



染色质结构蛋白变异与人类疾病

曹红^{1†}, 吴迪^{1,2†}, 魏瑜^{1,2†}, 林海燕^{1*}, 李国红^{1,2*}

1. 中国科学院生物物理研究所, 生物大分子国家重点实验室, 北京 100101;

2. 武汉大学生命科学院细胞稳态湖北省重点实验室, 泰康生命医学中心, 免疫与代谢前沿科学中心, 武汉 430072

† 同等贡献

* 联系人, E-mail: liguohong@sun5.ibp.ac.cn; linhy@ibp.ac.cn

收稿日期: 2023-09-21; 接受日期: 2023-10-09; 网络版发表日期: 2023-11-10

北京市科技计划项目(批准号: Z221100007022001)和国家自然科学基金(批准号: 32300435, 32270581)资助

摘要 层级复杂的三维染色质结构对于细胞命运决定和功能维持所需的多种DNA相关生物学过程的时空调控至关重要, 如DNA复制、转录、重组和损伤修复等. 三维染色质结构失调导致基因表达异常, 被认为是肿瘤或神经发育障碍等多种疾病的主要诱因. 本文重点阐述组蛋白及其变体、甲基CpG结合蛋白2在三维染色质高级结构及动态性调节中的作用, 总结疾病相关突变对基因功能的影响, 探讨肿瘤或神经发育障碍发生发展过程中染色质层面的病理学机制.

关键词 三维染色质结构, 染色质结构蛋白, 组蛋白及其变体, 甲基CpG结合蛋白2(MeCP2), 人类疾病

作为遗传信息的载体, 真核生物基因组DNA并非裸露存在, 而是与组蛋白(H1, H2A, H2B, H3, H4)、非组蛋白和少量RNA共同构成复杂的三维染色质高级结构. 染色质基本结构单元是核小体, 由基因组DNA缠绕核心组蛋白(H2A, H2B, H3, H4)八聚体形成, 在连接组蛋白H1和其他染色质结构蛋白的参与下, 折叠形成30纳米染色质纤维, 并在不同机制的调控下经过逐级组装和折叠形成更加复杂的染色质高级结构(图1)^[1]. DNA甲基化、组蛋白翻译后修饰(post-translational modifications, PTMs)、组蛋白变体和染色质重塑复合物等许多因素参与染色质高级结构的形成及动态调控.

利用体外染色质组装体系、冷冻电子显微镜(冷冻电镜)和单分子磁镊技术, 李国红和朱平团队^[2]在国

际上率先解析了30纳米染色质纤维的高分辨率三维冷冻电镜结构, 该结构揭示了30纳米染色质纤维是以“四核小体”为结构单元的左手双螺旋zig-zag构象; 尽管体内是否存在30纳米染色质纤维仍存在争议, 高分辨率30纳米染色质纤维结构的解析为理解染色质折叠压缩的基本原理提供了依据. 染色质高级结构研究方法的进步, 例如, 高通量染色体构象捕获(high-throughput chromosome conformation capture, Hi-C)、活细胞超分辨率成像技术以及基于CRISPR-Cas9或光遗传系统开发的可诱导染色质结构重塑工程技术, 为研究染色质高级结构在基因表达调控中的作用以及高级染色质失调如何导致疾病发生提供了有力手段^[3-7].

间期染色质是染色质最高层次的折叠, 它以非随机方式占据核质不同区域, 形成染色体疆域(chromo-

引用格式: 曹红, 吴迪, 魏瑜, 等. 染色质结构蛋白变异与人类疾病. 中国科学: 生命科学, 2023, 53: 1575-1594

Cao H, Wu D, Wei Y, et al. Variations in chromatin architectural proteins in human diseases (in Chinese). Sci Sin Vitae, 2023, 53: 1575-1594, doi: 10.1360/SSV-2023-0221

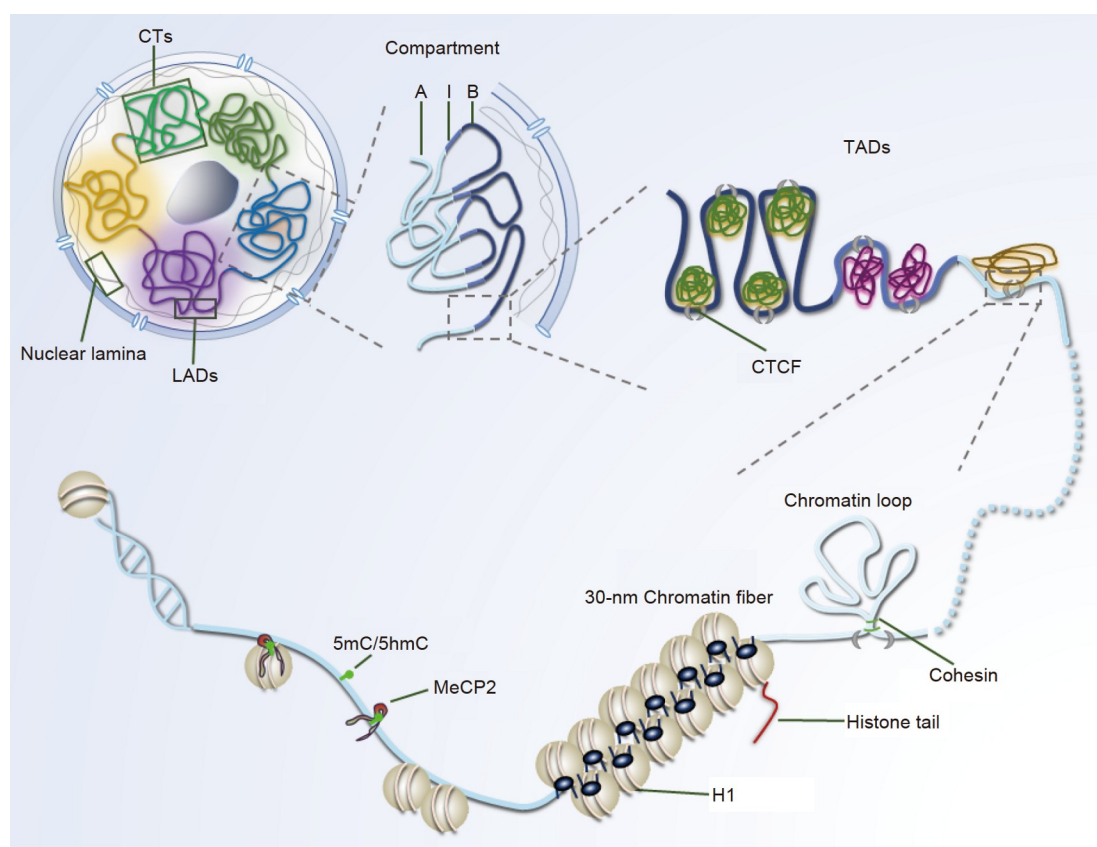


图 1 层级复杂的三维染色质结构. 从小到大依次为由DNA双螺旋和组蛋白组成的核小体、30纳米染色质纤维、染色质环、拓扑相关结构域、染色质区室和染色质疆域. 此外, 组蛋白尾部氨基酸发生翻译后修饰, 游离DNA或缠绕核小体的DNA发生甲基化修饰

Figure 1 The hierarchical three-dimensional (3D) chromatin architecture. The nucleosome is composed of the octameric histone core and the DNA double helix wrapped around it. Nucleosomes are packed together and hierarchically folded into 30-nm chromatin fibers and higher-order chromatin structures, including chromatin loops, TADs, chromatin compartments (compartment A, I, B) and CTs. In addition, histones undergo PTMs at different amino acids mainly on their unstructured tails. Methylation modifications occur on DNA that is either free or wrapped around nucleosomes

somal territories, CTs)^[8]. 每个染色体疆域进一步分隔成互不干扰的兆碱基(Mb)级别的不同区室, “A区室(compartment A)”中染色质开放并易于接近, 基因转录更容易被激活, 而“B区室(compartment B)”中的染色质高度压缩并难以接近, 基因转录是沉默的, A和B区室具有组织或细胞类型特异性^[9]. 最近, 根据染色质区域内H3K27me3富集程度将其定义为“中间区室I(compartment I)”^[10]. 染色质区室可以进一步由亚兆碱基(sub-Mb)级别的拓扑相关结构域(topologically associating domains, TADs)组成, 该结构具有细胞类型或物种之间高度的保守性, 表明其作为结构功能单元的重要性^[11]. TADs的边界由CTCF/Cohesin富集, TADs区域内DNA-DNA频繁发生相互作用, 例如, 控制基因时空准确表达的启动子和增强子之间形成染色质环

(chromatin loop)(图1)^[12].

越来越多的证据表明, 三维染色质结构重塑在发育过程中发挥关键作用. 在受精卵形成之初, 染色质结构松散, 在之后的合子发育过程中各级染色质结构逐步建立^[13]. 小鼠出生后一个月内, 大脑皮层和海马神经元中出现大量的染色质结构改变, 以保证神经发生、细胞发育和神经元投射^[14]. 三维染色质结构紊乱会导致严重的神经发育障碍, 如注意缺陷多动障碍(attention deficit/hyperkinetic disorder, ADHD)和孤独症谱系障碍(autism spectrum disorder, ASD), 尽管常见于儿童时期, 但可能是终身疾病^[15]. 此外, 伴随细胞周期的进行, 三维染色质结构会呈现出独特的结构图谱. 例如, 在经历快速增殖的生发中心B细胞或生发中心样弥漫性大B细胞淋巴瘤(diffuse large B cell lymphomas,

DLBCLs)中,三维染色质结构发生广泛重塑,即从压缩状态的B区室转变为更为开放和转录活跃的A区室,增强子和启动子之间的相互作用广泛增强^[10,16]。此外,生发中心B细胞分化为记忆B细胞或浆细胞的过程中,三维染色质结构发生可逆重塑,记忆B细胞的三维染色质结构与naïve B细胞非常相似^[10]。因此,三维染色质结构的高度动态性也为细胞中各种生命活动,如细胞命运决定、细胞周期调节以及免疫反应等有序进行提供保障。

近年来,大规模基因组测序技术筛选到许多以往未知的肿瘤或神经发育障碍关联基因突变^[17~19],其中,染色质结构蛋白,包括组蛋白及其变体、甲基CpG结合蛋白2(methyl-CpG binding protein 2, MeCP2),在30纳米染色质纤维和染色质高级结构的形成、维持以及动态调控中起决定性作用。本文将重点总结核心组蛋白及其变体、连接组蛋白H1以及MeCP2在三维染色质结构及动态调节中的作用,探讨疾病相关突变对其上述功能的影响,以加深对体内染色质结构和功能多样性的理解,为疾病病理学机制的揭示、诊断和治疗提供帮助。

1 核心组蛋白与肿瘤/神经发育障碍

组蛋白约占真核生物染色质质量的一半,是真核细胞中最丰富的碱性蛋白质家族,也是进化上最保守的蛋白家族之一。组蛋白由多拷贝基因编码,大部分组蛋白基因在染色体上成簇分布,这种基因的冗余可以确保细胞复制过程中组蛋白的连续来源。核心组蛋白是组成核小体最丰富的蛋白,以复制偶联的方式合成并掺入旧核小体之间,压缩DNA,降低染色质可及性。

组蛋白PTMs被称为组蛋白密码(histone code)^[20],广泛参与转录、DNA重组、复制、损伤修复以及基因组稳定性的调节,PTMs的多样性决定了表观遗传调控的可塑性和复杂性^[21]。组蛋白乙酰化大多与转录激活相关,而Lys或Arg可被单/双/三甲基化修饰,其对基因转录的调控作用更加复杂^[21]。H2A K119的单泛素化修饰(H2AK119ub1)参与基因转录沉默,抑制H3K4me2/me3的形成;H2BK120ub1参与H3K4me2, H3K4me3, H3K79甲基化的形成及基因转录激活^[22]。而且,组蛋白PTMs也通过调节核小体结构、招募修复

