

Review

Biology and Physics of Heterochromatin-Like Domains/Complexes

Prim B. Singh ^{1,2,*} , Stepan N. Belyakin ^{2,3}  and Petr P. Laktionov ^{2,3}

¹ Nazarbayev University School of Medicine, Nur-Sultan City 010000, Kazakhstan

² Epigenetics Laboratory, Department of Natural Sciences, Novosibirsk State University, 2 Pirogova St., 630090 Novosibirsk, Russia

³ Genomics laboratory, Institute of molecular and cellular biology SD RAS, Lavrentyev ave, 8/2, 630090 Novosibirsk, Russia; belyakin@mcb.nsc.ru (S.N.B.); laktionov@mcb.nsc.ru (P.P.L.)

* Correspondence: prim.singh@nu.edu.kz; Tel.: +7-7172-694706

Received: 9 July 2020; Accepted: 4 August 2020; Published: 11 August 2020



Abstract: The hallmarks of constitutive heterochromatin, HP1 and H3K9me_{2/3}, assemble heterochromatin-like domains/complexes outside canonical constitutively heterochromatic territories where they regulate chromatin template-dependent processes. Domains are more than 100 kb in size; complexes less than 100 kb. They are present in the genomes of organisms ranging from fission yeast to human, with an expansion in size and number in mammals. Some of the likely functions of domains/complexes include silencing of the donor mating type region in fission yeast, preservation of DNA methylation at imprinted germline differentially methylated regions (gDMRs) and regulation of the phylotypic progression during vertebrate development. Far *cis*- and *trans*-contacts between micro-phase separated domains/complexes in mammalian nuclei contribute to the emergence of epigenetic compartmental domains (ECDs) detected in Hi-C maps. A thermodynamic description of micro-phase separation of heterochromatin-like domains/complexes may require a gestalt shift away from the monomer as the “unit of incompatibility” that determines the sign and magnitude of the Flory–Huggins parameter, χ . Instead, a more dynamic structure, the oligo-nucleosomal “clutch”, consisting of between 2 and 10 nucleosomes is both the long sought-after secondary structure of chromatin and its unit of incompatibility. Based on this assumption we present a simple theoretical framework that enables an estimation of χ for domains/complexes flanked by euchromatin and thereby an indication of their tendency to phase separate. The degree of phase separation is specified by χN , where N is the number of “clutches” in a domain/complex. Our approach could provide an additional tool for understanding the biophysics of the 3D genome.

Keywords: HP1; H3K9me_{2/3}; epigenetic compartmental domains; block copolymers; Flory–Huggins parameter χ ; unit of incompatibility

1. Introduction

There is an intimate relationship between gene regulation, chromatin structure and genome organisation [1]. The kernel from which our understanding of this relationship grew can be found in studies on constitutive heterochromatin, especially with the phenomenon of position-effect variegation (PEV) in *Drosophila* (for reviews see [2–5]). Nine decades of work on PEV concluded that changing the position of a gene with respect to the heterochromatin–euchromatin boundary can affect its chromatin structure and that, in turn, affects its expression as manifest by phenotypic variegation and changes in transcription [6–10]. Saturation mutagenesis screens identified modifiers of PEV and molecular characterisation of their wild-type (wt) gene products showed they encode structural components and enzymatic activities that regulate the assembly of heterochromatin [5,11–13]. Two of

the modifiers encode Heterochromatin Protein 1 (HP1) and the Su(var)3-9 H3K9HMTase that generates the H3K9me_{2/3} histone modification bound by the chromo domain (CD) of HP1 [14] (Figure 1A,B; H3K9me_{2/3} will be termed H3K9me₃ hereafter [15]). HP1 and the H3K9me₃ modification are highly conserved across eukaryotes and represent hallmarks of constitutive heterochromatin [16,17] that are enriched at constitutively heterochromatic chromosomal territories of nearly all eukaryotic chromosomes. These territories include peri-centric heterochromatin surrounding the centromeres, (sub-)telomeric and (peri-)nucleolar organiser regions (NORs), with both hallmarks being found at these sites in organisms as distantly related as fission yeast through *Drosophila* to human [18–25]. Notable exceptions are the chromosomes of budding yeast where the silent information regulator (Sir) complex is assembled at heterochromatic territories (telomeres and NORs) by establishing and recognising a pattern of de-acetylated histones, especially hypo-acetylated H4K16 [26]. Co-localisation of HP1 and H3K9me₃ to constitutive heterochromatin arose early in the evolution of eukaryotes with the common ancestor of fission yeast and human living around one billion years ago [27]; budding and fission yeasts diverged from each other at around the same time both did from human [28].

The interaction of HP1 proteins with H3K9me₃ has been resolved at the atomic level and can be illustrated using mammalian HP1 β , an archetypal HP1 protein [29]. The primary structure of HP1 β is identical in human and mouse [30]. It is essential; HP1 β null mutant mice die at birth [31]. HP1 β is small at around 25 kD having an *N*-terminal CD and a sequence-related domain towards the *C*-terminus called the chromo shadow domain (CSD) [32] (c.f. Figure 1A vs. Figure 1C). These two domains are likely to have arisen by gene duplication [33] and are separated by a less-well conserved “hinge” region (HR) that is flexible and lacks a defined structure [34]. Both the CD and the CSD represent globular protein modules with a diameter of around 30 Å. The CD binds the methylated H3K9 tail [35] where three conserved aromatic residues, Tyr21, Trp42 and Phe45, form an “aromatic cage” around the methyl-ammonium moiety (Figure 1B). Most of the binding energy is driven by cation– π interactions where the cation methyl ammonium moiety is attracted to the negative electrostatic potential of the aromatic groups’ π -system [36]. The HP1 β CSD dimerises in solution with the dimer centring upon helix α 2 (Figure 1D), which interacts symmetrically and at an angle of 35° with helix α 2 of the adjacent CSD subunit and forms a non-polar pit that can accommodate penta-peptides with the consensus sequence motif PxVxL that is found in HP1-interacting proteins [37,38] (Figure 1C,D).

HP1 and H3K9me₃ co-localise *outside* constitutive heterochromatin as constituents of heterochromatin-like domains and complexes that are thought to regulate chromatin template-dependent processes [30,41,42]. There is evidence that domains/complexes are widespread in mammalian genomes [43] but the question of their number, size and function remains open. Here, we address this question and extend our answer using bioinformatics approach that interrogates the genomes of fission yeast, fruit fly, mouse and human. We focus on domains/complexes that regulate critical chromatin template-dependent processes during mammalian development and discuss how segregation of micro-phase separated domains/complexes could drive the compartmentalisation observed in Hi-C experiments. Treating domain/complexes as blocks in a block copolymer (BCP) that micro-phase separate from euchromatin implies that domains/complexes possess a value for χ , the Flory–Huggins parameter. We present a simple theoretical “clutch” model, where the “unit of incompatibility” of chromatin is an oligo-nucleosomal “clutch” of 2–10 nucleosomes, which could provide an approach for determining χ for domains/complexes experimentally. The magnitude of the product of χ and the number of clutches (N) in a domain/complex (χ N) specifies the degree of phase separation of the domain/complex from euchromatin.

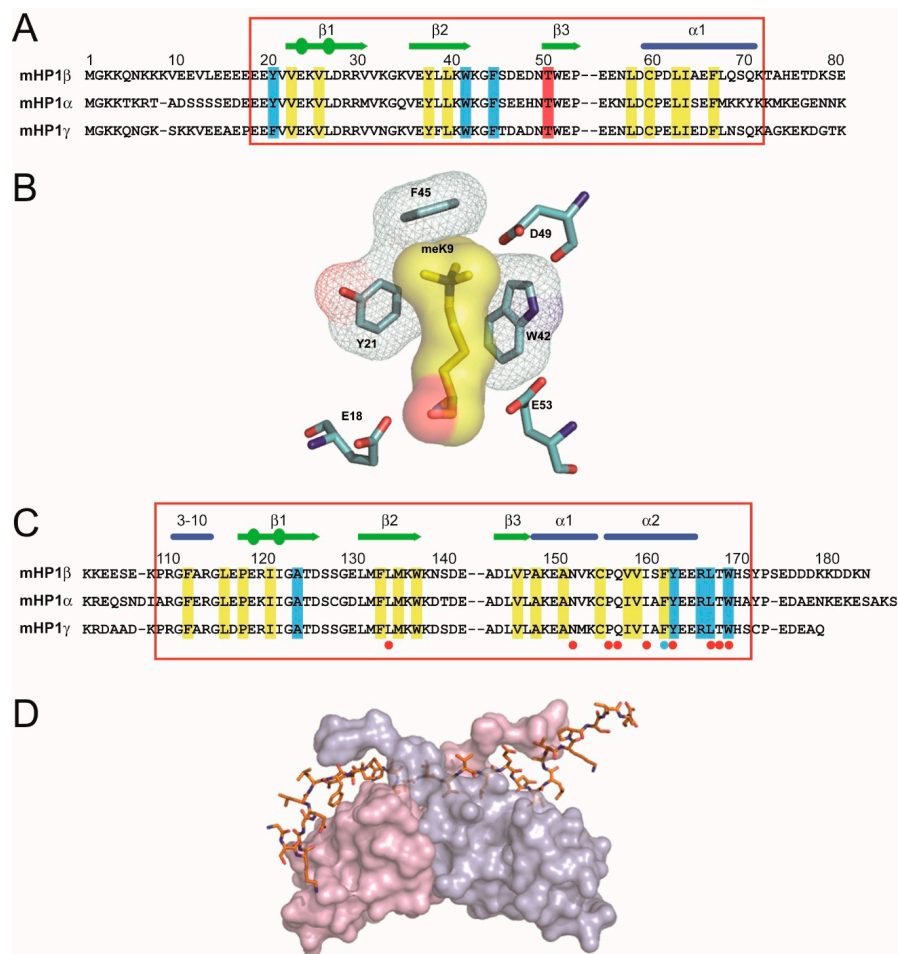


Figure 1. Alignment of peptide sequences of the murine HP1 CDs (A) and the CSDs (C), structure of the aromatic “cage” formed around the K9me3 moiety (B) and the CSD dimer bound to PxVxL motif in CAF-1 peptide (D). (A) The CDs from top to bottom are HP1β(amino-acids 1–80), HP1α (amino-acids 1–80) and HP1γ (amino-acids 1–80). The red box denotes the structured part of the CD. The secondary structure elements of the HP1β CD are displayed above the sequence: blue cylinders represent a α -helix (α 1) and green arrows represent β -strands; circles within the arrows indicate β -bulges. The residues that make up the hydrophobic core of the CD are shaded in yellow and the aromatic residues that form a notional “cage” around the methyl lysine are given in blue. The Thr51 residue that is phosphorylated after DNA damage [39] is shaded in red. (B) Binding of the HP1β CD to H3K9me3 a notional aromatic ‘cage’ is formed from three conserved aromatic residues: Tyr21, Trp42 and Phe45. The interaction between the methyl-ammonium moiety and the aromatic cage is largely electrostatic where the positively charged (cation) moiety is attracted to the negative electrostatic potential of the aromatic groups’ π -system [36]. (C) The CSDs from top to bottom are HP1β (amino-acids 103–185), HP1α (amino-acids 106–191) and HP1γ (amino-acids 97–173). The red box denotes the structured part of the CSD. The secondary structure elements of the HP1β CSD are given above the sequence, where the cylinders represent the α -helices (α 1 and α 2) and the arrows represent β -strands; circles within the arrows indicate β -bulges. The residues that make up the hydrophobic core of the CSD are shaded in yellow and show good alignment with those found in the CD indicating a similar overall structure of the CSD to the CD. Residues that are involved in binding the PxVxL motif are shaded in blue; there is a blue dot below the Phe163 residue, which is involved both in the structure of the CSD and in binding to the peptide. There are red dots below the residues that are involved in CSD:CSD dimerisation. (D) Surface view of the CSD homodimer (one monomer in pink and the other in blue) bound to the CAF-1 peptide (shown as a stick model) containing the PxVxL motif, which is involved in intermolecular β pairing with both monomers [37]. Taken and modified from [40].

2. Heterochromatin-Like Domains/Complexes in Eukaryotes

As a first step towards determining the number and size of heterochromatin-like domains/complexes we first investigated the degree to which HP1 and H3K9me3 co-localise *outside* constitutive heterochromatin. This was done for four (distantly related) genomes, namely man, mouse, *Drosophila melanogaster* and fission yeast. When constitutive heterochromatin is excluded, Pearson correlation coefficients for the co-localisation of HP1 α , β and γ with H3K9me3 across the rest of the human genome are 0.73 (α ; human H1 ES cells), 0.74 (β ; 293T cells) and 0.77 (γ ; human H1 ES cells) (Figure 2A–C). The same analysis in mouse ES cells gave correlations of 0.63 for HP1 α , 0.69 for HP1 β and 0.71 for HP1 γ (Figure 2D–F). The mouse results are in agreement with a previous study, which showed that HP1 β is preferentially targeted to genomic regions with high local concentrations of H3K9me3 in murine ES cells (correlation coefficient 0.77; [29]). In *Drosophila melanogaster* ovaries the correlation of HP1a with H3K9me3 outside constitutive heterochromatin has a coefficient of 0.92 (Figure 2G). In fission yeast the correlation of Swi6^{HP1} with H3K9me3 is weaker at 0.53 (Figure 2H). This likely reflects the finding that only a limited set of loci outside constitutive heterochromatin are marked by both H3K9me3 and Swi6^{HP1}, including the mating-type region, a variety of repeat elements and a number of meiotic genes [18]. If the correlation is made over the entire fission yeast genome the correlation coefficient increases to 0.93 (Figure 2I). For constitutive heterochromatin alone the coefficient approaches unity (0.99; Figure 2J).

These data indicate there is a correlation of HP1s with H3K9me3 outside constitutive heterochromatin but that correlation is not absolute. This is consistent with the observation that HP1 does not always “follow code” and can localise to (hetero)chromatin in the absence of H3K9me3 [44]. For example, in the mouse, the CHD4-ADNP-HP1 complex that contains HP1 represses gene expression locally by establishing inaccessible chromatin around its DNA-binding sites and does not depend on H3K9me3-modified nucleosomes [45]. Similarly, in *Drosophila*, a variety of protein partners can localise HP1a to euchromatic sites in the absence of H3K9me3 [46]; localisation of HP1a to telomeric constitutive heterochromatin in *Drosophila* is independent of the presence of H3K9me3 at the telomere [21,47]. Alternative modes of HP1 binding to chromatin that do not involve H3K9me3 have also been documented, for example, the HP1 CD or CSD with the H3 histone ‘core’ [48–50], HP1 binding to histone H1 [48,51], binding of the HR region to DNA and RNA [52–54] and a non-specific electrostatic interaction of the HP1 N-terminal extension with the H3 tail [29]. Of note is the recent demonstration that the Swi6^{HP1} CSD dimer binds to the H2B α 1 helix where it is thought to destabilise the nucleosome and promote phase separation of constitutive heterochromatin in fission yeast [55]. Our survey (Figure 2A–J) indicates that the hallmarks of constitutive heterochromatin (HP1 and H3K9me3) also co-localise at many sites outside constitutive heterochromatin where heterochromatin-like domains/complexes are likely to be assembled. HP1 proteins bound to H3K9me3-marked domains/complexes are constantly exchanging with unbound HP1 proteins in the nucleoplasm. Almost the entire pool of HP1 proteins outside constitutive heterochromatin turns over in around 10 s ($t_{1/2} = 1\text{--}10$ s; [56,57]); constant exchange maintains compaction of domains/complexes.

We next determined the number and size of the domains/complexes in the four organisms used for the correlation analysis where domains/complexes were put into three categories according to size. Domains are > 1 Mb and between 0.1 and 1 Mb with complexes less than 0.1 Mb down to 10 kb (Table 1). Our survey showed that the number and size of heterochromatin-like domains/complexes increases from fission yeast to fruit flies with a sharp increase from insects to mammals. For human cells we used two different cell lines to estimate the number of HP1 α / β / γ -containing domains/complexes. For the human H1 ES cell the HP1 α and γ distributions were intersected with H3K9me3 for the whole genome and the genome outside constitutive heterochromatin (Table 1). This showed that outside constitutive heterochromatin there were 163 HP1 α / γ -containing heterochromatin-like domains and 18,853 HP1 α / γ -containing complexes, with 90% of the complexes being in the range 10–30 kb (Table S1). In human 293T cells there were 859 HP1 α / β -containing domains and 32,292 HP1 α / β -containing complexes, with 75% of the complexes in the range 10–30 kb (Table S1). We conclude there are,

conservatively, between 163 and 859 heterochromatin-like domains and 18,853–32,292 complexes in humans depending on cell type. These values may be an underestimate because we could only mine data for two HP1 isoforms per human cell line. For the mouse, we obtained data for all three HP1 isoforms in a single ES cell line, which revealed no domains greater than 1 Mb but outside heterochromatin there were 622 HP1 $\alpha/\beta/\gamma$ -containing heterochromatin-like domains between 0.1 and 1 Mb and 10,227 HP1 $\alpha/\beta/\gamma$ -containing heterochromatin-like complexes, with around 60% of the complexes being in the range 10–30 kb (Table S1). In *Drosophila* there are two heterochromatin-like domains outside heterochromatin in addition to 161 complexes. In fission yeast a survey of the whole genome reveals only one heterochromatin-like domain that is larger than 0.1 Mb that most probably represents the centromeric constitutive heterochromatin of cen3 that is 110 kb in size [58]. Outside constitutive heterochromatin there are only 20 complexes of which the mating type region would be one [18].

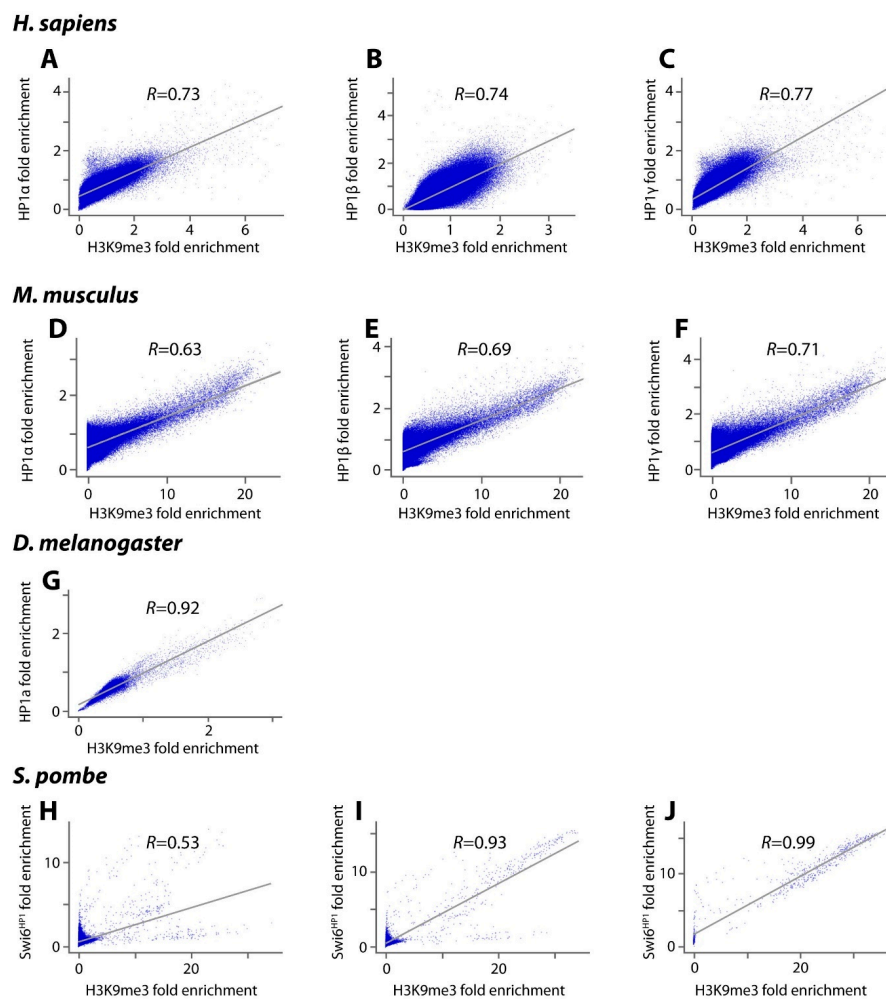


Figure 2. Plots of genome-wide correlation of HP1 proteins and H3K9me3 outside constitutive heterochromatin. The figure depicts fold enrichment of ChIP-seq genome profiles correlated using a 5 kb-sized window for *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster* and a 200 bp-sized window for *Schizosaccharomyces pombe*. Unless otherwise stated constitutive heterochromatin regions were excluded prior to analysis. (A–C) show correlations of HP1 α (H1 ES cells), HP1 β (293T cells) and HP1 γ (H1 ES cells) with H3K9me3 in human cells. (D–F) show correlations of HP1 α , HP1 β and HP1 γ with H3K9me3 in mouse ES cells; (G) shows correlation of *Drosophila* HP1 α with H3K9me3 in ovaries; (H) shows correlation of Swi6^{HP1} with H3K9me3 in *S. pombe*; (I) shows correlation of HP1 β with H3K9me3 over the whole genome of *S. pombe*; (J) shows correlation of Swi6^{HP1} with H3K9me3 over heterochromatin of *S. pombe*.

Table 1. Heterochromatin-like domains/complexes in *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster* and *Schizosaccharomyces pombe*.

Organism	Size, Mb	Cell Type Components	Whole Genome	Heterochromatic Regions Excluded
<i>H. sapiens</i>	>1	H1 ES cells (HP1 α + HP1 γ + H3K9me3)	48	4
	0.1–1	H1 ES cells (HP1 α + HP1 γ + H3K9me3)	345	159
	0.01–0.1	H1 ES cells (HP1 α + HP1 γ + H3K9me3)	19,550	18,853
<i>M. musculus</i>	>1	293T cells (HP1 α + HP1 β + H3K9me3)	24	4
	0.1–1	293T cells (HP1 α + HP1 β + H3K9me3)	1027	855
	0.01–0.1	293T cells (HP1 α + HP1 β + H3K9me3)	33,754	32,292
	>1	ES cells (HP1 α + HP1 β + HP1 γ + H3K9me3)	0	0
	0.1–1	ES cells (HP1 α + HP1 β + HP1 γ + H3K9me3)	1059	622
	0.01–0.1	ES cells (HP1 α + HP1 β + HP1 γ + H3K9me3)	12,675	10,227
<i>D. melanogaster</i>	>1	Ovaries (HP1a + H3K9me3)	7	0
	0.1–1	Ovaries (HP1a + H3K9me3)	27	2
	0.01–0.1	Ovaries (HP1a + H3K9me3)	183	161
<i>S. pombe</i>	>1	(Swi6 ^{HP1} + H3K9me3)	0	0
	0.1–1	(Swi6 ^{HP1} + H3K9me3)	1	0
	0.01–0.1	(Swi6 ^{HP1} + H3K9me3)	23	20

Heterochromatin-like domains/complexes in *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster* and *Schizosaccharomyces pombe*. In man, the size and number of domains (>1 Mb and between 0.1 and 1 Mb) and complexes (0.01–0.1 Mb) were determined using ChIP-seq data for H1 ES cells (HP1 α + HP1 γ + H3K9me3) [59,60] and 293T cells (HP1 α + HP1 β + H3K9me3) [61,62]. For the mouse, the same analysis was undertaken in a single cell line, namely mouse ES cells (HP1 α + HP1 β + HP1 γ + H3K9me3) [45]. Data from ovaries were used for the fly (HP1a + H3K9me3) [63]. A culture of fission yeast cells was used to generate the ChIP-seq (Swi6^{HP1} + H3K9me3) [64] that were mined for our survey. Domains and complexes were calculated for the whole genome (including constitutive heterochromatin) and the genome without constitutive heterochromatin.

