

Opinion

To incise or not and where: SET-domain methyltransferases know

Michelle N. Yancoskie ¹, Corina Maritz ¹, Patrick van Eijk ², Simon H. Reed ² and Hanspeter Naegeli ^{1,*}

The concept of the histone code posits that histone modifications regulate gene functions once interpreted by epigenetic readers. A well-studied case is trimethylation of lysine 4 of histone H3 (H3K4me3), which is enriched at gene promoters. However, H3K4me3 marks are not needed for the expression of most genes, suggesting extra roles, such as influencing the 3D genome architecture. Here, we highlight an intriguing analogy between the H3K4me3-dependent induction of double-strand breaks in several recombination events and the impact of this same mark on DNA incisions for the repair of bulky lesions. We propose that *Su(var)3–9*, *Enhancer-of-zeste* and *Trithorax (SET)*-domain methyltransferases generate H3K4me3 to guide nucleases into chromatin spaces, the favorable accessibility of which ensures that DNA break intermediates are readily processed, thereby safeguarding genome stability.

Structural determinants of genome functions

In 2022, the first truly gapless compilation of the human genome was released, with centromeres, satellite repeats, amplified elements, and other previously refractory regions finally assembled through long-read sequencing technologies [1]. Although now fully sequenced, the genome still holds myriad mysteries that cannot be deciphered solely from the linear succession of nucleotides. One important challenge is to understand the spatiotemporal regulation of genome-wide functions, including transcription, DNA replication, recombination, and repair.

Research over the past decades has shown that such dynamic nuclear transactions are intimately linked to how the genome engages in higher-order structures, whereby chromatin fibers build megabase pair-scale self-interacting regions, known as **topologically associating domains** (TADs; see [Glossary](#)). Each TAD encompasses multiple loops, or ‘sub-TADs’, formed by ring-like cohesin complexes that translocate along the chromatin fibers to bring distal genetic loci into proximity [2,3]. This **3D architecture** enables chromatin compaction and includes the demarcation of specialized compartments [4,5], the functions of which correlate with characteristic patterns of histone modifications (i.e., acetylation, methylation, phosphorylation, and ubiquitination). For example, active TADs are populated by **H3K4me1**, **H3K4me3**, and **H3K27ac** marks ([Figure 1](#)) and their boundaries are enriched for H3K4me3 [6]. Transcriptional promoters and enhancers display high and low ratios, respectively, of H3K4me3 to H3K4me1 [7], and key developmental genes are decorated by both H3K4me3 and **H3K27me3** [8].

In this opinion article, we examine H3K4me3 as one of the best-studied histone marks shaping the 3D genome and highlight the distinctive role of **SET**-domain methyltransferases depositing this modification in targeting nucleolytic DNA cleavages. The purpose is to demonstrate a striking

Highlights

The first gapless (‘telomere-to-telomere’) sequence of the human genome was recently published, but there are still knowledge gaps as to how the linear succession of nucleotides and the 3D partitioning of the genome affect genomic functions.

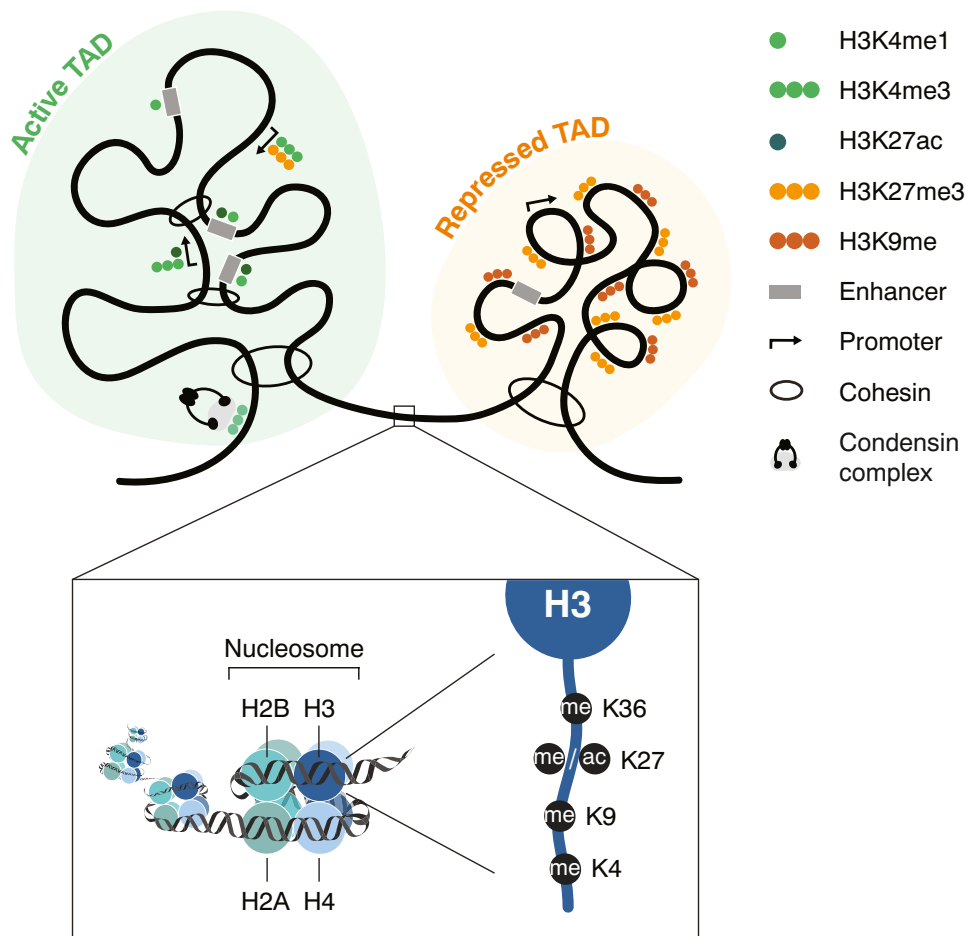
The strong inverse relationship in mammals between mutation rates and lifespan points to an evolutionary pressure to restrain mutagenesis even in non-coding ‘junk DNA’ regions, prompting a search for genome-wide markers of DNA damage that recruit enzymes to correct single- and double-strand breaks before they become mutagenic.

