在直接的相互作用。基于由 251 个氨基酸残基构成的 CBX7A 包含了 CBX7B 中的所有氨基酸序列,并且 CBX7A 与 CBX7B 具有相似的表观遗传学调控特性,因此在本论文中,我们主要选择 CBX7A 与 CBX7C 进行了比较分析。

GST-pull down 结果证实, 与 BiFC 结果相一致, CBX7A 能够直接结合 RING1B, 而缺失 PC 结构域的 CBX7C 蛋白亚型丧失了与 RING1B 的直接相互作用。

3.2.2 CBX7C 参与 cPRC1 复合体的组装

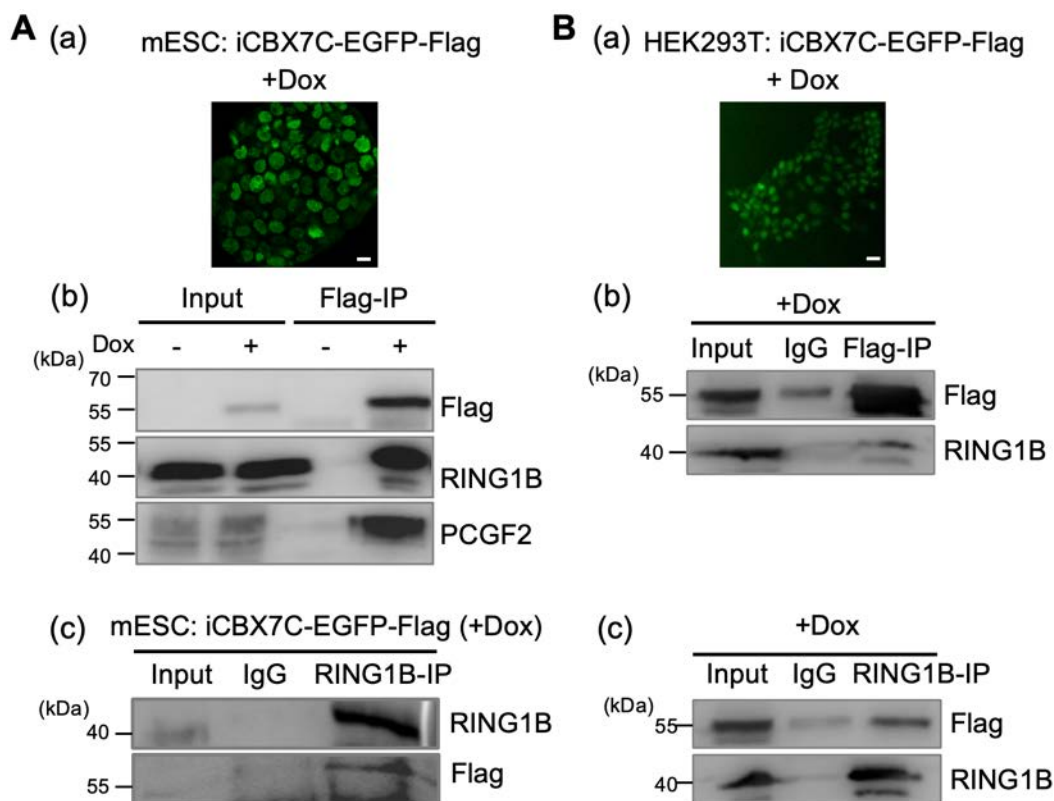


图 3-3 CBX7C 与 PRC1 组分的体内互作分析

Figure 3-3 The interaction analysis between CBX7C and PRC1 components *in vivo*

A-B(a): Dox 诱导 mESC 或 HEK293T 稳转细胞系表达 CBX7C-EGFP-Flag 后进行荧光成像检测。比例尺为分别 10 μ m (A(a))和 200 μ m (B(a)); A-B(b-c): CBX7C 与 RING1B 或 PCGF2 之间的 co-IP 体内互作分析。

CBX7C 中 PC 结构域的缺失导致其丧失了与 RING1B 的直接相互作用, 那么 CBX7C 是否还能参与 cPRC1 复合体的组装过程? 为了验证这一假设, 我们在诱导表达 CBX7C 的 mESC 系中通过 co-IP 实验检测了 CBX7C 与 cPRC1 核心酶学组分的生理性互作关系。与上述 BiFC 和 GST-pull down 测定的数据不同, co-IP 结果显示, 在稳定表达 CBX7C-EGFP-Flag 融合蛋白的 mESC 或 HEK293T 细胞 (天然性缺失内源 CBX7C) 中, CBX7C 均能够与 RING1B 和 PCGF2 结合 (图 3-3A-B)。这表明 CBX7C 与 cPRC1 核心酶学组分 RING1B 和 PCGF2 在生理条件下存在于同一复合体中, 并且这一组装过程不依赖于 CBX7 与 RING1B 的蛋白直接互作。这也表明 CBX7C 可能通过与其它组分的相互作用而参与 CBX7C-cPRC1 复合物的组装。在关于 CBX7C 和 RING1B 的互作研究结果中,

co-IP 与 BiFC 实验结果并不一致，这可能是由于 cPRC1 复合体中 CBX7C 和 RING1B 之间的空间距离大于 BiFC 实验系统的临界值导致的。

3.2.3 CBX7C 与 PHC2 的直接蛋白互作介导了 CBX7C-cPRC1 复合体的组装

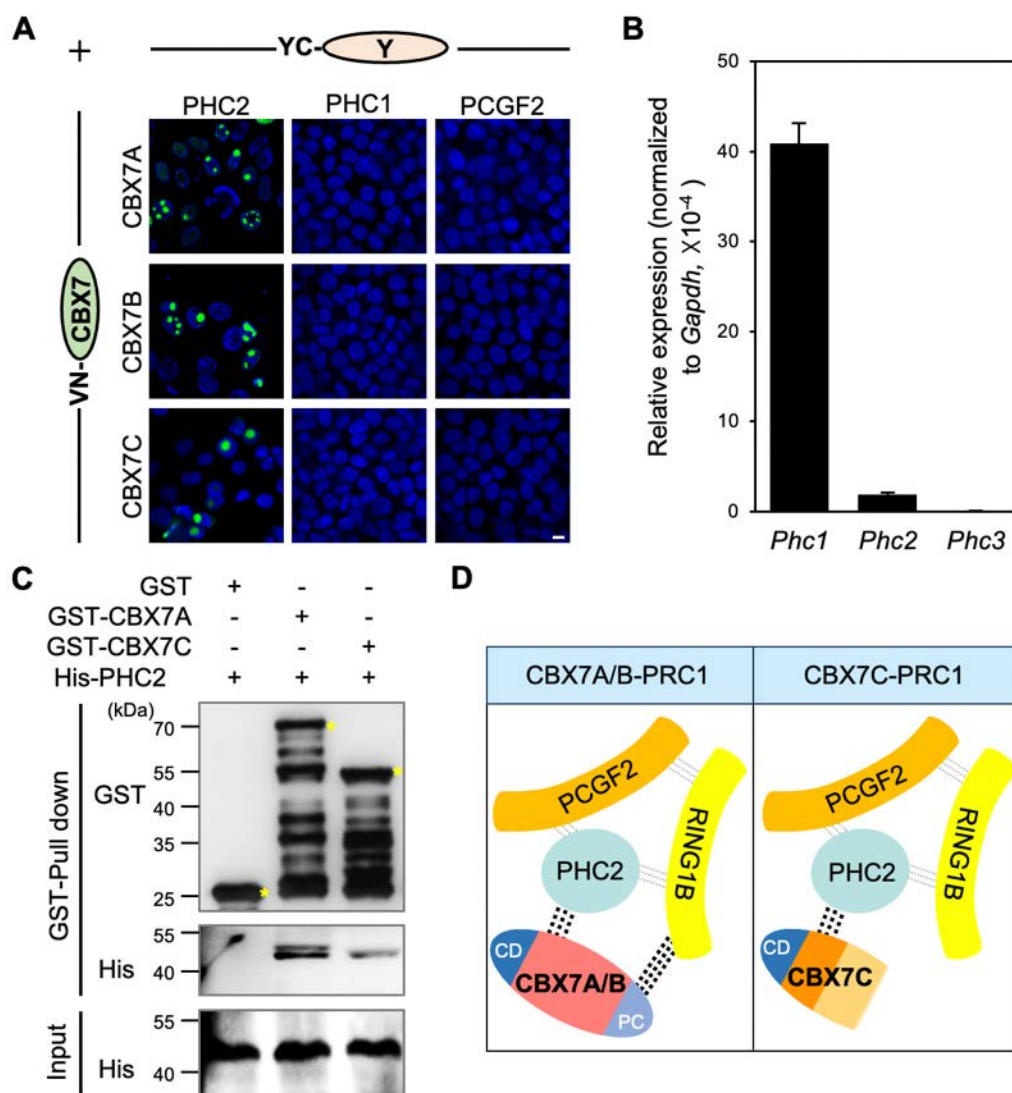


图 3-4 CBX7 各蛋白亚型与 PHC2 和 PCGF2 的互作检测

Figure 3-4 The interaction detection between CBX7 protein subtypes and PHC2 or PCGF2

A: BiFC 实验检测 VN-CBX7A、B 和 C 与 YC-PHC1、PHC2 和 PCGF2 在 HEK293T 细胞中的互作。Hoechst 信号指示细胞核。比例尺为 10 μ m；B: RT-qPCR 检测 mESC 中及 *Phc1-3* 的转录水平。C: GST-pull down 验证纯化蛋白 GST-CBX7A 和 C 与 His-PHC2 的互作。“*”指示目的蛋白条带。D: CBX7A/B 和 CBX7C 与 PRC1 其它组分互作关系示意图。

上文所述结果显示 CBX7C 能够独立于 RING1B 的介导而参与 cPRC1 复合体的组装过程。基于此，我们通过 BiFC 及 GST-pull down 实验进一步探究了介导复合体组装过程的 CBX7C 互作因子。BiFC 实验结果表明：在 HEK293T 细胞

中, 三种 CBX7 蛋白均未与 PCGF2 产生阳性 BiFC 互作信号, 意味着 cPRC1 复合物中的 CBX7 和 PCGF2 之间的物理距离可能相对较远。但是 CBX7 能够结合 mESC 中低水平表达的 PHC2 (图 3-4A 和 B), 这表明 PRC1 复合物中 PHC2 的构象可能更有利于建立与 CBX7 的相互作用。GST-pull down 结果也进一步证实, GST-CBX7A 和 GST-CBX7C 能够直接结合 His-PHC2 蛋白 (图 3-4A), 这表明 PHC2 可能作为关键性桥梁蛋白与 CBX7C 互作进而介导其参与 cPRC1 复合物的组装。结合上文所述结果综合分析, CBX7A 与 RING1B 和 PHC2 的互作共同介导了 CBX7A-cPRC1 复合体的组装。然而在 CBX7C-cPRC1 复合体的组装过程中, PHC2 替代了 RING1B 而作为 CBX7C 参与完整 cPRC1 复合体组装的桥接分子发挥功能 (图 3-4D)。

小结

CBX7C 缺失 PC 域将导致与 RING1B 的直接相互作用丧失, 并改变其参与完整的 cPRC1 复合体组装的蛋白互作模式, PHC2 可替代 RING1B 作为连接 CBX7C 与复合体之间的桥接分子, 以促进完整的 cPRC1 复合体在细胞内的组装。在 mESC 中, CBX7C 作为功能性 cPRC1 复合体的核心组分之一来驱动表观遗传学抑制功能的建立, 而不是作为独立调节因子发挥作用。

3.3 CBX7C-cPRC1 复合体识别靶基因并参与基因表达的表现遗传抑制调控

3.3.1 CBX7C 与染色质组蛋白 H3 相互作用

上文所述结果表明, 尽管缺失了与 RING1B 结合的 PC 结构域, 但 CBX7C 仍然能够参与完整的 cPRC1 复合体的组装。保留了 CD 结构域的 CBX7C 是否也具备了作为表观遗传调控因子的基本特征, 如同 CBX7A 和 CBX7B 一样能够结合染色质? 基于此, 我们通过 BiFC 实验检测了 CBX7A 和 CBX7C 与染色体组蛋白 H3 的互作关系, 并利用其截短体进一步确认其介导染色质结合的关键结构域。PRC1 蛋白复合体靶向到染色质主要依赖于 CBX 家族蛋白的 CD 结构域识别并结合被特异性修饰的组蛋白标签^[28, 29]。邻近 CD 结构域的 ATHL 结构域能够结合由富含 AT 的 DNA 序列形成的凹陷结构^[114]。CBX 蛋白的 CD 结构域和 ATH 结构域都能够结合包含有 H3K27me₃ 修饰标签的核小体和 AT 富集的卫星序列, 在 CBX 蛋白靶向至目的基因时需 CD 和 ATHL 结构域协同发挥作用^[117] (图 3-5A)。

与 CBX7A 和 CBX7B 一样的, CBX7C 同样具有与 H3 互作的能力 (图 3-5B), 这与预期结果相一致。与此同时, 对于互作结构域的研究显示, CD 或 ATHL 结构

域的单一缺失并未影响CBX7A和CBX7C与H3的结合，但是当同时缺失CD和ATHL结构域后，完全阻断了CBX7A和CBX7C与H3的结合(图3-5C)。表明CBX7C具有与CBX7A同样的染色质组蛋白结合能力，这也奠定了CBX7C可能具有基因表达调控功能的分子基础，并且位于CBX7的N端的CD和ATHL结构域共同介导了其与组蛋白的互作。

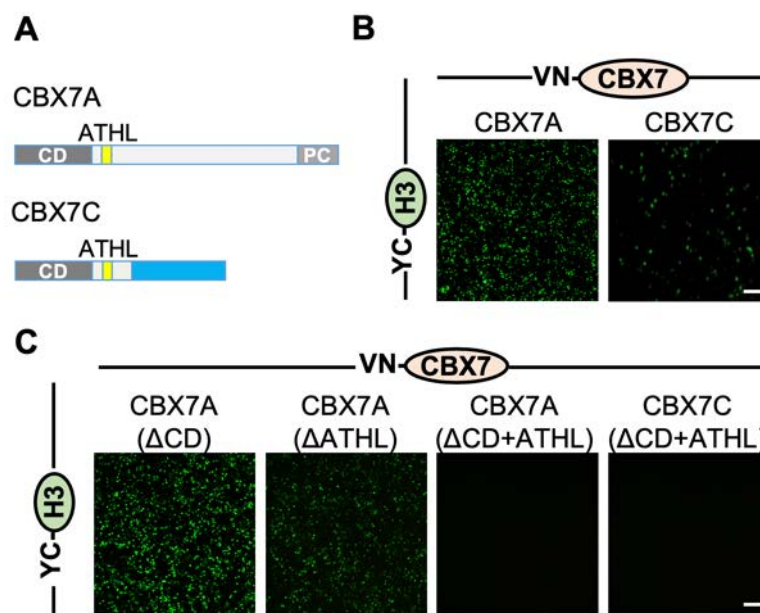


图 3-5 CBX7C 结合染色质组蛋白能力及互作结构域的分析

Figure 3-5 Analysis of CBX7C binding ability to chromatin histones and interaction domains

A: CBX7A 和 CBX7C 蛋白保守结构域示意图; B-C: 通过 BiFC 及荧光成像观察分析由 CBX7A 或 CBX7C 及其所示截断体与 H3 的互作。比例尺为 200 μm 。

3.3.2 CBX7C 识别并结合经典的 cPRC1 靶基因

为了进一步探究 CBX7C 是否也能够作为 cPRC1 的亚基被招募并结合到已知的 cPRC1 靶基因上,我们以 Dox 诱导并稳定表达 CBX7C-EGFP-Flag 的 mESC 为细胞材料,通过 ChIP-qPCR 实验检测分析了 CBX7C 靶向至目的基因的能力。结果显示, CBX7C-EGFP-Flag 能够在经典的 cPRC1 靶基因处富集,并且 CBX7C 与这些基因位点的结合普遍减少了此处 CBX7A 或 CBX7B 的结合,这表明 CBX7C 与其它 CBX7 蛋白亚型竞争性地结合靶基因(图 3-6)。相反,在大多数图示的靶基因位点上,随着 CBX7C 的结合,PCGF2 和 RING1B 的富集总体呈上升趋势,同时与其对应基因位点的酶学催化产物 H2AK119ub 修饰水平也相应增加(图 3-6)。这一结果也为 CBX7C 在靶基因上能够正常参与完整且具有酶活性的 PRC1 复合物的组装过程的结论提供了有效的证据。另外,在图示基因位点上,由 EZH2-PRC2 介导催化的 H3K27me3 的结合未发生显著变化,这说明 CBX7C

在基因位点的结合不影响此处 PRC2 的富集及催化活性。

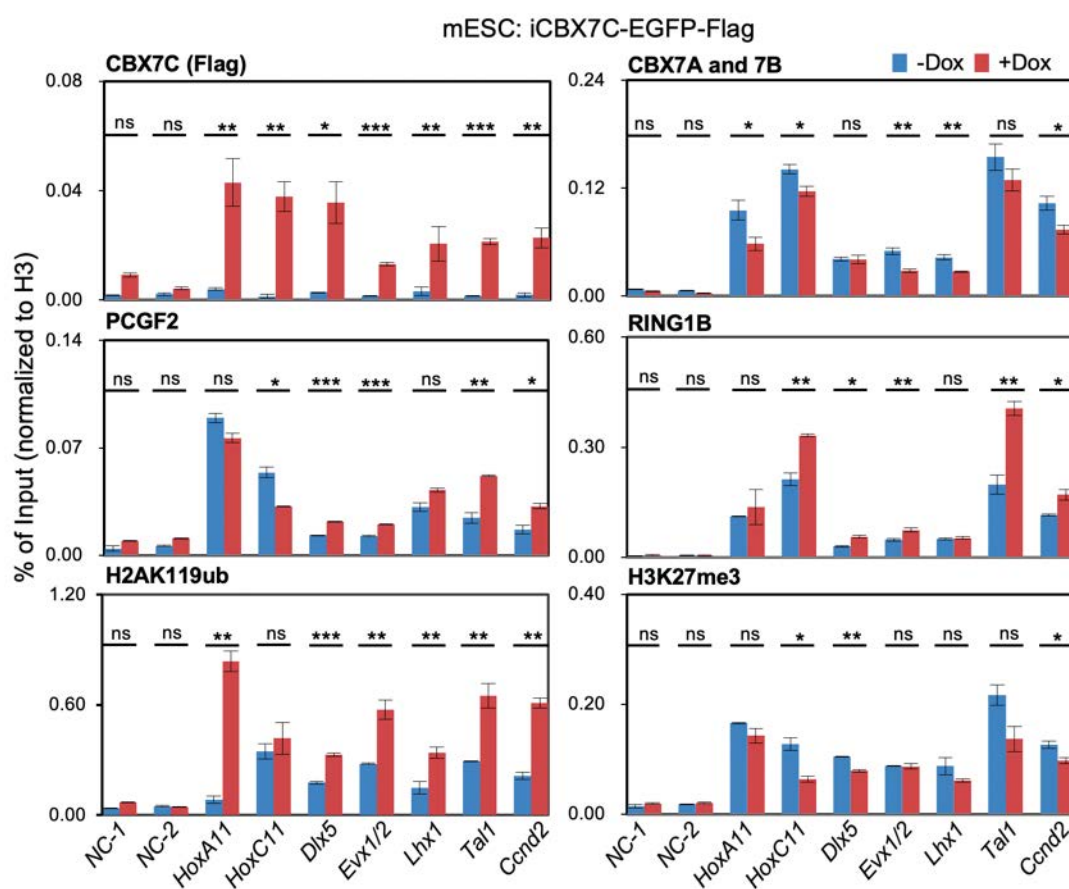


图 3-6 CBX7C 识别并结合 cPRC1 靶基因的能力的分析

Figure 3-6 Activity determination of CBX7C to recognize and bind cPRC1 target genes

在可诱导表达 CBX7C-EGFP-Flag 的 mESC 稳转细胞系中 (Dox 诱导), ChIP-qPCR 分析 CBX7C 在图示的 PRC1 靶基因上的富集及对 PRC1 其它组分和 H2AK119ub 修饰的影响。NC-1 (*Myc*) 和 NC-2 (*Klf4*) 是阴性对照组基因。通过 *t* 检验统计平均值之间的显著性差异 (ns, 不显著, $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)。

3.3.3 CBX7C 参与靶基因的转录抑制调控

在哺乳动物细胞中, 含 CD 结构域的 CBX 蛋白能够作为转录抑制因子发挥作用。为了进一步证实 CBX7C 是否也具有与 CBX7A/B 相似的转录抑制功能, 我们首先检测了在 mESC 中诱导表达 CBX7C-EGFP-Flag 时上文所述的 cPRC1 靶基因的转录水平。RT-qPCR 结果显示, 这些 PRC1 靶基因在 CBX7C-EGFP-Flag 诱导表达后的转录水平显著下降 (图 3-7)。虽然 *HoxC11* 和 *Dlx5* 则略微上调, 但无显著差异, 这可能是由于这两个发育相关基因在 mESC 中的本底转录水平较低引起的。以上数据表明 CBX7C 不仅参与了在靶基因处的 cPRC1 复合体的

组装，而且 CBX7C 也作为有效亚基参与了 cPRC1 复合体介导的转录抑制调控。

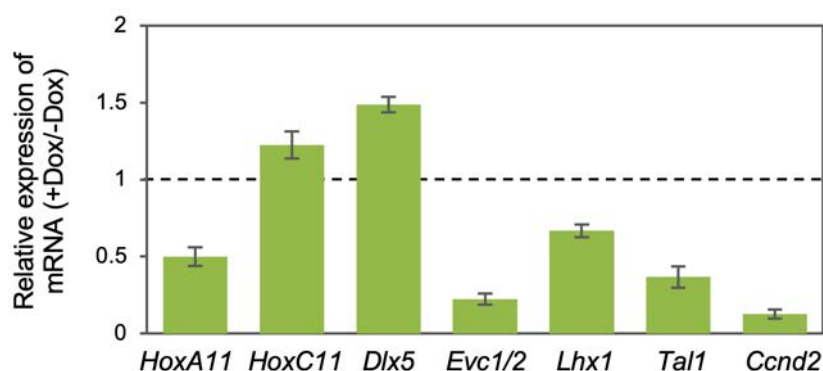


图 3-7 CBX7C 对结合的靶基因转录抑制调控能力的分析

Figure 3-7 Transcriptional repression regulation activity determination of CBX7C

transcriptional repression regulation activity in binding target genes

RT-qPCR 分析 CBX7C-PRC1 结合后相应靶基因的转录水平变化。

其次，我们检测并比较了 CBX7A 和 CBX7C 在体内靶向报告基因时介导转录沉默的能力，以探究 CBX7C 是否也具有直接抑制靶基因的能力（图 3-8A）。我们在 HK293T 细胞中瞬时转染等量的表达 GAL4 DBD (DNA binding domain) - CBX7A 或者 CBX7C 融合蛋白的载体。待转染 36 小时后细胞内积累一定量的 GAL4 DBD-CBX7 后，再次转染等量的表达 EGFP-Flag 的载体。该载体中 EGFP-Flag 的 DNA 序列前包含了 2000 bp 的 *Cyp26a1* 启动子（cPRC1 结合位点）以及相邻的 GAL4 DBD 的 UAS 结合位点。CBX7 被 GAL4 DBD 锚定到了 *Cyp26a1* 启动子处，进而通过检测 EGFP-Flag 的表达量来评估目的蛋白 CBX7 的转录调控能力。与对照组相比，GAL4 DBD-CBX7A 的表达导致 EGFP-Flag 蛋白水平显著下调；而表达了 GAL4 DBD-CBX7C 后，EGFP-Flag 蛋白的表达量与对照组相比也有显著减少，但是比 CBX7A 显示出较弱的沉默效应（图 3-8B）。

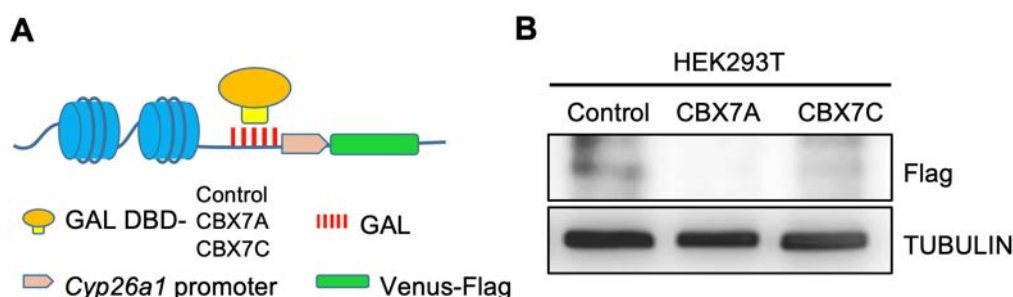


图 3-8 CBX7C 对含 *Cyp26a1* 启动子的报告基因转录抑制调控能力的测定

Figure 3-8 Transcriptional repression regulation activity determination of CBX7C

transcriptional repression regulation activity in report gene promoted by *Cyp26a1*

A: GAL 介导蛋白靶向报告基因系统原理示意图；B: 通过 GAL 介导蛋白靶向报告基因系统分析 CBX7A 和 CBX7C 对报告基因的转录调控功能。

此外,为了证实 CBX7 不同剪切体对靶基因的抑制效应的普遍性,我们通过双荧光素酶报告系统检测了 GAL4 DBD-CBX7A、C 及 PHC2 蛋白对细胞内编码萤火虫荧光素酶 (Firefly Luciferase, LUC) 的报告基因的表达调控,该 LUC 的表达受 *TK* 启动子控制 (图 3-9A)。同时,在此系统中我们使用 IRES 序列在同一表达载体中将编码 GAL4 DBD-CBX7 与海肾荧光素酶 (Renilla Luciferase, REN) 的序列隔开使二者独立表达,进而可将 REN 作为内参来标准化不同实验组间的转染效率。如图 3-9B 所示,我们发现 GAL4 DBD-CBX7 及 PHC2 的表达导致下游报告基因被抑制,其编码的 LUC 表达下调,导致下游的酶活反应受到显著抑制。以上结果表明,在相同的转录测定中,CBX7C 具有与 CBX7A 蛋白同样的基因沉默活性。

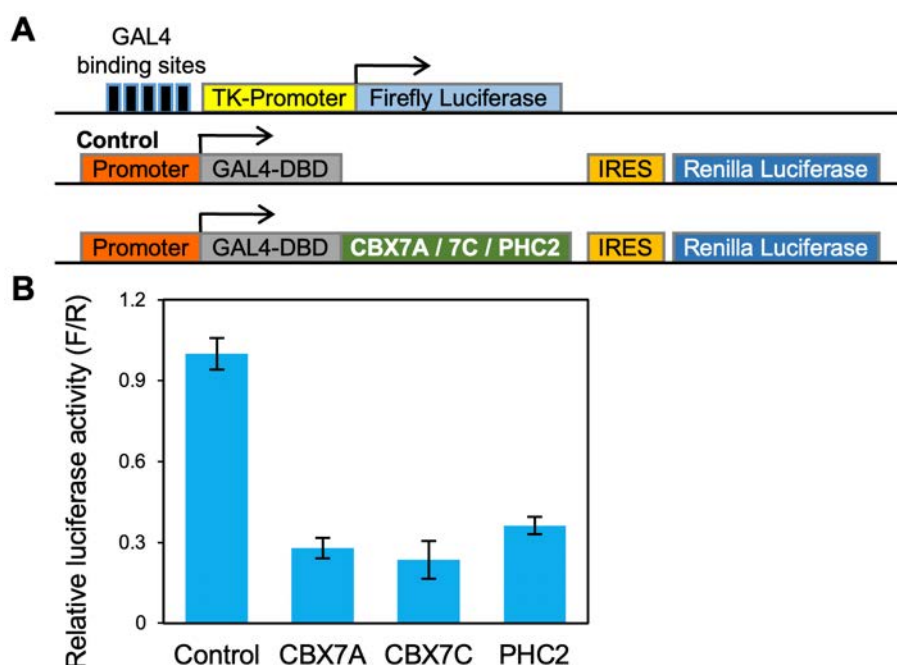


图 3-9 CBX7C 对含 *TK* 启动子靶基因转录抑制调控能力的测定

Figure 3-9 Transcriptional repression regulation activity determination of CBX7C

transcriptional repression regulation activity in target genes promoted by *TK*

A: 双荧光素酶报告系统原理示意图; B: 双荧光素酶报告系统分析 CBX7A、CBX7C 及 PHC2 蛋白对报告基因的转录调控功能。

小结

完整的 CBX7C-cPRC1 复合体在染色质上与 CBX7A 和 7B 竞争性地结合靶基因,并维持其催化组蛋白 H2AK119 泛素化的酶活性,进一步发挥对靶基因的表现遗传抑制调控。

3.4 CBX7C 驱动功能性 cPRC1 复合体在染色质上的组装起始

3.4.1 HTM 介导的蛋白异位靶向系统建立

尽管 CBX7C 参与完整的 cPRC1 复合体的组装,但是该过程精细调控的分子机制尚不清楚。在复合体组装过程中, CBX7 是否作为“驾驶员”主动招募作为“乘客”的其它组分? 我们通过由 HTM (heterochromatic targeting module) 介导的蛋白异位靶向实验体系来探究 PRC1 组装过程中的潜在驱动因素。将带有荧光标签的目的蛋白靶向至臂间异染色质 (pericentric-heterochromatin), 并通过免疫荧光或融合不同荧光标签而分析其它蛋白组分是否也能够在此位点上与其发生共定位^[57] (图 3-10)。PRC1 和 PRC2 复合物通常在兼性异染色质区域富集, 并负责产生 H2AK119ub 和 H3K27me3 修饰, 以驱动和维持染色质的缩合。相反, 臂间异染色质区域属于组成型异染色质区域, 其主要包含了组蛋白修饰 H3K9me3 和 H4K20me3^[220]。HTM 是 H4K20 的特异性甲基转移酶 SUV4-20H2 蛋白中一段由 88 个氨基酸残基构成的区域, 主要介导了 SUV4-20H2 蛋白与臂间异染色质的 H4K20me3 稳定结合^[221]。

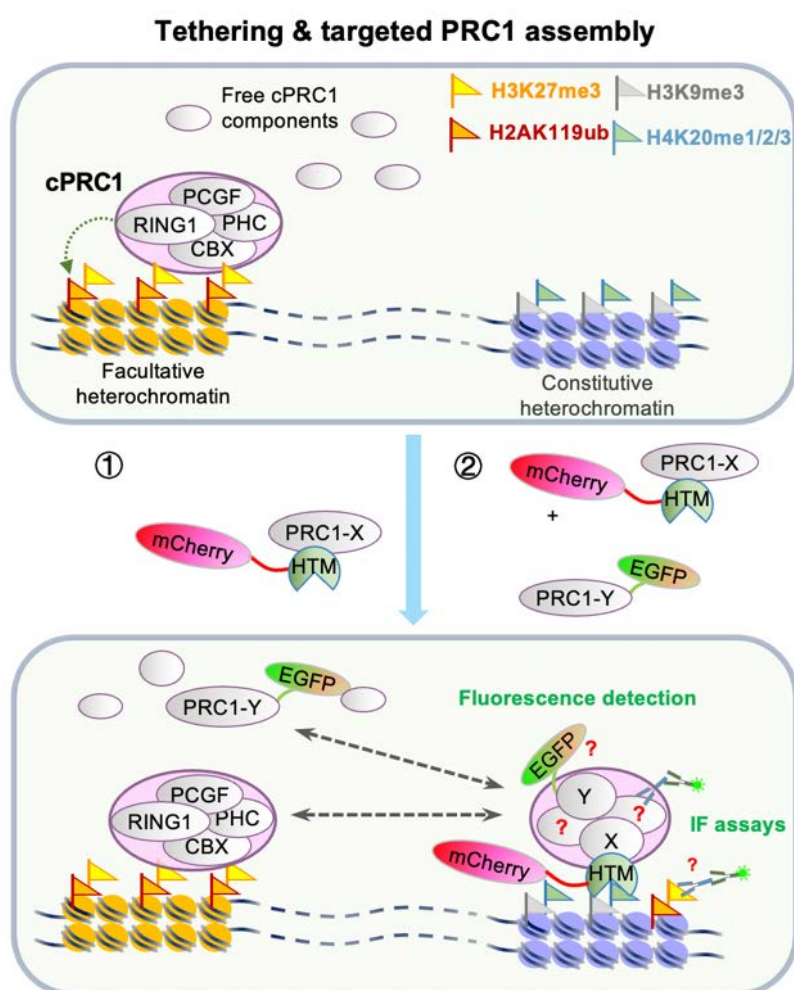


图 3-10 HTM 介导的蛋白异位靶向系统

Figure 3-10 HTM-based artificial tethering and PRC1 complex assembly assay

在 mESC 中单独表达的 mCherry-HTM 融合蛋白与 H3K9me3 标记的组成型异染色质完全共定位，但并未影响内源性 RING1B 或 H2AK119ub 的正常核分布（图 3-11）。该结果验证了该染色质异位靶向系统的特异性和可行性。因此，通过该结合将 CBX7A、CBX7C 或 PHC2 重新靶向至其非天然染色质结合位点，并检测 cPRC1 复合物在能否在新位点完成组装，进而有助于在复合体动态组装过程中区分驱动因素和从动因素。