Our survey indicates that heterochromatin-like domains/complexes are widespread in the genomes of (distantly related) eukaryotes (Table 1). At the outset they were proposed primarily as a *general* mechanism for regulating chromatin template-dependent processes outside constitutive heterochromatin and secondarily as a mechanism for a *special* case where chromosomes and genes exhibit allele-specific parent-of-origin-specific behaviour [30,41,42,65]. Initial support for the general case came from immunofluorescence studies [66] and soon thereafter from the discovery that the universal co-repressor of KRAB-zinc finger proteins (KRAB-ZFPs), KAP1, recruits HP1 proteins to form localised heterochromatin-like domains/complexes [42,67]. At around the same time it was shown that the function of HP1-containing heterochromatin-like domain/complexes was conserved from fission yeast to human [16]. Detailed studies on the KRAB-Zinc finger gene (KRAB-ZNF) clusters on human chromosome 19 have provided insight into how large heterochromatin-like domains (up to 4 Mb; [68,69]) can be nucleated at particular sites by small heterochromatin-like complexes (\approx 6 kb; [43,68]). The domains so assembled make far *cis*-contacts to generate the B4 heterochromatic sub-compartment detected in Hi-C maps [43,70]. An intriguing characteristic of the heterochromatin-like domains that encompass the KRAB-ZNF clusters is that the genes within the clusters remain expressible [71]. The domains are thought to ‘protect’ the KRAB-ZNF gene repeats as they have expanded during evolution by preventing illegitimate recombination [68], rather than to repress and silence the repeats. It would seem that heterochromatin-like domains/complexes have a variety of chromatin template-dependent functions an observation supported by studies on HP1, which have shown it to be associated with gene activation as well as with loci involved in other nuclear functions, including transcriptional elongation, RNA splicing and DNA repair [39,72–78].

Support for the special case alluded to above came from the observation, to be described, that KRAB-ZFPs assemble heterochromatin-like complexes at “imprinted genes” that exhibit parent-of-origin-specific gene expression [79].

3. Heterochromatin-Like Complexes and Preservation of DNA Methylation at Imprinted gDMRs during Pre-Implantation Embryogenesis

Evidence that phenotypic traits could be subject to parent-of-origin effects came not long after the re-discovery of Mendel's laws of inheritance in 1900 [80], although Mendel himself thought it indubitable that reciprocal crosses were equivalent, saying: "... it is perfectly immaterial whether the dominant character belongs to the seed-bearer or to the pollen parent; the form of the hybrid remains identical in both cases" [81]. We now know of many instances where this is not the case, where the behaviour of chromosomes and genes are dependent upon ancestry. Parent-of-origin-specific behaviour of chromosomes was first observed in insects, in *Sciara* (reviewed by Metz in 1938 [82]) and *Coccidae* [83]. In mammals, pronuclear transfer experiments confirmed that the parental contributions to the zygote were genetically but not functionally equivalent [84,85]. These experiments led to the suggestion that the expression of certain genes, called imprinted genes, was dependent upon parental origin. There are now known to be around 100 imprinted genes that exhibit mono-allelic parent-of-origin-specific expression in mouse and human [86,87]. Such genes are said to be subject to genomic imprinting [88,89]. Genomic imprinting results in genes (or gene clusters) that are either maternally or paternally imprinted. Maternally imprinted genes are associated with a maternal-specific "mark" that acts *in cis* such that there is a heritable (cell-to-cell) change in the behaviour of the gene. The same is true for paternally imprinted genes except that the "mark" is specific for the paternal allele. Genomic imprinting is necessarily reversible, thus epigenetic, because the parental alleles of an imprinted gene are marked differently in the soma, but the marks must be erased in the germ-line so that both alleles can then be marked again, this time according to the sex of the parent. Imprinting ("marking") takes place when the parental genomes are separate, which occurs in the respective germ-lines and during the brief period when the pro-nuclei lie separately in the ooplasm of the newly fertilised zygote.

In the mouse, depletion of DNA methylation of cytosine (5mC) in CpG dinucleotides leads to dysregulation of genomic imprinting [90]. Since differences in 5mC could be traced back to the sperm and egg it was concluded that the parent-of-origin-specific "mark" is DNA methylation [91]. Parent-of-origin-specific differences in CpG methylation are called gametic or germline differentially methylated regions (gDMRs) and fall within a broader category of CpG-rich genomic regions called CpG islands (CGIs; [92]) that are widespread in the genome, with $\approx 70\%$ of annotated gene promoters in human being associated with a CGI [93]; within the bounds of definition both maternally and paternally imprinted gDMRs are recognisable as CGIs [94,95]. CpGs that are part of CGIs are usually unmethylated, whether the associated gene is active or inactive [92], but imprinted gDMRs are exceptions where CGIs are methylated. Accordingly, classical imprinted gDMRs are methylated in either the female (maternally imprinted) or the male (paternally imprinted) germline and retain this parent-of-origin specific methylation following fertilisation and during pre-implantation development [89]. In the mouse, there are ≈ 26 (23 maternal and 3 paternal) definitive imprinted gDMRs [96,97] (Figure 3A).

The identification of a small number of imprinted gDMRs during the pre-genomic era prompted efforts directed towards identifying hypothetical specialised sequence elements that would be recognised by specific *trans*-acting factors that target the *de novo* DNA methylation machinery to imprinted gDMRs during gametogenesis. However, later whole genome studies revealed a very different picture, which led to the abandonment of the notion that specific imprinting machinery operates in the germline. Mining the sperm and oocyte methylomes showed that there were many more gDMRs in the gametes, far above the number of definitive imprinted gDMRs. Depending on the study, the mouse oocyte nucleus contains around 1–2000 oocyte-specific gDMRs, with the sperm-specific gDMRs numbering between 185 and 818 [98,99] (Figure 3A). The combined total is roughly 1600 (imprinted and non-imprinted) gDMRs in the pro-nuclei of the newly fertilised zygote [99]. This is almost two orders of magnitude greater than the number (≈ 26) of classical definitive imprinted gDMRs whose methylation is retained after fertilisation and through pre-implantation embryogenesis [96,97] (Figure 3A). Putting it short, imprinted gDMRs are not uniquely targeted for DNA methylation; imprinted and non-imprinted gDMRs are methylated by mechanisms common to CGIs [100,101].

The difference is that allele-specific methylation at imprinted gDMRs is preserved after fertilisation while the differential methylation at non-imprinted gDMRs is not.

Pre-implantation embryogenesis in mouse and human is characterised by a global DNA demethylation of the parental genomes [102,103]. Demethylation is thought to ensure that the embryonic epigenome is purged of any barriers to pluripotency, which is essential for those cells of the inner cellular mass (ICM) that will go on to form the tissues and cell types of the embryo proper [104]. DNA methylation at the imprinted gDMRs is preserved during this DNA demethylation phase through the assembly of localised heterochromatin-like complexes at imprinted gDMRs [105,106]. It is the sequence-specific assembly of heterochromatin-like complexes at imprinted gDMRs that is the key to understanding genomic imprinting in mammals, consistent with the earlier proposal [30].

Preservation of methylation at imprinted gDMRs requires binding of KRAB zinc-finger proteins (KRAB-ZFPs), ZFP57 [79,107] and ZFP445 [108]. Methylation of all imprinted gDMRs, except one, is lost in ZFP57/445 mouse double mutants [108]. Characterisation of ZFP57 in the mouse has shown that binding of ZFP57 is methylation sensitive with its binding being a hexamer motif TGCCGC found in imprinted gDMRs, where the central CpG dinucleotide is methylated [107,109,110] (Figure 2B). Assembly of ZFP57/445-directed heterochromatin-like complexes is likely to take place soon after fertilisation since most of the constituents are laid down maternally (Figure 2C). In mouse oocytes there are maternal stores of ZFP57 and KAP1 and their loss affects methylation at imprinted gDMRs [79,111]. The Setdb1 HMTase that interacts with KAP1 is localised to peri-nucleolar rims of the pro-nuclei in the zygote [112] and maternal deletion of Setdb1 leads to dramatic defects in preimplantation development [113]; deletion in ES cells leads to DNA demethylation of imprinted gDMRs [114]. Setdb1 generates the H3K9me3 binding site for HP1 proteins that are also found in the oocyte cytoplasm and pro-nuclei of the early embryo [115,116]. The maintenance DNA methyltransferase Dnmt1 that interacts with KAP1 [107,117] is barely detectable in the ooplasm and in the early embryo [118,119]. Yet these trace amounts are essential for methylation of imprinted gDMRs because loss of maternal-zygotic Dnmt1 leads to loss of methylation at all imprinted gDMRs [120]. H4K20me3 catalysed by H4K20HMTases during oogenesis is present in the maternal pronucleus but is undetectable by immunofluorescence in nuclei at later preimplantation stages [116,121,122]. It has yet to be shown whether the (very) low levels of H4K20HMTase mRNAs present in the zygote and pre-implantation stages [121] are translated into active proteins and generate H4K20me3 at imprinted gDMRs through interactions with the resident heterochromatin-like complexes.

Once assembled, the complex preserves methylation at imprinted gDMRs throughout the demethylation phase, which is complete at the blastocyst stage where the lowest levels of global DNA methylation are reached (Figure 3C) [101,102,123]. Biochemical and functional studies using ES cells (derived from the blastocyst) have shown that the structural proteins and enzymatic activities of the heterochromatin-like complexes are present at imprinted gDMRs [107,114,124–126] (Figure 3D). H4K20me3 is clearly detected at imprinted gDMRs in ES cells [124], indicating the recruitment of an H4K20HMTase [125], most likely through a known interaction of H4K20HMTases with HP1 [127] (Figure 3D).

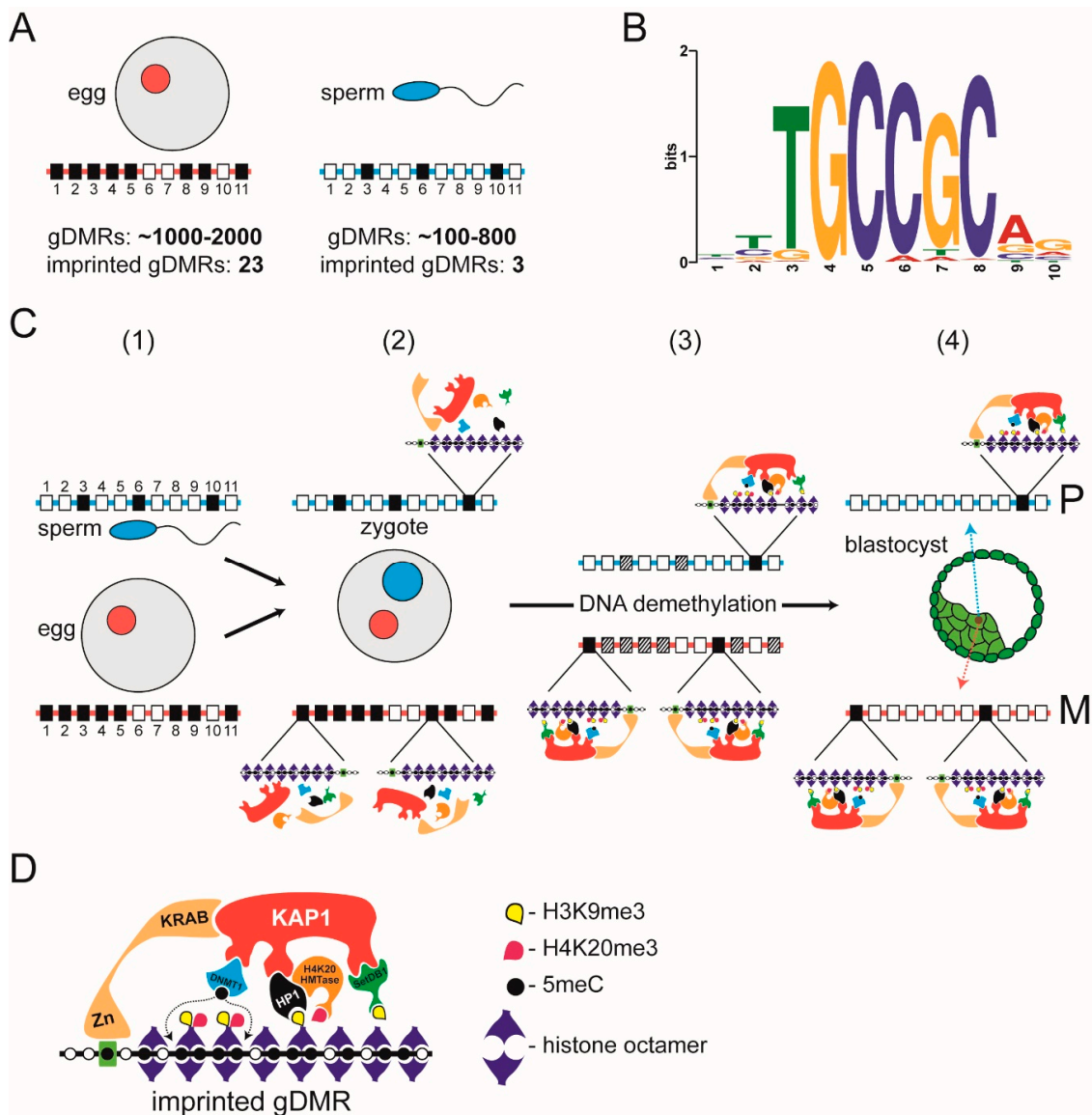


Figure 3. Preservation of DNA methylation at imprinted gDMRs by a localised heterochromatin-like complex. (A) There are many more (imprinted and non-imprinted) gDMRs than definitive imprinted gDMRs. The oocyte nucleus (red circle) contains 1–2000 oocyte-specific gDMRs of which 23 are definitive imprinted gDMRs. The sperm nucleus (blue circle) contains around 100–800 gDMRs of which three are definitive imprinted gDMRs. Below the oocyte and sperm are schematic maternal and paternal homologous chromosomes that carry CpG islands (CGIs) depicted as rectangles numbered 1 through to 11 on red (maternal homologue) and blue (paternal homologue) lines. Open rectangles represent non-methylated CGIs. Some methylated CGIs are shared (e.g., closed rectangle at position 3 on both parental homologues) and are not gDMRs. Some methylated CGIs are non-imprinted gDMRs (closed rectangles at position 6 on paternal chromosome and positions 2, 4, 5, 9 and 11 on the maternal chromosome) that will lose their methylation during the DNA demethylation that takes place as embryos pass through preimplantation development. A few methylated CGIs are imprinted gDMRs (closed rectangles at position 10 on the paternal chromosome and 1 and 8 on the maternal homologue) that will retain their methylation status through DNA demethylation. As explained, there are around 1–2000 non-imprinted and imprinted gDMRs present in the oocyte nucleus and around

100–800 non-imprinted and imprinted gDMRs will enter with the sperm. This difference in number in respective germ cells is reflected in the difference in closed rectangles on the maternal (red line) and paternal (blue line) homologues. **(B)** Consensus binding site of ZFP57. The TGCCGC hexamer motif is found in imprinted gDMRs and bound by ZFP57 when the central CpG dinucleotide is methylated. This motif shown (based on [124]) was downloaded from the Tomtom database (<http://meme-suite.org/tools/tomtom>) and has the searchable identifier ZFP57_MOUSE.H11MO.0.B. **(C)** Preservation of methylation at imprinted gDMRs. (1) The paternal (sperm nucleus in blue) and maternal (oocyte nucleus in red) nuclei contain homologous chromosomes that carry CpG islands (CGIs) depicted as described in **(A)**. (2) The maternal (in red) and paternal (in blue) pro-nuclei contain the homologous chromosomes (red and blue lines, respectively) described in (1). Of the ≈ 1600 non-imprinted and imprinted gDMRs in the zygote, only a small percentage—the imprinted gDMRs—will preserve DNA methylation in the face of the DNA demethylation that takes place during pre-implantation development [105]. The initial assembly of the heterochromatin-*like* complexes that preserve methylation at imprinted gDMRs takes place in the newly fertilized zygote (see text for details). It should be noted that H4K20HMTases have only been detected in the zygote by RT-PCR owing to lack of specific antibodies so it remains to be shown that the mRNAs are translated to give active proteins. (3) Preservation of methylation at imprinted gDMRs on the paternal (position 10) and maternal (positions 1 and 8) homologues is due to localised heterochromatin-*like* complexes at imprinted gDMRs. We have included H4K20HMTases as part of the heterochromatin-*like* complex that generates H4K20me₃ at the imprinted gDMRs because trace amounts have been detected by RT-PCR albeit translated protein has yet to be shown. The complexes preserve DNA methylation at imprinted gDMRs throughout pre-implantation development; non-imprinted gDMRs and methylated CGIs become de-methylated (stippled rectangles). (4) Global levels of DNA methylation reach their lowest point in embryonic nuclei of the blastocyst. However, methylation at imprinted gDMRs is preserved by the heterochromatin-*like* complexes shown in (3), on the paternal (position 10) and maternal (position 1 and 8) homologues. P denotes paternal homologue and M the maternal homologue. Taken and modified from [128]. **(D)** Assembly of localised heterochromatin-*like* complex at imprinted gDMRs. Methylation of cytosines in CpG dinucleotides (black circles) is preserved by the assembly of a heterochromatin-*like* complex at imprinted gDMRs. The complex is targeted by the KRAB zinc-finger protein ZFP57 that binds the hexamer motif TGCCGC when the cytosine in the CpG is methylated (black circle in green rectangle). This in turn recruits KAP1, which is a modular protein that acts as a focal point for the recruitment of Setdb1 histone methyltransferase, HP1 and Dnmt1. HP1 binds the H3K9me₃ generated by Setdb1 and recruits a H4K20me₃ histone methyl-transferase that generates H4K20me₃ thus forming the H3K9me₃:HP1:H4K20me₃ pathway. DNA methylation at the imprinted gDMR is maintained (dotted lines) by Dnmt1. Taken and modified from [128].

There are two ways by which heterochromatin-*like* complexes could preserve DNA methylation at imprinted gDMRs. The first is by protecting imprinted gDMRs from the activity of demethylating Tet dioxygenases that are present in the early embryo [104,105,129]. The second is mediated by the interaction of KAP1 with Dnmt1 [107,117] (Figure 3D), which is important because Dnmt1 is scarce during the DNA demethylation phase of pre-implantation development [118–120]. The KAP1–Dnmt1 interaction would have the effect of concentrating trace amounts of Dnmt1 in the vicinity of imprinted gDMRs thereby ensuring maintenance of 5mC (Figure 3C). Dnmt1 may also be recruited to imprinted gDMRs through a known interaction with HP1 proteins that has been shown to increase local DNA methylation levels [130].

The size of the heterochromatin-*like* complex at the imprinted gDMRs is around 6 kb [126]. The mechanism(s) by which the size of heterochromatin-*like* complexes is regulated is of interest because in other regions of the genome small, localised, heterochromatin-*like* complexes, such as that found at the imprinted gDMRs, act as nucleation sites for the assembly of much larger domains. This is the case for KRAB-ZFP-directed heterochromatin-*like* complexes (≈ 6 kb; [43,68]) that nucleate the assembly of KRAB-ZNF heterochromatin-*like* domains that can range up to 4 Mb in size [43,68].

After embryos implant into the maternal endometrium early post-implantation development is characterised by elongation of the primitive streak along the epiblast whereupon gastrulation begins

and cells of the embryo undergo finely orchestrated morphogenetic movements to form the three germ layers, ectoderm, mesoderm and endoderm [131]. Of note is that around this time H3K9me3-marked heterochromatin is transiently deployed in germ layer cells [132]. This transient deployment is likely involved in an evolutionary restriction observed during vertebrate development called the phylotypic period or progression, to which we now turn.

4. Heterochromatin-Like Domains/Complexes and the Phylotypic Progression during Vertebrate Development

In his popular work *Anthropogenie* Haeckel [133] published some of the most famous pictures in Biology: a series of comparative drawings showing different animals arising from near identical somite-stage embryos. For more than a century there has been controversy over what weight ought to be placed on the images [134], nevertheless, what they illustrated so graphically was there is a stage in development where an animal most closely resembles other species (Figure 4A,B). Ironically, this notion has become one of the central concepts in evolution and development because a similar embryonic stage can be identified in each phylum and has been termed the phylotypic stage [135]. In vertebrates, the identification of a precise phylotypic stage that is identical in all species has been elusive owing largely to the vagaries of heterochrony [136,137]. Rather, there is thought to be a phylotypic “period” [136] or “progression” [138] that roughly corresponds to organogenesis where numerous, undifferentiated, organ primordia are developing from the three germ layers (Figure 4A,B) [138–140]. Here, we use the term “phylotypic progression” because it most closely describes the character of the molecular mechanisms, to be described, that restrict the amount of evolution allowed during this developmental window.

The reduced inter-species variability of the phylotypic progression is flanked on either side by an earlier stage of ontogeny at which species differ markedly from one another, and a later stage that shows a progressive divergence among species (Figure 4A,B). The divergence of morphologies on either side of the phylotypic progression formed the basis of the hourglass model of vertebrate development [138–140]. Molecular explanations for the evolutionary “bottleneck” through which embryos pass are of two kinds. One is that there are signalling pathways among developmental modules in the mid-embryonic stages that are highly inter-dependent and make this period developmentally constrained, thus leading to evolutionary conservation [139]. The second relates to the mechanism(s) by which the anterior-posterior (A-P) axis (body plan; *Bauplan*) is laid down, specifically, the mechanisms that regulate the temporal and spatial collinearity of *Hox* cluster gene expression [138], where perturbations in the timing and/or extent of *Hox* gene expression are deleterious and this, again, leads to a restriction in the amount of evolution allowed. Support for this second explanation, as well as for the hourglass model, has come from cross-species transcriptome comparisons. Transcriptome profiling of mid-embryonic (around gastrula to organogenesis) stages of four vertebrate species (mouse, chicken, *Xenopus laevis* and zebrafish) showed conserved expression profiles during the above stages including conserved expression of *Hox* genes [141]. The relationship between the spatially restricted *Hox* gene expression patterns, the laying down of the *Bauplan* and the phylotypic progression reached its apotheosis in the remarkable observation that all three are intimately associated across different phyla. This led to the hypothesis that the association is a universal trait in animals and the defining characteristic of the zootype [142].

We suggest that the major cause for the phylotypic progression in vertebrates is the requirement for properly establishing epigenetic compartmental domains (ECDs; Box S1) so that by the end of the progression both position-specific and cell-type-specific cellular identities are safeguarded. It is a general mechanism that accommodates the role played by *Hox* gene expression in determining position-specific identities along the A-P axis [143]. ECDs contribute to ensuring that they remain so. ECDs are defined as contact enrichments seen in Hi-C maps that are generated by segregation of *micro*-phase separated HP1-containing heterochromatin-like and Pc-containing Pc-G domains/complexes (Box S1; [43]).

