Epigenetic mechanisms in development, inheritance and disease

Nikki Minnebo, Aleyde Van Eynde, Mathieu Bollen

Laboratory of Biosignaling & Therapeutics, Department of Cellular and Molecular Medicine, University of Leuven, Belgium

*Corresponding author: Prof. Dr. M. Bollen, Laboratory of Biosignaling & Therapeutics, Department of Cellular and Molecular Medicine, KULeuven, Campus Gasthuisberg, O&N1, Herestraat 49 Box 901, 3000 Leuven, Belgium. E-mail: Mathieu.bollen@med.kuleuven.be

Received: 02.12.2011 Accepted: 23.02.2012 Published: 05.03.2012

Abstract

Epigenetics is a relatively new and exciting field of the (bio)medical sciences. It confers an additional layer of information that controls gene expression by mechanisms involving DNA methylation, histone modifications, chromatin compaction and non-coding RNAs. As most epigenetic marks are mitotically transferable and sometimes even stable during meiosis, these modifications can potentially be passed on to future generations. Proper epigenetic signaling is essential for normal proliferation and differentiation, and epigenetic misregulation is a key feature of many common diseases including cancer, diabetes and cardiovascular disease. As most epigenetic marks are reversible, understanding the underlying disease mechanisms reveals the therapeutic potential for interference with aberrant epigenetic signatures. These therapies hold great promises for a wide range of applications. However, the creation of both potent and specific epigenetic therapies requires a detailed insight into the patients’ epigenomic landscape. Here, we provide an overview of the major epigenetic signaling pathways and their contribution to development and disease.

Keywords: epigenetics, chromatin, environment, inheritance, disease, therapy.

INTRODUCTION

We all originate from a single omnipotent stem cell, the fertilized egg. Many cycles of cell division, combined with a stepwise differentiation process, produces the thousands of billions of cells that constitute our body. With each round of cell division the genetic material or DNA is doubled and passed on to the daughter cells,
implying that each person is basically a clone of cells. Eukaryotes are faced with the problem to organize their large genome (human cells contain about 3 billion basepairs (bp), the equivalent of 2 meters of DNA) into the micrometer dimensions of a nucleus and at the same time keep the DNA sufficiently accessible to regulatory factors. Nature’s answer to this conundrum is the nucleosome. Each nucleosome is comprised of 146 bp of double-stranded DNA, wrapped around an octamer of 2 copies of each of the histone proteins H2A, H2B, H3 and H4. Successive nucleosomes are linked by a short fragment of DNA giving the appearance of a string of beads (Fig. 1). This string is further coiled and looped around scaffold proteins, giving rise to the protein-DNA complex called chromatin (1-5).

The human body contains about 210 different cell types that arise by changes in the gene expression profile during the stepwise differentiation of pluripotent stem cells into fully committed and terminally differentiated cells (Fig. 1). This process is driven by epigenetic mechanisms that determine which genes become repressed or activated during development (1;3;6;7). Although epigenetics is a rather new scientific discipline, the name epigenetics was already coined in 1938 by Conrad H. Waddington, an embryologist studying Drosophila, who used it to describe the link between a given phenotype and its associated genotype (now termed developmental biology) (8;9). He demonstrated that a temperature shock before puparium formation produced flies with cross veinless wings, a phenotype that could still be detected after 16 generations (10). He believed that the morphological and functional properties of an organism were dictated by a program controlled by the genome and influenced by the environment. Since then, epigenetics has been defined in many ways but the common theme is that it describes an additional, heritable layer of information, on top of the DNA nucleotide sequence, that influences the expression level of subsets of genes (11).
Figure 1. The epigenome determines cell fate. All cells of an organism contain an identical genome, comprised of an individual DNA sequence. Inside a cell, the DNA is condensed by its wrapping around a core of histone proteins, forming the nucleosomes (red ellipses) and giving rise to a DNA-protein complex called chromatin. Chromatin can be epigenetically modified to switch genes on or off. Examples of such epigenetic modifications are DNA methylation (green triangle) and histone modifications (orange sphere). In addition, the local chromatin structure can be changed by chromatin remodeling complexes (blue ellipse). ncRNAs (blue lines) have also been implicated in gene regulation. Although an organism has only one single genome, it contains many different epigenomes, which determine cell fate, function and phenotype. Cartoon cell types at the bottom from left to right: red blood cell, neuron, osteocyt, muscle cell and epithelium.

