Epigenetics, an emerging discipline with broad implications

Robert Feil

Institute of Molecular Genetics, CNRS UMR-5535 and University of Montpellier, 1919, route de Mende, 34293 Montpellier, France

Abstract

The field of epigenetics is young and quickly expanding. During the last year alone, thousands of research articles considered epigenetic mechanisms and their phenotypic consequences in different animal and plant species. Various definitions have been given, though, as to what precisely is epigenetics. Recent ones take into consideration that chromatin at genes and chromosomal regions can be structurally organised by covalent modifications and nuclear proteins, and via RNA molecules, in order to achieve defined expression states that can be perpetuated. Such somatically and meiotically heritable effects on gene function have diverse biological and medical implications. In particular, they are known to be important in development. A recent discussion meeting in Paris at the French Academy of Sciences reviewed our current understanding of ‘Epigenetics and Cellular Memory’ and where this novel discipline in life sciences is heading.

Keywords: Epigenetic; Development; Polycomb; Trithorax; DNA methylation; Histone methylation; Genomic imprinting; X-inactivation; Paramutation

1. Importance of epigenetic mechanisms

Although used much in recent years, the term ‘epigenetic’ is not new and has had different meanings over time. The word was originally used in the context of the development of the embryo. In the middle of the last century, Conrad Waddington applied it to describe the way genetic information determines how phenotypes arise during development, and to which extent this is a stochastic process [1,2]. Based on elegant work in flies, for instance, he determined that the expression of the genotype can be influenced by the environment and that the novel phenotypes obtained were sometimes heritably transmitted to the next generation. Waddington put forward theories on the genetic basis of such acquired developmental phenotypes, some of which were recently discovered to be linked to specific molecular processes [3]. The contemporary conception of epigenetics is a different one and puts much less emphasis on the genotype. In 1996, at a Cold Spring Harbor Conference on Quantitative Biology, an appealing definition was proposed by Arthur Riggs, Rob Martienssen and others in a book dedicated to Barbara McClintock, a pioneer in the field who studied specific phenotypes in maize to show that these were non-genetic in origin and were transmitted through meiosis [4]. The working definition proposed by Riggs and colleagues says that epigenetics is “the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence” [5]. Importantly,
this explanation puts the emphasis on the fact that epigenetic phenomena cannot be explained by genetic mutations, but are caused by `epimutations’, changes that occur at a level other than the DNA sequence. Lately, it is felt by some in the field that the emphasis on the heritability of gene function provides too stringent a definition of epigenetics and does not take into account the more transient chromosomal changes which contribute to the observed phenotypes. Adrian Bird (Institute of Cell Biology, Edinburgh), who was one of the speakers at the Paris conference, therefore proposed to describe epigenetic events as `the structural adaptation of chromosomal regions so as to register, signal, or perpetuate altered activity states’ [6].

Notwithstanding the need to carefully define what is `epigenetics’, extensive work during the last years has started to unravel the non-genetic changes that cause some of the phenomena. In animals and plants, for instance, DNA methylation is frequently involved in the acquisition, recognition and maintenance of novel phenotypes [7]. It plays also essential roles in embryonic development, for example, by shutting down genes that should not be expressed in specific cell lineages or tissues. For instance, recent work in mammals shows that many of the genes which are expressed specifically in the testis are silenced by heritable DNA methylation in all of the somatic cell lineages [8]. Interestingly, DNA methylation at specific loci can be recognised by proteins that bind to methylated cytosine residues. Such protein binding can affect the organisation of the associated chromatin. Thus, modifications at one level, in this case methylation on the genomic DNA, may have pronounced effects at other levels of organisation of the chromatin, a theme of growing importance in the field [9].

