

Editorial

The era of epigenetics

Historically, the term ‘epigenetics’ is attributed to Conrad Waddington (1905–75) who in the late 1930s remarked ‘It is, surely, obvious that the fertilized egg contains constituents which have definite properties which allow only a certain limited number of reactions to occur; in so far as this is true, one may say that development proceeds on a basis of the “preformed” qualities of the fertilized egg. But equally it is clear that the interaction of these constituents gives rise to new types of tissue and organ which were not present originally, and in so far development must be considered as “epigenetic”’ [1]. He further remarked that ‘One might say that the set of organizers and organizing relations to which a certain piece of tissue will be subject during development make up its “epigenetic constitution” or “epigenotype”; then the appearance of a particular organ is the product of the genotype and the epigenotype, reacting with the external environment’ [1].

However, the conceptual origins date back to the age of preformationism. In his book of ca. 350 BC (*On The Origins of Animals*) the Greek philosopher Aristotle proposed, as a rebuttal to preformationists, the notion of ‘epigenesis’ as the ‘development of individual organic form from unformed’ driven by a mechanism he termed a ‘vital cause’. That the adult form develops from the embryo through gradual steps (epigenesis), as opposed to being fully preformed in the zygote (preformation), was a debate that prevailed into the 20th century (for an excellent review see [2]). It was not until improvements in microscopy allowed visualization of embryos (and a failure to find evidence of a preformed body in the fertilized egg) and the advent of genetics in the early 20th century (and the distinction between genotype and phenotype) that the old controversies of centuries past were finally resolved. Thomas Hunt Morgan put it succinctly in 1910; ‘We have two factors determining characters: heredity and the modification during development’ [3]. The inherited preformed or predetermined genetic program provides information about what is possible, but

regulation of genetic expression involves interpretation. It is the latter that is epigenetic.

This union of genetics and developmental biology lead to Waddington proposing the concept of an epigenetic landscape in order to represent the processes of cellular decision making during development [4]. Since then the term epigenetic has gone through many refinements to reflect a broader biological focus than just development, and is now most commonly defined as ‘the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence’ [5]. However, more recent definitions proposed by Adrian Bird; ‘the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states’ [6] and Danny Reinberg; ‘those processes that ensure the inheritance of variation (-genetics) above and beyond (epi-) changes in the DNA sequence’ [7], convey more accurately the field of epigenetics in the 21st century. We now know that such processes and associated differential gene expression involve a complex interplay among transcription factors, chromatin regulators/remodelers, histone modifications, histone variants, DNA methylation and non-coding RNA molecules. Moreover, in addition to interpreting and directing the same genome to produce over 200 distinct cell types in mammals, epigenetic processes are involved in mono-allelic gene expression in genomic imprinting ([8] and references therein), X chromosome inactivation [9] and the transcriptional silencing of transposons [10].

The establishment, maintenance and modulation of such transcriptional programmes are reliant on the inherent ‘plasticity’ of chromatin. Nucleosomes, the fundamental repeating unit of chromatin, are composed of a multi-subunit complex of histones, around which 147 bp DNA is wrapped (which can be covalently modified at the position 5 of cytosine) [11]. While assembly of the genome into chromatin achieves the required DNA compaction to fit the genetic information into the nucleus, it paradoxically affects every DNA-based process including DNA

repair, DNA replication and gene transcription. For such processes to access the DNA sequence, chromatin is a dynamic structure.

In this special issue of *Briefings in Functional Genomics*, we provide an overview of epigenetic mechanisms, discuss state-of-the-art methodologies for interrogating epigenomes and describe the epigenomic landscape from fertilization through to the role of epigenetic changes in disease.

The establishment, maintenance and reversal of transcriptional programmes are fundamental to differentiation and specialization, but how does chromatin regulate these outputs? Eukaryotes have evolved multiple mechanisms to finely tune chromatin structure, a principle approach that involves the posttranslational modification (PTM) of histones. PTMs of histones are implicated in influencing gene expression and genome function by establishing and orchestrating DNA-based biological processes. PTMs, deposited by 'writer' enzymes, can either directly affect chromatin structure (via charge changes) or indirectly by recruiting specific effector molecules, referred to as 'readers'. With nearly 100 different modifications described and with advances in mass spectroscopy discovering more, the combinatorial complexity of these marks is vast. Annalisa Izzo and Robert Schneider provide a timely overview of how one type of modification, the covalent methylation of lysine and arginine residues in histones are established, removed, interpreted by effector molecules and how this establishes and orchestrates transcriptional outputs. It is becoming increasingly apparent that a variety of different types of histone PTMs can co-exist not only on a given histone, but also on different histones within the same nucleosome. Given the intricacy of histone PTMs and the myriad of possible combinations, such complexity increases the potential to encode or store epigenetic information. Moreover, the ability of different histone PTMs to 'talk' to each other in *cis* and in *trans* in order to coordinate and fine tune DNA-based processes is highlighted in this review.