The analysis of various nucleolytic transactions indicates that the histone mark trimethylation of lysine 4 of histone H3 (H3K4me3), thought primarily to trigger transcription, adopts a universal role to promote genome stability (and counteract mutations) by shaping dynamic 3D spaces for the remediation of DNA strand breaks.

¹Institute of Pharmacology and Toxicology, University of Zurich-Vetsuisse, Zurich, Switzerland

²Institute of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK

*Correspondence: hanspeter.naegeli@uzh.ch (H. Naegeli).



Trends in Biochemical Sciences

Figure 1. Histone modifications facilitate the compartmentalization of chromatin into specialized domains. The genome is organized into megabase pair-scale self-interacting neighborhoods termed ‘topologically associating domains’ (TADs). Interactions between functional elements are facilitated by ring-like cohesin complexes that translocate along chromatin fibers and generate extruding chromatin loops, bringing distal loci, such as gene enhancers and promoters, into proximity [2–4]. Transcriptionally active versus repressed TADs are partly distinguished by the modification marks that decorate the fundamental structural unit of chromatin, the nucleosome. At the highest level of resolution, chromatin comprises 147 base pairs of DNA wrapped around a nucleosome core of eight histones (an H3–H4 tetramer flanked by two H2A–H2B dimers), with linker DNA of 20–50 base pairs connecting these core particles along nucleosome repeats [75]. The 3D genome architecture is influenced by the histone code, which entails the interpretation of post-translational histone modifications on nucleosomes. For example, the N-terminal tail of histone H3 protruding from the nucleosome core can harbor methylation (‘me’) or acetylation (‘ac’) on lysine residues. Monomethylation of histone H3 at position lysine 4 (H3K4me1) histone marks are involved in cohesin recruitment [23,76], trimethylation of histone H3 at position lysine 4 (H3K4me3) in the demarcation of TAD boundaries [6,24], and H3K4me1/me3 together with acetylation/trimethylation of histone H3 at position lysine 27 (H3K27ac and H3K27me3) define enhancer and promoter sequences [7]. H3K4me3 is additionally implicated in the assembly of transcription factories comprising highly expressed genes [25–27]. Instead, transcriptionally repressed TADs contain H3K9 and H3K27 methylations [12].

parallelism between the formation of DNA double-strand breaks in programmed recombination events and the induction of DNA single-strand breaks for the repair of bulky base lesions.

Diverse functions of histone methylation marks

Histone modifications are deposited by epigenetic ‘writers’ and constitute a code that is interpreted into a multitude of responses by epigenetic ‘readers’ (reviewed in [9]). Understanding

Glossary

3D architecture: 3D organization of the chromatin fibers of chromosomes into domains, subdomains, and further compartments.

Absent, small, or homeotic 1 (ASH1): histone-lysine methyltransferase discovered in *Drosophila* as a member of the trithorax group of transcriptional activators involved in regulating *Hox* gene expression and body segment identity. The human homolog, ASH1L, is a 2969 amino-acid protein with a molecular weight of 333 kDa.

Chromosomal interacting domains (CIDs): regions of the genome in yeast that are more likely than random sequences to self-interact (comparable to TADs in mammals).

Class switch recombination (CSR): mechanism of somatic recombination to change the immunoglobulin subtype. During this process, the constant-region portion of the antibody heavy chain is switched from one subtype to another.

Cyclobutane pyrimidine dimers (CPDs): predominant and mutagenic bulky DNA lesions induced by exposure to UV radiation, including the UV-A and UV-B spectrum of sunlight.

Effector proteins: proteins recruited to chromatin to mediate downstream responses to histone modifications.

Global-genome nucleotide excision repair (GG-NER): global-genome subpathway of nucleotide excision repair responsible for the excision of bulky DNA lesions, including UV photoproducts and carcinogen–DNA adducts. Unlike TC-NER, this repair subpathway is able to handle bulky DNA lesions throughout the genome independently of ongoing transcription.

H3K4me1/3: mono/trimethylation of histone H3 at position lysine 4.

H3K27ac/me3: acetylation/trimethylation of histone H3 at position lysine 27.

Heterochromatin protein 1: family of proteins that establish and maintain repressive chromatin and gene silencing.

Polycomb repressive complexes 1 and 2: multiprotein transcription-repressive complexes containing methyltransferase and ubiquitin ligase activities responsible for the establishment of facultative heterochromatin.

PR/SET domain-containing protein 9 (PRDM9): unique histone lysine

histone codes has proved challenging, in part due to the abundance and redundancy of various marks and their writers. Misleading phenotypes may result from inadvertently mutating or deleting enzymes or readers the roles of which are not directly related to the establishment and interpretation of a given histone modification [10]. Even changing only the catalytic site of histone modifiers may produce effects unlinked to histones due to the potential of certain writers, including SET-domain methyltransferases, to methylate nonhistone substrates [11].

Histone acetylation is among the modifications the function of which has been largely elucidated. For example, H3K27ac contributes to transcription activation by directly destabilizing nucleosomes, thereby allowing transcription factors to access promoter and enhancer sequences [12]. Overall, the consequences of histone methylations are less clear because they involve the addition of nonpolar groups incapable of changing the histone charge; hence, they act through **effector proteins** [13]. Some histone methylations have been shown to downregulate gene expression. H3K9 methylation contributes to the formation of constitutive heterochromatin through the recruitment of **heterochromatin protein 1** [14–16]. H3K27me3 induces facultative heterochromatin during embryonic development, leading to X chromosome inactivation and gene silencing through **Polycomb repressive complexes 1 and 2** [17]. Such methylations at H3K9 and H3K27 result in partitioning of the genome into compartments that repress gene expression by physically excluding transcription factors (Figure 1). However, as mentioned above, other histone methylation marks activate genome functions, but the mechanisms remain elusive in many cases [12]. We discuss below H3K4me3 as the paradigm of a histone modification that is generally associated with active transcription.

The case of H3K4me3

This well-studied histone mark frequently decorates the promoters of active genes and, hence, has been surmised to stimulate gene expression [18]. Counterintuitively given its reputation as an active mark, H3K4me3 is dispensable for global transcription [19–21]. H3K4me3 only directly affects gene expression at a handful of loci, where it is required to maintain expression patterns established by specific transcription factors [22].

