REVIEW

Make way for the ‘next generation’: application and prospects for genome-wide, epigenome-specific technologies in endocrine research

Richard D Emes and William E Farrell¹

School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, College Road, Leicestershire LE12 5RD, UK

¹Human Disease and Genomics Group, School of Medicine, Institute of Science and Technology in Medicine, Keele University, Stoke on Trent, Staffordshire ST4 7QB, UK

(Correspondence should be addressed to W E Farrell; Email: w.e.farrell@keele.ac.uk)

Abstract

Epigenetic changes, which target DNA and associated histones, can be described as a pivotal mechanism of interaction between genes and the environment. The field of epigenomics aims to detect and interpret epigenetic modifications at the whole genome level. These approaches have the potential to increase resolution of epigenetic changes to the single base level in multiple disease states or across a population of individuals. Identification and comparison of the epigenomic landscape has challenged our understanding of the regulation of phenotype. Additionally, inclusion of these marks as biomarkers in the early detection or progression monitoring of disease is providing novel avenues for future biomedical research. Cells of the endocrine organs, which include pituitary, thyroid, thymus, pancreas ovary and testes, have been shown to be susceptible to epigenetic alteration, leading to both local and systemic changes often resulting in life-threatening metabolic disease. As with other cell types and populations, endocrine cells are susceptible to tumour development, which in turn may have resulted from aberration of epigenetic control. Techniques including high-throughput sequencing and array-based analysis to investigate these changes have rapidly emerged and are continually evolving. Here, we present a review of these methods and their promise to influence our studies on the epigenome for endocrine research and perhaps to uncover novel therapeutic options in disease states.

Journal of Molecular Endocrinology (2012) **49**, R19–R27

Introduction

Although The Human Genome Project, defining the sequences of bases that make up the human genome, was completed in 2003, more than 30 years have elapsed since the influence of DNA methylation on gene expression was first described (Holliday & Pugh 1975, Riggs 1975). This heritable epigenetic change targets CpG dinucleotides that together with post-replicative modification of histone proteins (Turner 1991, 1998) frequently leads to chromatin remodelling and thereby regulates access to the underlying genetic information by transcription factors (Fig. 1). Epigenetic remodelling plays essential roles in multiple pathways and systems that include development, physiological processes, X chromosome inactivation, genomic imprinting and tissue-specific gene expression (Razin & Cedar 1991, Li *et al.* 1993, Fraga 2009). These demands upon an essentially invariant genome are apparent as highly variable methylation patterns that together with the multi-combinatorial histone modifications in different

cell types comprise the epigenomic landscape. The characterisation of these modifications, to what is termed the methylome, presents us with considerable conceptual and methodological challenges for their comprehensive, genome-wide mapping in health and disease.

Detecting epigenetic change

DNA methylation

Initial efforts to characterise the epigenome began in the 1970s. For a comprehensive timeline of methods and applications, the reader is directed to recent excellent reviews (Baylin & Jones 2011, Harrison & Parle-McDermott 2011). However, it was not until 1992 that a now classic paper described the sodium bisulphite conversion technique, which when combined with conventional dideoxy sequencing allowed, for the first time, the analysis of DNA methylation patterns in genomic DNA (Frommer *et al.* 1992).

Figure 1 Schematic of epigenetic modifications associated with active and silent genes. Histones are shown as orange barrels, DNA (thick black lines) is wrapped around the histones to form nucleosomes. (A) An active gene with open, relaxed chromatin. DNA is unmethylated (yellow circles). H3K9Ac histone acetylation (red triangles) and H3K4Me (blue crosses) histone methylation are associated with open active chromatin. (B) A silent gene with closed chromatin structure. DNA is methylated (blue circles). Lysine 9 (H3K9Me3, red crosses) and/or lysine 27 (H3K27Me3, green crosses) trimethylation are associated with closed chromatin.

Improvements to the technique, emanating from the same laboratory, followed quickly, in this case through integration of a PCR amplification step, which increased assay sensitivity to some 10^4 -fold (Clark *et al.* 1994) and would spawn new methods and approaches to probe the epigenome. Although initial studies were confined to candidate genes, a plethora of other methods, many of which incorporated initial methylation-sensitive digestion and subsequent PCR amplification, quickly followed. These techniques would permit, for the first time, the whole genome analysis for identification of novel methylated genes (reviewed in Esteller (2007)). Indeed, our own studies exploiting methylation-sensitive digestion and subsequent PCR were successful at identifying novel and inappropriately methylated genes in pituitary adenomas (Bahar *et al.* 2004). However, a concern with these types of studies is that the majority of techniques rely on methylation-sensitive restriction digest, where incomplete digestion can confound the interpretation of derived data.

