



Epigenome: The Guide to Genomic Expression

5

Ajit Kumar and Gulshan Wadhwa

Abstract

Epigenetic research endeavours to empathize traditional gene regulation not under direct encoding of DNA sequence. Histone modifications, DNA methylation and binding of nonhistone proteins are well-identified mechanisms of epigenetic control of cellular phenotype by gene expression regulations. Environmental factors cause, wholly or partly, different human diseases. Environmental chemicals have long been accepted to cause many diseases through alterations in the genome or genetic effects. Epigenomics (i.e. beyond genomics) encompasses amalgamation of customary genomics with other branches of science like mathematics, computer science, biochemistry, chemistry, proteomics and molecular biology. It looks for the comprehensive analysis of heritable phenotypic changes, alterations in gene function/expression that are not independent of gene sequence. The epigenomic science offers and beckons novel opportunities to help and elevate our understanding of nuclear organization, regulation of transcription, developmental phenomena and diseases at molecular level. This article presents a comprehensive report about the existing computational strategies and approaches for studying the different factors of epigenetics, with special focus on important computational tools and biological databases. In addition, a brief introduction into epigenetics have also been outlined.

A. Kumar (✉)

Centre for Bioinformatics, Maharshi Dayanand University, Rohtak, India

G. Wadhwa

Department of Biotechnology, Apex Bioinformatics Centre, Ministry of Science & Technology, New Delhi, India

© Springer Nature Singapore Pte Ltd. 2018

G. Wadhwa et al. (eds.), *Current trends in Bioinformatics: An Insight*,

https://doi.org/10.1007/978-981-10-7483-7_5

89

Keywords

Computational epigenomics · DNA methylation · Epigenetics · Histone modification

5.1 Introduction

The term ‘epigenetic landscape’ was coined by Conrad Waddington and defined epigenetics as the synergistic interaction of genes with their products that bring out the phenotype into being. Epigenetics is the study of mitotically and meiotically heritable changes in gene function that does not bring out change in DNA sequence.

Genetically, it has been viewed for long that DNA expresses through genes that help us to survive, reproduce and develop, and it is all set in stone, i.e. this behaviour is settled in stone and the environment. Traditional genetics fails to explain many of our questions related to differential gene expressions in multicellular organisms.

The vertical flow of information, stored in DNA, is in general answered by classical genetics and genomics, but they fail to explain the control of horizontal flow of information within a cell, which has given rise to division of labour at multicellularity level, throughout the evolution of life. Epigenetics, as a name suggests, goes beyond the classical genetics and deals with the study about which genetic element is to be expressed along with where and when it is to be expressed. These questions are answered by the structural features of the genetic element and not by the sequence of the DNA as the later deals with what is to be expressed. To cut whole story short, we can say that it is epigenome – the real boss of DNA deciding its functionality. To detail the concept of epigenetics, we can take the example of identical twins. The identical twins have the characteristics like eye colour, hair colour, etc. depicted by the DNA sequences within their genome, and this is what we called genetics, but it is not like that way. Genetically identical twins can be very different; one might be normal while other may be suffering from any disease. This can’t be explained on the basis of genetics because the DNA is identical in both the twins. To look beyond and understand the regulation of gene expression that is not directly encoded in the DNA sequences is epigenetics (Bock and Lengauer 2007). In other words, it is the chemical modifications in a gene that is heritable and influences the activity of DNA but lies outside of genomic sequences and does not alter the gene expression. It actually muddles the nature and the nurture.

5.2 Epigenetic Machinery

Epigenetic mechanism involves a number of processes that are critical for diversity of cell type during embryogenesis and gametogenesis and also crucial for tissue-specific gene expression and global gene silencing. Crucial developmental process from

fertilization till the end of gastrulation phase can be altered and may result in structural abnormalities (Piplani et al. 2016). Epigenetic mechanisms encompass the mechanism that includes DNA and the nucleosomes and are known to be the carrier of this process (Espada and Esteller 2007; Rivera and Ren 2013). The three most common chemical and biological modifications that alter the genetic expressions are chromatin structure and modification, DNA methylation and miRNA.