There is increasing evidence that H3K4me3 may instead affect genome function through its impact on 3D genome architecture (Figure 1). Whereas H3K4me1 recruits cohesin to form loops between promoters and enhancers, delineating transcriptionally active sub-TADs [23], H3K4me3 engages condensin II complexes to tether active promoters at the origin of these loops, thus forming clusters of highly transcribed genes. The downregulation of a SET-domain methyltransferase complex depositing H3K4me3 resulted in a global H3K4me3 decrease accompanied by a loss of condensin II binding at the boundaries of topological domains, suggesting that H3K4me3 regulates gene expression through its contribution to chromatin loop formation [24]. Broad H3K4me3 peaks (where the H3K4me3 mark extends over the promoters into coding regions) are additionally able to assemble ‘transcription factories’, denoting phase-separated condensates with a high local content of histone modifiers, chromatin remodelers, topoisomerases and transcription factors [25–27]. These and other observations [28,29] lend support to the notion that H3K4me3 constitutes a topological determinant for the assembly of chromatin compartments or subcompartments as appropriate targets of dynamic transactions encompassing not only transcription, but also DNA recombination and repair, which are discussed in more detail below.

H3K4me3 in DNA recombination

At least three DNA recombination events [meiotic recombination, **variable, diversity and joining** (VDJ) gene segment rearrangement, and immunoglobulin **class switch recombination** (CSR)]

methyltransferase expressed in germ cells for the induction of meiotic recombination.

SET domain-containing protein 1A, 1B (Set1A/B): histone methyltransferases that deposit H3K4me3 in mammalian cells.

Su(var)3–9, Enhancer-of-zeste, and Trithorax (SET): catalytic domain of a family of histone methyltransferases, named after the *Drosophila* genes *Su(var)3–9*, *Enhancer-of-zeste*, and *Trithorax*.

Topologically associating domains (TADs): self-interacting chromatin neighborhoods discovered from high-throughput chromosome conformation capture assays that reveal the differential contact frequency between regions of the genome.

Transcription-coupled nucleotide excision repair (TC-NER):

transcription-coupled subpathway of nucleotide excision repair whereby bulky DNA lesions are detected by the elongating RNA polymerase II complex. This repair subpathway is restricted to the transcribed strand of active genes.

Variability, diversity and joining

(VDJ) rearrangement: mechanism of somatic recombination in developing B and T lymphocytes. This recombination between variability, diversity, and joining gene segments results in the highly diverse repertoire of antibodies and T cell receptors characteristic of the adaptive immune system.

Xeroderma pigmentosum groups A, B, C, etc. (XPA, XPB, XPC, etc.): NER factors named after the disease (xeroderma pigmentosum) caused by a defect in the respective genes.

use H3K4me3 to mark sites of double-strand break formation (Box 1). In this role, H3K4me3 ensures genome stability by regulating the induction of DNA breaks, possibly through its contribution to compartmentalizing the genome into specialized units (or loops) permissive for full restoration of the double helix.

In meiotic cells, the SET-domain methyltransferase **PR/SET domain-containing protein 9** (PRDM9) deposits H3K4me3 to initiate homologous recombination ending in genetic crossovers (Figure 2, left). This PRDM9-dependent reaction is suppressed in peripheral chromatin regions bound to the nuclear lamina, instead preferring accessible chromatin, including transcribed genes [30]. PRDM9-mediated double-strand breaks are also depleted in repetitive genetic elements marked by repressive H3K9 methylations [31], thus avoiding deleterious rearrangements between 'selfish' genetic elements. In the absence of PRDM9, meiotic mouse cells still generate double-strand breaks at pre-existing H3K4me3, especially in active gene promoters. However, these ectopic breaks fail to generate crossovers and are instead repaired through gene conversion. A possible explanation is that transcription or other transactions at such functional sequences compete for the same DNA substrate, preventing successful completion of meiosis [32–34].

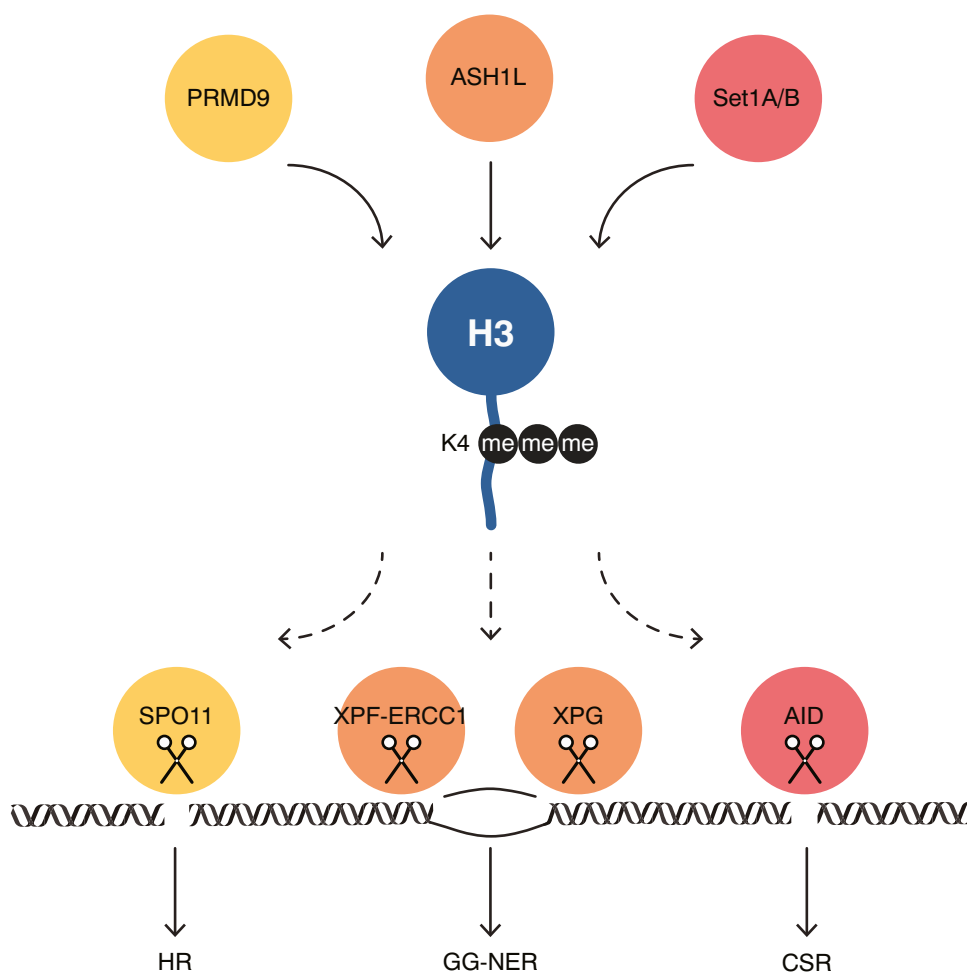
The H3K4me3 mark is also important for the VDJ rearrangement of antigen receptor genes in developing lymphocytes [35–38]. Similarly, H3K4me3 regulates and stimulates the immunoglobulin heavy chain CSR in mature B lymphocytes. Circumstantial evidence suggests that, in the case of CSR, H3K4me3 is deposited by the histone methyltransferases **SET domain-containing protein 1A and 1B** (Set1A/B) to guide the double-strand break induction [39] (Figure 2, right). Tight regulation of these DNA rearrangements in lymphocytes is important because unrepaired double-strand break intermediates give rise to oncogenic chromosomal translocations causing leukemias and lymphomas [40].