Histone modifications

The characterisation of epigenetic modification that targets histone tails has presented a still greater challenge for researchers (Esteller 2007). Commonly used techniques rely on antibodies that recognise specific histone tail modifications; however, there is some concern regarding their specificity (Esteller 2007). These analyses rely on chromatin immunoprecipitation (ChIP) of specific histone tail modification and subsequent quantitative PCR of enriched, gene-specific

DNA fragments. Antibody-mediated enrichment strategies, for either histone modification (ChIP) or DNA methylation, as example, methylated DNA immunoprecipitation (MeDIP), may also be used in combination with DNA tiling or CpG island promoter arrays (ChIP-Chip) for genome-wide analyses (reviewed in Harrison & Parle-McDermott (2011) and discussed in a subsequent section). However, a drawback of these techniques for the characterisation of change in DNA methylation is their inability to pinpoint DNA methylation at single-base-pair resolution (Beck & Rakyant 2008). Strategies to circumvent these limitations will be discussed in a subsequent section of this review.

Reversing epigenetic change

Indirect methods for detecting epigenetic changes have also been described. In these cases, strategies that either exploit knockout or knockdown of enzymes responsible for DNA methylation have been described. As an example, in endocrine pituitary cells, siRNA-mediated knockdown of a DNA methylase enzyme, *DNMT1*, has been used to identify novel methylated and silenced genes (Dudley *et al.* 2008). Other studies, also in pituitary cells, have used drug challenges of actively dividing cells that inhibit DNA methylation, histone deacetylation or a drug combination that targets both these modifications (Al-Azzawi *et al.* 2011, Yacqub-Usman *et al.* 2012). These types of studies, termed pharmacologic unmasking, are used in the candidate gene and genome-wide approaches, where the re-expressed transcripts are identified using transcript expression arrays (Dudley *et al.* 2009, Al-Azzawi *et al.* 2011, Yacqub-Usman *et al.* 2012). A caveat associated with these types of studies is that drug-induced re-expression or indeed silencing, as determined by transcript arrays, might be secondary to and a downstream consequence of the epigenetic silencing of an upstream gene (Al-Azzawi *et al.* 2011). Equally, the influence of microRNAs as epigenetic modulators of gene expression has become a recent focus for epigenetic research; however, they will not be part of this review.

Endocrine cells and epigenetics

In common with most other cell types, the developmental and physiological regulation of endocrine cells is influenced by genetic and epigenetic processes. The cell population where these changes are apparent includes, but are most likely not confined to, the pituitary, thyroid, thymus, pancreas, ovary and testes. The role of DNA methylation in these principal endocrine cell populations, in both health and disease, has been subject to recent extensive review

(Garcia-Carpizo *et al.* 2011). In these cases, aberration in epigenetic processes may result in profound local and systemic effects. The effects are frequently manifest as metabolic disease and/or syndromes and in some cases are consequent to a tumour within the particular cell population (Garcia-Carpizo *et al.* 2011). In addition to the specific genes subject to DNA methylation-mediated silencing reviewed by these authors, other studies on endocrine tumours have adopted genome-wide approaches to identify potential novel epigenetically silenced genes. These types of studies on testicular (Cheung *et al.* 2010) and ovarian cancer (Michaelson-Cohen *et al.* 2011) have utilised MeDIP enrichment and on pancreatic tumours and ovarian cancer have used pharmacologic unmasking approaches (Kang *et al.* 2010, Shimizu *et al.* 2011). Still other studies on pituitary and thyroid tumours have described drug-induced unmasking techniques of histone tail modification that impact upon gene expression in these tumour types (reviewed in Farrell (2006), Ezzat (2008) and Yacqub-Usman *et al.* (2012)).