Another epigenetic system in eukaryotes concerns the Polycomb group and Trithorax group of proteins, which are involved in the maintenance of repressed and active transcriptional states, respectively [10]. The names of these two groups of chromatin modifying proteins originate from developmental phenotypes observed in the fruitfly Drosophila melanogaster. Specifically, when there are defects in the opposing actions of the Polycomb group and Trithorax group proteins, homeotic genes that specify the different segments of the fly do not maintain their expression states between somatic cell generations as they should, leading to specific developmental phenotypes [11]. Importantly, the regulatory actions of Polycomb and Trithorax group proteins are conserved among different groups of animals and plants. Both regulatory systems act by putting specific covalent modifications onto the nucleosomal histones around which the DNA is packaged in the nucleus. Protein complexes comprising Polycomb group proteins mediate and recognise methylation of a lysine residue at position 27 on histone H3, a modification which is generally associated with gene repression. On the contrary, Trithorax group proteins mediate and recognise methylation on lysine-4 of histone H3, a covalent modification generally associated with transcriptional active states.

A third level of epigenetic control involves RNA molecules. In mammals, `dosage compensation’ is a process which compensates for the presence of two X chromosomes in females versus one in males [12]. In mammals, to compensate the expression levels of the genes on the X chromosome between females and males, one of the two X chromosomes is inactivated in females. This multi-step process occurs during early embryonic development and is called X chromosome inactivation. It involves a large non-coding RNA molecule produced from one of the two X chromosomes, called Xist, which induces the inactivation of the chromosome [13]. The allelic repression of autosomal genes by a phenomenon called genomic imprinting (see below) is in some cases controlled by non-coding RNAs as well. Also here, the non-coding RNA brings about covalent modifications on the chromatin at close-by chromosomal regions leading to a somatically heritable repression state [14]. Another epigenetic phenomenon in which RNA molecules play a role is called paramutation. This widespread mechanism was originally detected in the pea, but has been most extensively studied in maize [15]. More recently, it has been detected in mammalian species as well [16]. Paramutation can occur when there is the combination of two different copies of a specific gene sequence (allele) resulting in a heritable change in the expression of one of the two copies. This phenomenon may happen between homologous transgenes, or between transgenes and the homologous endogenous gene, and occurs not only at the endogenous position of the gene in the genome. Remarkably, the altered epigenetic state can sometimes be inherited through meiosis to the next generation. Recent studies on this intriguing phenomenon evoke, for certain cases, a model in which the interactions between homologous sequence elements (`homology sensing’) involves the production of small interfering RNAs (siRNAs) that mediate chromatin changes and thereby induce heritable gene repression [17].

The conference held at the Académie des sciences in Paris on 17 May 2008 captured the excitement in France about the recent progress in our understanding of epigenetic phenomena. The presentations were followed by
stimulating discussions, and emphasized that epigenetics is becoming a discipline in its own right. This was also clearly recognised by the members of the Académie des sciences who organised this successful event, Marcel Méchali from the Institute of Human Genetics in Montpellier, and Pascale Cossart and Moshe Yaniv from the Institut Pasteur in Paris. Below, I summarize the five lectures of the conference which covered several of the main themes in epigenetics.

2. An overview of epigenetics

After an introduction on the aims and context of the conference by Moshe Yaniv, a first lecture on the role and regulation Polycomb group (PcG) and Trithorax group (TrxG) proteins was presented by Giacomo Cavalli, from the Institute of Human Genetics in Montpellier. Giacomo Cavalli explained that PcG and TrxG protein complexes comprise many different components and that they are similarly organised in different species. The opposite actions of these two groups of protein complexes are evolutionarily conserved amongst different plant and animal species. The histone modifications that are mediated and recognised by PcG and TrxG proteins confer a ‘memory’ to the cell as to which chromosomal regions need to be maintained transcriptionally active, and which ones transcriptionally repressed [11,18]. Given the conservation of these essential systems, it is most interesting to compare findings between different model organisms. Most of the key discoveries have been made in the fly Drosophila melanogaster, though, the species in which the PcG and TrxG proteins were originally discovered based on the occurrence of specific developmental mutants. Giacomo Cavalli pointed out that in order to understand how these chromatin modifying complexes act, it should be important to know where precisely in the genome they bind to the chromatin. To address this question, his laboratory has been using large-scale chromatin immunoprecipitation (ChIP) approaches, followed by hybridisation of tiling micro-arrays with DNA purified from the precipitated chromatin fractions [19]. Data were presented to show that the sequence elements that are recognised by PcG proteins (i.e., the PcG responsive elements, PREs) are mostly confined to genes that code for transcription factors involved in development and cellular proliferation. Interestingly, similar binding profiles were observed in embryos and cultured differentiated cells. As concerns involvement in cellular proliferation, recent work was reviewed that suggests that some genes controlled by PcG proteins are prone to aberrant expression in different types of cancer. To genetically address this insight in Drosophila, the Cavalli group generated flies deficient for the PcG protein polyhomeotic (Ph) in which they have started to explore an interesting cancer phenotype in the eye.

