Chapter 1

Environment and Epigenetics
1.1 Introduction

The 1940s saw the emergence of the concept of Epigenetics when Sir Conrad Waddington described it as ‘the study of casual interaction between genes and their products which bring the phenotype into being’. Across several decades since then DNA methylation (Holliday and Pugh 1975), histone modifications (Z. X. Chen and Riggs 2005), chromatin structure (Turner 1998) and non-coding RNA (Pokholok et al. 2005) have been identified as tenets of an epigenome. Today epigenetics refers to ‘the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence’ and an epigenetic state as a coordinated interplay of Epigenator (environmental cue), Initiator (mediator of local chromation modification, for example, DNA-binding protein, ncRNA) and Maintainers (sustain the established epigenetic response through DNA methylation, histone modification, nucleosome repositioning etc) (Berger et al. 2009). The epigenome plays an important interface between the environment and the genome. This thesis explores one such interaction between the environment (mycobacterial infection) and the host genome through DNA methylation.

The proverbial epigenetic code is a complex interplay of DNA methylation, histone modification and non-coding RNA to generate specific chromatin conformation for specific genetic loci and is achieved by the actions of chromatin remodeling complexes. DNA methylation was the first epigenetic modification to be recognized and plays an important role in gene regulation (Holliday and Pugh 1975; Riggs 1975).
1.2 DNA Methylation

DNA methylation was first identified in prokaryotic genome as a defense mechanism against foreign DNA (virus or phage) (Arber and Dussoix 1962; Wion and Casadesus 2006). Subsequently, DNA methylation was also identified in mammals and since then various functions have been attributed to it, for example, transcriptional regulation of gene expression, silencing of transposable element, transmission of epigenetic memory, genomic imprinting and dosage compensation. Surprisingly, DNA methylation is present in prokaryotes, some insects, sea urchins, invertebrate chordates, all vertebrates and plants but absent or rare in yeast, fruit fly (Drosophila melanogaster) and worms (Caenorhabditis elegans) (S. Feng et al. 2010; Zemach et al. 2010). DNA methylation is thought to exert it effects on gene regulation by manipulating the chromatin structure via histone deacetylation, methylation and nucleosome compaction. While DNA methylation is established by de novo methyltransferases (DNMT3A, DNMT3B and DNMT3L), it is maintained by the actions of maintenance methyltransferase (DNMT1). Its regulatory effects are mediated by proteins with affinity for methylated DNA including proteins of the MBD and BTB/POZ family) (Bogdanovic and Veenstra 2009; Rose and Klose 2014).
In mammals, DNA is methylated at 5’ position of cytosine (5-mC) that accounts for approximately 1% of the total deoxy bases in somatic cells. 5mC is predominantly found in the CpG context with global 5mC levels varying in a tissue-specific manner. The distribution of CpG in the genome is not homogeneous; it is rather concentrated to regions referred to as CpG Islands (CGIs). Majority of the CGIs are located at 5’ ends of a gene or within repetitive sequences. CGIs within the 5’end of a gene are associated with regulation of transcription and are hypomethylated in most developmental stages and show tissue-specific methylation (A. P. Bird 1986). Highly repetitive DNA sequence CGIs are normally hypermethylated in all the cell types (Ehrlich et al. 1982). Hypermethylation of these repetitive elements has been suggested to be essential for maintaining genomic stability due to suppression of transcription from within these elements (Woodcock et al. 1997) (W. M. Liu et al. 1995).

Within a genetic locus DNA methylation is not just restricted to the promoter. Methylated cytosines are also present within a gene body and as discussed below, the resulting effect on transcription is dependent on the position of the methylated cytosine (Maunakea et al. 2010; Shukla et al. 2011; Stadler et al. 2011; Venolia and Gartler 1983).

1.2.1 Promoter Methylation

On the basis of CpG density, promoters have been categorized as High CpG Promoter (HCP), Intermediate CpG Promoter (ICP) and Low CpG Promoter (LCP) (Table 1.1). HCP are mostly associated with housekeeping genes, contain multiple GC-boxes and are mostly unmethylated (Zhu et al. 2008), LCP associated with tissue-specific genes in somatic cells contain the TATA box and are methylated (Schug et al. 2005). ICP contain intermediate CG-content and offer variability in its methylation status as well as in regulation of expression (Weber et al. 2007; Xie et al. 2013).
<table>
<thead>
<tr>
<th></th>
<th>HCP</th>
<th>ICP</th>
<th>LCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatic Cells</td>
<td>● Hypomethylated</td>
<td>● Hypermethylated</td>
<td>● Methylated</td>
</tr>
<tr>
<td></td>
<td>● Housekeeping Genes</td>
<td>● Negative correlation with gene activity</td>
<td>● Tissue-specific genes</td>
</tr>
<tr>
<td></td>
<td>● High gene expression</td>
<td></td>
<td>● No correlation with gene activity</td>
</tr>
<tr>
<td></td>
<td>● GC box (YY1)</td>
<td></td>
<td>● TATA box</td>
</tr>
<tr>
<td>Embryonic Cells</td>
<td>Early development genes</td>
<td>Germ-line specific genes</td>
<td>Lineage specific genes</td>
</tr>
<tr>
<td></td>
<td><em>SOX2, SOX1, NODAL</em></td>
<td></td>
<td><em>NANOG, POU5F1</em></td>
</tr>
</tbody>
</table>

**Table 1.1: Features of CGI Promoters**
The table shows the methylation status and type of associated genes with each class of promoter in somatic and embryonic cells.

Promoter DNA methylation dependent transcriptional regulation is effected by modulating the interaction and type of proteins binding to the promoter. For example, MBD or Methylbinding proteins (MeCP2 and MBD1-4) specifically associate with methylated CpG dinucleotides and also with chromatin remodeling complexes and thus couple DNA methylation and histone modification. While MeCP2 interacts and recruits Suv39h1/2 (histone methyltransferase) and Sin3a (histone deacetylase complex) (Fuks et al. 2003), MBD1 recruits SETDB1 (Sarraf and Stancheva 2004) at the target promoter leading to deposition of H3K9me3 at the methylated locus and hence transcriptional repression. DNA methylation at regulatory sites can also indirectly effect transcriptional repression at the gene promoters. This is well documented in case of the imprinted locus *H19/Igf2*. CTCF is a ubiquitous transcription factor that binds to the consensus sequence ‘CGCG(T/G)GGTGGCAG’. The binding of CTCF is disrupted by the presence of methylated cytosines. CTCF binding to H19 Imprinting control region leads to a hypomethylated state at the maternal locus. This hypomethylation of the ICR allows expression of the maternal *H19* allele. Simultaneously, this CTCF binding also blocks the enhancer from activating the promoter of *Igf2* on the maternal allele (Szabo et al. 2004).
1.2.2 Gene-body Methylation

Gene-body methylation is normally correlated with actively transcribed genes. It is also implicated in the prevention of spurious transcription and silencing of repetitive elements like LINE, SINE and Alu repeats (P. A. Jones 2012). It has been suggested that DNA methylation at the introns may be involved in regulating the enhancer activity. The level of DNA methylation across the gene body is not uniform with exons showing higher level of methylation than introns (Chodavarapu et al. 2010).