Detecting epigenetic change at the whole genome level

The experimental approaches, techniques and advances (see above) thus far described highlight the critical role and significance of epigenetic modification in cellular processes that include development, homeostasis and disease. These types of modifications, manifest as methylation of CpG dinucleotides in DNA and as histone tail modification, are apparent in all cell types including those within specific endocrine cells and organs. Although many of the described techniques for the characterisation of epigenetic change(s) have made important contributions to our understanding of this phenomenon and its impact on gene expression, each

has its own inherent limitations. In particular, they do not provide us with a complete overview of the epigenomic landscape, that is the methylome, in either normal or disease states. However, the rapid advances made in next-generation sequencing (NGS) technologies (see Table 1; Metzker (2010) for a review of sequencing technologies), in some cases combined with the techniques or approaches already described, make it possible to map DNA cytosine methylation at single-base resolution. In addition, these technological advances also facilitate single-base sequence information of DNA-associated chromatin that is itself associated with specific modifications. We now direct our attention to these emerging technologies.

The evolution of methods to detect genome-wide epigenetic modifications has predominantly been to increase the resolution of detection. For methylation, this has progressed from candidate genes to regions of the genome to specific CpG dinucleotides to all methylated cytosines in a genome. The most appropriate method to use in each experiment is a balance of the sample size and resolution required. These methods are introduced below.

Chips with everything

One approach for the detection of modifications at the whole genome level is by the application of microarray or DNA chip technology. Histone modifications can be detected by immunoprecipitation and analysed by array. In these ChIP-Chip experiments (Fig. 2), protein is reversibly cross-linked to DNA by formaldehyde. Protein–DNA is then fragmented by one of the several techniques that include sonication, enzymatic digestion or micrococcal nuclease digestion. Each of these techniques requires optimisation to yield fragment sizes that are typically between 200 and 500 bp. In our laboratory, we typically extract chromatin from

Table 1 Common second- and third-generation sequencing methods used in epigenomics

Technology	Feature generation	Typical read quantity	Common applications in epigenomics and their limitations
454	Emulsion PCR, sequencing by synthesis	$1 \times 10^6 \times 400$ bp	MeDIP-Seq, MBD-Seq, RRBS, MethylC-Seq. Longer read length improves mapping in repetitive regions but higher cost than other platforms. High error in homopolymeric repeats
SOLiD	Emulsion PCR, sequencing by ligation	$6 \times 10^8 \times 50$ bp	MeDIP-Seq, MBD-Seq. Two base (colourspace) encoding provides internal error correction
Illumina genome analyzer	Bridge PCR, sequencing by ligation	$2 \times 10^7 \times 100$ bp	MeDIP-Seq, MBD-Seq, RRBS, MethylC-Seq and WGSBS. Extensive range of uses and technical support mean this is the most widely adopted method
GridION	Single molecule, detection by nanopore-based sensing chemistry	Length limited by fragment length	Potential to detect modified bases in real time
PacBio RS	Single molecule, sequencing by synthesis in real time	10 kb reads reported	Potential to detect modified bases in real time. High error rate

Figure 2 Identification of regions of histone modifications in the genome using immunoprecipitation coupled with arrays (ChIP-Chip) and next-generation sequencing (ChIP-Seq). In both methods, proteins of interest (green and red shapes) are cross-linked to DNA by addition of formaldehyde and DNA is fragmented. Target protein-bound DNA is enriched by immunoprecipitation with specific antibodies (green shapes). Formaldehyde cross-linking is reversed by heating and DNA is then either hybridised to an array (ChIP-Chip) or directly sequenced and mapped to a reference genome (ChIP-Seq).

~0.7 mg tissue or tumour cells. Before analyses, we perform time course experiments of fragmentation by micrococcal nuclease digestion. An example of digestion following the reversal of cross-linking is shown in Fig. 3. Once optimal digestion conditions are established, the specific protein DNA complexes are precipitated using antibodies to proteins or modifications of interest. Released DNA is then either directly labelled or subject to a prior whole genome amplification (WGA) step and hybridised in competition with WGA input (non-immunoprecipitated) DNA (Ren *et al.* 2000). Simplistically, when the amount of fluorescence from immunoprecipitated DNA is greater than the input DNA, the presence of the protein or modification of interest can be inferred. For histone modifications, an antibody specific to the modification of interest can be used or alternatively histone-modifying enzymes that are associated with chromatin can be targeted (Collas 2010). Typically, in these types of studies, antibodies directed to histone tail modification, which

include methylation and/or acetylation, are used. In these cases, particular modifications are associated with active or silent genes. The resolution of these approaches is determined by the specificity of the antibody and the resolution of probes on the array. These can be overlapping DNA fragments covering the entire genome creating so-called tiling arrays. However, these require multiple arrays to cover a large genome such as that of humans. Additionally, identification of binding regions does not determine protein binding to individual bases but rather of the region covered by the tile. To overcome these problems, focused arrays covering specific genomic regions of perceived biological interest (commonly predicted genes and promoter regions) are more frequently used.