It is apt given the historical context of epigenetics, that this special issue contains an overview by Adam Burton and Maria-Elena Torres-Padilla of specific chromatin changes during the earliest stages of development: from shortly after fertilization through the first mitotic division to the differentiation of specific cell types in the blastocyst. Postfertilization the highly specialised sperm and oocyte are

epigenetically reprogrammed in order to allow formation of the totipotent zygote. This requires not only global changes in DNA methylation but also dynamic changes in chromatin and transcription. Interestingly, maternal and paternal chromatin undergoes distinct programmes and are epigenetically marked very differently upon fertilization, resulting in an asymmetric distribution of histone PTMs. The importance of chromatin modifications in the establishment of subsequent differentiation pathways is highlighted by the developmental defects that occur in the absence of functional chromatin remodelling proteins. Moreover, understanding the mechanisms involved in reversing epigenetic profiles and then establishing different ones in order to specify differentiation programmes, may also provide important insights into stem cell biology. This is manifest in the ability to induce pluripotency in mouse embryonic fibroblasts by the overexpression of certain transcription factors [12] and the seminal experiments in the 1950s and 1960s by John Gurdon, where transplantation of the nucleus from a fully differentiated somatic cell into an enucleated egg resulted in the development of an adult frog [13].

In order to understand and delineate how epigenetic 'units' define the functionality and plasticity of chromatin requires a comprehensive, unbiased and high-resolution co-localization analysis of epigenetic modifications across sequence features within the human genome. For this reason, epigeneticists have been particularly interested in mapping the locations of nucleosomes carrying specific histone PTMs and cytosines within DNA that are methylated. Histone PTMs can be studied using a technique called chromatin immunoprecipitation (ChIP). The use of modification-specific antibodies in ChIP has revolutionized the ability to infer the biological function of histone PTMs. In ChIP, DNA and associated proteins are chemically cross-linked and the DNA is fragmented. Proteins cross-linked to DNA are then immunoprecipitated using an antibody specific to the protein of interest. Cross-links are then reversed and the associated DNA purified. Determining which DNAs are enriched in the sample reflects where in the genome a modified nucleosome was bound. A variation on the ChIP methodology has also been employed to study DNA methylation patterns by using an antibody specific to 5-methylcytosine. Historically, quantitative PCR has been employed to query if specific

regions of DNA were co-purified with the modification of interest. More recently, ChIP-enriched DNA has been combined with microarray platforms (ChIP-on-chip or ChIP-chip) in order to define *en masse*, thousands of *in vivo* binding sites of a number of factors and 'epigenetic' chromatin features.

While the ChIP-on-chip approach has proved to be productive for the genome-wide mapping of DNA-binding proteins and epigenetic modifications, the costs incurred for complete tiling arrays of the human genome, or the requirement for custom arrays, has meant that the use of genome-wide ChIP-on-chip has been limited. In the analysis of mammalian genomes, ChIP-on-chip has been mainly restricted to chromosome wide analyses or promoter-based arrays. Profiling the epigenome in an unbiased way using next-generation sequencing methodologies offers numerous and profound advantages compared with microarray analysis [14]. Martin Hirst and Marco Marra provide an excellent overview of current techniques and methodologies in profiling the epigenome using different next-generation sequencing platforms, from sample preparation to data analyses. Such techniques form the basis for the NIH Roadmap Epigenomics initiative (<http://www.roadmapepigenomics.org/>). Importantly, Hirst and Marra highlight the need for standardization of protocols, reagents and bioinformatical tools in order to produce high-quality, comprehensive epigenomic maps.

The lag between the development of data analysis tools and the speed with which next-generation sequencing technologies are advancing is already creating a data bottleneck for many users. This was equally true during the early days of microarrays but, with time and an ever-increasing user-base, bioinformatics and data analysis support was forthcoming. A number of bioinformatic tools are already available for routine but important computational aspects of data sets generated by next-generation sequencing including base-calling, data thresholding and reference genome alignment. Lelu *et al.* provide an overview of currently established methods in the analysis of ChIP and next-generation sequencing (ChIP-Seq) data sets. They describe techniques for not only extracting meaningful and accurate biological data but also for interpreting DNA sequence features from enriched genomic regions.

