10. Chromatin remodeling and epigenetic regulation during development

Mayra Furlan-Margaril and Félix Recillas-Targa
Instituto de Fisiología Celular, Departamento de Genética Molecular
Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510, Mexico

Most of the cells that integrate a multicellular organism have the same “DNA genomic set”. Differences between cells arise very early in development due to a variety of processes that control which genes will be expressed in determined cells contributing to cell fate determination and maintenance. These processes gather in the term “epigenetics” and represent the first control of how does a genotype becomes a specific phenotype through the epigenome. Eukaryotic genomes reside inside the cell nucleus in a complex protein-DNA conformation called chromatin. All genomic metabolism takes place immerse in a chromatin environment and the molecular machinery that participates has evolved in tight relation with it. In recent years, this structure has shown to represent a primordial level of genome function regulation. On the other hand, the incorporation of epigenetics to the understanding of human diseases, as cancer, has opened a whole new analysis scheme. It has been recently postulated that the epigenotype represents a second code that is much more sophisticated and complex than the genetic code. The deciphering of such second code will certainly contribute to better design early diagnostic protocols and improve the prognostic, monitoring, treatment, and even prevention of cancer.

Correspondence/Reprint request: Dr. Félix Recillas-Targa, Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México, D.F. 04510, Mexico. Email: frecilla@ifc.unam.mx
Epigenetics have now turned into an essential framework to look at a myriad of processes during organism development as cellular plasticity, imprinting, differentiation, self-renewal, ageing and its abnormal counterparts.

**Introduction**

Organism development involves a sophisticated network of genetic information in combination with various and interdependent epigenetic events. Genetics and epigenetics are constantly interplaying in order to achieve proper expression programs in the cell and errors in any of these components could compromise cellular function (Fig 1).

In addition to its requirement for the assembly and packaging of the genome inside the nucleus, chromatin structure is a central component for several nuclear processes as DNA replication, DNA repair and recombination among others. The interplay between different epigenetic regulators such as the covalent modification of histones and ATP-dependent chromatin remodelers, are important in controlling gene expression and many other biological processes. The amino terminal domain of histones are exposed on the nucleosomal surface and are rich in post-translational modifications such as serine and threonine phosphorylation, lysine and arginine acetylation and deacetylation, methylation and demethylation of the same residues, lysine ubiquitination and sumoylation and ADP-ribosylation [1, 2, 3]. Additional complexity is added to the system when we consider that methylation of histones can be found in mono-, di- and tri-methylated forms, each of which is generated by a specific histone methyltransferase that can be removed in a regulated way by specific demethylases [3]. Furthermore, and with the exception of histone H4, all core histones have variants which in some cases differ in surprisingly few amino acids like the H3.3 histone variant which is associated to transcription sites and more recently, to nucleosomal dissociation [4, 5].

Several interdependent models have been proposed by Allis and collaborators, starting with the “histone code” referred to the combinatorial covalent modifications of histone tails, followed by the “cross-talk” model to explain the spreading of epigenetic modifications among poly-nucleosomes and the “barcode hypothesis” that incorporates histone variants to define different chromosomal domains and try to explain the concept of epigenetic memory [6, 7, 8]. Now the relevant question has turned into the mechanisms and targets of all these processes during early development.

Historically, transformation of a normal cell to a cancer cell involves multiple events where genetic defects emerge [9]. Those genetic
abnormalities can be grouped as gene mutations, loss of heterozygosity, translocations, deletions, recombination, and others. Knudson hypothesis based on the “two-hits”, explaining the origin and progression of retinoblastoma, attracted our attention on the genetic predisposition to certain types of cancers [10, 11]. Nowadays, the origins and progression of cancer cannot be understood exclusively from the genetic perspective. Epigenetics are coming of age supported by a large and novel amount of knowledge directly associated with chromatin structure.

The classical view that both alleles must be genetically inactivated to allow tumor progression is now evolving based on epigenetic silencing mechanisms, in particular of tumour suppressor gene expression [12, 13]. Bird defined epigenetic regulation as the mitotically and/or meiotically heritable changes in gene expression that cannot be explained by changes in DNA sequence [14]. This type of regulation incorporated an additional concept that lead to the notion that regulatory signals were transmitted from one cell generation to another and that represents a sort of “memory” that should be maintained post-mitotically. The “epigenetic memory” concept represents one of the most attractive and challenging research topics in the field and its understanding will certainly contribute to learn more about cell transformation and tumour progression with a much better position to design and develop novel therapeutic strategies.

Figure 1. Genetic and epigenetic elements regulate genome function and deregulation of both processes could lead to cancer.
Chromatin structure is intimately involved in epigenetic regulation and its understanding in terms of positive and negative modulation of gene expression will provide a good standpoint to address the origins and progression of cancer. Epigenetic regulatory elements can be grouped in different categories that include, DNA modifications like DNA methylation, chromatin structure modulation, non-coding RNA or intergenic transcripts and nuclear dynamics. All of them contribute to propagate genetic information in tight relationship with epigenetic regulatory mechanisms.

In the present chapter we will first describe the epigenetic mechanisms that control gene expression mainly based on chromatin remodeling process. The second part will analyze the role of epigenetics in early development and germ line specification. Finally, we discuss how epigenetic processes are involved in ageing and human diseases.