Similarly, in this case, to identify regions of the genome that are methylated, the methylation-enriched DNA can be isolated by immunoprecipitation, following fragmentation of the genome, using antibodies specific to 5-methylcytosine (5mC; MeDIP) or enriched by immobilisation (capture) to methyl binding domain (MBD) proteins. For MeDIP, we typically use ~4 µg input DNA and enriched fragments are WGA labelled and hybridised to an array in

Figure 3 Gel analysis of micrococcal nuclease digestion: from left to right. Lane 1 shows DNA before chromatin digestion and lane 2 shows a 100 bp DNA ladder. For clarity, only the 500 bp marker is arrowed. Subsequent lanes show a time course (3, 7 and 15 min) of DNA post-chromatin digestion. Micrococcal digestions preceded reversal of chromatin cross-linking and are achieved before electrophoresis. Lanes 3 and 4 show predominant digest products between 500 and 200 bp whereas longer digests 15 min show smearing.

arrays on a variety of research areas (e.g. Banister *et al.* 2011, Cotton *et al.* 2011, Fackler *et al.* 2011, Fryer *et al.* 2011, Martino *et al.* 2011). Although these arrays, post-sodium bisulphite conversion, report on methylation, it is important to bear in mind that they examine a discrete number of CpGs within a region and not methylation of multiple, contiguous CpGs within, for example, a CpG island. In these cases, although methylation at a single CpG is reported to reflect the methylation of other CpGs within that region, this is not an invariant finding (W E Farrell, personal communication). Of equal importance in approaches that are reliant on sodium bisulphite conversion is the efficiency of the conversion reaction. Several commercial kits for conversion are available and appear to be reliable, and it is relatively easy to design and include conversion-efficiency controls. In our own laboratories, post-conversion, we include a cytosine base that is not followed by a guanine, that is it is not a constituent of a CpG dinucleotide. As these cytosines by default will not be methylated, all should be converted and read as a T in a sequencing reaction.

To overcome the limitations of ChIP-Chip, techniques that combine ChIP with NGS have been developed (ChIP-Seq). In ChIP-Seq, protein-bound DNA of interest is isolated as with ChIP-Chip but is then directly sequenced and mapped to a reference genome to provide greater resolution of protein–DNA interaction. Directly sequencing bound DNA avoids the problems associated with array-based technologies where only those regions captured on the array can be interrogated. Sequencing approaches also increase the dynamic range of detection as theoretically a single to millions of copies of immunoprecipitated DNA can be measured (Park 2009). ChIP-Seq marks a considerable advance for characterisation of DNA at single-base resolution. However, it is important to note that the derived sequence information is consequent to either the enrichment or the depletion of that sequence or fragment achieved in the initial immunoprecipitation.

For methylation research, the ideal goal would be to directly determine the methylation status of all cytosines in a genome. This is achievable by whole genome shotgun sequencing of bisulphite-converted (WGSBS or MethylC-Seq) DNA. MethylC-Seq of a complete genome was first reported in the plant *Arabidopsis* (Cokus *et al.* 2008). This was closely followed by the first report of the methylome at single-nucleotide resolution in human cells (Lister *et al.* 2009). Whilst this is undoubtedly a hugely powerful tool, a limitation of sequencing bisulphite-converted DNA is that the approach cannot distinguish between 5mC and