The concept of dosage compensation was carefully introduced by Claire Rougeulle from the Pasteur Institute in Paris, taking mammalian X-chromosome inactivation as an example. The existence of X chromosome inactivation was hypothesised in 1961 by Mary Lyon based on her research on coat colour in cats and mice [20]. Claire Rougeulle explained that this mechanism is controlled by the X-linked gene Xist (X-inactive specific transcript), which produces a non-coding RNA that mediates the X chromosome inactivation. How, precisely, Xist RNA mediates chromatin silencing is not well understood, but this process involves specific histone modifications and acquisition of DNA methylation at the promoters of the repressed genes [13]. Once the inactive chromatin state has been achieved, its subsequent maintenance no longer requires Xist RNA, but necessitates other specific factors, for instance to maintain the DNA methylation at the silenced promoters [21].

The expression Xist RNA itself is highly complex. This non-coding RNA has to act on one of the two X chromosomes only, in female cells and not in male cells, and this, at the right time during development. This is one of the questions Claire Rougeulle’s group has been studying during the last years, with a particular emphasis on the involvement of another non-coding RNA, Tsix, which is transcribed in the anti-sense orientation relative to Xist. The initiation of X-inactivation depends on the coordinated expression of the sense/antisense pair Xist/Tsix and the group discovered that chromatin at the Xist promoter is controlled by the transcription of Tsix. When Tsix is transcribed through this region in embryonic stem cells, the Xist promoter acquires an inactive heterochromatin-like organisation. When Tsix expression becomes gradually reduced in differentiating stem cells, however, the Xist promoter acquires an open chromatin organisation, allowing the increased expression of Xist RNA that is required to mediate X-inactivation [22]. Recent work suggests that the formation of double-strand Xist-Tsix RNA could lead to the production of small RNAs. These could be involved in the chromatin repression at the Xist promoter region, via an RNA interference mechanism [23]. Claire Rougeulle emphasised the importance to now unravel which specific factors trigger the developmental switch leading to the inactivation of one of the two X chromosomes, a choice which is random in the embryo.

DNA methylation plays key roles in epigenetic events in mammals and is particularly relevant in the
context of human disease [9,24]. The latter was nicely illustrated by Adrian Bird, of the Institute of Cell Biology in Edinburgh. His laboratory has a long interest in mammalian proteins that bind to methylated cytosines in genomic DNA. One of these 5-methyl-cytosine binding proteins is MECP2, a transcriptional repressor which, when mutated in humans, leads to the neurological disorder Rett syndrome (RTT). Rett syndrome is an autism spectrum disorder in girls which is variable in its severity. The explanation for this variability is that the MECP2 gene is located on the X-chromosome and is subject to X-inactivation. In affected girls, therefore, when one of the two copies is mutated, due to random X-inactivation there will be mosaic expression of the remaining intact copy. Adrian Bird described several transgenic mouse models which recapitulate the pathological features of Rett syndrome. How, precisely, deficient expression of the protein gives rise to the neurological symptoms remains poorly understood. Several specific target sequences have been identified, though, including a gene which is important for mitochondrial function [25]. A significant question addressed by Adrian Bird is whether this neurological disease can be cured by re-establishing normal levels of MECP2. Using a transgenic mouse model, his group made the conceptually important discovery that the Rett phenotype can indeed be reversed, even in advanced states, when MeCP2 expression is reactivated [26].