DIFFERENT MECHANISMS OF EPIGENETIC REGULATION

In mammalian cells, the major part of chromatin is present in a highly condensed, heterochromatic state, which is largely devoid of transcription factors and thus transcriptionally inactive. Euchromatin is less condensed and mainly consists of actively transcribed genes (12). Both histones and DNA can be chemically modified with various epigenetic markers that change the affinity for chromatin-binding proteins and/or influence the electrostatic interaction between DNA and histones (13;14).

DNA methylation on CpG dinucleotides

DNA can be methylated by the chemical addition of a methyl (CH$_3$-) group to the C5 position of cytosine bases (15). In
mammals, mainly cytosines followed by a guanine (CpG) are highly methylated (Box 1). Whereas CpGs are statistically underrepresented and dispersed throughout the genome, there are regions of high CpG clustering called CpG islands, mostly in the promoter region of coding sequences. In fact, approximately 70% of all annotated gene promoters contain a CpG island which makes this one of the most common promoter elements in the genome. Virtually all housekeeping genes as well as some differentiation genes have an associated CpG island. Within islands, most cytosine bases are unmethylated to allow transcription of the associated gene. In contrast, CpGs that are dispersed throughout the genome are usually methylated (16;17).

In human cells, there are three main enzymes that catalyze the cytosine methylation reaction: DNA methyltransferase (DNMT)1, DNMT3a and DNMT3b. DNMT1 is a maintenance enzyme that copies the methylation pattern from mother to daughter strands after replication. The two other enzymes are de novo methyltransferases that establish the methylation pattern during early embryonic development (18). DNA methylation can have profound effects on the structure of large chromosome fragments, but is also implicated in the fine-tuning of the expression of single genes.

Dispersed CpG methylation stabilizes the integrity of the genome by silencing parasitic DNA sequences such as transposons. Although not firmly established for mammals, this process has been well documented for plants and fungi, where transposons are specific targets for DNA methylation. Methylated cytosines are prone to spontaneous deamination resulting in their conversion to thymine. Such point mutations can prevent transposons from changing their position in the genome (19;20). In addition, DNA methylation in the transposon promoter region can prevent its expression.

DNA methylation in promoter regions of protein coding genes represses their transcription. Two distinct underlying mechanisms have been proposed: [1] the methyl mark prevents the recruitment of DNA-binding transcription factors and RNA polymerase to target genes (21), and [2] some proteins specifically recognize methylated CpGs and themselves recruit histone modifying complexes that establish a transcriptionally inactive chromatin environment (18). Recently, it has been proposed that DNA methylation also directs nucleosome positioning in the promoter region of inactive genes (22;23).

For imprinted genes, the DNA methylation pattern differs between the maternal and paternal genome, allowing expression from only one of the two alleles. One of the first discovered imprinted genes was insulin-like growth factor 2 (IGF2). A loss of imprinting (LOI) at this region results in a doubling of IGF2 expression levels and is often associated with an increased frequency of intestinal tumors (24).

Finally, DNA methylation also plays a crucial role in X-chromosome
inactivation. In mammals, one of the two female X-chromosomes becomes silenced during early embryonic development to compensate for the single copy in male cells. In this way, X-linked genes will be expressed at an equal level in both male and female cells. X-inactivation is initiated at the X-inactivating centre from which a 17-kb long non-coding (nc)RNA, called Xist, is transcribed (25). This ncRNA then coats the future inactivated X-chromosome in cis, resulting in the recruitment of Polycomb group (PcG) proteins (see below) and DNMTs that silence most X-linked genes (26;27). Moreover, the key PcG protein EZH2 interacts directly with all three DNMT's, providing a link between DNA methylation and histone modifications (28).

Box 1: Techniques for studying epigenetic modifications

The development of several research tools has been crucial in the mapping of epigenetic modifications. The most thoroughly studied modification is DNA methylation. Initially, differences in methylation of a particular sequence could only be detected with methylation-specific restriction enzymes. This technique severely limited the range of sequences that could be investigated, a drawback that was solved by the discovery of bisulphite treatment (19). This technique reproducibly converts unmethylated cytosines into uracil bases but leaves methylated cytosines unchanged. Differences in DNA methylation at any given sequence can subsequently be investigated by PCR with methylation-specific primers.