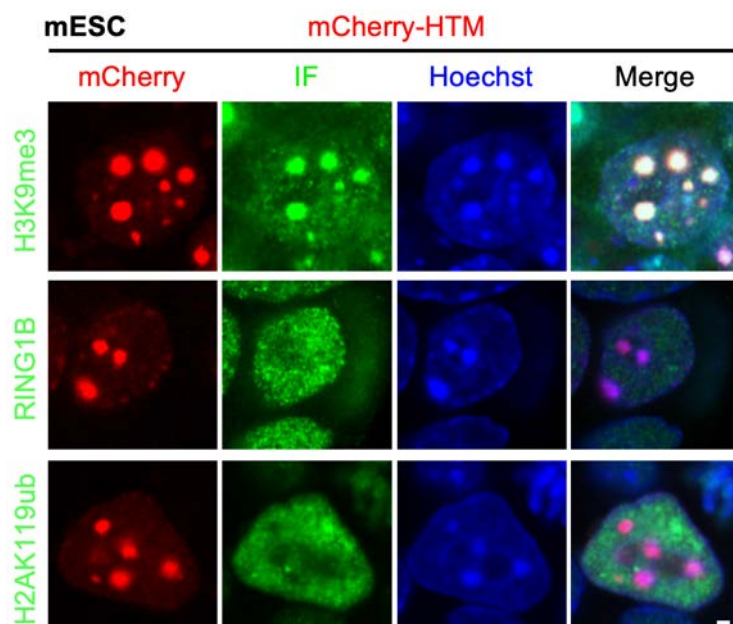


图 3-11 HTM 介导的蛋白异位靶向系统验证

Figure 3-11 Validation of HTM-based artificial tethering and PRC1 complex assembly assay
通过共聚焦成像分析 mCherry-HTM 形成的颗粒在 mESC 核内分布，以及与所示组蛋白修饰或 RING1B 的分布关系。Hoechst 信号指示细胞核。比例尺为 2 μm 。

3.4.2 CBX7C 和 CBX7A 均能作为功能性 cPRC1 复合体在异位组装的“驱动因素”

在 mESC 中，我们利用上述 HTM 介导的染色质异位靶向系统并结合共聚焦成像和 IF 实验分析探究 PRC1 复合体在组成型异染色质上的组装过程。首先，在表达 mCherry-HTM-CBX7A 的 mESC 中，我们观察到内源性 RING1B 也被招募至由 H3K9me3 标记的组成型异染色质处，同时 RING1B 催化产生的 H2AK119ub 在此处也显著聚集（图 3-12A 和 D），表明 CBX7A 可能作为具有酶活性的 PRC1 复合体在异位完成组装的“驱动因素”。同样的，mCherry-HTM-CBX7C 也能够作为“驱动因素”在异位募集内源性 RING1B 并产生 H2AK119ub 修饰（图 3-12B 和 D）。有研究报道称具有 E3 泛素连接酶活性的 RING1B-PCGF4 是催化

H2AK119ub 的 cPRC1 核心酶学组分^[18-20]，这也进一步印证了丧失了与 RING1B 结合活性的 CBX7C 仍然能够如同 CBX7A 一样参与完整的功能性 cPRC1 复合体在靶基因上的组装。此外，CBX7A、CBX7C 及 PHC2 在异位的组成型染色质区域的结合并未引起 H3K27me3 在此处的富集。这也意味着在此实验系统检测中，cPRC1 的组分 CBX7A、CBX7C 及 PHC2 的异位靶向不能驱动 PRC2 也发生异位靶向并在此发生酶促反应。

PHC2 能够直接结合 RING1B 及 CBX7，并且 PHC2 与 CBX7C 的互作作为分子桥梁介导了 CBX7C-cPRC1 在染色质上的构建。基于此，我们利用染色质异位靶向系统分析 PHC2 在 cPRC1 组装过程中的潜在作用。结果表明，PHC2 也可以作为起始 cPRC1 组装的“驱动器”，但效率低于 CBX7 蛋白（图 3-12C 和 D）。

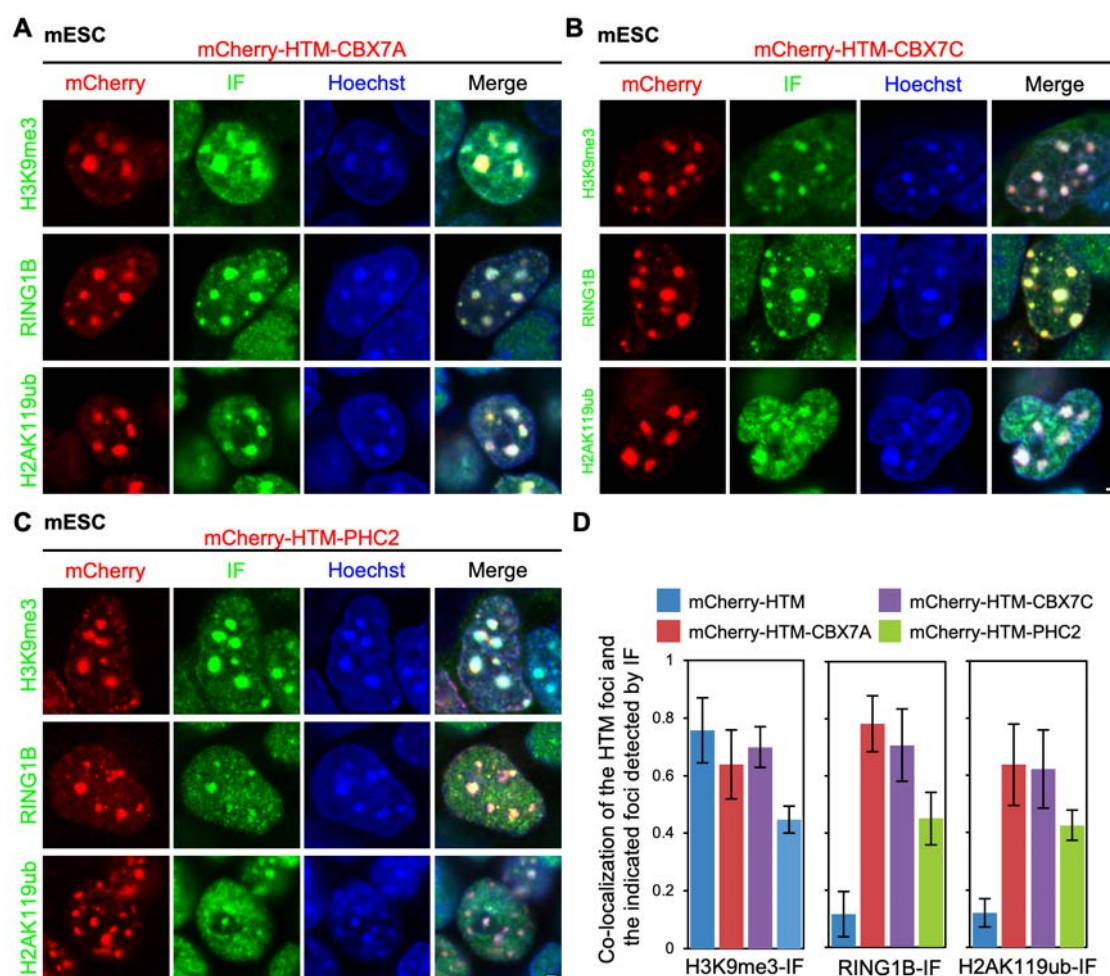


图 3-12 在 mESC 中 cPRC1 复合体组装过程中的潜在驱动因素的探究

Figure 3-12 Study of potential driving factors in cPRC1 complex assembly in mESCs

A-C: 通过共聚焦成像分析 mCherry-HTM-CBX7A、CBX7C 及 PHC2 形成的颗粒在 mESC 核内的分布，以及与所示组蛋白修饰或 RING1B 的分布关系。Hoechst 信号指示细胞核。比例尺为 2 μ m；D: 柱状图表示通过 ImageJ 分析 A-C 所示共定位颗粒与 mCherry-HTM 蛋白形成的颗粒数目的比例。

3.4.3 PHC2 在异位驱动复合体组装中优先招募 CBX7C

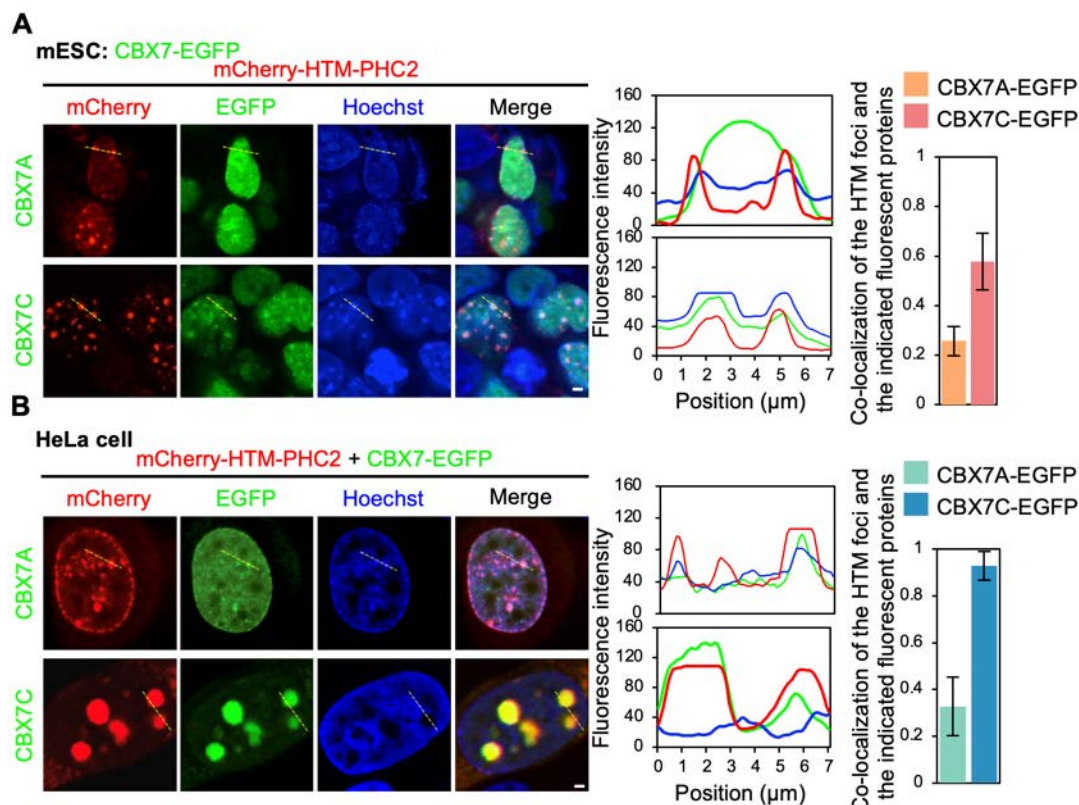


图 3-13 PHC2 对 CBX7A 和 CBX7C 招募能力的比较分析

Figure 3-13 Comparative analysis of PHC2's recruitment ability for CBX7A and CBX7C

A: 在稳定表达 CBX7A-EGFP 或 CBX7C-EGFP 的 mESC 中, 通过共聚焦成像观察分析 mCherry-HTM-PHC2 对 CBX7 亚型的募集(左); B: 在瞬时过表达 CBX7-EGFP 及 mCherry-HTM-PHC2 的 HeLa 细胞中, 通过共聚焦成像观察分析进一步验证 mCherry-HTM-PHC2 对 CBX7 亚型的募集(左)。曲线图表示图示虚线区域上的荧光强度(A 和 B 中)。柱状图表示通过 ImageJ 分析图示共定位颗粒与 mCherry-HTM-PHC2 形成颗粒数目的比例(A 和 B 右)。Hoechst 信号指示细胞核。比例尺分别为 4 μm (A) 和 2 μm (B)。

在稳定表达 CBX7A-EGFP 或 CBX7C-EGFP 的 mESC 中, 与 CBX7A 相比, HTM-PHC2 优先招募 CBX7C (图 3-13A)。结合上文所述结果, 建立了 CBX7C 和 PHC2 之间的相互招募模型, 并且这种协同的相互作用促进了完整 PRC1 在靶位点的进一步组装。为了进一步验证上述 PHC2 对 CBX7A 和 CBX7C 招募能力具有差异的结论, 我们在 HeLa 细胞中进行了检测以排除内源性 CBX7C 蛋白干扰的可能性(图 3-13B)。与在 mESC 中获得的结果一致, mCherry-HTM-PHC2 只能募集少量的 EGFP-CBX7A, 这说明 PRC1 在染色质上的组装过程中, CBX7A 和 PHC2 之间具有“驱动-从动”的关系。然而, 在 CBX7C 和 PHC2 共同表达时, 二者在核内完全共定位并形成了体积较大的凝集体, 这些巨大的核凝集体与之前

观察到的在臂间异染色质靶向的颗粒物截然不同，这意味着 HTM 的正常染色质锚定性质受到了影响。同时，这些凝集体所处的位置无法被 Hoechst 标记，即此处不含染色质成分，这可能是由于 CBX7C 和 PHC2 在 HeLa 细胞中瞬时表达量过高所致。综上，与 CBX7A 相比，尽管 CBX7C 失去了 RING1B 的结合能力，但其仍能驱动功能性 cPRC1 复合物的组装。

小结

通过独特的 HTM 介导的蛋白异位靶向系统对 cPRC1 复合体在其非靶向位点组装的监测，证明与 CBX7A 类似，CBX7C 也能够作为“驱动因子”而驱动具有有效催化酶活性的完整 cPRC1 复合物在染色质上的起始组装。此外，PHC2 也可以作为触发 cPRC1 组装的“驱动器”，并优先招募 CBX7C。CBX7C 和 PHC2 之间的相互及协调招募模型促进了完整的 CBX7C-cPRC1 在靶位点的进一步组装。

3.5 CBX7 蛋白亚型具有不同的相分离能力

3.5.1 体外纯化蛋白 CBX7C 具有比 CBX7A 更强的相分离能力

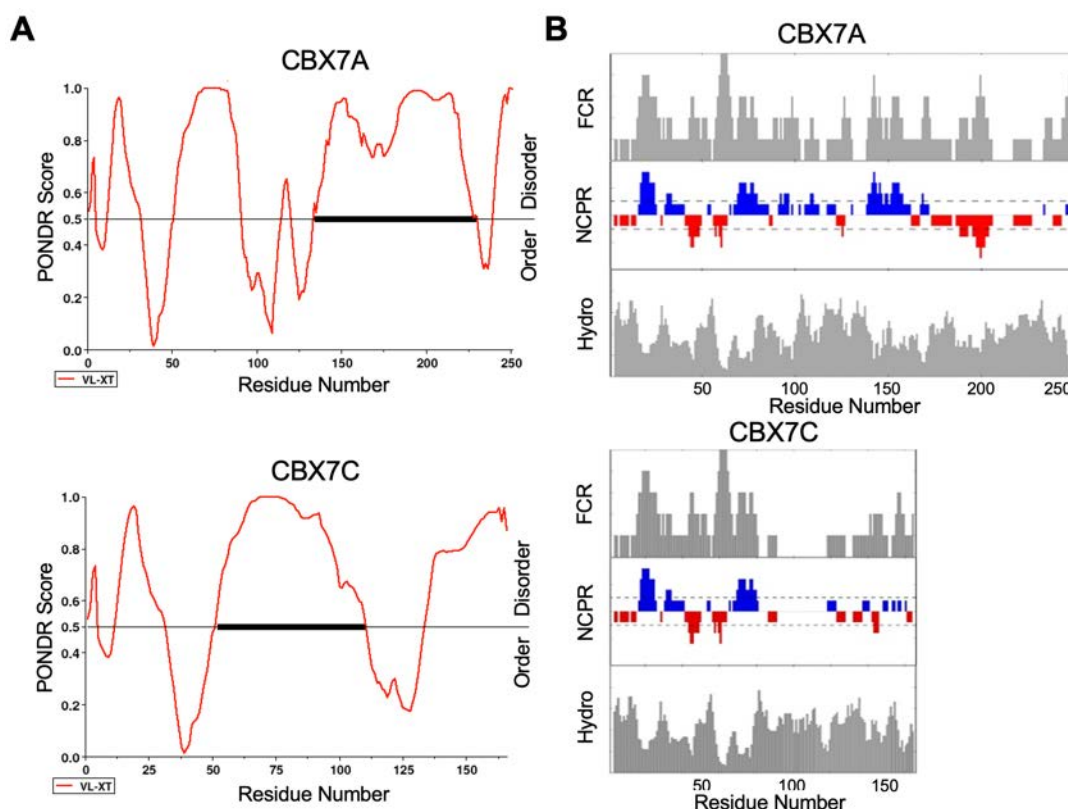


图 3-14 CBX7A 和 CBX7C 蛋白序列的 IDRs 及 LCD 预测

Figure 3-14 Prediction of IDR and LCD of CBX7A and CBX7C proteins

近年来越来越多的研究发现,在正常的细胞生理环境中,含有固有无序区域(IDRs)和/或低复杂区域(LCR)的蛋白质容易发生相分离而形成液滴状的凝集体。根据 IDR 预测,发现 CBX7A 含有六个不连续 IDR,约占整个蛋白氨基酸序列的 70.92%;类似地,CBX7C 含有四个 IDR,占整个蛋白氨基酸序列的 69.88%(图 3-14A)。此外,对 LCR 的分析还表明,CBX7A 和 CBX7C 都包含多个 LCR(图 3-14B),上述信息学分析结果暗示着 CBX7A 或 CBX7C 蛋白可能具有驱动相分离发生的能力。

为了探究和比较 CBX7A 和 CBX7C 在体外发生相分离的能力及其形成的凝集体的理化性质,我们利用大肠杆菌中原核表达并纯化了带有 GST 和 EGFP 双重标签的 CBX7A 或 CBX7C 融合蛋白(图 3-15A(a))。通过 BCA 测定两种蛋白质的浓度,并将其调节至相同浓度水平(12 μ M)。在此条件下,通常具有相分离潜力的蛋白质在溶液中会出现一定程度的浑浊^[215, 218, 222]。基于此,我们直接检测并比较了纯化蛋白样品在离心后含有 CBX7A 和 CBX7C 蛋白溶液是否产生凝集反应。如图 3-15A(b)所示,作为对照组的 GST-EGFP 蛋白本身在溶液中保持分散,而 GST-EGFP-CBX7A 或 CBX7C 均能够凝集为富含蛋白质的颗粒并集中在离心管的底部。这初步表明 CBX7A 和 CBX7C 能够形成与本体溶液分离的致密相。在此过程中,对 CBX7A 而言,只有少部分蛋白质参与了颗粒的形成,仍然相当多的蛋白质溶解在周围溶液中。与之相比,几乎所有的 CBX7C 蛋白都凝集到试管底部的颗粒中。1,6-己二醇(1,6-hexanediol)是一种能够通过破坏弱的疏水相互作用使得相分离液滴溶解的化学小分子,被广泛应用于鉴定蛋白相分离特性,其具有适度的疏水相互作用干扰能力,能够在不破坏生物膜结构的基础上干扰生物分子之间的弱多价相互作用,进而使相分离凝集体发生解聚^[223]。与其它已报道的能够进行相分离的蛋白质类似,我们观察到 10%的 1,6-己二醇溶液能够部分地溶解上述离心管底部的含有 CBX7A 或 CBX7C 的蛋白颗粒,而高浓度的 1,6-己二醇溶液(50%)几乎能够溶解所有的蛋白颗粒。这进一步说明 CBX7A 和 CBX7C 蛋白能够在体外发生相分离,且 CBX7C 蛋白驱动相分离发生的能力显著强于 CBX7A 蛋白。

随后,我们将上述离心管中的蛋白溶液进行充分重悬后将其滴加至载玻片上进行显微镜观察,并对每个蛋白样品的 4 个不同视野中的颗粒进行统计学分析(图 3-15B)。结果显示,含有 CBX7C 融合蛋白形成的荧光颗粒比 CBX7A 更显著,数量也更多;在浓度为 50%的 1,6-己二醇溶液中,大多数荧光颗粒逐渐溶解,然而仍有部分含有 CBX7C 的蛋白颗粒未溶解。已有研究表明蛋白分子发生相分离形成的凝集体在高浓度的 1,6-己二醇溶液中的溶解度能够被作为区分液体和固体凝集体的重要参数^[224]。因此,部分未溶解于 1,6-己二醇溶液的 CBX7C 蛋白

颗粒可能是通过液固相分离而产生的固体颗粒。综上所述表明,与 CBX7A 相比,体外高浓度 CBX7C 蛋白不仅具有驱动 LLPS 的能力,而且更倾向于引发液固转换。

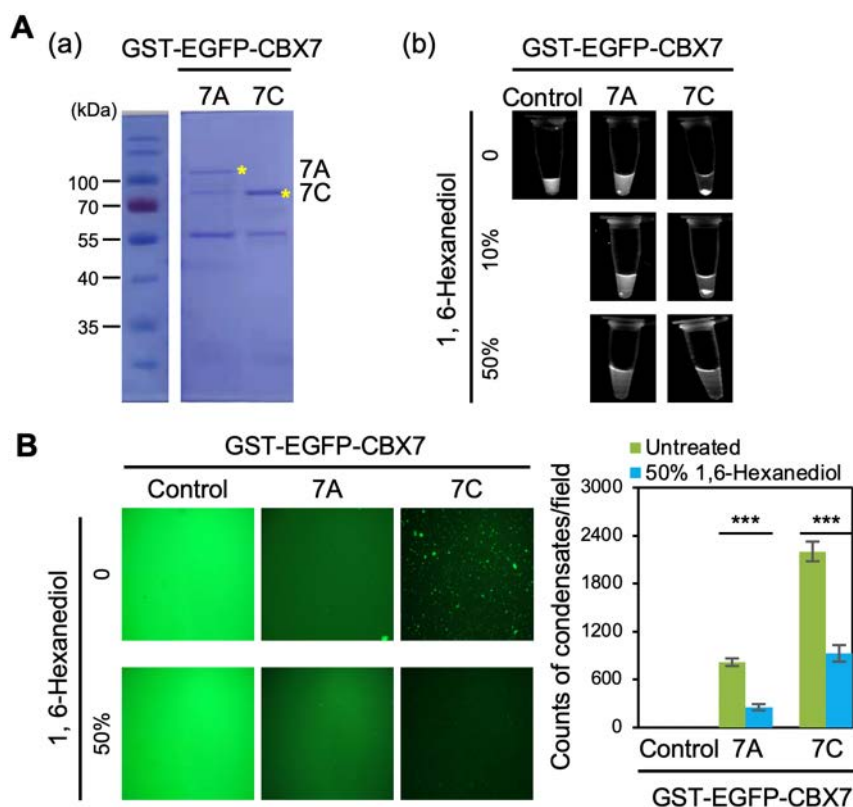


图 3-15 纯化蛋白 GST-EGFP-CBX7A 和 CBX7C 在体外发生相分离的比较分析

Figure 3-15 Comparative analysis of undergo phase separation between purified proteins GST-EGFP-CBX7A and CBX7C in vitro

A: 通过蛋白凝胶电泳检测纯化蛋白 GST-EGFP-Flag-CBX7A/C。“*”指示目的蛋白条带。其它条带为蛋白降解产物(a)。通过荧光显微镜观察分析高浓度纯化蛋白 GST-EGFP-CBX7A/C 在离心管内形成凝集体并分离的过程,以及不同浓度梯度的 1,6-己二醇对离心管底部凝集体的影响(b); **B:** 利用荧光显微镜观察高浓度纯化蛋白 GST-EGFP-Flag-CBX7A/C 在 1,6-己二醇处理前后形成的颗粒(左)。比例尺为 200 μm 。柱状图表示通过 ImageJ 统计分析多视野的平均颗粒数。通过 t 检验统计平均值之间的显著性差异 (***) $p < 0.001$ 。

3.5.2 CBX7C 在细胞内具有比 CBX7A 更强的相分离能力

CBX7-cPRC1 复合物能够通过相分离在细胞中产生的无膜结构 Pc 小体。为了进一步检测 CBX7 蛋白驱动相分离形成的凝集体在细胞内的理化特征,我们通过激光共聚焦显微镜对 HeLa 细胞中形成的含有 CBX7A-EGFP-Flag 或 CBX7C-EGFP-Flag 的 Pc 小体凝集体的图示选定区域进行了漂白后荧光恢复 (FRAP) 实验分析。结果显示,包含有 CBX7A 或 CBX7C 蛋白的 Pc 小体凝集体的荧光可以

在 488 nm 激光漂白后迅速恢复（图 3-16），说明 CBX7A 或 CBX7C 蛋白能够在凝集体中快速扩散，具有典型的液态动力学特征，符合 LLPS 凝集体的标准^[205]。

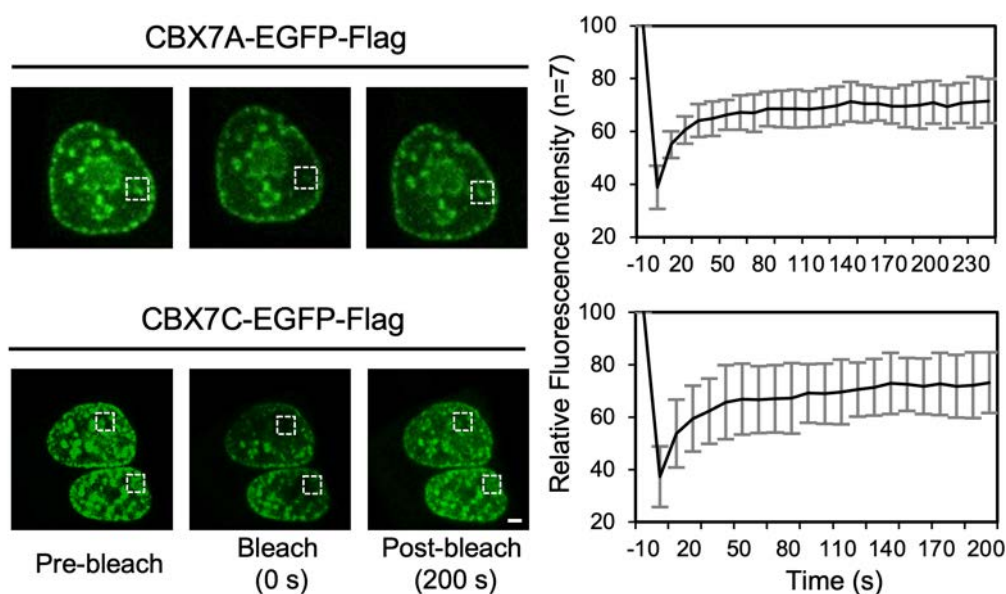


图 3-16 CBX7-cPRC1 在 HeLa 细胞中形成的 Pc 小体凝集体中的分子流动性测定

Figure 3-16 Measurement of molecular dynamic in Pc body condensates formed by CBX7-cPRC1 in HeLa cells

通过 FRAP 实验并结合共聚焦成像分析测定 CBX7-cPRC1 形成的 Pc 小体凝集体中的分子流动性（左）。折线图表示在光漂白（488nm 激光漂白）前后不同时间点上每个标记区域的荧光强度相对变化（右）。比例尺为 2 μm 。

为了进一步研究细胞中生物分子凝集物的性质，我们采用了近年来开发的基于植物光遗传学的 optoDroplet 系统，该系统可以使用蓝光调节蛋白的多价性，以控制体内生物分子凝集体的形成（图 3-17A）。Cry2 是一种拟南芥的蓝光受体，参与植物光形态发生的生物学过程，在蓝光激活后能够形成寡聚物^[209]。研究表明单独的 Cry2_{WT} 在蓝光激活时相对缓慢地显示出很少的聚集，而 IDR（如 FUS 的 N 末端）与 Cry2_{WT}（optoFUS）的融合表达导致大多数细胞中快速的蓝光依赖性凝集体组装^[225]。基于该系统，我们首先在 HeLa 细胞中表达 mCherry 标记的 CBX7-Cry2（optoCBX7），并测试蓝光介导的凝集反应。在蓝光持续刺激的过程中，每 30 s 采集一次图像。结果显示，在蓝光激发前，CBX7 的不同蛋白亚型在细胞中均能形成一定数量的凝集体，其中 CBX7C 产生的凝集体数量最高（图 17B）。在蓝光激发后，由所有 CBX7 蛋白亚型形成的凝集体的数量迅速增加。上述结果表明，所有三种 CBX7 蛋白亚型都具有一定的相分离能力，并且 CBX7C 的固有相分离能力较强。

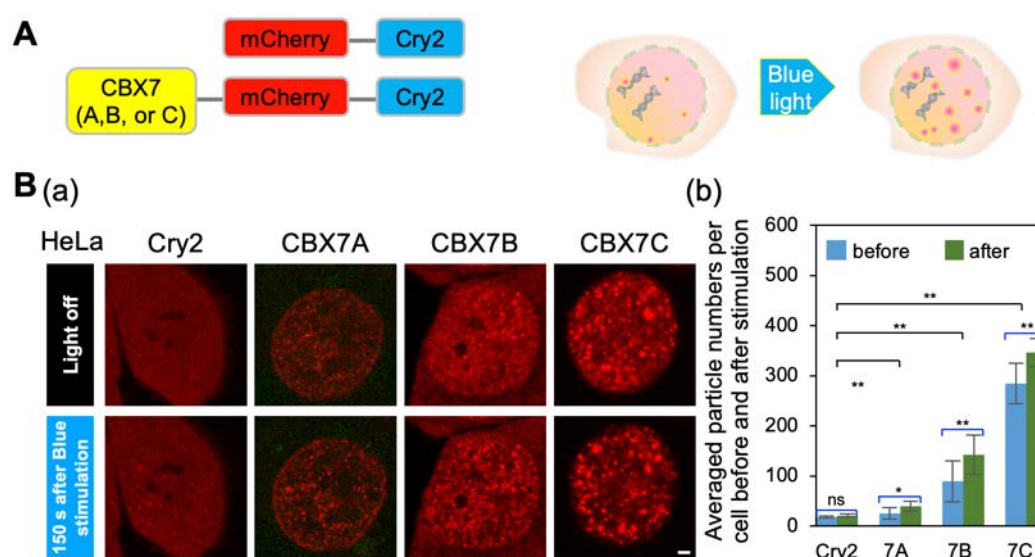


图 3-17 CBX7 不同蛋白亚型在 HeLa 中驱动相分离的能力分析

Figure 3-17 Ability analysis of different CBX7 subtypes to drive phase separation in HeLa

A: optoDroplet 系统原理示意图; B: 表达 CBX7-mCherry-Cry2 或 mCherry-Cry2 融合蛋白的 HeLa 细胞在蓝光激活前后的共聚焦成像分析 (b)。比例尺为 2 μm 。柱状图表示通过 ImageJ 分析图示蓝光激活前后的颗粒数。通过 t 检验统计平均值之间的显著性差异 (ns , 不显著, $p>0.05$; $*p<0.05$; $**p<0.01$)。

小结

在适当盐离子浓度、pH 及温度的体外溶液环境下, 含有 IDR 的 CBX7A 和 CBX7C 蛋白都能发生相分离。在活细胞中, CBX7A 和 CBX7C 也通过相分离参与核内无膜结构 Pc 小体在染色质上的构建。但是无论在成分单一的体外溶液中还是复杂的细胞内, CBX7C 表现出比其它 CBX7 亚型更强的驱动相分离的能力。

3.6 PHC2 与 CBX7A 或 CBX7C 的蛋白互作调节相分离特性

3.6.1 CBX7•PHC2 在细胞核内逐渐形成体积较大且分子流动性更低的凝集体

上文所述的 CBX7 与 cPRC1 各组分互作研究结果显示, CBX7 与 RING1B 的互作荧光信号分布于整个细胞核中 (图 3-2A), 而 CBX7 与 PHC2 互作后发生了相分离, 产生了显著的体积较大的致密荧光亮斑 (图 3-4A 和图 3-18A)。此外, 共聚焦成像结果显示, CBX7•PHC2 形成凝集体的区域 Hoechst 着色极浅, 甚至未检测到信号, 这说明该互作产生的相分离排斥其中的染色质组分 (图 3-18A)。与此同时, 在明场观察中, 核内对应于 CBX7•PHC2 的互作荧光亮斑的位置处形成了一个明显的液滴状结构 (图 3-18A), 进一步证实了在 CBX7 与 PHC2 互作

过程中发生了相分离。

进一步，我们通过 FRAP 实验分析了由 CBX7C•PHC2 形成的凝集体的理化特征。与图 3-16 中所示的由 CBX7A-EGFP 或 CBX7C-EGFP 形成的凝集体相比，CBX7C•PHC2 互作凝集体需要被更强的激光激发才能有效地漂白较强的荧光信号，并且荧光强度恢复速度缓慢，直至光漂白后近 15 分钟才恢复部分荧光信号（图 3-18B）。这说明与 CBX7A/C-cPRC1 形成的 Pc 小体相比，CBX7C•PHC2 互作凝集体中的分子流动性较低，其物理性质更接近于水凝胶状态。综上结果表明，当高浓度的 CBX7C 和 PHC2 蛋白共存于细胞中时，二者互作并驱动了相分离发生，并且在此互作蛋白凝集体形成的过程中逐渐排斥了其中的染色质组分，形成更加致密的由蛋白质组成的水凝胶状结构。这一凝集体的形成可能扰乱了胞内的染色质分布，进而阻碍了细胞的正常生长并引发了细胞的凋亡。

