

ISSN 1996-3343

Asian Journal of  
**Applied**  
Sciences

## **An Insight into the Use of Genome, Methylome and Gethylome in Synthetic Biology**

Mohammed A. Ibrahim

Princess Sumaya University for Technology, Jubaiha 11941, Amman, Hashemite Kingdom of Jordan

### **ABSTRACT**

The development of new molecular approaches in biology and biotechnology introduced new scientific terms which include genome, genomics, methylome and methylomics. These are currently used in the literature of molecular and synthetic biology to refer various molecular concepts and applications. The present study is about the requirement of new “ome” and “omics” terms. This is because there are unique structural and functional aspects associated with genome and methylome at specific regions in the DNA sequence which are not explained by the two currently used terms methylome and genome. There is a requirement for new ome/omics terms to refer specific structural and functional matters associated with particular specialized zones in the DNA sequences. These genomic regions contain both the fifth base “methyl cytosine” and the other four coding bases. Methylated and non methylated DNA sequences are involved in gene expression of RNA and proteins, beside their role in the structural organization of the DNA sequence and the chromosome. In this study I suggest introducing two words which are derived from the established scientific terms “genome, genomics, methylome and methylomics”. The two new derived words are: gethylome and gethylomics. I suggest that these two words might be useful and could be used to explain various issues related to specific regions in the DNA sequence and chromosomes of eukaryotic organisms which contain the fifth base. Furthermore, gethylomics will help to refer more precisely to new applications in synthetic biology and genomics to design or redesign specific gethylomic circuits.

**Key words:** Genome, genomics, methylome, methylomics, gethylome, gethylomics, synthetic biology, gethylomic circuits

### **INTRODUCTION**

During last two decades, active research works in molecular biology have introduced great achievements in the new basic and applied knowledge of molecular biotechnology in a rationally motivated process for understanding of various structural and functional aspects of genetic materials. As a result, molecular biologists have faced the challenge to find new proper words to define new concepts in the field of molecular biotechnology which have been continually introduced. These words are required for proper explanation of the newly gained knowledge, models, conceptions and views associated with the scientific efforts to investigate and understand the mechanisms of various genetic phenomena at the molecular level. Consequently new scientific terms have been introduced to define concepts related to new discoveries at the molecular level in biology and biotechnology. Molecular biologists and other scientists in particular those interested in bioinformatics started to use widely two morphemes. These are “-ome” and “-omic”, each is added as suffix at the end of many scientific terms to form the derivatives. Scientists working in the field

Table 1: Main four scientific terms with suffix (-ome or -omics) used in molecular genetics

| Terms with suffix (-ome) | Terms with suffix (-omics) | References  |
|--------------------------|----------------------------|---|
| Epigenome                | Epigenomics                | Lederberg and McCray (2001), Ibrahim (2010a, b), Zhang <i>et al.</i> (2010) and |
| Genome                   | Genomics                   | Ibrahim (2011)  |
| Methylome                | Methylomics                |   |
| Transcriptome            | Transcriptomics            |   |
| miRNAome                 | miRNAomics                 |   |

of molecular biology and bioinformatics use (-ome) suffix to refer to a totality of some sort. This comes from the Oxford English Dictionary which distinguishes three different fields of application for the (-ome) suffix. The one which is used in the field of molecular biology is forming nouns with the sense "all constituents considered collectively" (Lederberg and McCray, 2001). Reviewing the literature in the field of molecular biology and bioinformatics showed there is long list of scientific terms with (-ome) suffix. On the other hand, Omics.org defined "omics" as a "general term for a broad discipline of science and engineering for analyzing the interactions of biological information objects in various omes". Furthermore, Omics.org indicated that the main focus of omics is on 1) mapping information objects such as genes, proteins and ligands, 2) finding interaction relationships among the objects, 3) engineering the networks and objects to understand and manipulate the regulatory mechanisms and 4) integrating various omes and omics subfields. A more scholastic definition of "omics" was reported by Vallero (2010), who defined omics as a "shorthand term for computational, biological subfields for describing very large-scale data collection and analysis, all with the suffix -omics".

The data in Table 1 show the main terms with suffix -ome or -omics currently used in the literature of molecular genetics and biotechnology.

## GENOME AND GENOMICS

It is believed that the term "genome" was first presented in 1920 and it was attributed to Hans Winkler, who proposed the expression "genom" for the haploid chromosome set (Lederberg and McCray, 2001). On the other hand, the term "genomics" was introduced in September 1987 by McKusick and Ruddle (1987a, b) as a name for their newly established journal. Various definitions have been reported for genome and genomics; Genome org defined the genome as "the totality of the genetic material"; whereas the term "genomics" is defined by the same source as "the omics approach research of genome in biology". Another definition was reported by Nill (2000) stated the following: The scientific study of genes and their role in an organism's structure, growth, health, disease, resistance to disease and their contribution to the shape, function and the development of those whole organisms. Recently, Vallero (2010) defined the genome as an "entire genetic complement, i.e., all of the hereditary material possessed by an organism". The same author gave two definitions for genomics: 1) study of genes, including their functions 2) study of the molecular organization of genomes, their information content and the gene products they encode". In this context it is worth mentioning that Lederberg and McCray (2001) indicated in their article on the "ome" and "omics" that the word "genomics" has "the same narrower connotation today, of emphasis on linear gene mapping and DNA sequencing". At the end of this section I would like to sum up all above mentioned definitions of genome and genomics and say that authors do agree on the general concept which is concerned and emphasized on the basic linear molecule of DNA. Thus it is possible to state that genome and genomics are scientific terms which refer to the structure and function of DNA sequence which is composed of four nucleotides: adenine, cytosine, guanine and thymine.

