Forensic Science International 249 (2015) 255–265

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Review Article

DNA methylation and application in forensic sciences

Farzeen Kader, Meenu Ghai *

Department of Genetics, School of Life Sciences, University of KwaZulu Natal – Westville Campus, Private Bag X 54001, Durban, KwaZulu Natal, South Africa

ARTICLE INFO

Article history:
Received 18 July 2014
Received in revised form 22 January 2015
Accepted 29 January 2015
Available online 11 February 2015

Keywords:
DNA methylation
tDMRs
Forensic science
Body fluid identification

ABSTRACT

DNA methylation of cytosine residues is a stable epigenetic alteration, beginning as early as foetal development in the uterus and continuously evolving throughout life. DNA methylation as well as other epigenetic modifications such as chromatin remodelling and histone modifications are indispensable in mammalian development. Methylation is to a large extent influenced by the ageing process, diets and lifestyle choices. Our understanding of this crucial modification may even contribute to the treatment and prevention of age-related illnesses in the very near future. Genome-wide methylation analysis using high throughput DNA technologies has discovered numerous differentially methylated regions (tDMRs) which differ in levels of methylation in various cell types and tissues. TDMRs have been useful in various applications, particularly medicine and forensic sciences. Forensic scientists are constantly seeking exciting and novel methods to aid in the reconstruction of crime scenes, and the analysis of tDMRs represents a new and reliable technique to identify biological fluids and tissues found at the scene of a violent act. Not only has research been able to unequivocally identify various fluids and tissues, but methods to determine the sex, age and phenotype of donors has been developed. New tDMRs in genes are being searched for consistently to serve as novel markers in forensic DNA analysis.

Contents

* Corresponding author. Tel.: +27 31 260 8617.
E-mail addresses: farzeenkader68@gmail.com (F. Kader), ghai@ukzn.ac.za (M. Ghai).

http://dx.doi.org/10.1016/j.forsciint.2015.01.037
0379-0738/

1. Epigenetics

Epigenetics is a broad term that is used to describe various reversible modifications to the genome. The precise definition of epigenetics has baffled scientists for several years. On top of the genetic code, the epigenetic code comprises an additional layer of information. Whereas the former provides a framework for RNA and structure of protein; the epigenetic code controls packaging of DNA as well as gene regulation [79,88]. The term epigenetics was first introduced by Conrad Waddington in early 1940, who defined epigenetics as "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being" [93,94]. The definition of epigenetics has modified since then, with the development of genetics research. According to Riggs et al. [74], the definition of epigenetics was "the study of mitotically and/or miotically heritable changes in gene function that cannot be explained by changes in DNA sequence". Currently, a widely acknowledged definition is the "study of processes that produce a heritable phenotype that does not strictly depend on the DNA sequence" [54]. Epigenetic modifications include histone modifications, DNA methylation, chromatin remodelling and non-coding RNAs; all of which play a pertinent role in regulation of gene expression devoid of changes in DNA sequence [91,78]. The molecular basis of epigenetics is multifaceted and principally involves alterations in the activation of specific genes. Furthermore, chromatin proteins in association with DNA may be silenced or activated, thus ensuring that cells express only necessary genes required for an activity. Epigenetic programming is believed to begin as early as foetal development in the uterus. As DNA is inherited from one generation to the next, so too are epigenetic patterns preserved during cell division, yet modifications have been observed over an individual's lifetime. These changes have been found to occur in response to environmental exposure and various factors such as smoking and diet. Epigenetic processes include imprinting, reprogramming, gene silencing, X-chromosome inactivation and carcinogenesis. In mammals, a vital cell function regulated by epigenetic processes is cell differentiation wherein during embryogenesis, stem cells are completely differentiated [33,88,91].