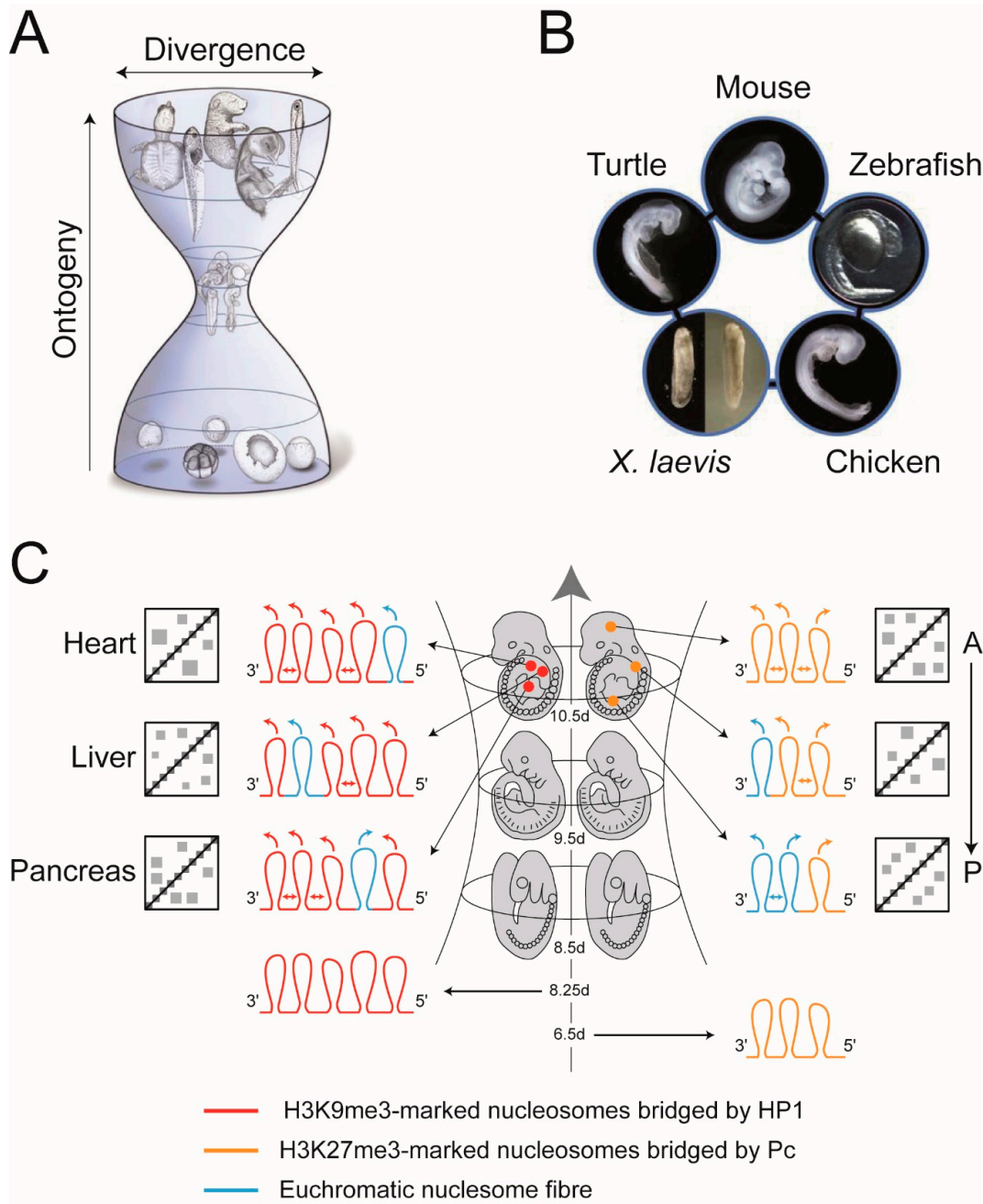


Figure 4. The phylotypic progression of vertebrate development and the generation of cell-type- and position-specific contact enrichments that contribute to epigenetic compartmental domains (ECDs). (A) The developmental hourglass model for vertebrate development. The model predicts that the mid-embryonic organogenesis stages (phylotypic progression) represent the developmental stages with highest morphological conservation across vertebrates. The phylotypic progression encompasses the developmental window when the anterior-posterior axis (body plan, *Bauplan*) is laid down. By the end of the phylotypic progression cell-type-specific patterns of gene expression have been initiated and position-specific (*Hox* code) patterns of gene expression are established and thereafter maintained for the rest of development. Taken and modified from [140]. (B) The phylotypic progression for vertebrates. This figure is taken from [140] where two stages of *Xenopus laevis* were shown because there was no statistically significant difference between these two stages. (C) Establishment of cell-type- and position-specific contact enrichments that contribute to ECDs. Depicted in the middle is

the “bottle-neck” of the hourglass depicted in (A), which shows the embryonic stages of the murine phylotypic progression where embryos show greatest similarity with other vertebrates. On the left of the progression is depicted the establishment of cell-type-specific contact enrichments. By day 8.25 of embryonic development H3K9me₃-marked heterochromatin is transiently deployed to compact and silence genes that regulate cell-type-specific differentiation. This is shown as five chromatin loops (red loops) that represent H3K9me₃-marked chromosome fibres that are “bridged” by HP1 proteins. As embryos pass through the phylotypic progression there is a progressive loss in H3K9me₃-marked heterochromatin. By the end of the phylotypic progression there is cell-type-specific loss of H3K9me₃-marked heterochromatin and differentiation-specific genes are expressed. Far *cis*- and *trans*-contact between micro-phase separated heterochromatin-like domains/complexes result in cell type-specific contact enrichments that emerge as cell-type-specific ECDs (double-headed arrows denote (far) *cis*-interactions; arrows on top of loops indicate *trans*-interactions; the cartoon Hi-C maps show, crudely, cell-type specific contact enrichments that become part of ECDs; A type compartments are not shown). The scenario is similar to the generation of position-specific contact enrichments that become part of ECDs that are shown on the right. On day 6.5 the *Hox* genes clusters are assembled as Pc-G domains (orange loops on the bottom right) that represent “closed” chromatin domains whose constituents are H3K27me₃-marked nucleosomes and the PRC1- and PRC2- complexes, where the Pc homologue in PRC1 “bridges” H3K27me₃-marked nucleosomes (Diagram A in Box S1, bottom row). As embryos progress through the phylotypic progression there is a progressive 3′ to 5′ activation of *Hox* genes along the *Hox* gene cluster (temporal collinearity). By the time embryos leave the phylotypic progression the spatially restricted patterns of *Hox* gene expression (spatial collinearity) have been established so that the *Hox* code for each region of the embryo are stable for the rest of development. In nuclei from the posterior trunk, much of the *Hox* gene cluster is in a euchromatic conformation that facilitates *Hox* gene expression (two blue “euchromatic” loops) with only a small region of the *Hox* cluster compacted into the remaining Pc-G domain (orange Pc-G loop). The *cis*- (denoted by double-headed arrows) and *trans*- (given by arrows on top of the loops) interactions that are mediated by this configuration of loops contribute position-specific contact enrichments to ECDs that are shown in the Hi-C maps (on the far right of Figure 4C). In nuclei from the mid-trunk region a smaller region of the *Hox* cluster is in a euchromatic conformation (one blue “euchromatic” loop) while a larger region remains compacted into a Pc-G domain (two orange Pc-G loops). This configuration of loops gives rise to position-specific contact enrichments that reflect the position of the cell along the A-P axis. These contact enrichments are shown in the middle Hi-C map cartoon to the right of the loops in Figure 4C. In the forebrain, where *Hox* genes are not expressed, the entire *HoxD* cluster is assembled into a Pc-G domain (three orange loops at top on right in Figure 4C). This configuration of loops give rise to contact enrichments that are specific for this anterior position along the A-P axis and are depicted in the cartoon Hi-C maps at the top right of Figure 4C. The position-specific contact enrichments are a small fraction of the contacts that contribute to ECDs. Not shown in the cartoon Hi-C maps are the myriad of additional contacts between ≈2000 Pc-G domains/complexes and between heterochromatin-like domains/complexes found elsewhere in the mouse genome.

A role for heterochromatin-like domains/complexes in regulating the phylotypic progression is indicated from two recent studies. The first, a study of the reprogramming of H3K9me₃-marked heterochromatin during early mouse development, showed that in recently implanted embryos (around day 6.5) there is a gradual increase in the association of H3K9me₃-marked heterochromatin with lineage-incompatible genes [144]. The second showed that levels continue to rise reaching their maximum in germ layer cells on day 8.25 where there is a net increase in association of H3K9me₃-marked heterochromatin with genes that regulate differentiation of adult cell types [132]. Repression of differentiation-specific genes is transient. Beyond day 8.25, as embryos traverse the phylotypic progression there is a progressive loss of H3K9me₃ and chromatin compaction at many sites in the genome and at those sites tissue-specific gene expression with concomitant differentiation begins [132]. By the time embryos exit the phylotypic progression (we give it as around day 10.5 in Figure 4C) there is cell-type-specific loss of H3K9me₃-marked heterochromatin, where previously

repressed genes take up euchromatic conformations that promote tissue specific gene expression (blue loops on the left in Figure 4C). Nuclei of differentiating cells still have H3K9me₃-marked heterochromatin domains/complexes elsewhere in their genomes and “bridging” within and between H3K9me₃-marked nucleosome fibres by HP1 contribute to the emergence of cell-type-specific contact enrichments in ECDs (red loops on the left in Figure 4C and associated cartoon Hi-C maps). ECDs safeguard cellular identity (Diagrams A–C in Box S1). Notably, RNAi “knock-down” screens for genes whose depletion destabilise cellular identity, identified genes that encode CAF-1, the SUMO-conjugating enzyme UBE2i, SUMO2, SETDB1, ATRX and DAXX proteins [145,146]. All are involved in either nucleation or replication of heterochromatin-like domains/complexes thus providing a link between safeguarding cellular identity and ECDs (Diagram C in Box S1; Figure 4C on left; [43]).

Pc-G domains/complexes also contribute contact enrichments to ECDs (Box S1; Figure 4C on the right) and the best described Pc-G domains are those that compact the *Hox* gene clusters. *Hox* genes are determinants of cellular fate and the positional identities of post-occipital tissues in the mouse (tissues below the skull, including the trunk and limbs) are determined by the collinear expression of *Hox* genes [147]. In cells where *Hox* genes are not expressed, such as ICM-derived ES cells and cells of the post-implantation epiblast, repressive Pc-G domains compact the ≈100 kb *Hox* gene clusters; constituents of Pc-G domains include the canonical H3K27me₃ histone modification and the PRC1 and PRC2 complexes [148,149]. In the forebrain, where *Hox* genes are not expressed, the same H3K27me₃-marked domain has been shown to encompass the *HoxD* and *HoxB* clusters [150]. Thus, in cells of the epiblast (around day 6.5 in the mouse), prior to entering the phylotypic progression, the *Hox* gene clusters are assembled into repressive Pc-G domains (depicted on the bottom right of Figure 4C as three orange Pc-G loops).

Hox gene expression is observed first on day 7.2 [151] and, once initiated, there is a gradual 3′ to 5′ activation along the *Hox* gene clusters with the 3′-most genes (group 1 genes) of the cluster being expressed first to be followed one after another by *Hox* genes that reside more 5′ finally ending with the 5′-located (group 13) genes (temporal collinearity; [152]). The 3′ to 5′ activation is associated with corresponding change in epigenetic modifications from a repressive H3K27me₃ modification to an activating H3K4me₃ modification [153]. The dynamical change in histone modification is associated with a progressive shift in the 3D compartmental organisation of the cluster. Accordingly, when a *Hox* cluster is transcriptionally inactive (enriched in H3K27me₃) it forms a single 3D compartment that can interact in *cis*- and *trans*- with distantly located loci that are also enriched in H3K27me₃ [150,154]. As the 3′ to 5′ transcription of a *Hox* cluster proceeds there is a switch in 3D organisation whereupon newly activated *Hox* genes beginning at the 3′ end of the cluster are progressively incorporated into a transcriptionally active compartment, while the rest remain in an inactive compartment [150]. By the time embryos exit the phylotypic progression temporal collinearity has established the spatially restricted patterns of *Hox* gene expression (spatial collinearity; day 10.5 for the somitic *Hox* code; Figure 4C on the top), which are stable for the rest of development [155]. In a nucleus taken from the posterior trunk of the day 10.5 embryo much of the *Hox* cluster has taken up an “open”, euchromatic conformation that is permissive for gene expression (two blue “euchromatic” loops at the bottom right of Figure 4C) and only a small part of the *Hox* cluster remain assembled into a compact, silent Pc-G domain (one orange loop at the bottom right of Figure 4C). More anteriorly, around the mid-point of the trunk, only the 3′-most genes of the *Hox* gene cluster are in an open conformation (one blue “euchromatic” and two orange (Pc-G) loops at the middle on right of Figure 4C). In the forebrain where *Hox* genes are not expressed the entire *HoxD* cluster is assembled into a Pc-G domain (three orange loops at top on right in Figure 4C). Segregation of micro-phase separated Pc-G domains associated with the *Hox* genes contributes position-specific contact enrichments to ECDs (Diagram C in Box S1; cartoons on right of Figure 4C). Position-specific contacts mediated by Pc-G domains/complexes associated with *Hox* genes are but one fraction of the contacts that contribute to ECDs for, as explained below, there are around 2000 Pc-G domains/complexes in mouse and human genomes [156] all of which are likely to contribute to ECDs along with heterochromatin-like domains/complexes.

Temporal collinearity has been posited as the *cause* of the phylotypic progression—an ineluctable *Einbahnstraße* [138] through which vertebrate embryos must pass in order for the spatially-restricted patterns to be established (spatial collinearity) thereby ensuring each position along the A-P has its own *Hox* code that is stably carried forward for the rest of development. We believe that the establishment of *Hox* gene spatial collinearity is taking place in the larger context of the deployment of heterochromatin-like domain/complexes during the phylotypic stage (Figure 4C on left). Together they contribute to the evolutionary restriction seen as the “bottleneck” in the hourglass model of vertebrate development (Figure 4A). Key questions remain. For example, what is the mechanism(s) by which the ground states of heterochromatin-like and Pc-G domains/complexes are achieved on days 8.25 and 6.5, respectively? A related question is: how is the *Einbahnstraße* [138] converted into a two-way-street during the animal cloning procedure? The A-P axis must be faithfully recapitulated in reconstructed embryos even when the transferred nucleus retains the memory of the *Hox* code specific for only one position along that axis.

The number of Pc-G domains/complexes is around 2000 in the genomes of mouse and human ES cells, where there is a preference of PRC1 and PRC2 complexes to localise to CGIs [156]. The domains/complexes contribute to the coarse-grained chromatin-state pattern that characterises mammalian genomes [157,158]. That is, they contribute to the “segmented” nature of mammalian genomes, where segments or “blocks” consisting of Pc-G and heterochromatin-like domains/complexes alternate with segments or “blocks” of euchromatin. As shown (Diagram C in Box S1), ECDs are generated by *cis*- and *trans*-contacts between both types of micro-phase separated domains/complexes. How the theory of micro-phase separation of block copolymers (BCPs) might explain the behaviour of the domains/complexes as “blocks” in a chromatin fibre is the subject of the next section where, in the interests of clarity, we concern ourselves with heterochromatin-like domains/complexes although the same arguments apply to Pc-G domains/complexes.

5. Heterochromatin-Like Domains/Complexes and Block Copolymers (BCPs)

Our survey of co-localisation of HP1s with H3K9me3 shows that depending on cell type there are (conservatively) around 163–855 heterochromatin-like domains (>0.1 Mb) and 18,853–32,292 complexes (<0.1 Mb) in the human genome (Table 1). The heterochromatin-like domains/complexes are contiguous with the nucleosome fibre and can be coarse-grained as alternating “blocks” of heterochromatin and euchromatin along the nucleosome fibre “polymer” (Box S1). This organisation is analogous to that of a BCP that contains a series of alternating blocks (e.g., A-type and B-type), each composed of multiple monomers (A monomers and B monomers). Where the monomers are incompatible the blocks segregate on the basis of like-with-like, with A-type blocks associating with A-type blocks and B-type associating with B-type. Accordingly, the BCP forms spatially segregated domains that are enriched in A or B [159]. We have drawn upon the basic polymer physics of micro-phase separation of bulk BCPs to explain how segregation of micro-phase separated heterochromatin-like domain/complexes is likely to generate contact enrichments that contribute to ECDs detected in Hi-C maps ([43]; Box S1). While we recognise that micro-phase separation seen with simple hydrocarbon BCPs (top of Box S2) is a useful analogy we argue in the next sections that the fundamental physics of micro-phase separation of heterochromatin-like domains/complexes is not the same as BCPs. In order to develop our argument, we begin with a simplified description of the physics underpinning micro-phase separation of BCPs.

BCPs can be configured into a wide variety of molecular architectures based on two, three or more monomer types [160]. Of these architectures, symmetric BCPs containing equal sized blocks A and B (AB di-BCPs) have been the focus of a very large number theoretical and experimental studies and our discussion of BCPs will deal almost exclusively with this simplest form of BCP. The physics of the phase behaviour of a bulk (undiluted) AB di-BCP centres upon the covalent bond that separates the two chemically dissimilar blocks and prevents macroscopic phase separation (for a review see [161]). The bond makes the entropy of mixing small and excess free energy generated by even minor chemical or structural differences between A and B blocks is sufficient to produce

contributions that are unfavourable to mixing. Put another way, phase separation of a bulk di-BCP is driven by an unfavourable mixing enthalpy coupled with small mixing entropy, with the covalent bond connecting the blocks preventing macroscopic phase separation. *Microscopic* phase separation of a di-BCP depends on three parameters [161]: (1) the volume fractions of the A and B blocks ($f_A + f_B = 1$), (2) the total degree of polymerisation ($N = N_A + N_B$), and (3) the Flory–Huggins parameter (χ_{AB}). The χ -parameter specifies the degree of incompatibility between the A and B blocks and this is what ultimately drives micro-phase separation. The relationship between χ_{AB} and temperature (T) is given in Equation (1), which is an application to BCPs of the original mean field lattice theory [162,163] for the thermodynamic behaviour of homo-polymers in a simple solvent (see Equation (S1) in Box S2):

$$\chi_{AB} = \frac{z}{k_B T} \left[\varepsilon_{AB} - \frac{1}{2} (\varepsilon_{AA} + \varepsilon_{BB}) \right] \quad (1)$$

Equation (1) describes the energy cost (in units of thermal energy $k_B T$) when A and B monomers make contact. z is the coordination number of an incompressible lattice (Box S2) and represents the number of nearest neighbours per lattice site that can be occupied by either an A-type monomer or a B-type monomer. ε_{AB} , ε_{AA} and ε_{BB} are the contact energies per repeat unit (monomer) of A-B, A-A and B-B, respectively. As with Equation (S1) in Box S2 the sign (positive or negative) and magnitude of χ_{AB} in Equation (1) is determined by the choice of monomer—the “unit of incompatibility” is the monomer. For a typical di-BCP consisting of two types of simple hydrocarbon monomer, such as polyisoprene-block-polystyrene (PI-b-PS), where electrostatic interactions are negligible (i.e., governed by dispersive interactions), the value for χ is positive and small (≈ 0.1). The *positive* value indicates that there is a net repulsion between the PI and PS blocks of the PI-b-PS and they have a tendency to micro-phase separate; a di-BCP that has a *negative* value of χ would indicate a free-energy drive towards mixing. In Equation (1) and Equation (S1) χ varies inversely with temperature. Increasing temperature or decreasing χ_{AB} through choice of monomers reduces the incompatibility between the constituent blocks and combinatorial entropy increases, resulting in mixing whereupon the di-BCP becomes disordered (i.e., homogeneous).

We have previously drawn on the observation that χ varies inversely with temperature and that micro-phase separation is dependent upon composition (volume fraction), to provide insight into how changes in the activity of chromatin-associated cohesin affects the compartmentalisation observed in Hi-C maps (Figure 5) [43]. As explained ([43]; Diagram B in Box S1; see also Sections 7 and 8 below), the incompatibility between the heterochromatin-like domains/complexes and euchromatin is owing to the “bridging” of the H3K9me3-marked nucleosomes and it is this bridging that drives micro-phase separation of the domains/complexes from euchromatin. Far *cis*- and *trans*-contacts between micro-phase separated heterochromatin-like complexes results in segregation and contributes to the emergence of the ECDs observed in Hi-C maps ([43]; Diagrams B and C Box S1). Mixing of heterochromatin-like complexes with euchromatin is caused by cohesin, which is a loop-extruding factor (LEF) [164,165]. LEFs attach to the chromatin fibre and reel it in from both sides, thereby extruding a progressively growing chromatin loop until the LEFs fall off, bump into each other, or bump into extrusion barriers such as CTCF, which define TAD boundaries [164–166]. Loop extrusion is an energy-driven, ATP-dependent, process [167]. Mixing is promoted by friction of the heterochromatin-like domain/complex with the nucleoplasm during loop extrusion, which converts the kinetic energy of loop extrusion into thermal energy (work done by ATP-hydrolysis is converted into heat). As a consequence, HP1-mediated “bridging” of H3K9me3-marked nucleosomes is disrupted making the domains/complexes less “heterochromatic” and more “euchromatic”, with smaller complexes undergoing greater mixing compared to the larger domains in keeping with the dependency of micro-phase separation on volume fraction.

The above provides a framework for understanding how mutants that affect cohesin activity can in turn affect ECDs observed in Hi-C maps (Figure 5). For example, when the cohesin subunit SCC1 is deleted [168] finer, better defined, ECDs emerge in Hi-C maps (Figure 5A, on the right; [169])

that are normally “masked” in wild-type (wt) cells (Figure 5A, on the left). The emergence of the finer compartmentalisation in SCC1 mutants can be understood in terms of the model in Figure 5B, that consists of two loops, loops 1 and 2. Loop 1 contains a large micro-phase separated heterochromatin-like domain. In loop 2 is a smaller heterochromatin-like complex. In wt cells, extrusion of the large heterochromatin-like domain in loop 1 by the cohesin complex (green rings on left in Figure 5B) has little effect on the mixing of the domain with euchromatin; the domain makes far *cis*- and *trans*-contacts (red arrows) that will be detected as contact enrichments in ECDs. Extrusion of the smaller complex in loop 2 by the cohesin complex results in extensive mixing with euchromatin leading to its dissolution in wt cells (depicted by the red dots on blue line). As a consequence, the smaller complex makes contacts (blue arrows) that will be detected as contact enrichments in A-type compartments. On the right in Figure 5B is an explanation of the emergence of the finer compartmentalisation in SCC1 deleted cells. Here energy-driven loop extrusion is absent and the HP1-mediated bridging of H3K9me3-marked nucleosomes in the small heterochromatin-like complex is reconstituted (red line in loop 2 on right of Figure 5B); the large domain that precedes it is unaffected. Reconstitution of the small complex promotes far *cis*- and *trans*-contacts (red arrows) that cause the finer, more defined, ECDs observed in SCC1 deleted cells (Figure 5A on right).

Mixing of larger heterochromatin-like domains with euchromatin by enhanced energy-driven loop extrusion could go some way to explaining the changes in compartmentalisation seen in Hi-C maps derived from WAPL/Pds5A/B compound mutant cells (Figure 5C on the right; [169]). By way of background, WAPL normally removes cohesion through binding of its YSR motifs to the regulatory subunit of cohesin, Pds5 [171,172]. The WAPL-pds5 interaction stabilises a transient, open, state of the cohesin ring that results from disruption of the interface between the SMC3 and Rad21/Sccl cohesin subunits, with consequent release of the sister chromatids [173,174]. In mutations or “knockdowns” of WAPL cohesin is retained along the length of the chromosome arms [175–177]. Retention of cohesin complexes results in continued loop extrusion, the formation of larger topologically associated domains (TADs) and the loss of compartmentalisation [170] (Figure 5C, on the right). As shown in Figure 5D (on the right), loss of compartmentalisation may, at least in part, be explained by enhanced loop extrusion that disrupts HP1-mediated bridging of H3K9me3-marked nucleosomes even within large heterochromatin-like domains (red dots on blue line) resulting in mixing of domains/complexes with euchromatin. Unrestrained loop extrusion eventually leads to an overall loss of interphase chromatin organisation where interphase chromatin takes on condensed mitotic-like chromatin state termed vermicelli [170]. This loss of interphase organisation will also affect compartmentalisation.