因子、调节其他PTMs位点等诸多途径参与DNA损伤应答。 γ H2AX(H2AXS139ph)是DNA双链断裂(double strand break, DSB)的分子标志,作为DSB感应起始信号招募一系列蛋白并形成DNA损伤反应功能复合物,启动激活DNA修复、细胞周期检查点等DNA损伤反应^[23,24]。H2A和H2AX的K63多聚泛素化参与DNA损伤修复;RING1B/BMI1和RNF20/RNF40也能被招募到DSB位点,分别催化生成H2A/H2AX K119ub1和H2BK120ub1,辅助DNA损伤反应复合物的装配^[22]。在同源重组修复(homologous recombinant repair, HRR)过程中, TIP60对H4K16进行乙酰化, RNF8泛素化修饰H2AK63,选择性招募BRCA1;在非同源末端连接(non-homologous end joining, NHEJ)中, HDAC1/2乙酰化修饰H4K16, 53BP1特异性结合H4K20me2和H2AK15ub^[25]。

除H4迄今只发现一种变体外,核心组蛋白存在多种不同形式的变体,尤其是H2A和H3,其一级序列之间存在显著差异,多以非DNA复制依赖的方式存在于整个细胞周期中,或以组织特异性的方式表达。大量组蛋白变体的存在,且不同物种之间具有进化的保守性,表明组蛋白在真核生物中的重要性^[26]。组蛋白变体可以与相应核心组蛋白之间发生动态交换,作为压缩真核基因组的支架蛋白改变染色质组成。含组蛋白变体的核小体在染色质上不是均匀分布的,组蛋白变体的掺入可以对核小体的结构和稳定性产生直接影响,赋予染色质特定状态,从而在染色质水平上调控基因表达。鉴于核心组蛋白尾部PTMs的重要性,核小体中组蛋白变体的募集也可能间接影响表观遗传标记,进而参与染色质高级结构的形成以及基因表达调控。

1.1 致癌组蛋白(oncohistones)

近年来,肿瘤全基因组测序研究发现,除组蛋白修饰相关酶以及染色质重塑复合物相关蛋白编码基因发生突变,在多种肿瘤类型中还发现了组蛋白编码基因的突变。这些突变具有肿瘤类型的偏好性、显性失活效应以及高达3.8%的突变率,与BRCA2和TET2等常见癌症相关基因突变率相近,因此,肿瘤中发现的带有突变的组蛋白被称为“致癌组蛋白”,在肿瘤发生发展中起关键作用^[19,27]。

组蛋白H3是第一个在儿童胶质母细胞瘤(pediatric glioblastoma, pGBM)中被发现的频繁突变的组蛋白,

现已扩展到软骨母细胞瘤、软骨肉瘤、骨肉瘤、头颈鳞状细胞癌、小儿软组织肉瘤、膀胱癌、黑色素瘤和急性髓系白血病等肿瘤中^[28]。研究最多的H3突变发生在H3.3变体N端结构域内的关键PTM位点或附近, 包括K27M, K36M和G34R/V/W, 目前认为这些突变主要通过破坏组蛋白的PTMs来发挥致癌作用^[29-31]。其中, H3K27M是儿童恶性胶质母细胞瘤(pediatric high-grade glioblastoma, pHGG)的高频突变, 主要发生在H3.3编码基因(*H3F3A*和*H3F3B*, 其中*H3F3A*>70%), 较少发生于H3.1编码基因(*HIST1H3B*或*HIST1H3C*), 极少发生于H3.2编码基因(*HIST2H3C*)^[32]。世界卫生组织(World Health Organization, WHO)将H3K27M作为肿瘤分类标志物并将H3K27M肿瘤单独划分为一个亚型——弥漫性中线神经胶质瘤(diffuse midline glioma, DMG)^[33]。此外, 对于软骨母细胞瘤(chondroblastoma), H3K36M是唯一的复发突变, 突出了致癌组蛋白在驱动肿瘤发生中的重要作用^[34]。

在不同发育时期或不同的细胞类型中, 组蛋白及其变体的特异性表达和修饰模式不同。肿瘤中组蛋白突变的发生也具有细胞谱系、组织及时空特异性, 仅在特定的发育时期或特殊的细胞类型中出现。儿童神经胶质瘤(pediatric gliomas)的转录组学分析表明, 不同肿瘤亚型存在不同的发育起源, H3.3G34R/V型肿瘤中具有早期神经前体细胞和抑制性神经元的特征, 而H3.3K27M型脑干瘤则显示后脑和晚期神经前体细胞的特征^[35]。骨巨细胞瘤(giant cell tumor of bone, GCT)中的G34W突变特定发生于H3F3A编码产物, 而软骨母细胞瘤中的K36M突变主要发生在H3F3B编码产物上^[34]。突变位点决定了分子伴侣的结合特异性及其在核小体装配和基因组定位的不同模式, 并影响组蛋白的重要PTMs位点^[32]。此外, 肿瘤中的组蛋白突变很少单独发生, 一般与特定的癌基因协同出现, 可能利于肿瘤细胞的增殖。例如, H3K27M在不同脑区中的频繁突变与特定生长因子受体的突变经常关联发生, 包括脑干中的H3.3K27M与血小板衍生生长因子(platelet derived growth factor alpha, PDGFRA)、丘脑中的H3.3K27M与成纤维细胞生长因子受体(fibroblast growth factor receptor 1, FGFR1)、脑干中的H3.1K27M与激活素A I型受体(activin A receptor type 1, ACVR1)^[35]。

致癌组蛋白H3K27M, K36M和G34R/V/W诱导肿

瘤发生的机制比较复杂, 简单理解为通过反式或顺式作用对组蛋白甲基转移酶的活性产生影响^[36]。具有显性效应的H3K27M和H3K36M通过反式作用抑制其相应Lys的甲基转移酶活性, 导致突变细胞中H3K27me3或H3K36me3整体水平下降。而H3G34R/V/W则顺式降低H3K36me2/3甲基化, 即仅含有H3G34R/V/W的核小体表现出H3K36me2/3水平降低。此外, 主要在恶性胶质瘤(high grade glioma, HGG)中发现的H3G34R/V也会导致H3K27me3整体水平的升高, 阻碍神经元分化(图2A和B)。另有机理研究显示, H3K36M肿瘤细胞中, H3K27me3发生全基因组水平的上调, 并沉默原有的H3K36甲基化富集的基因体(gene body)和增强子区^[37]; 另一方面, H3K27me3的上调导致其阅读器(reader)PRC1(polycomb repressive complex 1)^[38]在基因组上的重新分布以及染色质的解压缩和基因转录抑制的解除^[39]。在软骨母细胞瘤中, H3G34R/V顺式抑制的H3K36me3能影响DSB同源重组修复^[40]。而且, H3.3G34R能导致H3.3K36me3阅读器ZMYND11与H3.3G34R-核小体的结合下调、前脑核心调控因子(如forkhead box G1, FOXG1)的异常转录与剪接, 促进神经胶质瘤的发生^[39]; H3.3G34R/V也能导致染色质环的建立发生异常, 促进胶质瘤癌基因*PDGFRA*的转录激活^[41]。类似的, 与野生型胶质母细胞瘤相比, H3.3K27M突变的神经胶质瘤(diffuse intrinsic pontine glioma, DIPG)中TADs的边界及染色质环扩大, 并通过“增强子劫持”(enhancer hijacking)导致癌基因表达增强^[42]。

最近, 基于cBioPortal数据库的两项分析研究, 对涵盖几种肿瘤类型的数千名患者的组蛋白错义突变进行了分类和表征, 发现癌相关组蛋白突变不仅局限于组蛋白尾部PTM位点及其附近残基, 而且还分布在整个基因体中, 并且突变最严重的残基大部分位于组蛋白折叠结构域(histone fold domain)中, 如H2A E121, H2B G53/E76/E113, H3 E97/E105/R131以及H4 D68/R92等, 这些突变影响核小体的完整性以及DNA-核小体相互作用, 表明核小体或染色质高级结构的异常可能是致癌组蛋白介导致癌作用的主要机制^[19,43]。

H2B突变是近年来发现的肿瘤中最常见的组蛋白突变之一, 与H3组蛋白突变不同, H2B突变主要发生在折叠结构域, 并且不是已知的PTM位点。H2B E76K/Q突变常见于膀胱癌和头颈癌中, H2B E76K/Q

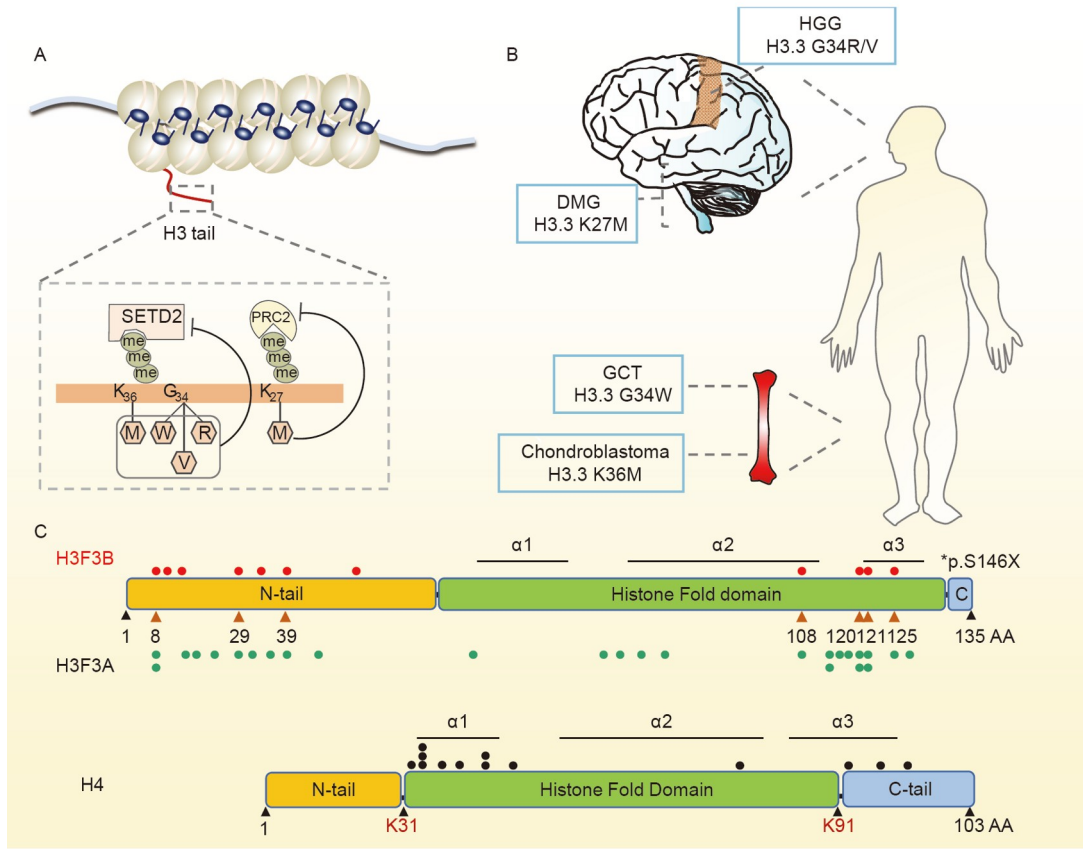


图 2 核心组蛋白与肿瘤/神经发育障碍。A和B: H3突变及诱发肿瘤机制。H3 K27M突变显性抑制PRC2复合物的催化功能, 导致H3K27me2和H3K27me3水平降低; H3 K36M突变显性抑制SETD2的催化功能, 降低H3K36me3水平; H3 G34R/V/W阻断SETD2的结合, 降低H3K36me3水平。C: 神经发育障碍患者携带的H3F3A/B和H4突变^[50,51]。H3F3A和H3F3B共同编码H3.3, 突变位点分布在整个编码序列中, 除*p.S146X(仅发生于H3F3B基因)外, 其他均为错义突变, 两个基因上共有的点突变用棕色三角形标注。H4的突变位点主要在 $\alpha 1$ 和C端区域