To investigate histone modification signatures or DNA-binding profiles of a particular nuclear factor, ChIP or chromatin immunoprecipitation is the method of choice (2). After crosslinking, genomic DNA is fragmented and an antibody is used to precipitate the protein of interest, together with fragments of associated DNA. After reversing the crosslinks, the precipitated proteins are digested and the associated DNA is purified. Any DNA of interest can then be quantified by qPCR, using appropriate primer sets.

A genome wide mapping of chromatin signatures uses the hybridization of fluorescently labeled DNA to a microarray. More recently, next-generation sequencing provides a greater coverage, superior sequencing resolution and requires less input material, which is often limited for ChIP experiments (2).
The histone code

Histones are globular proteins and their N-terminal tails protrude from the nucleosome core structure. These tails can carry different posttranslational covalent modifications including acetylation, methylation, phosphorylation, ubiquitylation and sumoylation. Histone modifications do not provide a simple on-off switch to regulate gene expression but rather act combinatorially as reflected by the term “histone code” (29-32).

Histone modifications can function as docking sites for other DNA-regulatory proteins, each associated with a specific biological process. In addition, they can have direct activating or repressing functions depending on the type and localization of the modification. Some histone modifications change the charge of the histone tail. For example, positively charged lysine residues lose their charge upon acetylation; this weakens the histone interaction with the negatively charged DNA, which makes the DNA more accessible for transcription factors, resulting in transcriptional activation. In addition, acetylated histones serve as docking sites for bromodomain-containing chromatin modifying proteins (30). In contrast, histone methylation on lysine or arginine residues does not change their charge. Rather, methylated residues are binding sites for chromodomain-containing transcriptional activators or repressors (32;33). For example, active genes are trimethylated on lysine 4 of histone H3 (H3K4me3) near the transcription start site. During transcription such genes also become trimethylated on H3K36 across the gene body. These methylated lysines subsequently attract chromodomain-containing transcription factors and chromatin remodelers. Methylation of H3K9 and H3K27 on the other hand are strongly associated with gene silencing and the establishment of a less accessible, facultative heterochromatic structure (32). For example, the H3K27me3 mark is recognized by the chromodomain-containing Polycomb protein that resides within a multiprotein complex named Polycomb Repressive Complex (PRC)1 (34). Another PRC1 protein called RING1, is an enzyme that monoubiquitilates histone H2A on K119 leading to a more condensed chromatin state. By a still unresolved mechanism, PRC1 can also impair the binding or progression of RNA polymerase along these genes leading to transcriptional silencing (35-37). Finally, DNMTs can be recruited by the H3K27 methyltransferase EZH2 to methylate the target DNA, providing a more stable and silent chromatin environment (28).

H3K27me3 is particularly important for the silencing of key differentiation factors during embryogenesis (38-40). This repressive mark is deposited by PRC2 which contains three core proteins: the methyltransferase EZH2 and its two co-activators SUZ12 and EED. PcG proteins were first discovered in Drosophila by mutations in a group of transcription factor genes, called homeotic (Hox) genes, that are important for early
fly development (1;41). The name Polycomb itself refers to the phenotype of male flies carrying mutations in the Polycomb gene. Male flies normally have a set of bristles or “sex combs” on their first pair of legs to help them during mating. Mutations in the Polycomb gene caused them to develop multiple sex combs on all their legs, hence the name Polycomb (42). In general, mutations in any of the PcG protein encoding genes cause deficient methylation of H3K27 and are associated with developmental defects (also called homeotic transformations), owing to their important role in development and differentiation (38;43). However, the PcG proteins also regulate the expression of hundreds of genes in differentiated cells, and therefore should be viewed as key regulators of transcription in all cells.

The biological counterpart of the PcG proteins are the Trithorax Group complexes that activate genes. As suppressors of PcG mutant phenotypes, some Trithorax complexes contain histone methyltransferases such as Ash1 that methylate H3K4. In addition, these complexes can also contain components of the ATP-dependent chromatin remodeling complexes NURF or SWI/SNF that facilitate the recruitment of the transcriptional machinery (44;45).