5-hydroxymethylcytosine (5hmC) modifications (Huang *et al.* 2010). 5HmC results from oxidation of 5mC and has been shown to be present in embryonic stem cells and Purkinje neurons and is of potential interest as a new epigenetic state that may influence gene expression and is also present as an intermediary between methylated and unmethylated states. A method to enrich and sequence 5hmC has recently been described (Song *et al.* 2011) and will be useful to discriminate 5hmC and 5mC at single-base resolution. In addition to these concerns, the number of sequence reads required to achieve recommended depth is beyond the budget and capabilities of many laboratories. For example, for the human genome, over 90 gb of mapped sequence would need to be achieved to fulfil the 30× coverage recommended by the NIH Roadmap Epigenomics Mapping Consortium guidelines (<http://www.roadmapepigenomics.org/protocols>). To overcome this limitation, multiple approaches for enrichment of the methylated fragment of the genome before NGS approaches have been developed. MeDIP sequencing (MeDIP-Seq (Jacinto *et al.* 2008)) and methylated DNA binding domain sequencing (MBD-Seq (Serre *et al.* 2010)) are the two techniques that are similar in principle. MeDIP-Seq uses antibodies raised against 5mC to enrich for methylated DNA, whereas MBD-Seq uses immobilised recombinant methylated CpG binding proteins (MBD and MBD2). The enriched fragment is then sequenced using NGS. Methylated regions of the

genome can then be determined by identifying regions of high-sequence density. These ‘peaks’ of sequence reads correspond to regions of methylation (Fig. 4). As MBD has been reported to not bind 5hmC (Frauer *et al.* 2011), this may provide an additional control to ensure that 5mC is specifically being assayed. In comparison of methods, Harris *et al.* showed that whilst the enrichment methods have the lowest cost per CpG covered, the number of CpGs covered was also low (Table 2). Additional limitations of these approaches are inability to call individual methylation events but rather the ‘average’ methylation status of each enriched fragment and that unless sequencing depth is sufficient, it is not possible to accurately determine whether a region with few sequence reads is unmethylated or simply has low coverage. It was also noted that both MeDIP-Seq and MBD-Seq tend to enrich for low CpG density regions with MeDIP more heavily biased in this manner (Harris *et al.* 2010). Alternate approaches have been developed to focus on higher density CpG regions such as CpG islands. The most widely used of these is reduced representation bisulphite sequencing (RRBS). This approach uses methylation-insensitive enzymes, e.g. *MspI*, which has a CCGG recognition sequence to digest at CG-rich regions, the fragments are then treated with bisulphite and sequenced allowing single-base resolution of methylation (Meissner *et al.* 2005). The original approach was modified to allow genome-scale analysis of samples (Gu *et al.* 2010). Benefits of the

Table 2 Comparison of resolution and coverage of methods discussed

	MethylC-Seq	MeDIP-Seq	MBD-Seq	RRBS	Infinium 27K	Infinium 450K
Resolution (bp)	1	100–1000	100–1000	1	1	1
Approximate theoretical coverage of genome-wide CpG (%)	100	100	100	10	0.1	1.6
Experimentally determined coverage of genome-wide CpGs (%)	14.73 ^a	0.09 ^a	1.77 ^a	0.37 ^a	0.08 ^b	1.6 ^c
Comments	Single base pair resolution and high coverage make this the gold standard. Lack of ability to determine that 5hmC methylation is possible concern	Good summary of genome-wide methylation	Good summary of genome-wide methylation. Ability to differentiate 5hmC and 5mC	Good summary of genome-wide methylation. Particularly focused to CpG islands	Good snap shot of methylation at key promoters. Reproducible results but lowest coverage of genome-wide CpGs. Largely superseded by Infinium 450K	Good snap shot of methylation at key positions. Reproducible results. Higher density of array increases coverage of genome-wide CpGs

^aHarris *et al.* (2010), probes identified with probability of detection <0.05.
^bFryer *et al.* (2011).
^cauthors' unpublished observations.

RRBS approach are the reduction of input DNA needed for the analysis, making it useful for clinical samples and the fact that the subset of the genome analysed is reproducible and can be generated *in silico* allowing mapping of sequence reads with greater ease.