Genomic imprinting is another epigenetic phenomenon which is controlled by DNA methylation. It was discovered following the observation that both a maternally and a paternally-derived genome are required for embryonic and postnatal mammalian development [27,28]. The requirement of both the parental genomes is due to imprinted genes. These unusual genes are expressed from one of their two alleles only, depending on the parental origin of the allele. Imprinting has evolved not only in mammals, but also in seed plants. In both phyla, germ cell derived methylation marks are involved [29]. In mammals, these oocyte- and sperm-derived methylation imprints are resistant to the global waves of DNA demethylation that occur following fertilisation of the egg, and bring about the allelic expression of imprinted genes during embryonic development [30]. So far, about hundred imprinted genes have been discovered in mice and humans. Many of these play key roles in development and cellular proliferation. Robert Feil, from the Institute of Molecular Genetics, Montpellier, presented recent studies which showed that also histone methylation is involved in genomic imprinting. The group discovered that, at a 900-kb imprinted domain on mouse chromosome 7, the silenced paternal chromosome is marked by repressive lysine-9 and lysine-27 methylation on histone H3. The H3 lysine-27 methylation was found to be controlled by the Polycomb repressive complex 2 [31]. The H3 lysine-9 methylation, on the other hand, is controlled by a lysine methyltransferase called G9a [32]. At another imprinted domain, H3 lysine-27 methylation is important in imprinted expression specifically in neurons [33]. How, precisely, germ-line derived imprints convey the parental allele-specific expression of imprinted genes during development, often in a tissue-specific manner, remains largely to be discovered.

For many years, it was unclear whether paramutation effects, as discovered in plants, existed in mammals as well [16]. A first glimpse into the occurrence of trans-allele cross-talk in mammals came from transgenic work in the mouse. For different mouse transgenes comprising imprinted gene sequences, it was found that they could induce aberrant DNA methylation and gene repression at the endogenous gene [34,35]. One of the clearest demonstrations of paramutation, however, is provided by recent research on the KIT receptor gene in the mouse. This was the topic of the last lecture, by Minoo Rassoulzadegan from the Sophia Antipolis University in Nice. Her research group created a null mutant of Kit that had a lacZ insertion and which gave rise to a white-tail phenotype in heterozygote animals. Unexpectedly, this specific phenotype was sometimes observed in offspring that were genetically wild-type, and had thus arisen in the complete absence of the mutated allele. The meiotic inheritance of the white-tail phenotype without transmission of the targeted allele could occur both through the male and female germ line [36]. Also in the genetically normal animals, however, the white-tail phenotype was caused by reduced expression of the Kit messenger RNA due to the accumulation of non-polyadenylated RNA molecules of aberrant sizes. After they saw that there was accumulation of abnormal Kit RNA in spermatozoa, Minoo Rassoulzadegan’s group went on to show that the white-tail phenotype could be induced simply by injecting the mutant Kit RNA into fertilised eggs. As explained during the lecture, this finding pinpoints a novel mode of inheritance in mammals linked to the zygotic transfer of RNA molecules. Could this be a more general mechanism, acting on other gene transcripts as well? Rassoulzadegan’s group tested this hypothesis for different genes including Cdk9, a key regulator of cardiac growth. Micro-injection of its regulating microRNA, MiR-1, led to strongly increased expression of the Cdk9 gene and, hence, cardiac hypertrophy [37]. Interestingly, the hereditary transmission of the cardiohypertrophy correlated with the presence of
MiR-1 in sperm nuclei. Therefore, also in this system there was an RNA-mediated epigenetic effect, the inheritance of which correlated with the transmission of RNA by germ cells. Minoo Rassoulzadegan hypothesised that this novel mode of inheritance could provide a paradigm for familial diseases whose segregation is not fully explained by Mendelian genetics. This intriguing idea needs testing, though, and it should also be interesting to unravel the molecular details of the underlying mechanism(s).

3. Future outlook

Not all themes could be included in this one-day event, but a longer international meeting on this topic is considered in the next years by the Académie des Sciences. One fundamental issue, for instance, is how histone modifications and DNA methylation patterns are set up at genes and chromosomal domains in embryonic cells and in the germ cells [38,39]. How important are these epigenetic patterns for cell fate in undifferentiated pluripotent cells, and in differentiating cells [40]? Furthermore, what are the developmental roles of histone variants, such as the H3 variant H3.3, which mark heritably active genes [41]? To address these key questions, future research will undoubtedly benefit from the novel technologies that allow mapping histone modifications and DNA methylation along chromosomes, and now even across entire genomes [42,43]. Thus-obtained ‘epigenetic landscapes’ can be explored by bio-informatics, an approach which recently led to novel insights in the developmental role of DNA methylation [44].