The level of intragenic DNA methylation correlates with the levels of H3K36me3, a mark for transcriptional elongation (Hahn et al. 2011). In addition, these two epigenetic marks are associated with higher RNA polIII density within the exonic regions. The suggestion that this could be related to splicing is supported by various studies. For example, in a transition from naïve to mature T-cells, a gain of methylation at exon 5 of CD45 correlates with the loss of CTCF binding and exclusion of exon 5 from the spliced product. CTCF mediates this by promoting RNA pol.II pausing, allowing co-transcriptional splicosome assembly at the splice site (Shukla et al. 2011).

1.3 Non-CpG methylation

Non-CpG methylation is present at very low levels in most of the differentiated cells in mammals (Truong et al. 2013; Ziller et al. 2011). Apart from germ cells and ES cells (Ichiyanagi et al. 2013; Ramsahoye et al. 2000; Tomizawa et al. 2011) human and mouse adult brain cells have also been shown to have non-CpG methylation (Guo et al. 2014; Lister et al. 2013)
1.3.1 Non-CpG methylation in Germ Cells

During fetal development in mice, sperm development is arrested at the prespermatogonia stage and is resumed after birth. An accumulation of non-CpG methylation has been observed around the B1 LINE repetitive sequences in these arrested cells (Ichiyanagi et al. 2013). The DNA was found to be hemimethylated at the symmetric CHG sites indicating that every non-CpG methylation was a \textit{de novo} event. Similarly, non-CpG methylation has been detected in oocytes (Kobayashi et al. 2012) at maternally methylated DMRs in mice (Tomizawa et al. 2011). Furthermore, non-CpG methylation was found to be abundant at CHH rather than at CAG and was localized to regions of high CG methylation. The functional significance, however, remains unknown.

1.3.2 Non-CpG methylation in Pluripotent Cells

In human embryonic stem cell line H1, 25% of methylated cytosine was found to exist in the non-CpG dinucleotide context and conservation of methylation at CHG and CHH sites was only 10-40%. The observed non-CpG methylation was predominantly observed on the anti-sense strand and was comparatively higher in exons as compared to introns and 3’ UTR. Highly methylated non-CpG sites were found to be conserved and were enriched in the motif TACAG. Low non-CpG methylation mostly occured at CpA dinucleotide (P. Y. Chen et al. 2011). A periodicity of 8-10 bases for methylated CHH sites, corresponding to a single turn of DNA helix, was also seen. However, a positive correlation was observed between non-CpG methylation and gene expression in H1 human embryonic stem cells (Lister et al. 2009).
1.3.3 Non-CpG methylation in Differentiated Cells

As mentioned before, most of the differentiated cells (except adult neurons) show very low levels of non-CpG methylation and it has been observed that non-CpG methylation levels decrease with increasing differentiation (Laurent et al. 2010; Ziller et al. 2011). Analysis of mature adult neurons from both human and mouse pre-frontal cortex revealed mCH to be the major form of methylated cytosine that is accumulated through development to adult stage (Lister et al. 2013). In differentiated cells other than neurons- hESC derived fibroblasts, neonatal fibroblasts and monocytes, the CpA methylation was found to be conserved (Laurent et al. 2010). The same CpA sites also showed methylation in hESCs, raising the possibility of association of conserved non-CpG methylation with a cellular memory transmitted through differentiation.

A comparison of pattern of non-CpG methylation in ESCs and differentiated brain cells has revealed some key features. While pluripotent stem cells showed majorly unconserved non-CpG methylation associated with highly expressed genes (P. Y. Chen et al. 2011), adult neurons exhibited abundant conserved non-CpG methylation associated with repression (Lister et al. 2013). Some of the methylated CpA sites in the pluripotent cells (hESCs) were also conserved in the derived lineage (hESC fibroblasts) (Laurent et al. 2010). However, the site-specific mCH that accumulated in adult mature neurons was absent in fetal neurons (Guo et al. 2014). Further, methylated non-CpG sites were enriched in ‘TACAG’ motif in the hESCs (P. Y. Chen et al. 2011) as compared to the preference for CHH methylation at ‘CAC’ in the neurons (Guo et al. 2014). A similar 8-bp spacing was observed in mCHH and mCHG in neurons in contrast to the 8-, 21-, 29- spacing pattern of mCHG in pluripotent stem cells (Guo et al. 2014). mCpA has been found to be coupled with mCpG in mouse oocytes (Tomizawa et
(Ziller et al. 2011) as well as in hESCs (Ziller et al. 2011). In prostate cancer cells, however, non-CG methylation is locus-specific and not associated with mCG (Truong et al. 2013).

<table>
<thead>
<tr>
<th>Pluripotent Cells</th>
<th>Mature Neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Rare</td>
<td>• Abundant</td>
</tr>
<tr>
<td>• Mostly unconserved mCA</td>
<td>• Conserved mCH</td>
</tr>
<tr>
<td>• Associated with highly expressed genes</td>
<td>• Associated with repression of expression</td>
</tr>
<tr>
<td>• Reduction in levels upon differentiation</td>
<td>• Accumulates in mature neurons</td>
</tr>
<tr>
<td>• 8-, 21-, 29- bp spacing of mCHG</td>
<td>• Similar 8-bp spacing of mCHH and mCHG</td>
</tr>
<tr>
<td>• Preference for TACAG at conserved sites</td>
<td>• Preference for CAC</td>
</tr>
<tr>
<td>• Role of DNMT3A and DNMT3B</td>
<td>• Role of DNMT3A</td>
</tr>
</tbody>
</table>

Table 1.2: Sallent features of non-CpG methylation
The table shows distinct patterns of non-CpG methylation between pluripotent and differentiated cells.