Recent work in fission yeast may provide additional support for loop extrusion as a mechanism for mixing of heterochromatin-like domain/complexes with euchromatin. Specifically, it was observed that the *Pds5* mutation in fission yeast alleviates heterochromatin-mediated silencing at the donor mating-type region [178]. The effect of the *Pds5* mutation can be reversed by introduction of the *eso1* mutation [178]. *eso1p* is an acetyltransferase required for establishing stable cohesin complexes on chromatin, which it does by acetylating the heads of the SMC3 cohesin subunit that protects against removal of cohesin by WAPL [179,180]. Genetic interaction of the two mutations can be explained by reversible mixing of the heterochromatin-like complex assembled at the silenced mating-type region with flanking euchromatin. Accordingly, loss of Pds5p activity enhances cohesin-driven loop extrusion leading to disruption of HP1-mediated bridging of H3K9me3-marked nucleosomes resulting in mixing with euchromatin (c.f. Figure 5C,D on right). Introduction of *eso1* returns cohesin activity to near normal levels by interrupting cohesin establishment and that, in turn, reverses mixing and reconstitutes the heterochromatin-like complex and silencing at the donor mating-type region.

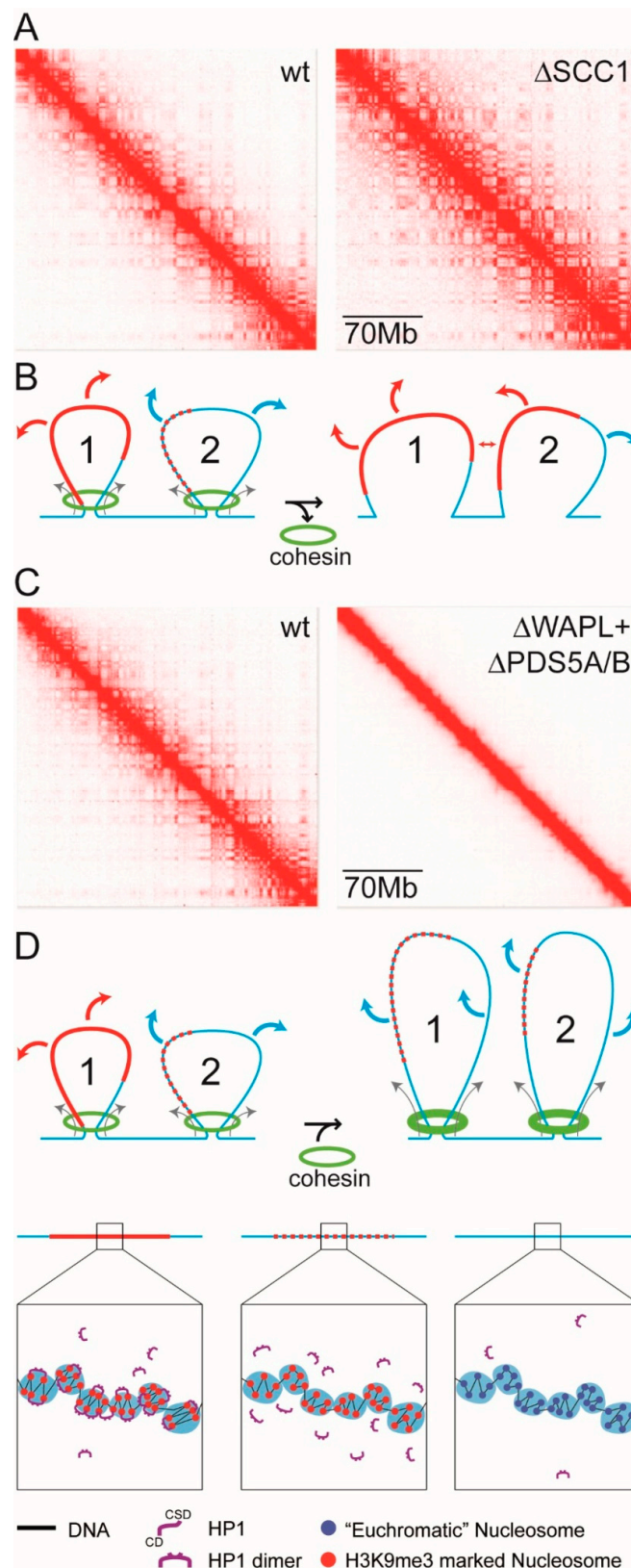


Figure 5. Mixing of heterochromatin-like domains/complexes with euchromatin by modulating chromatin-associated cohesin activity. (A) Hi-C maps for wt and SCC1 depleted cells at low resolution (70 Mb) to show the effect of cohesin depletion on compartmentalisation. The Hi-C map generated

from SCC1 deleted cells (on right) have a finer, better defined, compartmentalisation compared to the Hi-C map from wt cells (on left). Taken and modified from [169], with permission. (B) A BCP-based model that provides an explanation for the effect of SCC1 depletion on compartmentalisation. On the left are two loops that are being extruded by chromatin-associated cohesin (green rings at base of loops). Loop 1 contains a large micro-phase separated heterochromatin domain/complex (continuous red line) that is flanked by euchromatin (continuous blue line). The large domain/complex is resistant to mixing; bridging of H3K9me3-marked nucleosomes by HP1 is not disrupted (see key at bottom of figure). Loop 2 contains a smaller domain/complex (smaller “block”) that is subject to mixing (dotted line; see key at bottom of figure) by loop extrusion, which is a prediction of BCP theory, where a block of smaller volume fraction has greater tendency for mixing, with the BCP becoming “homogeneous” (one phase). Loop 1 makes far *cis*- (not shown) and *trans*-contacts (red arrows) with other HP1-containing domains/complexes which result in contact enrichments that emerge in ECDs. Loop 2 makes euchromatic contacts (blue arrows) that will fall into A-type compartments as a consequence of mixing because the smaller domain/complex is now more euchromatic and less heterochromatic. On the right are the same two loops after chromatin-associated cohesin is eliminated by loss of SCC1, whereupon loop extrusion ceases and mixing is reduced. This has little effect on the larger domain/complex in loop 1, but leads to the reconstitution of the smaller domain/complex in loop 2 (now a continuous red line) as a consequence of bridging H3K9me3-marked nucleosomes by HP1. The newly reconstituted domain/complex can then make far *cis*- (double-headed arrow) and *trans*-contacts (red arrows) that result in the finer, more defined, compartmentalisation seen Hi-C maps from SCC1 deleted cells. (C) Hi-C maps for wt and WAPL/PDS5A/B depleted cells at low resolution (70 Mb) to show effect of enhancing cohesin activity on compartmentalisation. The Hi-C map generated from WAPL/PDS5A/B depleted cells (on right) where cohesin activity is enhanced there is little or no compartmentalisation compared to the Hi-C map from wt cells (on left). Taken and modified from [169], with permission. (D) A BCP-based model that provides an explanation for the effect of WAPL/PDS5A/B depletion on compartmentalisation. The same two loops, loops 1 and 2, as seen on the left in (B), where loop extrusion by chromatin-associated cohesin leaves the larger heterochromatin-like domain/complex in loop1 intact, but the smaller domain/complex in loop 2 undergoes mixing. As before, in wt cells, loop 1 makes far *cis*- (not shown) and *trans*-contacts (red arrows) with other domains/complexes which result in contact enrichments that emerge in ECDs. Loop 2 makes euchromatic contacts (blue arrows) that will fall into A-type compartments. In WAPL/PDS5A/B depleted cells the activity of cohesin is *enhanced* (thicker green circles at base of loops). The effect on compartmentalisation is striking. Compartmentalisation is almost completely eliminated. One cause of this loss is that increased loop extrusion results in dissolution of the larger (loop 1) and smaller (loop 2) domains/complexes (dotted lines; see key at bottom of the figure) and mixing with euchromatin. If this occurs genome-wide there would be a loss of compartmentalisation. Moreover, continued loop extrusion leads to collapse of interphase chromatin organisation, whereupon chromatin takes on condensed mitotic-like chromatin state termed vermicelli [170]; compartmentalisation would also be affected by this loss of organisation. The “clutch” diagrams were modified from [43].

Micro-phase separation of a BCP is dependent upon the sign and magnitude of χ i.e., the degree of incompatibility of the monomer (“unit of incompatibility”; Box S2; Equation (1)) repeating units that make up the blocks. That BCP theory could provide an explanation for how cohesin activity might regulate mixing of heterochromatin-like domains/complexes with euchromatin (Figure 5; [43]) indicates that the same is true for heterochromatin-like and euchromatic “blocks” in the chromatin fibre. Put another way, there exists a “unit of incompatibility” that is the repeating unit from which heterochromatin-like and euchromatic “blocks” are assembled and it is the degree of incompatibility between these heterochromatin-like and euchromatic repeating units, specified by χ , that drives micro-phase separation of heterochromatin-like domains/complexes from euchromatin. Notably, virtually all modern theories of micro-phase separation employ this simple one-parameter thermodynamic description of the driving force for micro-phase separation [181,182]. Fortified by these

credentials, we move to exploring whether the approach taken for the determination of χ for di-BCPs can also be used to estimate χ for heterochromatin-like domains/complexes flanked by euchromatin.

6. Di-BCPs, χ and The Monomer as the “Unit of Incompatibility”

For di-BCPs, χ can be described by (1) where interactions between the hydrocarbon *monomer* repeating units (see top of Box S2) are dominated by dispersive (van der Waals) interactions. In a classical thermodynamic approach for determination of χ based on these dispersive interactions, Bates [183] showed that the contact energy (ε_{ij}) between i and j segments, each consisting of (dissimilar) monomer repeating units, could be quantified by the equation:

$$\varepsilon_{ij} = - \sum_{i,j} \frac{3}{4} \frac{I_i I_j}{I_i + I_j} \frac{\alpha_i \alpha_j}{r_{ij}^6} \quad (2)$$

r_{ij} is the segment-segment separation, α and I are the segment polarisability and the first ionisation potential, respectively. If there is neither a change in volume (Equation (1) assumes an incompressible lattice; see also Box S2) nor a preference for a particular segment orientation upon mixing, passing the binary interaction energies (for A-A, B-B and A-B) quantified in Equation (2) through Equation (1) gives the following result for χ :

$$\chi = \frac{3}{8} \frac{I}{k_B T} \frac{z}{V^2} (\alpha_A - \alpha_B)^2 \quad (3)$$

[Equation (3), as given in [183], contained a numerical error where the coefficient should be 3/8 (as it is here). The corrected equation was graciously provided by Professor FS Bates.]

Equation (3) is an approximation for χ owing to a number of assumptions (see [183]). A cubic lattice is assumed where $I_j = I_i \equiv I$; the latter is valid to within 10% for most hydrocarbons in a di-BCP. The volume of the cubic lattice site is also assumed as V , where $V = r^3$ (the r^6 term is represented as V^2 in (3)). Further, defining the number of segments that interact with each other by the co-ordination number z places all segments surrounding a particular segment within a single parameter, with an average interaction given by I and $(\alpha_A - \alpha_B)^2$. A more rigorous treatment that removes the assumption of an incompressible (cubic) lattice [162,163] would involve a calculation based upon summation over all neighbouring segments with the r^6 dependence (FS Bates, personal communication). Despite being an approximation, Equation (3) provides qualitative predictions, for example, for a bulk di-BCP governed solely by dispersive interactions a value of $\chi \geq 0$ is obtained indicating that the di-BCP will have a tendency to phase separate. In this (classical) treatment the magnitude of χ is fundamentally dependent upon the choice of repeating unit, the monomer. The monomer is the *unit of incompatibility* from which the “blocks” or “segments” of a di-BCP are made.

As might be expected, the assumptions required to derive Equation (3) are rarely satisfied in practice. Because of this, χ is usually determined empirically using a much simpler general formula:

$$\chi = \alpha T^{-1} + \beta \quad (4)$$

χ retains the temperature dependency. α and β are experimentally determined coefficients for enthalpy and excess entropy for a particular composition of a BCP. We believe that the utility of Equation (4) can be extended to enable an estimation of χ for a heterochromatin-like domain/complex vs. euchromatin once the *unit of incompatibility* has been defined. This is the subject of our next section.

7. The Oligo-Nucleosomal “Clutch” As the “Unit of Incompatibility” of Chromatin

Estimation of χ for heterochromatin-like domains/complexes flanked by euchromatin is likely to employ a different physics to that which developed for determination of χ for di-BCPs (Equations (1)–(3)). An obvious reason for this is that heterochromatin-like domains/complexes are assembled from the 11 nm “beads-on-a-string” nucleosome fibre “polymer” that is the primary structure of chromatin [1].

The “bead”, or *monomer* repeating unit, is the nucleosome, a multicomponent structure very different to the simple hydrocarbon *monomer* found in BCPs (see top of Box S2). By comparison the nucleosome is gigantic. The molecular weight of the nucleosome core particle, consisting of 147 bp of DNA that is wrapped 1.75 times in left-handed super-helical turns around a histone octamer, is over three orders of magnitude greater (≈ 200 k Da; [184]) than the hydrocarbon monomers of BCPs (≈ 100 Da; top of Box S2). In addition, monomers in BCPs are hydrophobic and dispersive interactions dominate (Equations (2) and (3)), while the nucleosome is a highly electrostatic multicomponent structure. A single histone octamer in the nucleosome has ≈ 220 positively charged lysine and arginine residues and ≈ 74 negatively charged aspartic acid and glutamic acid residues. The phosphate backbone of 200 bp of DNA that includes the 147 bp associated with the nucleosome core particle plus linker DNA adds a further 400 negative charges [185]. Given the obvious differences in physiochemical properties it is a wonder that BCP theory could be used as an analogy to understand how ECDs might emerge by micro-phase separation of heterochromatin-like complexes (Figure 5; [43]). This most likely has deep roots in their shared polymeric nature [186,187] but does not mean that the physics of micro-phase separation and segregation of heterochromatin-like domains/complexes and BCPs is the same. Nor may we expect that the degree of incompatibility, as designated by χ , is caused by the same physiochemical mechanisms.

For di-BCPs the sign and magnitude of χ is determined by the degree of incompatibility of the two dissimilar monomers; the monomer is the “unit of incompatibility” from which the micro-phase separated “blocks” or “segments” are composed (Box S2; Equations (1)–(3)). For heterochromatin-like domains/complexes the sign and magnitude of χ will be determined by the degree of incompatibility of the domains/complexes with euchromatin. We suggest that incompatibility between domains/complexes and euchromatin is the result of excess free energy generated by binding of HP1 to H3K9me3-marked nucleosomes in the 11 nm nucleosome fibre combined with excess entropy that results from compaction of the nucleosome fibre by the “bridging” effect [29,188,189]. This leads to an unfavourable enthalpy of mixing of the heterochromatin-like domain/complex with euchromatin. This begs the question: what is the “unit of incompatibility” for a heterochromatin-like domain/complex? On the face of it, it would seem that the unit of incompatibility is the H3K9me3-marked nucleosome *monomer* in the same way that the hydrocarbon *monomer* is the unit of incompatibility for di-BCPs. However, experimental and theoretical work indicates that (i) HP1 proteins drive the incompatibility rather than the H3K9me3-marked nucleosomes per se and (ii) the unit of incompatibility is a larger than the mono-nucleosome. For one, recent liquid Hi-C experiments directed towards identifying factors that generate and maintain compartmental domains indicated that ECDs (B-type compartmental domains) enriched in HP1 α and β were found to be equally stable after chromatin fragmentation followed by HP1 γ with ECDs enriched in the Polycomb CBX8 homologue being the most unstable [190]. K9me3 (to which HP1 α and HP1 β bind) is essentially indistinguishable compared to K27me3 (to which Polycomb CBX homologues bind) i.e., K9me3 and K27me3 are unlikely to contribute to differences in chemical potential between nucleosomes possessing these modifications, especially given the highly electrostatic environment of the nucleosome (see discussion above). Yet there are great differences in the dissociation kinetics of ECDs enriched in HP1 α and β compared to CBX8 indicating that it is the “bridging” protein that drives micro-phase separation and generation of ECDs rather than the histone modifications per se. Second, theoretical work using simple models where binding proteins (“binders”) cross-link polymer-specific binding sites have been able to recapitulate many features of Hi-C maps where binders cause folding of the polymer and phase transitions [191,192]. The binding sites on the polymer are neutral with respect to phase transitions and serve simply as binding sites for the “binders” that are the actual drivers of the transitions. Polymer simulations using a similar model have shown that it is HP1 binding to H3K9me3-marked nucleosomes that drives phase separation with the minimum H3K9me3-marked segment that can be phase separated by HP1 “bridging” being around 20 kb [193]. Finally, compaction of heterochromatin-like domains/complexes requires at least two H3K9me3-marked nucleosomes to be “bridged” by HP1 ([189]; Box S1, top row). This indicates

that the “unit of incompatibility” is more than a single H3K9me3-marked nucleosome, i.e., not a mono-nucleosome. It is likely to be even larger than two nucleosomes because *in vitro* studies show HP1-mediated bridging can compact nucleosome arrays into clusters of nucleosomes [29,188]. Similar observations have been made with Polycomb CBX proteins where, for example, the CBX2 protein, as part of the PRC1 complex, has intrinsic chromatin compaction activity where it is able to compact at least four nucleosomes ([194,195]; Diagram A in Box S1, bottom row). The unit of incompatibility is most likely a cluster or “clutch” of nucleosomes.

There is a weight of experimental evidence showing that “clutches” of 2–10 nucleosomes with variable degrees of zig-zag geometry are a ubiquitous motif within interphase chromatin; this organisation may represent the secondary structure of interphase chromatin [196]. Clear evidence came from super resolution microscopy, which showed that chromatin outside constitutive heterochromatin was characterised by the assembly of irregularly folded ‘clutches’ containing 2–10 nucleosomes while the density of larger ‘clutches’ was greater within constitutive heterochromatin [197]. A variety of electron microscopy approaches have confirmed the presence of small clumps of 2–10 nucleosomes *in vivo* without any evidence for longer stretches of an organised nucleosome fibre [198–201]. Nucleosomes in the “clutches” possess, to a lesser or greater extent, a local zig-zag organisation. Radiation-induced spatially correlated cleavage of DNA with sequencing revealed zig-zag geometry of short stretches of the nucleosome that was noticeably enriched in H3K9me3-marked nucleosome fibres in constitutive heterochromatin [202]. *In vivo* studies using controlled DNA breakages [203] or cross-linking of nucleosomes to one another followed by digestion and electron microscopy [204] have also given results consistent with there being short stretches of nucleosomes (3–10 nucleosomes) with zig-zag geometry in the nucleus. A substantial number of *in vitro* studies using synthetic nucleosomal templates have been confirmatory in nature and revealed a zig-zag motif for short stretches of (4–12) nucleosomes [205–207]. Notably, it has been shown in a wide variety of eukaryotes that nucleosomes are connected by linkers biased towards non-integer DNA double-helical turns (e.g., 0.5, 1.5, 2.5 turns) [208]. Such fibres possess zig-zag nucleosomal geometry [209] and exhibit enhanced phase separation [210].

Based on the evidence we suggest that the “unit of incompatibility” for chromatin is a dynamic multi-component structure unlike the indivisible hydrocarbon monomer (top of Box S2) that is the unit of incompatibility for BCPs. Specifically, for euchromatin the unit of incompatibility consists of an oligo-nucleosomal “clutch” of 2–10 nucleosomes where the nucleosomes within the “clutch” are disorganised with weak zig-zag geometry (Figure 6B,C). For heterochromatin-like domains/complexes the unit of incompatibility is an oligo-nucleosomal “clutch” of 2–10 H3K9me3-marked nucleosomes “bridged” by HP1 dimers that compact and stabilise the zig-zag geometry of nucleosomes within the “clutch” (Figure 6B,C). It is the incompatibility between the dissimilar repeating units (euchromatic “clutch” vs. heterochromatin-like “clutch”) that determines the magnitude of χ . This value will indicate whether there is a tendency for a heterochromatin-like domain/complex to micro-phase separate from the surrounding euchromatin. We now describe a theoretical model for the determination of χ for a heterochromatin-like “clutch” (HC) vs. euchromatic “clutch” based on Equation (4).

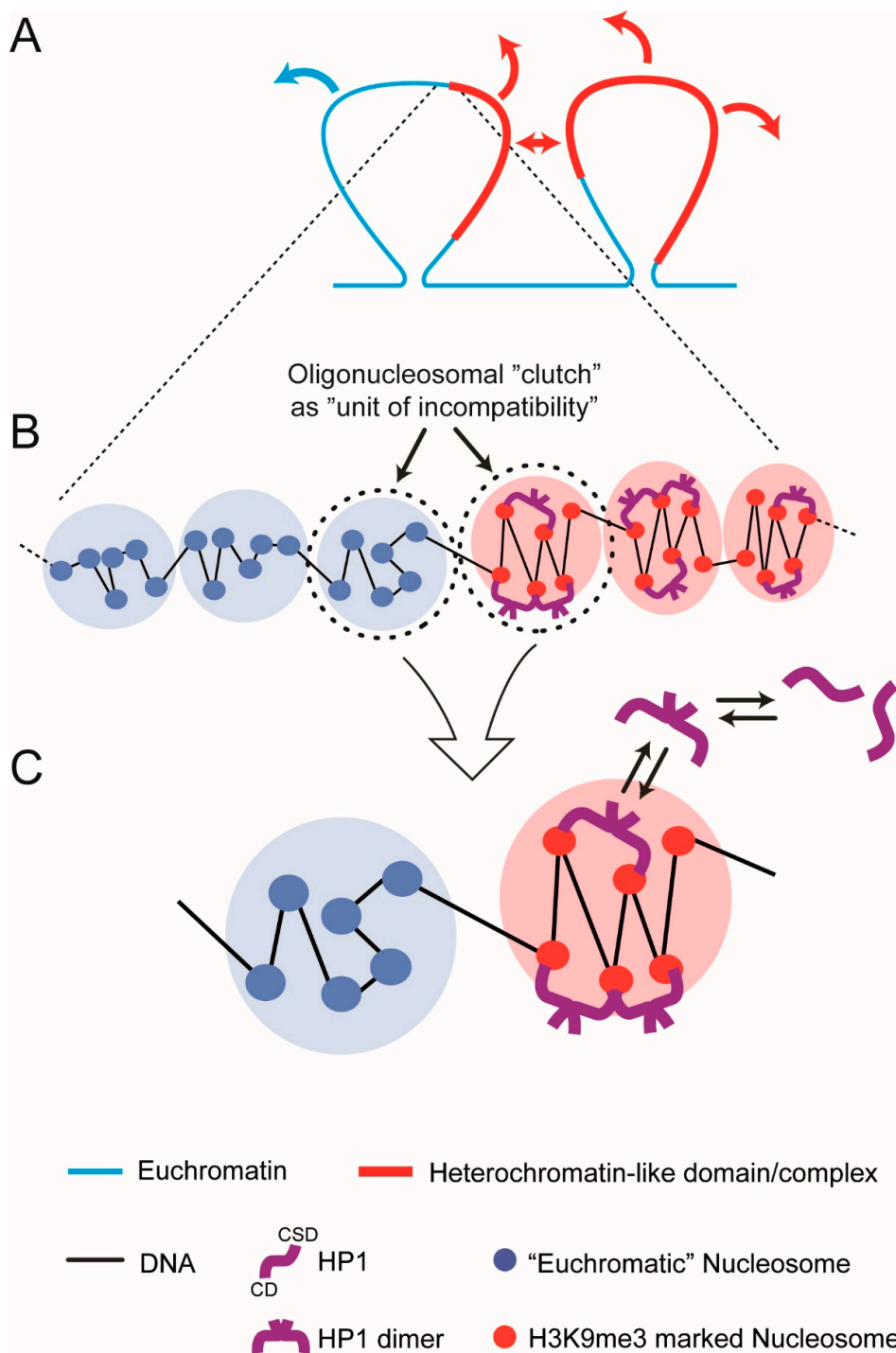


Figure 6. The oligo-nucleosomal “clutch” as the “unit of incompatibility” of chromatin. **(A)** Chromatin loops containing two heterochromatin-like domains/complexes (red line) flanked by euchromatin (blue line). Heterochromatin-like domains/complexes make *cis*- (doubled-headed arrow) and *trans*- (red arrows) contacts other micro-phase separated heterochromatin-like domains/complexes (Diagram B in Box S1). The *cis*- and *trans*-contacts between micro-phase separated heterochromatin-like domains/complexes generate contact enrichments seen in ECDs (Diagram C in Box S1). The euchromatic segments can also make *cis*- (not shown) and *trans*- (blue arrow) contacts with other euchromatic segments that are detected as A-type compartments in Hi-C maps. **(B)** A region of the loop in **(A)**

is magnified (dotted lines) to detail the junction between a heterochromatin-like domain/complex and a euchromatic segment. The heterochromatin-like domain/complex is made up of “clutches” (pink circles), where each clutch contains six H3K9me3-marked nucleosomes. The H3K9me3-marked nucleosomes are “bridged” by HP1 dimers where the nucleosomes have distinct zig-zag geometry; a clutch so organised is the “unit of incompatibility” of the heterochromatin-like complex. The “unit of incompatibility” of the euchromatic segment are “clutches” (blue circles) composed of six nucleosomes where the nucleosomes are more disorganised with only weak zig-zag geometry. (C) Depicts the theoretical model (see Section 8 and Equation (5)). Absent HP1 the HC containing H3K9me3-marked nucleosomes is thermodynamically equivalent to the euchromatic clutch. Incompatibility of the HC vs. euchromatic “clutch” is caused by HP1 bridging of H3K9me3-marked nucleosomes. Monomers of HP1 dimerise and the dimers “bridge” H3K9me3 marked nucleosomes, which has the effect of stabilising the zig-zag geometry and compacting the nucleosomes within the HC. “Clutch” diagrams were modified from [43].