5.2.1 DNA Methylation

DNA methylation is a crucial mechanism for cell differentiation and organism development. It alters the gene expression through methylation of DNA strand itself. This biochemical process involves the cytosine bases of eukaryotic DNA being converted into 5-methylcytosine and thus represses the transcription. This typically occurs at CpG island (Jones 2012; Virani et al. 2012).

This process thus passes on from generation to generation by the process of cell division. During zygote formation, DNA methylation is removed and re-established during cell divisions. The DNA methyltransferase (DNMT) enzymes are responsible for maintaining and establishing this unique pattern on CpG island (Dodge et al. 2002). Till date five DNMTs have been reported that are involved in catalysing this process (Roberts et al. 2003; Serman et al. 2006; Putiri and Robertson 2011) as summarized in Table 5.1.

5.2.2 Histone Modifications

These are another most common and complex epigenetic phenomenon that includes acetylation, phosphorylation, sumoylation, methylation and ubiquitination (Bannister and Kouzarides 2011). Histones are conserved proteins and can be modified at amino acid residues on N- and C- terminals. The C-terminus forms globular domains packed in core, while the N-terminal is comparatively flexible

Table 5.1 DNA methylating enzymes

Enzyme	Omim no.	Function
Dnmt1	126375	Maintain methyltransferase (Pradhan and Esteve 2003; Eastvaran et al. 2004)
Dnmt2	602478	Play role in DNA methylation but doesn't have DNA methyltransferase activity (Okano et al. 1998)
Dnmt3a	602769	Participate in de novo methylation (Gonzalzo Jones 1997; Okano et al. 1999)
Dnmt3b	602900	Participate in de novo methylation
Dnmt3L	606588	For methylation of most imprinted loci in germ cells (Kaneda et al. 2004)

Table 5.2 Histone modification enzymes

Enzyme	Function
Histone acetyltransferases (HATs)	Acetylase conserved lysine on histone protein by transfer acetyl group from acetyl Co-A to form ϵ -N-acetyllysine (Marmorstein 2001)
Histone deacetylase (HDAC)	Removes acetyl group from ϵ -N-acetyllysine on histone, allow histone to wrap DNA tightly (Song et al. 2005; Halkidou et al. 2004)
Histone methyltransferases (HMTs)	Catalyse the transfer of methyl groups to lysine and arginine of histone proteins(Campagna-Slater et al. 2011)
Histone demethylases (HDMs)	Removes methyl group from histone (Kawamura et al. 2010)

and interacts directly with DNA and other proteins within nucleus and is important to maintain chromatin stability. Unlike the rest of the modifications, the methylation and acetylation on arginine and lysine residues repress the transcription. These modifications are achieved by different histone-modifying enzymes as summarized in Table 5.2. The histone modifications alter the chromosome structure by altering the electrostatic properties or by producing binding affinities for protein recognition module (Iizuka and Smith 2003; Brooks and Shi 2014).