Figure 2 Core histones and tumor/neurodevelopmental disorders. A and B: H3 mutations and mechanisms of tumorigenesis. The H3K27M mutation causes reductions in H3K27me2/3 through inhibition of the polycomb repressive complex 2 (PRC2) *in trans*. Similarly, the reduction in H3K36me3 is caused by a dominant effect of the H3K36M mutation via inhibition of histone methyltransferase SET domain containing 2 (SETD2). In contrast, H3G34R/V/W mutations exert their effect on SETD2-mediated H3K36me3 in *cis*. C: H3F3A/B and H4 mutations in patients with neurodevelopmental disorders^[50,51]. H3F3A and H3F3B encode an identical protein H3.3, and the mutations are spread throughout the coding sequence. Except for *p.S146X which is only present in an H3F3B transcript, all others are missense mutations. Shared point mutations on H3F3A and H3F3B genes are marked with brown triangles. H4 mutations are mainly in the $\alpha 1$ and C-terminal domains

会破坏核小体的稳定性, 从而抑制H2A/H2B二聚体的形成。在正常乳腺上皮细胞系MCF10A中过表达H2B E76K/Q可以促进细胞增殖, 并与PIK3CA协同促进集落形成, 导致染色质可及性改变及基因表达失调^[43]。H2B E76位于组蛋白四聚体-二聚体界面处, 并与H4 R92或D68之间形成氢键, H2B E76K/Q突变导致氢键破坏^[44]。H2B E71和H4 K91之间形成盐桥, H4 K91的乙酰化和泛素化也会破坏H2A-H2B和H3-H4二聚体之间的相互作用, 影响染色质组装和DNA损伤修复^[44]。此外, 大约4.5%的胰腺导管腺癌中发现了H2B G53D

突变, G53D突变的癌细胞迁移能力增强, 推测该突变体可能促进染色质解压缩, 致使迁移相关基因异常表达^[45]。此外, 位于核小体锚定位点处——H2A和H2B残基形成的“酸性补丁”区域内的H2B E113是肿瘤中突变频率第二的氨基酸残基, 也是所有组蛋白残基中突变最严重的残基之一, E113K/Q突变可能通过参与核小体重塑、染色质折叠压缩、转录沉默以及DNA损伤修复等多种重要的生物学过程促进肿瘤发生^[46]。

H2A和H4基因并不是肿瘤中最常见的组蛋白突变基因, 然而, H2A的N端和C端尾部是非常特殊的, 突出

的H2A的C端尾部促进染色质高级结构的建立, 尤其是参与H2A-H2B与H3-H4二聚体之间以及H2A与连接DNA之间的相互作用, 且C端尾部的截短与镍(II)诱导的致癌作用相关^[47]. H2A C端最常见的突变残基是E121K/Q, E121突变可能参与破坏染色质高级结构, 但此类突变在肿瘤中的作用仍有待阐明. 不同于其他核心组蛋白, 迄今只发现一种人类特有的H4变体H4G(由H4C7基因编码)^[48]. 与H4不同, 它缺少5个C端氨基酸, 在98个氨基酸序列中只有85%的相似性. H4G定位于核仁, 通过与组蛋白伴侣核仁磷酸蛋白(nucleophosmin 1, NPM1)的相互作用促进核糖体DNA(ribosomal DNA, rDNA)转录. H4G不能形成稳定的核小体, 目前的假说认为, 其与DNA短暂结合并快速解离, 促进染色质的松弛, 从而提高rDNA转录的效率^[48]. H4G在几种癌症中表达上调, 特别是乳腺癌, 其中rDNA转录的增加可能有助于驱动肿瘤的发生^[49].

因此, 肿瘤中核心组蛋白频繁发生的突变可能代表了一种新的表观遗传功能障碍机制, 涉及核小体、染色质纤维以及染色质高级结构的不稳定, 导致染色质可及性的失调和基因表达的改变. 未来尚需进一步研究以充分阐明致癌组蛋白的作用机制, 并评估它们在癌症分类、预后和新疗法开发中的作用.

1.2 核心组蛋白与神经发育障碍疾病

DNA和组蛋白上的表观遗传修饰在细胞分裂时可遗传, 通过表观遗传学机制对神经元进行编程是神经发育的关键, 这一过程的中断或改变会导致神经发育障碍疾病, 包括注意力缺陷多动障碍、孤独症、学习障碍、智力缺陷(intellectual disability; 也称为智力迟滞)、行为障碍、脑瘫、视力和听力障碍以及其他发育迟缓. 最近的高通量测序发现, H3和H4的基因突变导致神经发育障碍疾病^[50,51]. 值得注意的是, 许多突变位于公认的PTM位点之外, 表明破坏组蛋白的核心功能可能是这些神经发育表型的主要分子机制.

体细胞表达的H3变体包括H3.1, H3.2, H3.3和CENP-A, 每种变体具有特定的功能. 其中, H3.1和H3.2是复制依赖的经典组蛋白, 仅在S期表达. H3.1/2在大多数高等真核生物中与H2A, H2B和H4的基因重复排列, 每种组蛋白基因共有10~50个无内含子拷贝^[52]. 而H3.3基因仅有单个拷贝, 通常包含内含子, 并产生经典的含poly(A)尾的mRNAs; 其表达不是复制依

赖的, 长期以来被认为是G0, G1和G2期细胞中主要的H3变体^[53]. 在终末分化细胞中, H3.3取代经典的H3.1和H3.2, 是主要的H3亚型, 例如, 在神经元中, H3.3可以积累到近90%^[54]. H3.3由H3F3A和H3F3B两个基因编码, mRNA序列不同但蛋白序列完全一致, 被特定的伴侣蛋白组装到染色质中^[55,56].

如前所述, 体细胞H3.3突变是已被广泛研究的致癌组蛋白. 2019年, 报道了一例患有继发性小头畸形并伴有严重发育迟缓、智力残疾、生长迟缓和畸形特征的患者携带H3.3 L61R的杂合新发突变(*de novo mutation*)^[57]. 2020年, 一项研究报道了46例携带H3F3A/B新发突变的病人, 他们患有进行性神经功能障碍和先天性异常(Bryant-Li-Bhoj neurodevelopmental syndrome 1, BRYLIB1; 在线人类孟德尔遗传数据库, Online Mendelian Inheritance in Man, OMIM #619720), 但无恶性肿瘤^[51]. 患者运动和社会发展能力延迟, 74%语言能力丧失, 40%不能独立行走, 癫痫发作(49%)、肌张力减退(68%)和眼部异常(55%)在整个队列中也很常见^[51]. 在近20%的病例中, 神经系统存在进行性损伤, MRI显示异常信号, 如胼胝体发育不全和/或脑萎缩^[58]. 这些H3F3A/B突变共有37种, 分布在整个编码序列中, 超过一半的突变位点位于H3.3的N端区域. 其中, R8, R17和K36是甲基化、乙酰化或瓜氨酸化PTM位点^[59], 而其他许多突变位点可能通过改变蛋白-蛋白相互作用而影响染色质重塑因子的结合. 分子建模显示, 37种突变体与DNA、其他组蛋白和组蛋白伴侣的相互作用会被显著破坏. 与体细胞H3.3致癌突变K27M/K36M不同, BRYLIB1相关突变仅导致局部PTM明显异常模式, 而非整体PTM失调, 且使细胞增殖能力增加^[47].

此外, H3.3 G34R基因编辑小鼠发生进行性小头畸形和神经退行性病变, 伴有小胶质细胞异常积聚和神经元损耗. G34R显著降低了H3.3尾部的H3K36me₂, 进而影响DNA甲基转移酶DNMT3A的募集及其在染色质上的重新分布; 与此同时, non-CG(CH)甲基化丢失, 异常CG甲基化沉默启动子, 导致补体和其他天然免疫基因持续表达, 引起神经炎症, 这可能是进行性神经退行性病变的原因^[60].

人类H4由14个H4基因编码, 它们在核苷酸序列上都不相同, 但编码一个相同的蛋白. 自2017年开始, 陆续报道了6个H4基因(H4C3, H4C4, H4C5, H4C6, H4C9和

H4C11)的新发错义突变。所有个体均表现为智力残疾、运动和/或发育迟缓的神经发育障碍特征,并具有多变的非神经系统特征^[50,61]。这些突变聚集在H4的 α -helix 1或C端尾部区域。这两个突变簇都可能破坏核小体内的相互作用以及H4与组蛋白伴侣的结合。在同一H4基因内和不同H4基因之间,突变位点的复发率都较高(图2C),表明特定位点/区域的变异敏感性更高,暗示了这些位点的重要性。某些位点突变后会破坏PTM,比如K31和K91等^[62,63]。而且,有的位点是与组蛋白伴侣相互作用的重要残基,或在DNA损伤修复中发挥重要作用^[64]。迄今,这些突变在神经发育障碍中的作用机制尚不清楚。此外,在H4J基因编码的蛋白上也发现了错义突变K91E,患者表现为严重的智力残疾、小头畸形和面部畸形,在遗传和表型上类似于由H4C基因编码的p.K91A和p.K91Q引起的神经发育障碍。高度重叠的患者表型突出了H4基因之间的功能相似性^[65]。

综上,核心组蛋白突变在神经发育障碍疾病中的作用机制可能涉及多个方面,包括改变PTM位点、核小体动力学的破坏以及与组蛋白伴侣相互作用的丢失等,从而影响基因转录和DNA复制等重要生物学过程。

2 连接组蛋白H1与肿瘤/神经发育障碍

在小鼠和人中,连接组蛋白H1有11种变体,其中5种为复制依赖型,包括HIST1H1A, HIST1H1B, HIST1H1C, HIST1H1D和HIST1H1E(也分别注释为H1-1, H1-5, H1-2, H1-3和H1-4, 或H1A, H1B, H1C, H1D和H1E),在体细胞中普遍表达;2种为复制非依赖型,包括H1FX和H1F0(也分别注释为H1-10和H1-0),主要在终末分化细胞中表达;其余4种包括HIST1H1T, H1FNT, H1LS1和H1FOO(也分别注释为H1-6, H1-7, H1-9和H1-8)则特异表达于生殖细胞^[66]。H1变体之间表现出高度的跨物种保守性,说明不同H1变体具有重要作用,并在进化压力中被保留下来。

H1主要有三个结构域:短的N端结构域(N-terminal domain, NTD)、中央球状结构域(globular domain, GD)和长的C端结构域(C-terminal domain, CTD)。球状结构域内带正电荷的Lys和Arg残基将H1以on-dyad或off-dyad的结合模式锚定在核小体的二元轴上或附近,对于染色质折叠和压缩非常重要^[67]。C端高度富含Lys,

是H1结合基因组DNA并压缩形成30纳米染色质纤维以及染色质高级结构的重要结构域^[68]。H1对于维持染色质高级结构中的染色质环、TADs间相互作用非常重要,并能调控局部染色质功能。相比于野生型, *H1c/H1d/H1e*敲除的小鼠胚胎干细胞中, TADs的位置及规模虽未发生显著改变,但TADs内染色质环数量、TADs内部结构域之间(inter-domain)以及TADs之间(inter-TAD)的相互作用显著增加; inter-TAD相互作用的增加与细胞表观遗传图谱改变相关性高,包括活跃转录的组蛋白标记H3K4me1和H3K4me3的改变, DNA酶高敏感位点(DNase-hypersensitivity sites, DHSs)的增加,以及DNA CpGs甲基化水平的降低^[69]。

H1变体受到复杂的PTMs调节,包括磷酸化、甲基化和乙酰化等,这些修饰可能对其二级结构或相互作用蛋白的结合产生影响,从而改变染色质的结构和功能^[68]。CTD内的保守序列S/TPXK/R可被丝氨酸激酶磷酸化,并参与调控细胞周期中染色质动态压缩过程^[70]。GD内也有多个影响相互作用蛋白招募的PTM位点,继而对H1-核小体复合物的稳定性产生影响,例如, H1E K26的甲基化有助于HP1募集,促进异染色质的形成^[71]。H1K26ac能够与组蛋白去乙酰化酶SIRT1结合,降低H3K9ac, H4K16ac和H1的乙酰化修饰,从而使H3K79me2的水平降低,促进兼性异染色质形成^[72]。