1.4 DNA methylation during development

In the mammalian germ line, during primordial germ cell specification, genome wide DNA demethylation occurs except for IAPs and a few LTRs that retain methylation (Hajkova et al. 2002; Lees-Murdock et al. 2003). Parent-specific remethylation occurs in a sex-specific manner with female gametes begin to reacquire methylation after meiosis I arrest and global levels being achieved only after sexual maturity (Smallwood et al. 2011). Male gametes, on the other hand, accumulate methylation during spermatogenesis and before birth (Oakes et al. 2007).

Upon fertilization, the paternal genome undergoes active DNA demethylation by the activities of TET1 and BER pathway (Iqbal et al. 2011; Santos et al. 2013). This is followed by a wave of replication-dependent passive demethylation of both the parental genomes that continues till the blastocyst stage (Reik et al. 2001). Global remethylation of the embryonic genome occurs in concert with gastrulation and is essential for ensuing lineage specification and differentiation (Borgel et al. 2010; Kafri et al. 1992).
1.5 Writers of DNA methylation

Mammalian genome reprograms its methylome during development and once the DNA methylation patterns are established, they are maintained throughout a particular cell lineage. Post-implantation, new methylation pattern is established in the embryo by the concerted activities of DNMT3A and DNMT3B. This de novo methylation helps to compartmentalize the genome during tissue-specific differentiation (Okano et al. 1999). The maintenance of post-implantation methylation marks is done by DNMT1 (Mertineit et al. 1998). DNMT3A, DNMT3B, DNMT3L and DNMT1 are the only known cytosine DNA methyltransferases in the human genome. These DNA methyltransferases are themselves regulated by various post-translation modifications that modulate their enzyme activity as well as their interacting partners. These PTMs are often dynamic and induced by specific environmental signals, allowing these proteins to relay the signal into an epigenetic response.

1.5.1 DNMT1

DNMT1 was the first eukaryotic DNA methyltransferase to be characterized. It is an 1620 amino acid long protein with a large N-terminal domain connected by a central domain to the conserved C-terminal methyltransferase domain. The N-terminal domain contains the nuclear localization signal (NLS), PCNA-interaction domain and replication-targeting domain. The central domain contains cysteine-rich Zinc binding motif, polybromo (BAH) motif and GK linker. The function of cysteine-rich region is still unknown but is present in DNA methyltransferases as well as methyl-C binding proteins like MBD1. The BAH motif is thought to be involved in mediating protein-protein interactions while the GK linker was proposed to be the site of PTMs (Bestor 1992; Leonhardt et al. 1992; Leonhardt and Bestor 1993; Turek-Plewa and Jagodzinski 2005).
DNMT1 functions to maintain the DNA methylation pattern on the daughter strand post-replication and hence has been termed as the ‘maintenance’ methyltransferase. In \textit{in vitro} assay, DNMT1 was found to have a 3- to 50-fold higher affinity towards DNA substrate hemimethylated at CpG (Goll and Bestor 2005). No specificity for flanking sequence for CpG methylation have been observed, though a preference for methylation at ‘CCGG’ has been reported for an unmethylated DNA substrate. Further, ‘GGCTC’ sites tend to terminate the processivity of DNMT1 (Goyal et al. 2006).

There are two isoforms of DNMT1- the oocyte-specific DNMT1o and the somatic isoform DNMT1s. DNMT1o represents the truncated isoform and is the result of germline-specific promoter activity and alternative splicing of 5’ exon (Mertineit et al. 1998). This isoform is mainly cytoplasmic and accumulates to high levels in the oocyte. The cytoplasmic staining of DNMT1o is observed till the 8-cell stage (Carlson et al. 1992), the blastocyst is free of DNMT1 and nuclear staining reappears post-implantation (Trasler et al. 1996). This nuclear staining corresponds to DNMT1s. The maternal store of DNMT1 has been shown to maintain the methylation at the IAPs and imprinted genes in the zygote when the rest of the genome undergoes demethylation. DNMT1s, on the other hand, has been shown to be involved in the wave of \textit{de novo} methylation post implantation (E. Li et al. 1992). The somatic isoform of DNMT1 is known to be exclusively nuclear. During G1-phase of cell cycle, DNMT1 was found to have a nucleoplasmic distribution that changed to discrete foci co-localized with Y-satellite DNA in S-phase in mouse fibroblast nuclei. This reorganization was found to be correlated to replication timing of the genomic regions (Leonhardt et al. 1992).

Recent research has shown that the activity of DNMT1 can be modulated by post-translational modifications like phosphorylation, methylation, acetylation and sumoylation (Figure 1.1). Phosphorylation occurs at multiple residues and is fostered by multiple kinases- AKT (Hodge
et al. 2007) and PKCζ (Lavoie et al. 2011). Set7, a histone methyltransferase, was shown to methylate DNMT1 at lys142 and induce its degradation via the proteasome that is required for clearance of DNMT1-bound inactive promoters (Esteve et al. 2009). Acetylation of DNMT1 by Tip60 targets it for degradation, which is counteracted by HDAC1 induced deacetylation (Du et al. 2010). SIRT1 mediated deacetylation of C-terminal lysines of DNMT1 accentuates its DNA methyltransferase activity while deacetylation of lysines in GK linker suppresses it (Peng et al. 2011). The multiple levels of regulation of DNMT1 imposed by post-translational modifications, illustrate the importance of this single protein in the dynamic coupling of the cellular response to environmental cues via the epigenome.

Figure 1.1: Regulation of DNMT1 activity by post-translational modifications

DNMT1 is phosphorylated by AKT and PKCζ, acetylated by Tip60, deacetylated by HDAC1 and methylated by Set7. Phosphorylation by PKCζ affects the interaction of DNMT1 with UHRF1. Both acetylation and methylation of DNMT1 leads to its degradation while deacetylation by HDAC1 reverses the effect.
1.5.2 DNMT3

The DNMT3 family consists of DNMT3A, DNMT3B and DNMT3L. While DNMT3A and DNMT3B have catalytic activity, DNMT3L is the catalytically inactive member of the family and is known for its regulatory nature. DNMT3A and DNMT3B share structural similarities in their three conserved domains. The N-terminal contains a PWWP domain involved in heterochromatin targeting, a PHD domain by which the proteins interact with histone methyltransferases, HDACs and other transcriptional repressors (Fuks et al. 2001). The C-terminus contains the catalytic methyltransferase domain which is similar to the prokaryotic cytosine methyltransferases (Okano et al. 1998a).