General mechanisms of epigenetic regulation

The central aspect of epigenetic regulation is undoubtedly related to eukaryote genomic organization into chromatin. Chromatin structure has evolved to allow large genomes to be incorporated in the cellular nuclei causing a natural repressive environment for transcription activity and the subsequent transmission of genetic traits. To allow regulated gene expression, such highly compacted chromatin structure needs to be coordinately remodeled [15]. Chromatin is defined as a set of DNA-protein and protein-protein interaction in which the fundamental unit is the nucleosome where the DNA is wrapped around encompassing 147 base pairs (bp) of genomic DNA [16]. Nucleosome is organized into a histone octamer composed of two copies of histone H3, H4, H2A and H2B corresponding to the minimal level of genome compaction [17, 18]. Nucleosomes contribute to the higher-order chromatin organization of the genome through a highly sophisticated and not completely understood mechanism. Importantly, the 30 nm chromatin fiber also known as solenoid constitutes the physiological substrate for the great majority of the epigenetic regulatory events [17, 19]. Thus, based on all the progress in understanding how the genome is organized into chromatin and how such natural occurring structure influences gene expression, normal and aberrant epigenetic regulation needs to be taken seriously into account.

To comprehend the relationship between epigenetics and human diseases it is first desirable to learn how chromatin structure is modulated. Chromatin structure is remodeled at different interdependent levels that include: 1) one of the most studied epigenetic modifications that is DNA methylation, 2) histone covalent modifications, 3) the expanded field of histone variants, 4) members of the Polycomb and Trithorax family of chromatin regulators, 5)
the novel research field incorporating non-coding RNA’s and RNA interference (RNAi) pathway and 6) the contribution of the nuclear architecture and dynamics. Due to the relevance of these remodeling activities we will describe their function during development and their relationship with human diseases.

**DNA methylation as a central epigenetic process**

One of the central components of epigenetic regulation, based on repression processes, is DNA methylation [14, 20, 21]. In mammals, DNA methylation occurs at the carbon-5 position of the cytosine in the context of symmetrical CpG dinucleotides and is present in 70-80% of all CpGs in a human somatic cell [22]. Evolutionary, DNA methylation has been proposed to contribute to genome stability repressing the chromatin structure and avoiding the mobilization of retrotransposons, endogenous viruses or repetitive sequences [23]. Interestingly and to a certain point contradictory, unmethylated CpG sequences are found normally into the genome corresponding to the so-called CpG-islands [24]. It has been estimated that around 60% of human gene promoters correspond to unmethylated CpG-islands and those promoters correspond either to housekeeping or tissue-specific genes [24]. In particular, the great majority of tumour suppressor gene promoters and surrounding genomic sequences correspond to CpG-islands [24, 25].

By definition a CpG-island corresponds to a DNA sequence between 300 and 650 bp with an average content of G+C of 60 to 70%. More recently, CpG-rich and CpG-poor promoters have been described to be both sensitive to DNA methylation [24]. Observations from our laboratory have shown that the CpG-abundance does not always correlate with their quantitative and qualitative silencing effects (De la Rosa et al., 2007; Soto-Reyes and Recillas-Targa 2010). At this point it is relevant to point out that one of the most frequent abnormal epigenetic events in human diseases is the DNA hypermethylation of tumour suppressor gene promoters [25, 26, 27].

DNA methylation has also been critical in early developmental stages and genomic imprinting [20, 28, 29]. DNA methylation participates on reprogramming specific epigenetic patterns of gene expression during early mouse development, of both maternal and paternal genomes. Even though several aspects remain unsolved, after fertilization and until blastocyst stage there is an active genome demethylation event (Fig 2). After implantation there is an active *de novo* DNA methylation [30, 31, 32, 33]. Interestingly, imprinted genes are not subject to such epigenetic variations in terms of DNA methylation. The active demethylation process remains a mystery in terms of
the enzymes responsible and the mechanisms of demethylation. Until today no-demethylation enzymatic activity has been discovered. At the present time it is well established that early DNA methylation patterns require both de novo and maintenance DNA methyltransferase enzymatic activities (see below).

Several models have been proposed to explain DNA methylation variations during early development. Based on the “epigenetic memory” theory it has been suggested that demethylation allows a rapid erase of specific parental methylation patterns, of course with the exception of imprinted genes. Furthermore, such demethylation will favor open chromatin environments facilitating the expression of a large set of genes that could be grouped in genomic clusters [34]. Such scenario should facilitate early patterns of gene expression and cell lineage differentiation. A complementary function for DNA methylation has been associated with an intense wave of DNA methylation beginning at birth. It has been discussed that such strategy facilitates a fast and stable silencing, in a large genomic scale, of embryonic genes that are no longer required for the subsequent developmental stages of an organism (Fig 2) [35, 36].

**Histone code and cross-talk**

As previously mentioned, histones can be post-translationally modified at both the core and the amino terminal tail. The combinatory of covalent modifications give rise to different biological outputs that reside at the centre of chromatin-regulated processes and are still under intense investigation. Histone covalent modifications can alter chromatin structure in different ways. One occurs when certain modification (i.e. acetylation) alters the histone electrostatic charge hence changing the local charge of nucleosomes. Such modifications can result in the alteration of histone-DNA and histone-histone interactions changing the properties of the chromatin fiber. A different phenomenon takes place when different post-translational modification patterns act as platforms to recruit effector proteins that modify chromatin structure. This final property has lead to the “histone code” hypothesis, in which histone modifications profiles provide recognition flags to specific nuclear factors [6]. Later on such concept has been extended into a global landscape in which different chromatin domains in a chromosome are determined by the differential amounts of histone modification profiles so that the code becomes meaningful when looking at broader chromatin regions. The previous idea supports that there is a cross-talk between nucleosomes that contributes to the definition and regulation of specific genomic regions [37].

Over the last decade, different chromatin-modifying enzymes have been identified. Some of these conserved protein complexes are able to establish
histone modifications like lysine acetylation and methylations, arginine methylation and serine or threonine phosphorylation. Other protein complexes can remove such modifications contributing to chromatin status reprogramming. Therefore, regulation of chromatin-modifying enzymes is required in order to maintain a balanced state that allows fine-tuning of gene expression.