H1与H3K27甲基化具有相关性, PRC2更倾向于选择含H1的寡核小体作为催化底物,并且H1还可以反过来刺激其酶活性以促进H3K27甲基化的建立。双核小体和寡核小体生化实验发现H1介导压缩染色质可以帮助PRC2接触到染色质中不相邻的核小体,进而促进H3K27甲基化在H1压缩的染色质纤维中的传播和继承^[73]。本团队研究发现, H1在介导形成30纳米染色质纤维高级结构的过程中,通过核小体-核小体配对原则和正反馈机制,促进DNA复制过程中由RYBP-PRC1介导的H2AK119ub1以及PRC2介导的H3K27me3的快速蔓延和恢复^[74,75]。此外, H1还可以直接与DNA甲基转移酶DNMT1和DNMT3B相互作用,促进增强子区域的DNA甲基化^[76]。

值得注意的是, H1与染色质的结合具有动态性。荧光漂白恢复(fluorescence recovery after photobleaching, FRAP)结果表明,相比核心组蛋白需要数小时的恢复时间, H1仅在3~4分钟内即可恢复。不同H1变体与染色质的亲和力也有显著差异,其中, CTD是影响

H1与核小体之间亲和力的主要因素, CTD越短, 亲和力越弱, 动态性越强^[77]。此外, H1与染色质的结合也会受到具有相似DNA结合结构域的蛋白家族的竞争抑制作用, 如具有相似的翼状螺旋DNA结合结构域的先鋒转录因子(pioneer transcriptional factors)和高迁移率家族蛋白(high mobility group proteins, HMG)^[78,79]。而且, MeCP2和H1之间存在共同的核小体结合位点, 并通过竞争方式调节染色质压缩状态^[80]。H1与染色质结合的动态性可以保证DNA相关的生物学过程(转录、DNA复制和损伤修复等)的顺利进行, 促进兼性异染色质瞬时或伴随发育过程的压缩。

2.1 连接组蛋白H1错义突变与生发中心样弥漫性大B细胞淋巴瘤

近年来, 多项研究相继观察到非霍奇金B细胞淋巴瘤, 特别是生发中心来源的DLBCLs中H1基因高频突变, 此类突变主要发生在HIST1H1B(8%), HIST1H1C(12%), HIST1H1D(8%)和HIST1H1E(18%)基因^[81], 而且, 这四种H1基因的非同义突变或插入突变与同义突变的比例是5.1:1, 远高于随机突变^[82~84]。此外, 泛癌数据库(The Cancer Genome Atlas, TCGA)中H1基因突变发生的主要肿瘤类型为B细胞淋巴瘤, 充分证明生发中心样DLBCLs与H1基因突变之间存在显著相关性^[83]。

生发中心反应B细胞表现为: 快速增殖、细胞周期检查点调控被抑制、诱导激活胞苷脱氨酶并引起体细胞超频突变、代谢重编程以及末端分化下调。这些典型的“癌症特征”促使B细胞对抗原刺激产生响应, 进而产生多样化的抗体并促进细胞增殖, 同时也使B细胞发生癌变的风险大大增加^[85]。近期, 根据其不同生物学特征以及突变图谱, DLBCLs分类方式也随之更新, 其中, MCD/5型DLBCLs在临床中最为难治, 具有MYD88, CD78B和H1等基因突变特征, 并表现为淋巴结外弥散以及增殖速度快的特点^[86]。

生发中心B细胞对H1基因突变表现出更高敏感性的原因: 首先, 体细胞中表达的7种H1变体中, 只有H1B, H1C, H1E在鼠B细胞中表达, 导致B细胞对H1基因突变的整体容错能力降低^[87]; 其次, B细胞分化发育过程较复杂, 并伴随大规模的三维基因组结构重塑和转录改变, 对H1的依赖性较强^[88,89]; 最后, B细胞中H1可以协同相互作用蛋白参与生发中心反应, 尤其是

已知的对生发中心反应比较重要的PRC1, PRC2以及与核周异染色质形成相关的核纤层蛋白Lamin B1^[90,91]。

H1基因突变促进生发中心B细胞恶性转化的机制总结为以下三个方面。

第一, H1错义突变可能改变染色质高级结构, 从而促进B细胞淋巴瘤的发生发展。近期有研究利用H1敲除小鼠模型, 发现H1缺失会破坏小鼠生发中心B细胞的三维基因组结构, 导致生发中心B细胞过度增殖, 并伴随染色质从B区室向A区室的转移以及染色质可及性增加, 表观遗传标记重塑, 干细胞相关基因异常表达。此外, 处于中间状态的I区室会在野生型B细胞进一步分化后恢复为B区室; 而H1缺失的生发中心B细胞中, A区室不断扩张, 导致H3K27me3的区域减少(图3B)^[10,83,87]。但是, 在人淋巴瘤中, H1主要为错义突变而非缺失突变, 并且突变位点主要集中在与核小体DNA结合的GD和CTD结构域(图3A), 因此, 推测这些突变可能对H1与核小体的结合产生影响, 从而导致染色质折叠的异常。体外液相色谱分析和细胞内荧光漂白恢复实验也发现, 位于H1 GD结构域内的GS突变会减弱其ASGS基序与基因组DNA大沟的结合紧密程度, 使其动态性显著增强, 推测该突变会破坏H1与染色质的结合, 导致功能丧失(loss-of-function, LOF)^[83,92]。但病理条件下发生的GS突变是否还通过其他机制促进B细胞淋巴瘤的发生发展仍有待研究。不同于H1 GD内的突变削弱其与基因组DNA的结合, CTD内的突变可能影响染色质动态压缩, 如H1E A164可以削弱其与DNMT3B之间的相互作用, 降低整体基因组DNA的甲基化水平^[82]。

第二, PRC复合物在经历生发中心反应的B细胞中发挥重要作用, H1可以通过与PRCs相互作用影响B细胞淋巴瘤发生发展。经历生发中心反应的B细胞主要利用非经典PRC1, 上调非经典PRC1.1的BCOR, PCGF1和KDM2B亚基, 降低经典PRC1亚基BMI1和PHC1-3。而且经典PRC1亚基BMI1或CBX7在恶性B细胞淋巴瘤中高表达, 导致抑癌基因被异常抑制, 并且成为淋巴瘤不良预后的标志^[93]。PRC2的完整性和动态性对于生发中心的形成也至关重要, PRC2催化亚基EZH2在中心母细胞中上调, 抑制细胞周期检查点基因(CDKN1A)和分化基因表达, 并伴随B细胞退出生发中心反应而下调^[94]。此外, H1也可以通过影响H2A尾部

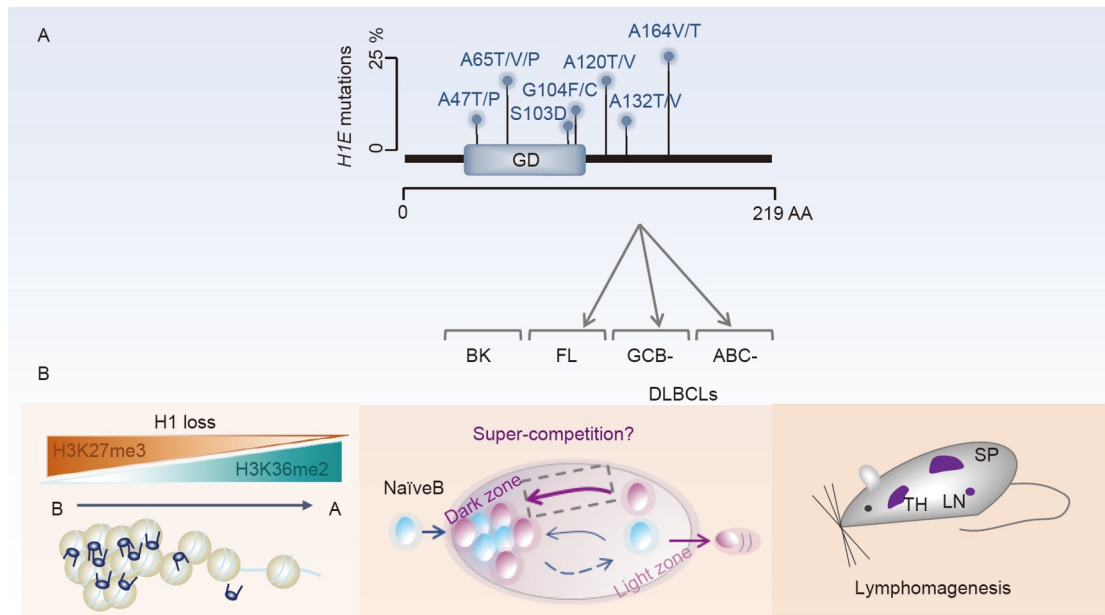


图 3 连接组蛋白H1错义突变与生发中心样弥漫性大B细胞淋巴瘤。A: 人B细胞淋巴瘤中的H1E错义突变。突变位点主要集中在与核小体DNA结合的中央球状结构域和C端结构域。H1E基因突变主要发生在滤泡性淋巴瘤, 生发中心B细胞样弥漫性大B细胞淋巴瘤和活化B细胞样弥漫性大B细胞淋巴瘤; 伯基特淋巴瘤中几乎不发生H1E基因突变; B: H1缺失会破坏小鼠生发中心B细胞中三维基因组结构, 使染色质从B区室转向A区室, 并伴随表观遗传标记的重塑以及干性基因上调, 细胞获得增殖优势进而促进小鼠脾脏(用SP表示)、淋巴结(用LN表示)和胸腺(用TH表示)等外周淋巴器官中淋巴瘤的发生

Figure 3 Linker histone H1 missense mutations and germinal center B-cell-like diffuse large B-cell lymphoma (GCB-DLBCLs). A: H1E missense mutations in human B-cell lymphoma. The mutation sites are mainly located in the GD and CTD. H1E mutations frequently occur in follicular lymphomas (FLs), GCB-DLBCLs and activated B-cell-like diffuse large B cell lymphomas (ABC-DLBCLs), but are rare in Burkitt's lymphomas (BLs); B: H1 deficiency in mice disrupts the three-dimensional chromosomal architecture in germinal center B cells and causes chromatin transition from compartment B to A, accompanied by remodeling of epigenetic landscape and upregulation of pluripotency genes, and lymphomagenesis in peripheral lymphoid organs such as spleen (represented as SP), lymph nodes (represented as LN) and thymus gland (represented as TH)

的空间位置, 便于PRC1复合体接近, 而且C端缺失的H2A也与淋巴瘤发生密切相关^[95]。PRC家族之间存在相互促进作用, 如H2AK119ub1可以招募PRC2并促进H3K27me3形成, PRC1和PRC2都对H1介导压缩的染色质具更高的亲和性, 而H1也可以促使H2AK119ub1和H3K27me3在压缩的染色质纤维中传播或继承^[74,75]。因此, 生发中心B细胞发育对H1和PRCs参与的表现遗传调控具有非常强的依赖性。

第三, H1错义突变可能影响细胞基因组的稳定性, 促进细胞异质性。基因组不稳定是肿瘤发展的内在机制, 其随机产生的突变是肿瘤异质性的原因^[96-98]。作为重要染色质结构蛋白, H1在核小体稳定、高级染色质压缩、基因表达抑制过程中发挥重要功能。DLBCLs不同亚群突变基因分析发现, 在含H1基因突变的4型DLBCLs亚群中, 整体基因突变密度显著升高($P<0.0001$)^[99], 推测H1错义突变可能影响细胞基因组

的稳定性, 促进细胞异质性。染色质中H1含量异常会影响DNA复制叉进程、基因转录及基因组稳定性。例如, 在H1缺失的黑腹果蝇、酵母和小鼠胚胎细胞中, 都观察到全基因组内DSB频率升高^[100-102]。此外, 不同H1修饰可以通过DNA损伤应答反应参与基因组稳定性, 染色质结构动态改变可以调节毛细血管共济失调突变基因(ataxia telangiectasia mutated, ATM)起始的信号级联反应激活^[103]。正常情况下, H1抑制ATM的招募和激活; 而DNA损伤后, H1.2的K64泛素化或PARP1介导的H1.2 CTD结构域内发生的聚ADP核糖基化, 可以促使其从染色质上解离, 染色质解压缩, 进而募集下游蛋白激活ATM反应^[104,105]。