Box 1. H3K4me3 regulates programmed nucleolytic activities initiating genomic recombination events

Site-specific DNA cleavage reactions are guided by the H3K4me3 epigenetic mark in at least three different recombination pathways (also see Figure 2 in main text).

- (i) During gametogenesis, double-strand breaks are purposefully introduced to initiate homologous recombination for the reciprocal exchange of genetic information (in the form of meiotic crossovers) between paired homologous chromosomes, one from each parent. The resulting new combinations of alleles in the gametes (egg or sperm) ensure genetic variation in the offspring. In humans, mice, and many other mammals, the locations of meiotic double-strand breaks induced by the Sporulation-specific 11 (SPO11) nuclease are predetermined through H3K4me3 depositions by the germline-expressed SET-domain methyltransferase PRDM9 [58,71–73]. This selection of meiotic recombination hotspots involves recognition of the histone methylations by dedicated H3K4me3 readers [56].
- (ii) The H3K4me3 mark is also important for somatic recombination mechanisms in developing lymphocytes initiated by the Recombination activating gene (RAG) complex, which rearranges the variable/diversity/joining (VDJ) segments of immunoglobulin and T cell receptor genes. VDJ recombination results in the highly diverse repertoire of antibodies and T cell receptors found in B cells and T cells, respectively. The Plant homeodomain (PHD) finger of the RAG2 recombinase subunit interacts with H3K4me3. This association with H3K4me3 is required for the binding of the RAG complex to recombination signal sequences that determine the sites of recombination [35–37].
- (iii) Similarly, H3K4me3 regulates the immunoglobulin class switch recombination (CSR) in mature B lymphocytes, which is necessary to convert the immunoglobulin heavy chain from IgM to another isotype (IgA, IgG, or IgE). Circumstantial evidence suggests that, in the case of immunoglobulin CSR, H3K4me3 is deposited by the histone methyltransferases Set1A and/or Set1B. In turn, H3K4me3 guides activation-induced cytidine deaminase (AID), which induces double-strand breaks at the immunoglobulin heavy chain loci by converting cytosines to uracil [39].

The biological relevance of H3K4me3 depositions by PRDM9 is demonstrated by the sterility of PRDM9^{-/-} mice due to a complete arrest of meiosis in both sexes [74]. Likewise, a tight regulation of RAG and AID-initiated DNA rearrangements in lymphocytes is important because any failure to repair the transient double-strand break intermediates results in oncogenic chromosomal translocations leading to leukemias and lymphomas [40].



Trends in Biochemical Sciences

Figure 2. Trimethylation of lysine 4 of histone H3 (H3K4me3) deposited by *Su(var)3–9*, *Enhancer-of-zeste*, and *Trithorax* (SET)-domain histone methyltransferases guides DNA endonucleases to cleave the DNA double helix. Left, In meiotic cells, H3K4me3 deposited by the histone methyltransferase PR/SET domain-containing protein 9 (PRDM9) predetermines the sites of double-strand DNA breaks, introduced by the sporulation-specific 11 (SPO11) nuclease to initiate genetic crossovers by homologous recombination HR). Right, During immunoglobulin heavy chain class switch recombination (CSR) in lymphocytes, Set1A/B deposits H3K4me3 to guide the formation of double-strand DNA breaks triggered by activation-induced cytidine deaminase (AID). H3K4me3 marks are also important for variable/diversity/joining (VDJ) recombination in developing lymphocytes, but in this process, the histone methyltransferase involved is unknown. By analogy, we hypothesize that the histone methyltransferase Absent, small or homeotic 1-like (ASH1L) deposits H3K4me3 to recruit the global-genome nucleotide excision repair GG-NER) complex and induce dual single-strand breaks by the structure-specific endonucleases XPF and XPG (xeroderma pigmentosum group F and G (XPF and XPG), which enable excision of the offending DNA damage.

From the above recombination pathways, we are of the opinion that H3K4me3 is able to regulate programmed nucleolytic activities in the genome. Notably, H3K4me3 associates with yet another nucleolytic enzyme that intentionally generates DNA breaks, topoisomerase 1, to relieve helical tensions [41], supporting the hypothesis that this histone mark exerts a universal role in the programmed induction of DNA cleavages. Next, we extend this analogy to another system, the **global-genome nucleotide excision repair** (GG-NER) reaction, which must efficiently process DNA incision intermediates.

H3K4me3 in DNA excision repair?

The GG-NER system uses a genome-wide cut-and-patch reaction that removes bulky DNA adducts by dual incision of damaged strands followed by excision of the lesions as part of single-strand oligomers (Figure 3) [42,43]. This repair response is particularly critical in the skin because exposure to the UV radiation of sunlight generates highly mutagenic and carcinogenic bulky lesions, primarily **cyclobutane pyrimidine dimers** (CPDs). Such UV lesions form rather uniformly throughout the genome, including in DNA wrapped around the histones of nucleosome cores [44]. Given that nucleosomes may constitute an obstacle to DNA damage repair, their reconfiguration by histone modifications (mainly acetylation) and chromatin remodelers is thought to render damaged DNA accessible to the GG-NER machinery [45]. However, we reasoned that nucleosomes primed by histone modifiers may provide a scaffold facilitating DNA damage recognition [46]. Intriguingly, the **xeroderma pigmentosum group** (XPC) initiator of GG-NER activity interacts with core histones, but tends to avoid their acetylated forms, including H3K27ac [47]. This observation was surprising because histone acetylation correlates with relaxed chromatin.