**ATP-dependent chromatin remodeling complexes**

Another mechanism to modify the intrinsically repressive chromatin status without covalent modifications is performed through the recruitment of “remodeling” complexes that use the energy from ATP hydrolysis to reorganize the chromatin fiber. Chromatin-remodeling enzymes can expose sequences facilitating the interaction of other proteins with nucleosomal DNA or stabilize repressive chromatin conformations [38]. Remodeling activities are necessary in different contexts such as transcription, replication, recombination, histone variants replacement and repair, and distinct complexes have been characterized linked to these processes.

Chromatin-remodeling enzymes can be categorized into different families depending on their ATPase subunit. The best characterized are the ISWI family that mobilizes nucleosomes along the DNA and the SWI/SNF family that alters nucleosome structure exposing DNA-histone contacts. Other complexes as Mi-2 and INO-80 have been linked with compact chromatin stabilization [39]. ATP-dependent remodeling complexes have several subunits besides the ATPase that allows them to carry out their function in a regulated and specific manner. One example is the mammalian NuRD complex in which a histone deacetylase (HDAC) forms part of the complex [40]. Another example comes from the *Drosophila* NURF complex that contains NURF301 subunit, which interacts with transcription factors [40].

Defects in chromatin-remodeling complexes have been implicated in cancer. For example, a loss of SNF5, a member of the human SWI/SNF complex, has been observed in pediatric cancer. In addition, the ATPase subunit BRM and BRG1 of the same complex, have been found mutated in a variety of cancer cell lines and primary tumours that are associated with a poorer prognosis in patients with non-small-cell lung cancer [41].

**Polycomb and Trithorax family of regulatory proteins**

Polycomb (PcG) and Trithorax (TrxG) group of proteins represent one of the main effectors that modify chromatin in response to cellular signals transforming developmental cues into epigenetic memory. These proteins were first identified in *Drosophila* were they participate in Hox gene cluster
regulation as their mutants cause homeotic transformations. Since then, PcG and TrxG complexes have been shown to be crucial regulators in essential cell processes as proliferation, cell identity (providing cellular memory) and lineage commitment. Additionally, PcG and TrxG proteins are involved in promoting chromatin transitions in response to a variety of signals as morphogens and growth factors.

PcG and TrxG proteins play antagonistic roles. PcG complexes are involved in maintaining a silenced chromatin state and TrxG proteins act preferentially propagating an open chromatin conformation that is transcriptionally permissive. This is achieved by a variety of proteins that are highly conserved between eukaryotes and perform diverse functions including most of the epigenetic mechanisms previously described. PcG genes encode products that include DNA-binding proteins (e.g. YY1), histone modifying enzymes (e.g. EZH2) and other chromatin associated factors with chromodomains that recognize repressive chromatin marks as H3K27me3 (e.g. PC). TrxG genes include products as ATP-dependent remodelers (e.g. Brahma), transcription factors (e.g. GAGA) and histone lysine methyl transferases (e.g. Ash1) [39]. These proteins are components of different complexes that stabilize either an open or closed chromatin conformation.

How do PcG/TrxG proteins are recruited to their chromatin target sites is still a prevailing question. In *Drosophila*, PcG proteins are localized to DNA elements named PREs (Polycomb Response Elements) located kilobases away from PcG target genes promoters. PREs are necessary in order to conserve the heritable chromatin repression exerted by PcG complexes. However, how do PcG proteins get to PREs and how do they repress their final target genes is currently unknown. Furthermore PRE sequences have not been found in mammalian organisms.

Current models propose PcG recruitment through the interaction with DNA-binding factors and further silencing stabilization through chromodomain containing proteins that associate with H3K27me3 modified nucleosomes [42]. However, evidences have emerged showing that PREs are nucleosome poor sequences and that PcG proteins can interact with unmodified histones in vitro [43, 44]. Thus, further research is needed to understand how these complexes are localized to their target genes and maintain their chromatin-mediated repression in different organisms.

Even less understood is the mechanism by which TrxG proteins are recruited to their final targets in the chromatin template. Some recent evidences have implicated a RNA-based pathway in order to initiate Ash1 (MLL in mammals) recruitment to their target promoters and intergenic transcription to preserve the permissive chromatin status [45].
Balance between heterochromatin and euchromatin is a key factor to conserve genome stability and cell functional integrity. Alterations in PcG and TrxG proteins have been reported in several developmental abnormalities and even cancer. For example, the increase in EZH2 and/or MLL is associated with increased risk of breast cancer, multiple myeloma or leukemia and prostate cancer [46, 47].

In summary, PcG and TrxG proteins are critical epigenetic modulators that participate in vital cell functions from keeping stem cell pluripotency, to specifying cell fates and balancing the chromatin status of the eukaryotic genome.

**Non-coding transcripts and RNA interference**

Eukaryotic and prokaryotic genomes are composed of DNA sequences some of which encode proteins and some of which does not. Eukaryotes have large abundance of non-coding sequences that are transcribed and perform diverse functions as RNA molecules (43% in the human genome) [48]. Besides transference RNA (tRNA) and ribosomal RNA (rRNA) there is a vast amount of non-coding transcripts that could be implicated in gene expression regulation. In humans, non-coding transcription generates the 98% of all cellular transcripts including intronic RNAs, intronic and exonic RNA from non-coding genes (as rRNA and tRNA) and intergenic RNAs [49].