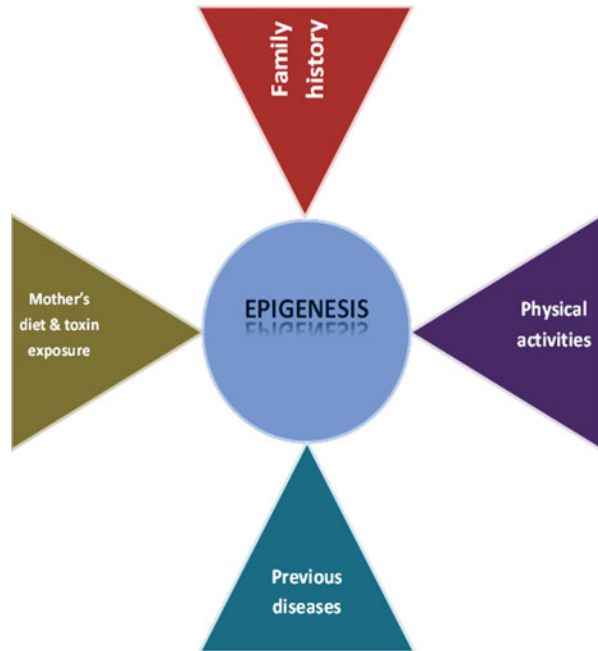
5.2.3 miRNA

miRNAs are ~22 nucleotide long non-coding RNAs that negatively control gene expression post transcriptionally. miRNAs are highly conserved in plant and animal species and play critical roles in a variety of biological processes including pattern formation and developmental timing, cell signalling, carcinogenesis, etc. (Sato et al. 2011). About 50% of miRNA genes are housed in the fragile genome regions and are very sensitive to deletion, translocation or duplication. miRNAs have already been looked upon as new therapeutic agents and/or targets for different diseases (Chuang and Jones 2007).

5.2.4 Binding of Nonhistone Proteins

The proteins in chromatin which remain even after the removal of histones are commonly known as nonhistone proteins. These proteins interact with the chromosome structure and remodel the chromatin structure, thus regulating the silencing of genes. The two nonhistone proteins, polycomb and trithorax, are epigenetic regulators and affect gene expressions. Polycomb protein induces gene silencing, while the trithorax protein induces gene inactivation (Pullirsch et al. 2010).

Fig. 5.1 External factors influencing epigenetic traits



5.2.5 Environmental Factors

Various dietary factors for mothers and environmental factors have been linked to epigenetic modifications. The conditions like stress, nutrition and environment deal with how our DNA behaves in present and next generation(s). Epigenetic mechanism regulates throughout the life. The nutrition of the mother affects the foetus; stress hormone also travels from mother to foetus. Social interactions, physical activities, exposure to toxins and diet are also the major factors shaping the epigenome (Kubota et al. 2012). The living place, consumption of alcohols and exposure to various drugs are potentially able to alter the changes in epigenetic status (Aguilera et al. 2010; Choi and Friso 2010). The external factors influencing epigenetic traits are depicted in Fig. 5.1.

5.3 Analysis of Epigenetic Modifications

5.3.1 In Vitro Methods

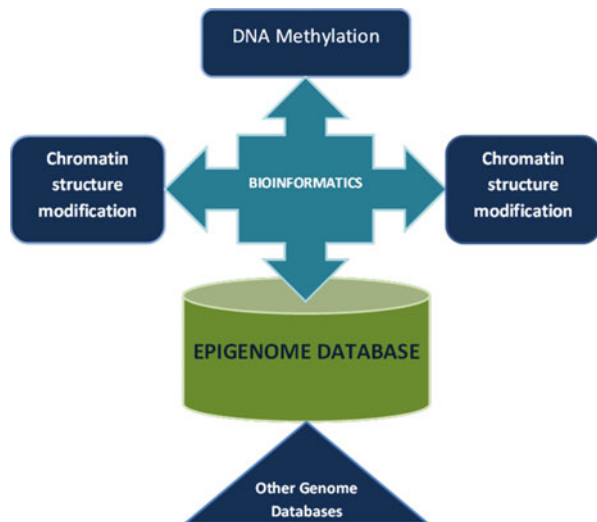
Various techniques are in extensive use for the interpretation of epigenetic data. Bisulphite sequencing method is used to determine the DNA methylation. This technique uses methylation restrictive enzymes that convert DNA with bisulphite to 5-methylcytosine (Jones and Takai 2001). Another commonly used method is

‘chromatin immunoprecipitation’ that determines the DNA binding sites on the genome for a particular protein (Collas 2010). It also predicts the DNA protein interactions, either inside the nucleus or within the living cell. The last decade witnessed rapid development of microarray-based technology. It analyses the bisulphite-treated DNA methylation sites (Schumacher et al. 2006) and utilizes the pairs of oligonucleotide hybridization probes for targeting the CpG sites.