While investigating whether H3K4me3 may promote the recruitment of GG-NER factors to damaged nucleosomes, we noted an analogy between the dual incisions for the removal of bulky lesions in the NER process and the site-directed DNA breaks that launch genetic recombination in meiotic cells as well as VDJ recombination and immunoglobulin CSR in lymphocytes (Figure 2). Here, we cover three aspects of this analogy between GG-NER and the above recombination processes.

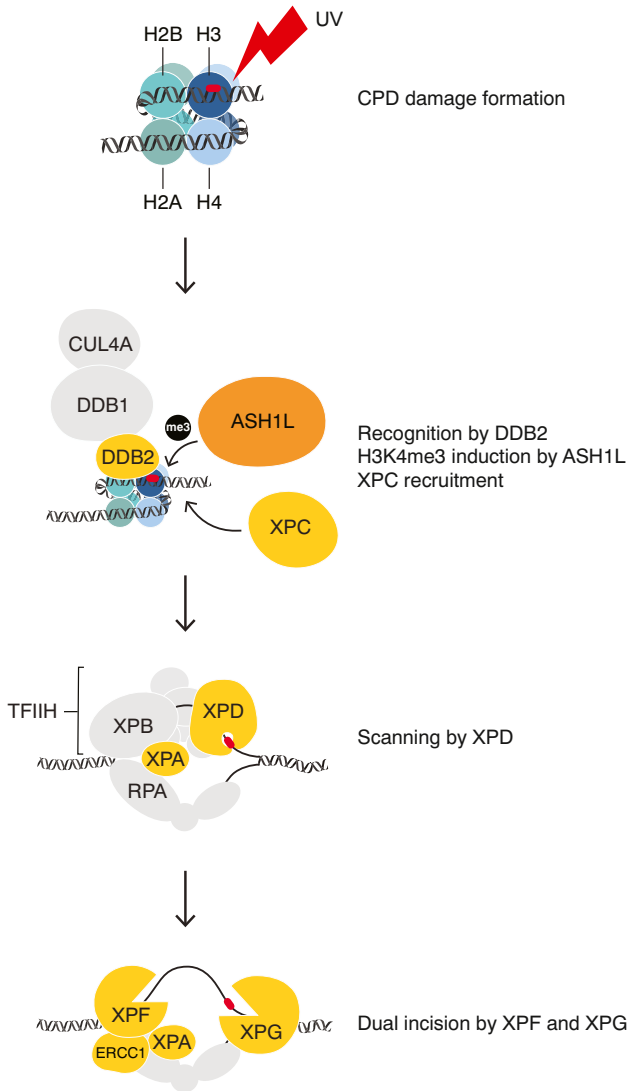
First, although restricted to damaged strands in the case of the GG-NER pathway, these reactions are all triggered by cleavage of an intact deoxyribose-phosphate backbone. The double helix is subjected to double-strand breaks in meiotic recombination, VDJ recombination, and CSR, and to single-strand breaks during the GG-NER reaction. Similar to many other bulky adducts, CPDs do not destabilize the DNA backbone on their own [48] and, therefore, incisions near these lesions only occur by the action of dedicated endonucleases (Figure 3).

Second, DNA cleavage must occur at sites that are readily amenable to the downstream processing of broken DNA intermediates. To this end, chromatin should be accessible for 'big enzyme' machines that repair double-strand or single-strand breaks. As outlined above, PRDM9 deposits H3K4me3 away from active gene promoters during meiosis to avoid DNA cleavages at functional sequences that would provoke conflicts between transcription and double-strand break repair [32,34]. Similar to recombination, the GG-NER complex must avoid potentially detrimental collisions with concurrent transactions, primarily transcription. Only in the case of **transcription-coupled nucleotide excision repair** (TC-NER) is the coordination between repair and transcription guaranteed *a priori* by delegating the initial damage sensing to elongating RNA polymerase II complexes [49].

Third, resolution of cleavage reactions in GG-NER and recombination is essential to maintain genome integrity [50,51]. In recombination, concluding the pathway is important to remove cytotoxic double-strand breaks and achieve the desired genetic rearrangements. In the GG-NER reaction, completion of the pathway is essential to avoid persistent single-strand breaks that, if unprocessed, would convert to double-strand breaks.

ASH1L methyltransferase for the repair of UV damage

In view of this multipart analogy between the initiation of recombination and the removal of bulky lesions, we asked whether, as demonstrated for PRDM9 in meiosis, a histone methyltransferase catalyzing the *de novo* formation of H3K4me3 may help prioritize sites of GG-NER activity. ChIP-



Trends in Biochemical Sciences

excision repair cross complementing 1 (XPF-ERCC1) and XPG [80]. Finally, the double helix is reconstituted by repair synthesis and ligation. XPC and DDB2 are dispensable for a second subpathway, known as transcription-coupled NER (TC-NER), which excises lesions only from transcribed strands and thereby relies on RNA polymerase II stalled at damage to alert the XPD verifier [49].

sequencing studies revealed that XPC protein is constitutively bound to the promoters of active genes in the absence of DNA damage [52] and that these constitutive sites of XPC occupancy overlap strongly with H3K4me3 marks [53]. By biochemical analyses, stronger binding of XPC protein to nucleosomes containing H3K4me3 (pulled down with anti-H3K4me3 antibodies) was observed compared with random nucleosomes (pulled down with generic pan-H3 antibodies) [54]. The H3K4me3 mark also correlates with early rapid excision of CPDs from active and poised gene promoters [44,55]. Following the analogy between PRDM9-induced recombination and the GG-NER pathway, the role of H3K4me3 in the repair of bulky lesions would be to mark chromatin sites that support the enzymatic machinery for DNA repair synthesis and ligation. One difference