**ATP-dependent chromatin remodeling**

Nucleosomes are not evenly spaced along the DNA and often the removal or sliding of nucleosomes along promoter areas is necessary for transcription factors and RNA polymerase to gain access to the DNA. Indeed, the small promoter region surrounding the transcription start site of active genes is usually depleted of nucleosomes (46). Nucleosome remodeling is also indispensable for DNA polymerase to copy the DNA along the nucleosomes during the S-phase of the cell cycle (47). Nucleosome remodeling complexes use ATP as an energy source to alter histone-DNA interactions. By relocating nucleosomes to available acceptor DNA in cis or even trans, nucleosome remodeling ATPases establish a dynamic chromatin state in which the overall packaging of DNA is maintained, but specific sequences can be transiently exposed. In mammals, many different remodeling ATPases have been discovered and they all associate with a wide range of other proteins giving rise to large multiprotein complexes with distinct functions (47;48). For example, the SWI/SNF family of ATPases can slide nucleosomes along the DNA axis and even eject histones from the nucleosomes, whereas the ISWI family is only capable of sliding nucleosomes (46). The Mi-2/NuRD complex is unique in that it combines a nucleosome sliding function with a histone deacetylase (HDAC) activity through interaction with HDAC1 and 2 (49;50). Another epigenetically related process can be found in the ISWI-type WICH complex, which binds to replication foci during S-phase. By keeping nucleosomes mobile, it provides access to the newly replicated DNA and prevents it from turning into heterochromatin (51). In doing so, this complex safeguards the transmission of the
Long non-coding RNAs

ncRNAs are functional transcripts that do not encode proteins. They represent a significant part of the mammalian transcriptome and regulate gene expression at multiple levels. Besides small (20-25 bp) ncRNAs, such as miRNA or siRNA, that mainly regulate translation by targeting mRNA for degradation (52), some of these ncRNAs are quite large, up to several hundred kb. These long ncRNAs have been shown to regulate epigenetic gene silencing, both in cis and in trans, by binding to complementary DNA and/or by association with chromatin-modifying complexes (53-55). The Xist RNA is only one example of a cis-inactivating ncRNA (26;27). Another well-known case is Air, a non-coding transcript of 108 kb that is transcribed from the second exon of the paternal allele of the IGF2 receptor gene (56). Air binds and silences a 490-kb region in cis and interacts with the H3K9 methyltransferase G9a (57). Yet another ncRNA is the 91-kb Kcnq1ot1, which is transcribed in the antisense direction from an intron of the Kcnq1 gene. The internal promoter for Kcnq1ot1 is only methylated on the maternal chromosome, accounting for the specific expression of Kcnq1ot1 from the paternal allele. Paternally expressed Kcnq1ot1 recruits the methyltransferase G9a and the PRC2 complex, which locally deposit the repressive H3K9 and H3K27 methylation marks, respectively, silencing protein encoding genes inside the 1-Mbp Kcnq1 domain. Intriguingly, paternal imprinting of Kcnq1ot1 is seen in placenta but not in liver and thereby contributes to lineage specific transcriptional silencing (58;59).

An example of a ncRNA-based epigenetic silencing mechanism in trans is HOTAIR. This 2.2-kb antisense transcript is generated from the HOXC cluster on chromosome 12 but represses transcription across a 40-kb region in the HOXD cluster on chromosome 2. HOTAIR also interacts with the PRC2 complex and is required for the deposition of the H3K27me3 mark across the HOXD region. Accordingly, a knockdown of HOTAIR results in an expression of genes in the HOXD locus and this is associated with a local drop in the level of H3K27me3 (60;61).

A recent genome-wide analysis identified over a thousand highly conserved mammalian ncRNAs (62). About 20% of these are bound by PRC2 in various cell types whereas some newly identified ncRNAs are bound by other chromatin-modifying complexes. Consistent with these findings, the siRNA-mediated depletion of PRC2-interacting ncRNAs results in the derepression of associated genes (63). Given these recent developments, ncRNAs can be predicted to play an important general role in directing chromatin-modifying complexes to their target genes.
EPIGENETIC PROCESSES ARE CRUCIAL FOR EMBRYONIC DEVELOPMENT

Studies in mice and Drosophila have shown that a depletion of epigenetic regulators often results in severe developmental defects or even early embryonic lethality. In mice for example, the lack of DNMT1 results in embryonic lethality at day E8.5-9 and DNMT1 heterozygous embryos show rudiments of the major organs but they are smaller than their wild type littermates and develop more slowly (64). Moreover, loss of any of the core PRC2 subunits results in severe defects and lethality around gastrulation (65-67). These effects suggest an essential role of these chromatin-modifying proteins in normal development.