8. Theoretical “Clutch” Model

As shown (Figure 6C) the model consists of the following:

1. A symmetric pair of “clutches”, a euchromatic “clutch” and heterochromatin-like “clutch” (HC) that are connected by linker DNA. Each clutch contains six nucleosomes, a value that lies in the middle of the range 2–10. The HC consisting of H3K9me3-marked nucleosomes represents the repeating unit (“unit of incompatibility”) of a heterochromatin-like domain/complex; likewise, for the euchromatic clutch and euchromatin. Both clutches are equally miscible in the nucleoplasm (a concentrated solution of chromatin fibres) reducing the problem to incompatibility between the clutches.
2. In the absence of HP1, the “clutches” are thermodynamically equivalent. Incompatibility between the HC and euchromatic “clutch” results from “bridging” of two K9me3 motifs on histone H3 proteins in separate nucleosomes by HP1 dimers; stacked nucleosomes are the preferred HP1 binding site. Bridging is maintained by constant exchange of bound HP1 dimers with free dimers in the nucleoplasm.
3. HP1 “bridging” stabilises the zig-zag geometry of H3K9me3-marked nucleosomes and leads to compaction. Nucleosomes in the euchromatic “clutch” are disorganised with limited (unstable) zig-zag geometry.

Using Equation (4), χ for HC vs. euchromatic “clutch” can be written as:

$$\chi_{\text{HC}} = [(\mathbf{H}_{\text{CD-H3K9me}})_{\text{HC}} + (\mathbf{H}_{\text{CSD-CSD}})_{\text{HC}} - \mathbf{H}_{\text{COMP}} - \mathbf{H}_{\text{TL}}]T^{-1} + \mathbf{S}_{\text{COMP}} \quad (5)$$

$\mathbf{H}_{\text{CD-H3K9me}}$ is the free energy contribution from the binding of the HP1 CD to H3K9me3-marked histone H3. CD-H3K9me3 binding is best understood for HP1 β [29,35] (Figure 1). Initial binding is via a non-specific electrostatic interaction of the N-terminal extension of HP1 β with the H3K9me3 tail [29]. This causes the CD N-terminal region to draw upwards and wrap around the peptide and, as a consequence, an aromatic “cage” is formed from three conserved aromatic residues: Tyr21, Trp42 and Phe45 (Figure 1B). The majority free energy contribution comes from the electrostatic cation– π interactions where the positively charged (cation) methyl ammonium moiety is attracted to the negative electrostatic potential of the aromatic groups’ π -system [36]. The HC sub-script indicates that it is the sum of CD-H3K9me3 interactions that “bridge” different nucleosomes within the clutch that gives the total free energy contribution.

$\mathbf{H}_{\text{CSD-CSD}}$ is the free energy contribution from the dimerisation of the HP1 CSD. It is the dimeric form of HP1 that “bridges” H3K9me3-marked nucleosomes and drives micro-phase separation. The HP1 β CSD forms a tight homodimer and the CSD-CSD interaction is of high affinity with sedimentation analysis indicating that the upper limit for the dissociation constant as <150 nM [211].

The HC sub-script indicates that it is the sum of CSD-CSD dimers that mediate “bridging” interactions between different nucleosomes within the clutch that gives the total free energy contribution.

H_{COMP} is potential energy owing to (i) elasticity of linker DNA (i.e., resistance to bending, twisting and stretching deformation) and (ii) steric exclusion between the nucleosomes and linker DNA [212]. H_{COMP} is a consequence of the “bridging” of H3K9me3-marked nucleosomes by HP1 dimers, which stabilises the zig-zag geometry and promotes compaction of nucleosomes in the “clutch”.

H_{TL} is elastic potential energy of the terminal linker DNA (i.e., resistance to bending, twisting and stretching deformation) that connects a heterochromatin-like clutch to a euchromatic clutch. For a heterochromatin-like domain/complex flanked by euchromatic segments the term would be $2H_{TL}$.

S_{COMP} is the excess entropy given up to the nucleoplasm after “bridging” of H3K9me3-marked nucleosomes by HP1 dimers, which stabilises the zig-zag geometry and promotes compaction.

Characteristics and Caveats of the Theoretical “Clutch” Model

1. In the form presented the model may not be applicable to *macro*-phase separation of cytologically-visible constitutive heterochromatin. There are additional factors that may need to be included that are together severally necessary and jointly sufficient in causing macro-phase separation of constitutive heterochromatin. They include repetitive DNAs, ncRNAs, proteins that bind (modified) DNA/histones [213]. Notably, there is a documented interaction of Swi6^{HP1} CSD dimer with the H2B α 1 helix that results in the deformation of the nucleosome core that is thought to drive phase separation of constitutive heterochromatin in fission yeast [55]. HP1 also binds to methylated lysine 26 of histone H1.4 (H1.4K26me) in constitutive heterochromatin [51], which may be significant in the context of the observation that H1-containing chromatin preferentially phase separates [210]. These factors are not included in our model since we are concerned with *micro*-phase separation of heterochromatin-like domains/complexes *outside* constitutive heterochromatin.
2. H3K9me3 epigenetic modification acts as an HP1 binding site only. Absent HP1 an H3K9me3-marked “clutch” is thermodynamically equivalent to the euchromatic “clutch”. Micro-phase separation results only after HP1 “bridging” between H3K9me3-marked nucleosomes. Experimentally, a reference “clutch” consisting of six H3K9me3-marked nucleosomes could be used to measure changes in free energy and excess entropy generated after addition of HP1 to the system. The H3K9me3-marked clutch would need to be tethered to mimic the connection through linker DNA to flanking euchromatic “clutches”.
3. HP1 dimer binding is modelled as “bridging” two H3K9-methylated histone H3 molecules in separate nucleosomes [37]. Accordingly, a nucleosome chain with stacked nucleosomes would be the preferred HP1 binding sites [214], which would be consistent with HP1 binding to a fibre with nucleosomes having a zig-zag geometry. Compaction would be promoted by allosteric cooperativity arising from changes in di-nucleosome conformation after HP1 bridging that would in turn enhance further HP1 bridging [214]. No direct HP1-HP1 co-operative binding is assumed, although there is evidence that such cooperative binding might contribute to “spreading” of the heterochromatic state [215,216]. This has been modelled previously [212] and could be incorporated as a free energy contribution ($H_{HP1-HP1}^{HC}$) to Equation (5).
4. Histone acetylation is not included in the model. Histone acetylation is an epigenetic modification generally associated with euchromatin and is known to neutralise the positive charge on the histone tails [217]. Inclusion of histone acetylation could affect the incompatibility between the euchromatic and heterochromatin-like clutches. Our reasoning for not including histone acetylation is that H3K9me3 and H3K27me3 are truly epigenetic and define ECDs (Box S1). They are associated with “write-and-read” activities that ensure their inheritance from one cell generation to the next [218]. Histone acetylation does not possess an analogous write-and-read activity [218].

5. The H_{TL} term reduces the magnitude of χ and acts against phase separation. Accordingly, release of a domain/complex from the constraints of the chromatin fibre by severing linker DNA that connects the domain/complex to flanking euchromatin would remove the H_{TL} term and enhance phase separation. Liquid Hi-C experiments have provided evidence that this is the case: chromatin fragmentation that releases domains/complexes from the chromatin fibre results in stronger compartmental segregation [190].

Using (5), χ can be calculated for *any* pair of clutches e.g., a HP1 $\alpha/\beta/\gamma$ -containing clutch (H3K9me3-marked nucleosomes) vs. a euchromatic clutch; a Pc-containing clutch (H3K27me3-marked nucleosomes) vs. a euchromatic clutch; a HP1 β -containing clutch vs. a HP1 γ -containing clutch, and any other pair-wise combinations. This is analogous to the approach for BCPs, where it is the degree of incompatibility between monomers (the “*unit of incompatibility*”) that determines the magnitude and sign of χ (Box S2; Equations (1)–(3)). Here, χ_{HC} specifies the degree of incompatibility between a heterochromatin-like clutch and a euchromatic clutch. The magnitude of χ will indicate whether a heterochromatin-like domain/complex consisting of repeating units of such clutches has a tendency to micro-phase separate from flanking euchromatin consisting of euchromatic clutches. The degree of phase separation of the domain/complex from the flanking euchromatin is specified by the segregation product χN , where N is the number of repeating units, that is: (i) the number of heterochromatin-like clutches in a heterochromatin-like domain/complex and (ii) the number of Pc-containing clutches in Pc-G domain/complex. The magnitude of χN will determine the character of the micro-phase separation actually observed (i.e., wavy, “liquid-like” or discrete, sharp interfaces) and this is our subject in the next section.

9. χN and the Order–Disorder Transition in Relation to Heterochromatin-Like and Pc-G Domains/Complexes

The sign and magnitude of χ specifies the *degree of incompatibility* between dissimilar monomers that make up the blocks in a symmetrical di-BCP and indicates whether the di-BCP has a tendency to micro-phase separate. The *degree of phase separation* of a symmetric di-BCP is specified by the segregation product χN , where N is the number of repeat units (monomers) that make up the polymer chain [183]. The seminal study that calculated the value of χN at which a bulk di-BCP would phase separate is that of Leibler [219] using self-consistent field theory (SCFT; [220]), where it was predicted that phase separation takes place when χN is around 10.5. This was confirmed experimentally [159]. Given that phase separation of a di-BCP takes place at $\chi N \approx 10.5$, when $\chi N \approx 10$ there is a delicate balance between energetic and entropic effects on segregation of the blocks (Box S3). When χN is increased there is a first-order phase transition (analogous to the phase transition from water to ice) that is called the order–disorder transition (ODT; $\chi N_{ODT} \approx 10.5$) where a disordered phase that is entropically favoured yet energetically costly is replaced by a periodic lamellae mesophase ($\chi N \geq 10$; Box S3). The character of the mesophase is dependent upon the magnitude of χN . When χN is close to or slightly greater than χN_{ODT} , i.e., $\chi N \geq 10$, the interfaces between the blocks are weak and “wavy” having the appearance of liquid–liquid phase separation (see ϕ_A vs. r_{\perp} for $\chi N \geq 10$ in Box S3). When χN is much greater than the χN_{ODT} , i.e., $\chi N \gg 10$, the interfaces are discrete and sharp (see ϕ_A vs. r_{\perp} for $\chi N \gg 10$ in Box S3). The evolution of micro-phase separation as the magnitude of χN is varied (Box S3) highlights how the interface between micro-phase separated “blocks” can change from wavy, “liquid-like”, to sharp and discrete depending on the magnitude of χN . This is relevant to the character (degree) of the micro-phase separations observed for heterochromatin-like compared Pc-G domains/complexes in the interphase nucleus.

The original SCFT approach to micro-phase separation of bulk BCPs treats polymers as flexible chains that coarse-grain as Gaussian random walks at sufficiently large (infinite) length scales [159,219]. However, the assumption of random-walk conformations of chains does not accurately describe the behaviour of domains/complexes in the interphase nucleus, where the majority of complexes in mouse and human are small, in the range of 10–30 kb (Table S1) albeit domains can be over 1 Mb

in size (Table 1). A more accurate description is that HP1-containing heterochromatin fibres behave like semi-flexible worm-like chains [193,221]. Semi-flexibility would mean that domains/complexes have their own values for χN_{ODT} distinct from ≈ 10.5 for bulk di-BCPs that behave as flexible chains governed by Gaussian statistics [159,222]. Accordingly, for the discussion below, the values for the order–disorder transition of heterochromatin-like and Pc-G domains/complexes are termed $\chi N_{\text{ODT_HC}}$ and $\chi N_{\text{ODT_PC}}$, respectively.

A bulk symmetric di-BCP is a simple one component system where segregation can be described analytically ([159,219]; Box S3), whereas segregation of heterochromatin-like domains/complexes involves a multicomponent system that takes place in the context of a concentrated solution of chromatin fibres where other domains/complexes roam, to a lesser or greater extent, the nuclear space as part of those fibres. Segregation of domains/complexes in interphase nuclei will, therefore, involve a different physics to that which has been described for bulk symmetrical di-BCPs ([159,219]; Box S3). Segregation of domains/complexes takes place in two stages: micro-phase separation then segregation. The magnitude of χ_{HC} —quantified by Equation (5)—will indicate the tendency of a heterochromatin-like domain/complex to micro-phase separate. The degree of phase separation of the domain/complex is specified by the product χN_{HC} (where N_{HC} is the number of “clutches” in the heterochromatin-like domain/complex). When χN_{HC} is greater than $\chi N_{\text{ODT_HC}}$ the domain/complex will become enriched owing to “bridging” by HP1 of H3K9me3-marked nucleosomes within and between clutches of the domain/complex (Diagram A in Box S1; top row, “compaction”) resulting in micro-phase separation from the flanking euchromatin (Diagram B in Box S1). Segregation is mediated by a like-with-like attraction (binding potential) between micro-phase separated heterochromatin-like domain/complexes that occurs in *cis* or *trans* (Diagrams B and C in Box S1). The binding potential results from an entropic effect. Since the heterochromatin-like domains/complexes have already given up entropy as a result of compaction (S_{COMP} in Equation (5)), they have less entropy to lose when they come into contact with each other compared to contact with euchromatic segments of chromatin fibres [221]. Once in contact, HP1 proteins can “bridge” between domains/complexes, forming inter-fibre bridges, which will stabilise far *cis*- and *trans*-contacts that emerge as ECDs in Hi-C experiments (Diagrams B and C in Box S1). Notably, work in *Drosophila* and mouse have shown that HP1 proteins form tightly localised domains in interphase nuclei that cannot be explained in terms of “liquid droplets” that form as a result of liquid–liquid phase separation [43,188,223]. Instead, the character (degree) of this phase separation is consistent with $\chi N_{\text{HC}} \gg \chi N_{\text{ODT_HC}}$, which would result in sharp interfaces between heterochromatin-like domains/complexes and euchromatin, reminiscent of what is observed for di-BCPs when χN is much greater than the χN_{ODT} ($\chi N \gg \chi N_{\text{ODT}}$; see ϕ_A vs. r_{\perp} for $\chi N \gg 10$ in Box S3). Equation (5) also shows how a value for χ_{PC} for a Pc-containing clutch can be quantified. The magnitude of χ_{PC} will indicate the tendency of a Pc-G domain/complex to micro-phase separate; the degree of micro-phase separation will be given by χN_{PC} (where N_{PC} is the number of “clutches” in the Pc-G domain/complex). Notably, the character of phase separated Pc bodies that contain Pc-G domains/complexes has been reported to be “liquid-like” [224,225]. It is tempting to speculate that this is because $\chi N_{\text{PC}} \geq \chi N_{\text{ODT_PC}}$ resulting in wavy, liquid-like, interfaces between the Pc-G domains/complexes and euchromatin. This would be similar to what is observed for di-BCPs when χN is close to χN_{ODT} ($\chi N \geq \chi N_{\text{ODT}}$; see ϕ_A vs. r_{\perp} for $\chi N \geq 10$ in Box S3).

The different degrees of phase separation observed with heterochromatin-like domains/complexes compared to Pc-G domains/complexes may reflect the finding that the binding affinity of HP1 for H3K9me3 is much higher than the affinity of Pc CBX homologues for H3K27me3 [226]. Given Equation (5) the difference in binding affinity would likely give a value of χ for a heterochromatin-like clutch that is greater than that for a Pc-G clutch ($\chi_{\text{HC}} > \chi_{\text{PC}}$). This would in turn determine the magnitude of χN_{HC} and χN_{PC} and thereby the character (degree) of phase separation (sharp interfaces vs. wavy, “liquid-like”) of the respective domains/complexes.

10. Conclusions and Perspectives

The hallmarks of constitutive heterochromatin, HP1 and H3K9me3, were present in the common ancestor of fission yeast and human around one billion years ago [16,17]. Both are enriched at canonical sites of constitutive heterochromatin, the peri-centric regions, (sub-)telomeres and (peri-)nucleolar organisers [18–25] and their function at these sites has been the subject of numerous studies in different organisms [213,227]. Less well studied are heterochromatin-like domains/complexes that share the hallmarks, HP1 and H3K9me3, with constitutive heterochromatin but lie *outside* the canonical constitutively heterochromatic territories [30,65]. Our survey shows that domains/complexes are likely to be present in most eukaryotic genomes (Figure 2; Table 1). One of the best characterised heterochromatin-like complexes is that which encompasses the 20 kb donor mating type region in fission yeast [228], indicating that heterochromatin-like domains/complexes represent an ancient mechanism for epigenetically regulating chromatin template-dependent processes in euchromatic regions of the genome outside canonical constitutive heterochromatin [16].

Heterochromatin-like domains/complexes have expanded in mammals where the number in the human genome is, conservatively, around 163–855 heterochromatin-like domains (>0.1 Mb) and around 18,853–32,292 complexes (<0.1 Mb) (Table 1). How heterochromatin-like domains/complexes are nucleated and assembled at specific sites in the genome has been reviewed recently using as an exemplar the B4 sub-compartment that is generated by far *cis*-contacts between the KRAB-ZNF heterochromatin-like domains [43,70]. Briefly, the KRAB-ZNF heterochromatin-like domains (up to 4 Mb) are nucleated by smaller complexes (\approx 6 kb) assembled at specific sites within the domains by sequence specific KRAB-ZFPs, where the density of nucleation sites is around 30 nucleosomes per 200 nucleosomes [43]. Targeting of nucleation sites by KRAB-ZFPs may be a common mechanism for assembly of larger domains. A recent survey of 222 of the 350 human KRAB-ZFPs showed that the number of genomic sites bound per protein ranged from more than 10,000 to around 15 [229], indicating that KRAB-ZFP-directed heterochromatin-like complexes could nucleate larger domains/complexes at many sites within the genome. A majority, 159 of the 222, of KRAB-ZFPs were found to bind at least one type of transposable element (TE) and their binding to TEs is thought assemble local heterochromatin-like complexes that repress transposon expression [229,230]. It is common for TE regulatory sequences to undergo exaptation and take on new functions [230,231]. It is tempting to speculate that a number of TE-specific KRAB-ZFP binding sites may have been subject to exaptation and now act as nucleation sites that assemble larger heterochromatin-like domains. Other genomic sites include imprinted gDMRs where heterochromatin-like complexes are assembled through binding of ZFP57/445 to their methylation-sensitive recognition sequences. Assembly of complexes at imprinted gDMRs preserves parent-of-origin-specific 5mC against the global DNA demethylation that characterises pre-implantation development in mouse and human (Figure 3) [102,123]. In addition to the heterochromatin-like domains/complexes, the related Pc-G domains/complexes number in the region of 2000 in mouse and human genomes [156]. Both heterochromatin-like and Pc-G domains are probably involved in regulating the phylotypic progression (Figure 4). As explained (Figure 4), during the phylotypic progression segregation of heterochromatin-like domains/complexes likely contribute cell-type-specific contact enrichments to ECDs that safeguard cellular identity, while Pc-containing domains/complexes contribute position-specific contact enrichments. ECDs represent the “epigenetic” component of cellular identity (Box S1; [43]).

Another aspect of cellular identity is age. In this context, it will be of interest to investigate the relationship between ECDs (Box S1) and the recently described “epigenetic clock” [232]. For humans, the “clock” is based on age-dependent changes in DNA methylation at 353 ‘clock’ CpGs, where methylation of 193 of the 353 CpGs increase with age while the remaining 160 CpGs decrease with age. The epigenetic clock has a high ticking rate until adulthood (\approx 20 years), after which it slows to a constant, steady, ticking rate that can be used to predict the age (or epigenetic age, eAge) of multiple tissues with a median error of 3.6 years [232]. It is known that HP1 proteins interact with DNA methyltransferases [130] and, as explained, HP1 proteins are associated with permissive as well

as repressive chromosomal regions [72,73] indicating that domains/complexes could regulate both the increasing (193 CpGs) and decreasing (160 CpGs) methylation changes observed with the 353 “clock” CpGs. This remains to be tested.

A key question concerns the contribution of DNA methylation to the generation of ECDs. DNA methylation is an essential epigenetic mechanism that regulates the cell-to-cell inheritance of gene repression patterns [233]. Its role in folding the genome in the interphase nucleus and the emergence ECDs detected in Hi-C maps may be indirect through an effect on the distribution of H3K27me3. DNA hypomethylation can affect the distribution H3K27me3 in the genome [234] most likely owing to targeting of PRC2 and PRC1 to CGIs [235,236] and the generalised affinity of PRCs for chromatin [237]. Notably, changing the distribution of H3K27me3 by DNA hypo-methylation affects compartmentalisation as detected in Hi-C maps; introduction of DNA methyltransferase activity reconstitutes both H3K27me3 distribution and the Hi-C maps [238].

BCP theory has been used to provide plausible explanations for changes in compartmentalisation seen in Hi-C maps when chromatin-associated cohesin activity is experimentally manipulated [43,164] (Figure 5). Implicit in these explanations there is a value of χ for heterochromatin-like domains/complexes flanked by euchromatin. A recent attempt at determining the degree of incompatibility between A- and B-type homo-polymers (euchromatin vs. heterochromatin) estimated χ as $0.03 \pm 0.01/\text{nucleosome}$ [190]. This value was thought crude because it was calculated based on phase separation of bulk homo-polymers [186], contrary to the known nuclear environment of a concentrated solution of chromatin “polymer” fibres. We suspect that the problem lies at a more fundamental level. As explained, mean field lattice theory [162,163] used to define χ , assumes the *unit of incompatibility* to be a monomer (Box S2; Equations (1)–(3)). Similarly, the seminal exposition on scaling concepts in polymer physics [186] used to estimate χ above [190] subsumes the monomer into a statistical segment (persistence length, l_p); the fundamental unit remains the monomer nonetheless. Instead, we suggest that the *unit of incompatibility* for chromatin is a larger more dynamic structure that probably represents the secondary structure of chromatin, namely the oligo-nucleosomal “clutch”, where a clutch contains from 2 to 10 nucleosomes (see Section 8 for details; Figure 6). In our “clutch” model incompatibility is due not to the epigenetic modifications of the nucleosomes per se, rather incompatibility is owing to proteins that “bridge” nucleosomes within a clutch, examples of which are HP1 and Polycomb CBX homologues that bridge H3K9me3- and H3K27me3-marked nucleosomes, respectively (Diagram A in Box S1). Our approach (Section 8; Figure 6; Equation (5)) could be used to quantify χ for both a heterochromatin-like “clutch” and a Pc-G “clutch”. For large domains/complexes consisting of N number of clutches the magnitude of χN will determine the character (degree) of phase separation (Section 9 for details). Each type of domain/complex has its own value for the order–disorder transition (χN_{ODT}) and when $\chi N \gg \chi N_{\text{ODT}}$ the interfaces between the domain/complex and euchromatin will be sharp and well-defined as observed for heterochromatin-like complexes [43,188,223]. Where $\chi N \geq \chi N_{\text{ODT}}$ the interfaces between the domain/complex and euchromatin will be wavy, “liquid-like”, as observed for Pc bodies that contain Pc-G domains/complexes [224,225].