Intergenic transcription was considered the result of an excess of RNA polymerases activity, however there is increasing amount of evidence that these transcripts have a fundamental role in gene expression regulation at both the chromatin state and the spatial organization of the genome inside the nucleus [50]. One of the most conserved examples of non-coding RNAs, are small RNAs. Two main types have been broadly studied, interference RNAs (RNAi) and microRNAs. Both of them are processed and interact with different machinery ultimately leading to gene silencing and chromatin compaction.

RNAi was first described as a conserved host defense mechanism that breaks down dsRNA species into small RNA molecules known as short interference RNA (siRNA). dsRNA is processed by the enzyme DICER and then, the enzymatic complex (RISC) (RNA Induced Silencing Complex) interacts with the siRNAs and degrades the corresponding mRNA. A different mechanism occurs when the siRNAs inhibit translation through post-transcriptional silencing (PTGS).

More recently, siRNA has been implicated in gene silencing at transcriptional level through heterochromatin formation from yeast to mammals and to be a central player in maintaining genome stability [51, 52]. The clearest picture of how intergenic RNAs can contribute to heterochromatin formation comes from studies in *Schizosaccharomyces pombe*. First evidences that linked
RNAi machinery and heterochromatin was that mutants in Ago1 (member of RITS complex) presented segregation abnormalities. Subsequently it was found that mutants in all Dcr1, Rdpl and Ago1 lead to defects in heterochromatin formation at centromeres and loss of H3K9me at the mating-type locus [53]. The proposed model is that intergenic transcription promotes dsRNA processing by DICER generating siRNAs that interact with RITS complex. siRNA-RITS complex recruits RDRC (a complex composed by an RNA dependent RNA polymerase and other proteins) to the nascent transcript. Silencing propagation seems to be mediated by Clr4 (Suv-3-9h) interacting with RNA Polymerase II (Pol II). This interaction promotes H3K9 methylation by Clr4 as transcription progresses. H3K9 methylation provokes Swi6 (HP1 in mammals) binding and hence heterochromatin formation [53]. siRNAs interact with other remodeling machineries promoting DNA methylation of repetitive sequences in telomeres and centromeres. In the case of *Tetrahymena* siRNAs are implicated in DNA elimination. After conjugation of two organisms a new macronucleus is formed and the old one is eliminated through a complex recruited by RNA molecules [54].

Several evidences have accumulated supporting the idea that intergenic transcripts are also involved in chromosome dynamics, chromatin domain organization, epigenetic memory and imprinting [53]. Long non-coding transcripts direct X inactivation in mammals and X activation in *Drosophila*. In both cases the non-coding RNA interacts with remodeling factors and histone modifying enzymes in order to generate heterochromatin and euchromatin, respectively. In addition non-coding transcription has been implicated in contributing to form active chromatin domains in vivo. Such is the case for the human β-globin gene locus in which intergenic transcripts all along the domain have been reported to play a critical role in preserving the open chromatin conformation of the domain when globin gene expression is needed [55].

In conclusion, the RNA molecule is widely involved in different epigenetic mechanisms that range from maintaining repetitive sequences silenced and compact to avoid incorrect recombination events, degrading unwanted mRNAs, and acting at the domain or chromosome level in order to promote an open or closed chromatin conformation.

**Cell nucleus and epigenetics**

For many years cell nucleus was thought to be a passive genome container. In the last decade, evidences from different groups have dramatically changed this notion and has positioned the cell nucleus architecture as a key epigenetic factor involved in almost all genomic
Chromatin remodeling and epigenetic regulation during development

regulatory processes. Chromatin inside the nucleus appears to be organized in chromosomal territories (CT) [56]. CTs are observed in interphase nucleus where each chromosome occupies a defined three-dimensional space. Differences in the compaction levels of autosomal CTs have been reported among several chicken haematopoietic cell lines, reflecting distinct differentiation stages [57].

The initial proposal of this model suggested that chromosome territories surfaces were rather smooth and separated by interchromosome domains (ICD). Genes localized to the periphery of the CT in order to approach the machinery in ICDs such as nuclear speckels, nuclear bodies and transcription factories (see below). These view has know turn into a more complex landscape in which CTs resemble a “chromatin sponge” with DNA free channels interspersed in between like lacunae. This space has been named interchromatin compartment (IC), which eliminates the “interchromosome” localization idea of ICDs [58]. The observation of different CTs has shown that chromosome distribution inside the nucleus depends on genic density having gene-reach chromosomes localized to the nuclear center and gene-poor chromosomes to the nuclear periphery. Additionally the CT-IC model suggests that nuclear architecture is highly dynamic where genes within a chromosome territory are localized differently depending on their expression levels. Coding sequences and their regulatory elements must relocate to the CT periphery and loop out to the IC in order to interact with transcriptional machinery (or others depending on the process) and be correctly expressed [58].

Simultaneously to the understanding of chromosome territories inside the nucleus and their link with the regulation of distinct epigenetic processes, nuclear specific compartments containing different molecular machineries were described. An outstanding example is the notion of “transcription factories”. Microscopic analysis of transcriptionally active sites in HeLa cells nucleus revealed \(\sim10^4\) sites of ongoing transcription, 8000 of them containing RNA Pol II and the rest RNA Pol III. These clusters measure \(\sim80\) nm and where named transcription factories [reviewed in 59]. More recently, a series of three-dimensional techniques have provided major insights into transcription factories function. Using a combination of these methods, Dr. Peter Fraser’s group has shown that active genes can localize to the sites of active transcription (and not vice-versa) and furthermore, that widely separated active genes can co-localized in one factory [60]. Moreover it has been shown that factories remain in the absence of transcription arguing in favor of a more “stable” structure than previously recognized [61].