The activities of Dnmt3a and Dnmt3b were found to be regulated by an accessory factor, later identified to be Dnmt3l. Recently, Dnmt3l has been reported to form complex with and direct the DNA methyltransferase activity of Dnmt3a by binding to unmethylated H3K4 through its PHD-domain (Ooi et al. 2007). The complex involves interaction of C-terminal domain of Dnmt3l with the catalytic domain of Dnmt3a and shows further dimerization to form a tetramer (Dnmt3l-Dnmt3a-Dnmt3a-Dnmt3l). This tetramer can methylate two CpGs separated by one helical turn simultaneously (Jia et al. 2007). Surprisingly, Dnmt3l also antagonizes the activities of both Dnmt3a and Dnmt3b by preventing their association with H3K27me3 promoters. Dnmt3l knockdown in ES cells leads to loss of methylation in gene bodies of actively transcribing genes while a gain in methylation is observed at bivalent promoters of developmental genes. Also, Dnmt3l interacts directly with PRC2 through Ezh2. Thus at bivalent promoters, PRC2 recruits Dnmt3l which inhibits the recruitment of Dnmt3a and Dnmt3b (Neri et al. 2013).
Both Dnmt3a and Dnmt3b are highly expressed during gametogenesis and early embryonic development in mice but present at very low levels in adult somatic tissue. Inactivation of both these proteins leads to early embryonic lethality (Okano et al. 1999). Dnmt3a and Dnmt3b exhibit similar localization patterns but Dnmt3a has a stronger affinity for heterochromatin as compared to the diffuse nucleoplasmic localization of Dnmt3b. As cellular differentiation proceeds in a developing embryo, lineage-specific genes are turned on and genes associated with the undifferentiated state, Oct4 and Nanog, are turned off. In mouse, this repression of Oct4 and Nanog is brought about by the concerted actions of Dnmt3a and Dnmt3b. In the absence of one of the proteins, methylation still happens but with much reduced efficiency (J. Y. Li et al. 2007).

Similar to Dnmt1, post-translation modification add another layer of regulation to Dnmt3a and Dnmt3b activities (Figure 1.2). Till date, sumoylation has been observed for both Dnmt3a (Ling et al. 2004) and Dnmt3b (Kang et al. 2001) while citrullination has been reported only for Dnmt3a. Sumoylation disrupts the interaction of Dnmt3a with HDACs but does not affect its ability to interact with Dnmt3b. Citrullination, on the other hand, increases the stability of DNMT3A but its functional significance still needs to be investigated. Citrullination by PADI4 occurs at arginines located upstream of the PWWP domain (Deplus et al. 2014). These arginines are involved in nuclear localization of DNMT3A while PWWP domain is involved in targeting to the pericentric heterochromatin (T. Chen et al. 2004). Thus, citrullination may also affect the localization of DNMT3A and thus indirectly modulate its DNA methylation activity.


Figure 1.2: Regulation of DNMT3A activity by post-translational modifications
DNMT3A is citrullinated by PADI4 that affects its chromatin localization. Sumoylation of DNMT3A disrupts its interaction with HDACs.

Role in non-CpG methylation: As mentioned earlier, non-CpG methylation is predominantly detected in embryonic stem cells. Given the predominant expression of DNMT3A and DNMT3B proteins during embryogenesis, it is plausible that non-CpG methylation may in fact be effected by these de novo methyltransferases. DNMT3 knockdown in hESC was found to lead to a reduction in the levels of non-CpG methylation but molecular pluripotency remained unaffected (Ziller et al. 2011). This indicated that non-CpG methylation could be a result of high levels DNMT3A, 3B in ESCs and hence may actually be dispensable for pluripotency. But the presence of conserved non-CpG methylation argues against this hypothesis. Moreover, mCH methylation in adult neurons also appears to be marked by Dnmt3a and read by MeCP2.

1.5.3 DNMT2

DNMT2, the second member of the DNMT family, has been functionally established to be a tRNA methyltransferase. The protein was found to have sequence similarity to other members of the DNMT family and even use the conserved DNA methyltransferase mechanism to methylate cytosine in the anticodon loop of tRNAs (Goll et al. 2006; Okano et al. 1998b). DNMT2 has been shown to possess a weak DNA methylation activity and the residual CG
methylation was found in a loose ttnCGga(g/a) consensus (Hermann et al. 2003). The protein is conserved from *Schizosaccharomyces pombe* to humans but the organisms containing DNMT2 as the sole DNA methyltransferase - *Schizosaccharomyces pombe, Dictyostelium discoideum, Entamoeba histolytica, Schistosoma mansoni, and Drosophila melanogaster* - were found to have very low levels of global DNA methylation (Jeltsch et al. 2006). However, a few studies have indicated the importance of Dnmt2-dependent DNA methylation. For example, in *Drosophila*, genome stability was shown to be regulated by Dnmt2-dependent methylation of transposons (Phalke et al. 2009). Another study in *S. mansoni*, parasitic blood fluke, has demonstrated that the genome was epigenetically modified by DNA methylation and was important for regulation of oviposition and egg development (Geyer et al. 2011). Nonetheless, whole genome bisulfite sequencing revealed both *Drosophila* and *S. mansoni* genomes to be unmethylated, thus contradicting the earlier two reports (Raddatz et al. 2013). The contribution of Dnmt2 towards DNA methylation remains controversial.