PcG proteins also maintain the pluripotency and self-renewal capacity of embryonic stem (ES) cells by controlling key developmental regulatory genes. For example, ES cells lacking SUZ12 and EED show a more differentiated phenotype (39;68), and ES cells devoid of EZH2 are unable to reprogram B cells towards pluripotency (69). In ES cells, many genes that are important for development contain both active (H3K4me3) and repressive (H3K27me3) chromatin marks and are therefore named “bivalent” genes. During further development, specific signaling processes determine cell fate decisions by activation or repression of these bivalent domains (70). A nice example of such lineage commitment was described for skin development (71;72). EZH2 is highly expressed in basal progenitor cells and promotes their proliferation by repressing the Ink4A-Ink4B locus, which encodes inhibitors of cell cycle progression. As skin cells differentiate and move to the surface of the skin, EZH2 expression levels decrease resulting in the expression of the Ink4A-Ink4B locus and a proliferation arrest. This study reveals that the EZH2-containing PRC2 complex controls epigenetic modifications both temporally and spatially in tissue-restricted stem cells. In this way, the PRC2 complex ensures that the proliferation potential is maintained and that undesirable differentiation processes are repressed (73). Similar conclusions have been drawn from studies on the role of EZH2 in neurogenesis (74), myogenesis (75;76), hematopoiesis (77) and adipogenesis (78).

Another appealing example of developmental epigenetics comes from the honeybee (Apis mellifera). If a new queen is needed in a honeybee population, one larva is chosen to be fed large quantities of royal jelly as an exclusive food source during the first four days of its growth. This type of early feeding triggers the development of queen morphology, including fully developed ovaries that are needed to propagate the species. Intriguingly, the silencing of DNMT3 in honeybees causes similar effects as the feeding of royal jelly, suggesting that this food source contains a DNA methylation inhibitor (79).
TRANSGENERATIONAL EPIGENETIC INHERITANCE

The transmission of genetic information from parent to progeny is well understood and generally accepted. However, the transfer of a non-genetic determination of phenotype has also been documented and is now known as “soft inheritance” or transgenerational epigenetic inheritance (TEI) (80).

A first clear example of TEI came from Överkalix, a desolate small town in northern Sweden where, during the 19th century, its small population would starve if the harvest was poor. Several decades later, Bygren and colleagues found that pregnant women, suffering from food deprivation, delivered children with an increased risk of cardiovascular disease (81). Similarly, the children of women who were pregnant during the Dutch hunger winter of 1944, suffered from impaired glucose tolerance as an adult which was correlated with less DNA methylation of the imprinted IGF2 gene (82).

Unfortunately, large scale studies of epigenetic inheritance in humans are scarce and most information on TEI comes from studies on plants and animals (80). For example, a study in rats showed that maternal grooming caused offspring to be more fearful and this was linked to promoter hypermethylation of the Glucocorticoid Receptor gene in the hippocampus. This stress-induced phenotype was epigenetic but not transferred through germ cells as cross-fostering pups from one mother to the other during the first week caused pups to adopt the stress-phenotype of their new mother (80;83).

Mice have been engineered with a uniquely regulated agouti viable yellow gene (A<sup>vy</sup>) that can be used as an easy epigenetic read-out system (84). The A<sup>vy</sup> gene gives mice a yellow coat color and is correlated with a propensity for obesity and diabetes when expressed constitutively. This model system was used to show that ethanol consumption by pregnant females led to an increase in DNA methylation in developing embryos at the A<sup>vy</sup> locus, correlating with a change in coat color from yellow to pseudoagouti (brown). A similar switch in coat color was observed when feeding pregnant mice a methyl-rich diet, changing the phenotype of the pups towards the healthy, longer-lived direction (85).

Only very recently, a study of the roundworm Caenorhabditis elegans (86) showed that methylation of H3K4 plays a key role in longevity. Deficiencies in the H3K4 methyltransferases ASH-2 or WDR-5 in the parental generation extends their lifespan with about 20%, an effect that is observed up until the third generation. This transgenerational transmission depends on the H3K4 demethylase RBR-2 and is associated with epigenetic changes that influence the expression level of key genes involved in longevity.