How does DNA template gets into transcription factories is still under investigation. Several proposals have been made. One possibility is that RNA
Pol II gives the driving force in order to pull in its DNA template. Alternatively molecular motors as nuclear actin and myosin can be involved in recruiting sequences into factories. Finally, intergenic transcription at regulatory elements has been proposed as a critical step to set the factories gears into motion [reviewed in 59]. More recently, transcription initiation and elongation seem to be responsible for the relocation of DNA templates to transcriptional active foci [61]. For all the previous, it has become indispensable to analyze all epigenetic processes into the light of nuclear architecture and its impact on gene regulation.

Based in all this background it is not surprising that epigenetic processes are involved in several aspects of development, going from the early to the late stages of an organism life. In the next section we will address some of the most important epigenetic features.

II- Epigenetic reprogramming

Maternal and paternal genomes undergo a dramatic epigenetic reprogramming after fertilization. Gametes are terminally differentiated cells that are capable of initiating all the genetic and epigenetic processes necessary for a new organism development [62, 63]. Immediately after fertilization both maternal and paternal genomes suffer a progressive and generalized DNA demethylation by until now an unclear mechanism (Fig 2) [32]. However, imprinted genes are an exception to this global demethylation process keeping intact their parental epigenetic imprints. This allows specific mono-allelic expression of a selected group of genes required for the earliest stages of development (Fig 2) [29]. At the blastocyst stage, in particular in the inner cell mass of the blastocyst, differential patterns of de novo DNA methylation occurs in tight association with cell differentiation programs that give rise to the entire organism [33]. During early embryo development, both the de novo (Dnmt3a and Dnmt3b) and maintenance (Dnmt1) DNA methyltransferases are needed to establish proper post-blastocyst patterns of DNA methylation [64]. Why is so important to erase, with the exception of the imprinted genes, all DNA methylation patterns? One possibility comes from the need to erase some of the epigenetic marks coming from the inherited gametes. Such scenario may certainly allow, at the genome scale, the desilencing of the great majority of genes, facilitating the coordinated establishment of genetic and epigenetic programs early during cell differentiation. This model is consistent with the need of complementary epigenetic processes like histone modifications, the action of PcG proteins, and ATP-dependent remodeling complexes during early embryogenesis [65]. For example, H3K27me3, a characteristic repressive PcG-dependent
modification, has been found co-localizing with the H3K4me3 chromatin open mark in ES cells [66, 67]. This co-localization suggested the term “bivalent chromatin” distribution on embryonic stem cells referring to a chromatin state capable of rapid activation or repression of specific group of genes in response to precise developmental cues. For all the previous observations, it seems that DNA methylation and other epigenetic patterns should be transmitted from cell to cell ensuring the integrity of each genomic program in post-implantation embryos in order to achieve a proper embryonic development.

Figure 2. CpG methylation status during early mouse development. After fertilization the bulk genome undergoes demethylation. The lowest level of demethylation is reached at the preimplantation blastocyst stage at E 3.5. After implantation the whole genome becomes de novo hypermethylated in embryonic ectoderm and mesoderm (blue and red respectively) whereas the genome of extra-embryonic cells so as the primitive endoderm (green) and throphoectoderm (Light blue), remain hypomethylated. The parental imprinted genes escape demethylation. X inactivation is imprinted in extraembryonic tissues and random in embryonic ones [modified from 20].

III- Embryonic development: Epigenetics of the zygote

One of the most important features of post-fertilization and pre-implantation stages, are the zygote totipotency and the embryonic stem cells pluripotency, respectively. The zygote is a totipotent cell meaning that it can give rise to the whole embryonic and extra-embryonic tissues of the
organism. This cell divides forming blastomeres (34-64 cell stage), which retain totipotency but lack self-renewal capacity. In contrast in a later stage of development the cells originated within the inner cell mass of the blastocyst are capable to self-renew when cultured and are pluripotent, meaning that they can differentiate into every cell type of the embryo but not to extra-embryonic tissues. This takes place through the interdependency of genetic and epigenetic information that specify distinct differentiation profiles.

In addition, in the post-implantation embryo the pluripotent epiblast cells differentiate into somatic and pluripotent germ cells [68]. This latter case is very interesting since activating or repressive signals are needed for pluripotent epiblast cells to undertake a differentiation decision to establish a somatic cell program or to maintain the pluripotency identity as a germ cell (Fig 3) [65]. Evidences coming from different studies have brought to conclude that development and cell fate acquisition requires coordinated action between genetic and epigenetic instructions to activate or repress specific set of genes.

The zygote and its totipotency

Maternal and paternal epigenomes are surprisingly different. For instance, the paternal genome is structured in a highly compacted arrangement, through the incorporation of protamines, which are rapidly replaced by histones. Then, one of the earliest epigenetic events is the incorporation of the histone variant H3.3, also associated to active transcription sites into the genome [69]. As mentioned before, the zygote genome becomes rapidly demethylated in association with diverse patterns of histone post-translational modifications (Fig 2). From the maternal point of view, the zygote receives a well defined heritage composed by chromatin remodeling components as the PcG proteins EZH2 and EED among others, Brg1 that is a sub-unit of an ATP-dependent chromatin remodeler and key transcription factors associated to the pluripotency like Oct3/4 and Sox2 [70]. In fact, Brg1 loss can cause zygote arrest at the two-cell stage [70]. In addition, the epigenetic identity of the zygote is gradually reached with the progressive acquisition of H3K9me2 chromatin mark [71, 72]. All these epigenetic changes are critical for the determination of zygote’s totipotency needed for the establishment after a finite number of cell divisions, of the pluripotent embryonic stem cells and primordial germ cells (see below).