1.6 The Readers

1.6.1 The MBD family

The MBD (Methyl Binding Domain) family has five members MeCP2, MBD1, MBD2, MBD3 and MBD4. The structural similarities of these members are restricted to the MBD motif (Hendrich and Bird 1998). All the MBD proteins, except MBD4, recruit HDACs and are part of chromatin remodeling complexes. MeCP2 has been reported to interact with Sin3a (P. L. Jones et al. 1998; Ng et al. 1999) and SWI/SNF complex Brahma (Harikrishnan et al. 2005), MBD2 is a part of MeCP1 complex (Ng, Zhang et al. 1999) while MBD3 is a component of Mi2-NuRD complex (Wade et al. 1999; Zhang et al. 1999). MBD4 has N-glycosylase activity and is involved in DNA repair (Hendrich and Bird 1998).
The binding of MBD proteins to DNA is not sequence specific and seems to be dictated by density and distribution of CpG and the localized chromatin organization (Ballestar and Wolffe 2001). MeCP2 binds to symmetrically methylated CpG and can recruit Suv39h1/2 histone methyltransferase and HDAC1. Thus, transcriptional repression through MeCP2 is brought by overlaying of DNA methylation with histone methylation and histone deacetylation (Fuks et al. 2003). MeCP2 is also known to associate with Dnmt1, independent of its association with HDAC1 and this complex is involved in maintenance of DNA methylation at newly synthesized daughter strand (Kimura and Shiota 2003). MBD1 associates with histone methyltransferase SETDB1 and facilitates H3K9 methylation on newly assembled histones to stably maintain the heterochromatin state (Sarraf and Stancheva 2004). MBD2 is a component of MeCP1 histone deacetylase complex while MBD3 is a part of the Mi-NuRD complex (Q. Feng and Zhang 2001).

1.6.2 Kaiso and Kaiso-like proteins

Apart from MBD family of proteins, Kaiso-like proteins are also known to bind to methylated DNA. Kaiso is a zinc finger protein that binds to two consecutive methylated CpG dinucleotides (Daniel et al. 2002). ZBTB4 and ZBTB38, two recently discovered Kaiso-like zinc finger proteins, can bind single methylated CpG dinucleotide and localize to heavily methylated satellite DNA (Filion et al. 2006).
1.7 Environment induced alteration of the Epigenome

A signal, whether developmental or environmental, elicits an epigenetic response as a change in DNA methylation and gain or loss of histone acetylation and/or methylation, which eventually lead to nucleosome repositioning, open or closed chromatin structures and hence a modulation of cell transcriptome, replication and repair (Jaenisch and Bird 2003). Developmental cues initiate temporal as well as spatial regulation of an epigenome. Environmental cues, on the other hand, trigger persistent modulation or a dynamic modulation, depending on the strength and duration of the stressor (Feil and Fraga 2011).

The plasticity of an epigenome towards environment first came into view through studies on monozygotic twins. These twins have exactly the same DNA sequence and share the same *in utero* environment. Thus, any difference or similarity in their epigenome tends to reflect the effect of environment (Bell and Spector 2012). Monozygotic (MZ) twins are further distinguished as dichorionic (splitting of blastomeres within four days of fertilization) or monochorionic (splitting of blastomeres after four days of fertilization) (Kaminsky et al. 2009). Thus monochorionic MZ twins would show a higher concordance of DNA methylation than dichorionic MZ twins, which in turn was found to be higher than that for dizygotic twins. So far this concordance has been observed only at specific loci for examples, at genes functional in the respective tissue (Kaminsky et al. 2009). Thus even with the same *in utero* environment the methylome conservation was very low, indicating that stochastic events may be more influential than environment itself (Gordon et al. 2012). On similar lines was the discordance observed in monozygotic twins for bipolar disorder or schizophrenia (Bell and Spector 2012).

Another interesting observation was epigenetic drift and convergence. Studies found that methylation level variability increased (epigenetic drift) with age in some twin pairs while it became less discordant (convergence) in others. Epigenetic Drift was proposed to be a
cumulative effect of stochastic events as well as environmental influence, which indicated that DNA methylation is rather dynamic and does change over time. Convergence, on the other hand, was mostly observed in cases where the twins differed considerably in birth weights. Convergence is best explained by the regression model, wherein individual phenotypes turn to population mean over time (Martino et al. 2013). Thus, these MZ twin methylome studies paved the way for a deeper understanding of DNA methylation dynamics in development and disease.

Aberrant DNA methylation has also been implicated in various diseases either as a cause or a consequence. Cancer, a multifactorial event, has best exemplified the importance of DNA methylation in cellular homeostasis. Attempts are now being made to identify driver epigenetic events (De Carvalho et al. 2012). Global hypomethylation has been seen mainly at repetitive sequences and leads to chromosomal instability and translocations. Loss of imprinting due to promoter hypomethylation was seen at $IGF2$ locus in breast, liver and lung cancers (Ito et al. 2008). $DNMT3L$, known to be involved in reprogramming, was also overexpressed in cervical cancers due to loss of methylation at the promoters (Gokul et al. 2007). Apart from hypomethylation, hypermethylation was also observed at several gene loci like $BRCA1$, $RARB2$, $RASSFIA$, $RB$, $p53$, $DAPK1$ among others (Portela and Esteller 2010).

Moreover, alterations in DNA methylation have been documented in response to environmental agents like metal ions, temperature, exposure to harmful chemicals, pollutants, pathogens and nutritional insufficiency/deficiency. While nutritional effects are most prominent during peri-conceptional and early embryonic development; chemical pollutants can affect at any stage of life (Feil and Fraga 2011). Some of these alterations are discussed below.
1.7.1 Nutrients and other compounds