2. DNA methylation

DNA methylation is an epigenetic mark of paramount importance for normal development in the human genome. The loss of DNA methylation leads to apoptosis or growth arrest in normal cells. DNA methylation involves the addition of a methyl group ($-CH_3$) at the 5′ position of cytosine residues. Most DNA methylation occurs in CpG dinucleotides, although methylation outside of these dinucleotides has been reported in human DNA in recent years [56,103]. Methyl-cytosine was thought to be as the only chemical modification of the mammalian genomic DNA. However, the existence of hydroxymethyl-cytosine in mammalian cells was proven by Kriaucionis and Heintz [50] and Tahiliani et al. [87]. 5′-Hydroxymethyl-cytosine is an oxidation product of 5′-methyl-cytosine and the conversion of 5′-methyl-cytosine into 5′-hydroxymethyl-cytosine could be the first step in a pathway leading towards DNA demethylation. Due to its probable regulatory role in gene transcription, not unlike methyl-cytosine, hydroxymethyl-cytosine has been termed the 'sixth base' [62,88]. The entire human genome contains about 30 million CpG dinucleotides; these may exist in an unmethylated or methylated state. Within the genome, it is estimated that 60–90% of CpGs are methylated in mammals; unmethylated CpGs are frequently found grouped in regions referred to as CpG islands. Such islands are 300–3000 bp long and have >55% GC content. They are located at the 5′ end (regulatory region) of human genes [67,79,88,91]. The prominent property of CpG islands is that they are unmethylated in germ-line, and most somatic tissues. CpG islands are believed to be protected from methylation by *cis*-acting elements; this ensures continued existence in the presence of strong mutagenic pressure of methyl-cytosine deamination. CpG islands generally act as strong promoters and are also thought to serve as replication origins. Rather than on regions in which majority of methylation is frequently found, most investigations on the role of DNA methylation in mammals have focussed on CpG islands [45,79]. DNA methyltransferases are enzymes responsible for *de novo* methylation and maintenance of methylation. (For details, refer to Supplementary Information Section 1.)

2.1. DNA methylation and gene expression

A key role of DNA methylation is to control gene expression; it is long believed that there is an indirect correlation between gene expression and DNA methylation of CpG island promoter regions. DNA methylation has been associated with silencing of gene expression and condensed nuclease-resistant heterochromatin. Majority of the CpG islands of the inactive X chromosome display methylation, and mono-allelic methylation of imprinted genes is linked with mono-allelic gene expression. DNA methylation status of the CpG-rich promoters of the *GATA2* (gene encoding GATA binding protein 2; transcription factor) and *MASPIN* (mammary serine protease inhibitor) genes correlates well with gene silencing [84,70,37,85,91].

Conversely, there have been instances in which DNA methylation showed no correlation with transcriptional regulation and gene expression, or was even associated with gene activation. Straussman et al. [85] identified 50 loci where host genes were expressed in the same cell-type only when the regions were methylated. This was mainly observed in non-CpG islands, possibly a result of selective demethylation of inactive genes [70,37,73,85,91]. Research by Walsh and Bestor [95] to investigate the methylation status of seven genes found no relationship with gene expression and studies by Warnecke and Clark [97] also found that expression of skeletal α-actin genes in adult mice did not correlate with methylation status of the promoter. There are various explanations for the lack of correlation observed. Several genes are able to engender numerous transcripts by using other transcription start sites. Despite the high degree of methylation displayed by the promoter of the *PARP12* gene which encodes poly (ADP-ribose) polymerase 12, gene expression was observed, as demonstrated in the study by Rauch et al. [73], yet rapid amplification of cDNA ends revealed transcription initiation from downstream of the methylated CpG island, via an intragenic promoter. It has been speculated that when intragenic islands are not associated with a known transcription start site, the methylation status could hinder spurious body transcription which may obstruct correct expression of parent genes [37].

Studies by Sleutels et al. [82] provided evidence that intragenic CpG islands are able to localise to sites of antisense ncRNA transcription initiation; this results in negative regulation of the sense transcripts. Both *Tsix* and *Air* ncRNA are derived from CpG islands and partake in the regulation of the sense transcripts [82,85]. A ncRNA that is transcribed from the HOXC locus (of the Homeobox super family), *HOTAIR*, represses the *HOXD* cluster *in trans* [75]. In all cases described, methylation of CpG islands led to the derepression of genes that were silenced by ncRNA.

Numerous hypermethylated CpG islands are known to have no regulatory roles in gene transcription as they are located outside coding sequences, in intergenic DNA. Yet, monoallelic expression of the *H19*/*IGF2* imprinted locus is determined by methylation of an intergenic CpG island that is situated upstream of the *H19* gene that encodes a 2.3 kb spliced, capped, and polyadenylated long