5.3.2 In Silico Methods

With the developing technologies and advancement in the field of computational biology, several techniques have now developed to focus on identification of epigenetic modifications, such as Support Vector Machine (SVM), artificial neural networks (ANNs), hidden Markov model (HMM) and clustering analysis. SVM is used to diagnose the cytosine methylation in CpG nucleotides (Bhasin et al. 2005; Robinson et al. 2014). It is a successful machine learning technique to evaluate the pattern, but the problem is the lack of experimentally available public data. ANN performs the prediction algorithm for the human-specific methylation sites (de Pretis and Pelizzola 2014). This method is designed to mimic the architecture of the brain. The data that is needed to be evaluated is differentiated into processing units called neurons. These neurons process the data using a variety of mathematical evaluations (Marchevsky et al. 2004). Another widely used computational technique is HMM which is used to detect the CpG islands (Robinson and Pelizzola 2015). This method is deliberately used for the sequence analysis (Wu et al. 2010). Clustering analysis helps in reflecting the true distribution of gene space (Fazzari and Grealley 2004). Many reported attempts have been successfully made to analyse epigenetic modifications using in silico methods (Fig. 5.2) (Gitan et al. 2002; Collins et al. 2003;

Fig. 5.2 Integration of technology to generate epigenome database



Laird 2003; Yan et al. 2004; Meissner et al. 2005; Pfister et al. 2007; Petrossian and Clarke 2009).

5.4 Epigenomics: The Computational View of Epigenetics

Post-‘Human Genome Project’ era has witnessed a tsunami of epigenetic data due to rapid development in technological applications in biological sciences. Bioinformatics has evolved at a parallel rate to support, store, manage, manipulate and exploit the biological data flood. As a result, a new jargon has been added to bioinformatics vocabulary, namely, ‘epigenomics’. It deals in the computational analysis of experimentally determined epigenetic data and is emerging as a separate but prominent frontier of biological sciences. It involves in silico collection of data, maintaining databases and analysing the stored data in a computationally intensive manner, related to alterations of genetic expressions and gene activity independent of gene sequence. Epigenetics deals in stable alterations in genetic expressions and gene activity independent of gene sequence, both in the same generation (horizontal flow) and long-term changes in transcriptional potential of a cell that are generally not heritable. Epigenomics differs from epigenetics in emphasizing on global analyses of sequence-independent genetic changes throughout the genome, while the latter looks for studying the same for a gene or a gene set (Tammen et al. 2013). There have been many epigenomic efforts being carried out worldwide of medium/large scale; a major few are as listed in Table 5.3.

Various bioinformatics methods help in identifying gene expression by studying the contribution of DNA methylation, identifying the CpG islands and studying the in silico modelling and dynamics of epigenetic processes. Above all, the most important is the epigenomic data collection for solving the mystery. Algorithms like BLAST (Altschul et al. 1990), BLAT (Kent 2002) and Clustal W (Thompson et al. 1994) allow for the sequence analysis for inference of functional, structural and evolutionary relationships. The scientific literature database like PubMed and molecular databases like DDBJ, EMBL and GenBANK serves as repositories for nucleotide and protein sequence of different species. The databases, like GEO, GENSAT, StemBase, etc., allow for the identification of dynamic changes in gene expression in different cell types. The epigenomic data is getting enriched every day; a few important resources are summarized in Table 5.4.

5.5 Recent Insights into Disease Epigenomics

The epigenomics supersedes epigenetic studies in the manner that the former has comprehensive approach as compared to single-gene association to a disease in epigenetics (Feinberg 2010). Candidate gene-disease association studies have now been replaced by whole-genome disease association studies and are accepted with more scientific appreciations (Lieberman-Aiden et al. 2009). Methylation of cytosine residues of DNA has been attributed as the central stage player in epigenomic