综上所述, H1在染色质折叠和压缩过程中发挥的结构性或功能调控作用揭示了生发中心B细胞对H1的功能具有特殊的依赖性。进一步对不同H1错义突变在单核小体和染色质不同层次结构的动态调控进行分

析, 同时在细胞和个体水平上研究H1错义突变对B细胞分化发育或B细胞淋巴瘤发生发展的影响, 对于理解B细胞淋巴瘤的病理学机制、临床诊断分型以及新型治疗策略和创新药物的研发具有重要意义。

2.2 连接组蛋白H1与HIST1H1E综合征

自2017年以来, 临床上陆续报道了一类与连接组蛋白HIST1H1E基因突变相关的神经系统疾病, 并被命名为HIST1H1E综合征(HIST1H1E syndrome, 也称Rahman综合征, OMIM #617537), 这些患者主要表现为不同程度的智力缺陷、焦虑、恐惧症、强迫性行为、注意力缺陷、多动障碍和孤独症谱系障碍等特征^[17,106,107]。迄今已报道的HIST1H1E综合征患儿携带21种移码突变, 均位于H1E CTD结构域的99-bp区域内(图4A)。其中, 20种突变新产生含相同38个氨基酸的C端截短突变蛋白。在唯一1例中枢性睡眠呼吸暂停症的病人中, 新形成25个氨基酸序列^[17]。与野生型蛋白相比, 移码突变的H1E C端氨基酸序列净正电荷大量减少, 新产生的38个氨基酸序列能否对H1压缩染色质的功能产生影响, 或者导致功能获得(gain-of-function, GOF)是重要的科学问题。目前的机制研究表明, 患者全基因组DNA甲基化水平降低; 在患者皮肤成纤维细胞中, H1E移码突变不影响蛋白在核内与染色质的结合, 但H3K4me2, H3K9me3, H3K27me3以及异染色质标记分子HP1 β 表达下降, 染色质更松散, 细胞增殖受影响, 很难进入S期并伴随加速衰老^[108,109]。HP1可以结合H3K9me3, 自我寡聚并招募H3K9甲基转移酶; H3K9me3, HP1 α 和HP1 β 标记的异染色质具有明显的区域特征, 这种异染色质会在DNA甲基化抑制下丧失, 转变为富含H3K9me2的状态, 从而显著改变基因组的结构域组成状态^[110]。此外, 在重度抑郁患者中也发现了H1家族其他变体的表达水平下降(H1A和H1E), 进一步提示H1介导的染色质结构失调在神经系统性疾病的发生发展中具有重要作用^[111]。

H1E C端移码突变除影响其染色质压缩功能外, 也影响其与其他蛋白之间的相互作用, 如FACT, Nap1L1, ANP32E等组蛋白伴侣等, 使染色质高级结构发生异常^[112]。而且, 移码突变也会影响H1E C端PTMs的功能, 例如, Ser187的磷酸化可以促进rDNA的转录, 并促进RNA pol I和RNA pol II依赖的雌激素受体相关基因的特异性激活^[113]。在大鼠神经元中过表

达含致病突变的H1E会导致大鼠神经元细胞转录异常, 差异表达基因分析发现其主要富集于突触信号传导、生物黏附、行为、运动和离子转运调节等相关生物学过程中^[114]。

神经发育障碍疾病的病因非常复杂, 其潜在的致病机制尚不清楚。神经干细胞中染色质的异常折叠导致神经发育过程中兴奋性和抑制性神经元的失衡被认为是潜在的病理机制之一^[115]。随着更多的HIST1H1E综合征患者被发现, 疾病表型和基因变异图谱将得到更好的描述, 深入探究染色质动态调控在疾病发生发展中的病理学功能将具有重要意义。

3 甲基CpG结合蛋白2与神经发育障碍

作为一种重要的表观遗传调控机制, DNA甲基化被甲基化CpG结合域(methyl-CpG DNA binding domain, MBD)蛋白家族识别, 进而招募特定的染色质修饰因子发挥作用。MeCP2是MBD家族成员, 研究最为广泛。人MeCP2基因位于X染色体上, 由于翻译起始位点不同和可变剪接, MeCP2蛋白有两种亚型: MeCP2e1和MeCP2e2, 这两种亚型的序列仅在N端略有不同。MeCP2蛋白由N端结构域、甲基化CpG结合域、中间结构域(intervening domain, ID)、转录抑制域(transcription repression domain, TRD)和C端结构域组成。另外, 在ID, TRD和CTD中有三个AT-钩状结构域(AT-hook domain), 该区域通过核心短序列与富含AT碱基的DNA小沟结合(图4B)^[116]。MBD特异性识别和结合5-甲基胞嘧啶(5-methylcytosine, 5mC)和5-羟甲基胞嘧啶(5-hydroxymethylcytosine, 5hmC), 而TRD能够结合多个转录抑制因子, 从而抑制基因表达。TRD结构域中有与转录共抑制因子NCoR/SMRT相互作用区域(NCoR/SMRT interaction domain, NID)^[117]。然而, MeCP2也被证明在下丘脑中结合在基因启动子区并与CREB1相互作用激活基因表达^[118]。NTD中的部分区域也能自主结合DNA, MBD结构域对其与DNA结合的亲和力也有增强作用。CTD包含DNA结合域和染色质特异结合域, 主要负责介导核小体阵列的压缩。而ID和TRD表现出非甲基化依赖的DNA结合能力^[119]。

MeCP2在神经发育过程中具有广泛的功能, MeCP2基因突变引起的功能丧失与雷特综合征(Rett syndrome, RTT; OMIM #312750)密切相关。RTT是一种

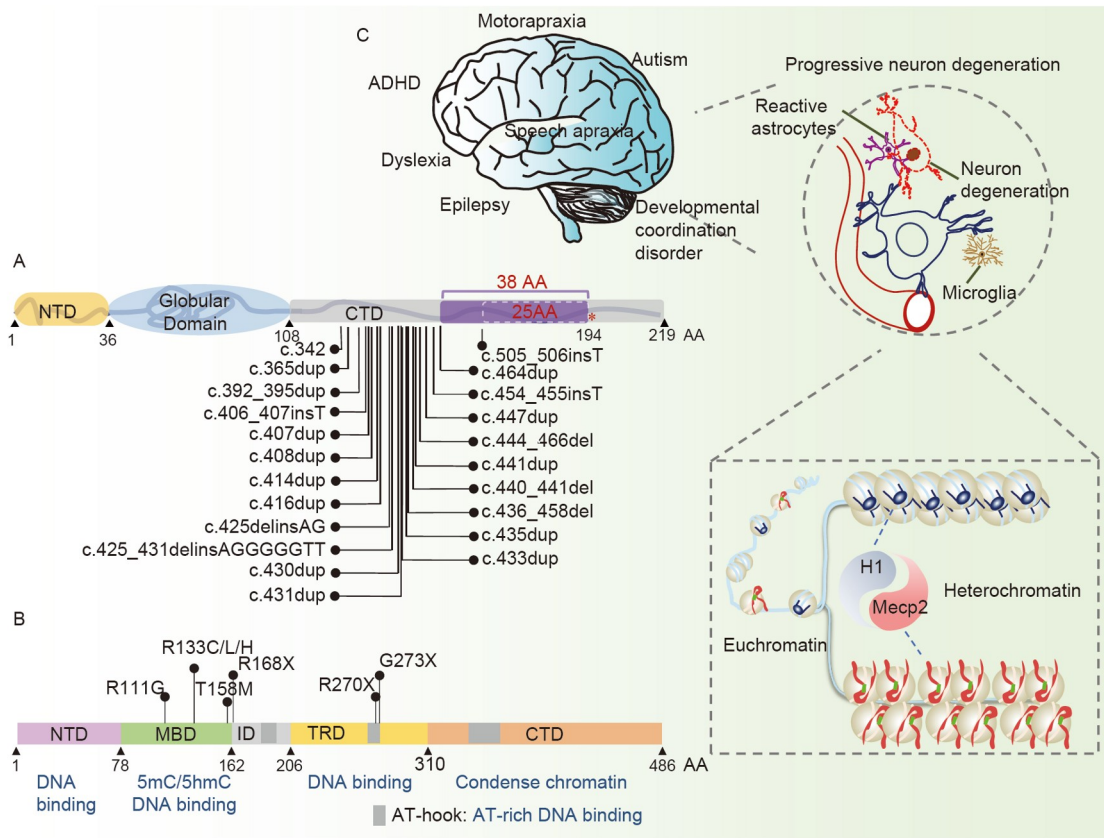


图 4 连接组蛋白H1、甲基CpG结合蛋白2与神经发育障碍。A: HIST1H1E致病变异示意图。野生型H1E蛋白大小为219 AA, 移码蛋白大小均为194 AA。38 AA: 单碱基的插入或缺失导致H1E C末端阅读框发生移码突变的同时新产生一段含有38个氨基酸的相同序列。25 AA: 一例病例的突变蛋白的C末端产生了25个与其他突变相同的氨基酸序列^[17]; B: MeCP2蛋白结构域及RTT致病点突变。MeCP2 MBD结构域特异性识别5mC和5hmC DNA, 其他结构域也都有结合DNA的能力。RTT致病突变集中在MBD和TRD结构域; C: 神经发育障碍可能与神经元进行性衰退有关, 神经元富含H1与MeCP2蛋白, 两者在异染色质区域可以竞争结合核小体

Figure 4 Linker histone H1, MeCP2 and neurodevelopmental disorders. A: Schematic diagram of HIST1H1E pathogenic mutations. Wild-type *HIST1H1E* gene encodes a protein with 219 amino acids. Each frameshift mutation in the CTD domain introduces a premature stop codon resulting in a truncated protein with 194 amino acids. 38 AA: CTD variants caused by single-base insertion or deletion share the same last 38 amino acids in the C terminus. 25 AA: The recently identified variant creates a new 25 aa sequence which is shared with the 38 aa sequence^[17]; B: schematic diagram of MeCP2 domains and Rett-causing mutations. The MBD domain binds preferentially to 5mC- and 5hmC-containing DNA, and other domains also have DNA binding capacities. Rett-causing mutations are mainly localized in the MBD and TRD domains; C: neurodevelopmental disorders are associated with progressive neurodegeneration. H1 and MeCP2, highly enriched in neurons, can compete for binding to nucleosomes in heterochromatin regions

多发生于女性的严重的进行性神经发育障碍疾病, 而在男性中, MeCP2功能障碍通常会导致患者在两岁前死亡。大约每一万名新生女婴中有1人发生这种疾病, 患者在出生后6~18个月才会出现症状。这种疾病的特征为发育倒退, 包括认知和运动发育受阻, 如丧失语言功能和刻板的手部运动。患者还表现出小头畸形、癫痫、呼吸紊乱、肌张力低下和孤独症等表型。MeCP2基因还会发生重复, 导致MeCP2重复综合征 (MeCP2 duplication syndrome, MDS; OMIM #300260), 主要表现为运动发育落后、语言缺失或落

后、严重智力障碍、癫痫、孤独症样表现等^[120]。因此, MeCP2功能的丧失或冗余都表现为严重的神经系统异常。

在成人脑中, MeCP2是含量最丰富的MBD家族蛋白。在小鼠胚胎发育早期MeCP2蛋白表达水平很低, 而在发育后期增加, 在成熟神经元中达到较高水平。人MeCP2的表达从妊娠中期开始, 并持续增加到10岁^[121]。MeCP2的表达增加发生在神经发生和分化完成后, 并与神经元成熟阶段相一致^[122], 这种表达模式被认为是RTT患者早期发育正常的原因。虽然MeCP2