**Nutrients:** Nutrients and other bioactive agents may influence gene expression by directly altering the enzyme activity of epigenetic effectors like DNA methyltransferases and histone modifying enzymes or by limiting the availability of cofactors like SAM (folate), Biotin, acetyl CoA (Pantothenic Acid) and NAD (Niacin) (Figure 1.3). Folate and Vitamin B12 are the major regulators of one-carbon metabolic pathway for generation of S-adenosyl methionine (SAM). Folate deficiency during pregnancy has been associated with locus-specific epigenetic alterations in the fetus. These include glucocorticoid receptor (Lillycrop et al. 2007), insulin-like growth factor 2 (Heijmans et al. 2008), insulin, leptin, IL10 (Tobi et al. 2009). The methylated form of Vitamin B-12 (methylCobalamin) is utilised as a cofactor by methionine synthase to transfer a methyl group from 5-methyltetrahydrofolate to homocysteine to generate tetrahydrofolate (THF) and methionine. This reaction is the major source of regenerating THF and the methionine is the precursor for SAM (Selhub 1999). Studies in rats (Choi et al. 2004) and mice (Fernandez-Roig et al. 2012) have shown global hypomethylation in animals with Vitamin B12 deficiency as compared to control animals. Another study in rats has demonstrated promoter hypomethylation in the cystathionine beta-synthase gene in Vitamin B12 deficient mice that did not recover even after methionine supplementation (Uekawa et al. 2009). Betaine, the oxidation product of choline, also acts a methyl group donor in an alternate pathway for SAM synthesis. Increased uptake of choline during gestation in mice has been shown to cause alterations in DNA cytosine and histone lysine methylation levels which in turn have been correlated to improved cognitive functions later in life. Gestational choline was also shown to modulate the expression of Dnmt1, Dnmt3a, G9a, Suv39h1 and Kmt1a, and affecting the genes involved in synaptic plasticity (Blusztajn and Mellott 2012).
**Naturally Occuring Biocompounds:** Bioactive agents like phytoestrogens (resveratrol and genistein), polyphenols (bioflavonoids) and Curcumin have been implicated in cancer prevention by modulation of DNA methylation and DNMT activity. Resveratrol and genistein are known to regulate oestrogen-responsive genes by interaction with estrogen receptor. Resveratrol treatment of breast cancer cell line was found to result in hypomethylation and reactivation of tumor suppressor genes $RAR\beta_2$ and $PTEN$ (Stefanska et al. 2012). A partial reversal of promoter hypermethylation of $RAR\beta_2$ and $p16$ genes was observed in genistein treated oesophageal squamous carcinoma cell line (Papoutsis et al. 2012). Bioflavinoids, coffee polyphenols and tea catechins are known to act as substrate for Catechol-O-methyltransferase, depleting SAM and accumulating S-adenosyl homocysteine (SAH). SAH acts as a feedback inhibitor of human DNMTs (W. J. Lee and Zhu 2006). Curcumin, a naturally occurring compound in the spice turmeric, binds to the catalytic thiol group of Cys1226 of DNMT1 leading to inhibition of its activity and an indirect global hypomethylation (Z. Liu et al. 2009).

![Figure 1.3: Role of micronutrients in generation of SAM pool](image)

Methyltetrahydrofolate and betaine provide methylgroup to the SAM pool via generation of methionine.
Chemical pollutants in the environment: Among environmental pollutants, chemicals like asbestos, silica and benzene have been associated with altered DNA methylation. Low-dose exposure of airborne benzene has been correlated to methylation changes in blood DNA including hypomethylation of LINE-1 and Alu elements, hypomethylation of MAGE1 and hypermethylation of p15 (Hou et al. 2012). Among metal ions, Arsenic exposure has been shown to cause SAM insufficiency and altered activities of DNMTs (Hou et al. 2012). Nickel has been shown to seed chromatin compaction leading to DNA methylation and hence silencing of senescence and tumor-suppressor genes (Y. W. Lee et al. 1995). Exposure to chemicals like aluminum (Pogue et al. 2009) and cadmium (Bollati et al. 2010) has been shown to alter miRNA pool in human cells.

Exposure to pesticides including vinclozolin and methoxychlor has been reported to lead to sterility by affecting ovarian and testicular function. This dysfunction has been attributed to induced alterations in DNA methylation in the germline (Guerrero-Bosagna et al. 2010; Zama and Uzumcu 2009).

Bisphenol A and phthalates, the chemicals widely used in plastic manufacturing industries and known to leech into the environment through industrial waste, have been shown to have weak estrogenic properties and mice studies have highlighted their role in in utero induced alterations in DNA methylation. In mouse model, exposure to BPA was found to alter the methylation status of Hox10a in utero (Bromer et al. 2010). In human breast cancer cell lines, BPA was found to increase the expression of EZH2 (histone methyltransferase) while in placental cell lines, it altered the miRNA levels specifically miR-146a (Avissar-Whiting et al. 2010; Doherty et al. 2010).
1.7.2 Pathogens

Pathogen induced host epigenetic modifications have mainly been studied in the context of histone modifications. Alterations of DNA methylation have been reported in a handful of cases (Figure 1.4, Table 1.3) but whether it is a direct effect of infection has not been ascertained.

**Indirect effect: Epigenetic modifications in host**

**Helicobacter pylori**: *H. pylori* is a gram negative bacteria that induces chronic gastric inflammation leading to cancer. The infection of gastric cells was found to induce hypermethylation of E-cadherin gene, *CDH1* (Qian et al. 2008). This effect, however, may have been mediated by IL-1β that is known to be upregulated in *H. pylori* infections (Hmadcha et al. 1999).

**Campylobacter rectus**: A study in murine model involving infection of pregnant mice with *C. rectus* showed hypermethylation of P0-promoter of *Igf2* gene and repression of *Igf2* expression (Bobetsis et al. 2007).

**Leishmania donovani**: The intracellular parasite *L. donovani* is the causative agent of visceral leishmaniasis and its life cycle involves two hosts, gut of sandfly and immune cells of humans. A recent report has shown that cytosine methylation changes occur at CpG dinucleotides of genes involved in host defense response in macrophages infected with the parasite. A majority of the regions that gained methylation were found to be low CpG density sites in contrast to intermediate CpG sites that showed loss of methylation upon infection (Marr et al. 2014).

**Toxoplasma gondii**: This protozoan is an obligatory intracellular parasite that infects a variety of immune and non-immune cells in human host. IFN-γ secreted by macrophages is the key player in controlling *T. gondii* pathogenesis. However, the protozoan itself is known to
downregulate the expression of IFN-γ inducible genes like MHC, iNOS and MIP in infected macrophages. This inhibition has been attributed to decreased histone acetylation and impaired recruitment of chromatin remodeling complexes at IFN-γ responsive promoters (Lang et al. 2012). Interestingly, HDAC inhibitors were found to reverse the inhibition of IFN-γ regulated H2-A/E and CIITA genes.

*Mycobacterium leprae:* A very novel and interesting finding has been the reprogramming of host Schwann cells to progenitor/Stem like cells by *M. leprae*. The mycobacterium infected cells showed loss of nuclear Sox10 (master regulator of Schwann cell homeostasis) and concomitant mesenchymal transition by upregulating several Hox and EMT-associated transcription factors. This reprogramming was associated with DNA hypomethylation in the promoter region of mesodermal/EMT genes and hypermethylation in SOX10 gene. Further, these pSLCs secreted all chemokines/cytokines known to be produced by mesenchymal cells. The immunomodulatory factors released by these pSLCs were found to have increased the chemotactic properties of tissue macrophages that in turn aided in the dissemination of mycobacterium (Masaki et al. 2013).