With the range of methods available, deciding on the most appropriate method is not trivial. With different methods achieving optimal results for different targets of the genome, no approach will be appropriate in all situations. Therefore, integration of multiple approaches is recommended to achieve coverage over a range of cytosines in different densities (Harris *et al.* 2010). Additionally, limitations of the sequencing-based methods are the cost associated with data generation, size of the data produced and the interpretation of results obtained. As with other technologies, guiding principles for epigenetic experiments have been proposed (see www.roadmapepigenomics.org/files/protocols/). These guidelines suggest both the information regarding samples and experimental design required when reporting and also the recommendations for conducting experiments, e.g. for MethylC-Seq, a minimum of duplicate experiments is suggested and least 30× coverage (where each base of the genome is sequenced on average 30 times) of the genome when reads from the biological replicates is combined. For ChIP-Seq again, a minimum of biological replicates should be conducted and 20 million aligned 36 bp reads per replicate are recommended. This guideline raises an important point of read length and hence which NGS method is most appropriate to generate data (see Table 1). At the extremes of current technology are high number of short reads as generated with SOLiD (50 bp) in contrast to the relatively low number of long read lengths generated by 454 technology (400 bp). Perhaps, as it is a compromise between these extremes, the Illumina sequencing platform (100 bp) is currently the most widely used method. However, the choice of method will depend on the organism studied and the region of genome targeted. For a well-annotated and 'complete' genome, such human short reads from SOLiD will be mapped with greater efficiency than for a genome with a lower quality draft genome which may require longer reads. Whilst costs for sequencing are decreasing, these experiments remain a major investment for research groups and are, therefore, frequently preceded by exploratory research using fragmentation methods such as RRBS or Illumina arrays (only available for human) for methylation analysis. In addition to the costs of generating the raw data, the generation and storage of many gigabytes to terabytes of sequence read data may require considerations of infrastructure or improvement of computational power. For these reasons, we recommend that these experiments are best conducted in collaboration with bioinformaticians with experience in data handling and the methods described.

As with detection technologies, development of methods for the analysis of data is a very active field of research and beyond the scope of this review; for an excellent introduction, see recent reviews such as Laird (2010) and Krueger *et al.* (2012).

Whilst we have attempted to provide an overview of the current state of the art, the technologies described are continually evolving and improving. Recently, announcements from third-generation technologies such as PacBio RS from Pacific Biosciences and GridION from Oxford Nanopore Technologies have emerged, which promise to measure modifications on single DNA strands. The potential of these approaches will be of huge interest in the near future.

Summary and perspective

This review has presented an overview of the key advances in next-generation technologies applicable to investigation of the epigenome. It is clear that there has been, and will continue to be, advances in technological approaches for the characterisation of changes to the epigenomic landscape that are apparent in both health and disease. There is little doubt that the description of a chemical conversion method for the discrimination of the methylation status at cytosine bases in DNA represented the prerequisite forerunner to many of the next-generation technologies described in this review. As with many technological approaches (developed and developing), each has its own particular advantages and disadvantages. However, the genome-wide characterisation of a frequently dynamic epigenome as opposed to the essentially invariant genome presents significantly greater problems for their identification and characterisation that are further compounded by the size and volume of the derived data sets.

The impact of epigenetic modification on endocrine cells in health and disease and those apparent in endocrine-related syndromes has been subject to recent review (Zhang & Ho 2011). In their review, these authors reiterate and reinforce the general consensus opinion that epigenetic change is perhaps the pivotal mechanism of interaction between genes and the environment. Our knowledge of the plethora of inappropriate/aberrant epigenetic modifications in these types of cells or their target organs is far from complete. The prospects, therefore, for 'next generation' technologies in, for example, diabetes, osteoporosis, menopause, Cushing's syndrome, endocrine tumours and obesity will provide significant advances for disease management. In this context, and distinct from genetic aberration, epigenetic changes are potentially reversible and as such have potential as therapeutic drug targets.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the review reported.

Funding

This work was supported through Keele University Acorn funding (grant ID: KU/KY-U 1) to W E F and a Nottingham University new appointment funding award (grant award number: A13885) to R D E.