Finally, Christopher Bell and Stephan Beck explore the utility and application of integrated

epigenomic data sets in the context of disease. A number of diseases, most notably cancer, are characterized by altered epigenetic profiles. Furthermore, they discuss the interplay between the environment and the genome in the context of disease, and how this may be mediated by epigenetic mechanisms. Such epigenomic profiling is fundamental to understanding how alterations in the epigenome play a role in the susceptibility and pathogenesis of human disease.

So what does the future hold for epigenetics? The interaction of the environment with the genome, mediated through the epigenome, is of particular interest and biological significance. Recent studies in mammals have shown that maternal environment can lead to *in utero* transmission of epigenetic information, which can affect the offspring's phenotype [15]. Human monozygotic twin studies have revealed epigenomic differences due to non-shared environment and a possible explanation for some elements of phenotypic discordance [16]. Perhaps most fascinating of all, however, are studies emerging from the Hymenoptera, where a single genome can specify more than one distinct organism that display dramatically different physiologies, phenotypes and behaviors [17, 18]. The best-studied example is the honeybee (*Apis mellifera*), where environmental cues such as diet have been shown to dictate developmental trajectory through epigenetic mechanisms, resulting in the development of a reproductive queen or a sterile female worker bee [19]. Moreover, differential and distinct DNA methylation and transcriptional patterns are observed in the brains of queen and worker honeybees [20]. It remains to be seen what the significance of epigenetics is with respect to human behavior and whether information can be stored in neurons as epigenetic marks or signals. The era of epigenetics is upon us, and in the words of James D. Watson; 'What determines whether a given piece of DNA along the chromosome is functioning, since it's covered with the histones? You can inherit something beyond the DNA sequence. That's where the real excitement of genetics is now.' [21].

Paul J. Hurd
Queen Mary University of London,
School of Biological and Chemical Sciences,
Mile End Road,
London E1 4NS. UK

References

1. Waddington CH. *Introduction to Modern Genetics*. London: Allen and Unwin, 1939.
2. Mayr E. *This is Biology: The Science of the Living World*. Harvard: Harvard University Press, 1997.
3. Morgan TH. Chromosomes and heredity. *Am Nat* 1910;**64**: 449–96.
4. Waddington CH. *The Strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology*. London: Allen and Unwin, 1957.
5. Riggs AD, Martienssen RA, Russo VEA. Introduction. In: Russo VEA, Martienssen RA, Riggs AD (eds). *Epigenetic Mechanisms of Gene Regulation*. New York: Cold Spring Harbor Laboratory Press, 1996.
6. Bird A. Perceptions of epigenetics. *Nature* 2007;**447**:396–8.
7. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. *Science* 2010;**330**:612–6.
8. Murrell A. Beyond genomic imprinting. *Brief Funct Genomics* 2010;**9**:279–80.
9. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol* 2009;**21**:359–66.
10. Bestor TH, Bourch'his D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb Symp Quant Biol* 2004;**69**:381–7.
11. Luger K, Mäder AW, Richmond RK, *et al*. Crystal structure of the nucleosome at 2.8 Å resolution. *Nature* 1997;**389**: 251–60.
12. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006;**126**:663–76.
13. Gurdon JB, Uehlinger V. “Fertile” intestine nuclei. *Nature* 1966;**210**:1240–1.
14. Hurd PJ, Nelson CJ. Advantages of next-generation sequencing versus the microarray in epigenetic research. *Brief Funct Genomics Proteomic* 2009;**8**:174–83.
15. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 2007;**8**:253–62.
16. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 2010;**465**: 721–7.
17. Bonasio R, Zhang G, Ye C, *et al*. Genomic comparison of the ants *Camponotus folridanus* and *Harpegnathos saltator*. *Science* 2010;**329**:1068–71.
18. Maleszka R. Epigenetic integration of environmental and genomic signals in honeybees: the critical interplay of nutritional, brain and reproductive networks. *Epigenetics* 2008;**3**:188–92.
19. Kucharski R, Maleszka J, Foret S, *et al*. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 2008;**319**:1827–30.
20. Lyko F, Foret S, Kucharski R, *et al*. The honeybee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* 2010;**8**:e1000506.
21. Watson JD. Celebrating the genetic jubilee: a conversation with James D. Watson. Interviewed by John Rennie. *Sci Am* 2003;**288**:66–9.