Table 5.3 Large- and medium-scale epigenomic efforts

Projects	Aim	Web address
European epigenomic project	Epigenome Network of Excellence fosters the epigenetic research community in Europe, on technical aspects of epigenomics and development of tools	http://www.epigenome-noe.net/WWW/researchtools/protocols.php http://www.epigenome-noe.net/WWW/index.php
Canadian and Australian projects	The Canadian Institutes of Health and Research (Ottawa) and Australian Alliance for Epigenetics have come forward to elevate the understanding on 'epigenetics, environment and health'	http://www.epialliance.org.au/
Asian projects	Yonsei University (Seoul), the Japanese National Cancer Center, the Shanghai Cancer Institute and the Genome Institute of Singapore developed various technological platforms to generate large sequencing data and to promote interactions and collaborations in epigenomic research	
The ENCODE project	The Encyclopedia of DNA Elements (ENCODE) and modENCODE projects use a wide array of different assays to identify functional elements, and epigenomic profiling is thus an important component of the programmes	http://www.genome.gov/10005107
ICGC projects	International Cancer Genome Consortium (ICGC) investigating genomic changes with the goal of obtaining a comprehensive description of genomic, transcriptomic and	http://www.icgc.org/

	<p>epigenomic changes in 50 different tumour types and/or subtypes</p>	
<p>NIH roadmap epigenomic programme</p>	<p>The consortium is developing standards and best practices for individual and integrative analyses of the different data types to provide a reference for the larger epigenomic community. Another role is to include development of various new techniques of epigenomic analysis and imaging, identification of new epigenetic modifications and investigations</p>	<p>http://www.roadmapepigenomics.org</p>
<p>The International Human Epigenome Consortium</p>	<p>IHEC builds on the NIH effort in epigenomics and created a truly global epigenomic project. IHEC develops best practices and standards for epigenomic data generation and analysis</p>	<p>http://ihec-epigenomes.org</p>

Table 5.4 Information resources for epigenomics

Database	Information	URL
MethDB	Contains information on 19,905 DNA methylation content data and 5382 methylation patterns for 48 species, 1511 individuals, 198 tissues and cell lines and 79 phenotypes. Provides valuable search for patterns and profiles, 5mC content, phenotypes	http://www.methdb.de/
PubMeth	Contains over 5000 records on methylated genes in various cancer types	http://www.pubmeth.org
REBASE	Contains over 22,000 DNA methyltransferases genes derived from GenBank	http://rebase.neb.com/rebase/rebase.html
MeInfoText	Contains gene methylation information across 205 human cancer types	http://mit.lifescience.ntu.edu.tw/
MethPrimerDB	Contains 259 primer sets from human, mouse and rat for DNA methylation analysis	http://medgen.ugent.be/methprimerdb
The HistoneDB	Contains 254 sequences from histone H1, 383 from histone H2, 311 from histone H2B, 1043 from histone H3 and 198 from histone H4, altogether representing at least 857 species	http://research.nhgri.nih.gov/histones/
ChromDB	Contains 9341 chromatin-associated proteins, including RNAi-associated proteins, for a broad range of organisms	http://www.chromdb.org/
CREMOFAC	Contains 1725 redundant and 720 nonredundant chromatin-remodelling factor sequences in eukaryotes	http://www.jncasr.ac.in/cremofac/
CAMH: The Krembil Family Epigenetics Laboratory	Contains DNA methylation data of human chromosomes 21, 22, male germ cells and DNA methylation profiles in monozygotic and dizygotic twins	http://www.camh.ca/en/
MethylomeDB	Experimental information from over 14,000 entries and 175 high-throughput data sets from a wide number of sources also includes gene-centric methylation data of 72 human diseases	http://bioinfo.hrbmu.edu.cn/diseasemeth

modification and plays a pivotal role in cellular processes including regulation of genes, organism development and disease (Lister et al. 2009). The recent studies on different cancer-associated genes and its epigenetic machinery have revealed the requisites of epigenomic correlation search for unexplained diseases (Frigola et al. 2006; Irizarry et al. 2009; Fraga et al. 2005a, b; Lister et al. 2009; Doi et al. 2009).