What determines the extent of the chromosomal domains that are marked by specific epigenetic patterns? The way DNA replication is orchestrated during development across domains in pluripotent versus differentiating cells may contribute to this differential expression [45]. Replication and transcriptional machineries could also provide potent forces to structurally organise chromosomal domains in the three-dimensional space of the nucleus [46–49]. During every cell cycle the chromatin organisation needs to be faithfully copied to the daughter cells so that the epigenetic landscapes are maintained. To which extent is this process linked to DNA replication [50]? What is the role of chromatin remodelling complexes in the control of chromatin modifications across genes and chromosomal domains, and which novel players are yet to be discovered [51]?

Recent reviews have emphasized the growing importance of epigenetic alterations, in particular of DNA methylation, in human pathologies [52,53]. Some epigenetic diseases are caused by changes in DNA methylation at specific DNA sequences. For instance, the aberrant foetal growth disorders Beckwith–Wiedemann syndrome (BWS) and Silver–Russell syndrome (SRS) are caused by somatic changes in methylation at regulatory sequences that control imprinted gene expression [54]. Epigenetic deregulation could also be consistent with various non-Mendelian features in complex diseases in which multiple genes are suspected to be involved. In a recent study, for instance, changes in DNA methylation profiles in the frontal cortex were found to be associated with major psychosis [55]. In terms of potential treatment, pathological epigenetic alterations would have the advantage over genetic mutations that, in principle, they could be reversed. Epigenetic reactivation of tumour suppressor genes that are repressed in cancer is being attempted as a therapeutic approach by treating patients with inhibitors of DNA methylation and chemicals that affect the action of histone deacetylases. Although these rather non-specific components act on the chromatin in general, and therefore generate undesirable side-effects, promising results have been obtained in certain types of cancer, particularly in myeloid leukaemia [56]. Future research may lead to the development of novel drugs with higher specificities that can correct aberrant epigenetic alterations, ideally in a gene-specific manner.

There is a growing interest in the possibility that environmentally-induced changes at levels other than the genetic information could have long-lasting phenotypic consequences. Particularly in plants, this phenomenon could readily create heritable phenotypic differences within species as a means to adapt to different environments. At the molecular level, the link between the environment and epigenetic modifications has remained largely unexplored. Nevertheless, recent publications indicate that environmental factors, diet and ageing can perturb the way genes are controlled by DNA methylation and histone modifications [57,58]. One striking example is provided by a study on monozygotic twins, in which aged twins were found to have accumulated differences in DNA methylation and histone acetylation at many gene loci [59]. Intriguingly, environmentally-induced alterations and their phenotypic consequences can sometimes be passed on to the next generation, at least in plant and animal models [60]. There is a growing awareness that, also in humans, environmental factors and diet influence the epigenetic organisation of genes and could thus have long-term phenotypic effects. In human assisted reproduction technologies, for instance, germ cells and early embryos are taken from their natural environment and are manipulated and cultured in vitro. These procedures are thought to give rise to increased frequencies of certain
developmental abnormalities, such as the Beckwith-Wiedemann syndrome. Fertility clinics should ensure that the in vitro conditions used have no or little effect on the epigenetic organisation of genes. This issue will not be easy to address, though, given the ethical and technical limitations associated with human assisted reproduction. Again here, novel molecular insights will likely be obtained from studies on animal models.

In conclusion, the field of epigenetics provides us with unprecedented opportunities to address some of the most fundamental questions in developmental biology and medicine. It has generated powerful novel technologies to explore to which extent gene expression and phenotype are heritably influenced by external factors, via non-genetic alterations. The latter should be relevant for understanding common metabolic and behavioural diseases, and cancer, in which susceptibility is strongly influenced by dietary and environmental factors [58]. Given the many initiatives in epigenetical research in France and other countries, it seems likely that the Académie des sciences will continue to follow with interest what happens in this emerging discipline.

Acknowledgements

I thank Marcel Méchali, Michael Weber and Philippe Arnaud for critical reading of the manuscript, and the ANR, ARC and INCa for grant support.

References