The mesoscale organisation of chromatin as oligo-nucleosomal “clutches” containing 2–10 nucleosomes with a variable zig-zag organisation of nucleosomes appears to be a ubiquitous motif of interphase chromatin [196]. What causes this organisation to emerge from the “sea of nucleosomes” [196,239,240] that characterises interphase chromatin is not known. Given our “clutch” model described by Equation (5) the mesoscale organisation could arise from (transient) “bridging” of nucleosomes by proteins that possess two (or more) chromatin-binding motifs [241,242]. The sum of competing free energy contributions (for a heterochromatin-like clutch this would be $(H_{\text{CD-H3K9me}})_{\text{HC}}$ and $(H_{\text{CSD-CSD}})_{\text{HC}}$ in Equation (5)) and potential energies (H_{COMP} and H_{TL} in Equation (5)) could generate, in the dynamic environment of the interphase nucleus, oligo-nucleosomal “clutches” that vary from 2 to 10 nucleosomes. In this context, a simple application of Equation (5) to constitutive heterochromatin, where HP1 and H3K9me3-marked nucleosomal fibres are highly enriched [29],

predicts the formation of larger clutches with more pronounced zig-zag geometry. This is, in fact, what is observed [197,202].

11. Coda

- The oligo-nucleosomal “clutch” is the “unit of incompatibility” of chromatin.
- When a domain/complex assembles from “clutches” consisting of “bridged” nucleosomes with zig-zag geometry it will have a tendency, specified by χ , for micro-phase separation from flanking euchromatin.
- A qualitative prediction of the degree of micro-phase separation (wavy vs. sharp interfaces) is specified by the magnitude of χ_N compared to $\chi_{N_{ODT}}$ for a given domain/complex.
- Segregation of micro-phase separated domains/complexes facilitates folding of the genome into the confines of the nucleus while retaining an environment that enables (regulation of) chromatin template-dependent processes.

Availability of Data and Materials

Data	Original Paper	GEO Accession
HP1 α (CBX5), HP1 γ (CBX3), H3K9me3 in Human E1 ES cells	ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. <i>Nature</i> 2012 , <i>489</i> , 57-74, [59]. http://www.roadmapepigenomics.org/ [60]	GSE101135 GSE123229 GSE95868 GSE16256
HP1 α (CBX5), HP1 β (CBX1), H3K9me3 in HEK293 cells	LeRoy, G.; Chepelev, I.; DiMaggio, P.A.; Blanco, M.A.; Zee, B.M.; Zhao, K.; Garcia B.A. Proteogenomic characterization and mapping of nucleosomes decoded by Brd and HP1 proteins. <i>Genome Biol.</i> 2012 , <i>13</i> , R68, PMID: 22,897,906 [61]. Hattori, T.; Lai, D.; Dementieva, I.S.; Montano, S.P.; Kurosawa, K.; Zheng, Y.; Akin, L.R.; Świst-Rosowska, K.M.; Grzybowski, A.T.; Koide, A.; Krajewski, K.; Strahl, B.D.; Kelleher, N.L.; Ruthenburg, A.J.; Koide, S. Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation. <i>Proc Natl Acad Sci U S A.</i> 2016 , <i>113</i> , 2092-2097, doi:10.1073/pnas.1522691113 [62].	GSE39579 GSE66530
HP1 α (CBX5), HP1 β (CBX1), HP1 γ (CBX3), H3K9me3 in murine ES cells	Ostapczuk, V.; Mohn, F.; Carl, S.H.; Basters, A.; Hess, D.; Iesmantavicius, V.; Lampersberger, L.; Flemr, M.; Pandey, A.; Thomä, N.H.; Betschinger, J.; Bühler M. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. <i>Nature</i> 2018 , <i>557</i> , 739-743, doi:10.1038/s41586-018-0153-8. [45].	GSE97945
HP1 α and H3K9me3 in <i>D. melanogaster</i>	Klenov, M.S.; Lavrov, S.A.; Korbut, A.P.; Stolyarenko, A.D.; Yakushev, E.Y.; Reuter, M.; Pillai, R.S.; Gvozdev, V.A. Impact of nuclear Piwi elimination on chromatin state in <i>Drosophila melanogaster</i> ovaries. <i>Nucleic Acids Res.</i> 2014 , <i>42</i> , 6208-18. PMID: 24,782,529 [63].	GSE56347
Swi6 ^{HP1} and H3K9me3 in <i>S. pombe</i>	Iglesias, N.; Paulo, J.A.; Tatarakis, A.; Wang, X.; Edwards, A.L.; Bhanu, N.V.; Garcia, B.A.; Haas, W.; Gygi, S.P.; Moazed, D. Native Chromatin Proteomics Reveals a Role for Specific Nucleoporins in Heterochromatin Organization and Maintenance. <i>Mol. Cell</i> 2020 , <i>77</i> , 51-66.e8. PMID: 31,784,357 [64].	GSE125935

Data Analysis: We first trimmed ChIP-seq raw data (see Table 1 for source of the ChIP-Seq data for each organism) for quality using Trimmomatic SE and then aligned the trimmed data to GRCh38 (Ensembl91, Human), GRCm38 (Ensembl95, Mouse), BDGP6 (Ensembl95, *D. melanogaster*) and ASM294 (*S. pombe*) using Bowtie2 [243]. Peak calling was performed with the MACS2 tool using as the parameters: (i) –broad; (ii) –max-gap 5000 [244]. After peak calling, the fold enrichment profiles of sample signal versus input were calculated using the *bdgcmp* function from the MACS2 tool (method-fold enrichment (-m FE; [244])). The profiles obtained were converted to bigWig files using the bedGraphToBigWig tool with default parameters (<https://www.encodeproject.org/software/bedgraph-tobigwig/>). The bigWig files were then used for the pairwise Pearson correlation analysis of H3K9me3 and HP1 isoforms (see Figure 2) using the multiBigWigSummary from deepTools [245]. The parameters were: (i) bin size 5 kb

for human, mouse and *D. melanogaster* and (ii) 200 bp for *S. pombe*, (iii) coordinates of all sequences that have characteristics of constitutive heterochromatin were supplied as a blacklist (–blackListFileName) and were excluded from the analysis. To define regions with constitutive heterochromatin properties we used ENCODE blacklists—lists of genome regions that demonstrate anomalous signal in next-generation sequencing mostly due to repetitive sequences—and matched it with H3K9me3-enrichment binding profile in IGV-browser to obtain coordinates of pericentromeric and subtelomeric heterochromatin regions [246]. Unplaced scaffolds enriched with repetitive sequences were also considered as related to constitutive heterochromatin and were excluded from the analysis. The correlations were visualised using the R package *ggpubr*. In order to calculate the number and sizes of heterochromatin-like domains/complexes we determined the number of continuous segments of genome outside of constitutive heterochromatin regions that were enriched in both H3K9me3 and HP1 isoform (in human, mouse, *D. melanogaster* and *S. pombe*; see Table 1) with ChromHMM tool in 200 bp bins (for all HP1s) [247]. For human data 5-state model, for mouse data 4-state model, and for fly and yeast 2-state models were built. After coordinates of regions enriched in H3K9me3 and HP1 isoforms were determined we calculated the number and sizes of domains (>1 Mb and 0.1–1 Mb) and complexes (0.01–0.1 Mb) in R.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4409/9/8/1881/s1>, Box S1: Epigenetic compartmental domains (ECDs) in mammals; Box S2: The Flory-Huggins parameter (χ) and the monomer as the “unit of incompatibility”; Box S3: The segregation product χ_N and order-disorder (ODT) transition of BCPs; Table S1: Size range of heterochromatin-like complexes in man (H1 ES cells and 293T cells) and mouse (ES cells).

Author Contributions: P.B.S. conceived of the synthesis presented, wrote the paper and drafted all figures. S.N.B. drew the figures except for Figure 2. P.P.L. undertook the bioinformatic analyses and assembled Figure 2. All authors made comments on the manuscript and approved the final version. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from the Ministry of Education and Science of Russian Federation (grant no. 14.Y26.31.0024).

Acknowledgments: We are grateful to Andrew Newman for advice on bioinformatics and many helpful discussions. We thank Andrew Newman, Jafar Sharif and FS Bates for comments on the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

5mC: 5-methylcytosine; Å: Angstrom; α : segment polarisability; α : coefficient for enthalpy; ADD: ATRX-DNMT3-DNMT3L domain; A-P: anterior-posterior; ATRX: Alpha Thalassemia/Mental Retardation Syndrome X-Linked; β : coefficient for excess entropy; BCP: block copolymer; bp: base pair; CBX: Chromobox homologue; CD: chromodomain; *cen3*: fission yeast centromere 3; CGIs: CpG islands; CpG: CpG di-nucleotide; CSD: chromo shadow domain; CTCF: CCCTC-binding factor; DAXX; Death Domain associated protein; di-BCP: di-block copolymer; DNMT1: maintenance DNA methyltransferase 1; ϵ_{AB} , ϵ_{AA} and ϵ_{BB} ; contact energies per repeat unit of A-B, A-A and B-B, respectively; ECD: epigenetic compartmental domain; ϵ_{ij} : contact energy; ϵ_{ps} , ϵ_{pp} and ϵ_{ss} ; contact energies of polymer–solvent, polymer–polymer and solvent–solvent, respectively; Eso1: SMC3 acetyltransferase; ES: embryonic stem; FRAP: fluorescence recovery after photo-bleaching; gDMRs: germline differentially methylated regions; f_A : volume fraction of A; f_B : volume fraction of B; H1.4K26me: methylated lysine 26 of histone H1.4; H3K4me3: tri-methylated lysine 4 on histone H3; H3K9me2: di-methylated lysine 9 on histone H3; H3K9me3: tri-methylated lysine 9 on histone H3; H3K27me3: tri-methylated lysine 27 on histone H3; H4K16: lysine 16 of histone H4; H4K20me3: tri-methylated lysine 20 on histone H4; HC: heterochromatin-like “clutch”; HMTases: histone methyltransferases; HP1: Heterochromatin Protein 1; *Hox*: homeobox; HR: hinge region; ICM: inner cellular mass; KAP1: KRAB-associated protein 1; kb: kilobases; k_B : Boltzmann’s constant; $k_B T$: thermal energy; kDa; kiloDalton; I : first ionisation potential; i and j segments; “blocks” of a di-BCP each consisting of (dissimilar) monomer repeating units; KRAB-ZNPs: Krüppel-associated box (KRAB) domain zinc-finger proteins; KRAB-ZNF: KRAB domain-zinc finger; LEF: loop-extruding factor; Mb: Mega-bases; ncRNA: non-coding RNA; $N = N_A + N_B$; total degree of polymerisation of AB di-BCP; NORs: nucleolar organiser regions; ODT: order–disorder transition; Pc-G: Polycomb-Group consisting of H3K27me3, PRC1 and PRC2 complexes; Pds5A/B; sister chromatin cohesin protein 5A/B; l_p : persistence length; PEV: position-effect variegation; PI-b-PS: polyisoprene-block-polystyrene; PRC1: Polycomb-repressive complex; PRC2: Polycomb-repressive complex 2; RICC-Seq; Radiation-induced spatially correlated cleavage of DNA with sequencing; PxVxL: Proline/Any/Valine/Any/Leucine penta-peptide motif; r_{ij} : segment-segment separation; ϕ_A : composition of A; $\langle r \rangle$: ensemble average of bond vectors; r_{\perp} : ensemble bond vectors perpendicular to periodic mesophase; SETDB1: SET Domain Bifurcated 1 K9H3 HMTase; Sir: silent information regulator; SCC1: cohesin subunit Scc1; SMC3: cohesin subunit SMC3; SUMO2: Small ubiquitin-related modifier 2; Suv39h1/2: mammalian suvar K9H3 HMTase 1 and 2; Swi6^{HP1}: HP1 homologue swi6p from fission yeast; T: temperature; TAD: topologically associated domain; TEs: transposable elements; Tet: Ten-eleven translocation dioxygenase; UBE2i: Ubiquitin conjugating enzyme 2i; WAPL: WAPL cohesin release factor; χ : Flory–Huggins parameter; χ_{HC} : specifies the degree of incompatibility between a heterochromatin-like “clutch” and a euchromatic “clutch”; χ_N : segregation product that specifies the degree of phase separation; $\chi_{N_{ODT_{HC}}}$: magnitude of segregation product at the order–disorder transition for a heterochromatin-like domain/complex. $\chi_{N_{ODT_{PC}}}$: magnitude of segregation product at the order–disorder transition for a Pc-G domain/complex.

References

1. Wolffe, A.P. *Chromatin: Structure and Function*, 3rd ed.; Academic Press: London, UK, 2012.
2. Lewis, E.B. The phenomenon of position effect. *Adv. Genet.* **1950**, *3*, 73–115. [[PubMed](#)]
3. Baker, W.K. Position-effect variegation. In *Advances in Genetics*; Caspari, E.W., Ed.; Elsevier: Amsterdam, The Netherlands, 1968; Volume 14, pp. 133–169.
4. Spofford, J.B. Position-effect variegation in *Drosophila*. In *The Genetics and Biology of Drosophila*; Ashburner, M., Novitski, E., Eds.; Academic Press: New York, NY, USA, 1976; Volume 1, pp. 955–1018.
5. Elgin, S.C.; Reuter, G. Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a017780. [[CrossRef](#)] [[PubMed](#)]
6. Schultz, J. The relation of the heterochromatic chromosome regions to the nucleic acids of the cell. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 1956; Volume 21, pp. 307–328.
7. Henikoff, S. Position-effect variegation and chromosome structure of a heat shock puff in *Drosophila*. *Chromosoma* **1981**, *83*, 381–393. [[CrossRef](#)] [[PubMed](#)]
8. Zhimulev, I.F.; Belyaeva, E.S.; Fomina, O.V.; Protopopov, M.O.; Bolshakov, V.N. Cytogenetic and molecular aspects of position effect variegation in *Drosophila melanogaster*. *Chromosoma* **1986**, *94*, 492–504. [[CrossRef](#)]
9. Weiler, K.S.; Wakimoto, B.T. Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet.* **1995**, *29*, 577–605. [[CrossRef](#)] [[PubMed](#)]
10. Rudolph, T.; Yonezawa, M.; Lein, S.; Heidrich, K.; Kubicek, S.; Schäfer, C.; Phalke, S.; Walther, M.; Schmidt, A.; Jenuwein, T.; et al. Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU (VAR) 3–3. *Mol. Cell* **2007**, *26*, 103–115. [[CrossRef](#)]
11. Locke, J.; Kotarski, M.A.; Tartof, K.D. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **1988**, *120*, 181–198.
12. Wustmann, G.; Szidonya, J.; Taubert, H.; Reuter, G. The genetics of position-effect variegation modifying loci in *Drosophila melanogaster*. *Mol. Gen. Genet.* **1989**, *217*, 520–527. [[CrossRef](#)]
13. Hayashi, S.; Ruddell, A.; Sinclair, D.; Grigliatti, T. Chromosomal structure is altered by mutations that suppress or enhance position effect variegation. *Chromosoma* **1990**, *99*, 391–400. [[CrossRef](#)]
14. Jenuwein, T.; Allis, C.D. Translating the histone code. *Science* **2001**, *293*, 1074–1080. [[CrossRef](#)]
15. Becker, J.S.; Nicetto, D.; Zaret, K.S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet.* **2016**, *32*, 29–41. [[CrossRef](#)] [[PubMed](#)]
16. Wang, G.; Ma, A.; Chow, C.M.; Horsley, D.; Brown, N.R.; Cowell, I.G.; Singh, P.B. Conservation of heterochromatin protein 1 function. *Mol. Cell. Biol.* **2000**, *20*, 6970–6983. [[CrossRef](#)] [[PubMed](#)]
17. Elgin, S.C.; Grewal, S.I. Heterochromatin: Silence is golden. *Curr. Biol.* **2003**, *13*, R895–R898. [[CrossRef](#)] [[PubMed](#)]
18. Cam, H.P.; Sugiyama, T.; Chen, E.S.; Chen, X.; FitzGerald, P.C.; Grewal, S.I. Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **2005**, *37*, 809–819. [[CrossRef](#)]
19. Kellum, R.; Alberts, B.M. Heterochromatin protein 1 is required for correct chromosome segregation in *Drosophila* embryos. *J. Cell Sci.* **1995**, *108*, 1419–1431.
20. Schotta, G.; Ebert, A.; Krauss, V.; Fischer, A.; Hoffmann, J.; Rea, S.; Jenuwein, T.; Dorn, R.; Reuter, G. Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J.* **2002**, *21*, 1121–1131. [[CrossRef](#)]
21. Fanti, L.; Giovinazzo, G.; Berloco, M.; Pimpinelli, S. The heterochromatin protein 1 prevents telomere fusions in *Drosophila*. *Mol. Cell* **1998**, *2*, 527–538. [[CrossRef](#)]
22. Cowell, I.G.; Aucott, R.; Mahadevaiah, S.K.; Burgoyne, P.S.; Huskisson, N.; Bongiorno, S.; Prantera, G.; Fanti, L.; Pimpinelli, S.; Wu, R.; et al. Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **2002**, *111*, 22–36. [[CrossRef](#)]
23. Peng, J.C.; Karpen, G.H. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat. Cell Biol.* **2007**, *9*, 25–35. [[CrossRef](#)]
24. García-Cao, M.; O’Sullivan, R.; Peters, A.H.; Jenuwein, T.; Blasco, M.A. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat. Genet.* **2004**, *36*, 94–99. [[CrossRef](#)]

25. Chakrabarti, R.; Sanyal, S.; Ghosh, A.; Bhar, K.; Das, C.; Siddhanta, A. Phosphatidylinositol-4-phosphate 5-kinase 1 α modulates ribosomal RNA gene silencing through its interaction with histone H3 lysine 9 trimethylation and heterochromatin protein HP1- α . *J. Biol. Chem.* **2015**, *290*, 20893–20903. [[CrossRef](#)] [[PubMed](#)]
26. Johnson, A.; Li, G.; Sikorski, T.W.; Buratowski, S.; Woodcock, C.L.; Moazed, D. Reconstitution of heterochromatin-dependent transcriptional gene silencing. *Mol. Cell* **2009**, *35*, 769–781. [[CrossRef](#)] [[PubMed](#)]
27. Doolittle, R.F.; Feng, D.F.; Tsang, S.; Cho, G.; Little, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **1996**, *271*, 470–477. [[CrossRef](#)] [[PubMed](#)]
28. Schmidt, M.W.; Houseman, A.; Ivanov, A.R.; Wolf, D.A. Comparative proteomic and transcriptomic profiling of the fission yeast *Schizosaccharomyces pombe*. *Mol. Syst. Biol.* **2007**, *3*, 79. [[CrossRef](#)]
29. Hiragami-Hamada, K.; Soeroes, S.; Nikolov, M.; Wilkins, B.; Kreuz, S.; Chen, C.; De La Rosa-Velázquez, I.A.; Zenn, H.M.; Kost, N.; Pohl, W.; et al. Dynamic and flexible H3K9me3 bridging via HP1 β dimerization establishes a plastic state of condensed chromatin. *Nat. Commun.* **2016**, *7*, 11310. [[CrossRef](#)]
30. Singh, P.B.; Miller, J.R.; Pearce, J.; Kothary, R.; Burton, R.D.; Paro, R.; James, T.C.; Gaunt, S.J. A sequence motif found in a *Drosophila* heterochromatin protein is conserved in animals and plants. *Nucleic Acids Res.* **1991**, *19*, 789–794. [[CrossRef](#)]
31. Aucott, R.; Bullwinkel, J.; Yu, Y.; Shi, W.; Billur, M.; Brown, J.P.; Menzel, U.; Kioussis, D.; Wang, G.; Reiser, I.; et al. HP1-beta is required for development of the cerebral neocortex and neuromuscular junctions. *J. Cell Biol.* **2008**, *183*, 597–606. [[CrossRef](#)]
32. Aasland, R.; Stewart, A.F. The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res.* **1995**, *23*, 3168–3173. [[CrossRef](#)]
33. Epstein, H.; James, T.C.; Singh, P.B. Cloning and expression of *Drosophila* HP1 homologs from a mealybug *Planococcus citri*. *J. Cell Sci.* **1992**, *101*, 463–474.
34. Ball, L.J.; Murzina, N.V.; Broadhurst, R.W.; Raine, A.R.; Archer, F.J.; Stott, F.J.; Murzin, A.G.; Singh, P.B.; Dmaille, P.J.; Laue, E.D. Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *EMBO J.* **1997**, *16*, 2473–2481. [[CrossRef](#)]
35. Nielsen, P.R.; Nietlispach, D.; Mott, H.R.; Callaghan, J.; Bannister, A.; Kouzarides, T.; Murzin, A.G.; Murzina, N.V.; Laue, E.D. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **2002**, *416*, 103–107. [[CrossRef](#)] [[PubMed](#)]
36. Mahadevi, A.S.; Sastry, G.N. Cation- π interaction: Its role and relevance in chemistry, biology, and material science. *Chem. Rev.* **2013**, *113*, 2100–2138. [[CrossRef](#)] [[PubMed](#)]
37. Thiru, A.; Nietlispach, D.; Okuwaki, M.; Lyon, D.; Nielsen, P.R.; Hirshberg, M.; Verreault, A.; Murzina, N.V.; Laue, E.D. Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J.* **2004**, *23*, 489–499. [[CrossRef](#)] [[PubMed](#)]
38. Smothers, J.F.; Henikoff, S. The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Mol. Cell. Biol.* **2001**, *21*, 2555–2569. [[CrossRef](#)] [[PubMed](#)]
39. Ayoub, N.; Jeyasekharan, A.D.; Bernal, J.A.; Venkitaraman, A.R. HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature* **2008**, *453*, 682–686. [[CrossRef](#)] [[PubMed](#)]
40. Billur, M.; Bartunik, H.D.; Singh, P.B. The essential function of HP1 beta: A case of the tail wagging the dog? *Trends Biochem. Sci.* **2010**, *35*, 115–123. [[CrossRef](#)]
41. Singh, P.B. Molecular mechanisms of cellular determination: Their relation to chromatin structure and parental imprinting. *J. Cell Sci.* **1994**, *107*, 2653–2668.
42. Jones, D.O.; Cowell, I.G.; Singh, P.B. Mammalian chromodomain proteins: Their role in genome organisation and expression. *Bioessays* **2000**, *22*, 124–137. [[CrossRef](#)]
43. Singh, P.B.; Newman, A.G. On the relations of phase separation and Hi-C maps to epigenetics. *R. Soc. Open Sci.* **2020**, *7*, 191976. [[CrossRef](#)]
44. Li, Y.; Kirschmann, D.A.; Wallrath, L.L. Does heterochromatin protein 1 always follow code? *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16462–16469. [[CrossRef](#)]
45. Ostapcuk, V.; Mohn, F.; Carl, S.H.; Basters, A.; Hess, D.; Iesmantavicius, V.; Lampersberger, L.; Flemr, M.; Pandey, A.; Thomä, N.H.; et al. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* **2018**, *557*, 739–743. [[CrossRef](#)] [[PubMed](#)]
46. Cryderman, D.E.; Grade, S.K.; Li, Y.; Fanti, L.; Pimpinelli, S.; Wallrath, L.L. Role of *Drosophila* HP1 in euchromatic gene expression. *Dev. Dyn.* **2005**, *232*, 767–774. [[CrossRef](#)] [[PubMed](#)]