Figure 3. Absent, small or homeotic 1-like (ASH1L) may facilitate DNA adduct removal from the genome by depositing trimethylation of lysine 4 of histone H3 (H3K4me3). Sunlight-induced cyclobutane pyrimidine dimers (CPDs) are DNA adducts processed exclusively by the global-genome nucleotide excision repair pathway (GG-NER). In a cut-and-patch reaction, the damaged strands are incised on both sides of the lesion followed by release of the damage as part of 24–32-nucleotide oligomers [42,43]. Specialized GG-NER factors are shown in yellow. In detail, this pathway uses xeroderma pigmentosum (XP) C as a generic sensor that detects unpaired bases [77]. XPC alone is ineffective at recognizing CPDs, because they cause minimal distortions of the DNA helix. The histone methyltransferase ASH1L, shown in orange, appears to modulate the interaction of XPC with damaged sites by depositing H3K4me3 [54]. CPD recognition by XPC further depends on an auxiliary lesion receptor termed ‘damaged DNA-binding 2’ (DDB2) [78]. In turn, XPC recruits transcription factor IIH (TFIIH) to carry out a damage verification mechanism whereby XPD, the DNA helicase subunit of TFIIH, aided by XPB unwinds the double helix and scans the damaged strand for abnormalities [79]. This helicase activity of XPD is required for repair but not transcription. Demarcation of the lesion by XPD generates an unwound intermediate that is stabilized by XPA and replication protein A (RPA) and subjected to DNA incision by the endonucleases

would be the sequence-specific action of PRDM9, determined by the binding preference of its multiplex zinc-finger domain [56–58], whereas the GG-NER process displays DNA damage selectivity, conferred for UV lesions by the damaged DNA-binding 2 (DDB2) lesion receptor (Figure 3).

The aforementioned preference of XPC protein for sites marked by H3K4me3 strengthens the case for a SET-methyltransferase regulating this repair pathway. The human genome encodes multiple SET-domain methyltransferases that can methylate histone H3 [59]. The histone methyltransferase absent, small or homeotic 1 (ASH1) was originally identified in *Drosophila* as an antagonist of the Polycomb-induced silencing of *Hox* genes governing body patterning during development [60,61]. An essential function in embryonic development and cell differentiation has also been attributed to the mammalian homolog **ASH1-like** (ASH1L) [62,63]. Additionally, ASH1L has been implicated in the homeostasis of adult mammalian tissues, including the skin [64], instigating the hypothesis that this histone methyltransferase may participate in cellular responses to UV radiation, to which the human skin is constantly exposed. Besides methylating H3K36 [65,66], ASH1L also generates H3K4me3 [63,67].

Depletion of ASH1L causes UV hypersensitivity and substantially reduces the excision of CPDs in UV-irradiated human cells in culture. This repair inhibition appears to be entirely due to a less efficient GG-NER subpathway, whereas the TC-NER counterpart is not dependent on ASH1L [54], although the exact role of ASH1L and its histone methylation products during the GG-NER reaction remains under investigation. Protein dynamics experiments, based on protein trafficking between different-sized spots of UV damage in the nuclei of human cells, suggest that ASH1L, by methylating histones, modulates the interaction of XPC protein with damaged sites in such a way as to stimulate the GG-NER pathway [54] (Figure 3). In the absence of ASH1L, XPC still binds to chromatin and retains the ability to relocate to UV lesions but persists at UV-damaged sites without being able to recruit downstream factors leading to the engagement of endonucleases [54]. These findings are consistent with the hypothesis that ASH1L, through its histone methyltransferase activity, provides a docking site for GG-NER factors to control their associated nucleolytic activity.

Concluding remarks

The purpose of this opinion article is to reconcile the function of H3K4me3 marks in seemingly unrelated processes involving the induction of DNA breaks as reaction intermediates (Figure 2). One of these nucleolytic processes is the GG-NER response for the removal of DNA adducts. Indeed, exposure to physical or chemical carcinogens may generate tens of thousands of such adducts in each copy of the genome [68]. If not tightly regulated, the GG-NER response constitutes, in principle, a deleterious reaction to genotoxic insults, because its mechanism involves the transient formation of ubiquitous single-strand breaks with a potential to be converted to cytotoxic double-strand breaks. Thus, an uncontrolled GG-NER reaction could be pathological by causing chromosomal breaks that would ultimately jeopardize rather than protect genome stability. We postulate that the risk of excessive DNA cleavage is mitigated by a gatekeeper function of H3K4me3 marks, which determine and prioritize the sites of GG-NER activity. This hypothesis accommodates the observation that XPC, the initiator of the GG-NER pathway, displays a preference for H3K4me3, that gene promoters rich in H3K4me3 are rapidly cleared of DNA adducts, and that the ASH1L writer stimulates GG-NER activity. Key experiments to verify this hypothesis are outlined in the Outstanding questions (see Outstanding questions).

We hypothesize that, just as PRDM9 reroutes meiotic recombination away from gene promoters, the *de novo* deposition of histone methylation marks by ASH1L at sequences outside the typical H3K4me3 enrichment in promoters helps to prioritize other regions for GG-NER activity to

Outstanding questions

To test the hypothesis that SET-domain methyltransferase-deposited H3K4me3 has a role in mediating DNA repair, future experiments should entail comparisons of H3K4me3 deposition as well as repair factor binding distributions (obtainable, for example, by chromatin-immunoprecipitation combined with sequencing) in cells proficient or deficient for a specific methyltransferase activity.

Cohesin-dependent loop extrusion mediates the juxtaposition of functional elements, thus facilitating complex genomic processes. However, it is generally not clear how the 3D genome compaction resulting from TADs, sub-TADs, and other chromatin structures is reconcilable with the need for DNA accessibility during the execution of genome-wide DNA repair programs, such as GG-NER.

How the deposition of H3K4me3 or other methylation marks may reshape the 3D genome in response to genotoxic stress merits further research. For example, it is not clear how XPC and its damage co-sensors may be diverted away from constitutive binding sites in promoters or other functional sequences, once these sites have been repaired with highest priority, to allow for the subsequent repair of other genomic regions.

ASH1L may be responsible for depositing not only H3K4me3, but also H3K36me2 or H3K36me3, which are histone marks with a broader distribution than H3K4me3. Do these H3K36me2/3 modifications further support the reorganization of the 3D genome architecture for GG-NER to take place?