References

- Al-Azzawi H, Yacub-Usman K, Richardson A, Hofland LJ, Clayton RN & Farrell WE 2011 Reversal of endogenous dopamine receptor silencing in pituitary cells augments receptor-mediated apoptosis. *Endocrinology* **152** 364–373. (doi:10.1210/en.2010-0886)
- Bahar A, Simpson DJ, Cutty SJ, Bicknell JE, Hoban PR, Holley S, Mourtada-Maarabouni M, Williams GT, Clayton RN & Farrell WE 2004 Isolation and characterization of a novel pituitary tumor apoptosis gene. *Molecular Endocrinology* **18** 1827–1839. (doi:10.1210/me.2004-0087)
- Banister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA & Marsit CJ 2011 Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. *Epigenetics* **6** 920–927. (doi:10.4161/epi.6.7.16079)
- Baylin SB & Jones PA 2011 A decade of exploring the cancer epigenome – biological and translational implications. *Nature Reviews Cancer* **11** 726–734. (doi:10.1038/nrc3130)
- Beck S & Rakan VK 2008 The methylome: approaches for global DNA methylation profiling. *Trends in Genetics* **24** 231–237. (doi:10.1016/j.tig.2008.01.006)
- Bibikova M & Fan JB 2009 GoldenGate assay for DNA methylation profiling. *Methods in Molecular Biology* **507** 149–163.
- Bibikova M & Fan JB 2010 Genome-wide DNA methylation profiling. *Wiley Interdisciplinary Reviews Systems Biology and Medicine* **2** 210–223.
- Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL *et al.* 2011 High density DNA methylation array with single CpG site resolution. *Genomics* **98** 288–295. (doi:10.1016/j.jgeno.2011.07.007)
- Cheung HH, Lee TL, Davis AJ, Taft DH, Rennert OM & Chan WY 2010 Genome-wide DNA methylation profiling reveals novel epigenetically regulated genes and non-coding RNAs in human testicular cancer. *British Journal of Cancer* **102** 419–427. (doi:10.1038/sj.bjc.6605505)
- Clark SJ, Harrison J, Paul CL & Frommer M 1994 High sensitivity mapping of methylated cytosines. *Nucleic Acids Research* **22** 2990–2997. (doi:10.1093/nar/22.15.2990)
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M & Jacobsen SE 2008 Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452** 215–219. (doi:10.1038/nature06745)
- Collas P 2010 The current state of chromatin immunoprecipitation. *Molecular Biotechnology* **45** 87–100. (doi:10.1007/s12033-009-9239-8)
- Cotton AM, Lam L, Affleck JG, Wilson IM, Penaherrera MS, McFadden DE, Kabor MS, Lam WL, Robinson WP & Brown CJ 2011 Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. *Human Genetics* **130** 187–201. (doi:10.1007/s00439-011-1007-8)
- Dudley KJ, Revill K, Whitby P, Clayton RN & Farrell WE 2008 Genome-wide analysis in a murine Dnmt1 knockdown model identifies epigenetically silenced genes in primary human pituitary tumors. *Molecular Cancer Research* **6** 1567–1574. (doi:10.1158/1541-7786.MCR-08-0234)
- Dudley KJ, Revill K, Clayton RN & Farrell WE 2009 Pituitary tumours: all silent on the epigenetics front. *Journal of Molecular Endocrinology* **42** 461–468. (doi:10.1677/JME-09-0009)
- Esteller M 2007 Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Reviews Genetics* **8** 286–298. (doi:10.1038/nrg2005)
- Ezzat S 2008 Epigenetic control in pituitary tumors. *Endocrine Journal* **55** 951–957. (doi:10.1507/endocrj.K08E-082)
- Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF, Teo WW, Zhang Z, Huang P, Visvanathan K *et al.* 2011 Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. *Cancer Research* **71** 6195–6207. (doi:10.1158/0008-5472.CAN-11-1630)
- Farrell WE 2006 Pituitary tumours: findings from whole genome analyses. *Endocrine-Related Cancer* **13** 707–716. (doi:10.1677/erc.1.01131)
- Fraga MF 2009 Genetic and epigenetic regulation of aging. *Current Opinion in Immunology* **21** 446–453. (doi:10.1016/j.coi.2009.04.003)
- Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I & Leonhardt H 2011 Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE* **6** e21306. (doi:10.1371/journal.pone.0021306)
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL & Paul CL 1992 A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *PNAS* **89** 1827–1831. (doi:10.1073/pnas.89.5.1827)
- Fryer AA, Emes RD, Ismail KM, Haworth KE, Mein C, Carroll WD & Farrell WE 2011 Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. *Epigenetics* **6** 86–94. (doi:10.4161/epi.6.1.13392)
- Garcia-Carpizo V, Ruiz-Llorente L, Fraga M & Aranda A 2011 The growing role of gene methylation on endocrine function. *Journal of Molecular Endocrinology* **47** R75–R89. (doi:10.1530/JME-11-0059)
- Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, Gnirke A, Lander ES & Meissner A 2010 Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nature Methods* **7** 133–136. (doi:10.1038/nmeth.1414)
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y *et al.* 2010 Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology* **28** 1097–1105. (doi:10.1038/nbt.1682)
- Harrison A & Parle-McDermott A 2011 DNA methylation: a timeline of methods and applications. *Frontiers in Genetics* **2** 74. (doi:10.3389/fgene.2011.00074)
- Holliday R & Pugh JE 1975 DNA modification mechanisms and gene activity during development. *Science* **187** 226–232. (doi:10.1126/science.1111098)
- Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR & Rao A 2010 The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5** e8888. (doi:10.1371/journal.pone.0008888)
- Jacinto FV, Ballestar E & Esteller M 2008 Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome. *Biotechniques* **44** 35, 37, 39 passim. (doi:10.2144/000112708)
- Kang S, Dong SM & Park NH 2010 Frequent promoter hypermethylation of TGFBI in epithelial ovarian cancer. *Gynecologic Oncology* **118** 58–63. (doi:10.1016/j.ygyno.2010.03.025)
- Krueger F, Kreck B, Franke A & Andrews SR 2012 DNA methylome analysis using short bisulfite sequencing data. *Nature Methods* **9** 145–151. (doi:10.1038/nmeth.1828)
- Laird PW 2010 Principles and challenges of genomewide DNA methylation analysis. *Nature Reviews Genetics* **11** 191–203. (doi:10.1038/nrg2732)
- Li E, Beard C & Jaenisch R 1993 Role for DNA methylation in genomic imprinting. *Nature* **366** 362–365. (doi:10.1038/366362a0)

- Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM *et al.* 2009 Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462** 315–322. (doi:10.1038/nature08514)
- Martino DJ, Tulic MK, Gordon L, Hodder M, Richman T, Metcalfe J, Prescott SL & Saffery R 2011 Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics* **6** 1084–1094. (doi:10.4161/epi.6.9.16401)
- Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES & Jaenisch R 2005 Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research* **33** 5868–5877. (doi:10.1093/nar/gki901)
- Metzker ML 2010 Sequencing technologies – the next generation. *Nature Reviews Genetics* **11** 31–46. (doi:10.1038/nrg2626)
- Michaelson-Cohen R, Keshet I, Straussman R, Hecht M, Cedar H & Beller U 2011 Genome-wide *de novo* methylation in epithelial ovarian cancer. *International Journal of Gynecological Cancer* **21** 269–279. (doi:10.1097/IGC.0b013e31820e5cda)
- Park PJ 2009 ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics* **10** 669–680. (doi:10.1038/nrg2641)
- Razin A & Cedar H 1991 DNA methylation and gene expression. *Microbiological Reviews* **55** 451–458.
- Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E *et al.* 2000 Genome-wide location and function of DNA binding proteins. *Science* **290** 2306–2309. (doi:10.1126/science.290.5500.2306)
- Riggs AD 1975 X inactivation, differentiation, and DNA methylation. *Cytogenetics and Cell Genetics* **14** 9–25. (doi:10.1159/000130315)
- Serre D, Lee BH & Ting AH 2010 MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Research* **38** 391–399. (doi:10.1093/nar/gkp992)
- Shimizu H, Horii A, Sunamura M, Motoi F, Egawa S, Unno M & Fukushima S 2011 Identification of epigenetically silenced genes in human pancreatic cancer by a novel method “microarray coupled with methyl-CpG targeted transcriptional activation” (MeTA-array). *Biochemical and Biophysical Research Communications* **411** 162–167. (doi:10.1016/j.bbrc.2011.06.121)
- Song CX, Clark TA, Lu XY, Kislyuk A, Dai Q, Turner SW, He C & Korlach J 2011 Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nature Methods* **9** 75–77. (doi:10.1038/nmeth.1779)
- Turner BM 1991 Histone acetylation and control of gene expression. *Journal of Cell Science* **99** (Pt 1) 13–20.
- Turner BM 1998 Histone acetylation as an epigenetic determinant of long-term transcriptional competence. *Cellular and Molecular Life Sciences* **54** 21–31. (doi:10.1007/s000180050122)
- Yacqub-Usman K, Alan Richardson A, Duong CV, Clayton RN & Farrell WE 2012 The pituitary tumour epigenome: aberrations and prospects for targeted therapy. *Nature Reviews Endocrinology* (In press). (doi:10.1038/nrendo.2012.54)
- Zhang X & Ho SM 2011 Epigenetics meets endocrinology. *Journal of Molecular Endocrinology* **46** R11–R32. (doi:10.1677/JME-10-0053)

Received in final form 16 March 2012

Accepted 23 April 2012

Made available online as an Accepted Preprint 23 April 2012