47. Perrini, B.; Piacentini, L.; Fanti, L.; Altieri, F.; Chichiarelli, S.; Berloco, M.; Turano, C.; Ferraro, A.; Pimpinelli, S. HP1 controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in *Drosophila*. *Mol. Cell* **2004**, *15*, 467–476. [CrossRef] [PubMed]
48. Nielsen, A.L.; Oulad-Abdelghani, M.; Ortiz, J.A.; Remboutsika, E.; Chambon, P.; Losson, R. Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol. Cell* **2001**, *7*, 729–739. [CrossRef]
49. Lavigne, M.; Eskeland, R.; Azebi, S.; Saint-André, V.; Jang, S.M.; Batsché, E.; Fan, H.Y.; Kingston, R.E.; Imhof, A.; Muchardt, C. Interaction of HP1 and Brg1/Brm with the globular domain of histone H3 is required for HP1-mediated repression. *PLoS Genet.* **2009**, *5*, e1000769. [CrossRef] [PubMed]
50. Richart, A.N.; Brunner, C.I.; Stott, K.; Murzina, N.V.; Thomas, J.O. Characterization of chromoshadow domain-mediated binding of heterochromatin protein 1 α (HP1 α) to histone H3. *J. Biol. Chem.* **2012**, *287*, 18730–18737. [CrossRef]
51. Daujat, S.; Zeissler, U.; Waldemann, T.; Happel, N.; Schneider, R. HP1 binds Specifically to Lys²⁶-methylated Histone 1.4, whereas Simultaneous Ser²⁷ Phosphorylation Blocks HP1 binding. *J. Biol. Chem.* **2005**, *280*, 38090–38095. [CrossRef]
52. Maison, C.; Bailly, D.; Peters, A.H.; Quivy, J.-P.; Roche, D.; Taddei, A.; Lachner, M.; Jenuwein, T.; Almouzni, G. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **2002**, *30*, 329. [CrossRef]
53. Muchardt, C.; Guillemé, M.; Seeler, J.-S.; Trouche, D.; Dejean, A.; Yaniv, M. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1 α . *EMBO Rep.* **2002**, *3*, 975–981. [CrossRef]
54. Meehan, R.R.; Kao, C.-F.; Pennings, S. HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. *EMBO J.* **2003**, *22*, 3164–3174. [CrossRef]
55. Sanulli, S.; Trnka, M.J.; Dharmarajan, V.; Tibble, R.W.; Pascal, B.D.; Burlingame, A.L.; Griffin, P.R.; Gross, J.D.; Narlikar, G.J. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature* **2019**, *575*, 390–394. [CrossRef] [PubMed]
56. Cheutin, T.; McNairn, A.J.; Jenuwein, T.; Gilbert, D.M.; Singh, P.B.; Misteli, T. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **2003**, *299*, 721–725. [CrossRef] [PubMed]
57. Schmiedeberg, L.; Weisshart, K.; Diekmann, S.; Meyer zu Hoerste, G.; Hemmerich, P. High- and low mobility populations of HP1 in heterochromatin of mammalian cells. *Mol. Biol. Cell* **2004**, *15*, 2819–2833. [CrossRef] [PubMed]
58. Pidoux, A.L.; Allshire, R.C. Kinetochores and heterochromatin domains of the fission yeast centromere. *Chromosome Res.* **2004**, *12*, 521–534. [CrossRef] [PubMed]
59. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74. [CrossRef]
60. Roadmap Epigenomics Project. Available online: <http://www.roadmapepigenomics.org/> (accessed on 15 March 2020).
61. LeRoy, G.; Chepelev, I.; DiMaggio, P.A.; Blanco, M.A.; Zee, B.M.; Zhao, K.; Garcia, B.A. Proteogenomic characterization and mapping of nucleosomes decoded by Brd and HP1 proteins. *Genome Biol.* **2012**, *13*, R68. [CrossRef]
62. Hattori, T.; Lai, D.; Dementieva, I.S.; Montañó, S.P.; Kurosawa, K.; Zheng, Y.; Akin, L.R.; Świst-Rosowska, K.M.; Grzybowski, A.T.; Koide, A.; et al. Antigen clasp by two antigen-binding sites of an exceptionally specific antibody for histone methylation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 2092–2097. [CrossRef]
63. Klenov, M.S.; Lavrov, S.A.; Korbut, A.P.; Stolyarenko, A.D.; Yakushev, E.Y.; Reuter, M.; Pillai, R.S.; Gvozdev, V.A. Impact of nuclear *Piwi* elimination on chromatin state in *Drosophila melanogaster* ovaries. *Nucleic Acids Res.* **2014**, *42*, 6208–6218. [CrossRef]
64. Iglesias, N.; Paulo, J.A.; Tatarakis, A.; Wang, X.; Edwards, A.L.; Bhanu, N.V.; Garcia, B.A.; Haas, W.; Gygi, S.P.; Moazed, D. Native Chromatin Proteomics Reveals a Role for Specific Nucleoporins in Heterochromatin Organization and Maintenance. *Mol. Cell* **2020**, *77*, 51–66.e8. [CrossRef]
65. Singh, P.B. Heterochromatin and the molecular mechanisms of ‘parent-of-origin’ effects in animals. *J. Biosci.* **2016**, *41*, 759–786. [CrossRef]

66. Horsley, D.; Hutchings, A.; Butcher, G.W.; Singh, P.B. M32, a murine homologue of *Drosophila* heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin. *Cytogenet Cell Genet.* **1996**, *73*, 308–311. [[CrossRef](#)] [[PubMed](#)]
67. Ryan, R.F.; Schultz, D.C.; Ayyanathan, K.; Singh, P.B.; Friedman, J.R.; Fredericks, W.J.; Rauscher, F.J. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: A potential role for Krüppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **1999**, *19*, 4366–4378. [[CrossRef](#)] [[PubMed](#)]
68. Vogel, M.J.; Guelen, L.; de Wit, E.; Peric-Hupkes, D.; Lodén, M.; Talhout, W.; Feenstra, M.; Abbas, B.; Classen, A.K.; van Steensel, B. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res.* **2006**, *16*, 1493–1504. [[CrossRef](#)] [[PubMed](#)]
69. Groner, A.C.; Meylan, S.; Ciuffi, A.; Zangger, N.; Ambrosini, G.; Déneraud, N.; Bucher, P.; Trono, D. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* **2010**, *6*, e1000869. [[CrossRef](#)]
70. Rao, S.S.; Huntley, M.H.; Durand, N.C.; Stamenova, E.K.; Bochkov, I.D.; Robinson, J.T.; Sanborn, A.L.; Machol, I.; Omer, A.D.; Lander, E.S.; et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **2014**, *159*, 1665–1680. [[CrossRef](#)] [[PubMed](#)]
71. Iyengar, S.; Farnham, P.J. KAP1 protein: An enigmatic master regulator of the genome. *J. Biol. Chem.* **2011**, *286*, 26267–26276. [[CrossRef](#)]
72. Hediger, F.; Gasser, S.M. Heterochromatin protein 1: Don't judge the book by its cover! *Curr. Opin. Genet. Dev.* **2006**, *16*, 143–150. [[CrossRef](#)]
73. Kwon, S.H.; Workman, J.L. The heterochromatin protein 1 (HP1) family: Put away a bias toward HP1. *Mol. Cell* **2008**, *26*, 217–227.
74. Piacentini, L.; Fanti, L.; Berloco, M.; Perrini, B.; Pimpinelli, S. Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. *J. Cell Biol.* **2003**, *161*, 707–714. [[CrossRef](#)]
75. Smallwood, A.; Hon, G.C.; Jin, F.; Henry, R.E.; Espinosa, J.M.; Ren, B. CBX3 regulates efficient RNA processing genome-wide. *Genome Res.* **2012**, *22*, 1426–1436. [[CrossRef](#)]
76. Vakoc, C.R.; Mandat, S.A.; Olenchock, B.A.; Blobel, G.A. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **2005**, *19*, 381–391. [[CrossRef](#)] [[PubMed](#)]
77. Yearim, A.; Gelfman, S.; Shayevitch, R.; Melcer, S.; Glaich, O.; Mallm, J.P.; Nissim-Rafinia, M.; Cohen, A.H.; Rippe, K.; Meshorer, E.; et al. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.* **2015**, *10*, 1122–1134. [[CrossRef](#)] [[PubMed](#)]
78. Dinant, C.; Luijsterburg, M.S. The emerging role of HP1 in the DNA damage response. *Mol. Cell. Biol.* **2009**, *29*, 6335–6340. [[CrossRef](#)] [[PubMed](#)]
79. Li, X.; Ito, M.; Zhou, F.; Youngson, N.; Zuo, X.; Leder, P.; Ferguson-Smith, A.C. A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Dev. Cell* **2008**, *15*, 547–557. [[CrossRef](#)] [[PubMed](#)]
80. Bateson, W. *Mendel's Principles of Heredity*; Cambridge University Press: Cambridge, UK, 1909.
81. Mendel, G. Versuche über Pflanzen-Hybriden. *Verh. Naturforsch. Ver. Brunn* **1866**, *4*, 3–47, English translation in *J. R. Hort. Soc.* **1901**, *26*, 1–32.
82. Metz, C.W. Chromosome behavior, inheritance and sex determination in *Sciara*. *Am. Nat.* **1938**, *72*, 485–520. [[CrossRef](#)]
83. Brown, S.W.; Nur, U. Heterochromatic chromosomes in the coccids. *Science* **1964**, *145*, 130–136. [[CrossRef](#)]
84. McGrath, J.; Solter, D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **1984**, *37*, 179–183. [[CrossRef](#)]
85. Surani, M.A.; Barton, S.C.; Norris, M.L. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **1984**, *308*, 548–550. [[CrossRef](#)]
86. Henckel, A.; Arnaud, P. Genome-wide identification of new imprinted genes. *Brief. Funct. Genomics* **2010**, *9*, 304–314. [[CrossRef](#)]
87. Williamson, C.M.; Blake, A.; Thomas, S.; Beechey, C.V.; Hancock, J.; Cattanach, B.M.; Peters, J.; MRC Harwell; Oxfordshire. World Wide Web Site-Mouse Imprinting Data and References. Available online: http://www.har.mrc.ac.uk/research/genomic_imprinting (accessed on 16 July 2020).
88. Ferguson-Smith, A.C. Genomic imprinting: The emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **2011**, *12*, 565–575. [[CrossRef](#)] [[PubMed](#)]

89. Barlow, D.P.; Bartolomei, M.S. Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a018382. [[CrossRef](#)]
90. Li, E.; Beard, C.; Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **1993**, *366*, 362–365. [[CrossRef](#)] [[PubMed](#)]
91. Sasaki, H.; Matsui, Y. Epigenetic events in mammalian germ-cell development: Reprogramming and beyond. *Nat. Rev. Genet.* **2008**, *9*, 129–140. [[CrossRef](#)]
92. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **2011**, *25*, 1010–1022. [[CrossRef](#)]
93. Saxonov, S.; Berg, P.; Brutlag, D.L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1412–1417. [[CrossRef](#)]
94. Tomizawa, S.; Kobayashi, H.; Watanabe, T.; Andrews, S.; Hata, K.; Kelsey, G.; Sasaki, H. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* **2011**, *138*, 811–820. [[CrossRef](#)] [[PubMed](#)]
95. Kobayashi, H.; Suda, C.; Abe, T.; Kohara, Y.; Ikemura, T.; Sasaki, H. Bisulfite sequencing and dinucleotide content analysis of 15 imprinted mouse differentially methylated regions (DMRs): Paternally methylated DMRs contain less CpGs than maternally methylated DMRs. *Cytogenet Genome Res.* **2006**, *113*, 130–137. [[CrossRef](#)]
96. Wang, L.; Zhang, J.; Duan, J.; Gao, X.; Zhu, W.; Lu, X.; Yang, L.; Zhang, J.; Li, G.; Ci, W.; et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **2014**, *157*, 979–991. [[CrossRef](#)]
97. Hanna, C.W.; Kelsey, G. The specification of imprints in mammals. *Hered.* **2014**, *113*, 176–183. [[CrossRef](#)] [[PubMed](#)]
98. Smallwood, S.A.; Tomizawa, S.; Krueger, F.; Ruf, N.; Carli, N.; Segonds-Pichon, A.; Sato, S.; Hata, K.; Andrews, S.R.; Kelsey, G. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **2011**, *43*, 811–814. [[CrossRef](#)] [[PubMed](#)]
99. Kobayashi, H.; Sakurai, T.; Imai, M.; Takahashi, N.; Fukuda, A.; Yayoi, O.; Sato, S.; Nakabayashi, K.; Hata, K.; Sotomaru, Y.; et al. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet.* **2012**, *8*, e1002440. [[CrossRef](#)] [[PubMed](#)]
100. Stewart, K.R.; Veselovska, L.; Kelsey, G. Establishment and functions of DNA methylation in the germline. *Epigenomics* **2016**, *8*, 1399–1413. [[CrossRef](#)]
101. Sendžikaitė, G.; Kelsey, G. The role and mechanisms of DNA methylation in the oocyte. *Essays Biochem.* **2019**, *63*, 691–705. [[CrossRef](#)]
102. Smith, Z.D.; Chan, M.M.; Mikkelsen, T.S.; Gu, H.; Gnirke, A.; Regev, A.; Meissner, A. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **2012**, *484*, 339–344. [[CrossRef](#)]
103. Guo, H.; Zhu, P.; Yan, L.; Li, R.; Hu, B.; Lian, Y.; Yan, J.; Ren, X.; Lin, S.; Li, J.; et al. The DNA methylation landscape of human early embryos. *Nature* **2014**, *511*, 606–610. [[CrossRef](#)]
104. Leseva, M.; Knowles, B.B.; Messerschmidt, D.M.; Solter, D. Erase-Maintain-Establish: Natural Reprogramming of the Mammalian Epigenome. *Cold Spring Harb. Symp. Quant. Biol.* **2015**, *80*, 155–163. [[CrossRef](#)]
105. Messerschmidt, D.M.; Knowles, B.B.; Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **2014**, *28*, 812–828. [[CrossRef](#)]
106. Tucci, V.; Isles, A.R.; Kelsey, G.; Ferguson-Smith, A.C.; Erice Imprinting Group. Genomic Imprinting and Physiological Processes in Mammals. *Cell* **2019**, *176*, 952–965. [[CrossRef](#)]
107. Quenneville, S.; Verde, G.; Corsinotti, A.; Kapopoulou, A.; Jakobsson, J.; Offner, S.; Baglivo, I.; Pedone, P.V.; Grimaldi, G.; Riccio, A.; et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol. Cell* **2011**, *44*, 361–372. [[CrossRef](#)]
108. Takahashi, N.; Coluccio, A.; Thorball, C.W.; Planet, E.; Shi, H.; Offner, S.; Turelli, P.; Imbeault, M.; Ferguson-Smith, A.C.; Trono, D. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* **2019**, *33*, 49–54. [[CrossRef](#)] [[PubMed](#)]
109. Anvar, Z.; Cammisa, M.; Riso, V.; Baglivo, I.; Kukreja, H.; Sparago, A.; Girardot, M.; Lad, S.; De Feis, I.; Cerrato, F.; et al. ZFP57 recognizes multiple and closely spaced sequence motif variants to maintain repressive epigenetic marks in mouse embryonic stem cells. *Nucleic Acids Res.* **2016**, *44*, 1118–1132. [[CrossRef](#)] [[PubMed](#)]
110. Liu, Y.; Toh, H.; Sasaki, H.; Zhang, X.; Cheng, X. An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. *Genes Dev.* **2012**, *26*, 2374–2379. [[CrossRef](#)] [[PubMed](#)]

111. Messerschmidt, D.M.; de Vries, W.; Ito, M.; Solter, D.; Ferguson-Smith, A.; Knowles, B.B. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* **2012**, *335*, 1499–1502. [[CrossRef](#)]
112. Cho, S.; Park, J.S.; Kwon, S.; Kang, Y.K. Dynamics of Setdb1 expression in early mouse development. *Gene Expr. Patterns* **2012**, *12*, 213–218. [[CrossRef](#)]
113. Kim, J.; Zhao, H.; Dan, J.; Kim, S.; Hardikar, S.; Hollowell, D.; Lin, K.; Lu, Y.; Takata, Y.; Shen, J.; et al. Maternal Setdb1 Is Required for Meiotic Progression and Preimplantation Development in Mouse. *PLoS Genet.* **2016**, *12*, e1005970. [[CrossRef](#)]
114. Leung, D.; Du, T.; Wagner, U.; Xie, W.; Lee, A.Y.; Goyal, P.; Li, Y.; Szulwach, K.E.; Jin, P.; Lorincz, M.C.; et al. Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6690–6695. [[CrossRef](#)]
115. Arney, K.L.; Bao, S.; Bannister, A.J.; Kouzarides, T.; Surani, M.A. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int. J. Dev. Biol.* **2002**, *46*, 317–320.
116. Probst, A.V.; Santos, F.; Reik, W.; Almouzni, G.; Dean, W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **2007**, *116*, 403–415. [[CrossRef](#)]
117. Zuo, X.; Sheng, J.; Lau, H.T.; McDonald, C.M.; Andrade, M.; Cullen, D.E.; Bell, F.T.; Iacovino, M.; Kyba, M.; Xu, G.; et al. Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. *J. Biol. Chem.* **2012**, *287*, 2107–2118. [[CrossRef](#)]
118. Cirio, M.C.; Ratnam, S.; Ding, F.; Reinhart, B.; Navara, C.; Chaillet, J.R. Preimplantation expression of the somatic form of Dnmt1 suggests a role in the inheritance of genomic imprints. *BMC Dev. Biol.* **2008**, *8*, 9. [[CrossRef](#)] [[PubMed](#)]
119. Kurihara, Y.; Kawamura, Y.; Uchijima, Y.; Amamo, T.; Kobayashi, H.; Asano, T.; Kurihara, H. Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Dev. Biol.* **2008**, *313*, 335–346. [[CrossRef](#)] [[PubMed](#)]
120. Hirasawa, R.; Chiba, H.; Kaneda, M.; Tajima, S.; Li, E.; Jaenisch, R.; Sasaki, H. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* **2008**, *22*, 1607–1616. [[CrossRef](#)] [[PubMed](#)]
121. Eid, A.; Rodriguez-Terrones, D.; Burton, A.; Torres-Padilla, M.E. SUV4-20 activity in the preimplantation mouse embryo controls timely replication. *Genes Dev.* **2016**, *30*, 2513–2526. [[CrossRef](#)]
122. Wongtawan, T.; Taylor, J.E.; Lawson, K.A.; Wilmut, I.; Pennings, S. Histone H4K20me3 and HP1 α are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J. Cell Sci.* **2011**, *124*, 1878–1890. [[CrossRef](#)]
123. Smith, Z.D.; Chan, M.M.; Humm, K.C.; Karnik, R.; Mekhoubad, S.; Regev, A.; Eggan, K.; Meissner, A. DNA methylation dynamics of the human preimplantation embryo. *Nature* **2014**, *511*, 611–615. [[CrossRef](#)]
124. Strogantsev, R.; Krueger, F.; Yamazawa, K.; Shi, H.; Gould, P.; Goldman-Roberts, M.; McEwen, K.; Sun, B.; Pedersen, R.; Ferguson-Smith, A.C. Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. *Genome Biol.* **2015**, *16*, 112. [[CrossRef](#)]
125. Pannetier, M.; Julien, E.; Schotta, G.; Tardat, M.; Sardet, C.; Jenuwein, T.; Feil, R. PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. *EMBO Rep.* **2008**, *9*, 998–1005. [[CrossRef](#)]
126. Regha, K.; Sloane, M.A.; Huang, R.; Pauler, F.M.; Warczok, K.E.; Melikant, B.; Radolf, M.; Martens, J.H.; Schotta, G.; Jenuwein, T.; et al. Active and repressive chromatin are interspersed without spreading in an imprinted gene cluster in the mammalian genome. *Mol. Cell* **2007**, *27*, 353–366. [[CrossRef](#)]
127. Schotta, G.; Lachner, M.; Sarma, K.; Ebert, A.; Sengupta, R.; Reuter, G.; Reinberg, D.; Jenuwein, T. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **2004**, *18*, 1251–1262. [[CrossRef](#)]
128. Singh, P.B.; Shloma, V.V.; Belyakin, S.N. Maternal regulation of chromosomal imprinting in animals. *Chromosoma* **2019**, *128*, 69–80. [[CrossRef](#)]
129. Zeng, T.B.; Han, L.; Pierce, N.; Pfeifer, G.P.; Szabó, P.E. EHMT2 and SETDB1 protect the maternal pronucleus from 5mC oxidation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 10834–10841. [[CrossRef](#)]
130. Smallwood, A.; Estève, P.O.; Pradhan, S.; Carey, M. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev.* **2007**, *21*, 1169–1178. [[CrossRef](#)] [[PubMed](#)]
131. Wolpert, L. *The Triumph of the Embryo*; Oxford University Press: Oxford, UK, 1991; ISBN 0-19-854243-7.