ASH1L has been identified as a player in the cellular response to UV lesions, whereby it acts to stabilize exposed tissues including the skin. Are ASH1L or any other histone methyltransferases involved in the processing of other GG-NER substrates, primarily bulky DNA adducts, induced by chemical carcinogens?

What is the spatiotemporal distribution and extension of accessible repair spaces or repair hotspots in chromatin formed as a consequence of histone

maintain overall genome integrity. The localization of the GG-NER machinery to **chromosomal interacting domains** (CIDs) in budding yeast, although orders of magnitude shorter than the megabase-long TADs in mammals, illustrates that the reorganization of the genome into specialized and permissive substructures may facilitate the repair of DNA adducts [69]. Thus, in response to UV irradiation, a surplus of H3K4me3 may transiently arise outside of promoters and other constitutively H3K4me3-rich elements to temporally reshape the 3D genomic architecture, reorganizing TADs, sub-TADs or other compartments to generate open spaces or at least local hotspots for 'big-enzyme' machines, such as the GG-NER complex.

The evolutionary pressure, particularly in long-lived species, to keep low mutation rates [70] and, hence, detect and remove DNA adducts throughout the entirety of the genome raises several questions the resolution of which holds crucial implications for future research on the role of histone marks and the 3D genome in guiding DNA repair and other nucleolytic processes.

Acknowledgments

We apologize to all colleagues whose important and interesting findings could not be cited because of reference number limitations. Research in the authors' laboratory is supported by the Swiss National Science Foundation (grant 310030-189125 to H.N.). S.H.R and P.V.E. are supported by the BBSRC (BB/R00756X/1).

Declaration of interests

None declared.

References

- Nurk, S. *et al.* (2022) The complete sequence of a human genome. *Science* 376, 44–53
- Pombo, A. and Dillon, N. (2015) Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* 16, 245–257
- Misteli, T. (2020) The self-organizing genome: principles of genome architecture and function. *Cell* 183, 28–45
- Rao, S.S.P. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680
- Emerson, D.J. *et al.* (2022) Cohesin-mediated loop anchors confine the locations of human replication origins. *Nature* 606, 812–819
- Nanni, L. *et al.* (2020) Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries. *Genome Biol.* 21, 197
- Robertson, A.G. *et al.* (2008) Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res.* 18, 1906–1917
- Bernstein, B.E. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326
- Prakash, K. and Fournier, D. (2018) Evidence for the implication of the histone code in building the genome structure. *BioSystems* 164, 49–59
- Calo, E. and Wysocka, J. (2013) Modification of enhancer chromatin: what, how, and why? *Mol. Cell* 49, 825–837
- Fu, W. *et al.* (2016) Structural basis for substrate preference of SMYD3, a SET domain-containing protein lysine methyltransferase. *J. Biol. Chem.* 291, 9173–9180
- Talbert, P.B. and Henikoff, S. (2021) The yin and yang of histone marks in transcription. *Annu. Rev. Genom. Hum. Genet.* 22, 147–170
- Deshpande, N. *et al.* (2022) Histone 3 lysine 4 monomethylation supports activation of transcription in *S. cerevisiae* during nutrient stress. *Curr. Genet.* 68, 181–194
- Bannister, A.J. *et al.* (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124
- Lachner, M. *et al.* (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120
- Nakayama, J. *et al.* (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292, 110–113
- Francis, N.J. *et al.* (2004) Chromatin compaction by a Polycomb group protein complex. *Science* 306, 1574–1577
- Santos-Rosa, H. *et al.* (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411
- Hödl, M. and Basler, K. (2012) Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr. Biol.* 22, 2253–2257
- Douillet, D. *et al.* (2020) Uncoupling histone H3K4 trimethylation from developmental gene expression via an equilibrium of COMPASS, Polycomb and DNA methylation. *Nat. Genet.* 52, 615–625
- Park, S. *et al.* (2020) Broad domains of histone H3 lysine 4 trimethylation in transcriptional regulation and disease. *FEBS J.* 287, 2891–2902
- Kassis, J.A. *et al.* (2017) Polycomb and Trithorax group genes in *Drosophila*. *Genetics* 206, 1699–1725
- Yan, J. *et al.* (2018) Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. *Cell Res.* 28, 204–220
- Yuen, K.C. *et al.* (2017) Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Sci. Adv.* 3, e1700191
- Chen, K. *et al.* (2015) Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* 47, 1149–1157
- Cao, F. *et al.* (2017) Super-enhancers and broad H3K4me3 domains form complex gene regulatory circuits involving chromatin interactions. *Sci. Rep.* 7, 2186
- Thibodeau, A. *et al.* (2017) Chromatin interaction networks revealed unique connectivity patterns of broad H3K4me3 domains and super enhancers in 3D chromatin. *Sci. Rep.* 7, 14466
- Sommermeier, V. *et al.* (2013) Spp1, a member of the Set1 complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Mol. Cell* 49, 43–54
- Chen, Y. *et al.* (2020) Refined spatial temporal epigenomic profiling reveals intrinsic connection between PRDM9-mediated H3K4me3 and the fate of double-stranded breaks. *Cell Res.* 30, 256–268
- Walker, M. *et al.* (2015) Affinity-seq detects genome-wide PRDM9 binding sites and reveals the impact of prior chromatin

methylation in response to genotoxic stress?

Are there relevant species differences in the adjustments of the 3D genome in response to genotoxic insults? For example, are there more extensive rearrangements of the 3D genome in species with low mutation rates, such as humans, compared with species with high overall mutation rates, such as mice?