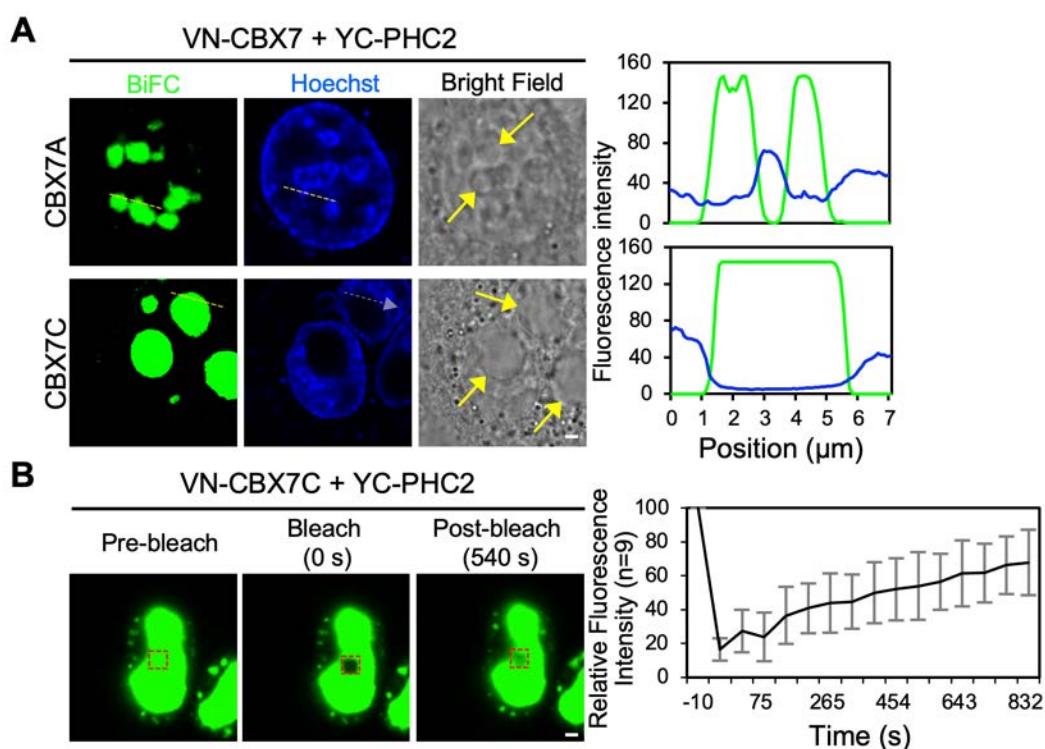


图 3-18 CBX7•PHC2 互作凝集体的观察及分子流动性分析

Figure 3-18 Observation and dynamic analysis of CBX7•PHC2 condensates

A: 通过共聚焦成像观察分析 VN-CBXA 和 CBX7C 与 YC-PHC2 互作后形成凝集体和明场中对应的液滴状结构，及其与 Hoechst 标记的染色质区域的互斥关系。明场图片中黄色箭头指示呈液滴状结构的 CBX7•PHC2 凝集体。曲线图表示图示虚线区域上的荧光强度。Hoechst 信号指示细胞核；B: 通过 FRAP 实验并结合共聚焦成像分析测定 VN-CBX7C 与 YC-PHC2 互作后形成的凝集体中的分子流动性。折线图表示在光漂白（488nm 激光漂白）前后不同时间点上每个标记区域的荧光强度相对变化。比例尺为 2 μm 。

此外，我们通过高分辨率活细胞成像系统(DeltaVision)观察了 CBX7C•PHC2

互动时相分离发生的动态过程，以及互动凝集体的形成对细胞状态的影响。如图 3-19 所示，在 VN-CBX7C 和 YC-PHC2 双质粒转染至 HeLa 细胞后 18-30 小时内每小时采集系列共聚焦图像。在此过程中，随着 CBX7C•PHC2 互动凝集体的开始形成，其体积和荧光强度逐渐增加，形成了由小及大的互动荧光亮斑，而不是由若干个凝集体通过彼此融合而形成的（图 3-19 左）。同时，该过程还伴随着在凝集体形成的区域中 Hoechst 信号逐渐减弱，直至消失（图 3-19 左）。随着时间推移，与同视野中未表达 VN-CBX7C 与 YC-PHC2 的对照组 HeLa 细胞相比（图 3-19 右下），CBX7C•PHC2 互动凝集体的形成导致细胞形态发生异常，并脱离培养皿底悬浮于培养基中，引发了细胞的生长异常（图 3-19 右上）。这表明高浓度的 CBX7C 与 PHC2 共存于细胞后，二者互动并驱动相分离的发生，形成致密的仅由蛋白质组成的水凝胶状结构的凝集体，进而排斥凝集体中染色质组分，扰乱了细胞内染色质分布，进而阻碍了细胞正常生长。

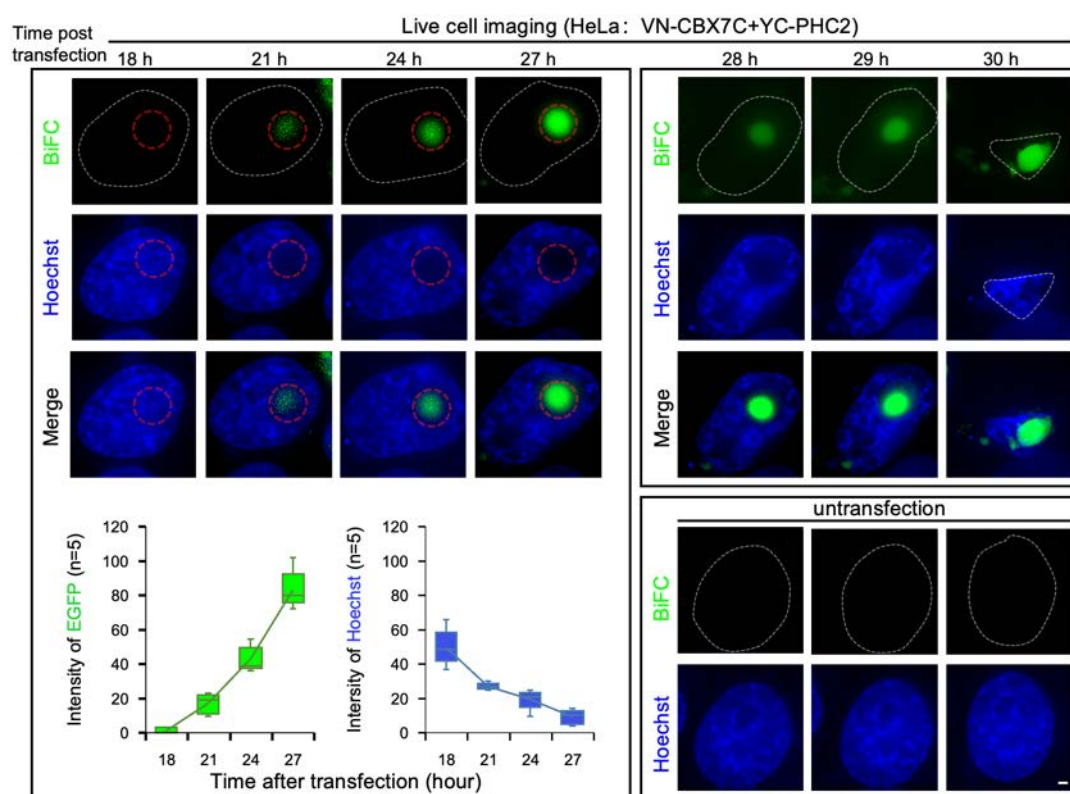


图 3-19 CBX7C•PHC2 形成凝集体并排斥染色质过程及其对细胞状态影响的分析

Figure 3-19 Analysis of the process of CBX7C•PHC2 forming condensates and rejecting chromatin and its impact on cell state

通过 DeltaVision 实时成像技术分析 VN-CBX7C 与 YC-PHC2 形成互动凝集体并排斥染色质过程，以及凝集体形成引发细胞凋亡。Hoechst 信号指示细胞核。箱形图表示在 CBX7C•PHC2 凝集体形成的不同时间点，图示红色圆圈标记区域的荧光强度或 Hoechst 信号变化。图中虚线区域指示细胞核轮廓。比例尺为 2 μ m。

3.6.2 CBX7•PHC2 形成的凝集体包含了完整的 cPRC1 核心组分

我们通过 IF 实验结合共聚焦成像分析探究了排斥染色质的 CBX7•PHC2 互作凝集体中是否能够招募 cPRC1 的其它核心组分。在未处理的 HeLa 细胞中，内源表达的 RING1B、PCGF2 以及 CBX7A 蛋白均匀的分布在细胞核中，并且形成经典的颗粒状的 Pc 小体（图 3-20A）。当在细胞中瞬时表达 VN-CBX7 与 YC-PHC2 蛋白互作并形成凝集体后，绝大多数内源性的 RING1B 和 PCGF2 蛋白被显著集中至 CBX7•PHC2 互作凝集体中，表明完整的 cPRC1 复合体在 CBX7•PHC2 相分离形成的互作凝集体中大量积累（图 3-20B 和 C）。在图像数据分析时，我们还发现 CBX7C•PHC2 形成的互作凝集体虽然包含了 cPRC1 复合体的其它组分，但是它却显著排斥了内源性的 CBX7A 蛋白（图 3-20B 和 C）。结合上文 CBX7C 在染色质上的结合降低了 CBX7A 在该位点上富集的数据（图 3-6），进一步证明 CBX7C 在染色质上参与 cPRC1 复合体组装的过程中与 CBX7A 是相互排斥或拮抗的关系。

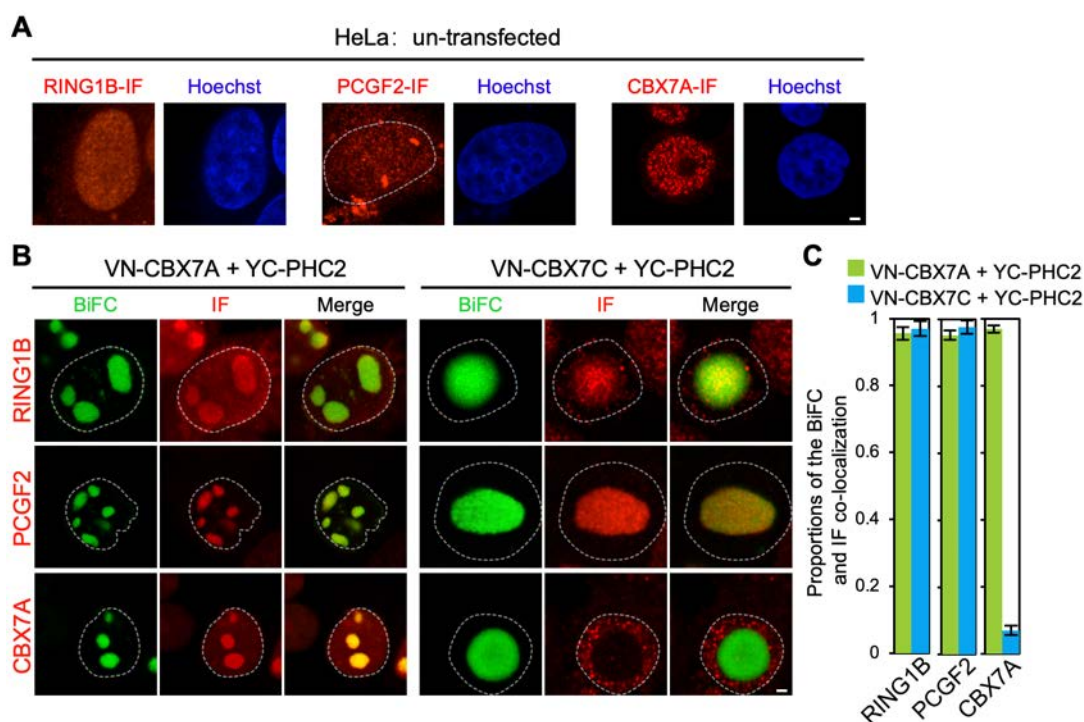


图 3-20 CBX7•PHC2 互作凝集体中的蛋白组分分析

Figure 3-20 Analysis of protein components in CBX7•PHC2 interacting condensates

A: 通过 IF 实验结合共聚焦成像检测未处理的 HeLa 细胞中内源 cPRC1 酶学核心组分 RING1B 和 PCGF2 的亚细胞定位。Hoechst 信号指示细胞核；B: 通过 IF 实验结合共聚焦成像检测 HeLa 细胞中 CBX7•PHC2 互作凝集体是否含有 cPRC1 核心酶学组分 RING1B 和 PCGF2，以及 CBX7C•PHC 互作凝集体与 CBX7A 的定位关系。图中虚线区域指示细胞核轮廓。比例尺为 2 μm；C: 柱状图表示通过 ImageJ 分析 BiFC 和图示的 IF 信号能够共定位的细胞数目总 BiFC 信号阳性的细胞数目的比例。

上述结果表明通过相分离形成的 CBX7•PHC2 互作凝集体中包含了完整的 cPRC1 组分。基于此，我们进一步通过 IF 实验结合共聚焦成像分析了包含有核心酶学组分 RING1B 和 PCGF2 蛋白的 CBX7•PHC2 互作凝集体是否具有酶学活性，是否也能够与 H2AK119ub 共定位。结果显示，包括 H2AK119ub 在内的四种组蛋白修饰均被排斥在凝集体之外（图 3-21），该结果与图 3-19 和 3-20 所示的结果是一致的，这可能是由于高浓度的 CBX7-PHC2 驱动了液-固相分离的发生：一方面，形成的致密水凝胶状的凝集体将染色质组分从凝集体中排斥出去，仅剩余高浓度的蛋白组分；另一方面，致密水凝胶状的凝集体将具有酶学活性的 RING1B•PCGF2 包裹起来，进而使其与底物组蛋白隔离。总之，尽管 CBX7•PHC2 凝集体中存在完整的 cPRC1 复合体，但是这种高浓度蛋白形成的水凝胶状凝集体可能产生了空间阻隔，进而影响了 cPRC1 复合体的组蛋白修饰的催化功能。

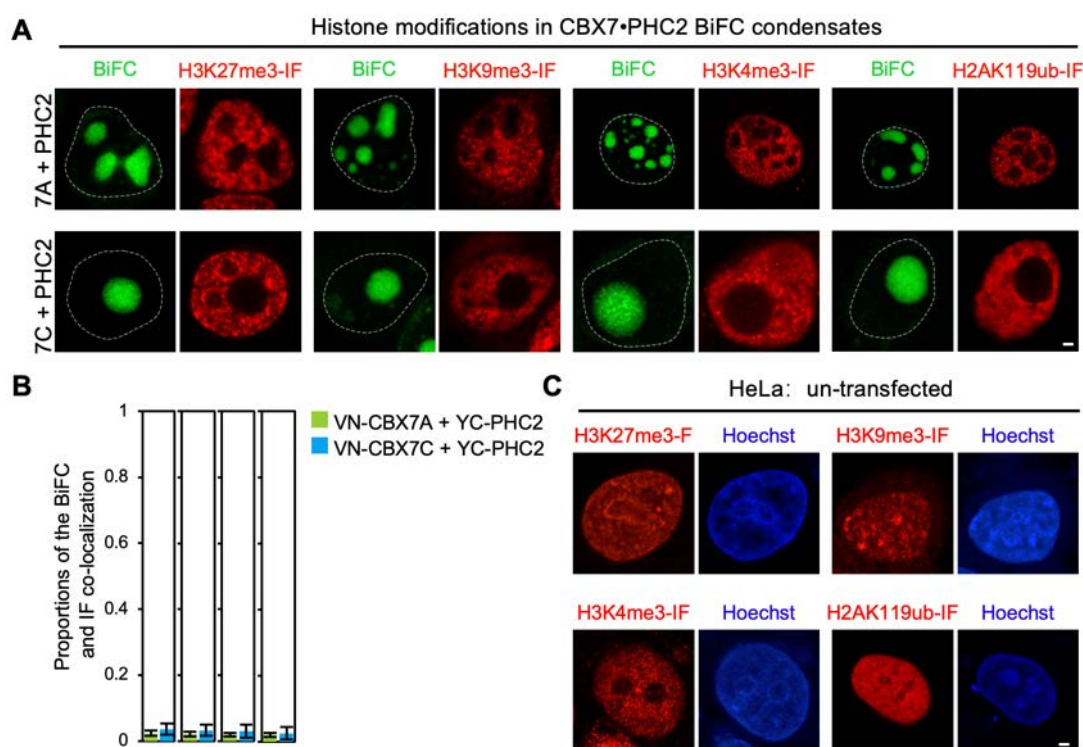


图 3-21 CBX7•PHC2 互作凝集体的形成对组蛋白修饰的影响分析

Figure 3-21 The affect of CBX7•PHC2 condensates formation on histone modification

A: 通过 IF 实验结合共聚焦成像检测 HeLa 细胞中形成的 CBX7•PHC2 互作凝集体与图示染色质组蛋白修饰的定位关系；**B:** 柱状图表示通过 ImageJ 分析 BiFC 和图示的 IF 信号能够共定位的细胞数目总 BiFC 信号阳性的细胞数目的比例。**C:** 通过 IF 实验结合共聚焦成像检测在未处理的 HeLa 细胞中图示染色质组蛋白修饰的亚细胞定位。Hoechst 信号指示细胞核。图中虚线区域指示细胞核轮廓。比例尺为 2 μ m。

3.6.3 PHC2 作为 CBX7 相变促进因子而决定相分离状态

为了排除 BiFC 系统中 VN 与 YC 的不可逆结合对结论的影响，我们通过

HTM 系统结合 FRAP 实验进一步探究了 PHC2 对 HeLa 细胞中 CBX7A 或 CBX7C 的相变能力的影响。与 mESC 中所得结果相同, mCherry-HTM-CBX7 发生异位靶向并形成较小颗粒, EGFP-PHC2 均匀分布在 HeLa 细胞核中 (图 3-22A)。然而高水平表达的 CBX7A/C 和 PHC2 在核内互作后改变了原有的定位分布, 形成体积较大的凝集体, 并排斥其中的染色质成分 (图 3-22B)。CBX7C•PHC2 形成的凝集体的平均体积显著大于 CBX7A•PHC2。此外, CBX7C•PHC2 互作复合物形成的凝集体的体积也普遍大于单独表达的 CBX7C 或 PHC2 (图 3-22C)。上述结果进一步证明了 CBX7C 的相分离能力强于 CBX7A 的结论, 并且 CBX7C 和 PHC2 能够相互调节彼此的相分离能力。

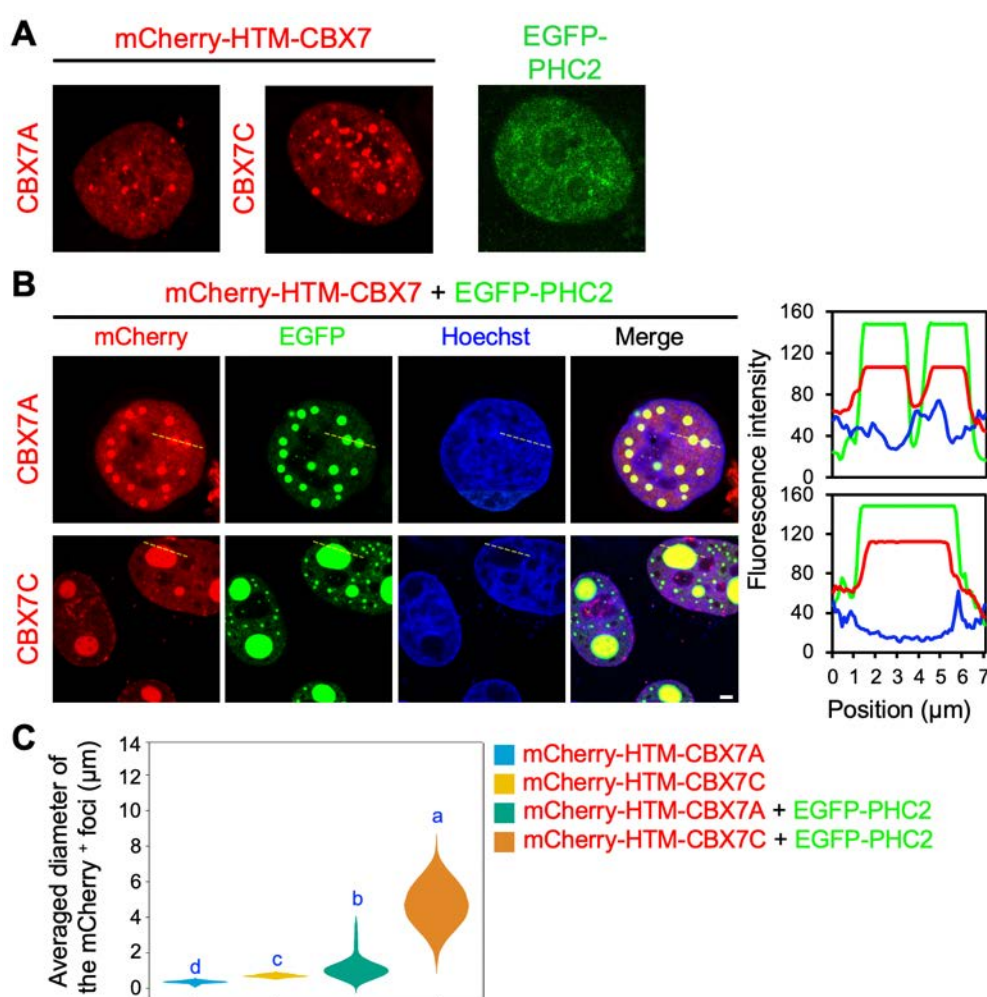


图 3-22 PHC2 对 CBX7 相分离能力的影响

Figure 3-22 The affect of PHC2 on the phase separation activity of CBX7

A: mCherry-HTM-CBX7 或 EGFP-PHC2 在 HeLa 细胞中的共聚焦成像分析; B: mCherry-HTM-CBX7 与 EGFP-PHC2 在 HeLa 细胞共表达的共聚焦成像分析。Hoechst 信号指示细胞核。曲线图表示图示虚线区域的荧光强度 (右); C: 小提琴图表示通过 ImageJ 分析具有 mCherry 荧光的颗粒直径, 图示不同的小写字母表示 ANOVA 分析中 Duncan 检验的显著性差异 ($p < 0.05$)。比例尺为 $2 \mu\text{m}$ (A) 或 $4 \mu\text{m}$ (B)。

此外，我们通过 FRAP 测定，对上述 CBX7A·PHC2 和 CBX7C·PHC2 互作复合物形成的凝集体的分子流动性进行比较分析。如图 3-23 所示，mCherry-HTM-CBX7A 或 CBX7C 的荧光信号在光漂白后迅速恢复，说明其形成了分子流动性较高的液态凝集体。相反，CBX7A·PHC2 和 CBX7C·PHC2 互作复合物形成的体积较大的凝集体中的荧光信号需要更长的时间才能恢复。上述结果表明，CBX7C/CBX7A 与 PHC2 的互作将显著改变所得凝集体的特性，促进其更加致密的凝集，并最终将其转变为水凝胶状态。

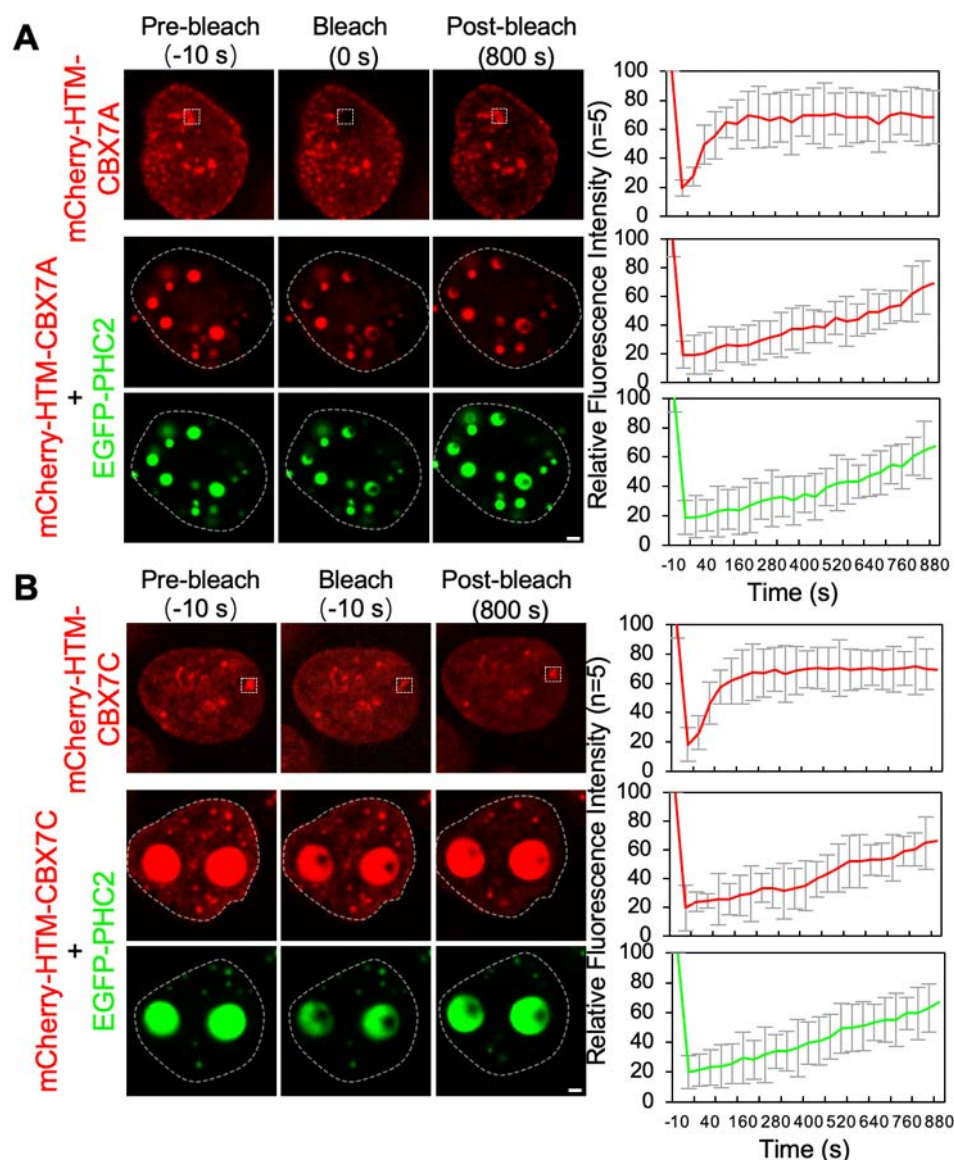


图 3-23 PHC2 对 CBX7 凝集体分子流动性的影响

Figure 3-23 The effect of PHC2 on the molecular dynamic of CBX7 condensates

A-B: FRAP 实验分析测定 HeLa 细胞中图示各组形成的颗粒的分子流动性。曲线图表示通过 ImageJ 分析图示具有荧光的细胞与 Hoechst 信号阳性细胞数的比例。右边的折线图表示在光漂白（488nm 激光漂白）前后不同时间点上每个标记区域的荧光强度相对变化。图中虚线区域指示细胞核轮廓。比例尺为 2 μ m。

3.6.4 CBX7C 的 N 端区域是介导其与 PHC2 互作的主要区域

为了进一步明确介导 CBX7•PHC2 互作及相分离的蛋白关键区域,以及比较 CBX7A 和 CBX7C 在这一过程中的差异,我们根据不同策略将 CBX7A 和 CBX7C 进行截短并与 GST 标签融合表达,通过体外 His-pull down 实验探究了 CBX7 与 PHC2 的互作区域。N 端(1-82 aa)序列是 CBX7A 和 CBX7C 相同的区域,而由于 mRNA 的选择性剪接,它们的 C 末端完全不同。因此,我们首先将 CBX7A 或 CBX7C 在第 82 位氨基酸残基后进行截短表达,产生 GST 标记的相同截短体 GST-CBX7(1-82 aa)和差异截短体 CBX7A(83-251 aa)与 GST-CBX7C(83-166 aa)。His-pull down 实验结果表明,如同 CBX7A 或 CBX7C 的全长,以上三种截短体在体外均能够与 PHC2 蛋白直接互作(图 3-24A)。

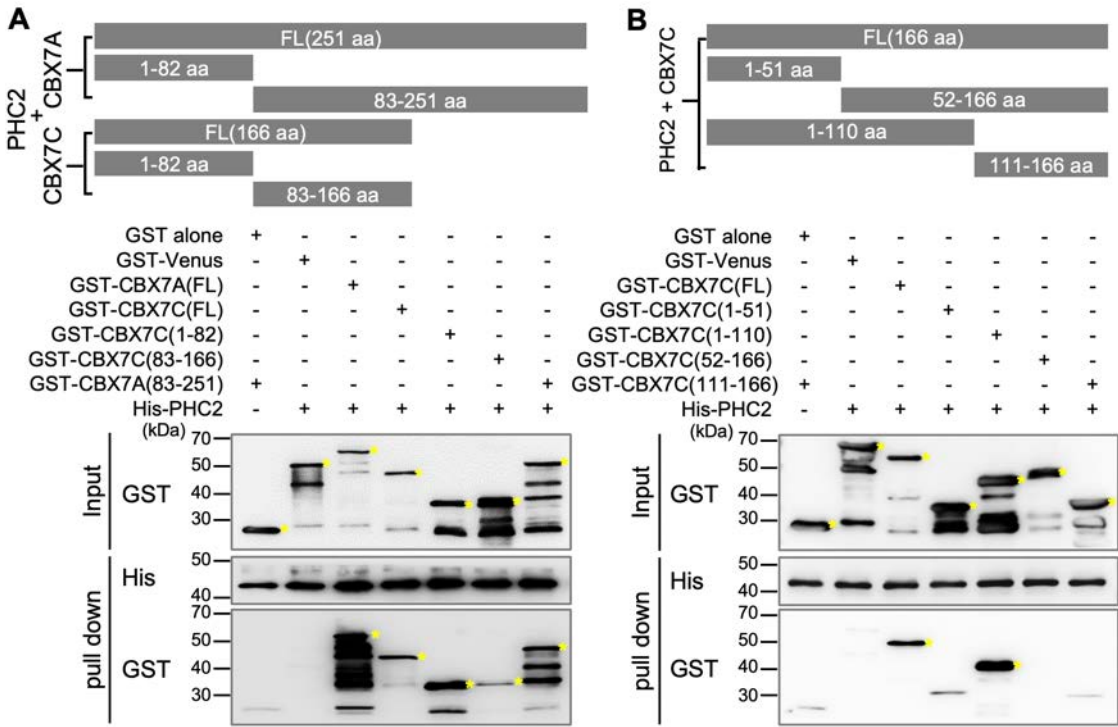


图 3-24 介导与 PHC2 互作的 CBX7C 蛋白区域的鉴定

Figure 3-24 Identification of CBX7C protein regions that mediate interactions with PHC2
GST-pull down 及 WB 检测原核表达蛋白 GST-CBX7A 或 C 及其截断体与 His-PHC2 的互作。
“*”指示目的蛋白条带。

其次,我们重点关注了 CBX7C 的不同区域与 PHC2 的互作关系,并采用第二种截短策略进一步分析。结合图 3-16A 的分析,对于 CBX7C 的 IDR 区域预测结果显示,CBX7C 的第 52 至 110 位氨基酸残基区域是其最长的 IDR 区域,我们将第 52 位或第 110 位作为截短点,表达了四个 CBX7C 的截短体,包括 1-51 aa、52-166 aa、1-110 aa 和 111-166 aa。结果表明,CBX7C(1-110 aa)截短体在体

外与 PHC2 存在直接互作。而 WB 结果显示被 His-PHC2 结合的 CBX7C(1-51 aa) 或(111-166 aa)截短体被抗体标记的条带较浅。同时, 几乎未检测到 CBX7C(52-166 aa)的蛋白条带, 说明 CBX7C 的 52-166aa 截短体丧失了与 PHC2 互作的能力(图 3-24B)。在上述结果中, 虽然 CBX7C(1-110 aa)与 PHC2 能够直接互作, 但是包含其中的 1-52 截短体仅有少量被 His-PHC2 结合, 这可能是由于该截短策略破坏了参与相互作用的区域。上述结果表明 CBX7C 的 N 端(1-110 aa)是介导 CBX7C 与 PHC2 互作的主要区域。

3.6.5 CBX7 的 C 端是参与 CBX7•PHC2 发生相分离形成凝集体的关键区域

探究复杂的凝集体是如何在细胞中形成的, 以及识别驱动其形成的关键蛋白质, 并确定这些蛋白质的序列如何对相分离的发生做出贡献是至关重要的。PHC2 蛋白具有 LLPS 的性质, 其 SAM 结构域在这一过程中发挥了重要作用^[51, 226]。为了探究参与 CBX7C 与 PHC2 互作的主要区域 CBX7C(1-110 aa)是否也介导了 CBX7-PHC2 互作蛋白的相分离过程, 我们通过 BiFC 和共聚焦显微成像技术分析了 CBX7A 和 CBX7C 的 N 端((1-82 aa)和(1-110 aa))和 C 端((83-251 aa)和(111-166aa))对相分离的影响(图 3-25)。