132. Nicetto, D.; Donahue, G.; Jain, T.; Peng, T.; Sidoli, S.; Sheng, L.; Montavon, T.; Becker, J.S.; Grindheim, J.M.; Blahnik, K.; et al. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. *Science* **2019**, *363*, 294–297. [[CrossRef](#)] [[PubMed](#)]
133. Haeckel, E. *Anthropogenie oder Entwicklungsgeschichte des Menschen*, 4th ed.; Engelmann: Leipzig, Germany, 1891.
134. Hopwood, N. *Haeckel's Embryos: Images, Evolution and Fraud*; The University of Chicago Press: Chicago, IL, USA, 2015; ISBN 978-0-226-04694-5.
135. Sander, K. The evolution of patterning mechanisms: Gleanings from insect embryogenesis and spermatogenesis. In *Development and Evolution*; Goodwin, B.C., Holder, N., Wylie, C.G., Eds.; Cambridge University Press: Cambridge, UK, 1983; pp. 137–159.
136. Richardson, M.K. Heterochrony and the phylotypic period. *Dev. Biol.* **1995**, *172*, 412–421. [[CrossRef](#)] [[PubMed](#)]
137. Richardson, M.K. A phylotypic stage for all animals? *Dev. Cell* **2012**, *22*, 903–904. [[CrossRef](#)] [[PubMed](#)]
138. Duboule, D. Temporal colinearity and the phylotypic progression: A basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev. Suppl.* **1994**, 135–142.
139. Raff, R.A. Developmental mechanisms in the evolution of animal form: Origins and evolvability of body plans. In *Early Life on Earth*; Bengtson, S., Ed.; Columbia: New York, NY, USA, 1994; pp. 489–500.
140. Irie, N.; Kuratani, S. The developmental hourglass model: A predictor of the basic body plan? *Development* **2014**, *141*, 4649–4655. [[CrossRef](#)] [[PubMed](#)]
141. Irie, N.; Kuratani, S. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat. Commun.* **2011**, *2*, 248. [[CrossRef](#)]
142. Slack, J.M.; Holland, P.W.; Graham, C.F. The zootype and the phylotypic stage. *Nature* **1993**, *361*, 490–492. [[CrossRef](#)]
143. Young, T.; Deschamps, J. Hox, Cdx, and anteroposterior patterning in the mouse embryo. *Curr. Top. Dev. Biol.* **2009**, *88*, 235–255. [[CrossRef](#)] [[PubMed](#)]
144. Wang, C.; Liu, X.; Gao, Y.; Yang, L.; Li, C.; Liu, W.; Chen, C.; Kou, X.; Zhao, Y.; Chen, J.; et al. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.* **2018**, *20*, 620–631. [[CrossRef](#)] [[PubMed](#)]
145. Cheloufi, S.; Elling, U.; Hopfgartner, B.; Jung, Y.L.; Murn, J.; Ninova, M.; Hubmann, M.; Badeaux, A.I.; Euong Ang, C.; Tenen, D.; et al. The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* **2015**, *528*, 218–224. [[CrossRef](#)] [[PubMed](#)]
146. Borkent, M.; Bennett, B.D.; Lackford, B.; Bar-Nur, O.; Brumbaugh, J.; Wang, L.; Du, Y.; Fargo, D.C.; Apostolou, E.; Cheloufi, S.; et al. A serial shRNA screen for roadblocks to reprogramming identifies the protein modifier SUMO2. *Stem Cell Rep.* **2016**, *6*, 704–716. [[CrossRef](#)]
147. Noordermeer, D.; Duboule, D. Chromatin looping and organization at developmentally regulated gene loci. *Wiley Interdiscip. Rev. Dev. Biol.* **2013**, *2*, 615–630. [[CrossRef](#)]
148. Illingworth, R.S.; Botting, C.H.; Grimes, G.R.; Bickmore, W.A.; Eskeland, R. PRC1 and PRC2 are not required for targeting of H2A.Z to developmental genes in embryonic stem cells. *PLoS ONE* **2012**, *7*, e34848. [[CrossRef](#)]
149. Zheng, H.; Huang, B.; Zhang, B.; Xiang, Y.; Du, Z.; Xu, Q.; Li, Y.; Wang, Q.; Ma, J.; Peng, X.; et al. Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol. Cell* **2016**, *63*, 1066–1079. [[CrossRef](#)]
150. Noordermeer, D.; Leleu, M.; Splinter, E.; Rougemont, J.; De Laat, W.; Duboule, D. The dynamic architecture of Hox gene clusters. *Science* **2011**, *334*, 222–225. [[CrossRef](#)]
151. Neijts, R.; Simmini, S.; Giuliani, F.; van Rooijen, C.; Deschamps, J. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Dev. Dyn.* **2014**, *243*, 88–98. [[CrossRef](#)]
152. Deschamps, J.; Duboule, D. Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* **2017**, *31*, 1406–1416. [[CrossRef](#)]
153. Soshnikova, N.; Duboule, D. Epigenetic temporal control of mouse Hox genes in vivo. *Science* **2009**, *324*, 1320–1323. [[CrossRef](#)] [[PubMed](#)]
154. Vieux-Rochas, M.; Fabre, P.J.; Leleu, M.; Duboule, D.; Noordermeer, D. Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 4672–4677. [[CrossRef](#)] [[PubMed](#)]

155. Burke, A.C.; Nelson, C.E.; Morgan, B.A.; Tabin, C. Hox genes and the evolution of vertebrate axial morphology. *Development* **1995**, *121*, 333–346. [[PubMed](#)]
156. Ku, M.; Koche, R.P.; Rheinbay, E.; Mendenhall, E.M.; Endoh, M.; Mikkelsen, T.S.; Presser, A.; Nusbaum, C.; Xie, X.; Chi, A.S.; et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **2008**, *4*, e1000242. [[CrossRef](#)] [[PubMed](#)]
157. Roadmap Epigenomics Consortium Roadmap; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317. [[CrossRef](#)] [[PubMed](#)]
158. Molitor, J.; Mallm, J.-P.; Rippe, K.; Erdel, F. Retrieving chromatin patterns from deep sequencing data using correlation functions. *Biophys. J.* **2017**, *112*, 473–490. [[CrossRef](#)] [[PubMed](#)]
159. Matsen, M.W. Self-consistent Field Theory and Its Applications. *Soft Matter* **2007**, *1*, 87–178. [[CrossRef](#)]
160. Riess, G. Micellization of block copolymers. *Prog. Polym. Sci.* **2003**, *28*, 1107–1170. [[CrossRef](#)]
161. Bates, F.S.; Fredrickson, G.H. Block copolymer thermodynamics: Theory and experiment. *Annu. Rev. Phys. Chem.* **1990**, *41*, 525–557. [[CrossRef](#)]
162. Flory, P.J. Thermodynamics of High Polymer Solutions. *J. Chem. Phys.* **1941**, *9*, 660. [[CrossRef](#)]
163. Huggins, M.L. Solutions of Long Chain Compounds. *J. Chem. Phys.* **1941**, *9*, 440. [[CrossRef](#)]
164. Nuebler, J.; Fudenberg, G.; Imakaev, M.; Abdennur, N.; Mirny, L.A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. USA* **2017**, *115*, E6697–E6706. [[CrossRef](#)] [[PubMed](#)]
165. Mirny, L.A.; Imakaev, M.; Abdennur, N. Two major mechanisms of chromosome organization. *Curr. Opin. Cell Biol.* **2019**, *58*, 142–152. [[CrossRef](#)] [[PubMed](#)]
166. Bonev, B.; Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **2016**, *17*, 661–678. [[CrossRef](#)]
167. Vian, L.; Pękowska, A.; Rao, S.S.P.; Kieffer-Kwon, K.R.; Jung, S.; Baranello, L.; Huang, S.C.; El Khattabi, L.; Dose, M.; Pruett, N.; et al. The energetics and physiological impact of cohesin extrusion. *Cell* **2018**, *173*, 1165–1178. [[CrossRef](#)] [[PubMed](#)]
168. Rao, S.S.P.; Huang, S.C.; Glenn St Hilaire, B.; Engreitz, J.M.; Perez, E.M.; Kieffer-Kwon, K.R.; Sanborn, A.L.; Johnstone, S.E.; Bascom, G.D.; Bochkov, I.D.; et al. Cohesin Loss Eliminates All Loop Domains. *Cell* **2017**, *171*, 305–320.e24. [[CrossRef](#)]
169. Parmar, J.J.; Woringer, M.; Zimmer, C. How the genome Folds: The Biophysics of Four-Dimensional Chromatic Organisation. *Ann. Rev. Biophys.* **2019**, *48*, 231–253. [[CrossRef](#)]
170. Wutz, G.; Várnai, C.; Nagasaka, K.; Cisneros, D.A.; Stocsits, R.R.; Tang, W.; Schoenfelder, S.; Jessberger, G.; Muhar, M.; Hossain, M.J.; et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **2017**, *36*, 3573–3599. [[CrossRef](#)]
171. Zhou, L.; Liang, C.; Chen, Q.; Zhang, Z.; Zhang, B.; Yan, H.; Qi, F.; Zhang, M.; Yi, Q.; Guan, Y.; et al. Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell* **2010**, *143*, 737–749. [[CrossRef](#)]
172. Ouyang, Z.; Zheng, G.; Tomchick, D.R.; Luo, X.; Yu, H. Structural Basis and IP6 Requirement for Pds5-Dependent Cohesin Dynamics. *Mol. Cell* **2016**, *62*, 248–259. [[CrossRef](#)]
173. Buheitel, J.; Stemmann, O. Prophase pathway-dependent removal of cohesin from human chromosomes requires opening of the Smc3-Scc1 gate. *EMBO J.* **2013**, *32*, 666–676. [[CrossRef](#)]
174. Eichinger, C.S.; Kurze, A.; Oliveira, R.A.; Nasmyth, K. Disengaging the Smc3/kleisin interface releases cohesin from Drosophila chromosomes during interphase and mitosis. *EMBO J.* **2013**, *32*, 656–665. [[CrossRef](#)] [[PubMed](#)]
175. Gandhi, R.; Gillespie, P.J.; Hirano, T. Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr. Biol.* **2006**, *16*, 2406–2417. [[CrossRef](#)] [[PubMed](#)]
176. Haarhuis, J.H.; Elbatsh, A.M.; van den Broek, B.; Camps, D.; Erkan, H.; Jalink, K.; Medema, R.H.; Rowland, B.D. WAPL-mediated removal of cohesin protects against segregation errors and aneuploidy. *Curr. Biol.* **2013**, *23*, 2071–2077. [[CrossRef](#)] [[PubMed](#)]
177. Kueng, S.; Hegemann, B.; Peters, B.H.; Lipp, J.J.; Schleiffer, A.; Mechtler, K.; Peters, J.M. Wapl controls the dynamic association of cohesin with chromatin. *Cell* **2006**, *127*, 955–967. [[CrossRef](#)]
178. Folco, H.D.; McCue, A.; Balachandran, V.; Grewal, S.I.S. Cohesin Impedes Heterochromatin Assembly in Fission Yeast Cells Lacking Pds5. *Genetics* **2019**, *213*, 127–141. [[CrossRef](#)] [[PubMed](#)]

179. Rudra, S.; Skibbens, R.V. Cohesin codes—Interpreting chromatin architecture and the many facets of cohesin function. *J. Cell Sci.* **2013**, *126*, 31–41. [[CrossRef](#)] [[PubMed](#)]
180. Lin, S.J.; Tapia-Alveal, C.; Jabado, O.J.; Germain, D.; O’Connell, M.J. An acetyltransferase-independent function of Eso1 regulates centromere cohesion. *Mol. Biol. Cell* **2016**, *27*, 4002–4010. [[CrossRef](#)]
181. Hamley, I.W. *The Physics of Block Copolymers*; Oxford University Press: Oxford, UK, 1998.
182. Bates, F.S.; Frederickson, G.H. Block Copolymers—Designer Soft Materials. *Physics Today* **1999**, *52*, 32. [[CrossRef](#)]
183. Bates, F.S. Polymer-polymer phase behavior. *Science* **1991**, *251*, 898–905. [[CrossRef](#)]
184. Chua, E.Y.; Vogirala, V.K.; Inian, O.; Wong, A.S.; Nordenskiöld, L.; Plitzko, J.M.; Danev, R.; Sandin, S. 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* **2016**, *44*, 8013–8019. [[CrossRef](#)] [[PubMed](#)]
185. Maeshima, K.; Imai, R.; Tamura, S.; Nozaki, T. Chromatin as dynamic 10-nm fibers. *Chromosoma* **2014**, *123*, 225–237. [[CrossRef](#)] [[PubMed](#)]
186. De Gennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, USA, 1979.
187. Grosberg, A.; Khokhlov, A.R. *Giant Molecules: Here, There, and Everywhere*; Academic Press: New York, NY, USA, 1997; ISBN1 0123041309, ISBN2 9780123041302.
188. Azzaz, A.M.; Vitalini, M.W.; Thomas, A.S.; Price, J.P.; Blacketer, M.J.; Cryderman, D.E.; Zirbel, L.N.; Woodcock, C.L.; Elcock, A.H.; Wallrath, L.L.; et al. Human heterochromatin protein 1 α promotes nucleosome associations that drive chromatin condensation. *J. Biol. Chem.* **2014**, *289*, 6850–6861. [[CrossRef](#)] [[PubMed](#)]
189. Machida, S.; Takizawa, Y.; Ishimaru, M.; Sugita, Y.; Sekine, S.; Nakayama, J.; Wolf, M.; Kurumizaka, H. Structural basis of heterochromatin formation by human HP1. *Mol. Cell* **2018**, *69*, 385–397. [[CrossRef](#)]
190. Belaghzal, H.; Borrman, T.; Stephens, A.D.; Lafontaine, D.L.; Venev, S.V.; Weng, Z.; Marko, J.F.; Dekker, J. Compartment-dependent chromatin interaction dynamics revealed by liquid chromatin Hi-C. *bioRxiv* **2019**, 704957. [[CrossRef](#)]
191. Barbieri, M.; Chotalia, M.; Fraser, J.; Lavitas, L.M.; Dostie, J.; Pombo, A.; Nicodemi, M. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16173–16178. [[CrossRef](#)]
192. Nicodemi, M.; Pombo, A. Models of chromosome structure. *Curr. Opin. Cell Biol.* **2014**, *28*, 90–95. [[CrossRef](#)]
193. MacPherson, Q.; Beltran, B.; Spakowitz, A.J. Bottom-up modeling of chromatin segregation due to epigenetic modifications. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 12739–12744. [[CrossRef](#)]
194. Grau, D.J.; Chapman, B.A.; Garlick, J.D.; Borowsky, M.; Francis, N.J.; Kingston, R.E. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev.* **2011**, *25*, 2210–2221. [[CrossRef](#)]
195. Kim, J.; Kingston, R.E. The CBX family of proteins in transcriptional repression and memory. *J. Biosci.* **2020**, *45*, 16. [[CrossRef](#)]
196. Krietenstein, N.; Rando, O.J. Mesoscale organization of the chromatin fiber. *Curr. Opin. Genet. Dev.* **2020**, *61*, 32–36. [[CrossRef](#)] [[PubMed](#)]
197. Ricci, M.A.; Manzo, C.; García-Parajo, M.F.; Lakadamyali, M.; Cosma, M.P. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **2015**, *160*, 1145–1158. [[CrossRef](#)] [[PubMed](#)]
198. Ou, H.D.; Phan, S.; Deerinck, T.J.; Thor, A.; Ellisman, M.H.; O’Shea, C.C. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **2017**, *357*, eaag0025. [[CrossRef](#)] [[PubMed](#)]
199. Chen, C.; Lim, H.H.; Shi, J.; Tamura, S.; Maeshima, K.; Surana, U.; Gan, L. Budding yeast chromatin is dispersed in a crowded nucleoplasm in vivo. *Mol. Biol. Cell* **2016**, *27*, 3357–3368. [[CrossRef](#)]
200. Dubochet, J.; Adrian, M.; Schultz, P.; Oudet, P. Cryo-electron microscopy of vitrified SV40 minichromosomes: The liquid drop model. *EMBO J.* **1986**, *5*, 519–528. [[CrossRef](#)]
201. Eltsov, M.; MacLellan, K.M.; Maeshima, K.; Frangakis, A.S.; Dubochet, J. Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19732–19737. [[CrossRef](#)]
202. Risca, V.I.; Denny, S.K.; Straight, A.F.; Greenleaf, W.J. Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping. *Nature* **2017**, *541*, 237. [[CrossRef](#)]
203. Rydberg, B.; Holley, W.R.; Mian, I.S.; Chatterjee, A. Chromatin conformation in living cells: Support for a zig-zag model of the 30 nm chromatin fiber. *J. Mol. Biol.* **1998**, *284*, 71–84. [[CrossRef](#)]

204. Grigoryev, S.A.; Arya, G.; Correll, S.; Woodcock, C.L.; Schlick, T. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13317–13322. [[CrossRef](#)]
205. Schalch, T.; Duda, S.; Sargent, D.F.; Richmond, T.J. X-ray structure of a tetranucleosome and its implications for the chromatin fiber. *Nature* **2005**, *436*, 138–141. [[CrossRef](#)]
206. Ekundayo, B.; Richmond, T.J.; Schalch, T. Capturing Structural Heterogeneity in Chromatin Fibers. *J. Mol. Biol.* **2017**, *429*, 3031–3042. [[CrossRef](#)] [[PubMed](#)]
207. Song, F.; Chen, P.; Sun, D.; Wang, M.; Dong, L.; Liang, D.; Xu, R.M.; Zhu, P.; Li, G. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **2014**, *344*, 376–380. [[CrossRef](#)]
208. Lohr, D.; Van Holde, K.E. Organization of spacer DNA in chromatin. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 6326–6330. [[CrossRef](#)] [[PubMed](#)]
209. Norouzi, D.; Zhurkin, V.B. Topological polymorphism of the two-start chromatin fiber. *Biophys. J.* **2015**, *108*, 2591–2600. [[CrossRef](#)]
210. Gibson, B.A.; Doolittle, L.K.; Schneider, M.W.G.; Jensen, L.E.; Gamarra, N.; Henry, L.; Gerlich, D.W.; Redding, S.; Rosen, M.K. Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **2019**, *179*, 470–484.e21. [[CrossRef](#)] [[PubMed](#)]
211. Brasher, S.V.; Smith, B.O.; Fogh, R.H.; Nietlispach, D.; Thiru, A.; Nielsen, P.R.; Broadhurst, R.W.; Ball, L.J.; Murzina, N.V.; Laue, E.D. The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J.* **2000**, *19*, 1587–1597. [[CrossRef](#)] [[PubMed](#)]
212. Mulligan, P.J.; Koslover, E.F.; Spakowitz, A.J. Thermodynamic model of heterochromatin formation through epigenetic regulation. *J. Phys. Condens Matter* **2015**, *27*, 064109. [[CrossRef](#)]
213. Saksouk, N.; Simboeck, E.; Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **2015**, *8*, 3. [[CrossRef](#)]
214. Teif, V.B.; Kepper, N.; Yserentant, K.; Wedemann, G.; Rippe, K. Affinity, stoichiometry and cooperativity of heterochromatin protein 1 (HP1) binding to nucleosomal arrays. *J. Phys. Condens Matter* **2015**, *27*, 064110. [[CrossRef](#)]
215. Al-Sady, B.; Madhani, H.D.; Narlikar, G.J. Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Mol. Cell* **2013**, *51*, 80–91. [[CrossRef](#)]
216. Canzio, D.; Chang, E.Y.; Shankar, S.; Kuchenbecker, K.M.; Simon, M.D.; Madhani, H.D.; Narlikar, G.J.; Al-Sady, B. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell* **2011**, *41*, 67–81. [[CrossRef](#)] [[PubMed](#)]
217. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **2011**, *21*, 381–395. [[CrossRef](#)]
218. Reinberg, D.; Vales, L.D. Chromatin domains rich in inheritance. *Science* **2018**, *361*, 33–34. [[CrossRef](#)]
219. Leibler, L. Theory of Microphase Separation in Block Copolymers. *Macromolecules* **1980**, *13*, 1602. [[CrossRef](#)]
220. Edwards, S.F. The statistical mechanics of polymers with excluded volume. *Proc. Phys. Soc.* **1965**, *85*, 613–624. [[CrossRef](#)]
221. MacPherson, Q.; Beltran, B.; Spakowitz, A.J. Chromatin Compaction Leads to a Preference for Peripheral Heterochromatin. *Biophys. J.* **2020**, *118*, 1479–1488. [[CrossRef](#)]
222. Mau, S.; MacPherson, Q.; Spakowitz, A.J. Polymer Semiflexibility Induces Nonuniversal Phase Transitions in Diblock Copolymers. *Phys. Rev. Lett.* **2018**, *120*, 067802–067805. [[CrossRef](#)]
223. Shin, Y.; Chang, Y.-C.; Lee, D.S.; Berry, J.; Sanders, D.W.; Ronceray, P.; Wingreen, N.S.; Haataja, M.; Brangwynne, C.P. Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* **2018**, *175*, 1481–1491.e13. [[CrossRef](#)]
224. Plys, A.J.; Davis, C.P.; Kim, J.; Rizki, G.; Keenen, M.M.; Marr, S.K.; Kingston, R.E. Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev.* **2019**, *33*, 799–813. [[CrossRef](#)] [[PubMed](#)]
225. Tatavosian, R.; Kent, S.; Brown, K.; Yao, T.; Duc, H.N.; Huynh, T.N.; Zhen, C.Y.; Ma, B.; Wang, H.; Ren, X. Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J. Biol. Chem.* **2019**, *294*, 1451–1463. [[CrossRef](#)]
226. Kaustov, L.; Ouyang, H.; Amaya, M.; Lemak, A.; Nady, N.; Duan, S.; Wasney, G.A.; Li, Z.; Vedadi, M.; Schapira, M.; et al. Recognition and specificity determinants of the human cbx chromodomains. *J. Biol. Chem.* **2011**, *286*, 521–529. [[CrossRef](#)] [[PubMed](#)]

227. Allshire, R.C.; Madhani, H.D. Ten principles of heterochromatin formation and function. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 229. [[CrossRef](#)] [[PubMed](#)]
228. Hall, I.M.; Shankaranarayana, G.D.; Noma, K.; Ayoub, N.; Cohen, A.; Grewal, S.I. Establishment and maintenance of a heterochromatin domain. *Science* **2002**, *297*, 2232–2237. [[CrossRef](#)] [[PubMed](#)]
229. Imbeault, M.; Helleboid, P.Y.; Trono, D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **2017**, *543*, 550–554. [[CrossRef](#)]
230. Helleboid, P.Y.; Heuse, I.M.; Duc, J.; Piot, C.; Thorball, C.W.; Coluccio, A.; Pontis, J.; Imbeault, M.; Turelli, P.; Aebersold, R.; et al. The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. *EMBO J.* **2019**, *38*, e101220. [[CrossRef](#)]
231. Yang, P.; Wang, Y.; Macfarlan, T.S. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends Genet.* **2017**, *33*, 871–881. [[CrossRef](#)]
232. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **2013**, *14*, R115. [[CrossRef](#)]
233. Bestor, T.H.; Edwards, J.R.; Boulard, M. Notes on the role of dynamic DNA methylation in mammalian development. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6796–6799. [[CrossRef](#)]
234. Reddington, J.P.; Perricone, S.M.; Nestor, C.E.; Reichmann, J.; Youngson, N.A.; Suzuki, M.; Reinhardt, D.; Dunican, D.S.; Prendergast, J.G.; Mjoseng, H.; et al. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol.* **2013**, *14*, R25. [[CrossRef](#)]
235. Farcas, A.M.; Blackledge, N.P.; Sudbery, I.; Long, H.K.; McGouran, J.F.; Rose, N.R.; Lee, S.; Sims, D.; Cerase, A.; Sheahan, T.W.; et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife* **2012**, *1*, e00205. [[CrossRef](#)]
236. Riising, E.M.; Comet, I.; Leblanc, B.; Wu, X.; Johansen, J.V.; Helin, K. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol. Cell* **2014**, *55*, 347–360. [[CrossRef](#)] [[PubMed](#)]
237. Blackledge, N.P.; Farcas, A.M.; Kondo, T.; King, H.W.; McGouran, J.F.; Hanssen, L.L.P.; Ito, S.; Cooper, S.; Kondo, K.; Koseki, Y.; et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **2014**, *157*, 1445–1459. [[CrossRef](#)] [[PubMed](#)]
238. McLaughlin, K.; Flyamer, I.M.; Thomson, J.P.; Mjoseng, H.K.; Shukla, R.; Williamson, I.; Grimes, G.R.; Illingworth, R.S.; Adams, I.R.; Pennings, S.; et al. DNA Methylation Directs Polycomb-Dependent 3D Genome Re-organization in Naive Pluripotency. *Cell Rep.* **2019**, *29*, 1974–1985.e6. [[CrossRef](#)] [[PubMed](#)]
239. Maeshima, K.; Hihara, S.; Eltsov, M. Chromatin structure: Does the 30-nm fiber exist in vivo? *Curr. Opin. Cell Biol.* **2010**, *22*, 291–297. [[CrossRef](#)]
240. Maeshima, K.; Tamura, S.; Hansen, J.C.; Itoh, Y. Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Curr. Opin. Cell Biol.* **2020**, *64*, 77–89. [[CrossRef](#)] [[PubMed](#)]
241. Ruthenburg, A.J.; Li, H.; Patel, D.J.; Allis, C.D. Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 983. [[CrossRef](#)]
242. Wang, Z.; Patel, D.J. Combinatorial readout of dual histone modifications by paired chromatin-associated modules. *J. Biol. Chem.* **2011**, *286*, 18363–18368. [[CrossRef](#)]
243. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
244. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoutte, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [[CrossRef](#)]
245. Ramírez, F.; Ryan, D.P.; Grüning, B.; Bhardwaj, V.; Kilpert, F.; Richter, A.S.; Heyne, S.; Dündar, F.; Manke, T. deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **2016**, *44*, W160–W165. [[CrossRef](#)]
246. Amemiya, H.M.; Kundaje, A.; Boyle, A.P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci. Rep.* **2019**, *9*, 9354. [[CrossRef](#)] [[PubMed](#)]
247. Ernst, J.; Kellis, M. ChromHMM: Automating chromatin-state discovery and characterization. *Nat. Methods* **2012**, *9*, 215–216. [[CrossRef](#)] [[PubMed](#)]