Although clear examples of TEI exist, several issues concerning this phenomenon remain unclear. For an epigenetic mark to be transmitted to future progeny, these modifications must be
present in the germ cells. However, the paternal DNA becomes demethylated prior to the first cell division of the zygote and the maternal genome is demethylated after several cleavage divisions. De novo methylation is subsequently established in the inner cell mass of the blastocyst (87;88). Additionally, in sperm cells, the histones are replaced by smaller protamines without modifiable tails that would disable any histone-linked epigenetic information to be passed through the male germ line (89). However, it was recently found that some histone-containing nucleosomes are retained in sperm and these appear to be specifically present at loci important for embryonic development (90). Moreover, parts of the genome seem to escape the genome-wide demethylation process during embryogenesis, potentially mediating epigenetic inheritance. TEI can also be partially explained by the presence of developing germ cells inside the embryo of a pregnant mother. Therefore, environmental conditions acting on the pregnant mother can also induce epigenetic changes in the developing germ cells of the embryo, accounting for the transmission of a phenotype from grandparent to grandchild (80;89).

In summary, environmental conditions (food, stress, toxins,...) can cause epigenetic changes, resulting in an altered phenotype (Fig. 2) and in some cases these effects can be transferred to the next generation(s). However, this does not imply that these changes are stable throughout evolution. The notion of soft inheritance is still rather controversial and is often linked with the work of Jean-Baptiste Lamarck (1744-1829), who proposed an evolutionary theory based on the inheritance of properties acquired during the lifetime of an organism that can be summarized in contemporary terms as “the environment contributes to an inherited phenotype”. This theory contrasts with that of Charles Darwin (1809-1882), which is based on the competition between individuals that show a (natural) genetic variation and can be translated as “the (epi-)genotype determines the phenotype”. Although the Darwinistic view of evolution is now broadly accepted, the notion of epigenetic inheritance has reopened the scientific debate on the contribution of the Lamarckianistic view (89;91). Hard evidence of true epigenetic “inheritance” is still missing as most observed transgenerational effects tend to fade away after a few generations (92). However, there is evidence that environmentally induced epigenetic changes contribute to the variability between individuals in a population, thereby increasing the speed of evolution in a Darwinistic manner. Indeed, in contrast to genetic mutations, epigenetic signals can be easily modified and provide a powerful mechanism for rapid adaptations to sudden environmental changes (93).
Figure 2. Environmental factors can induce both genetic and epigenetic alterations. An organism’s phenotype is determined by the combination of genetic and epigenetic determinants. Environmental factors such as radiation or smoking can influence a phenotype by introducing genetic alterations. The mutation of the DNA sequence is irreversible and sometimes alters the amino acid sequence of the encoded protein which may be the cause of specific “inherited” diseases. Epigenetic alterations (epimutations) can also contribute to disease development and even influence an organism’s lifespan, but in contrast to genetic mutations, they are reversible. Diseases caused by genetic mutations are hard to cure and often require surgery or chemotherapy to remove the affected cells. In contrast, diseases that originate from epigenetic alterations hold great promise for conventional treatment, as the epimutation can be restored to normal by inhibiting specific chromatin-modifying enzymes.

ABERRANT EPIGENETIC SIGNALING IN DISEASE DEVELOPMENT

The last decade of research has tremendously increased our understanding of the mechanisms underlying epigenetic phenomena, both in normal development and disease. Currently, it is well known that environmental factors and lifestyle choices not only cause DNA sequence alterations but also change the epigenetic marks on our chromatin. Aberrant epigenetic signals can contribute to disease development and can even be passed on to future generations. Importantly, epigenetic marks, unlike genetic mutations, are in se reversible and scientists are beginning to understand how they can manipulate these
epigenetic signals to revert them to their original states (Fig. 2).

Cancer comprises over 200 different diseases of abnormal cell growth induced by a series of mutations, but also involving epigenetic changes. DNA methylation in cancer has been extensively studied and the onset of cancer is generally accompanied by a global demethylation of the genome. This leads to a destabilization of the genome and thus increases the risk of gaining additional mutations (94). Hypomethylation is also observed locally, for example at the promoter of oncogenes, increasing their expression levels. Conversely, hypermethylation is observed in the promoter region of tumor suppressor genes, as well as genes involved in cell-cycle regulation, DNA repair, apoptosis and angiogenesis (95;96). Importantly, many PcG proteins are themselves oncogenes and their expression levels are often elevated in cancer samples. For example, SUZ12 is found to be overexpressed in colon, liver and breast cancer, while EZH2 is overexpressed in 16 different types of cancer and is correlated with increased angiogenesis, metastasis and a poor patient outcome (97). Indeed, EZH2 can be considered a true oncogene as removal of EZH2 from cancer cells results in growth arrest whereas ectopic expression of EZH2 in normal cells promotes invasion and cell proliferation in vitro (98). Detection of aberrant DNA methylation and overexpression of PcG proteins therefore hold great promise as biomarkers for the diagnosis and prognosis of cancer (99).