结果显示, CBX7A 和 CBX7C 的 N-或 C-末端足以与 PHC2 协同驱动相分离。但是当 PHC2 中 SAM 缺失时, CBX7•PHC2 介导的相分离形成被显著遏制, 大量 BiFC 信号弥散分布在整个细胞核中(图 3-25)。此外, 在 CBX7 亚型的 N-末端和 C-末端驱动相分离的能力具有显著差异。CBX7_N•PHC2_{ΔSAM} 形成的绝大多数凝集体能够像正常的 Pc 小体一样与染色质结合(图 3-26)。相反, 具有 CBX7_C•PHC2_{ΔSAM} 仍然能够形成体积较大并排斥染色质的凝集体(图 3-25B)。与 CBX7A_C•PHC2_{ΔSAM} 相比, CBX7C_C•PHC2_{ΔSAM} 的 BiFC 信号仍然被完全的集中在形成的凝集体中(图 3-25B)。这意味着 CBX7C 的 C 末端比 CBX7A 在驱动相分离过程中发挥了更重要的作用, 这也很好地解释了 CBX7C 和 CBX7A 之间的不同相分离能力。上述结果表明, CBX7 的 N 末端在染色质靶向和功能性 PRC1 的建成中起着至关重要的作用, 而 PHC2 的 SAM 结构域更易影响无染色质 CBX7•PHC2 凝集体的协同相分离性质。综上所述, CBX7•PHC2 介导的 Pc 小体形成可能是一个具有不同物理状态的可控相分离事件。当核心因子以低剂量或中等剂量表达时, 功能性 Pc 小体凝集体主要通过 CBX7C 的 N 末端和 PHC2 之间的分子间互作形成, 并以分子流动性较高的液态存在; 在高表达核心成分的情况下, 在 CBX7C 的 C 末端和 PHC2 的 SAM 结构域的共同参与下促进其形成功能失调并排斥染色质的水凝胶状凝集体。

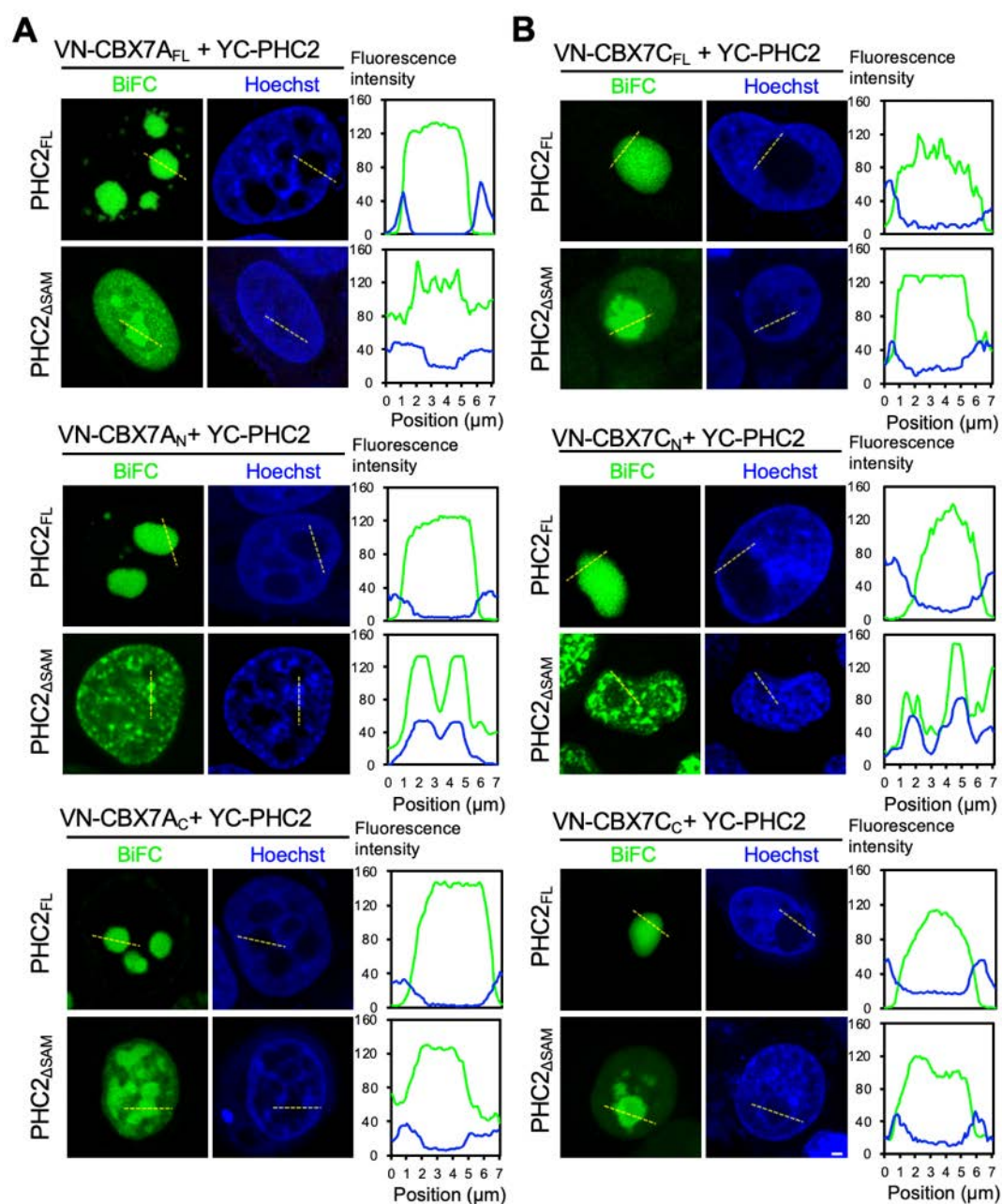


图 3-25 CBX7•PHC2 协同相分离的相关区域的鉴定

Figure 3-25 Identification of relevant regions for CBX7 • PHC2 synergistic phase separation
A-B: 通过共聚焦成像分析 HeLa 细胞中图示 CBX7A (A) 和 CBX7C (B) 不同截断体的相分离特性的分析。曲线图表示图示虚线区域上的荧光强度。Hoechst 信号指示细胞核。比例尺为 2 μm 。

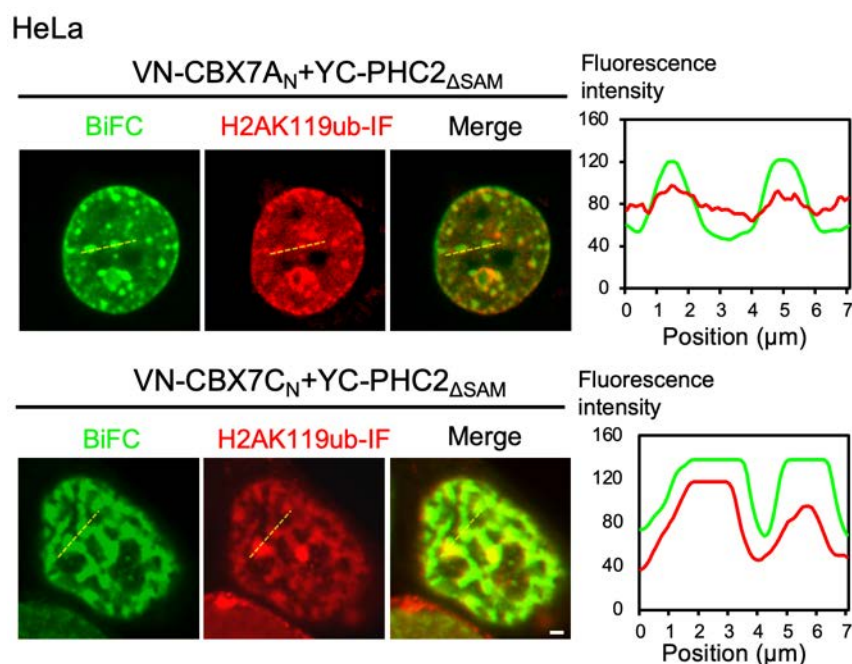


图 3-26 CBX7 和 PHC2 截断体产生的凝集体对 H2AK119ub 分布的影响

Figure 3-26 The affect of CBX7•PHC2 truncation condensates on H2AK119ub distribution
通过 IF 实验结合共聚焦成像分析 HeLa 细胞中 H2AK119ub 与图示形式的 CBX7 和 PHC2 截断体产生的凝集体的分布关系。曲线图表示图示虚线区域上的荧光强度。Hoechst 信号指示细胞核。比例尺为 2 μm。

与此同时，我们还探究了 SAM 结构域对 CBX7•PHC2 互作凝集体形成的主导作用是否与其特殊的分子结构相关。已发表的研究表明，PHC2 具有 4 个保守的结构域（图 3-27A (a)），其中 HD1 结构域介导了 PHC2 与 PCGF2 或 RING1B 的互作，而 SAM 结构域能够引发 PHC2 的蛋白自聚^[49]。SAM 域包含 ML 和 EH 结合面，这两个部分能够以首尾相接的形式介导蛋白自聚^[51]。并且 EH 结合面的突变（L307R）或 ML 结合面的突变（L293R/H298R）都能导致 SAM 结构域介导的蛋白自聚能力的丧失^[51]。为了探究 SAM 结构域对 CBX7•PHC2 互作凝集体形成的影响，我们在 HEK293T 细胞中观察分析了不同的 SAM-PHC2 突变体及 HD1 缺失的 PHC2 截短体与 CBX7 的相互作用（图 3-27A(b)）。结果显示，突变体 PHC2（L307R）及截短体 PHC2(ΔHD1)仍然能够与 CBX7 相互作用，但是 ML 结合面的突变（L293R/H298R）导致了 PHC2 与 CBX7 的互作被破坏。这说明 PHC2 的 ML 结合面是介导 PHC2 与 CBX7 互作的关键区域。

IDR 是相分离蛋白中一种常见的结构域，其一级序列往往决定了这些蛋白的相变能力、相变的驱动因素、相变的临界浓度及黏弹性^[227]。此外，氨基酸的疏水性和非离子相互作用对于相分离也具有重要作用^[228]。基于此，我们将 CBX7 的 C 端(83-166 aa)中的包含较多的疏水性氨基酸亮氨酸（L）突变为带正电荷的

亲水性氨基酸精氨酸 (R)，以期探究是否能够影响 CBX7C 的 C 端参与相分离的能力。然而结果显示，突变体 CBX7C(83-166, 10L to R) 与 PHC2 的互作及相分离并未受到影响 (图 3-27B)。在本课题后续的研究中，还需进一步明确 CBX7C 参与相分离的氨基酸位点。

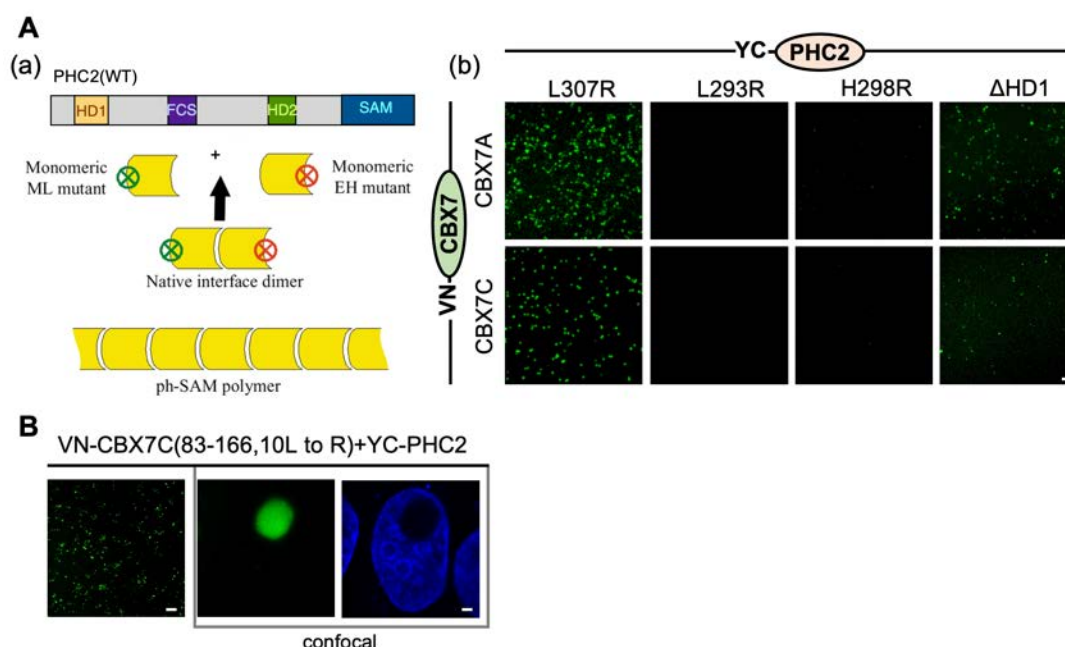


图 3-27 PHC2 的 SAM 结构域及 CBX7C_C 的疏水性对 CBX7•PHC2 凝集体形成的影响

Figure 3-27 The affect of SAM domain of PHC2 and hydrophobicity of CBX7C_C on the formation of CBX7•PHC2 interacting condensates

A: PHC2 的保守结构域示意图以及 ML 和 EH 介导蛋白自聚的模式图^[49, 52] (a)。通过 BiFC 检测 PHC2 的突变体对 CBX7 与 PHC2 互作的影响 (b)。比例尺为 200 μm; B: 通过 BiFC 和共聚焦成像分析 CBX7C 的突变体对 CBX7C 与 PHC2 互作及相分离的影响。Hoechst 信号指示细胞核。比例尺分别为 200 μm (左, 普通倒置荧光显微镜成像) 和 2 μm (右, 共聚焦成像)。

3.6.6 PHC2 与其它 CBX 蛋白或 cPRC1 复合体组分互作后发生相分离

CBX7 是 mESC 中 cPRC1 复合体组装的 CBX 家族中的主要成员之一，其在细胞分化时被家族其它 CBX 蛋白所取代^[136]。因此，我们同时还探究了其它 CBX 蛋白及 cPRC1 组分是否与 PHC2 也能够互作并发生相分离。如图 3-28 所示，在 Hela 细胞中，CBX2、CBX4 和 CBX8 与 PHC2 互作后也能够形成与 CBX7 类似的凝集体。尤其是 CBX2 与 PHC2 的互作信号全部集中于中央体积较大的荧光亮斑中，而 CBX4-PHC2 形成的互作凝集体的体积相对较小。同时，绝大部分的 CBX8 与 PHC2 的互作信号集中于大体积的荧光亮斑中，但是仍有极少部分荧光信号分散在细胞核中，说明这部分互作蛋白可能未发生相分离。对于 CBX6，尽

管它与 PHC2 在 BiFC 实验系统中能够产生互作荧光信号, 但是 CBX6•PHC2 并未发生相分离现象, 而是 BiFC 互作信号在细胞核中均匀分布。上述结果表明 PHC2 在促进几种不同 CBX 蛋白的凝集体形成的过程中发挥了重要作用。此外, RING1B•PHC2 在细胞核中也能够完全凝聚并发生相分离而形成体积较大的凝集体, 而仅部分 PCGF2•PHC2 互作蛋白发生了相分离, 其余互作信号仍然弥散分布于细胞的核质中。综上所述, PHC2 作为 CBX 蛋白和潜在的其它 cPRC1 成分的通用相变增强剂, 它们协调相互作用的浓度或强度可能决定它们的相分离。

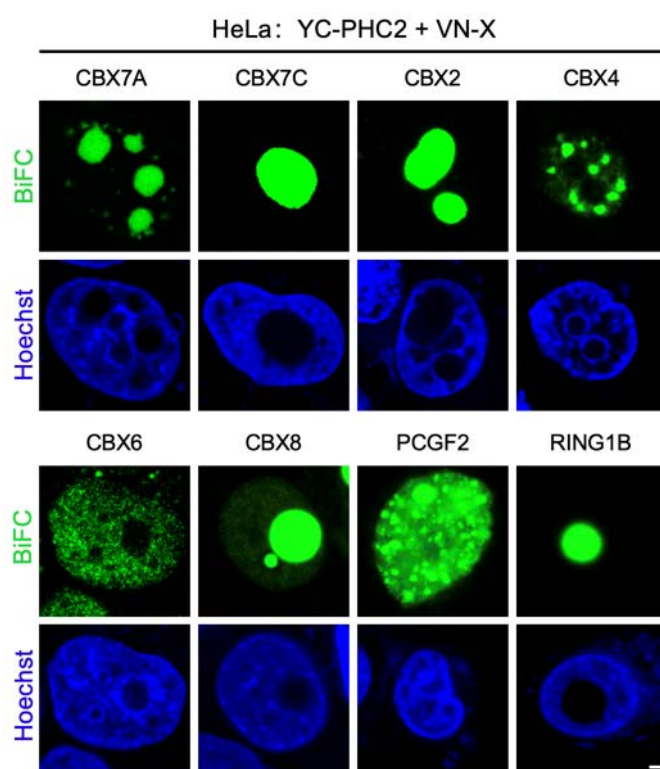


图 3-28 PHC2 与其它 CBX 蛋白及 cPRC1 组分互作及相分离现象的分析

Figure 3-28 Analysis of the interaction and phase separation between PHC2 and other CBX proteins or cPRC1 components

通过 BiFC 实验结合共聚焦成像检测 HeLa 细胞中 PHC2 与所示其它 CBX 蛋白及 cPRC1 组分互作及相分离特性。Hoechst 信号指示细胞核。比例尺为 2 μm 。

小结

正常生理条件下, CBX7A、CBX7C 或 PHC2 在组成型异染色质上通过相分离参与形成具有催化活性的 Pc 小体。当 PHC2 与 CBX7A/CBX7C 在细胞中过量共表达后进一步促进相分离而形成的致密蛋白凝集体。尽管招募了大部分内源性 cPRC1 的酶学组分 RING1B 和 PCGF2, 但是该水凝胶状凝集体由于空间分隔而排斥染色质成分并失去了催化组蛋白修饰的酶活性。在此过程中, CBX7C 仍表

现出强于 CBX7A 的相分离能力, 并且二者无法共存于同一凝集体中。CBX7C 和 PHC2 的协同作用有力地控制了相变和 Pc 小体的功能。PHC2 还可作为 CBX 蛋白和潜在的其他 cPRC1 成分的通用相分离“增强子”, 通过协调相互作用的浓度或强度而决定相分离形成的凝集体的物理状态。

3.7 CBX7C 在 mESC 分化过程中上调表达并参与调控 EB 的形成

3.7.1 CBX7C 蛋白的敲低导致 mESC 状态失稳

为了深入研究 CBX7C 在 mESC 的增殖及分化中的生物学功能, 我们针对于 CBX7 的 C 端(83-166 aa)区域设计了 3 条 shRNA 靶向序列, 在 mESC 中构建 CBX7C 敲低的稳定细胞系(图 3-29A)。我们发现在利用慢病毒感染并筛选阳性细胞的过程中, CBX7C 蛋白的表达沉默(图 3-29B), 对 mESC 的生长产生严重影响, 使其处于失稳状态。如图 3-30 所示, 在加入嘌呤霉素筛选七天后, 与对照组细胞相比, CBX7C 表达沉默的细胞形态发生明显变化, 不再呈集落状生长, 且无法形成饱满立体的克隆。同时, 细胞增殖显著减慢, 甚至停滞, 致使无法得到能够正常培养生长的 CBX7C 敲低稳定细胞系。基于此, 我们认为尽管 CBX7C 在 mESC 中的表达水平低于 CBX7A 与 CBX7B, 但是 CBX7C 在 mESC 的增殖与稳态中可能具有重要生物学功能。

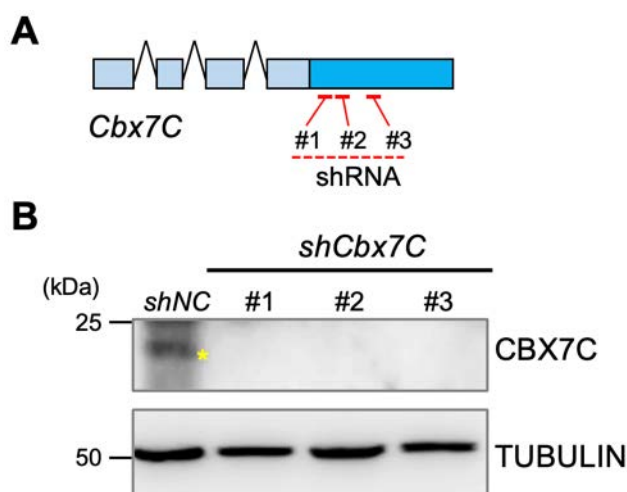


图 3-29 CBX7C 敲低的 mESC 细胞系鉴定

Figure 3-29 Identification of mESCs with CBX7C knockdown

A: 用于敲低 mESC 内 CBX7C 的 shRNA 靶向位置示意图; B 通过 WB 鉴定 CBX7C 敲低的 mESC 细胞系。“*”指示目的蛋白条带。

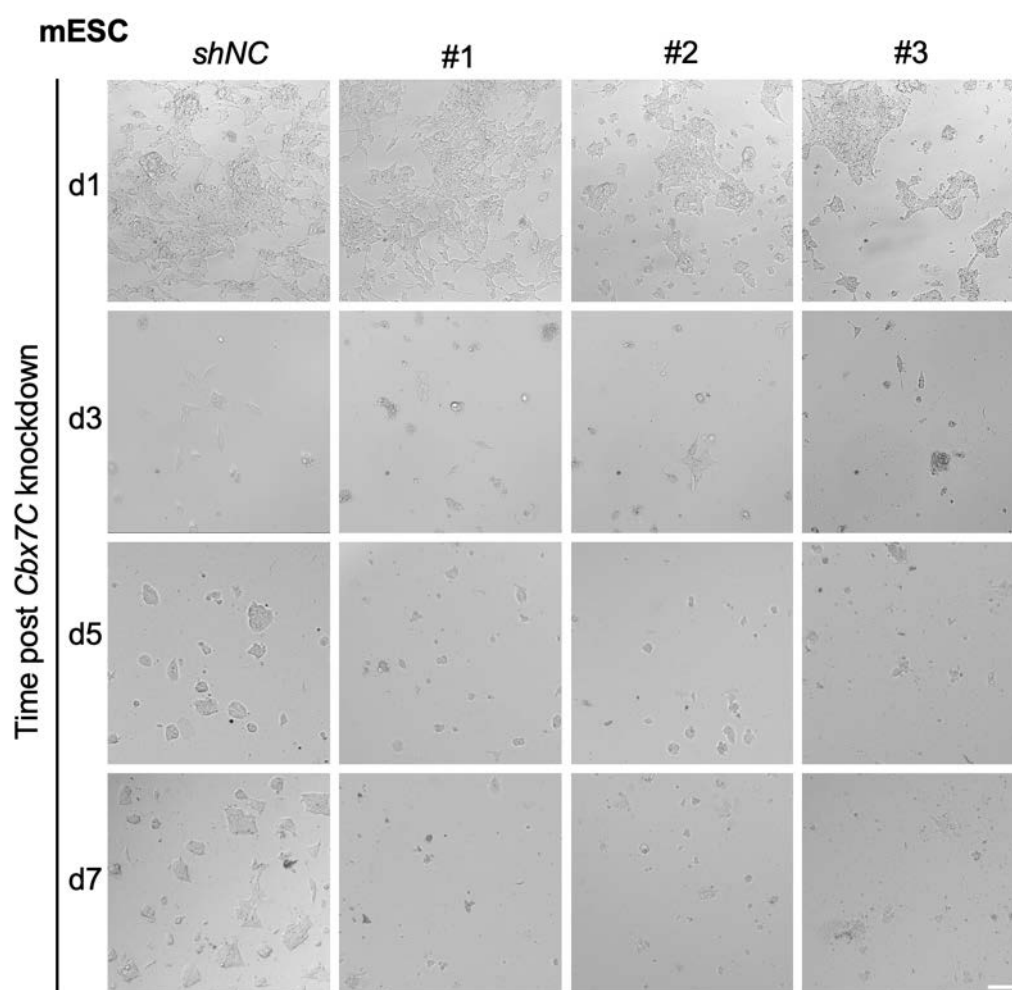


图 3-30 CBX7C 敲低的 mESC 细胞形态及增殖观察

Figure 3-30 The morphological observation of mESCs with CBX7C knockdown

CBX7C 敲低后不同时间点的 mESC 相应形态变化观察。表达 shNC 或 shCbx7C#1-3 的 mESC 在所示天数的图像显示在 20× 放大倍数。比例尺为 100 μm 。

3.7.2 CBX7C 蛋白参与转录调控、多组织分化调控等相关基因的转录调控

为了进一步探究 CBX7C 是否具有与 CBX7A 和 CBX7B 相似的抑制基因转录的功能以及 CBX7C 参与转录调控的分子网络，我们利用 RNA-Seq 对野生型及可诱导且稳定表达 mCbx7C-EGFP-Flag 的 mESC 进行转录组测序，获得相应的转录本信息。对野生型、无 Dox 诱导和 Dox 诱导的细胞系进行差异表达基因的筛选发现有 1377 和 1606 个基因表达发生变化，其中 Dox 诱导的细胞系相对于野生型细胞有 781 个基因的转录表达下调，而相对于无 Dox 诱导的细胞系则有 562 个基因的转录表达下调（图 3-31A）。我们进一步对两组基因表达下调的数据进行集合分析，筛选出一个共有集合，包含了 243 个在两组数据中均下调表达的基因集合。在这些基因位点上 CBX7C 表现出了经典的转录调控抑制能力。

同时，我们对这一共有集中表达下调的基因进行了 GO（Gene Ontology）富集（KEGG）分析，结果显示这些差异基因主要参与“转录调控”、“多组织分化调控”等细胞功能（图 3-31B），该结果与 CBX7-cPRC1 所涉及的部分经典生物学功能相一致。其次，有部分基因在 CBX7C-EGFP-Flag 诱导表达后转录水平上调，暗示 CBX7C 可能也参与基因激活的表观遗传调控。

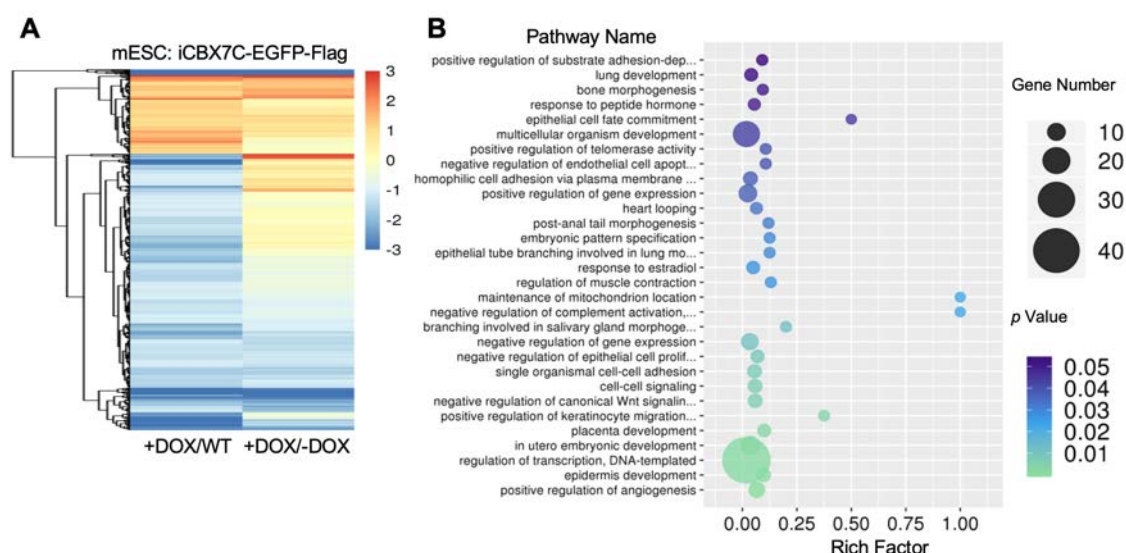


图 3-31 CBX7C 的诱导表达对 mESC 中全基因组表达模式的影响

Figure 3-31 Effect of induced CBX7C on the whole genome expression pattern in mESC

A: CBX7C 诱导表达前后，mESC 中差异表达基因聚类热图；B: CBX7C 诱导表达前后，mESC 中差异表达基因 GO 富集分析散点图。

此外，我们利用 qRT-PCR 验证了相关候选基因的表达，结果显示 qRT-PCR 数据与 RNA-Seq 数据相一致（图 3-32A），并且结合图 3-6 以及图 3-32B 的 ChIP-qPCR 实验结果进一步分析这些靶基因是否为 CBX7C 直接靶向的下游目的基因。结果显示，所示的大部分基因均有不同程度的 CBX7C 富集。此外，我们还发现，CBX7C 能够靶向的 *Cdh22* 基因能够由 CBX7C 的诱导表达而上调，结合 ChIP-Seq 相关数据库分析结果，该基因座上作为转录抑制标志的 H3K27me3 的富集度较低，而作为转录激活标志的 H3K4me3 的富集度相对较高。同时，该处也检测到 RING1B 的富集。因此，我们推测 CBX7C 可能也参与部分基因的转录激活调控。以上结果表明 CBX7C 蛋白具有转录调控的功能并通过表观修饰引起靶基因表达差异，但是其具体的分子机制仍需进一步深入研究。

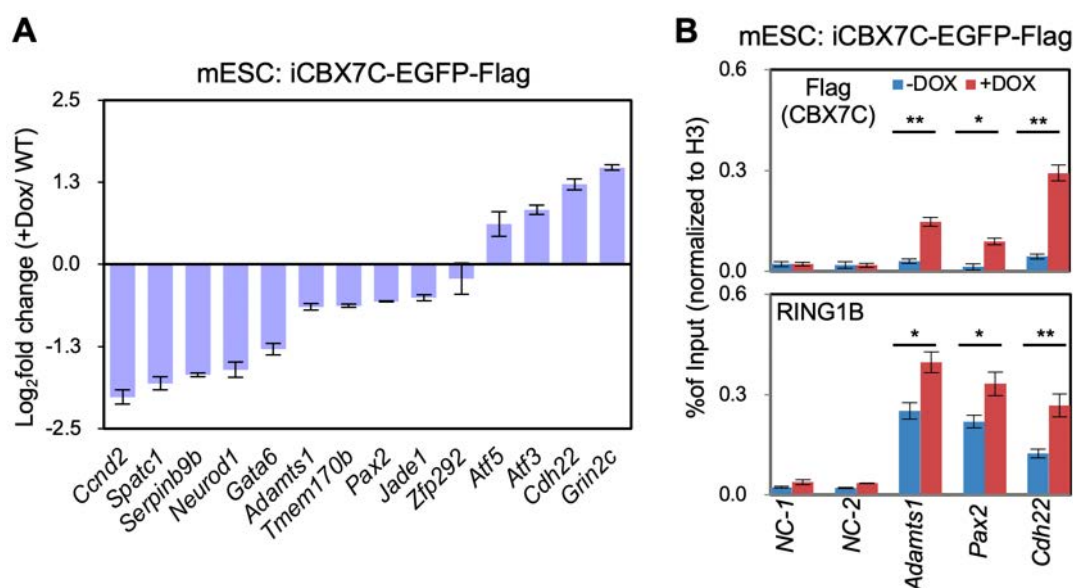


图 3-32 CBX7C 对下游靶基因的调控及结合的验证

Figure 3-32 Verification of CBX7C regulation and binding of downstream target genes

A: 通过 RT-qPCR 验证 CBX7C 诱导表达前后图示基因的表达差异; B: ChIP-qPCR 分析 CBX7C-EGFP-Flag 在图示差异表达基因位点上的富集。NC-1 (*Myc*) 和 NC-2 (*Klf4*) 是的阴性对照组基因。通过 *t* 检验统计平均值之间的显著性差异 ($p>0.05$; $*p<0.05$; $**p<0.01$)。

3.7.3 在 mESC 诱导分化为 EB 过程中内源性 CBX7C 蛋白表达水平显著上调

在 mESC 中, 表达水平相对较低的 CBX7C 在细胞分化过程中是否发挥潜在的生物学功能? 据此, 我们首先检测了通过诱导 mESC 分化形成 EB 的过程中 CBX7A、CBX7C 及在 mESC 中同样表达水平较低的 PHC2 的转录及蛋白表达水平 (图 3-33)。结果显示, 与已报道的研究结果一致, 随着 mESC 的分化, *Cbx7A* 的转录水平呈现逐渐下调的变化趋势, 而 *Phc2* 的转录水平则逐渐上调。然而, 同为 *Cbx7* 转录剪接体的 *Cbx7C* 的转录水平呈现显著增加 (图 3-33A)。同时, 上述蛋白的表达水平随着 mESC 的分化也具有相类似的变化趋势。伴随着 CBX7A 在细胞分化时下调, CBX7C 在同一时间显著增加。在 EB 形成的第 6 天, CBX7C 显著开始积累, 而 CBX7A 几乎消失 (图 3-33B)。PHC2 的表达水平在早期 (第 6 天之前) 也有所增加, 但是在后期减少 (图 3-33B 和 C)。CBX7C 和 PHC2 在细胞分化过程中共同上调表达的数据表明它们可能在 mESC 分化早期阶段协同发挥作用。