首先被描述为一种经典的转录抑制因子^[123],但是现有的一些研究表明,MeCP2也通过三维基因组和表观基因组等非常复杂的机制发挥调控作用^[124,125]。在神经发育和成年期,MeCP2的表达受严格调节。有研究发现,MeCP2广泛地结合在神经元基因组中,并且在纯化的小鼠脑细胞核中,MeCP2的表达水平接近组蛋白八聚体水平;在MeCP2缺乏的神经元中,H1的水平相比于野生型增加了一倍^[126]。而且,体外研究证明与核小体摩尔比为1:1的MeCP2能够压缩折叠核小体串珠形成规则的椭圆形结构,展现出与H1相似的压缩染色质的能力^[127],表明MeCP2是一个类似于H1的染色质结构蛋白。同时,MeCP2与H1结合核小体的位置一致,荧光各向异性研究发现,MeCP2可以取代染色质中的H1,并整体改变染色质状态^[80]。MeCP2通过改变染色质高级结构调控印记基因*DLX*(distal-less homeobox)的表达,野生型小鼠脑中MeCP2介导的沉默染色质环位于基因*Dlx5-Dlx6*处,而MeCP2缺失的小鼠脑中没有这种结构;同时在RTT患者的淋巴母细胞中,*DLX5*基因印迹丢失^[128]。MeCP2缺乏通过增加H3乙酰化影响整体染色质的组成和状态^[126]。MeCP2还可以增加*SIRT1*基因启动子上的H3K9me2^[129],抑制MeCP2可以降低沉默基因启动子H3K27me3的水平^[130],表明了MeCP2在染色质折叠中的作用。

MeCP2的RTT致病点突变包括错义突变、无义突变和移码突变,主要位于MBD和TRD结构域(图4B)^[18],这些突变造成的影响与突变的类型和位置有关。本团队^[131]研究证明,MeCP2可以诱导体外核小体串珠的压缩以及液液相分离,而RTT突变会破坏MeCP2介导的染色质相分离,良性突变对这一过程几乎没有影响;MeCP2和H1分别诱导形成不相容的染色质相分离,说明二者可以竞争结合核小体,从而揭示RTT的病理学机制(图4C)。MBD中的大部分突变降低了MeCP2与5mC的结合亲和力,从而导致细胞内异染色质的组织和功能受损^[132]。位于MBD的R111和R133是结合5mC的关键氨基酸残基,其中任何一个位点的突变都会影响MeCP2在异染色质中的定位。MeCP2 R111G是仅在一患者中发现的罕见RTT突变,它使MeCP2不再定位于异染色质^[132]。MeCP2 R133的不同突变有不同的影响,R133C和R133L会减少MeCP2在异染色质的富集,而R133H则对MeCP2在异染色质的富集起促进作用^[133]。MeCP2 R270X和MeCP2 G273X的小鼠有不同

的发育速度和疾病表型。MeCP2 G273X突变小鼠的存活时间比MeCP2 R270X小鼠长,这是由于MeCP2 G273X保留了AT钩状结构域(aa 264~273)。AT钩状结构域的丢失导致在神经系统中染色质重塑蛋白ATRAX的错误定位^[134],从而导致疾病的发生。MeCP2 R168X缺失TRD和CTD区域,雄性和雌性小鼠均表现为典型的RTT表型,但其具体机制尚不清楚。虽然整个MBD被保留,但MeCP2 R168X折叠核小体串珠形成染色质高级结构的能力受损^[127]。所以MeCP2对染色质高级结构的形成、稳定和功能起着至关重要的作用。

4 其他染色质结构蛋白与肿瘤/神经发育障碍

在非典型RTT综合征中也发现了除MeCP2外其他基因的突变,如FOXG1。携带FOXG1突变的患者从婴儿早期就发病,且没有发育倒退的表型,与MeCP2突变引起的RTT综合征的临床表型有很大差别,称为FOXG1综合征(FOXG1 syndrome)^[135]。FOXG1是FOX家族成员,含高度保守的、与H1 GD高度相似的DNA结合结构域,可作为先锋转录因子识别H1压缩的染色质中的结合位点,打开局部染色质或核小体区域^[136]。在小鼠海马中,FOXG1主要结合于增强子区域,与NEUROD1协同作用,通过改变H3K27ac/H3K4me3标记和染色质可及性影响突触发生、轴突形成和神经元成熟^[137]。FOXG1与染色质的结合有静态和动态两种状态,静态结合时,FOXG1作为结构蛋白参与染色质组装从而抑制基因表达,而通过动态结合,FOXG1直接参与基因表达调控。即使是FoxG1的小部分缺失,例如截短突变FoxG1 Y400X,也会导致其与染色质的亲和力急剧下降,随着缺失的延长,FoxG1染色质结合的亲和力进一步减少;然而,这种亲和力的进行性降低似乎与患者的症状严重程度无关,可能是由于稳定结合部分不受染色质亲和力的降低和蛋白缺失程度的影响^[138]。因此,FoxG1与染色质结合动力学的变化可能与FOXG1突变的致病机制有关。

HMG蛋白家族是一种广泛存在于真核细胞中的非组蛋白染色质结构蛋白,包括HMGN, HMGB和HMGA,与H1竞争结合染色质,参与染色质解压缩,增加染色质可及性。HMG蛋白与染色质结合的位置与H1重叠,也与核小体连接DNA结合,且定位于核小体的

进出口位置. 不同于H1, HMG蛋白介导了较为松散的染色质结构, 从而导致局部染色质不稳定, 同时可能招募其他蛋白, 导致转录激活^[139]. HMGB1在肿瘤发生和治疗中发挥双重作用. 一方面, 慢性炎症反应导致的HMGB1过表达促进肿瘤发生. 另一方面, HMGB1在肿瘤放化疗和免疫治疗中具有保护作用. 由于HMGB1基因缺失能导致基因组不稳定和肿瘤发生, 其在染色体结构中的作用可能为癌症治疗提供新的治疗靶点^[140].

FACT(facilitates chromatin transcription)是与转录延伸相关的组蛋白伴侣, 由SPT16和SSRP1亚基组成, 具有拆散和组装核小体的能力, 在基因转录、DNA复制和修复等过程中发挥保持染色质结构完整性的作用^[141]. 本团队的研究表明, FACT在单核小体水平上破坏并保持其完整性^[142], 并且也可以作用于四核小体单

元促进核小体的解聚^[143], 揭示FACT对核小体结构稳定性和完整性的双重调控功能, 并在维持或保留旧的组蛋白表观遗传标记中发挥重要作用. FACT在不同细胞类型中表达差异很大, 在干细胞和癌细胞中的表达很高. 大量研究表明, 多种肿瘤中FACT的表达增高, 特别在侵袭性癌症中表达显著升高, 并与肿瘤耐药和不良预后相关^[144].

5 总结与展望

综上所述, 组蛋白及其变体、MeCP2和其他染色质结构蛋白参与核小体、30纳米染色质纤维、染色质环、染色质区室和异染色质等各个层次三维染色质结构的形成、稳定和动态调节, 这些染色质结构蛋白的变异导致肿瘤或神经发育障碍疾病(表1). 目前已知的

表 1 染色质结构蛋白的主要突变和疾病类型

Table 1 The main mutations and disease types of chromatin architectural proteins

蛋白	基因	突变	结构域	疾病类型或特征	参考文献
H3.3	H3F3A	K27M	N端结构域	弥漫性中线神经胶质瘤	[33]
		G34R/V/W/L		恶性胶质瘤	[30]
		R8S/G, R17G, K36E			[51]
	H3F3B	L61R	组蛋白折叠结构域	进行性神经功能障碍和先天性异常	[57]
		R8C			[51]
		G34R/V	N端结构域	严重神经发育综合征	[60]
H2B	HIST1H2BH/J/E	K36M	组蛋白折叠结构域	软骨母细胞瘤	[34]
		G53D		胰腺导管腺癌	[19,45]
	HIST1H2BB/C/D/E/F/G/H/I/K/L/O	E76K/Q		膀胱癌症与头颈部鳞状细胞癌	[19,43]
		E113K/Q			[19,46]
H2A	HIST1H2AI/K/L/B/C/A/H/J/M	R29L/P/Q/W	C端结构域	癌肉瘤	[19]
		A103D/E/G/P/S/T			
	HIST1H2AC/E/J/K/L/M/B/C/D/H	E121D/K/Q			
H4	H4C	K31T	N端结构域	神经发育障碍	[37]
		K91A/Q	组蛋白折叠结构域	严重的智力残疾、小头畸形和面部畸形	[50]
	H4J	K91E			
H1	HIST1H1B/C/D/E	A47T/P, A65T/V/P, S103D, G104F/C	组蛋白折叠结构域	滤泡性淋巴瘤、弥漫性大B细胞淋巴瘤	[63,64]
H1E		A120T/V, A132T/V, A164V/T	C端结构域	拉赫曼综合征	[17,76,77]
MeCP2	MeCP2	R111G, R133C/L/H R168X R270X, G273X	甲基化CpG结合域 中间结构域 转录抑制域	雷特综合征	[18]

疾病发生相关机制主要包括三个方面: (i) 组蛋白的突变破坏PTMs位点, 导致异常PTM模式, 抑制组蛋白甲基转移酶的活性, 影响DNA甲基转移酶、组蛋白甲基转移酶和/或组蛋白伴侣等在染色质上的募集; (ii) 染色质结构蛋白的突变影响核小体的完整性、稳定性和DNA-核小体相互作用, 破坏核小体动力学, 导致染色质纤维和三维染色质结构的异常, 以及染色质可及性的失调; (iii) 染色质结构蛋白的突变影响基因组稳定性、DNA复制、转录、损伤修复, 进而导致基因表达异常。

基于领域进展, 未来需要进一步研究: 染色质结构蛋白、表观遗传因子(组蛋白PTMs和DNA甲基化

等)与基因表达之间相关性的深入分子机制; 染色质结构蛋白突变如何启动病理性基因表达程序, 例如, 致癌组蛋白H3K27M对致癌信号通路的激活和调节机制; H3.3的体细胞突变和生殖细胞原发突变分别导致肿瘤和神经发育障碍疾病的原因; 突变位点所在的结构域虽不同, 但引起的疾病表型相似, 通过影响核小体稳定性或相互作用蛋白导致疾病发生的分子机制的共性。总之, 对染色质结构蛋白变异的更全面深入的理解将对发育和人类疾病中的染色质生物学产生广泛而深远的影响, 对于疾病的临床分型、精准诊治, 以及开发表观遗传编辑和干预手段具有重要的意义。

参考文献

- 1 Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev*, 2011, 21: 175–186
- 2 Song F, Chen P, Sun D, et al. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science*, 2014, 344: 376–380
- 3 Jerkovic I, Cavalli G. Understanding 3D genome organization by multidisciplinary methods. *Nat Rev Mol Cell Biol*, 2021, 22: 511–528
- 4 Bintu B, Mateo L J, Su J H, et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, 2018, 362: eaau1783
- 5 Wang H, Xu X, Nguyen C M, et al. CRISPR-mediated programmable 3D genome positioning and nuclear organization. *Cell*, 2018, 175: 1405–1417
- 6 Kim J H, Rege M, Valeri J, et al. LADL: light-activated dynamic looping for endogenous gene expression control. *Nat Methods*, 2019, 16: 633–639
- 7 Li G. The 3D organization of genome in the nucleus: from the nucleosome to the 4D nucleome. *Sci China Life Sci*, 2020, 63: 791–794
- 8 Cremer T, Cremer M. Chromosome Territories. *Cold Spring Harb Perspect Biol*, 2010, 2: a003889
- 9 Wang S, Su J H, Beliveau B J, et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*, 2016, 353: 598–602
- 10 Vilarrasa-Blasi R, Soler-Vila P, Verdager-Dot N, et al. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. *Nat Commun*, 2021, 12: 651
- 11 Dixon J R, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 2012, 485: 376–380
- 12 Dekker J, Mirny L. The 3D genome as moderator of chromosomal communication. *Cell*, 2016, 164: 1110–1121
- 13 Collombet S, Ranisavljevic N, Nagano T, et al. Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature*, 2020, 580: 142–146
- 14 Tan L, Ma W, Wu H, et al. Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development. *Cell*, 2021, 184: 741–758
- 15 Thapar A, Cooper M, Rutter M. Neurodevelopmental disorders. *Lancet Psychiatry*, 2017, 4: 339–346
- 16 Bunting K L, Soong T D, Singh R, et al. Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity*, 2016, 45: 497–512
- 17 Indugula S R, Ayala S S, Vetrini F, et al. Exome sequencing identified a novel HIST1H1E heterozygous protein-truncating variant in a 6-month-old male patient with Rahman syndrome: a case report. *Clin Case Rep*, 2022, 10: e05370
- 18 Krishnaraj R, Ho G, Christodoulou J. RettBASE: Rett syndrome database update. *Hum Mutat*, 2017, 38: 922–931