Later the epigenetic environment that is generated within the zygote activates the synthesis and function of critical transcription factors for blastocyst development. Oct4 and Cdx2 are two of the most important factors needed for inner cell mass (Oct4) and outer cells (Cdx2) development [73]. Another key factor is Nanog which is a homeodomain protein whose function is restricted to the inner cell mass [74, 75]. The epigenetic requirements at this early
developmental stage are exemplified by nanog expression regulation by Carm1 that is a histone-specific arginine methyltransferase [76]. Based on stage-specific synthesis of regulatory factors and epigenetic modifications, cell fate decisions are progressively made until the identity of the pluripotent embryonic stem cells is reached in the blastocyst at E3.5. Then, dramatic epigenetic differences are seen between the inner cell mass and trophoectoderm cells in the blastocyst at both the histone post-translational modifications and the incorporation of different DNA methylation patterns (Fig 3). DNA methylation occurs only in the ES cells, instead in the trophoectoderm, the genome stays basically hypomethylated [65].

![Figure 3](image_url)

**Figure 3.** Epigenetic reprogramming in mammalian development and pluripotent cells derivatives. After fertilization the paternal and maternal chromatin are differently packed. The maternal chromatin contains H3K9me2 and H3K27me3 and 5MeC. The paternal chromatin lacks these histone modifications and rapidly loses DNA methylation. Passive loss of 5MeC occurs until the blastocyst stage. Then, the inner cell mass (ICM) starts to acquire high levels of all three marks. The trophoectoderm (TE) derived tissue (placenta) remains hypomethylated. PGCs undergo 5MeC and H3K9 demethylation. At later stages de novo methylation including parental imprinting occurs in germ cells. Pluripotent cell lines can be obtained from the ICM, PGCs and spermatogonia stem (SS) cells [modified from 39].
The pluripotency of the embryonic stem cells

It has been proposed that the inner cell mass represents a “niche” where signaling molecules from the surrounding cells converge and contribute to the pluripotency of the embryonic stem cells. Furthermore, those signals may contribute to a general euchromatinization of the ES cells epigenome leading to the plasticity needed for pluripotency (Fig 4). It is well established that Oct4, Nanog and Sox2 are critical transcription factors necessary for the activation and repression of a large number of target genes contributing to the pluripotency of ES cells [77]. From the chromatin structure point of view the ES cells are highly dynamic with a generalized relaxed organization that tends to compact during most differentiation pathways (Fig 4). Also and as previously mentioned, ES cells present “bivalent” histone marks [66, 78]. Interestingly, these “bivalent” domains frequently overlap genomic regions in which the binding sites for Oct3/4, Sox2 and Nanog are clustered [79]. Then this two features of ES cells chromatin, the “bivalent” domains and the association of key transcription factors, may participate in the establishment and maintenance of pluripotency on the one hand, and in the ability of ES cells to respond to differentiation signals in order to acquire specific cell identities on the other. In addition to the histone H3K27me3 repressive mark co-existing with the H3K4me3 (bivalent domain), two other members of the Polycomb PRC2 complex, Esd and Suz12, are also present on ES cells chromatin and it has been suggested that they are ready to induce repression of a subset of genes as soon as the ES cells start to undergo differentiation [77, 80, 81]. Furthermore, mutations on epigenetic regulators including EZH2, Eset, MBD3 and Dicer, perturb the pluripotency of ES cells [65, 82].

An unexplored epigenetic aspect on ES cells is the role of the recent discovered reversible status of histone methylation through the action of the LSD and Jumonji domain-containing histone demethylases [83]. The window of possibilities is now bigger since in ES cells, histone methyltransferases and demethylases can coordinate their action to keep an undifferentiated state or make decisions leading to the activation of specific differentiation programs. According to this, recent work has shown that when Jmjd1a is knocked-down in ES cells it causes the activation of a differentiation pathway with an increase in H3K9me2. In contrast, depletion of Jmjd2c induces differentiation with a global increase in H3K9me3 [84].

On ES cells, domain specific chromatin conformations can also be reached by the incorporation of histone variants as part of a pluripotent “barcode” as proposed by Allis [8]. This hypothesis suggests that the relative amount of histone variants (in particular of H3), defines genomic domains with particular epigenetic features. In addition, the different characteristics of
histone variants contribute to determine nucleosome stability and thus to the flexibility of chromatin fiber [5]. This has been suggested on the basis of mutants of different histone chaperon components (HirA and CAF-1) that bring aberrant consequences in ES cells pluripotency [85, 86, 87].

Another aspect associated to ES cells epigenetics is the participation of ATP-dependent chromatin remodeling complexes, in particular, the case of NURD (nucleosome remodeling and histone deacetylation complex) and its components MBD3, that are required in vitro for differentiation of ES cells [88].

Figure 4. Chromatin status of ES cells and its derivatives. ES cells are pluripotent stem cells able to give rise all embryonic cell types. ES cells contain an “open” chromatin conformation with mostly euchromatin. As differentiation proceeds heterochromatin starts to be more abundant and finally in the majority of terminally differentiated cells, chromatin is mostly compacted and just some regions are still permissive.

Epigenetic regulation of germ cells

At the post-implantation embryonic stages E6-E6.5 epiblast cells remain pluripotent and have the capacity to generate all somatic tissues including germ cells. In contrast to the action of key transcription factors at the early blastocyst stage, at the post-implantation time, the establishment of programmed repressive procedures is needed. Different processes that include DNA methylation, histone methyltransferase activities, repressors
and co-repressors and non-coding RNAs achieve this repression in order to guide the transition of proximal epiblast cells to primordial germ cells (PGC), which will give rise to germ cells (sperm and oocytes) in the adult organism (Fig 3).