Prenatal overgrowth, hypoglycemia, abdominal wall defects and a high frequency of tumors are all characteristic of the Beckwith-Wiedemann Syndrome (BWS). Being a rare familial disorder caused by epigenetic changes in several genes, it is a true paradigm for understanding the epigenetics of cancer. Large-scale screenings of affected families identified that LOI at the IGF2 gene is one of the main determinants of this disease. Normally, the maternal IGF2 allele is methylated, restricting expression to the paternal allele, but in BWS it is expressed from both paternal and maternal alleles, leading to a doubling of IGF2 expression. LOI at the maternal locus was also observed to be 5-fold more frequent in colorectal cancer patients showing that LOI of IGF2 confers a general cancer risk, even though this epigenetic change accounted for only 15% of all BWS patients (88;100).

Various pathologies have been linked to a disruption of the epigenetic state, often referred to as epimutations. Two imprinted-gene disorders associated with mental retardation, Prader-Willy Syndrome and Angelman Syndrome, involve two adjacent reciprocally imprinted genes. Usually, these syndromes are associated with mutations in this locus, located on chromosome 15, but some patients have been identified where an aberrant DNA methylation at the same locus causes the disease (100). This epimutation is the result of an allele that has passed through the male germline without proper clearing of the silent
epigenetic state that was previously established in the patients’ grandmother, indicative of TEI (101).

Intriguingly, a controversial association has been found between the occurrence of epigenetic disorders (e.g., BWS) and assisted reproduction techniques, such as in vitro fertilization (102;103). However, the number of reported cases is too low to draw firm conclusions and it has not yet been established how this increased incidence of imprinting disorders is caused. Aberrant epigenetic signals may arise by in vitro manipulation of embryos or the unnatural harvesting of egg and sperm cells. For example, embryo culture medium composition has been shown to affect DNA methylation (104) and strongly modifies the placental expression profile long after embryo manipulations, showing that stress of an artificial environment is memorized after implantation (105). Alternatively, abnormalities could already have been present in the parents, leading to reduced fertility and subsequent birth defects. In this regard, an association between reduced sperm concentration and abnormal imprinting in spermatozoa has provided a link between imprinting defects and impaired gametogenesis (106). Therefore, it has been hypothesized that spermatozoa from infertile males, used for in vitro fertilization, contain a higher number of gametes with chromosomal abnormalities and that a similar mechanism might account for infertility in women.

Maternal nutritional imbalance and metabolic disturbances that take place during early development can have profound effects on the health of offspring and may even be transmitted to the next generation. This theory is called “fetal programming” and many common diseases such as obesity, diabetes, cardiovascular disorders, cancer, asthma and even schizophrenia, take root in nutrition and environmental effects during early embryonic development (107). The epigenome is most vulnerable to environmental factors during embryogenesis, when DNA synthesis rates are high and the de novo DNA methylation pattern is shaped for normal tissue development (89). Moreover, oocytes have a more open chromatin structure and a longer life time than male germ cells, making oocytes more vulnerable to epigenetic modifications by environmental factors (87). In addition, sperm nuclei are much more likely to be cleared of any epigenetic errors because their histones are largely replaced by protamines (108).

THE PROMISE OF EPIGENETIC DRUGS

Epigenetic diseases can (in part) be prevented by adopting a healthy lifestyle. In addition, since epigenetic modifications are reversible, they can potentially be reversed by targeting chromatin-modifying enzymes. Some epigenetic drugs have already been approved by the FDA and are being used clinically. Decitabine or 5-aza-2’-deoxycytidine is an example of a DNMT inhibitor that is currently used to treat myelodysplastic syndrome (MDS), a bone
disease that can lead to leukemia (109). The drug is a nucleoside analogue with a modified cytosine ring that is incorporated in the DNA of replicating cells and therefore preferentially targets rapidly growing cancer cells. The drug stoichiometrically binds to and thereby inhibits DNMTs, counteracting hypermethylation of eg. tumor suppressor genes (110). Low drug doses are needed to limit toxic side effects and have been shown to increase overall survival in patients with MDS and prolong the time of conversion from MDS to leukemia. However, being incorporated in the genome, decitabine can cause mutations in daughter cells if the parental cell does not die. Another problem is that decitabine is toxic to the bone marrow, where blood cells are constantly being synthesized.