- 19 Nacev B A, Feng L, Bagert J D, et al. The expanding landscape of ‘oncohistone’ mutations in human cancers. *Nature*, 2019, 567: 473–478
- 20 Strahl B D, Allis C D. The language of covalent histone modifications. *Nature*, 2000, 403: 41–45
- 21 Millán-Zambrano G, Burton A, Bannister A J, et al. Histone post-translational modifications — cause and consequence of genome function. *Nat Rev Genet*, 2022, 23: 563–580
- 22 Cao J, Yan Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front Oncol*, 2012, 2: 26
- 23 Zhao H Y, Xu D Y. Pathway choice for DNA double strand break repair (in Chinese). *Sci Sin Vitae*, 2021, 51: 56–69 [赵华煜, 徐冬一. DNA双链断裂修复途径的选择与调控. 中国科学: 生命科学, 2021, 51: 56–69]
- 24 Chen J Y, Lu L Y. BRCA1 and DNA damage response (in Chinese). *Sci Sin Vitae*, 2022, 52: 1763–1772 [陈绩源, 陆林宇. BRCA1与DNA损伤修复调控网络. 中国科学: 生命科学, 2022, 52, 12: 1763–1772]
- 25 Audia J E, Campbell R M. Histone modifications and cancer. *Cold Spring Harb Perspect Biol*, 2016, 8: a019521
- 26 Buschbeck M, Hake S B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol*, 2017, 18: 299–314
- 27 Gonzalez-Perez A, Jene-Sanz A, Lopez-Bigas N. The mutational landscape of chromatin regulatory factors across 4,623 tumor samples. *Genome Biol*, 2013, 14: r106
- 28 Kallappagoudar S, Yadav R K, Lowe B R, et al. Histone H3 mutations—a special role for H3.3 in tumorigenesis? *Chromosoma*, 2015, 124: 177–189
- 29 Brien G L, Bressan R B, Monger C, et al. Simultaneous disruption of PRC2 and enhancer function underlies histone H3.3-K27M oncogenic activity in human hindbrain neural stem cells. *Nat Genet*, 2021, 53: 1221–1232
- 30 Jain S U, Khazaei S, Marchione D M, et al. Histone H3.3 G34 mutations promote aberrant PRC2 activity and drive tumor progression. *Proc Natl Acad Sci USA*, 2020, 117: 27354–27364
- 31 Fang D, Gan H, Lee J H, et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science*, 2016, 352: 1344–1348
- 32 Lowe B R, Maxham L A, Hamey J J, et al. Histone H3 mutations: an updated view of their role in chromatin deregulation and cancer. *Cancers*, 2019, 11: 660
- 33 Louis D N, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol*, 2016, 131: 803–820
- 34 Behjati S, Tarpey P S, Presneau N, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet*, 2013, 45: 1479–1482
- 35 Jones C, Baker S J. Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. *Nat Rev Cancer*, 2014, 14: 651–661
- 36 Weinberg D N, Allis C D, Lu C. Oncogenic mechanisms of histone H3 mutations. *Cold Spring Harb Perspect Med*, 2017, 7: a026443
- 37 Rajagopalan K N, Chen X, Weinberg D N, et al. Depletion of H3K36me2 recapitulates epigenomic and phenotypic changes induced by the H3.3K36M oncohistone mutation. *Proc Natl Acad Sci USA*, 2021, 118: e2021795118
- 38 Zhou H X, Wang Z X. Molecular mechanisms of the Polycomblike family proteins in epigenetic regulation and their biological functions (in Chinese). *Sci Sin Vitae*, 2022, 52: 1118–1128 [周海霞, 王占新. Polycomblike家族蛋白参与表观遗传调控的分子机制及其生物学意义. 中国科学: 生命科学, 2022, 52: 1118–1128]
- 39 Sahu V, Lu C. Oncohistones: hijacking the histone code. *Annu Rev Cancer Biol*, 2022, 6: 293–312
- 40 Ferrand J, Plessier A, Polo S E. Control of the chromatin response to DNA damage: histone proteins pull the strings. *Semin Cell Dev Biol*, 2021, 113: 75–87
- 41 Chen C C L, Deshmukh S, Jessa S, et al. Histone H3.3G34-mutant interneuron progenitors co-opt PDGFRA for gliomagenesis. *Cell*, 2020, 183: 1617–1633
- 42 Lewis N A, Klein R H, Kelly C, et al. Histone H3.3 K27M chromatin functions implicate a network of neurodevelopmental factors including ASCL1 and NEUROD1 in DIPG. *Epigenet Chromatin*, 2022, 15: 18
- 43 Bennett R L, Bele A, Small E C, et al. A mutation in histone H2B represents a new class of oncogenic driver. *Cancer Discov*, 2019, 9: 1438–1451
- 44 Arimura Y, Ikura M, Fujita R, et al. Cancer-associated mutations of histones H2B, H3.1 and H2A.Z.1 affect the structure and stability of the

- nucleosome. *Nucleic Acids Res*, 2018, 46: 10007–10018
- 45 Wan Y C E, Liu J, Zhu L, et al. The H2BG53D oncohistone directly upregulates ANXA3 transcription and enhances cell migration in pancreatic ductal adenocarcinoma. *Signal Transduct Target Ther*, 2020, 5: 106
- 46 Bagert J D, Mitchener M M, Patriotis A L, et al. Oncohistone mutations enhance chromatin remodeling and alter cell fates. *Nat Chem Biol*, 2021, 17: 403–411
- 47 Karaczyn A A, Bal W, North S L, et al. The octapeptidic end of the C-terminal tail of histone H2A is cleaved off in cells exposed to carcinogenic nickel(II). *Chem Res Toxicol*, 2003, 16: 1555–1559
- 48 Long M, Sun X, Shi W, et al. A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res*, 2019, 47: 8399–8409
- 49 Pang M Y H, Sun X, Ausió J, et al. Histone H4 variant, H4G, drives ribosomal RNA transcription and breast cancer cell proliferation by loosening nucleolar chromatin structure. *J Cell Physiol*, 2020, 235: 9601–9608
- 50 Tessadori F, Duran K, Knapp K, et al. Recurrent *de novo* missense variants across multiple histone H4 genes underlie a neurodevelopmental syndrome. *Am J Hum Genet*, 2022, 109: 750–758
- 51 Bryant L, Li D, Cox S G, et al. Histone H3.3 beyond cancer: Germline mutations in *Histone 3 Family 3A and 3B* cause a previously unidentified neurodegenerative disorder in 46 patients. *Sci Adv*, 2020, 6: eabc9207
- 52 Marzluff W F, Gongidi P, Woods K R, et al. The human and mouse replication-dependent histone genes. *Genomics*, 2002, 80: 487–498
- 53 Wu R S, Tsai S, Bonner W M. Patterns of histone variant synthesis can distinguish go from G1 cells. *Cell*, 1982, 31: 367–374
- 54 Frank D, Doenecke D, Albig W. Differential expression of human replacement and cell cycle dependent H3 histone genes. *Gene*, 2003, 312: 135–143
- 55 Drané P, Ouararhni K, Depaux A, et al. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev*, 2010, 24: 1253–1265
- 56 Lewis P W, Elsaesser S J, Noh K M, et al. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci USA*, 2010, 107: 14075–14080
- 57 Maver A, Čuturilo G, Ruml S J, et al. Clinical next generation sequencing reveals an *H3F3A* gene as a new potential gene candidate for microcephaly associated with severe developmental delay, intellectual disability and growth retardation. *Balkan J Med Genet*, 2019, 22: 65–68
- 58 Alves C A P F, Sherbini O, D’Arco F, et al. Brain abnormalities in patients with germline variants in *H3F3*: novel imaging findings and neurologic symptoms beyond somatic variants and brain tumors. *AJNR Am J Neuroradiol*, 2022, 43: 1048–1053
- 59 Latham J A, Dent S Y R. Cross-regulation of histone modifications. *Nat Struct Mol Biol*, 2007, 14: 1017–1024
- 60 Khazaei S, Chen C C L, Andrade A F, et al. Single substitution in H3.3G34 alters DNMT3A recruitment to cause progressive neurodegeneration. *Cell*, 2023, 186: 1162–1178
- 61 Tessadori F, Giltay J C, Hurst J A, et al. Germline mutations affecting the histone H4 core cause a developmental syndrome by altering DNA damage response and cell cycle control. *Nat Genet*, 2017, 49: 1642–1646
- 62 Sindikubwabo F, Ding S, Hussain T, et al. Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites. *eLife*, 2017, 6: e29391
- 63 Ye J, Ai X, Eugeni E E, et al. Histone H4 lysine 91 acetylation: a core domain modification associated with chromatin assembly. *Mol Cell*, 2005, 18: 123–130
- 64 Selvam K, Rahman S A, Li S. Histone H4 H75E mutation attenuates global genomic and Rad26-independent transcription-coupled nucleotide excision repair. *Nucleic Acids Res*, 2019, 47: 7392–7401
- 65 Tessadori F, Rehman A U, Giltay J C, et al. A *de novo* variant in the human HIST1H4J gene causes a syndrome analogous to the HIST1H4C-associated neurodevelopmental disorder. *Eur J Hum Genet*, 2020, 28: 674–678
- 66 Izzo A, Kamieniarz K, Schneider R. The histone H1 family: specific members, specific functions? *Biol Chem*, 2008, 389: 333–343
- 67 Zhou B R, Feng H, Kale S, et al. Distinct structures and dynamics of chromatosomes with different human linker histone isoforms. *Mol Cell*, 2021, 81: 166–182
- 68 Roque A, Ponte I, Suau P. Post-translational modifications of the intrinsically disordered terminal domains of histone H1: effects on secondary structure and chromatin dynamics. *Chromosoma*, 2017, 126: 83–91
- 69 Fyodorov D V, Zhou B R, Skoultschi A I, et al. Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol*

- [Cell Biol](#), 2018, 19: 192–206
- 70 Alexandrow M G, Hamlin J L. Chromatin decondensation in S-phase involves recruitment of Cdk2 by Cdc45 and histone H1 phosphorylation. [J Cell Biol](#), 2005, 168: 875–886
 - 71 Daujat S, Zeissler U, Waldmann T, et al. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. [J Biol Chem](#), 2005, 280: 38090–38095
 - 72 Vaquero A, Scher M, Lee D, et al. Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. [Mol Cell](#), 2004, 16: 93–105
 - 73 Leicher R, Ge E J, Lin X, et al. Single-molecule and in silico dissection of the interaction between Polycomb repressive complex 2 and chromatin. [Proc Natl Acad Sci USA](#), 2020, 117: 30465–30475
 - 74 Zhao J, Wang M, Chang L, et al. RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. [Nat Cell Biol](#), 2020, 22: 439–452
 - 75 Liu C, Yu J, Song A, et al. Histone H1 facilitates restoration of H3K27me3 during DNA replication by chromatin compaction. [Nat Commun](#), 2023, 14: 4081
 - 76 Yang S M, Kim B J, Norwood Toro L, et al. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. [Proc Natl Acad Sci USA](#), 2013, 110: 1708–1713
 - 77 Catez F, Ueda T, Bustin M. Determinants of histone H1 mobility and chromatin binding in living cells. [Nat Struct Mol Biol](#), 2006, 13: 305–310
 - 78 Taube J H, Allton K, Duncan S A, et al. Foxa1 functions as a pioneer transcription factor at transposable elements to activate Afp during differentiation of embryonic stem cells. [J Biol Chem](#), 2010, 285: 16135–16144
 - 79 Postnikov Y V, Bustin M. Functional interplay between histone H1 and HMG proteins in chromatin. [Biochim Biophys Acta](#), 2016, 1859: 462–467
 - 80 Ghosh R P, Horowitz-Scherer R A, Nikitina T, et al. MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. [Mol Cell Biol](#), 2010, 30: 4656–4670
 - 81 Papin A, Cesarman E, Melnick A. 3D chromosomal architecture in germinal center B cells and its alterations in lymphomagenesis. [Curr Opin Genet Dev](#), 2022, 74: 101915
 - 82 Li H, Kaminski M S, Li Y, et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. [Blood](#), 2014, 123: 1487–1498
 - 83 Yusufova N, Kloetgen A, Teater M, et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. [Nature](#), 2021, 589: 299–305
 - 84 Morin R D, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. [Blood](#), 2013, 122: 1256–1265
 - 85 Soshnev A A, Allis C D, Cesarman E, et al. Histone H1 mutations in lymphoma: a link(er) between chromatin organization, developmental reprogramming, and cancer. [Cancer Res](#), 2021, 81: 6061–6070
 - 86 Wright G W, Huang D W, Phelan J D, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. [Cancer Cell](#), 2020, 37: 551–568
 - 87 Willcockson M A, Heaton S E, Weiss C N, et al. H1 histones control the epigenetic landscape by local chromatin compaction. [Nature](#), 2021, 589: 293–298
 - 88 Cuartero S, Stik G, Stadhouders R. Three-dimensional genome organization in immune cell fate and function. [Nat Rev Immunol](#), 2023, 23: 206–221
 - 89 Scourzic L, Salataj E, Apostolou E. Deciphering the complexity of 3D chromatin organization driving lymphopoiesis and lymphoid malignancies. [Front Immunol](#), 2021, 12: 669881
 - 90 Béguelin W, Popovic R, Teater M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. [Cancer Cell](#), 2013, 23: 677–692
 - 91 Klymenko T, Bloehdorn J, Bahlo J, et al. Lamin B1 regulates somatic mutations and progression of B-cell malignancies. [Leukemia](#), 2018, 32: 364–375
 - 92 Zhou B R, Jiang J, Feng H, et al. Structural mechanisms of nucleosome recognition by linker histones. [Mol Cell](#), 2015, 59: 628–638
 - 93 Scott C L, Gil J, Hernando E, et al. Role of the chromobox protein CBX7 in lymphomagenesis. [Proc Natl Acad Sci USA](#), 2007, 104: 5389–5394