Gene-silencing events were observed to be spanning through a large regions of genome in colorectal cancer. The DNA-methylated and adjoining un-methylated

genes were also observed to be coordinately suppressed across the entire chromosome (Frigola et al. 2006). Irizarry et al. (2009) showed that in colon cancer, the epigenetic alterations due to DNA methylation occur in promoters and CpG islands, extending to sequences up to 2Kb distance, termed as 'CpG island shores'. With CpG island hypermethylation, the global hypomethylation across genome has also been a common epigenomic character of cancerous cells (Irizarry et al. 2009). Tumour cells in humans have been attributed with the hallmark of mono-acetylation loss and tri-methylation of H4 histone proteins at global genomic level (Fraga et al. 2005a). A widespread epigenomic difference has been attributed responsible for differential susceptibility to diseases and other anthropomorphic variations observed in the lifetime of monozygotic twins (Fraga et al. 2005b).

While mapping whole genome at single-base resolution, significant differences were observed in pattern and composition of cytosine methylation of embryonic and differentiated cells (Lister et al. 2009). About one-fourth part of total genomic methylations of embryonic stem cells have been identified in a non-CG context reflecting to other methylation mode to regulate gene. Higher degree of non-CG methylations were observed in gene-coding areas, while the same were observed to be depleted in binding sites of proteins and regulatory regions (Lister et al. 2009). There has been suggestion about different mechanisms of epigenetic reprogramming resulting in differential methylation of tissue and disease-specific CpG islands in differentiated cells (fibroblasts), pluripotent stem cells and embryonic stem cells (Doi et al. 2009).

The role of epigenetic mechanism in altered embryonic developments and its manifestations in adulthood diseases, like cardiovascular diseases, is still unknown, but there is a growing volume of evidences supporting the epigenetic regulations (Martinez et al. 2015). Since the early 1980s, several studies have indicated the possible involvement of adverse intrauterine condition to adulthood diseases like cardiovascular diseases (Barker and Osmond 1988), diabetes (type 2), metabolic syndrome, ischemic heart disease, hypertension, etc. (Chen and Zhang 2011). These studies formed the basis of 'thrifty phenotype hypothesis' (Hales and Baker 2001) stating the abnormal intrauterine situations like toxins, hypoxia, undernutrition, chemicals, etc., pushes the developing embryo to undergo irreversible changes to adapt and survive the suboptimal environment, resulting in increased susceptibility of neonate or adult towards developing a disease.

Since the completion of 'Human Genome Project' in the early twenty-first century and next-generation sequencing technology, there has been a flood of genomic and epigenomic data. These developments have resulted in a paradigm shift in our concepts of cell and cellular function. The idea of reprogramming a cell in form of stem cell development has triggered a race for understanding the cellular function beyond genomics, i.e. epigenomics. Several diseases have been and will be associated with epigenetic and epigenomic basis and needs to exploit more to develop therapeutic applications of the same (Kanherkar et al. 2014). After the development of induced pluripotent stem cells (Park et al. 2012), there have been several successful attempts to modulate them epigenetically for generating differential potentials. These have led to many device novel therapies and develop novel

disease models for neurodegenerative disorders, cardiovascular diseases, metabolic disorders (PCOS), etc. (Huang and Wu 2013).

5.6 Conclusion

Epigenetic modifications provide a link between nature and nurture. The epigenomics paves the way through a new research that is valuable for predicting and diagnosing various diseases. Epigenomics has the potential to overthrow the disease-causing genes and identify the altered gene expression. Recent research by American Association for Cancer Research (AACR) offers the epigenetic therapy of killing cancer cells by methylating without disrupting its pathway (Jones 2012). The epigenetic tags and their response make it a valuable technology. Various other stocks are in queue for targeting different tissues affected by the disease. The epigenetic data adds a flavour to both the computational- and wet-lab work, and we can say ‘epigenomics: an era beckoning’.