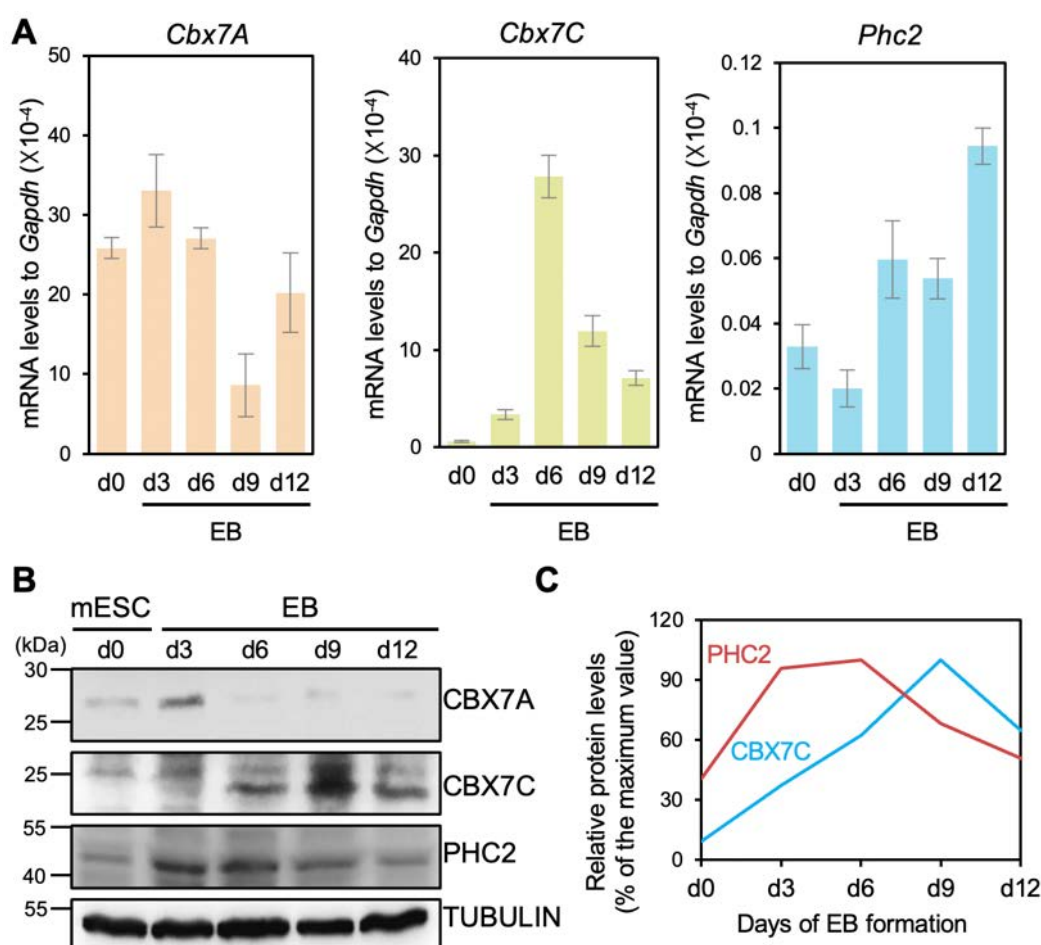


图 3-33 诱导 mESC 分化为 EBs 过程中 CBX7A、CBX7C 及 PHC2 表达水平的检测

Figure 3-33 Detection of CBX7A, 7C, and PHC2 during mESCs differentiation into EBs

A: 通过 RT-qPCR 检测 mESC 诱导分化为 EBs 的不同时间点时 *Cbx7A*、*Cbx7C* 和 *Phc2* 的转录水平; B: WB 检测诱导 mESC 诱导分化为 EBs 的不同时间点所示蛋白的表达水平; C: 曲线图表示图 B 中 CBX7C 和 PHC2 的目的条带灰度值分析。

上文所述的数据表明, CBX7C 在结合的靶基因处可能部分地取代了 CBX7A/B (图 3-6), 并且 CBX7C 可能与 CBX7A 竞争 PHC2 结合和 cPRC1 形成 (图 3-13)。为了获得直接证据, 我们通过 BiFC 竞争实验检测了 CBX7C 是否与 CBX7A 竞争性结合 PHC2。如图 3-34 所示, 在 HeLa 细胞中共转染 CBX7A 和 PHC2 的 BiFC 质粒及表达竞争者 mCherry-CBX7C 的不同梯度的质粒。通过监测 BiFC 和 mCherry 信号的变化来鉴定 CBX7A 和 CBX7C 之间对 PHC2 结合的竞争。与对照组相比, 随着 mCherry-CBX7C 表达的增加, CBX7A•PHC2-BiFC 信号逐渐消失, 这表明 CBX7C 确实能够与 CBX7A 竞争性结合 PHC2。另外, CBX7C 能够靶向到与 CBX7A/B 相同的 cPRC1 靶基因处并部分取代 CBX7A/B, 而在 mESC 诱导分化为 EB 的过程中二者表现出相反的表达调控模式, 因此 CBX7C

可能通过部分替代 CBX7A 和 CBX7B 而参与靶基因的转录调控。

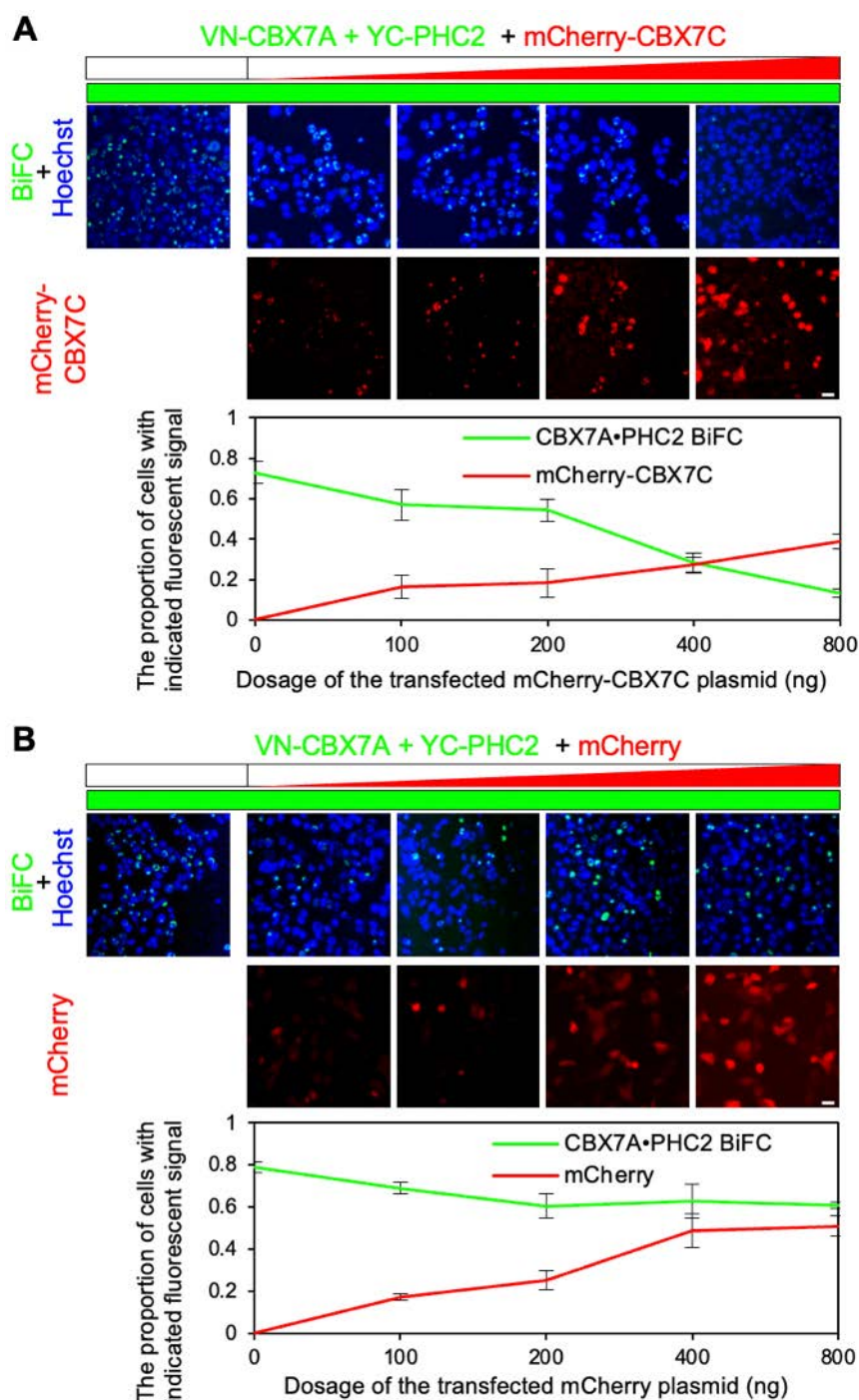


图 3-34 CBX7C 与 CBX7A 竞争性结合 PHC2 的检测

Figure 3-34 Detection of PHC2 in competitive combination of CBX7C and CBX7A

A: 通过 BiFC 竞争实验检测 HeLa 细胞中瞬时表达的 CBX7C 与 CBX7A 竞争性结合 PHC2;
 B: 用于竞争性检测的 BiFC 实验系统阴性对照。表达 mCherry-CBX7C (A) 或 mCherry (B) 质粒的转染量依次为 BiFC 质粒的 0.2、0.4、0.8 和 1.6 倍 (12 孔板的每个孔转染 400ng 所示 BiFC 质粒)。Hoechst 信号指示细胞核。比例尺为 200 μ m。下方的曲线图表示通过 ImageJ 分析图示具有荧光的细胞与 Hoechst 信号阳性细胞数的比例。

3.7.4 CBX7C 是 mESC 分化过程中类胚体 EB 形成所需的促进因子

为深入探究 CBX7C 在 mESC 分化过程中的潜在生物学功能，我们通过过表达或 RNAi 调节 *Cbx7C* 在 mESC 中的表达水平，并测定其对 EB 形成的影响。结果表明，CBX7C 的高表达促进了 mESC 产生 EB 的过程，与野生型 mESC 相比，CBX7C 过表达后形成的 EB 普遍生长更快，形态也相比更加圆润光滑（图 3-35A）。相反，CBX7C 被敲低的 mESC 在诱导形成 EB 的过程中分化细胞团直接附着在培养皿底部并发生了异常分化（图 3-35B）。上述数据表明 CBX7C 可能是 mESC 诱导分化形成 EB 过程中所需的促进因子。

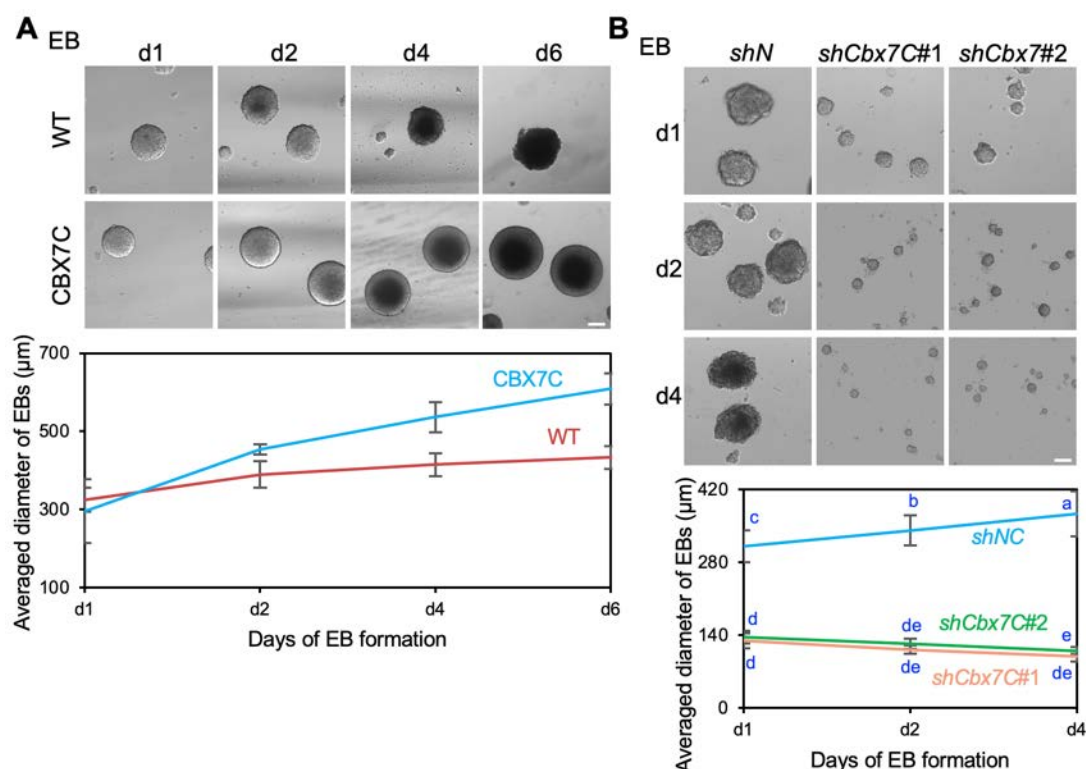


图 3-35 CBX7C 的表达水平对 mESC 分化形成 EBs 的影响

Figure 3-35 The effect of the CBX7C expression on the differentiation of mESCs into EBs

A: 稳定表达 CBX7C 或野生型 mESC 分化为 EBs 的成像分析（上）；B: CBX7C 敲低的 mESC 分化为 EBs 的成像分析（上）。比例尺为 200 μm。曲线图表示通过 ImageJ 测量形成 EBs 的平均直径（B 和 C 下）。曲线图中不同的小写字母表示 ANOVA 分析中 Duncan 检验的显著性差异 ($p < 0.05$)。

小结

在 mESC 诱导分化为 EB 的过程中，随着 CBX7A 的表达下调，CBX7C 在同一时间显著增加。与此同时，PHC2 也在此阶段有所增加，二者在 ESC 分化的早期阶段共同发挥作用。此外，CBX7C 作为促进因子在 mESC 分化过程中影响了 EB 形成的状态及生长速率。

第二节 CBX7 与 CBX2 蛋白互作并协同参与靶基因的表现遗传抑制调控

3.8 CBX7 与 CBX2 蛋白存在相互作用

研究表明, CBX7 在 ESCs 中作为主要的 CBX 蛋白存在形式占主导地位, 除了抑制某些分化发育相关的基因外, 还直接参与了对其它 CBX 蛋白编码基因的转录抑制 (包括 *Cbx2*, 4 和 8) [136, 142]。随着 ESC 的分化, *Cbx7* 基因表达下调, *Cbx2*, 4 和 8 表达上调, 靶基因位点上结合的 CBX7 逐渐被其它 CBX 蛋白所取代[142]。另外, 已报道的亚细胞定位实验结果显示不同的 CBX 蛋白在 HEK293T 细胞中无共定位的现象[229]。相关蛋白质组学分析结果也表明 cPRC1 蛋白复合体中的某些同源 CBX 蛋白组分之间相互排除, 不共存于同一复合体内。然而, CBX2 分别与 CBX4 和 CBX8 之间仍然存在相对较弱的亲和性[230], 并且 CBX4 蛋白自身之间能够直接相互作用而形成寡聚体[230]。本论文前期 ChIP-qPCR 的检测结果显示, 在 mESC 中, CBX7 并不是唯一存在的 CBX 蛋白, 此时 CBX2 也有相对较高的表达水平, 并且 CBX2 与 CBX7 能够共同募集到 PRC1 的靶基因上。据此, 我们推测 CBX7 与 CBX2 之间可能存在相互作用。

基于上述分析, 我们将 BiFC 与流式细胞分析技术相结合, 对 CBX7 与其它 CBX 蛋白之间的相互作用进行了定量分析。首先, 由于 RING1B 的 C 端(221-336 aa)包含了能够与 CBX7 互作的 RAWUL 结构域[33], 因此在此实验体系中将 VN-RING1B(1-220 aa)与 YC-CBX7 作为阴性对照, VN-RING1B(221-336 aa)与 YC-CBX7 作为阳性对照, 并且将表达 CFP 的质粒与 BiFC 相关质粒同时转染 HEK293T 细胞作为参照, 校正各组的 BiFC 互作荧光信号。如图 3-34 所示, 产生 CBX2 与 CBX7 互作的 BiFC 荧光信号的细胞群体所占比例显著高于阴性对照和其它组 (图 3-36 (A))。此外, 比较各组 CBX 蛋白与 CBX7 的 BiFC/CFP 的值发现, CBX2 与 CBX7 的比值也是显著高于阴性对照及其它各组 (图 3-36 (B) 中红线所示)。

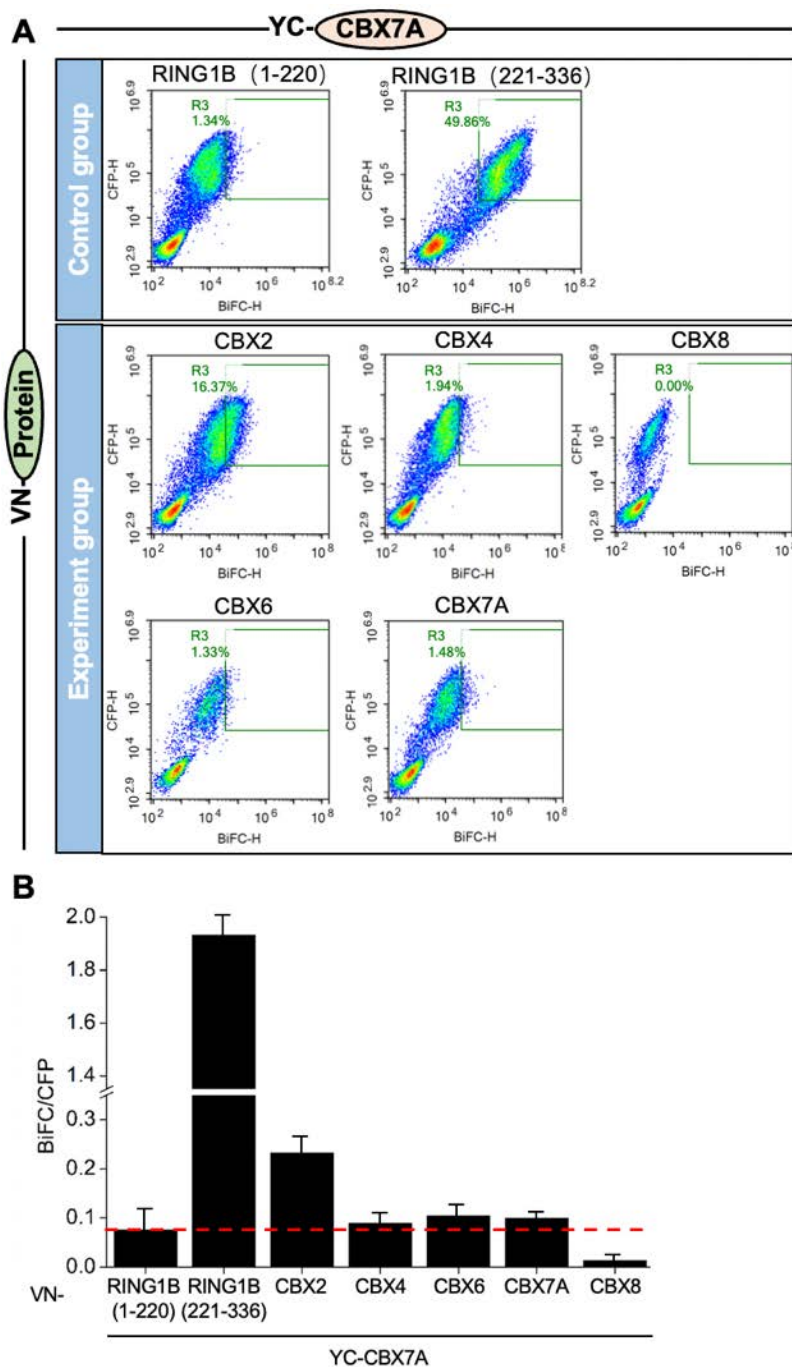


图 3-36 通过 CBX7A 与其它 CBX 蛋白互作的定量分析

Figure 3-36 Quantitative analysis of the interaction between CBX7 and other CBX proteins

A: 通过 BiFC 实验结合流式细胞分析技术定量分析 HEK293T 细胞中 CBX7A 与所示其它 CBX 蛋白及 cPRC1 组分的互作关系; B: 柱状图表示 BiFC 互作信号阳性的细胞数目与 CFP 荧光信号阳性 (内参) 的细胞数目的比值。

接下来,我们通过共聚焦成像观察到 CBX7 与 CBX2 的互作信号同 CBX7 与 RING1B 一样均匀分布于细胞核中, 并且形成若干可见的颗粒状 Pc 小体 (图 3-37 (A))。与此同时, 我们还利用 co-IP 实验进一步证实上述相互作用, 结果显

示, 在稳定表达 CBX7-EGFP-Flag 的 HEK293T 细胞中瞬时表达 Flag-CBX2 蛋白, 二者能够发生免疫共沉淀 (图 3-37 (B))。综上所述, 在不同的 CBX 蛋白之间, CBX7 仅与 CBX2 具有相互作用, 这可能也是 CBX2 与 CBX7 能够共同募集到 PRC1 靶基因上的分子基础。

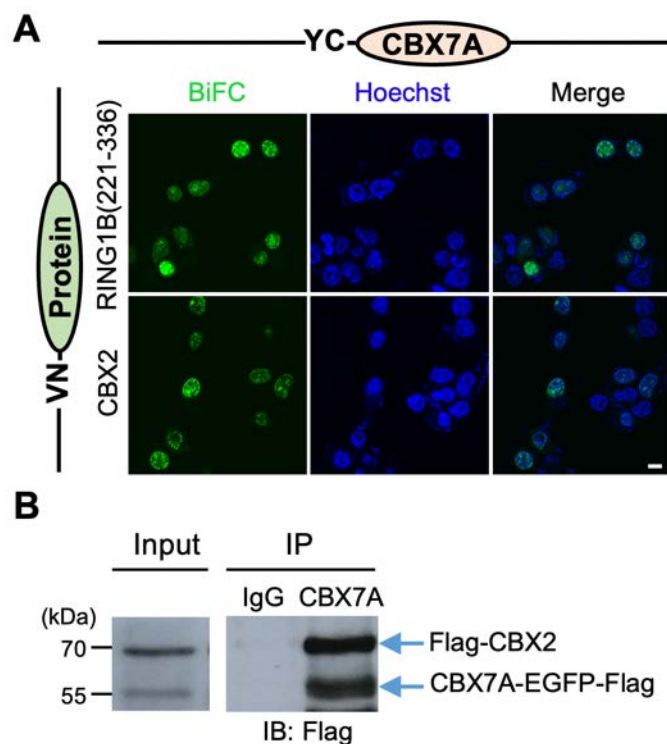


图 3-37 CBX7 与 CBX2 的互作分析

Figure 3-37 Analysis of the interaction between CBX7 and CBX2

A: 通过 BiFC 实验结合共聚焦成像分析 CBX7 与 CBX2 的互作。CBX7 与 RING1B (221-336) 作为阳性对照。蛋白及 cPRC1 组分的互作关系。比例尺为 10 μm; **B:** 通过 co-IP 实验分析在 HEK293T 细胞中瞬时表达的 Flag-CBX2 与 CBX7A-EGFP-Flag 的互作关系。

3.9 ATH 或 ATHL 结构域介导了 CBX7 与 CBX2 蛋白的相互作用

蛋白质互作位点或区域的识别对于揭示蛋白质功能机制有着重要的意义。因此, 我们通过对 CBX7 和 CBX2 的结构域进行不同组合的截短表达, 探究介导二者互作的区域。首先, CBX7 包含有三个保守的结构域: CD、PC 和 ATHL 结构域。我们利用 BiFC 及流式细胞分析技术检测了上述三个结构域对互作的影响 (图 3-38)。结果显示, 尽管 CD 和 PC 结构域在介导 CBX7 与染色质结合及与 RING1B 互作中发挥关键作用, 但是它们的缺失并未影响 CBX7 与 CBX2 的互作。然而, 当缺失了 ATHL 结构域后, CBX7 完全丧失了与 CBX2 互作的能力, 即 ATHL 是

CBX7 与 CBX2 相互作用的关键区域。

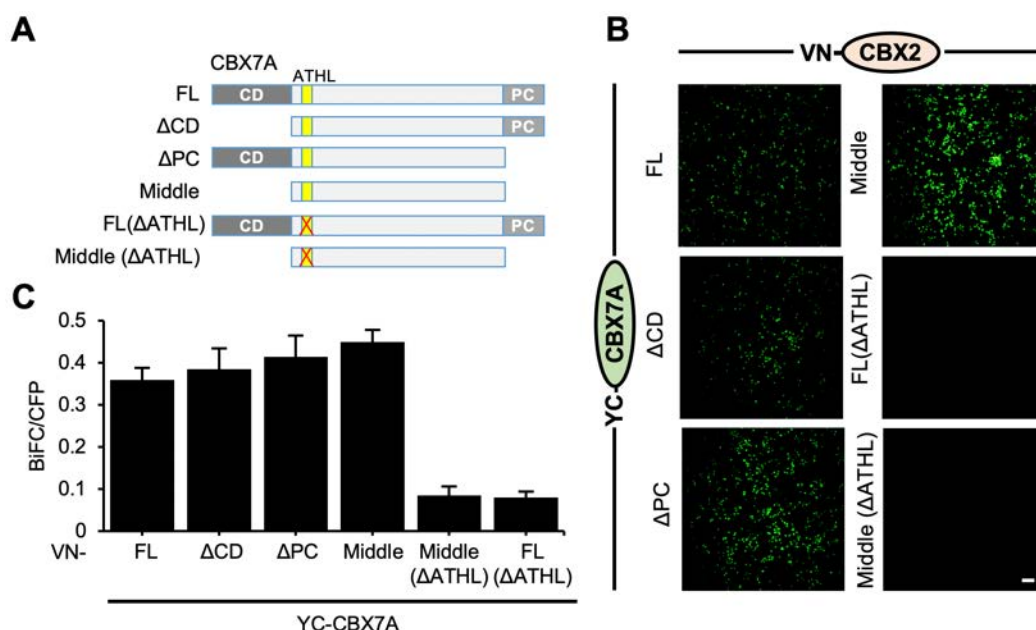


图 3-38 CBX7 及其截短体与 CBX2 互作的定量分析

Figure 3-38 Quantitative analysis of the interaction between CBX7 and truncation with CBX2

A: CBX7 全长及截短体区域示意图。比例尺为 200 μm; B: 通过 BiFC 实验分析 HEK293T 细胞中 CBX7A 及其截短体与 CBX2 的互作关系; C: 通过 BiFC 实验结合流式细胞分析技术定量分析 CBX7A 及其截短体与 CBX2 的互作。柱状图表示 BiFC 互作信号阳性的细胞数目与 CFP 荧光信号阳性 (内参) 的细胞数目的比值。

与此同时, 我们还将 CBX2 也进行了不同组合的截短表达。如图 3-39 所示, 与 CBX7 相同, CBX2 的 CD 和 PC 结构域并不是介导二者互作的区域 (图 3-39A)。与 CBX7 不同的是, 邻近 CBX2 的 CD 结构域的区域有 1 个 ATH 和 2 个 ATHL (ATHL1 和 ATHL2) 结构域。将缺失 CD 和 PC 结构域的中间区域 (Middle) 截短为包含 ATH 或 ATHL 的 Middle1 与剩余的 Middle2 (图 3-39B(a))。结果表明, 仍然是 CBX2 的 ATH 或 ATHL 结构域在二者的互作中发挥关键作用 (图 3-39B(b))。此外, 在与 CBX7 互作的检测中, ATH、ATHL1 与 ATHL2 对 CBX2 与 CBX7 互作产生的影响不同。图 3-39C 结果显示, ATHL1 与 ATHL2 结构域的缺失并未阻断 CBX2 与 CBX7 的互作, 二者仍然能够在细胞内产生 BiFC 互作荧光信号。但是当缺失 ATH 后, CBX2 几乎完全丧失了与 CBX7 的互作能力。以上结果表明, ATH 结构域是参与 CBX2 与 CBX7 互作的关键区域。

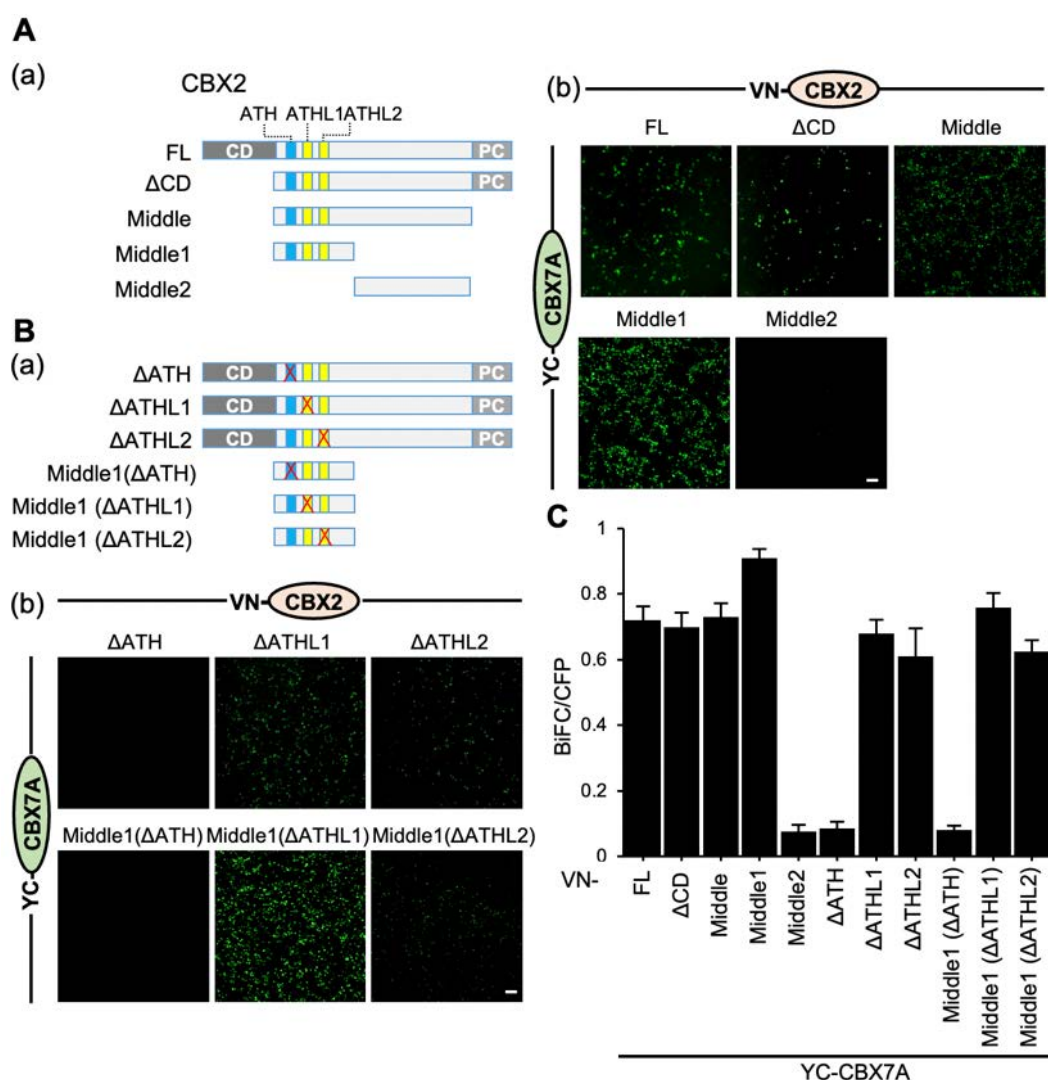


图 3-39 CBX7 与 CBX2 及其截短体互作的定量分析

Figure 3-39 Quantitative analysis of the interaction between CBX2 and truncation with CBX7
A-B: CBX2 全长及截短体区域示意图 (a)。通过 BiFC 实验分析 HEK293T 细胞中 CBX2 及其截短体与 CBX7 的互作关系 (b)。比例尺为 200 μ m; C: 通过 BiFC 实验结合流式细胞分析技术定量分析 CBX7A 与 CBX2 及其截短体的互作。柱状图表示 BiFC 互作信号阳性的细胞数目与 CFP 荧光信号阳性 (内参) 的细胞数目的比值。

此外，我们利用 GST-pull down 实验进一步证实了 CBX7 的 ATHL 或 CBX2 的 ATH 介导了 CBX2 与 CBX7 的直接相互作用。如图 3-40 所示，在不同的 CBX7 和 CBX2 的截短体组合中，仅检测到了 CBX7 与包含了 ATH 结构域的 CBX2 截短体 Middle1 的在体外的直接相互作用，而当缺失了 CBX7 的 ATHL 或 CBX2 的 ATH 结构域后，二者无法再体外发生直接的互作。综上所述，CBX2 的 ATH 结构域与 CBX7 的 ATHL 结构域介导了 CBX2 与 CBX7 的直接相互作用。