PGCs are capable to sense and respond to extra-embryonic signals to undergo differentiation. These cells are highly specialized and distinct to ES cells since they are incapable to create chimeras when introduced into blastocysts but they retain particular pluripotent characters. PGCs originated from the proximal epiblast generate germ cells and direct specific responses to block somatic programs, in particular through the action of the transcriptional repressor Blimp1 [65, 89, 90]. Deficient Blimp1 cells cause aberrant development of PGCs and the cease of cell proliferation.

From the epigenetic point of view, Blimp1 can interact and form complexes with several remodeling complexes including the histone methyltransferase G9a, HDAC2 and the arginine-histone methyltransferase Prmt5 [91]. Therefore the covalent modifications H2A and H4R3me2 seem to be relevant for PGC formation and germ cell specification [65]. Consistent with the role of Prmt5, mutation of its Drosophila homologue, Casuleen/daut5, causes dramatic deleterious effects on germ cells [92,93]. Then, one key epigenetic aspect of the germ cell lineage is the maintenance of some of the features of pluripotency therefore, the great majority of the epigenetic processes should focus on the generation of such specific type of pluripotency. Accordingly, it has been proposed that PGCs have some of the zygote and some of the ES cells epigenetic features. In the future it will be interesting to decipher which are the epigenetic programs that are exclusive to the germ line.

IV- Genomic imprinting during development

Genomic imprinting is the phenomenon associated to the differential gene expression of paternally and maternally inherited alleles. Imprinted genes are often required for tissue and developmental processes during early development [29]. Notoriously, imprinted genes are distributed in clustered domains that often include both paternally and maternally silenced genes. Mono-allelic gene expression is in part determined by parental epigenetic marks defined during gametogenesis [94]. Importantly, imprinting is not present in all organisms and it has been suggested to appear in marsupials around 201 million years ago [35, 95].

One key component of imprinted loci that is tightly associated to epigenetic regulation, is the allele differentially marked cis element defined
as the imprinted control region (ICR), also named as differential methylation domain (DMD) [29]. It is now well established that differential DNA methylation of paternal or maternal alleles at the DMD is one of the key processes leading to mono-allelic expression of imprinted loci. In the present chapter we will not go through all the aspects associated to genomic imprinting, instead we would describe the latest epigenetic mechanisms associated to allele-specific gene expression.

CTCF and imprinting mechanisms

The \(Igf2/H19\) locus has been the paradigm of imprinted domains. In this domain, the fetal growth factor insulin-like growth factor 2 gene (\(Igf2\)) is paternally expressed and the \(H19\) gene is maternally expressed. Both genes are located around 100 kb from each other and are regulated by enhancers downstream of the \(H19\) gene (Fig 5) [29]. These genes are mainly expressed in mesodermal, endodermal and extraembryonic tissues in the developing fetus, with a downregulation around 3 weeks of postnatal development. Between \(Igf2\) and \(H19\) genes there are 2 kbs corresponding to the imprinted control region. The DMD is DNA methylated from the paternal derived germ line, in contrast to the unmethylated state on the maternal allele. The differential DNA methylation of the DMD is essential to achieve monoallelic gene expression. Interestingly at the maternal allele, the DMD is acting as an enhancer-blocking element that is dependent on multiple CTCF binding sites [97]. CTCF is exclusively bound to the maternal allele and its binding is sensitive to DNA methylation. Therefore, when CTCF is binding to the DMD, the downstream enhancer only trans-activates the \(H19\) gene, while in the paternal allele, the DMD is methylated, CTCF can not bound and then the enhancer now acts over the \(Igf2\) gene (Fig 5).

More recently, Reik and collaborators demonstrated that in addition to the enhancer blocking mechanism, there is differential formation of chromosome loops that explain allelic-specific expression at the \(Igf2/H19\) locus [98]. This was performed through the application of a recent developed method known as chromosome conformation capture (3C). On the maternal allele the unmethylated \(H19\) DMD, which is bound by CTCF, interacts with the \(Igf2\) DMR1 (Fig 5). This conformation results in the formation (by looping) of two topological independent chromatin domains, with the \(H19\) gene in an active domain where the enhancers can activate its promoter. In contrast, the \(Igf2\) gene is placed on an inactive domain unable to be activated by the enhancers (Fig 5) [98]. Then, the \(Igf2/H19\) imprinted locus exemplifies how higher-order chromatin structure can actively participate on epigenetic imprinting regulation.
Surprisingly, interchromosomal interactions are not only involved in regulating gene expression in cis [59]. Recent observations demonstrate that the epigenetic role of the DMD of the Igf2/H19 locus located on the murine chromosome 7 interacts in trans with the imprinted locus Wsb1/Nf1 on chromosome 11 [99]. Interestingly, and supporting its topological role, those interchromosome interactions are dependent on CTCF-binding to the maternal allele. This is consistent with the regulated loop formation at the murine β-globin locus in which CTCF mediates the long-range contacts and three-dimensional conformation of the locus [100]. Then, the epigenetic processes occurring in early stages of development do not only require the action of highly specific transcription factors; long-range contacts, chromatin fiber topology and nuclear dynamics are also participating as epigenetic mechanisms that control different networks of gene expression.