*Mycobacterium tuberculosis:* Control of *M. tuberculosis* infection is determined by IFN-γ levels in the host cells. The bacillus subverts this defense by counteracting the activation of IFN-γ inducible genes, particularly those relating to MHC class II complex. Infection with the mycobacterium up-regulates the expression of Sin3A that in concert with HDACs suppresses transcription of MHC class II genes (Wang et al. 2005). Mycobacterium activated TLR2/MAPK pathway also leads to recruitment of the repressive C/EBP to the promoter of CIITA gene (master regulator of MHC class II) (Pennini et al. 2006).

*Porphyromonas gingivalis:* This bacterium uses its metabolite to inhibit host proteins and in turn leads to reactivation of latent viruses. Butyrate produced by this bacteria leads to up-
regulation of anti-inflammatory genes in the host as well as inhibits host HDACs. This leads to reactivation of viruses like HIV and EBV that were silenced by the activity of HDACs (Imai et al. 2009; Imai et al. 2012).

**Direct Effect: Pathogen proteins modulating the host epigenome**

**RomA:** *Legionella pneumophilia* is a gram negative bacterium that causes Legionnaires’ disease, a severe form of pneumonia. The bacterium is known to modify multiple host pathways by inducing modification of host proteins like ubiquitination, phosphorylation, lipidation, glycosylation, AMPylation, deAMPylation, phosphocholination and dephosphorylcholination. The bacterium is also known to contain a plethora of eukaryotic-like proteins, one of them being a SET-domain containing protein RomA. This protein has been reported to be a host histone methyltransferase that methylates histone H3 at lysine 14. Following infection with this bacterium, global increase in H3K14me3 levels in THP1 macrophages was observed as early as 4hrs post-infection and this modification was found to be repressive in nature (Rolando et al. 2013).

**NuE:** *Chlamydia trachomatis*, unlike other bacterial pathogens, is able to survive only within a host cell and thus must subvert the host immune response by any means necessary. NuE, a SET-domain containing protein of the bacterium, has been characterized as a histone methyltransferase. NuE shows a strong affinity for H4 but also methylates H3 and H2B. The protein was shown to localize to the host nucleus upon infection but the specific target regions have not been identified yet (Pennini et al. 2010).

**LLO:** The intracellular pathogen *Listeria monocytogens* is known to use molecular mimicry for exploiting host defense pathways. Listeriolysin O (LLO) secreted by the bacterium belongs to a family of pore-forming toxins and is known to be involved in its escape from the
phagolysosome. Independent of this activity, LLO has been shown to induce dephosphorylation of H3S10 during early infection (Hamon et al. 2007). This dephosphorylation has been found to be triggered by LLO-induced K⁺ efflux from host cytoplasm upon infection (Hamon and Cossart 2011).

**OspF:** *Shigella flexneri* is an intracellular pathogen that causes bacillary dysentery. The bacterium contains a set of T3SS effector proteins, Osp (OspB-OspG) that modulate host cell signaling and transcription. OspF has been shown to inhibit H3S10 phosphorylation by dephosphorylating MAP kinases in the nucleus and hence interfering with activation of NF-κB responsive genes (Arbibe et al. 2007). OspB, another protein of the same family, has also been shown to localize to the host nucleus. Both OspB and OspF have been shown to repress IL8 secretion by infected epithelial cells and this effect has been proposed to be mediated by chromatin remodeling at cytokine genes by association of OspB-OspF with retinoblastoma protein (Rb) (Zurawski et al. 2009).

Recently, OspF has been shown to be directly associated with chromatin at the *IL8* promoter where it interfered with the binding of HP1γ. OspF also inhibited the phosphorylation of HP1γ by MSK-1, thus regulating its effect on transcription of immune genes and moderating the innate immune response (Harouz et al. 2014).

**AnkA:** The rickettsial pathogen *Anaplasma phagocytophilum* is known to infect the short-lived host neutrophils and alter its function. The infection results in inhibition of oxidative burst, apoptosis and phagocytosis but promotes degranulation and cytokine/chemokine production. In *A. phagocytophilum* infected cells, the expression of defense genes is downregulated. This repression was found to be mediated by decrease in levels of diacetylated H3 at the promoters and a concomitant increase in expression of HDAC1 and HDAC2 levels (Garcia-Garcia et al. 2009a). CYBB, cytochrome B-245, is a component of phagocyte oxidase and plays an
important role in defense against *A. phagocytoplilum*. This gene was one of the genes repressed in infected cells. AnkA (microbial ankyrin protein) protein of the pathogen was detected in the host cell nuclei where it was associated with binding sites at and upstream of the proximal promoter region of *CYBB* gene. These binding sites were AT-rich and were detected throughout the genome indicating a possibility of global host epigenetic modulation (Garcia-Garcia et al. 2009b).

**BaSET:** *Bacillus anthracis* is a spore-forming gram negative pathogen, notorious for its lethal anthrax toxin (a zinc dependent metalloproteinase). Recently, a SET-domain containing protein, BaSET, was characterized as a virulent factor, deletion of which resulted in the bacteria being avirulent. The protein was found to localize to the host nucleus where it trimethylated histone H1 at eight lysine residues. Moreover, BaSET acted as a transcriptional repressor of NF-κB responsive genes in a dose-dependent manner. Compared to *BaΔSET* bacillus, the wild type bacteria caused significant downregulation of TNF-α, IL6, c-fos and c-jun (Mujtaba et al. 2013).

**LPS:** Lipopolysaccharide (LPS) is a major component of outer cell wall of gram negative bacteria and a potent stimulator of inflammatory cytokines from immune cells. Binding of LPS to TLR4 was found to lead to histone acetylation and H3S10 phosphorylation at *IL12* promoter in T-cells (Weinmann et al. 2001). In another study LPS stimulation of epithelial cells was found to be associated with transient and dynamic changes in H3-acetyl, H3K4me2 and H3K9me2 at *IL8* promoter (Angrisano et al. 2010).

**6b:** *Agrobacterium tumifaciens* is a plant pathogen that transfers a part of its own DNA (T-DNA) into host cells. The integration of this T-DNA (carrying oncogenes for auxin and cytokinin enzymes) is facilitated by the interaction of the bacterial protein VirE2 with host
chromatin (Lacroix et al. 2008). Another protein, 6b, has been proposed to act as histone chaperone and binds to miRNA and histones in the host nucleus (Terakura et al. 2007).