5.7 Future Directions

The initial phase of computational epigenetic research got impetus from the ever-progressing experimental ways of data generation. The high-throughput epigenomic data thus generated need proper computational analyses for low-level data processing and quality control. This has led to epigenome predictions as a way out which understand the epigenetic information distributed throughout genome. In conclusion, exciting times are ahead for research in epigenetics with high computational input requirement.

References

- Aguilera O et al (2010) Epigenetics and environment: a complex relationship. *J Appl Physiol* 109 (1):243–251
- Altschul SF et al (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21 (3):381–395
- Barker DJ, Osmond C (1988) Low birth weight and hypertension. *BMJ Br Med J* 297(6641):134
- Bhasin M et al (2005) Prediction of methylated CpG in DNA sequences using a support vector machine. *FEBS Lett* 579(20):4302–4308
- Bock C, Lengauer T (2007) Computational epigenetics. *Bioinformatics* 24(1):1–10
- Brookes E, Shi Y (2014) Diverse epigenetic mechanisms of human disease. *Ann Rev Gen* 48:237–268
- Campagna-Slater V et al (2011) Structural chemistry of the histone methyltransferases cofactor binding site. *J Chem Inf Model* 51(3):612–623
- Chen M, Zhang L (2011) Epigenetic mechanisms in developmental programming of adult disease. *Drug Discov Today* 16(23):1007–1018

- Choi SW, Friso S (2010) Epigenetics: a new bridge between nutrition and health. *Am Soc Nut Adv Nutr* 1:8–16
- Chuang JC, Jones PA (2007) Epigenetics and microRNAs. *Pediatr Res* 61(5):24R–29R
- Collas P (2010) The current state of chromatin immunoprecipitation. *Mol Biotechnol* 45(1):87–100
- Collins FS et al (2003) A vision for the future of genomics research. *Nature* 422(6934):835–847
- de Pretis S, Pelizzola M (2014) Computational and experimental methods to decipher the epigenetic code. *Front Genet* 5:335
- Dodge JE et al (2002) de novo methylation of MMLV provirus in embryonic stem cells: CpG versus non- CpG methylation. *Gene* 289(1–2):41–48
- Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, Miller J (2009) Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* 41(12):1350–1353
- Eastvaran HP et al (2004) Replication- independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep* 5(12):118
- Espada J, Esteller M (2007) Epigenetic control of nuclear architecture. *Cell Mol Life Sci* 64(4):449–457
- Fazzari MJ, Grealley JM (2004) Epigenomics: beyond CpG islands. *Nature* 5(6):446–455
- Feinberg AP (2010) Epigenomics reveals a functional genome anatomy and a new approach to common disease. *Nat Biotechnol* 28(10):1049–1052
- Fraga MF et al (2005a) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 37(4):391–400
- Fraga MF et al (2005b) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 102(30):10604–10609
- Frigola J et al (2006) Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nat Genet* 38(5):540–549
- Gitan RS et al (2002) Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res* 12(1):158–164
- Gonzalzo M, Jones P (1997) Mutagenic and epigenetic effects of DNA methylation. *Mutat Res* 386(2):107–118
- Hales CN, Barker DJ (2001) The thrifty phenotype hypothesis. *Br Med Bull* 60(1):5–20
- Halkidou K et al (2004) Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer. *Prostate* 59(2):177–189
- Huang C, Wu JC (2013) Epigenetic modulations of induced pluripotent stem cells: novel therapies and disease models. *Drug Discov Today Dis Model* 9(4):e153–e160
- Iizuka M, Smith MM (2003) Functional consequences of histone modifications. *Curr Opin Genet Dev* 13(2):154–160
- Irizarry RA et al (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 41(2):178–186
- Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Gen* 13:484–492
- Jones PA, Takai D (2001) The role of DNA methylation in mammalian epigenetics. *Science* 293(5532):1068–1070
- Kaneda M et al (2004) Essential role for de novo methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 429(6994):900–903
- Kanherkar RR et al (2014) Cellular reprogramming for understanding and treating human disease. *Front Cell Dev Biol* 2:67
- Kawamura A et al (2010) Development of homogeneous luminescence assays for histone demethylase catalysis and binding. *Anal Biochem* 404(1):86–93
- Kent WJ (2002) BLAT – the BLAST-like alignment tool. *Genome Res* 12:656–664
- Kubota T et al (2012) Epigenetic understanding of gene-environment interactions in psychiatric disorders: a new concept of clinical genetics. *Clin Epigenetics* 4(1):1