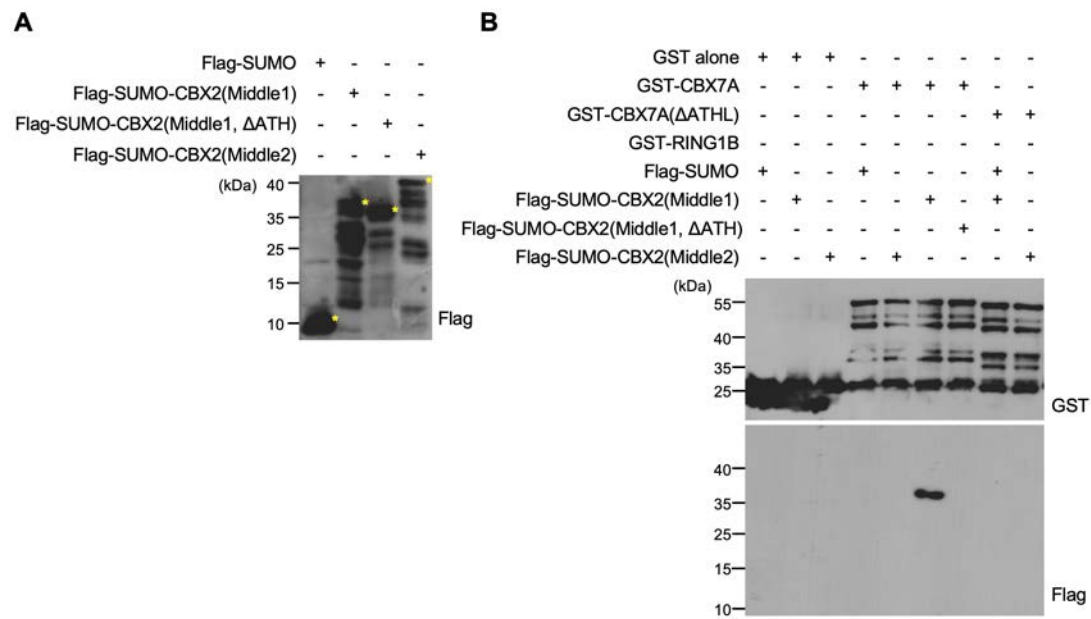


图 3-40 CBX2 与 CBX7 互作结构域的鉴定

Figure3-40 Identification of the domains mediating the interaction between CBX2 and CBX7
A: 通过 WB 检测与 GST 融合蛋白孵育前的图示原核表达蛋白; B: GST-pull down 及 WB 检测原核表达蛋白 GST-CBX7A 及其截断体与 Flag-SUMO-CBX2 及其截短体的互作。“*”指示目的蛋白条带。

3.10 ATH 或 ATHL 结构域的缺失降低了 CBX7 或 CBX2 蛋白在靶基因上的富集

为了探究 CBX7 和 CBX2 的蛋白相互作用的生物学意义, 我们通过缺失 CBX2 的 ATH 结构域而影响 CBX7 与 CBX2 的蛋白互作, 进而比较因蛋白互作的破坏所引起的 CBX7 和 CBX2 对共同靶基因的识别并结合能力的差异。如图 3-41 所示, ATH 的缺失显著降低了 CBX2 自身结合靶基因的活性。并且, 与对照组相比, CBX7 丧失了与 CBX2 的蛋白互作后其结合靶基因的能力也显著减弱。上述结果表明, 在 mESC 中, CBX2 与 CBX7 以蛋白互作的方式共同结合至靶基因而发挥特定的生物学功能。

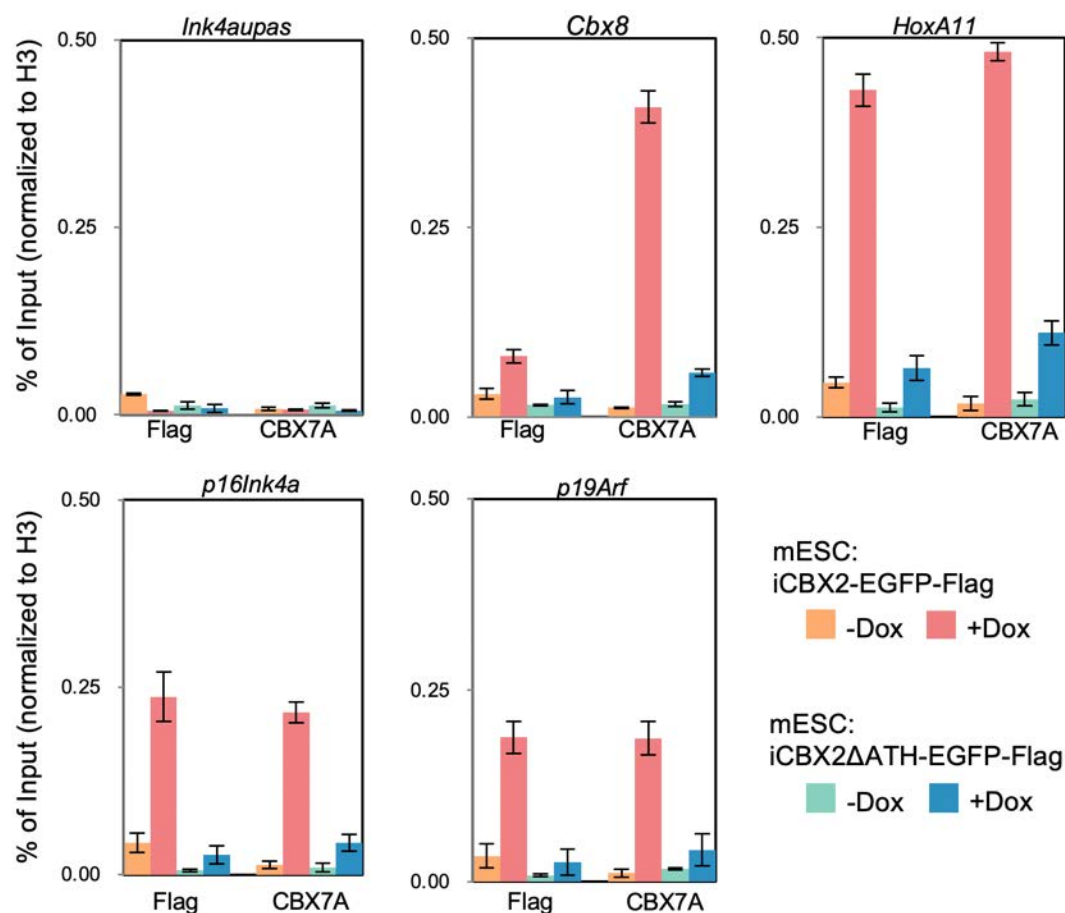


图 3-41 CBX7 与 CBX2 的互作对靶基因结合能力的影响

Figure 3-41 The effect of CBX7 and CBX2 interaction on the binding ability of target genes
在可诱导表达 CBX2-EGFP-Flag 或 CBX2ΔATH-EGFP-Flag 的 mESC 稳转细胞系中 (Dox 诱导), ChIP-qPCR 分析互作能力丧失后 CBX2 和 CBX7 对图示的共同靶基因结合活性的影响。Ink4aupas 是阴性对照组基因。

小结

在 mESC 中共存的 CBX2 和 CBX7C 通过 ATH 或 ATHL 结构域而发生相互作用, 并以此为分子基础共同结合至靶基因而发挥生物性功能。

第四章 讨论与展望

在细胞命运决定中，多梳蛋白家族（PcG）作为重要的表观遗传调控蛋白，通过复杂而精细的调控网络参与调控了诸多发育、信号传导及癌症发生等生物学和病理学过程相关的基因表达沉默。PcG 主要包含了能够直接作用于染色质的 PRC1 和 PRC2 两类复合体，它们能够抑制关键的发育相关基因，并在整个发育过程中维持这种抑制状态。其中，CBX 蛋白作为 cPRC1 的核心组分之一参与了对 H3K27me3 的识别与结合，并驱动 cPRC1 其它核心组分的招募。在靶基因座位上完成组装的 cPRC1 蛋白复合体通过压缩染色质结构和催化形成 H2AK119ub 修饰的途径促进了兼性异染色质的形成，并以此维持了 PRC1 靶基因的沉默状态^[26]。

4.1 CBX 蛋白亚型具有结构保守性与功能多样性

在物种进化过程中，PcG 蛋白家族通过压缩染色质结构或隔离沉默基因，在染色质水平上发挥多梳蛋白复合体的作用。从果蝇到脊椎动物，PRC1 和 PRC2 复合体的组装、作用机制及潜在的生物学功能都是相对保守的。然而，通过基因复制和同源基因的产生等途径增加了 PcG 蛋白家族的复杂性和多样性，多种同源蛋白的存在为其潜在的组合多样性提供了巨大的空间。在特定的基因位点上，不同组分组合的复合体产生了功能的复杂性。此外，由于转录过程中的可变剪接而产生不同的蛋白亚型也是增加复合体多样性与复杂性的途径之一。但是这些不同亚型的特有功能仍然是未知的。例如，PHC2 的长短亚型都是 PRC1 复合体的一部分^[231]，而多梳样蛋白 PCL3 的长短亚型则是不同 EZH2-PRC2 复合体的组成部分^[232]。

本论文中，我们比较分析了三种鼠源 CBX7 蛋白亚型的功能作用及分子生物学机制。哺乳动物中表达的 CBX7 是五种 Pc 同源基因之一，它包含有三个保守的结构域：介导 H3K27me3 识别并结合的 N 端 CD 结构域；位于 C 端的 PC 结构域，它是参与与 Sce 的同源蛋白 RING1B 直接相互作用的关键区域；结合双链 DNA 小沟中富含 AT 碱基区域的 ATHL 结构域，其功能目前未知。基于选择性磷酸化多腺苷化事件的调控，鼠源的 CBX7 蛋白有三种不同的亚型：分别为由含 6 个外显子转录本编码 251 个氨基酸的 CBX7A 和 158 个氨基酸的 CBX7B 蛋白亚型，这两个亚型均包含保守的 CD、ATHL 和 PC 结构域，另一个是由含 4 个外显子转录本编码 166 个氨基酸的缺少 PC 结构域的 CBX7C 蛋白亚型。近年研究

表明,在斑马鱼的 *Cbx* 基因位点也会发生这种基于选择性剪接聚腺苷酸化事件的调节而导致产生缺失 PC 结构域的 *Cbx* 蛋白亚型,例如 *cbx4* 和 *cbx6a* 基因^[123]以及人源的 *Cbx2* 基因。在斑马鱼发育过程中, *Cbx4* 或 *Cbx6* 的由于可变剪接产生的不同蛋白亚型表现出共表达模式;相比之下, *Phc2* 的两种亚型 *Phc2a1* 和 *Phc2a2* 则表现出此消彼长的替代表达模式^[134]。本研究中,我们发现鼠源 CBX7 的不同蛋白亚型在 mESC 及不同组织中共表达,但是其表达水平存在差异。然而,在 mESC 的分化过程中,低水平的 CBX7C 逐渐上调表达,与 CBX7A 蛋白亚型表现出表达替代模式。在进化过程中, *Cbx* 基因座上发生选择性磷酸化多聚腺苷酸化事件的保守性突显了这些缺乏 PC 结构域的 CBX 蛋白亚型可能具有重要的生物学功能。例如斑马鱼 CBX4 和 CBX6^[123]及人源 CBX2^[57]的缺失 PC 结构域的蛋白亚型维持一定的表达水平,说明这类缺失 PC 结构域但是含有特殊 C 端区域的蛋白亚型在生命活动中具有不可或缺的生物意义。尽管我们利用位于人源 CBX7 的第四个内含子起始区域的多对引物未能从人脾脏组织(因为小鼠 *Cbx7C* 在脾脏组织中表达)提取的 cDNA 样本中检测到人源缺失 PC 结构域的 CBX7 蛋白亚型的潜在表达,但是由于很难预测和测试所有的替代剪接变体及其表达场景,我们不能排除人源 *Cbx7* 基因的其它内含子和外显子交界位置是否会发生类似的选择性剪接聚腺苷酸化事件而缺失 PC 结构域的 CBX7 蛋白亚型的可能性。本论文关于鼠源 CBX7C 蛋白亚型的复合体组装机制、调控细胞分化及其相分离特性的研究结果对于哺乳动物中潜在表达的类似缺失 PC 结构域的 CBX 蛋白亚型的作用机制及功能探索具有重要的借鉴意义。

关于人源 CBX2 的研究表明,由于不同的多聚腺苷酸化位点和选择性剪接事件,人源 CBX2 出现两种蛋白亚型:一种是包含了保守的 CD 和 PC 结构域的 CBX2-1 亚型,另一种是较短的缺失 PC 结构域的 CBX2-2 亚型。与 CBX2-1 相比, CBX2-2 不参与 PRC1 蛋白复合体的组装,而是在体内自聚并形成高分子量的寡聚复合物,但是其仍然能够在没有 PRC1 复合体形成的情况下介导靶基因的转录抑制^[57]。与之相比,在基因转录调控中, CBX7C 表现出截然不同的分子机制。虽然缺失了 PC 结构域,但 CBX7C 仍然能作为有效组分驱动功能性 PRC1 复合体的组装,并介导了靶基因位点 H2AK119ub 修饰的产生及转录抑制,这也意味着 PRC1 复合体可以通过不同途径完成组装。虽然在 *Cbx* 基因座上发生选择性磷酸化多聚腺苷酸化事件具有保守性,但不同的缺失 PC 结构域的 CBX 蛋白亚型在机体中的表达模式、参与调控的分子机制及生物学功能存在潜在的多样性。

4.2 CBX7 的相分离影响 PRC1 功能及细胞生长状态

为了有序的组织生化反应的进行,真核细胞中的细胞内成分在时间和空间上划分为不同的膜结合的或无膜的细胞器。而核仁、应激颗粒 SGs、Cajal 小体等都是细胞内典型的无膜结构。近年来,相分离发生的分子机理及其生化功能的研究越来越受到关注。LLPS 通常也是这些无膜结构形成的理化基础之一。在过去的几年中,相分离无膜凝集体被认为与转录激活或转录抑制密切相关。负责形成兼性异染色质相的 PcG 蛋白是否分离形成凝集体逐渐成为研究热点。早在 1998 年, Saurin AJ 等人在 U2-OS 细胞中观察到了 PcG 蛋白在染色质上的聚集,将其定义为多梳小体 PB^[233]。PcG 蛋白在细胞核中形成凝集体是其通过缩合的形式介导兼性异染色质上靶基因沉默的物理位点。然而, PcG 凝集体形成的理化原理及压缩染色质的分子机制尚不完全清楚, 仍然有待深入探究。近年来, CBX2 和 CBX4 的相分离性质被进一步报道。研究表明, CBX2-PcG 的相分离是兼性异染色质形成的基础, 并且 CBX2 能够在染色质上独立发生成核反应^[152]。此外, CBX2 和 DNA 之间的相互作用是 CBX2-PRC1 在染色质上形成凝聚体的重要条件^[218]。研究表明 CBX2 驱动相分离的能力有助于提高相关生化反应的效率和特异性,加速了 CBX2 对目标基因的搜索过程^[150, 168]。此外, 具有相分离性质的 CBX4 蛋白能够同时招募 RING1B 和 EZH2 至其驱动形成的凝集体中, 并利用其 SUMO E3 连接酶活性增强 EZH2 的 H3K27 甲基转移酶活性。CBX4 作为 PRC1 和 PRC2 之间的桥梁, 促进了二者的协同作用以维持艾滋病致病病毒 HIV-1 的潜伏状态, 而 CBX4 驱动的相分离的逆转可能是重新激活潜伏的 HIV-1 的潜在干预措施^[219]。

尽管目前相分离在 CBX 蛋白参与基因转录调控中的功能研究取得了进展, 但仍需更多的直接的实验证据以揭示 CBX 蛋白参与的相分离在活细胞中的作用机制, LLPS 如何组织核酸代谢的生化反应的分子机制仍未可知。我们在研究 CBX7 不同剪接体蛋白功能的过程中发现, CBX7A 和 CBX7C 与 PHC2 形成的互作复合物在细胞核内出现明显的相分离现象, 并且 CBX7A 和 CBX7C 自身便可在核内发生成核反应并形成凝集体。应用已开发的能够预测靶蛋白进行相分离的能力的工具或数据库 (PONDER, Predictor of Natural Disordered Regions) 对 CBX7A/C 进行分析后, 我们发现 CBX7A 和 CBX7C 包含了参与相分离的无序区域 (IDRs, intrinsically disordered protein regions)。生化分析数据表明 CBX7C 比 CBX7A 在体内和体外具有更强的发生相分离的驱动力 (图 3-10 和图 3-12)。在正常生理条件下, CBX7C 与 PHC2 及其它组分组装成功能性 PRC1 复合体并凝集形成 PB, 进一步催化靶基因位点的 H2AK119ub 并抑制其转录。当高浓度的

CBX7C 与 PHC2 蛋白共存于细胞中时, 它们往往会驱动发生液-固相转变, 产生的凝集体流动性差, 在其形成的过程中逐渐排斥了其中的染色质组分并招募了 cPRC1 的其余组分, 形成更加致密的由蛋白质组成的水凝胶状结构。在此过程中, CBX7C 与 PHC2 的互作强化了 CBX7C 驱动相分离的能力。然而, 由于 CBX7C-PHC2 互作并缩合形成的致密凝集体虽然包含了 PRC1 的各组分, 但由于空间阻隔使其丧失了酶学活性, 在外扰乱了胞内的染色质分布, 进而阻碍了细胞的正常生长, 甚至引发了细胞的凋亡。此外, 本论文的研究结果还证实了 CBX7A 的 N 端 (1-82) 或 CBX7C 蛋白的 N 端 (1-110) 是介导 CBX7 与 PHC2 互作的主要区域, 但是二者结构不同的 C 端 (83-251) 或 (111-166) 是参与 CBX7-PHC2 相分离的关键区域, 这也解释了 CBX7A 与 CBX7C 均能与 PHC2 互作的原因, 但是二者驱动相分离的能力却有所不同。在体外环境中, Elizabeth S. Jaensch 等人的结果表明纯化的 CBX7 蛋白可能无法发生相分离而形成混浊溶液。在此基础上, 我们继续增加了 CBX7 纯化蛋白的浓度后, 溶液的混浊度显著增加, 并形成了典型的具有相分离特征的凝集体。由于结构差异, 在上述过程中 CBX7 的不同蛋白亚型表现出差异, 即 CBX7C 比 CBX7A 对相分离的形成具有更强的驱动力。此外, 由于蛋白在体内或体外的高浓度表达会使得凝集体突破临界点, 部分地由液态转变被胶状的固态, 并且在细胞内强烈地排斥了染色质, 破坏了原本的微环境。总之, 上述研究结果为一些 *Cbx7* 异常高表达的疾病模型研究提供了重要信息。但 CBX7 相分离的性质及其不同亚型差异的生物学功能及病理性意义尚不完全清楚, 有待进一步深入研究。

4.3 CBX7C 的蛋白表达及其相分离影响 mESC 的生长状态及分化

在果蝇和哺乳动物中, CBX 蛋白是从早期发育到成体阶段过程中参与基因表观遗传控制的关键蛋白之一。Polycomb 在脊椎动物中的同源蛋白 CBX2、4、6、7 和 8 具有不同的基序组合, 在发育过程中, 它们表现出交替的时空表达模式^[136]。在哺乳动物 ESC 的 PRC1 复合物中, CBX7 是占主导地位的 Polycomb 同源蛋白, CBX7 蛋白的表达能够抑制 ESC 的分化并促进其自我更新。在此过程中, CBX7 能够靶向至 *Cbx2*、*Cbx4* 和 *Cbx8* 基因位点并抑制其转录。随着 ESCs 的分化, *Cbx7* 的表达下调, 而 *Cbx2*、*Cbx4* 和 *Cbx8* 会逐渐上调。Ana O'Loughlen 等人的研究显示虽然 CBX7 在 ESCs 中表达, 并在分化过程中急剧下调, 但是在此短暂的过程中, 其它 CBX 的表达未发生显著变化, 而是在分化后逐渐增加。在 ESC 分化过程中, miR-125 和 miR-181 家族是 *Cbx7* 的重要调节因子。miR-

125b 和 miR-181a/b 主要在 ESC 分化过程中诱导产生, 并通过直接靶向 *Cbx7* 的 3'UTR 区而介导 *Cbx7* 的转录下调^[142]。我们的研究表明, 在 ESCs 诱导形成 EBs 的过程中, *Cbx7C* 转录本和其编码的蛋白水平急剧增加, 这一变化与 *Cbx7A* 和 *Cbx7B* 的表达模式是相反的, 这意味着 *Cbx7c* 可能逃逸了 miR-125 和 miR-181 家族的抑制。与此同时, 我们还发现在 ESCs 中, CBX7C 在靶基因位点上的富集在一定程度上排斥了 CBX7A 与 CBX7B 的结合。同时, 在亚细胞共定位的检测结果中也显示出相同的“替代效应”, 即 CBX7C-PRC1 凝集体中 CBX7A 和 CBX7B 被排斥在外, 而 CBX7C 同样具有结合 PRC1 靶基因并抑制其转录的功能。同时, 我们发现 CBX7A 和 CBX7B 的急剧下调并未引起其靶基因表达水平的显著变化。我们推测可能存在某种转换机制, 即在 ESCs 中 *Cbx7A* 与 *Cbx7B* 是作为主导来调控基因表达; 随着分化的进行, 增加的 CBX7C 可能逐渐替代了被 miR-125 和 miR-181 家族抑制的 CBX7A 和 CBX7B, 进而参与对靶基因的抑制和状态维持。此外, 在 mESC 诱导分化为 EB 的过程中, 随着 CBX7C 的表达显著增加, 与其直接互作的 PHC2 也在此阶段有所增加, 二者在 ESC 分化的早期阶段共同发挥作用。CBX7C 作为促进因子在 mESC 分化过程中影响了 EB 形成的状态及生长速率。

在后续研究中, 我们还需深入探究在分化过程中, 急剧增加的 CBX7C 是否通过驱动相分离, 改变染色质的空间位置, 进而抑制不同靶基因群的转录, 通过更加精准和高效的方法以揭示 ESCs 细胞的自我更新和分化中 *Cbx7* 的不同剪接体的协同或拮抗关系及它们表达模式变化的机制和生物学功能。在此基础上, 利用转基因小鼠 (CBX7C 过表达及敲除或敲低小鼠) 深入探究这种存在时空表达特异性的 CBX7C 蛋白亚型在小鼠胚胎发育及脾脏功能中是否也具有类似的功能及影响。此外, 在目前获得的敲低 CBX7C 导致 mESC 状态失稳的现象的基础上, 我们还需在后续研究中深入探究 CBX7C 影响 mESC 状态稳定性的潜在功能及其机制。

基于以上实验证据, 我们得出一个模式图来阐明 CBX7C 和 7A 在 mESC 和分化过程中参与 PRC1 组装和相分离的工作模型 (图 4-1)。在该模型中, PHC2 可代替 RING1B 作为 CBX7C 参与复合体组装的桥接分子, CBX7C•PHC2 二聚体作为核心结构促进 cPRC1 在染色质上的原位组装。此外, 含不同蛋白亚型组分的 cPRC1 复合物聚集在一起, 通过 LLPS 形成 Pc 小体, 以实现更有效的转录共抑制。因此, CBX7C 作为有效的 cPRC1 组装的驱动因子发挥重要作用, 以驱动表观遗传学抑制功能的建立, 而不是独立调节因子参与基因表达调控。在 CBX7C-cPRC1 复合体组装过程中, CBX7C 可以作为识别靶基因并驱动 cPRC1 组装的引发剂, 而 PHC2 的加入进一步增强了 Pc 小体的形成。这表明, 在进化

过程中, cPRC1 组分不同剪接异构体的出现可能会大大增加 PRC1 的多样性和复杂的分裂。此外, CBX7C•PHC2 凝集体的固有性质因其表达剂量而异。CBX7C•PHC2 在正常生理水平下形成具有催化活性的 Pc 小体, 而当表达剂量高于阈值时, CBX7C•PHC2 凝集体趋于发生液-固相分离, 并汇聚成致密的水凝胶状结构, 同时排除染色质。与此同时, 致密的凝集体像分子海绵一样不断地从周围的核内环境中吸收 cPRC1 组分蛋白。然而, 由于组蛋白修饰酶在空间上被包裹在这些凝集体中, 与其酶促底物发生空间隔离, 因此即使包含所有复合体组分, 所形成的 Pc 小体也几乎发生功能失调。CBX7C 的 N 末端主要介导 CBX7C•PHC2 的二聚体形成, 并足以在没有 SAM 的情况下通过 LLPS 驱动具有酶活性的 Pc 小体的构建。然而, CBX7C 的 C 末端主要在促进具有水凝胶性质的 CBX7C•PHC2 凝集体的形成中发挥作用, 这可能与疾病的发展有关。综上所述, CBX7C•PHC2 互作的构象可能在定义相关结构域方面发挥重要作用, 从而决定相分离的选择。

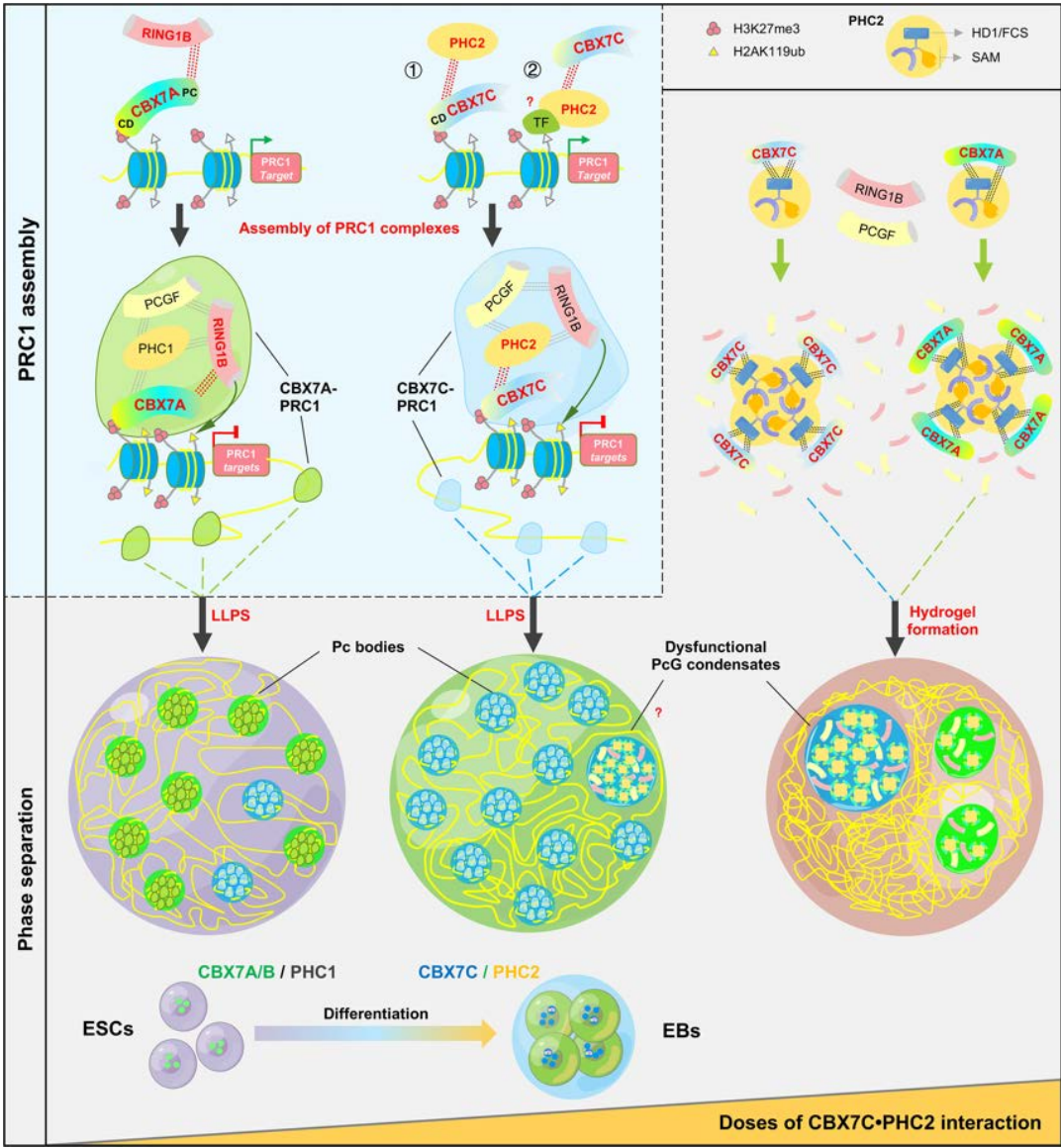


图 4-1 CBX7C 和 7A 在 mESC 和分化过程中参与 PRC1 组装和相分离的工作模型
Figure 5-1 A working model of CBX7C and 7A in participating PRC1 assembly and phase separation in mESCs and during differentiation

第五章 结论

在多细胞生物个体发育过程中,复杂而精确的基因表达调控网络决定了不同细胞类型中特定的基因表达谱。基因表达的表观遗传学调节是调节细胞命运决定和发育过程中许多重要细胞生物学过程的关键机制之一。多梳蛋白家族 PcG (包括 PRC1 和 PRC2) 在机体中控制数百个参与发育的基因的表达。在生物进化过程中, PRC1 的不同亚基在脊椎动物中都衍生出多个同源物, 以及在基因转录过程中的选择性剪接多腺苷酸化事件引起单个基因产生不同的异构体, 为 PRC1 复合体潜在的组合多样性提供了巨大的空间。由此引发的功能及作用机制多样性的产生有待深入探究。

在本论文中, 我们讨论了鼠源 CBX7 的三种不同蛋白亚型 CBX7A、CBX7B 和以及缺失保守的 PC 结构域的 CBX7C 的结构、相分离特质及参与基因调控的分子机制和生物学功能。结合多种生化及分子生物学方法及显微成像分析, 我们发现了缺失 PC 结构域的 CBX7C 仍然能够作为 cPRC1 的功能亚基而参与基因转录抑制调控, 并在 mESC 分化过程中表达上调, 可能替代了在该过程中下调的 CBX7A 和 CBX7B。与 CBX7A 相比, CBX7C 在体内和体外具有更强的驱动相分离的能力。此外, 在 mESC 中共表达且能够共靶向到相同靶基因上的 CBX2 和 CBX7A 之间能够发生直接互作, 并且 ATH-CBX2 与 ATHL-CBX7 (CBXA、B 和 C 共有结构域) 共同介导了二者互作的发生。

本论文得到的主要结论如下:

(1) 鼠源 *Cbx7* 基因产生的三种不同的转录本及其编码的蛋白在 mESC 及小鼠组织中表现出不同的表达模式, 并且我们关注的 CBX7C 主要在 mESC 和小鼠脾脏组织中表达。尽管由于缺失了 PC 结构域而丧失了与 RING1B 的直接相互作用, 但是与 CBX7A/B 不同的是 CBX7C 能够通过与 PHC2 之间的直接互作在细胞中与复合体其它核心组分结合并驱动功能性的 cPRC1 复合体的组装, 催化靶向基因位点的 H2AK119ub 修饰

(2) 含有 IDRs 和 LCD 的 CBX7A 和 CBX7C 在体内和体外均能够发生相分离, 并且 CBX7C 表现出比 CBX7A 更强的驱动相分离的能力。此外, CBX7A/C 与 PHC2 的互作能够进一步强化 CBX7A/C 驱动相分离的能力, 形成更加致密的由蛋白质组成的水凝胶状结构。在此过程中, CBX7A 的 N 端 (1-82) 或 CBX7C 蛋白的 N 端 (1-110) 是介导 CBX7A/C 与 PHC2 互作的主要区域, 而其 C 端 (83-251) 或 (111-166) 是参与 CBX7A/C-PHC2 相分离的关键区域。另外, PHC2 在此互作并相分离的过程中起主导作用, 并且其 SAM 结构域的完整性显著影响相

相互作用蛋白的相分离特性;

(3) 缺失了 PC 结构域的 CBX7C 在其 N 端的 CD 和 ATHL 结构域共同介导下仍然保留了与染色质组蛋白的互作能力。在 mESC 中, 与 CBX7A/B 一样, CBX7C 也能够识别并结合至已知的 cPRC1 靶基因位点, 进而导致其转录被抑制。并且 CBX7C 的诱导表达引起了涉及转录调控、多组织分化调控等发育相关的生物学过程的基因群的表达差异。此外, 在 mESC 诱导分化为 EBs 的过程中 *Cbx7c* 的转录及编码蛋白表达水平显著上调, 可能通过部分替代 CBX7A 和 CBX7B 而参与靶基因的转录调控, 在干细胞分化进程中 CBX7-PRC1 与其他 CBX 蛋白-PRC1 发生功能转换时起到衔接作用。

(4) 在 mESC 中共表达且能够共靶向到相同靶基因上的 CBX2 和 CBX7A 之间能够发生直接互作, 并且 ATH-CBX2 与 ATHL-CBX7 (CBXA、B 和 C 共有结构域) 共同介导了二者互作的发生。

本论文的创新点:

(1) 本论文首次发现 *Cbx7C* 是一种在小鼠 ESCs 中特异性表达的新型 *Cbx7* 剪接异构体。

(2) 本研究确定了 CBX7C 通过 PHC2 募集 RING1B, 并参与组装功能性 PRC1。

(3) 本研究首次确定了 CBX7C 的相分离特性, 并发现 PHC2 显著提高了 CBX7C 的相分离性能。

(4) 本研究发现 CBX7C 在 mESC 分化过程中上调并在 mESC 分化过程中发挥重要作用。

对 *Cbx7C* 新剪接体的深入研究不仅为阐明 CBX 蛋白的相分离特性在胚胎发育等生理过程中的特殊生物学功能提供了理论依据, 而且有助于解释 CBX 失稳后导致相关致病的分子机制。本项目不仅具有重要的学术价值, 而且蕴藏着潜在的应用价值, 研究结果为未来相关疾病的早期诊断、治疗和预后检测提供新的有效靶点。

参考文献

- [1] Helin K, Dhanak D. Chromatin proteins and modifications as drug targets[J]. *Nature*, 2013, **502**: 480-488.
- [2] Klemm S L, Shipony Z, Greenleaf W J. Chromatin accessibility and the regulatory epigenome[J]. *Nature Reviews Genetics*, 2019, **20**: 207-220.
- [3] Lewis E B. A gene complex controlling segmentation in *Drosophila*[J]. *Nature*, 1978, **276**: 565-70.
- [4] Ingham P W. Differential Expression of Bithorax Complex Genes in the Absence of the Extra Sex Combs and Trithorax Genes[J]. *Nature*, 1983, **306**: 591-593.
- [5] Ingham P W. Genetic-Control of the Spatial Pattern of Selector Gene-Expression in *Drosophila*[J]. *Cold Spring Harbor Symposia on Quantitative Biology*, 1985, **50**: 201-208.
- [6] Struhl G, Akam M. Altered Distributions of Ultrabithorax Transcripts in Extra-Sex-Combs Mutant Embryos of *Drosophila*[J]. *Embo Journal*, 1985, **4**: 3259-3264.
- [7] Kennison J A, Tamkun J W. Dosage-Dependent Modifiers of Polycomb and Antennapedia Mutations in *Drosophila*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1988, **85**: 8136-8140.
- [8] Muller J, Verrijzer P. Biochemical mechanisms of gene regulation by polycomb group protein complexes[J]. *Current Opinion in Genetics & Development*, 2009, **19**: 150-158.
- [9] Richly H, Aloia L, Di Croce L. Roles of the Polycomb group proteins in stem cells and cancer[J]. *Cell Death & Disease*, 2011, **2**: e204.
- [10] Aloia L, Di Stefano B, Di Croce L. Polycomb complexes in stem cells and embryonic development[J]. *Development*, 2013, **140**: 2525-2534.
- [11] Schuettengruber B, Chourrout D, Vervoort M, *et al.* Genome regulation by polycomb and trithorax proteins[J]. *Cell*, 2007, **128**: 735-745.
- [12] Schuettengruber B, Bourbon H M, Di Croce L, *et al.* Genome Regulation by Polycomb and Trithorax: 70 Years and Counting[J]. *Cell*, 2017, **171**: 34-57.
- [13] Schuettengruber B, Martinez A M, Iovino N, *et al.* Trithorax group proteins: switching genes on and keeping them active[J]. *Nature Reviews Molecular Cell Biology*, 2011, **12**: 799-814.
- [14] Aranda S, Mas G, Di Croce L. Regulation of gene transcription by Polycomb proteins[J]. *Science Advances*, 2015, **1**: e1500737.
- [15] Schwartz Y B, Pirrotta V. Ruled by Ubiquitylation: A New Order for Polycomb Recruitment[J]. *Cell Reports*, 2014, **8**: 321-325.
- [16] Morey L, Helin K. Polycomb group protein-mediated repression of transcription[J]. *Trends in Biochemical Sciences*, 2010, **35**: 323-332.
- [17] Isono K, Fujimura Y, Shinga J, *et al.* Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes[J]. *Molecular and Cellular Biology*, 2005, **25**: 6694-6706.
- [18] Buchwald G, Van Der Stoop P, Weichenrieder O, *et al.* Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b[J]. *Embo Journal*, 2006, **25**: 2465-2474.
- [19] Li Z Z, Cao R, Wang M, *et al.* Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex[J]. *Journal of Biological Chemistry*, 2006, **281**: 20643-20649.
- [20] De Napoles M, Mermoud J E, Wakao R, *et al.* Polycomb group proteins Ring1A/B link ubiquitylation

- of histone H2A to heritable gene silencing and X inactivation[J]. *Developmental Cell*, 2004, **7**: 663-676.
- [21] Francis N J, Kingston R E, Woodcock C L. Chromatin compaction by a polycomb group protein complex[J]. *Science*, 2004, **306**: 1574-1577.
- [22] Guenther M G, Levine S S, Boyer L A, *et al.* A chromatin landmark and transcription initiation at most promoters in human cells[J]. *Cell*, 2007, **130**: 77-88.
- [23] Zeitlinger J, Stark A, Kellis M, *et al.* RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo[J]. *Nature Genetics*, 2007, **39**: 1512-1516.
- [24] Wang H B, Wang L J, Erdjument-Bromage H, *et al.* Role of histone H2A ubiquitination in polycomb silencing[J]. *Nature*, 2004, **431**: 873-878.
- [25] Stock J K, Giadrossi S, Casanova M, *et al.* Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells[J]. *Nature Cell Biology*, 2007, **9**: 1428-U174.
- [26] Gao Z H, Zhang J, Bonasio R, *et al.* PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes[J]. *Molecular Cell*, 2012, **45**: 344-356.
- [27] Chittock E C, Latwiel S, Miller T C R, *et al.* Molecular architecture of polycomb repressive complexes[J]. *Biochemical Society Transactions*, 2017, **45**: 193-205.
- [28] Cao R, Wang L J, Wang H B, *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing[J]. *Science*, 2002, **298**: 1039-1043.
- [29] Min J R, Zhang Y, Xu R M. Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys 27[J]. *Genes & Development*, 2003, **17**: 1823-1828.
- [30] Kaustov L, Hui O Y, Amaya M, *et al.* Recognition and Specificity Determinants of the Human Cbx Chromodomains[J]. *Journal of Biological Chemistry*, 2011, **286**: 521-529.
- [31] Fischle W, Wang Y, Jacobs S A, *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains[J]. *Genes Dev*, 2003, **17**: 1870-81.
- [32] Bernstein E, Duncan E M, Masui O, *et al.* Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin[J]. *Molecular and Cellular Biology*, 2006, **26**: 2560-9.
- [33] Sanchez-Pulido L, Devos D, Sung Z R, *et al.* RAWUL: A new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs[J]. *Bmc Genomics*, 2008, **9**: 308.
- [34] Bezsonova I, Walker J R, Bacik J P, *et al.* Ring1B contains a ubiquitin-like docking module for interaction with Cbx proteins[J]. *Biochemistry*, 2009, **48**: 10542-8.
- [35] Wang R J, Taylor A B, Leal B Z, *et al.* Polycomb Group Targeting through Different Binding Partners of RING1B C-Terminal Domain[J]. *Structure*, 2010, **18**: 966-975.
- [36] Czypionka A, De Los Panos O R, Mateu M G, *et al.* The isolated C-terminal domain of Ring1B is a dimer made of stable, well-structured monomers[J]. *Biochemistry*, 2007, **46**: 12764-76.
- [37] Wang R, Ilangovan U, Robinson A K, *et al.* Structural transitions of the RING1B C-terminal region upon binding the polycomb cbox domain[J]. *Biochemistry*, 2008, **47**: 8007-15.
- [38] Deshaies R J, Joazeiro C A. RING domain E3 ubiquitin ligases[J]. *Annual Review of Biochemistry*, 2009, **78**: 399-434.
- [39] Metzger M B, Pruneda J N, Klevit R E, *et al.* RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination[J]. *Biochim Biophys Acta*, 2014, **1843**: 47-60.
- [40] Cao R, Tsukada Y, Zhang Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing[J]. *Molecular Cell*, 2005, **20**: 845-854.
- [41] Yadav A K, Sahasrabudhe A A, Dimri M, *et al.* Deletion analysis of BMI1 oncoprotein identifies its negative regulatory domain[J]. *Molecular Cancer*, 2010, **9**: 158.

- [42] Ben-Saadon R, Zaaroor D, Ziv T, *et al.* The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity[J]. *Molecular Cell*, 2006, **24**: 701-711.
- [43] Alkema M J, Bronk M, Verhoeven E, *et al.* Identification of Bmi-1-interacting proteins as constituents of a multimeric mammalian Polycomb complex[J]. *Genes & Development*, 1997, **11**: 226-240.
- [44] Gunster M J, Satijn D P E, Hamer K M, *et al.* Identification and characterization of interactions between the vertebrate polycomb-group protein BMI1 and human homologs of polyhomeotic[J]. *Molecular and Cellular Biology*, 1997, **17**: 2326-2335.
- [45] Gray F, Cho H J, Shukla S, *et al.* BMI1 regulates PRC1 architecture and activity through homo- and hetero-oligomerization[J]. *Nature Communications*, 2016, **7**: 13343.
- [46] Cohen K J, Hanna J S, Prescott J E, *et al.* Transformation by the Bmi-1 oncoprotein correlates with its subnuclear localization but not its transcriptional suppression activity[J]. *Molecular and Cellular Biology*, 1996, **16**: 5527-35.
- [47] Dimri G P, Martinez J L, Jacobs J J L, *et al.* The Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells[J]. *Cancer Research*, 2002, **62**: 4736-4745.
- [48] Itahana K, Zou Y, Itahana Y, *et al.* Control of the replicative life span of human fibroblasts by p16 and the polycomb protein Bmi-1[J]. *Molecular and Cellular Biology*, 2003, **23**: 389-401.
- [49] Isono K, Endo T A, Ku M C, *et al.* SAM Domain Polymerization Links Subnuclear Clustering of PRC1 to Gene Silencing[J]. *Developmental Cell*, 2013, **26**: 565-577.
- [50] Colombo M, Pessey O, Marcia M. Topology and enzymatic properties of a canonical Polycomb repressive complex 1 isoform[J]. *Febs Letters*, 2019, **593**: 1837-1848.
- [51] Kim C A, Gingery M, Pilpa R M, *et al.* The SAM domain of polyhomeotic forms a helical polymer[J]. *Nature Structural Biology*, 2002, **9**: 453-457.
- [52] Kim C A, Bowie J U. SAM domains: uniform structure, diversity of function[J]. *Trends in Biochemical Sciences*, 2003, **28**: 625-628.
- [53] Gambetta M C, Muller J. O-GlcNAcylation Prevents Aggregation of the Polycomb Group Repressor Polyhomeotic[J]. *Developmental Cell*, 2014, **31**: 629-639.
- [54] Tsuboi M, Kishi Y, Yokozeki W, *et al.* Ubiquitination-Independent Repression of PRC1 Targets during Neuronal Fate Restriction in the Developing Mouse Neocortex[J]. *Developmental Cell*, 2018, **47**: 758-772.
- [55] Cheutin T, Cavalli G. Progressive Polycomb Assembly on H3K27me3 Compartments Generates Polycomb Bodies with Developmentally Regulated Motion[J]. *Plos Genetics*, 2012, **8**: e1002465.
- [56] Wani A H, Boettiger A N, Schorderet P, *et al.* Chromatin topology is coupled to Polycomb group protein subnuclear organization[J]. *Nature Communications*, 2016, **7**: 10291.
- [57] Volkel P, Le Faou P, Vandamme J, *et al.* A human Polycomb isoform lacking the Pc box does not participate to PRC1 complexes but forms protein assemblies and represses transcription[J]. *Epigenetics*, 2012, **7**: 482-91.
- [58] Fauvarque M O, Dura J M. Polyhomeotic Regulatory Sequences Induce Developmental Regulator-Dependent Variegation and Targeted P-Element Insertions in *Drosophila*[J]. *Genes & Development*, 1993, **7**: 1508-1520.
- [59] Muller J, Bienz M. Long-Range Repression Conferring Boundaries of Ultrabithorax Expression in the *Drosophila* Embryo[J]. *Embo Journal*, 1991, **10**: 3147-3155.
- [60] Simon J, Chiang A, Bender W, *et al.* Elements of the *Drosophila* Bithorax Complex That Mediate Repression by Polycomb Group Products[J]. *Developmental Biology*, 1993, **158**: 131-144.

- [61] Entrevan M, Schuettengruber B, Cavalli G. Regulation of Genome Architecture and Function by Polycomb Proteins[J]. Trends in Cell Biology, 2016, **26**: 511-525.
- [62] Wang L, Brown J L, Cao R, *et al.* Hierarchical recruitment of polycomb group silencing complexes[J]. Molecular Cell, 2004, **14**: 637-46.
- [63] Kahn T G, Stenberg P, Pirrotta V, *et al.* Combinatorial Interactions Are Required for the Efficient Recruitment of Pho Repressive Complex (PhoRC) to Polycomb Response Elements[J]. Plos Genetics, 2014, **10**: e1004495.
- [64] Schuettengruber B, Elkayam N O, Sexton T, *et al.* Cooperativity, Specificity, and Evolutionary Stability of Polycomb Targeting in Drosophila[J]. Cell Reports, 2014, **9**: 219-233.
- [65] Kahn T G, Dorafshan E, Schultheis D, *et al.* Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements[J]. Nucleic Acids Research, 2016, **44**: 10132-10149.
- [66] Frey F, Sheahan T, Finkl K, *et al.* Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC[J]. Genes & Development, 2016, **30**: 1116-1127.
- [67] Boyer L A, Plath K, Zeitlinger J, *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells[J]. Nature, 2006, **441**: 349-353.
- [68] Tanay A, O'donnell A H, Damelin M, *et al.* Hyperconserved CpG domains underlie Polycomb-binding sites[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, **104**: 5521-5526.
- [69] Mendenhall E M, Koche R P, Truong T, *et al.* GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells[J]. Plos Genetics, 2010, **6**: e1001244.
- [70] Riising E M, Comet I, Leblanc B, *et al.* Gene Silencing Triggers Polycomb Repressive Complex 2 Recruitment to CpG Islands Genome Wide[J]. Molecular Cell, 2014, **55**: 347-360.
- [71] Farcas A M, Blackledge N P, Sudbery I, *et al.* KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands[J]. Elife, 2012, **1**: e00205.
- [72] Bernstein B E, Mikkelsen T S, Xie X H, *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells[J]. Cell, 2006, **125**: 315-326.
- [73] Ku M, Koche R P, Rheinbay E, *et al.* Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains[J]. Plos Genetics, 2008, **4**: e1000242.
- [74] Klose R J, Cooper S, Farcas A M, *et al.* Chromatin Sampling-An Emerging Perspective on Targeting Polycomb Repressor Proteins[J]. Plos Genetics, 2013, **9**: e1003717.
- [75] Plath K, Fang J, Mlynarczyk-Evans S K, *et al.* Role of histone H3 lysine 27 methylation in X inactivation[J]. Science, 2003, **300**: 131-135.
- [76] Davidovich C, Cech T R. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2[J]. Rna, 2015, **21**: 2007-2022.
- [77] Blackledge N P, Farcas A M, Kondo T, *et al.* Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation[J]. Cell, 2014, **157**: 1445-1459.
- [78] Cooper S, Dienstbier M, Hassan R, *et al.* Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment[J]. Cell Reports, 2014, **7**: 1456-1470.
- [79] Almeida M, Pintacuda G, Masui O, *et al.* PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation[J]. Science, 2017, **356**: 1081-1084.
- [80] Kalb R, Latwiel S, Baymaz H I, *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression[J]. Nature Structural & Molecular Biology, 2014, **21**: 569-571.
- [81] Cooper S, Grijzenhout A, Underwood E, *et al.* Jarid2 binds mono-ubiquitylated H2A lysine 119 to

- mediate crosstalk between Polycomb complexes PRC1 and PRC2[J]. *Nature Communications*, 2016, **7**: 13661.
- [82] Lee W, Teckie S, Wiesner T, *et al.* PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors[J]. *Nature Genetics*, 2014, **46**: 1227-1232.
- [83] Kang H, McElroy K A, Jung Y L, *et al.* Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in *Drosophila*[J]. *Genes & Development*, 2015, **29**: 1136-1150.
- [84] Musselman C A, Gibson M D, Hartwick E W, *et al.* Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility[J]. *Nature Communications*, 2013, **4**: 2969.
- [85] Ballare C, Lange M, Lapinaite A, *et al.* Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity[J]. *Nature Structural & Molecular Biology*, 2012, **19**: 1257-65.
- [86] Margueron R, Justin N, Ohno K, *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks[J]. *Nature*, 2009, **461**: 762-U11.
- [87] Mozzetta C, Pontis J, Fritsch L, *et al.* The Histone H3 Lysine 9 Methyltransferases G9a and GLP Regulate Polycomb Repressive Complex 2-Mediated Gene Silencing[J]. *Molecular Cell*, 2014, **53**: 277-289.
- [88] McCabe M T, Graves A P, Ganji G, *et al.* Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27)[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**: 2989-2994.
- [89] Loubiere V, Delest A, Thomas A, *et al.* Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development[J]. *Nature Genetics*, 2016, **48**: 1436-1442.
- [90] Schaaf C A, Misulovin Z, Gause M, *et al.* Cohesin and Polycomb Proteins Functionally Interact to Control Transcription at Silenced and Active Genes[J]. *Plos Genetics*, 2013, **9**: e1003560.
- [91] Schwartz Y B, Kahn T G, Nix D A, *et al.* Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*[J]. *Nature Genetics*, 2006, **38**: 700-705.
- [92] Glancy E, Wang C, Tuck E, *et al.* PRC2.1- and PRC2.2-specific accessory proteins drive recruitment of different forms of canonical PRC1[J]. *Molecular Cell*, 2023, **83**: 1393-1411 e7.
- [93] Wotton D, Merrill J C. Pc2 and SUMOylation[J]. *Biochemical Society Transactions*, 2007, **35**: 1401-1404.
- [94] Satijn D P E, Olson D J, Vandervlag J, *et al.* Interference with the expression of a novel human polycomb protein, hPc2, results in cellular transformation and apoptosis[J]. *Molecular and Cellular Biology*, 1997, **17**: 6076-6086.
- [95] Senthilkumar R, Mishra R K. Novel motifs distinguish multiple homologues of Polycomb in vertebrates: expansion and diversification of the epigenetic toolkit[J]. *Bmc Genomics*, 2009, **10**: 549.
- [96] Shin D, Kim G, Kim G, *et al.* Structural and biochemical characterizations of an intramolecular tandem coiled coil protein[J]. *Biochemical and Biophysical Research Communications*, 2014, **455**: 339-346.
- [97] Aravind L, Landsman D. AT-hook motifs identified in a wide variety of DNA binding proteins[J]. *Nucleic Acids Research*, 1998, **26**: 4413-4421.
- [98] Ma R G, Zhang Y, Sun T T, *et al.* Epigenetic regulation by polycomb group complexes: focus on roles of CBX proteins[J]. *Journal of Zhejiang University-Science B*, 2014, **15**: 412-428.
- [99] Bernstein E, Duncan E M, Masui O, *et al.* Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin[J]. *Molecular and Cellular Biology*, 2006, **26**: 2560-2569.

- [100] Vincenz C, Kerppola T K. Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences[J]. Proceedings of the National Academy of Sciences of the United States of America, 2008, **105**: 16572-16577.
- [101] Pemberton H, Anderton E, Patel H, *et al.* Genome-wide co-localization of Polycomb orthologs and their effects on gene expression in human fibroblasts[J]. Genome Biology, 2014, **15**: R23.
- [102] Klauke K, Radulovic V, Broekhuis M, *et al.* Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation[J]. Nature Cell Biology, 2013, **15**: 353-62.
- [103] Dietrich N, Bracken A P, Trinh E, *et al.* Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus[J]. Embo Journal, 2007, **26**: 1637-1648.
- [104] Connelly K E, Dykhuizen E C. Compositional and functional diversity of canonical PRC1 complexes in mammals[J]. Biochim Biophys Acta Gene Regul Mech, 2017, **1860**: 233-245.
- [105] Khalil A M, Guttman M, Huarte M, *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression[J]. Proc Natl Acad Sci U S A, 2009, **106**: 11667-72.
- [106] Popov N, Gil J. Epigenetic regulation of the INK4b-ARF-INK4a locus In sickness and in health[J]. Epigenetics, 2010, **5**: 685-690.
- [107] Yang L Q, Lin C R, Liu W, *et al.* ncRNA- and Pc2 Methylation-Dependent Gene Relocation between Nuclear Structures Mediates Gene Activation Programs[J]. Cell, 2011, **147**: 773-788.
- [108] Breiling A, Bonte E, Ferrari S, *et al.* The Drosophila Polycomb protein interacts with nucleosomal core particles in vitro via its repression domain[J]. Molecular and Cellular Biology, 1999, **19**: 8451-8460.
- [109] Muller J. Transcriptional Silencing by the Polycomb Protein in Drosophila Embryos[J]. Embo Journal, 1995, **14**: 1209-1220.
- [110] Bunker C A, Kingston R E. Transcriptional Repression by Drosophila and Mammalian Polycomb Group Proteins in Transfected Mammalian-Cells[J]. Molecular and Cellular Biology, 1994, **14**: 1721-1732.
- [111] Schoorlemmer J, Marcosgutierrez C, Were F, *et al.* Ring1A is a transcriptional repressor that interacts with the Polycomb-M33 protein and is expressed at rhombomere boundaries in the mouse hindbrain[J]. Embo Journal, 1997, **16**: 5930-5942.
- [112] Sowpati D T, Ramamoorthy S, Mishra R K. Expansion of the polycomb system and evolution of complexity[J]. Mechanisms of Development, 2015, **138**: 97-112.
- [113] Messmer S, Franke A, Paro R. Analysis of the Functional-Role of the Polycomb Chromo Domain in Drosophila-Melanogaster[J]. Genes & Development, 1992, **6**: 1241-1254.
- [114] Reeves R, Nissen M S. The A.T-DNA-Binding Domain of Mammalian High Mobility Group-I Chromosomal-Proteins - a Novel Peptide Motif for Recognizing DNA-Structure[J]. Journal of Biological Chemistry, 1990, **265**: 8573-8582.
- [115] Fonfria-Subiros E, Acosta-Reyes F, Saperas N, *et al.* Crystal Structure of a Complex of DNA with One AT-Hook of HMGA1[J]. Plos One, 2012, **7**: e37120.
- [116] Tardat M, Albert M, Kunzmann R, *et al.* Cbx2 Targets PRC1 to Constitutive Heterochromatin in Mouse Zygotes in a Parent-of-Origin-Dependent Manner[J]. Molecular Cell, 2015, **58**: 157-171.
- [117] Zhen C Y, Tatavosian R, Huynh T N, *et al.* Live-cell single-molecule tracking reveals co-recognition of H3K27me3 and DNA targets polycomb Cbx7-PRC1 to chromatin[J]. Elife, 2016, **5**: e17667.
- [118] Tie F, Banerjee R, Fu C, *et al.* Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain[J]. Proceedings of the National Academy of Sciences of the United States of America,

2016, **113**: E744-E753.

[119] Pan Q, Shai O, Lee L J, *et al.* Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing[J]. *Nature Genetics*, 2008, **40**: 1413-1415.

[120] Ren C Y, Morohashi K, Plotnikov A N, *et al.* Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain[J]. *Chemistry & Biology*, 2015, **22**: 161-168.

[121] Kim M S, Pinto S M, Getnet D, *et al.* A draft map of the human proteome[J]. *Nature*, 2014, **509**: 575-81.

[122] Maertens G N, El Messaoudi-Aubert S, Racek T, *et al.* Several distinct polycomb complexes regulate and co-localize on the INK4a tumor suppressor locus[J]. *Plos One*, 2009, **4**: e6380.

[123] Le Faou P, Volkel P, Angrand P O. The zebrafish genes encoding the Polycomb repressive complex (PRC) 1[J]. *Gene*, 2011, **475**: 10-21.

[124] Gil J, Bernard D, Martínez D, *et al.* Polycomb CBX7 has a unifying role in cellular lifespan[J]. *Nature Cell Biology*, 2004, **6**: 67-U19.

[125] Scott C L, Gil J, Hernando E, *et al.* Role of the chromobox protein CBX7 in lymphomagenesis[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**: 5389-5394.

[126] Zhang X W, Zhang L, Qin W, *et al.* Oncogenic role of the chromobox protein CBX7 in gastric cancer[J]. *Journal of Experimental & Clinical Cancer Research*, 2010, **29**: 114.

[127] Bernard D, Martinez-Leal J F, Rizzo S, *et al.* CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the Ink4a/Arf locus (Retraction of Vol 24, Pg 5543, 2005)[J]. *Oncogene*, 2022, **41**: 4980-4980.

[128] Pallante P, Federico A, Berlingieri M T, *et al.* Loss of the gene expression correlates with a highly malignant phenotype in thyroid cancer[J]. *Cancer Research*, 2008, **68**: 6770-6778.

[129] Karamitopoulou E, Pallante P, Zlobec I, *et al.* Loss of the CBX7 protein expression correlates with a more aggressive phenotype in pancreatic cancer[J]. *European Journal of Cancer*, 2010, **46**: 1438-1444.

[130] Forzati F, Federico A, Pallante P, *et al.* Tumor suppressor activity of CBX7 in lung carcinogenesis[J]. *Cell Cycle*, 2012, **11**: 1888-1891.

[131] Kim H Y, Park J H, Won H Y, *et al.* CBX7 inhibits breast tumorigenicity through DKK-1-mediated suppression of the Wnt/ β -catenin pathway[J]. *Faseb Journal*, 2015, **29**: 300-313.

[132] Shinjo K, Yamashita Y, Yamamoto E, *et al.* Expression of chromobox homolog 7 (CBX7) is associated with poor prognosis in ovarian clear cell adenocarcinoma via TRAIL-induced apoptotic pathway regulation[J]. *International Journal of Cancer*, 2014, **135**: 308-318.

[133] Cho K W, Andrade M, Zhang Y, *et al.* Mammalian CBX7 isoforms p36 and p22 exhibit differential responses to serum, varying functions for proliferation, and distinct subcellular localization[J]. *Sci Rep*, 2020, **10**: 8061.

[134] Kawamura A, Yamada K, Fujimori K, *et al.* Alternative transcripts of a polyhomeotic gene homolog are expressed in distinct regions of somites during segmentation of zebrafish embryos[J]. *Biochem Biophys Res Commun*, 2002, **291**: 245-54.

[135] Montgomery N D, Yee D, Chen A, *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation[J]. *Current Biology*, 2005, **15**: 942-947.

[136] Morey L, Pascual G, Cozzuto L, *et al.* Nonoverlapping Functions of the Polycomb Group Cbx Family of Proteins in Embryonic Stem Cells[J]. *Cell Stem Cell*, 2012, **10**: 47-62.

[137] Morey L, Santanach A, Blanco E, *et al.* Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms[J]. *Cell Stem Cell*, 2015, **17**:

300-315.

- [138] Shema E, Jones D, Shores N, *et al.* Single-molecule decoding of combinatorially modified nucleosomes[J]. *Science*, 2016, **352**: 717-721.
- [139] Piunti A, Shilatifard A. Epigenetic balance of gene expression by Polycomb and COMPASS families[J]. *Science*, 2016, **352**: aad9780.
- [140] Ren X J, Vincenz C, Kerppola T K. Changes in the distributions and dynamics of Polycomb repressive complexes during embryonic stem cell differentiation[J]. *Molecular and Cellular Biology*, 2008, **28**: 2884-2895.
- [141] Sparmann A, Van Lohuizen M. Polycomb silencers control cell fate, development and cancer[J]. *Nature Reviews Cancer*, 2006, **6**: 846-856.
- [142] O'loghlen A, Munoz-Cabello A M, Gaspar-Maia A, *et al.* MicroRNA Regulation of Cbx7 Mediates a Switch of Polycomb Orthologs during ESC Differentiation[J]. *Cell Stem Cell*, 2012, **10**: 33-46.
- [143] Kloet S L, Makowski M M, Baymaz H I, *et al.* The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation[J]. *Nature Structural & Molecular Biology*, 2016, **23**: 682-690.
- [144] Ram O, Goren A, Amit I, *et al.* Combinatorial Patterning of Chromatin Regulators Uncovered by Genome-wide Location Analysis in Human Cells[J]. *Cell*, 2011, **147**: 1628-1639.
- [145] Lau M S, Schwartz M G, Kundu S, *et al.* Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning[J]. *Science*, 2017, **355**: 1081-1084.
- [146] Santanach A, Blanco E, Jiang H, *et al.* The Polycomb group protein CBX6 is an essential regulator of embryonic stem cell identity[J]. *Nature Communications*, 2017, **8**: 1235.
- [147] Grau D J, Chapman B A, Garlick J D, *et al.* Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge[J]. *Genes Dev*, 2011, **25**: 2210-21.
- [148] Core N, Bel S, Gaunt S J, *et al.* Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice[J]. *Development*, 1997, **124**: 721-9.
- [149] Katoh-Fukui Y, Tsuchiya R, Shiroishi T, *et al.* Male-to-female sex reversal in M33 mutant mice[J]. *Nature*, 1998, **393**: 688-92.
- [150] Kent S, Brown K, Yang C H, *et al.* Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging[J]. *Cell Reports*, 2020, **33**: 108248.
- [151] Eskeland R, Leeb M, Grimes G R, *et al.* Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination[J]. *Molecular Cell*, 2010, **38**: 452-464.
- [152] Tatavosian R, Kent S, Brown K, *et al.* Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation[J]. *Journal of Biological Chemistry*, 2019, **294**: 1451-1463.
- [153] Xie H F, Xu J, Hsu J H, *et al.* Polycomb Repressive Complex 2 Regulates Normal Hematopoietic Stem Cell Function in a Developmental-Stage-Specific Manner[J]. *Cell Stem Cell*, 2014, **14**: 68-80.
- [154] Camahort R, Cowan C A. Cbx Proteins Help ESCs Walk the Line between Self-Renewal and Differentiation[J]. *Cell Stem Cell*, 2012, **10**: 4-6.
- [155] Huang Z N, Yan Y L, Zhu Z, *et al.* CBX7 suppresses urinary bladder cancer progression via modulating AKR1B10-ERK signaling[J]. *Cell Death & Disease*, 2021, **12**: 537.
- [156] Wang B Q, Tang J J, Liao D, *et al.* Chromobox Homolog 4 Is Correlated with Prognosis and Tumor Cell Growth in Hepatocellular Carcinoma[J]. *Annals of Surgical Oncology*, 2013, **20**: S684-S692.
- [157] Zhang Y X, Chen H, Zhu H C, *et al.* CBX8 promotes tumorigenesis and confers radioresistance in esophageal squamous cell carcinoma cells through targeting APAF1[J]. *Gene*, 2019, **711**: 143949.
- [158] Zhang Y, Kang M, Zhang B, *et al.* m(6)A modification-mediated CBX8 induction regulates

- stemness and chemosensitivity of colon cancer via upregulation of LGR5[J]. *Molecular Cancer*, 2019, **18**: 185.
- [159] Iqbal M A, Siddiqui S, Rehman A U, *et al.* Multiomics integrative analysis reveals antagonistic roles of CBX2 and CBX7 in metabolic reprogramming of breast cancer[J]. *Molecular Oncology*, 2021, **15**: 1450-1465.
- [160] Colaprico A, Silva T C, Olsen C, *et al.* TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data[J]. *Nucleic Acids Research*, 2016, **44**: e71.
- [161] Cerami E, Gao J, Dogrusoz U, *et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data (vol 2, pg 401, 2012)[J]. *Cancer Discovery*, 2012, **2**: 960-960.
- [162] Zhang Y J, Zhao L Y, He X, *et al.* CBXs-related prognostic gene signature correlates with immune microenvironment in gastric cancer[J]. *Aging-Us*, 2022, **14**: 6227-6254.
- [163] Tian P, Zhang C, Ma C L, *et al.* Decreased chromobox homologue 7 expression is associated with epithelial-mesenchymal transition and poor prognosis in cervical cancer[J]. *Open Medicine*, 2021, **16**: 410-418.
- [164] Ueda S, Kanda M, Sato Y, *et al.* Chromobox 2 Expression Predicts Prognosis After Curative Resection of Oesophageal Squamous Cell Carcinoma[J]. *Cancer Genomics & Proteomics*, 2020, **17**: 391-400.
- [165] Li W, Chen H H, Wang Z G, *et al.* Chromobox 4 (CBX4) promotes tumor progression and stemness via activating CDC20 in gastric cancer[J]. *Journal of Gastrointestinal Oncology*, 2022, **13**: 1058-1072.
- [166] Yu T, Wu Y, Hu Q, *et al.* CBX7 is a glioma prognostic marker and induces G1/S arrest via the silencing of CCNE1[J]. *Oncotarget*, 2017, **8**: 26637-26647.
- [167] Nawaz Z, Patil V, Arora A, *et al.* Cbx7 is epigenetically silenced in glioblastoma and inhibits cell migration by targeting YAP/TAZ-dependent transcription[J]. *Scientific Reports*, 2016, **6**: 27753.
- [168] Banani S F, Lee H O, Hyman A A, *et al.* Biomolecular condensates: organizers of cellular biochemistry[J]. *Nature Reviews Molecular Cell Biology*, 2017, **18**: 285-298.
- [169] Larson A G, Narlikar G J. The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation[J]. *Biochemistry*, 2018, **57**: 2540-2548.
- [170] Shin Y, Brangwynne C P. Liquid phase condensation in cell physiology and disease[J]. *Science*, 2017, **357**: eaaf4382.
- [171] Boeynaems S, Alberti S, Fawzi N L, *et al.* Protein Phase Separation: A New Phase in Cell Biology[J]. *Trends in Cell Biology*, 2018, **28**: 420-435.
- [172] Swain P, Weber S C. Dissecting the complexity of biomolecular condensates[J]. *Biochemical Society Transactions*, 2020, **48**: 2591-2602.
- [173] Wilson E B. The Structure of Protoplasm[J]. *Science*, 1899, **10**: 33-45.
- [174] Brangwynne C P, Eckmann C R, Courson D S, *et al.* Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation[J]. *Science*, 2009, **324**: 1729-1732.
- [175] Li P L, Banjade S, Cheng H C, *et al.* Phase transitions in the assembly of multivalent signalling proteins[J]. *Nature*, 2012, **483**: 336-U129.
- [176] Kato M, Han T N W, Xie S H, *et al.* Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels[J]. *Cell*, 2012, **149**: 753-767.
- [177] Su X L, Ditlev J A, Hui E F, *et al.* Phase separation of signaling molecules promotes T cell receptor signal transduction[J]. *Science*, 2016, **352**: 595-599.
- [178] Molliex A, Temirov J, Lee J, *et al.* Phase Separation by Low Complexity Domains Promotes Stress