**NS1:** Influenza A (H3N2) virus protein NS1 is known to contain a sequence resembling histone tail that enables it to act as a histone mimic in host cells. Specifically, it contains ‘ARSK’ motif that resembles the ‘ARTK’ motif comprising of lysine 4 of H3. The lysine in this motif was found to be methylated by host Set1 protein *in vitro*. This motif allows the protein to target human PAF1 transcription complex and suppress transcription elongation. PAF1 complex regulates expression of antiviral genes and hence its depletion by NS1 allows the virus to replicate and survive (Marazzi et al. 2012).

**Virus:** Viruses are more commonly known to affect DNA methylation than bacteria. KSH virus encodes an antigen LANA that can interact with cellular DNMTs (Shamay et al. 2006) as well as the methyl-C binding protein MeCP2 (Matsumura et al. 2010). Similarly, E7 protein from HP virus was shown to modulate DNMT1 levels (Laurson et al. 2010) as well as interact directly with DNMT1 (Burgers et al. 2007).
Figure 1.4: Interaction of pathogen proteins with host chromatin
The proteins highlighted in red interact directly with histones and modify them. Proteins highlighted in yellow trigger a signaling cascade that leads to histone modification. AnkA protein directly binds to promoter elements of target gene.

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Pathogen Effector protein</th>
<th>Host Epigenetic Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect Effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>H. pylori</em></td>
<td></td>
<td>DNA hypermethylation at promoter of <em>CDH1</em></td>
</tr>
<tr>
<td><em>C. rectus</em></td>
<td></td>
<td>DNA hypermethylation at promoter of <em>Igf2</em></td>
</tr>
<tr>
<td><em>L. donovani</em></td>
<td></td>
<td>Differential DNA methylation of immune response genes</td>
</tr>
<tr>
<td><em>T. gondii</em></td>
<td></td>
<td>Decreased histone acetylation at IFN-γ responsive genes</td>
</tr>
<tr>
<td><em>M. leprae</em></td>
<td></td>
<td>Reprogramming of Schwann cells to pSLCs.</td>
</tr>
<tr>
<td><em>M. tuberculosis</em></td>
<td></td>
<td>Recruitment of repressive complex to promoters of IFN-γ</td>
</tr>
<tr>
<td><em>P. gingivalis</em></td>
<td></td>
<td>Induction of HDACs by butyrate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. pneumophila</em></td>
<td>RomA</td>
<td>Methylation of histone H3 at lysine 14</td>
</tr>
<tr>
<td><em>C. trachomatis</em></td>
<td>NuE</td>
<td>Methylation of H4</td>
</tr>
<tr>
<td><em>L. monocytogenes</em></td>
<td>LLO</td>
<td>Dephosphorylation of H3S10</td>
</tr>
<tr>
<td><em>C. perfringes</em></td>
<td>PFO</td>
<td>Dephosphorylation of H3S10</td>
</tr>
<tr>
<td><em>S. pneumoniae</em></td>
<td>PLY</td>
<td>Dephosphorylation of H3S10</td>
</tr>
<tr>
<td><em>A. hydrophila</em></td>
<td>Aerolysin</td>
<td>Dephosphorylation of H3S10</td>
</tr>
<tr>
<td><em>S. flexneri</em></td>
<td>OspF, OspB</td>
<td>Dephosphorylation of H3S10 and HP1-γ</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td>AnkA</td>
<td>Direct association with promoter of <em>CYBB</em> gene</td>
</tr>
<tr>
<td><em>B. anthracis</em></td>
<td>BaSET</td>
<td>Trimethylates histone H1</td>
</tr>
<tr>
<td><em>A. tumifaciens</em></td>
<td>VirE2, 6b</td>
<td>Direct interaction with host chromatin</td>
</tr>
<tr>
<td>Influenza virus A</td>
<td>NS1</td>
<td>Histone H3 mimic</td>
</tr>
<tr>
<td>KSH virus</td>
<td>LANA</td>
<td>Interacts with host DNMTs</td>
</tr>
<tr>
<td>HP virus</td>
<td>E7</td>
<td>Interacts with host DNMT1</td>
</tr>
</tbody>
</table>

Table 1.3: Epigenetic modifications of host induced by pathogens
Pathogen induced indirect and direct epigenetic modifications of the host are listed. The name of the pathogen factor is provided wherever proof of direct interaction with the host epigenome is available.
1.8 Aim of the thesis: Role of DNA methylation in Host-Pathogen Interaction

Changes in the transcriptome of a cell closely parallel changes in its epigenome highlighting the pliability of the epigenome towards signals emanating from the environment. Amongst the various epigenetic modifications, DNA methylation is an important component of a mammalian cell’s epigenome (A. Bird 2002) and several studies have exemplified the dynamic nature of DNA methylation and its contribution in translating an environmental cue to a cellular response (Meaney and Szyf 2005). Interaction with an infectious agent also invokes a response from the infected mammalian cell that manifests as molecular changes including modulation of the epigenome. The susceptibility of a cell’s methylome to manipulation by an infectious agent has been reported but there is no direct evidence for modulation of host DNA methylation by the bacterial proteins. While histone modifications are dynamic, DNA methylation changes are associated with persistent states (Barter et al. 2012; Bonasio et al. 2010). For intracellular pathogens that need to stay undetected in the host cells, modulation of host methylome would serve that cause. That is to say, changes in host DNA methylation would eventually affect the cell transcriptome and hence the host defense response.

*Mycobacterium tuberculosis* is an ideal model pathogen to study the epigenetic modulation of host genome in response to infection. It has a GC-rich genome and cytosine methylation at non-canonical *Dcm* sites. Still a cytosine methyltransferase has not been reported till date. Absence of Dcm-like methyltransferase indicates mycobacterium DNMT might be distinct from other bacteria. It is possible that they may resemble hDNMT more closely and hence may be able to tweak the host methylome and interact with host epigenetic circuitry. This ability to affect the host methylome may in turn aid the mycobacteria in creating a microenvironment more conducive to its survival.
Based on these observations, this doctoral thesis aims to study the host DNA methylation changes that occur upon infection with *Mycobacterium tuberculosis* as well as to investigate the existence of a mycobacterial DNMT involved in such a modification of the host. The work is based on the following two objectives:

1. Identification of genomic loci within the mammalian host cells that show change in DNA methylation profile upon *M. tuberculosis* infection.

2. Characterization of a novel *Mycobacterium* methyltransferase and its role in inducing epigenetic changes in the host.