In addition to coding transcription, a large portion of the genome composed by non-coding sequences, are constantly transcribed. Then, a vast variety of non-coding RNAs (ncRNAs) represent a common feature of mammalian gene regulation with consequences during development. The classical example of ncRNA is the Xist RNA, which is responsible for the initial epigenetic steps leading to X-chromosome inactivation in female cells [101]. Today, there are clear evidences showing that long ncRNAs participate as well, on the silencing of imprinted genes. A growing list of imprinted loci is emerging including the Igf2/H19, Kcnq1, Gnas, Air (antisense of Igf2) and others [102]. It has been proposed that in some way the imprinted control regions may have something to do with the regulation and expression of ncRNAs. In all the cases, the non-coding RNAs are acting through epigenetic silencing but the mechanisms are still not clear [102, 103]. Several models can be proposed to explain the mechanism of action of this kind of RNAs; for example, the participation of the RNA interference machinery to induce repressive chromatin conformation (a model that requires the formation of double-stranded RNA intermediates) [104]. An alternative model could come from the influence of transcription of ncRNAs affecting imprinted genes in an allelic-specific manner by either opening or closing the chromatin structure. In fact in both cases we can imagine that ncRNA transcription is contributing to the formation of an open chromatin configuration that can be of more easy access to regulatory factors, either positive or negative.

In summary, imprinting regulation and its influence during early stages of development is clearly a complex process. Despite the large amount of knowledge generated around imprinting there are novel and exciting results clearly illustrating the participation of epigenetic mechanisms on the regulation of imprinted loci with a determined role during development.
Chromatin remodeling and epigenetic regulation during development

Figure 5. *Igf2/H19* locus. In the maternal allele CTCF binds the unmethylated DMD and the DMR1 forming a loop that excludes *Igf2* gene from the enhancer’s activation so that only *H19* is activated. In the paternal allele, the DMD is methylated so CTCF cannot bind. Therefore, a different loop is formed in which the enhancer can trans activate *Igf2* gene. *H19* gene is silenced by DNA methylation on its promoter [modified from 96].

V- Senescence

The development of an organism is a progressive process and after reaching adulthood another progressive process begins that leads to senescence and ageing. Ageing may be understood as the decay in the regenerative capacity of tissues and these phenomena is directly linked with a change in the progenitor cells of each tissue. Based on such view, cellular senescence in an adult organism can be a caused by the combination of the physiological loss of regenerative competence and the unbalance of tissue homeostasis that lead to cell ageing.

Cellular senescence can be described as a state of permanent and irreversible cell cycle arrest with no capacity of response to serum or growth factors. Senescent cells exhibit a large and flat morphology and are positive for the senescence-associates β-galactosidase assay. It is known that p53 and retinoblastoma (Rb) tumor suppressor genes participate in the onset and maintenance of the senescent cell state [105]. However, the precise mechanisms are still not well understood.
As mentioned, a currently unsolved question resides on which are the processes that lead to cell senescence and ageing. Different evidences have emerged arguing that senescent cells present distinctive epigenetic features thus making the epigenome analysis an appealing start point to look into senescence. As cells go through senescence, heterochromatin foci called senescence-associated heterochromatin foci (SAHF) are formed [106]. In human senescent cells there are between 30 and 50 SAHFs representative of highly compacted chromosomes [107]. These foci consist on heterochromatic spots in which histones are hypoacetylated, there is an enrichment of H3K9me3 and incorporation of the heterochromatin protein 1, HP1. In addition, SAHFs are depleted of linker histone H1 and the present deposition of the histone variant macroH2A and the high mobility group A (HMGA) proteins [for review see 108]. It has been proposed that SAHFs must represent the repressive foci in which proliferative promoting genes are silenced in senescent cells [109].

Senescence represents an attractive scenario to understand certain epigenetic mechanisms linked to tumor suppressor processes. Conceptually, a senescent cell can be viewed as incapable to respond to mitogenic stimuli thus presenting anti-proliferative properties and SAHFs are important structures in order to maintain a this state. Therefore and to certain extent, understanding the genetic and epigenetic mechanisms of senescence could give rise to alternative therapeutic strategies to obstruct malignant cell proliferation in cancer.

Discussion and prospects

Epigenetics comprehend a broad range of processes involved in every aspect of organism development, life and ageing. One of the most outstanding evidences about the importance of epigenetics during early stages of development and cell fate determination has emerged from the examination of cloned mammalian embryos [110]. A specific study revealed drastic differences on H3K9 methylation and global DNA methylation when comparing a zygote from cloned and a wild-type bovine [110]. This observation clearly demonstrates an incomplete reprogramming of epigenetic marks and that such marks are critical for preimplantation embryos, their development and adulthood. An intriguing aspect that has yet to be resolved is to understand the epigenetic mechanisms occurring at the developmental transition between pre-implantation and post-implantation. Particular attention is needed to unravel the genetic and epigenetic programs that maintain the undifferentiated state of ES cells and the initial steps they take towards the establishment of particular differentiation programs.
The interdependency between genetic and epigenetic mechanisms requires from the genetic perspective, the expression of specific network of transcription factors that blocks differentiation and promotes continuous self-renewal capacity of ES cells. From the epigenetic point of view, those transcription factors and their associated co-factors should have the responsibility to attract chromatin remodeling activities that allows the construction of the particular scenario (open and close chromatin conformation) for self-renewal and differentiation of ES cells.

Moreover, the understanding of epigenetic processes involved in cell senescence, ageing and disease may provide a new platform to better approach therapeutic protocols. A characteristic feature of epigenetic defects is that apparently, and the majority of the data supports this idea, they tend to be acquired in a progressive way rather than abruptly as compared to genetic ones. In the next future, these must be capitalized in terms of improved strategies for early diagnosis in diverse pathologies as cancer.

Acknowledgments

This work was supported by grants from the Dirección General de Asuntos del Personal Académico-UNAM (IN209403 and IN214407), Consejo Nacional de Ciencia y Tecnología, CONACyT (42653-Q and 58767), and the Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA). MF-M is a fellowship recipient from CONACyT.

References

Chromatin remodeling and epigenetic regulation during development

64. Okano, M, Bell D,W, Haber, D, A, and Li E, 1999, Cell, 99, 247-257.