- 94 Béguelin W, Rivas M A, Calvo Fernández M T, et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. [Nat Commun](#), 2017, 8: 877
- 95 Chew G L, Bleakley M, Bradley R K, et al. Short H2A histone variants are expressed in cancer. [Nat Commun](#), 2021, 12: 490
- 96 Hanahan D, Weinberg R A. The hallmarks of cancer. [Cell](#), 2000, 100: 57–70
- 97 Hanahan D, Weinberg R A. Hallmarks of cancer: the next generation. [Cell](#), 2011, 144: 646–674
- 98 Hanahan D. Hallmarks of cancer: new dimensions. [Cancer Discov](#), 2022, 12: 31–46
- 99 Chapuy B, Stewart C, Dunford A J, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. [Nat Med](#), 2018, 24: 679–690
- 100 Vujatovic O, Zaragoza K, Vaquero A, et al. Drosophila melanogaster linker histone dH1 is required for transposon silencing and to preserve genome integrity. [Nucleic Acids Res](#), 2012, 40: 5402–5414
- 101 Downs J A, Kosmidou E, Morgan A, et al. Suppression of homologous recombination by the *Saccharomyces cerevisiae* linker histone. [Mol Cell](#), 2003, 11: 1685–1692
- 102 Murga M, Jaco I, Fan Y, et al. Global chromatin compaction limits the strength of the DNA damage response. [J Cell Biol](#), 2007, 178: 1101–1108
- 103 Bakkenist C J, Kastan M B. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. [Nature](#), 2003, 421: 499–506
- 104 Thorslund T, Ripplinger A, Hoffmann S, et al. Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. [Nature](#), 2015, 527: 389–393
- 105 Li Z, Li Y, Tang M, et al. Destabilization of linker histone H1.2 is essential for ATM activation and DNA damage repair. [Cell Res](#), 2018, 28: 756–770
- 106 Burkardt D D C, Zachariou A, Loveday C, et al. *HIST1H1E* heterozygous protein-truncating variants cause a recognizable syndrome with intellectual disability and distinctive facial gestalt: a study to clarify the HIST1H1E syndrome phenotype in 30 individuals. [Am J Med Genet A](#), 2019, 179: 2049–2055
- 107 Zhao J, Lyu G, Ding C, et al. Expanding the mutational spectrum of Rahman syndrome: a rare disorder with severe intellectual disability and particular facial features in two Chinese patients. [Mol Genet Genomic Med](#), 2022, 10: e1825
- 108 Flex E, Martinelli S, Van Dijk A, et al. Aberrant function of the C-terminal tail of HIST1H1E accelerates cellular senescence and causes premature aging. [Am J Hum Genet](#), 2019, 105: 493–508
- 109 Ciolfi A, Aref-Eshghi E, Pizzi S, et al. Frameshift mutations at the C-terminus of HIST1H1E result in a specific DNA hypomethylation signature. [Clin Epigenet](#), 2020, 12: 7
- 110 Spracklin G, Abdennur N, Imakaev M, et al. Diverse silent chromatin states modulate genome compartmentalization and loop extrusion barriers. [Nat Struct Mol Biol](#), 2023, 30: 38–51
- 111 Belzeaux R, Bergon A, Jeanjean V, et al. Responder and nonresponder patients exhibit different peripheral transcriptional signatures during major depressive episode. [Transl Psychiatry](#), 2012, 2: e185
- 112 Prendergast L, Reinberg D. The missing linker: emerging trends for H1 variant-specific functions. [Genes Dev](#), 2021, 35: 40–58
- 113 Zheng Y, John S, Pesavento J J, et al. Histone H1 phosphorylation is associated with transcription by RNA polymerases I and II. [J Cell Biol](#), 2010, 189: 407–415
- 114 Tremblay M W, Green M V, Goldstein B M, et al. Mutations of the histone linker *H1-4* in neurodevelopmental disorders and functional characterization of neurons expressing C-terminus frameshift mutant H1.4. [Hum Mol Genet](#), 2022, 31: 1430–1442
- 115 Bozzi Y, Provenzano G, Casarosa S. Neurobiological bases of autism-epilepsy comorbidity: a focus on excitation/inhibition imbalance. [Eur J Neurosci](#), 2018, 47: 534–548
- 116 Lyst M J, Connelly J, Merusi C, et al. Sequence-specific DNA binding by AT-hook motifs in MeCP2. [FEBS Lett](#), 2016, 590: 2927–2933
- 117 Lyst M J, Ekiert R, Ebert D H, et al. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. [Nat Neurosci](#), 2013, 16: 898–902
- 118 Chahrour M, Jung S Y, Shaw C, et al. MeCP2, a key contributor to neurological disease, activates and represses transcription. [Science](#), 2008, 320: 1224–1229
- 119 Ghosh R P, Nikitina T, Horowitz-Scherer R A, et al. Unique physical properties and interactions of the domains of methylated DNA binding protein 2. [Biochemistry](#), 2010, 49: 4395–4410

- 120 Collins B E, Neul J L. Rett syndrome and MECP2 duplication syndrome: disorders of MeCP2 dosage. *Neuropsychiatr Dis Treat*, 2022, Volume 18: 2813–2835
- 121 Shahbazian M D. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum Mol Genet*, 2002, 11: 115–124
- 122 Kishi N, Macklis J D. Dissecting MECP2 function in the central nervous system. *J Child Neurol*, 2005, 20: 753–759
- 123 Nan X, Campoy F J, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, 1997, 88: 471–481
- 124 Marano D, Fioriniello S, D’Esposito M, et al. Transcriptomic and epigenomic landscape in Rett syndrome. *Biomolecules*, 2021, 11: 967
- 125 Good K V, Vincent J B, Ausió J. MeCP2: the genetic driver of Rett syndrome epigenetics. *Front Genet*, 2021, 12: 620859
- 126 Skene P J, Illingworth R S, Webb S, et al. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol Cell*, 2010, 37: 457–468
- 127 Georgel P T, Horowitz-Scherer R A, Adkins N, et al. Chromatin compaction by human MeCP2. *J Biol Chem*, 2003, 278: 32181–32188
- 128 Horike S, Cai S, Miyano M, et al. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet*, 2005, 37: 31–40
- 129 Wang C, Wang F, Li Z, et al. MeCP2-mediated epigenetic regulation in senescent endothelial progenitor cells. *Stem Cell Res Ther*, 2018, 9: 87
- 130 Pandey S, Simmons Jr G E, Malyarchuk S, et al. A novel MeCP2 acetylation site regulates interaction with ATRX and HDAC1. *Genes Cancer*, 2015, 6: 408–421
- 131 Wang L, Hu M, Zuo M Q, et al. Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Res*, 2020, 30: 393–407
- 132 Agarwal N, Becker A, Jost K L, et al. MeCP2 Rett mutations affect large scale chromatin organization. *Hum Mol Genet*, 2011, 20: 4187–4195
- 133 Casas-Delucchi C S, Becker A, Bolius J J, et al. Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization. *Nucleic Acids Res*, 2012, 40: e176
- 134 Baker S A, Chen L, Wilkins A D, et al. An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell*, 2013, 152: 984–996
- 135 Wong L C, Singh S, Wang H P, et al. FOXG1-related syndrome: from clinical to molecular genetics and pathogenic mechanisms. *Int J Mol Sci*, 2019, 20: 4176
- 136 Dai S, Qu L, Li J, et al. Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations. *Nucleic Acids Res*, 2021, 49: 10235–10249
- 137 Akol I, Izzo A, Gather F, et al. Multimodal epigenetic changes and altered NEUROD1 chromatin binding in the mouse hippocampus underlie FOXG1 syndrome. *Proc Natl Acad Sci USA*, 2023, 120: e2122467120
- 138 De Filippis R, Pancrazi L, Bjørge K, et al. Expanding the phenotype associated with *FOXG1* mutations and *in vivo* FoxG1 chromatin-binding dynamics. *Clin Genet*, 2012, 82: 395–403
- 139 Thomas J O, Stott K. H1 and HMGB1: modulators of chromatin structure. *Biochem Soc Trans*, 2012, 40: 341–346
- 140 Wang S, Zhang Y. HMGB1 in inflammation and cancer. *J Hematol Oncol*, 2020, 13: 116
- 141 Jeronimo C, Robert F. The histone chaperone FACT: a guardian of chromatin structure integrity. *Transcription*, 2022, 13: 16–38
- 142 Chen P, Dong L, Hu M, et al. Functions of FACT in breaking the nucleosome and maintaining its integrity at the single-nucleosome level. *Mol Cell*, 2018, 71: 284–293
- 143 Li W, Chen P, Yu J, et al. FACT remodels the tetranucleosomal unit of chromatin fibers for gene transcription. *Mol Cell*, 2016, 64: 120–133
- 144 Bhakat K K, Ray S. The FAcilitates Chromatin Transcription (FACT) complex: its roles in DNA repair and implications for cancer therapy. *DNA Repair*, 2022, 109: 103246

Variations in chromatin architectural proteins in human diseases

CAO Hong¹, WU Di^{1,2}, WEI Yu^{1,2}, LIN HaiYan¹ & LI GuoHong^{1,2}

1 National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;

2 Frontier Science Centre of Immunology and Metabolism, TaiKang Center for Life and Medical Sciences, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan 430072, China

The hierarchical three-dimensional (3D) chromatin architecture is central to spatiotemporal control of multiple biological processes inherent to DNA that are required for the maintenance of cell identity and function, including DNA replication, transcription, recombination, damage repair. Aberrant alterations in 3D chromatin architecture, which lead to pathological gene expression programmes, have been delineated as prominent causes of many diseases, such as tumors and neurodevelopmental disorders. In this review, we summarize the roles of histones, histone variants and methyl CpG-binding protein 2 in the regulation of 3D chromatin structure and dynamics, explore the impacts of human diseases-related mutations on their functions, and provide insights into the underlying pathological mechanisms of tumors and neurodevelopmental disorders from the aspect of chromatin structure.

three-dimensional (3D) chromatin architecture, chromatin architectural proteins, histones and histone variants, methyl CpG-binding protein 2 (MeCP2), human diseases

doi: [10.1360/SSV-2023-0221](https://doi.org/10.1360/SSV-2023-0221)