- Laird PW (2003) The power and the promise of DNA methylation markers. *Nat Rev Cancer* 3 (4):253–266
- Lister R et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462(7271):315–322
- Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragozcy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326 (5950):289–293
- Marchevsky AM et al (2004) Classification of individual lung cancer cell lines based on DNA methylation markers: use of linear discriminant analysis and artificial neural networks. *J Mol Des* 6(1):28–36
- Marmorstein R (2001) Structure and function of histone acetyltransferases. *Cell Mol Life Sci* 58:693–703
- Martinez SR et al (2015) Epigenetic mechanisms in heart development and disease. *Drug Discov Today* 20(7):799–811
- Meissner A et al (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 33(18):5868–5877
- Okano M et al (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* 19(3):219–220
- Okano M et al (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99(3):245–257
- Park LK, Friso S, Choi SW (2012) Nutritional influences on epigenetics and age-related disease. *Proc Nutr Soc* 71(1):75–83
- Petrossian T, Clarke S (2009) Bioinformatics identification of novel methyltransferases. *Epigenomics* 1(1):163–175
- Pfister S et al (2007) Array-based profiling of reference-independent methylation status (aPRIMES) identifies frequent promoter methylation and consecutive downregulation of ZIC2 in pediatric medulloblastoma. *Nucleic Acids Res* 35(7):e51
- Piplani S et al (2016) Homology modelling and molecular docking studies of human placental cadherin protein for its role in teratogenic effects of anti-epileptic drugs. *Comp Biol Chem* 60:1–8
- Pradhan S, Esteve PO (2003) Mammalian DNA (cytosine-5) methyltransferases and their expression. *Clin Immunol* 109(1):6–16
- Pullirsch D et al (2010) The trithorax group protein Ash 2l and Saf – A are recruited to inactive X-chromosome at the onset of stable X inactivation. *Development* 137(6):935–943
- Putiri EL, Robertson KD (2011) Epigenetic mechanisms and genome stability. *Clin Epigenetics* 2 (2):299–314
- Rivera CM, Ren B (2013) Mapping human epigenomes. *Cell* 155(1):39–55
- Roberts RJ et al (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* 31(7):1805–1812
- Robinson MD, Pelizzola M (2015) Computational epigenomics: challenges and opportunities. *Front Genet* 6(88):1–3
- Robinson MD et al. (2014) Statistical methods for detecting differentially methylated loci and regions. *Front Genet*. 5:324: eCollection-2014
- Sato F et al (2011) MicroRNAs and epigenetics. *FEBS J* 278:1598–1609
- Schumacher A et al (2006) Microarray based DNA methylation profiling: technology and application. *Nucleic Acids Res* 34(2):528–542
- Serman A et al (2006) DNA methylation as a regulatory mechanism for gene expression in mammals. *Coll Antropol* 30(3):665–671
- Song J et al (2005) Increased expression of histone deacetylase 2 is found in human gastric cancer. *APMIS* 113(4):264–268
- Tammen SA et al (2013) Epigenetics: the link between nature and nurture. *Mol Asp Med* 34 (4):753–764

-
- Thompson JD et al (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22(22):4673–4680
- Virani S et al (2012) Cancer epigenetics: a brief review. *ILAR J* 53(3–4):359–369
- Wu H et al (2010) Redefining CpG islands using hidden Markov Models. *Biostatistics* 11(3):499–514
- Yan PS et al (2004) Methylation-specific oligonucleotide microarray. *Methods Mol Biol* 287:251–260