- modifications on mammalian recombination hotspot usage. *Epigenetics Chromatin* 8, 31
31. Tock, A.J. and Henderson, I.R. (2018) Hotspots for initiation of meiotic recombination. *Front. Genet.* 9, 521
 32. Brick, K. *et al.* (2012) Genetic recombination is directed away from functional genomic elements in mice. *Nature* 485, 642–645
 33. Mihola, O. *et al.* (2019) Histone methyltransferase PRDM9 is not essential for meiosis in male mice. *Genome Res.* 29, 1078–1086
 34. Schwarzkopf, E.J. and Cornejo, O.E. (2022) PRDM9-directed recombination hotspots depleted near meiotically transcribed genes. *Gene* 813, 146123
 35. Liu, Y. *et al.* (2007) A plant homeodomain in Rag-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* 27, 561–571
 36. Matthews, A.G.W. *et al.* (2007) RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450, 1106–1110
 37. Ramón-Maiques, S. *et al.* (2007) The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18993–18998
 38. Shimazaki, N. *et al.* (2009) H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in *trans* in addition to tethering in *cis*: implications for translocations. *Mol. Cell* 34, 535–544
 39. Stanlie, A. *et al.* (2010) Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22190–22195
 40. Daniel, J.A. and Nussenzweig, A. (2012) Roles for histone H3K4 methyltransferase activities during immunoglobulin class-switch recombination. *Biochim. Biophys. Acta Gene Regul. Mech.* 1819, 733–738
 41. Husain, A. *et al.* (2016) Chromatin remodeler SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. *Nat. Commun.* 7, 10549
 42. Huang, J.C. *et al.* (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc. Natl. Acad. Sci. U. S. A.* 89, 3664–3668
 43. Araújo, S.J. *et al.* (2000) Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev.* 14, 349–359
 44. Hu, J. *et al.* (2017) Dynamic maps of UV damage formation and repair for the human genome. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6758–6763
 45. Ramanathan, B. and Smerdon, M.J. (1989) Enhanced DNA repair synthesis in hyperacetylated nucleosomes. *J. Biol. Chem.* 264, 11026–11034
 46. Gsell, C. *et al.* (2020) A chromatin scaffold for DNA damage recognition: how histone methyltransferases prime nucleosomes for repair of ultraviolet light-induced lesions. *Nucleic Acids Res.* 48, 1652–1668
 47. Kakumu, E. *et al.* (2017) Xeroderma pigmentosum group C protein interacts with histones: regulation by acetylated states of histone H3. *Genes Cells* 22, 310–327
 48. Kim, J.K. *et al.* (1995) Contrasting structural impacts induced by *cis-syn* cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity. *Photochem. Photobiol.* 62, 44–50
 49. Spivak, G. (2016) Transcription-coupled repair: an update. *Arch. Toxicol.* 90, 2583–2594
 50. Kuzminov, A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8241–8246
 51. Okano, S. *et al.* (2003) Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol. Cell Biol.* 23, 5472
 52. Le May, N. *et al.* (2010) NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol. Cell* 38, 54–66
 53. Bidon, B. *et al.* (2018) XPC is an RNA polymerase II cofactor recruiting ATAC to promoters by interacting with E2F1. *Nat. Commun.* 9, 2610
 54. Balbo Pogliano, C. *et al.* (2017) ASH1L histone methyltransferase regulates the handoff between damage recognition factors in global-genome nucleotide excision repair. *Nat. Commun.* 8, 1333
 55. Adar, S. *et al.* (2016) Genome-wide kinetics of DNA excision repair in relation to chromatin state and mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* 113, E2124–E2133
 56. Huang, T. *et al.* (2020) The histone modification reader ZCWPW1 links histone methylation to PRDM9-induced double-strand break repair. *eLife* 9, e53459
 57. Grey, C. *et al.* (2011) Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. *PLoS Biol.* 9, e1001176
 58. Diagouraga, B. *et al.* (2018) PRDM9 methyltransferase activity is essential for meiotic DNA double-strand break formation at its binding sites. *Mol. Cell* 69, 853–865
 59. Crump, N.T. and Milne, T.A. (2019) Why are so many MLL lysine methyltransferases required for normal mammalian development? *Cell. Mol. Life Sci.* 76, 2885–2898
 60. Shearn, A. (1989) The ash-1, ash-2 and trithorax genes of *Drosophila melanogaster* are functionally related. *Genetics* 121, 517–525
 61. Klymenko, T. and Müller, J. (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* 5, 373–377
 62. Brinkmeier, M.L. *et al.* (2015) The histone methyltransferase gene absent, small, or homeotic discs-1 like is required for normal Hox gene expression and fertility in mice. *Biol. Reprod.* 93, 121
 63. Yin, B. *et al.* (2019) Epigenetic control of mesenchymal stem cell fate decision via histone methyltransferase Ash1L. *Stem Cells* 37, 115–127
 64. Li, G. *et al.* (2017) The histone methyltransferase Ash1L is required for epidermal homeostasis in mice. *Sci. Rep.* 7, 45401
 65. Zhu, L. *et al.* (2016) ASH1L links histone H3 lysine 36 dimethylation to MLL leukemia. *Cancer Discov.* 6, 770–783
 66. Miyazaki, H. *et al.* (2013) Ash1L methylates lys36 of histone H3 independently of transcriptional elongation to counteract Polycomb silencing. *PLoS Genet.* 9, e1003897
 67. Gregory, G.D. *et al.* (2007) Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol. Cell Biol.* 27, 8466–8479
 68. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* 362, 709–715
 69. van Eijk, P. *et al.* (2019) Nucleosome remodeling at origins of global genome-nucleotide excision repair occurs at the boundaries of higher-order chromatin structure. *Genome Res.* 29, 74–84
 70. Cagan, A. *et al.* (2022) Somatic mutation rates scale with lifespan across mammals. *Nature* 604, 517–524
 71. Myers, S. *et al.* (2010) Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* 327, 876–879
 72. Parvanov, E.D. *et al.* (2010) Prdm9 controls activation of mammalian recombination hotspots. *Science* 327, 835
 73. Baudat, F. *et al.* (2013) Meiotic recombination in mammals: localization and regulation. *Nat. Rev. Genet.* 14, 794–806
 74. Hayashi, K. *et al.* (2005) A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* 438, 374–378
 75. Komberg, R.D. and Lorch, Y. (2020) Primary role of the nucleosome. *Mol. Cell* 79, 371–375
 76. Dixon, J.R. *et al.* (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380
 77. Min, J.H. and Pavletich, N.P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* 449, 570–575
 78. Sugawara, K. *et al.* (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* 121, 387–400
 79. Mathieu, N. *et al.* (2013) DNA quality control by a lesion sensor pocket of the Xeroderma pigmentosum group D helicase subunit of TFIIH. *Curr. Biol.* 23, 204–212
 80. Staresinic, L. *et al.* (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J.* 28, 1111–1120