- Granule Assembly and Drives Pathological Fibrillization[J]. *Cell*, 2015, **163**: 123-133.
- [179] Sabari B R, Dall'agnese A, Boija A, *et al.* Coactivator condensation at super-enhancers links phase separation and gene control[J]. *Science*, 2018, **361**: eaar3958.
- [180] Strom A R, Emelyanov A V, Mir M R, *et al.* Phase Separation Drives Heterochromatin Domain Formation[J]. *Biophysical Journal*, 2018, **114**: 445a-445a.
- [181] Sheu-Gruttadauria J, Macrae I J. Phase Transitions in the Assembly and Function of Human miRISC[J]. *Cell*, 2018, **173**: 946-957.
- [182] Zhang G M, Wang Z, Du Z, *et al.* mTOR Regulates Phase Separation of PGL Granules to Modulate Their Autophagic Degradation[J]. *Cell*, 2018, **174**: 1492-1506.
- [183] Banjade S, Rosen M K. Phase Transitions of Multivalent Proteins Can Promote Clustering of Membrane Receptors[J]. *Elife*, 2014, **3**: e04123.
- [184] Brangwynne C P, Tompa P, Pappu R V. Polymer physics of intracellular phase transitions[J]. *Nature Physics*, 2015, **11**: 899-904.
- [185] Nott T J, Petsalaki E, Farber P, *et al.* Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles[J]. *Molecular Cell*, 2015, **57**: 936-947.
- [186] Gilks N, Kedersha N, Ayodele M, *et al.* Stress granule assembly is mediated by prion-like aggregation of TIA-1[J]. *Molecular Biology of the Cell*, 2004, **15**: 5383-5398.
- [187] Jiang H, Wang S S, Huang Y J, *et al.* Phase Transition of Spindle-Associated Protein Regulate Spindle Apparatus Assembly[J]. *Cell*, 2015, **163**: 108-122.
- [188] Crick S L, Jayaraman M, Frieden C, *et al.* Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**: 16764-16769.
- [189] Pak C W, Kosno M, Holehouse A S, *et al.* Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein[J]. *Molecular Cell*, 2016, **63**: 72-85.
- [190] Das R K, Pappu R V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, **110**: 13392-13397.
- [191] Quiroz F G, Chilkoti A. Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers[J]. *Nature Materials*, 2015, **14**: 1164-71.
- [192] Harmon T S, Holehouse A S, Rosen M K, *et al.* Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins[J]. *Elife*, 2017, **6**: e30294.
- [193] Franzmann T M, Alberti S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior[J]. *Journal of Biological Chemistry*, 2019, **294**: 7128-7136.
- [194] Ruff K M, Roberts S, Chilkoti A, *et al.* Advances in Understanding Stimulus Responsive Phase Behavior of Intrinsically Disordered Protein Polymers[J]. *Journal of Molecular Biology*, 2018, **430**: 4619-4635.
- [195] Widom B. Note on the Interfacial-Tension of Phase-Separated Polymer-Solutions[J]. *Journal of Statistical Physics*, 1988, **52**: 1343-1351.
- [196] Hyman A A, Weber C A, Juelicher F. Liquid-Liquid Phase Separation in Biology[J]. *Annual Review of Cell and Developmental Biology*, Vol 30, 2014, **30**: 39-58.
- [197] Murray D T, Kato M, Lin Y, *et al.* Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains[J]. *Cell*, 2017, **171**: 615-627 e16.
- [198] Riback J A, Katanski C D, Kear-Scott J L, *et al.* Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response[J]. *Cell*, 2017, **168**: 1028-1040.

- [199] Linding R, Jensen L J, Diella F, *et al.* Protein disorder prediction: Implications for structural proteomics[J]. *Structure*, 2003, **11**: 1453-1459.
- [200] Dosztanyi Z, Csizmok V, Tompa P, *et al.* The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins[J]. *Febs Journal*, 2005, **272**: 360-360.
- [201] Oates M E, Romero P, Ishida T, *et al.* (DP2)-P-2: database of disordered protein predictions[J]. *Nucleic Acids Research*, 2013, **41**: D508-D516.
- [202] Wang B, Zhang L, Dai T, *et al.* Liquid-liquid phase separation in human health and diseases[J]. *Signal Transduction and Targeted Therapy*, 2021, **6**: 290.
- [203] Asano S, Engel B D, Baumeister W. In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology[J]. *Journal of Molecular Biology*, 2016, **428**: 332-343.
- [204] Iserman C, Roden C A, Boerneke M A, *et al.* Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid[J]. *Molecular Cell*, 2020, **80**: 1078-1091 e6.
- [205] Alberti S, Gladfelter A, Mittag T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates[J]. *Cell*, 2019, **176**: 419-434.
- [206] Carlson C R, Asfaha J B, Ghent C M, *et al.* Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions[J]. *Molecular Cell*, 2020, **80**: 1092-1103.
- [207] Ishikawa-Ankerhold H C, Ankerhold R, Drummen G P C. Advanced Fluorescence Microscopy Techniques-FRAP, FLIP, FLAP, FRET and FLIM[J]. *Molecules*, 2012, **17**: 4047-4132.
- [208] Shin Y, Berry J, Pannucci N, *et al.* Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets[J]. *Cell*, 2017, **168**: 159-171 e14.
- [209] Bugaj L J, Choksi A T, Mesuda C K, *et al.* Optogenetic protein clustering and signaling activation in mammalian cells[J]. *Nature Methods*, 2013, **10**: 249-252.
- [210] Kroschwald S, Munder M C, Maharana S, *et al.* Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery[J]. *Cell Reports*, 2018, **23**: 3327-3339.
- [211] Shin Y, Chang Y C, Lee D S W, *et al.* Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome (vol 175, pg 1481, 2018)[J]. *Cell*, 2019, **176**: 1518.
- [212] Boehning M, Dugast-Darzacq C, Rankovic M, *et al.* RNA polymerase II clustering through carboxy-terminal domain phase separation[J]. *Nature Structural & Molecular Biology*, 2018, **25**: 833-840.
- [213] Boija A, Klein I A, Sabari B R, *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains[J]. *Cell*, 2018, **175**: 1842-1855.
- [214] Cho W K, Spille J H, Hecht M, *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates[J]. *Science*, 2018, **361**: 412-415.
- [215] Larson A G, Elnatan D, Keenen M M, *et al.* Liquid droplet formation by HP1 alpha suggests a role for phase separation in heterochromatin[J]. *Nature*, 2017, **547**: 236-240.
- [216] Sanulli S, Trnka M J, Dharmarajan V, *et al.* HP1 reshapes nucleosome core to promote phase separation of heterochromatin[J]. *Nature*, 2019, **575**: 390-394.
- [217] Fan H T, Lu J W, Guo Y R, *et al.* BAHCC1 binds H3K27me3 via a conserved BAH module to mediate gene silencing and oncogenesis[J]. *Nature Genetics*, 2020, **52**: 1384-1396.
- [218] Plys A J, Davis C P, Kim J, *et al.* Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2[J]. *Genes & Development*, 2019, **33**: 799-813.
- [219] Wu L Y, Pan T, Zhou M, *et al.* CBX4 contributes to HIV-1 latency by forming phase-separated

- nuclear bodies and SUMOylating EZH2[J]. *Embo Reports*, 2022, **23**: e53855.
- [220] Fioriniello S, Marano D, Fiorillo F, *et al.* Epigenetic Factors that Control Pericentric Heterochromatin Organization in Mammals[J]. *Genes*, 2020, **11**: 595.
- [221] Souza P P, Volkel P, Trinel D, *et al.* The histone methyltransferase SUV420H2 and Heterochromatin Proteins HP1 interact but show different dynamic behaviours[J]. *Bmc Cell Biology*, 2009, **10**: 41.
- [222] Schwartz J C, Wang X Y, Podell E R, *et al.* RNA Seeds Higher-Order Assembly of FUS Protein[J]. *Cell Reports*, 2013, **5**: 918-925.
- [223] Brangwynne C P. Phase transitions and size scaling of membrane-less organelles[J]. *Journal of Cell Biology*, 2013, **203**: 875-881.
- [224] Liu X Y, Jiang S S, Ma L, *et al.* Time-dependent effect of 1,6-hexanediol on biomolecular condensates and 3D chromatin organization[J]. *Genome Biology*, 2021, **22**: 230.
- [225] Shin Y, Berry J, Pannucci N, *et al.* Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets[J]. *Cell*, 2017, **168**: 159-171.
- [226] Seif E, Kang J J, Sasseville C, *et al.* Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity[J]. *Nature Communications*, 2020, **11**: 5609.
- [227] Dong R Y, Granick S. Reincarnations of the phase separation problem[J]. *Nature Communications*, 2021, **12**: 911.
- [228] Krainer G, Welsh T J, Joseph J A, *et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions[J]. *Nature Communications*, 2021, **12**: 1085.
- [229] Sowpati D T, Ramamoorthy S, Mishra R K. Expansion of the polycomb system and evolution of complexity[J]. *Mech Dev*, 2015, **138 Pt 2**: 97-112.
- [230] Van Den Boom V, Rozenveld-Geugien M, Bonardi F, *et al.* Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells[J]. *Blood*, 2013, **121**: 2452-61.
- [231] Vandamme J, Volkel P, Rosnoblet C, *et al.* Interaction Proteomics Analysis of Polycomb Proteins Defines Distinct PRC1 Complexes in Mammalian Cells[J]. *Molecular & Cellular Proteomics*, 2011, **10**: M110.002642.
- [232] Boulay G, Rosnoblet C, Guerardel C, *et al.* Functional characterization of human Polycomb-like 3 isoforms identifies them as components of distinct EZH2 protein complexes[J]. *Biochemical Journal*, 2011, **434**: 333-342.
- [233] Saurin A J, Shiels C, Williamson J, *et al.* The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain[J]. *Journal of Cell Biology*, 1998, **142**: 887-898.

附录

引物名称	引物序列 (5'到 3')
RT-qPCR 引物:	
β -Actin-F	CTAAGGCCAACCGTGAAAAG
β -Actin-R	ACCAGAGGCATACAGGGACA
mCbx7a-F1	AGGAGAGAGACCGAGCCTCG
mCbx7a-R1	CGCGCCAGAGAGAAGCAGA
mCbx7a+b-F2	CCGGAGACTGGGAGCCTATGGA
mCbx7a+b-R2	ACGGAGTTGGCGGTGATGTCAG
mCbx7c-F3	CTCTCTACCCTACGCAGCACC
mCbx7c-R3	TGGCCTTTGCAAGATCCCTTCT
Hoxa11-F	TGATGGATTTTGATGAGCGTGGTC
Hoxa11-R	TGGGGCAGGTTGGAGGAGT
Hoxc11-F	GAGCTTATGCACCGGGAGT
Hoxc11-R	CGTAGGAGCCTTCATTTTCA
Dlx5-F	GGA CTGACGCAAACACAGG
Dlx5-R	GCGAGTTACACGCCATAGG
Evx1/2-F	CAGGGAGAACTACGTTTCAAGAC
Evx1/2-R	GCCGGTTCTGAAACCACA
Lhx1-F	TGGGCAGTCTGACACGCACAC
Lhx1-R	AAGACGCCAAGAGCTCAGGATTC
Tal1-F	CGCCTCACTAGGCAGTGG
Tal1-R	CTCTTCACCCGGTTGTGTT
Ccnd2-F	CAAGCATGCGCAGACCTTCATC
Ccnd2-R	GGGTACATGGCAAACCTGAAGTCG
Spatc1-F	AGCTCCACCTCCAAAGTCAATG
Spatc1-R	TCCATCTCCACGACATTCTTCCG
Serpib9b-F	GGTCTCCAAACAGACTGAAGGC
Serpib9b-R	ATGGCATTTCCTGGTGGATTC
Neurod1-F	ACAACAGGAAGTGGAAACATGACC
Neurod1-R	ACACTCATCTGTCCAGCTTGGG
Gata6-F	ATTACACGACGCGACTAGCAG
Gata6-R	TTTGATTCTCGAGCGATGTGC
Adamts1-F	TGTCAGTGGCAAGTGCCTGAAC
Adamts1-R	CCCAGCTTCCATGAACAGGAGTAG
Tmem170b-F	TGAGGAACCTCACGGAGATGTG
Tmem170b-R	AGCACCGTGGACAAACAGAGAG
Pax2-F	AAGAGGAAACGCGAGGAAGTCG
Pax2-R	ACCTCCACCTCCTCTAATGTGG
Jade1-F	TCATGAAACGAGGTCGCCTTCC

引物名称	引物序列 (5'到 3')
Jade1-R	GGACCATGTCGTTGACAAGCTG
Zfp292-F	GTGCTGGAACGCTTGGCATTAAAG
Zfp292-R	TGATAGCTCAACAGGCAGACACAG
Atf5-F	ACTGTGGGCTATTTGGCAGAGTC
Atf5-R	AGGGCCTTGCTTTCTCAGTTGC
Atf3-F	AGACAGAGTGCCTGCAGAAAGAG
Atf3-R	GCCTTCAGCTCAGCATTACAC
Cdh22-F	TTCTTCGTGGTGGAGGAGTACACG
Cdh22-R	CGTCAGAATCCGAGTGGATCTTGC
Grin2c-F	TCTCACCAAGGGCAAGAAGTCAG
Grin2c-R	CAACCACACGGACTTGCCAATG
ChIP-qPCR 引物:	
NC-1-Myc-F	TTAGGTAAGAATTGGCAAGGATA
NC-1-Myc-R	AGGGTTAGGGCACAGGTGAGAAA
NC-1-Klf4-F	GAACGGCTCTGCTGGAGAAATTA
NC-1-Klf4-R	GGCTCTGTTGAAACCCAAAGGAT
Hoxa11-F	TGATGGATTTTGATGAGCGTGGTC
Hoxa11-R	TGGGGCAGGTTGGAGGAGT
Hoxc11-F	GAGCTTATGCACCGGGAGT
Hoxc11-R	CGTAGGAGCCTTCATTTTCA
Dlx5-F	GTTGGACGAGTTAGGGTGTG
Dlx5-R	GTGGATCGGGTTCCATTG
Evx1/2-F	TGTGTGTGTCTCTCTGTTTTCACA
Evx1/2-R	AATGGCTCGGCAGTGATG
Lhx1-F	TGGGCAGTCTGACACGCACAC
Lhx1-R	AAGACGCCAAGAGCTCAGGATTC
Tal1-F	GTCCGTTTCCTCCGTCTTT
Tal1-R	GGGGAAGGAAGGAAGAGAGTC
Ccnd2-F	CCGCGTTGGCACTTTGGCAGGAG
Ccnd2-R	GCGGTTCTTTTTCAGATTGCCCTAA
Grin2c-F	ACGGTGGAGACAGACGCGGGACA
Grin2c-R	CCAACAGTCCGGCTCCCAAGGGT
Adamts1-F	ACTCTGGCACGGTGAACGGTGAT
Adamts1-R	CAGGCGCTGGCTGAATGAAGAAC
Pax2-F	GGCGAATCACAGAGTGGTGAAT
Pax2-R	CCGGGAGATGGATGACTTGTCAG
Cdh22-F	GCAACAGCTGAAGCGCGCCCATC
Cdh22-R	GCGCCGGGGAAAGTTTGACGTG
分子克隆引物:	
pBiFC-CBX7A-F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATCGGCG
pBiFC-CBX7A-R	CGGGATCCGCTAGCTCAGAAGTCCCACTGCGGTCTCGG

引物名称	引物序列 (5'到 3')
pBiFC-CBX7B-F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATAGGC
pBiFC-CBX7B-R	CGGGATCCGCTAGCTCACAGCTTCTCGTTGCGGTCTC
pBiFC-CBX7C-F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATAGGC
pBiFC-CBX7C-R	CGGGATCCTCATTTTAAGCTCTCAGAGGTTGGCTGTC
pBiFC-CBX7A (1-82)-F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATAGGC
pBiFC-CBX7A (1-82)-R	CGGGATCCTCACTGTAGCAGAAGCCGCTGGGTTT
pBiFC-CBX7A (83-251) -F	GCTCTAGAGCAGTCTCTATGCGGCTGTACAGCATGGAC
pBiFC-CBX7A (83-251) -R	CGGGATCCGCTAGCTCAGAAGTCCCACTGCGGTCT
pBiFC-CBX7A (Δ CD) -F	GCTCTAGAGCAGTCTCTATGGCCTACGAGGAGAAGGA
pBiFC-CBX7A (Δ CD) -R	CGGGATCCTCAGAAGTCCCACTGCGGTCTCGGAAG
pBiFC-CBX7A (Δ CD+ATHL) -F	GCTCTAGAGCAGTCTCTATGCTTCTGCTGCAGCGGCTG
pBiFC-CBX7A (Δ CD+ATHL) -R	CGGGATCCGCTAGCTCAGAAGTCCCACTGCGGTCTCGG
pBiFC-CBX7C (1-110) -F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATAGGC
pBiFC-CBX7C (1-110) -R	CGGGATCCGCTAGCTCATGGGCCAAGGGCCAGTGCA
pBiFC-CBX7C (111-166) -F	GCTCTAGAGCAGTCTCTATGCCACTCTCTACCCCTACGCA
pBiFC-CBX7C (111-166) -R	CGGGATCCGCTAGCTCATTTTAAGCTCTCAGAGGTTG
pBiFC-CBX7C (Δ CD+ATHL) -F	GCTCTAGAGCAGTCTCTATGCTTCTGCTGCAGCGGCTG
pBiFC-CBX7C (Δ CD+ATHL) -R	CGGGATCCTCATTTTAAGCTCTCAGAGGTTGGCTGTC
pBiFC-PHC2 (Δ SAM) -F	GCTCTAGAGCAGTCTCTATGACCTCAGGGAACGGAAGT
pBiFC-PHC2 (Δ SAM) -R	CGGGATCCTCATGGCAGGAAGTGGTGCCATGCCACCA
pBiFC-CBX7A(Δ ATHL) -1F	GCTCTAGAGCAGTCTCTATGGAGCTGTCAGCCATCGGCG
pBiFC-CBX7A(Δ ATHL) -1R	GTACAGCCGCTGCAGCAGAAGATACCCCGATGCTCG
pBiFC-CBX7A (Δ ATHL) -2F	CGAGCATCGGGTATCTTCTGCTGCAGCGGCTGTAC
pBiFC-CBX7A (Δ ATHL) -2R	CGGGATCCGCTAGCTCAGAAGTCCCACTGCGGTCTCGG
PGEX-6p-1-CBX7A-F	TCCAGGGGCCCCCTGGGATCCATGGAGCTGTCA
PGEX-6p-1-CBX7A-R	GATGCGGCCGCTCGAGTCGACTCAGAAGTTC
PGEX-6p-1-CBX7A (1-82) -F	TCCAGGGGCCCCCTGGGATCCATGGAGCTGTCA
PGEX-6p-1-CBX7A (1-82) -R	GATGCGGCCGCTCGAGTCGACTCACTGTAGCAGAA
PGEX-6p-1-CBX7A (83-251) -F	TCCAGGGGCCCCCTGGGATCCATGCGGCTGTAC
PGEX-6p-1-CBX7A (83-251) -R	GATGCGGCCGCTCGAGTCGACTCAGAAGTTC
PGEX-6p-1-CBX7C-F	TCCAGGGGCCCCCTGGGATCCATGGAGCTGTCA
PGEX-6p-1-CBX7C-R	GATGCGGCCGCTCGAGTCGACTCATTTTAAGCTCTCAG
PGEX-6p-1-CBX7C (1-51)-F	TCCAGGGGCCCCCTGGGATCCATGGAGCTGTCA
PGEX-6p-1-CBX7C (1-51)-R	GATGCGGCCGCTCGAGTCGACTCAAGGGTCCAAGATGT
PGEX-6p-1-CBX7C (1-110)-F	TCCAGGGGCCCCCTGGGATCCATGGAGCTGTCA
PGEX-6p-1-CBX7C (1-110)-R	GATGCGGCCGCTCGAGTCGACTCATGGGCCAAG
PGEX-6p-1-CBX7C (52-166)-F	TCCAGGGGCCCCCTGGGATCCATGCGCCTTGTCAT
PGEX-6p-1-CBX7C (52-166)-R	GATGCGGCCGCTCGAGTCGACTCATTTTAAGCTCTCAG
PGEX-6p-1-CBX7C (111-166)-F	TCCAGGGGCCCCCTGGGATCCATGCCACTCTCTCA
PGEX-6p-1-CBX7C (111-166)-R	GATGCGGCCGCTCGAGTCGACTCATTTTAAGCTCTCAG
PGEX-6p-1-EGFP-F	TCCAGGGGCCCCCTGGGATCCATGGTGAGCAAG

引物名称	引物序列 (5'到 3')
PGEX-6p-1-EGFP-R	GATGCGGCCGCTCGAGTCGACTCACTTGTACAGCTCGT
PGEX-6p-1-EGFP-protein-F	TCCAGGGGCCCCCTGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTT CACCGGGGTGGTGCC
PGEX-6p-1-EGFP-protein-R	TGACAGCTCCATTCTAGAGTCGACTGGTACCGATATCAGATCTATCG ATGAATTCTTGTACAGCTC
PGEX-6p-1-EGFP-CBX7A-F	GAGCTGTACAAGAATTCATCGATAGATCTGATATCGGTACCACTCGA CTCTAGAATGGAGCTGTCA
PGEX-6p-1-EGFP-CBX7A-R	GATGCGGCCGCTCGAGTCGACTCAGAACTTCCCACTGCGGTCTCGG AAGAAGCCCTCAG
PGEX-6p-1-EGFP-CBX7C-F	GAGCTGTACAAGAATTCATCGATAGATCTGATATCGGTACCACTCGA CTCTAGAATGGAGCTGTCA
PGEX-6p-1-EGFP-CBX7C-R	GATGCGGCCGCTCGAGTCGACTCATTTTAAGCTCTCAG
pET28a-RING1B-F	GCCGCGCGGCAGCCATATGATGTCTCAGGCTGTGCAG
pET28a-RING1B-R	GGTGGTGGTGCTCGAGTTCATTGTGCTCCTTTGTAGG
pET28a-PHC2-F	GCCGCGCGGCAGCCATATGATGACCTCAGGGAACGGAAA
pET28a-PHC2-R	GGTGGTGGTGCTCGAGTCTAGGAGTCCTTGAGCATG
pcDNA3.1-TRE-CMV-CBX7A-EGFP-Flag-F	TGCTAGCGGCCCGGAATTCATGGAGCTGTCAGCCAT
pcDNA3.1-TRE-CMV-CBX7A-EGFP-Flag-R	TTGCTCACCATATCGATCCCGCCACCTCCGGATCCCG
pcDNA3.1-TRE-CMV-CBX7C-EGFP-Flag-F	TGCTAGCGGCCCGGAATTCATGGAGCTGTCAGCCATA
pcDNA3.1-TRE-CMV-CBX7C-EGFP-Flag-R	TTGCTCACCATATCGATCCCGCCACCTCCGGATCCCG
pcDNA3.1-CMV-CBX7C-EGFP-Flag-F	TAGCGTTTAACTTAAGCTTATGGAGCTGTCAGCCATA
pcDNA3.1-CMV-CBX7C-EGFP-Flag-R	TTGCTCACCATATCGATCCCGCCACCTCCGGATCCCG
pBiFC-mCherry-PHC2-F	AAGACGATGACGACAAGCTTATGGTGAGCAAGGGCGAGGAG
pBiFC-mCherry-PHC2-R	ATCTATCGATGAATTCTTGTACAGCTCGTCCATGCCG
pcDNA3.1-CMV-mCherry-HTM-1F	AAGACGATGACGACAAGCTTATGGTGAGCAAGGGCGAGGAGG
pcDNA3.1-CMV-mCherry-HTM-1R	GCAGCGTGGTGGGTGGACATCCCGCCACCTCCGGATCCCGGGCC
pcDNA3.1-CMV-mCherry-HTM-2F	GGCCCGGATCCGGAGGTGGCGGGATGTCCACCCACCACGCTGC
pcDNA3.1-CMV-mCherry-HTM-2R	ACAGGTTTTCTCTAGATCAGAAGGCGAGGGCCTGCTTCGG
pcDNA3.1-CMV-mCherry-HTM-protein-F	AAGACGATGACGACAAGCTTATGGTGAGCAAGGGCGAGGAGGATA ACATGGCCATCATCAAGGA
pcDNA3.1-CMV-mCherry-HTM-protein (A)-R	CTGACAGCTCCATTCTAGAGTCGACTGGTACCGATATATCGATTCCCG GGCCGAAGGCGAGGGCCT
pcDNA3.1-CMV-mCherry-HTM-CBX7A-F	AGGCCCTCGCCTTCGGCCCGGAATCGATATATCGGTACCACTCGAC TCTAGAATGGAGCTGTCAG
pcDNA3.1-CMV-mCherry-HTM-CBX7A-R	ACAGGTTTTCTCTAGATCAGAACTTCCCACTGCGGTCTCGGAAGAA GCCCTCAGCT
pcDNA3.1-CMV-mCherry-HTM-protein (C)-R	ATGGCTGACAGCTCCATTCTAGAGTCGACTGGTACCGATATATCGATT CCCGGGCCGAAGGCGAGGGC
pcDNA3.1-CMV-mCherry-HTM-CBX7C-F	GCCCTCGCCTTCGGCCCGGAATCGATATATCGGTACCACTCGACTC TAGAATGGAGCTGTCAGCCAT
pcDNA3.1-CMV-mCherry-HTM-CBX7C-R	ACAGGTTTTCTCTAGATCATTTTAAGCTCTCAGAGGTGGCTGTCTGT

引物名称	引物序列 (5'到 3')
	CTACCCTCAGGC
pcDNA3.1-CMV-mCherry-HTM-protein (PHC2)-R	TTCCGTTCCCTGAGGTCAATTCTAGAGTCGACTGGTACCGATATATCGA TTCCCGGGCCGAAGGCGAGGGC
pcDNA3.1-CMV-mCherry-HTM-PHC2-F	GCCCTCGCCTTCGGCCCGGAATCGATATATCGGTACCAGTCGACTC TAGAATGACCTCAGGGAACGGAA
pcDNA3.1-CMV-mCherry-HTM-PHC2-R	ACAGGTTTTCTCTAGACTAGGAGTCCTTGAGCATGCTGATGCGTGCG TAGATCTTCAGGGCGGGC
pcDNA3.1-CMV-mCherry-Cry2oligo-1F	TAGCGTTTAACTTAAGCTTATGGTGAGCAAGGGCGAGGAGGATAA CATGGCCATCAT
pcDNA3.1-CMV-mCherry-Cry2oligo-1R	ATAGTCTTTTTGTCCATCTTCATTGATATCAGATCTATTGATGAATTCT TGTACAGC
pcDNA3.1-CMV-mCherry-Cry2oligo-2F	GCTGTACAAGAATTCATCAATAGATCTGATATCAATGAAGATGGACA AAAAGACTAT
pcDNA3.1-CMV-mCherry-Cry2oligo-2R	ACAGGTTTTCTCTAGATGCTGCTCCGATCATGATCTGTGCTCCACGG GTTCTTGAAATAGCT
pcDNA3.1-CMV-protein-mCherry-Cry2oligo-R	ACAGGTTTTCTCTAGATGCTGCTCCGATCATGATCTGTGCTCCACGG GTTCTTGAAATAGCT
pcDNA3.1-CMV-protein(A)-mCherry-Cry2oligo-F	AGTGGGAAGTTCGGCCCGGATCCGGAGGTGGCGGGATCGATATGG TGAGCAAGGGCGAGGA
pcDNA3.1-CMV-CBX7A-mCherry-Cry2oligo-F	TAGCGTTTAACTTAAGCTTATGGAGCTGTCAGCCATCGGCGAGCAG GTGTTCCGCGT
pcDNA3.1-CMV-CBX7A-mCherry-Cry2oligo-R	TCCTCGCCCTTGCTCACCATATCGATCCCGCCACCTCCGGATCCCGG GCCGAACCTCCCACT
pcDNA3.1-CMV-protein(C)-mCherry-Cry2oligo-F	GAGAGCTTAAAAGGCCCGGATCCGGAGGTGGCGGGATCGATATGG TGAGCAAGGGCGAGGA
pcDNA3.1-CMV-CBX7C-mCherry-Cry2oligo-F	TAGCGTTTAACTTAAGCTTATGGAGCTGTCAGCCATAGGCGAGCAG GTGTTTGCGGT
pcDNA3.1-CMV-CBX7C-mCherry-Cry2oligo-R	TCCTCGCCCTTGCTCACCATATCGATCCCGCCACCTCCGGATCCCGG GCCTTTTAAGCTCTC
pTripZ-5×UAS-CYP26A1-EGFP-Flag-1F	TGGATCCATCATCGACGCGTTCGGAGTACTGTCTCCGTA
pTripZ-5×UAS-CYP26A1-EGFP-Flag-1R	GTACACGCCTACCTCGACATACGTTCTCTATCACTGATAG
pTripZ-5×UAS-CYP26A1-EGFP-Flag-2F	CTATCAGTGATAGAGAACGTATGTCGAGGTAGGCGTGTAC
pTripZ-5×UAS-CYP26A1-EGFP-Flag-2R	ACAGGTTTTCTCTAGACTTGACAGCTCGTCCATGCCGAG
GAL4 DBD-CBX7A-F	GCCCCGGCTGGAATTCTAGAAATGGAGCTGTGAGCCATCGGCGA
GAL4 DBD-CBX7A-R	TACTTATCAAGATCTTCAGAACTTCCCACTGCGGTCTC
GAL4 DBD-CBX7C-F	GCCCCGGCTGGAATTCTAGAAATGGAGCTGTGAGCCATA
GAL4 DBD-CBX7C-R	TACTTATCAAGATCTTCATTTTAAGCTCTCAGAGGTT
GAL4 DBD-IRES-Renilla Luciferase-1F	TTCTAGAAATCGATAGATCTGACTACAAAGACGATGACGACAAGTG AGCCCTCTCCCTCCCCCCCCCT
GAL4 DBD-IRES-Renilla Luciferase-1R	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATGGTATTAT CGTGTTTTTCAAAGGAAACCAGT

引物名称	引物序列 (5'到 3')
GAL4 DBD-IRES-Renilla Luciferase-2F	ACGTGGTTTTCTTTGAAAAACACGATAATACCATGACTTCGAAAGT TTATGATCCAGAACAAAGGAAACGGA
GAL4 DBD-IRES-Renilla Luciferase-2R	CTTAAAAATTGGATCTACTTATTCATCATTGTTCATTTTGTAGAACTC GCTCAACGAACGATTTGATATATTTCCC
GAL4DBD-CBX7A-IRES-Renilla Luciferase-F	GCCCCGGGCTGGAATTCTAGAAATGGAGCTGTCAGCCATCGGCGAGC A
GAL4DBD-CBX7A-IRES-Renilla Luciferase-R	CTTTGTAGTCAGATCTGAACCTCCCACTGCGGTCTCGGAAGAA
GAL4DBD-CBX7C-IRES-Renilla Luciferase-F	GCCCCGGGCTGGAATTCTAGAAATGGAGCTGTCAGCCATAGGCGAGC A
GAL4DBD-CBX7C-IRES-Renilla Luciferase-R	CTTTGTAGTCAGATCTTTTAAAGCTCTCAGAGGTGGCTGTCTGT
GAL4DBD-PHC2-IRES-Renilla Luciferase-F	GCCCCGGGCTGGAATTCTAGAAATGACCTCAGGGAACGGAACTCTG CCT
GAL4DBD-PHC2-IRES-Renilla Luciferase-R	CTTTGTAGTCAGATCTGGAGTCCTTGAGCATGCTGATGCGGGCGT