Two HDAC inhibitors, Vorinostat and Romidepsin, are remarkably efficient for treating cutaneous T-cell lymphoma (96). However, the precise molecular mechanisms for patient response still need to be determined. Moreover, the HDAC inhibitor trichostatin-A impairs blood vessel formation by indirectly repressing vascular endothelial growth factor, a key regulator of tumor vascularization (111). Another possible use of HDAC inhibitors lies in their potential to overcome resistance of cancer to conventional therapies. Indeed, multiple HDAC inhibitors reverse therapeutic resistance in a subpopulation of cancer stem cells in culture. Also the combination of DNMTs and HDAC inhibitors is being clinically tested as treatment for many different types of cancer. Epigenetic cancer therapy can thus be seen as a promising approach because of its synergy with conventional chemotherapy, reversing tumor chemoresistance and increasing the efficiency of radiotherapy (112). Finally, there is a growing interest in therapies targeted against PcG proteins, such as EZH2, that are overexpressed in many cancer samples. Recently, 3-deazaneplanocin-A was reported to rather selectively inhibit EZH2 and thereby decrease repressive methylation of H3K27 as well as reactivate genes that became aberrantly silenced in cancer (113).

Other epigenetic drugs rather downregulate the expression of specific sets of genes. For example, patients with Chronic Obstructive Pulmonary Disease have increased histone acetylation levels on genes involved in inflammation. Treatment of this disease includes the use of corticosteroids that are thought to act, in part, by promoting the recruitment of HDAC2 to the promoter of the active inflammatory genes. Conversely, the antidepressant Imipramine is known to induce acetylation of histones and reverse depression-induced repressive chromatin signatures. In rats, Imipramine inhibits HDAC5 in the hippocampus (114).

Another promising epigenetic curative tool is the use of induced pluripotent stem cells (iPSCs). Stem cells are cells with the potential to differentiate into any kind of tissue. Therefore, they are very attractive for the generation of replacement tissues for a wide range of disease conditions. In theory, stem cells
from the patient can then be used to regenerate his/her own damaged tissue, bypassing the problem of finding suitable donors. However, natural sources of stem cells are limited and so bioengineered iPSCs are rapidly gaining interest as a new tool for regenerative medicine. Patient-specific iPSCs are differentiated somatic cells that are experimentally reprogrammed to resemble embryonic stem cells, typically by introducing stem cell transcription factors that will reprogram the cell (114;115). iPSCs can then give rise to any cell type given the right mix of differentiation stimuli. So far, iPSC technology has demonstrated therapeutic benefits by generating patient-specific replacement tissues for a wide-spectrum of disease conditions. Upon transplantation in model systems, it was shown beneficial for the treatment of sickle cell anemia, ischemic heart disease, Parkinson’s disease and hemophilia. However, more work is needed to fully understand the molecular mechanisms underlying iPSC formation and differentiation. For example, failure to completely reverse repressive epigenetic modifications might limit the potential of iPSC formation. Moreover, it was reported that iPSCs contain epigenetic marks that resemble those found in cancer cells. Therefore, iPSC need to be better characterized before they can be used therapeutically (114;115).

CONCLUSIONS AND FUTURE PERSPECTIVES

The importance of a correct epigenetic regulation is highlighted by its key functions in both embryonic development and the etiology of many pathologies. Chromatin modifying enzymes are vital to establish a correct lineage commitment during terminal differentiation of stem cells in the embryo. Both lifestyle choices and environmental factors can influence the epigenome and deregulation of epigenetic processes can be the cause of many diseases. It has also become clear that epigenetic changes that accumulate during the lifespan of an organism can be passed on to future generations. However, our knowledge about the epigenome is still in its infancy and a greater understanding of the function and regulation of epigenetics is needed to identify attractive drug targets and design novel therapeutic strategies. Deep-sequencing technologies (Box 1) have been developed that are capable of mapping the epigenomic landscape in high resolution and have already been applied to identify DNA methylation markers for cancer patient diagnosis. Using this technology to map combinations of epigenetic markers in patient samples will provide a greater understanding of gene regulation in disease epigenomics and identify the factors that mediate changes to important loci. This will subsequently translate into the development of more successful disease treatments (2;116;117).
ACKNOWLEDGEMENTS

N.M. is a research fellow of the FWO-Flanders.

LIST OF REFERENCES

22. Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G. DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied


89. Faulk C, Dolinoy DC. Timing is everything: the when and how of environmentally induced