Table 2: The main functions of DNA methylation

| Functions of DNA methylation                       | References                       |
|--|----------------------------------|
| Embryonic development and cellular differentiation | Robertson (2005) and Szyf (2009) |
| Genomic imprinting                                 | Tucker <i>et al.</i> (1996)      |
| Aging  | Fraga and Esteller (2007)        |
| Chronic diseases                                   | Ogoshi <i>et al.</i> (2011)      |
| Histone-methylome interaction                      | Razin (1998)                     |
| Hot spots of mutations                             | Lutsenko and Bhagwat (1999)      |

## METHYLOME AND METHYLOMICS

Recent molecular studies have indicated that methylation of genomic DNA is widely prevailed in most eukaryotic species and it is an ancient property of these organisms which is characterized by conservative phylogenetic features and it is one of important epigenetic mechanisms which control gene expression (Ibrahim *et al.*, 2009; Feng *et al.*, 2010; Ibrahim, 2010a-c; Jeltsch, 2010; Zemach *et al.*, 2010; Ibrahim, 2011). DNA methylation is a result of addition of a methyl group at position 5' of the cytosine pyrimidine ring next guanine in CpG dinucleotides. Consequently, DNA methylation might disrupt the binding of transcription factors and draws methyl-binding proteins which are linked with gene silencing and chromatin packaging (Weissbach, 1993; Stratthdee and Brown 2002; Ibrahim, 2010a, b). Thus, methylated cytosine is considered the fifth nitrogen base in the eukaryotic genome beside cytosine, guanine, thymine and adenine (Adams, 1990; Lister and Ecker, 2009). The wide-spread interest in investigation the DNA methylation profile (methylome) of human's genome followed the great achievement in sequencing of human genome. Currently research is ongoing to elucidate how the genome executes the information it holds and the role of DNA methylation in this process (McKusick and Ruddle, 1987a, b; Beck and Rakyan, 2008; Lister and Ecker, 2009; Bibikova *et al.*, 2011). The term methylome is now accepted in the scientific literature, it refers to "the totality of methylated DNA sites in a genome, cells and tissues" (Methylome.Org). On the other hand, the term "methylomics" refers to "systematic research that maps the histone codes and methylation patterns of healthy genomes" (Methylomics.org).

One of interesting points related to the concepts of methylome and methylomics is that although all human's nucleated cells effectively contain the same genome, they contain very different DNA methylation profiles (Guil and Esteller, 2009). In addition, it has been reported that DNA methylation is associated with changes in the cells phenotypes (Baron *et al.*, 2006), products of gene expression and in development of various types of cancers (Ibrahim, 2010a, b; Ogoshi *et al.*, 2011), aging (Brunet and Rando, 2007; Fraga and Esteller, 2007), fragile X syndrome (Oostra and Willemsen, 2002), Beckwith-Weideman syndrome (Maher and Reik, 2000) and pathogenesis of psychiatric disorders (Mill *et al.*, 2008; Ibrahim, 2010c). Investigation in this field indicated that DNA methylation might be stable and reflect long term characteristics and persistent commitment along a cell type and lineage (Baron *et al.*, 2006). The data presented in Table 2 summarize some of critical roles which are played by DNA methylation.

## GETHYLOME AND GETHYLOMICS

In the previous two sections I reviewed and discussed the concepts and definitions of genome, genomics, methylome and methylomics. It is worth mentioning that both genome and methylome share same structure, i.e., the double helix. However, the only difference between two is the fifth nitrogen base, the methylated cytosine (Adams, 1990; Lister and Ecker, 2009). And one more important point is that the DNA sequences which contain methyl cytosine are few as compared with

non methylated regions. The reported results showed that approximately 1% of bases in a somatic human genome are methyl-cytosines which equates to 70-80% of all CpG dinucleotides in the genome (Ehrlich *et al.*, 1982). In this respect, it is important to distinguish between genome and methylome by noting that there are differences in functional and structural aspects of both. The main function of the genome is coding the information for protein and RNA production, whereas methylome has completely different function which is associated with certain aspects of control gene expression. Furthermore, methylated cytosines are the sites on the DNA sequence for structural changes of genome. The changes at these sites are a result of occurrence of high rates of mutations as compared with mutation rates which might occur in the normal dominant sites which are composed of other four bases. These sites which contain methyl cytosine are known as “hot spots” (Lutsenko and Bhagwat, 1999). Another important point in this context is that hypomethylated DNA is associated with active regions of the chromatin which enable gene expression, whereas hypermethylated DNA is found in inactive chromatin (Razin and Cedar, 1977; Razin, 1998). Taking into account these considerations which are related to structure and function of methylome and genome, it might be possible to understand the need for new “ome” and “omics” to explain various basics and applications of the uniqueness of each form of DNA and the interaction between both. Another point which is in support of this argument is about the direct or indirect interactive states of methylome, genome and histone(s) which might have very useful applications in synthetic biology. Accordingly, it is worth considering the possibility of using new derived word from genome and methylome to define and refer to specific molecular regions which are composed of interactive structural and functional forms of methylome, genome and histone(s). Hereby I suggest the word “gethylome” which refers to specific interactive part(s) of genome and methylome with histone and possibly other molecules. These molecular regions might form specific sets and systems and are found in various regions of chromosomes. An example of gethylomic regions are methylated CpG islands in the promoter regions and hot spots. Various molecular tools (molecular biology, nano-biotechnology, bioinformatics and synthetic biology tools) could be used to investigate the structural and functional properties of gethylomic regions. Such studies might be conducted in the field of “gethylomics”. Accordingly, I propose the following definition for gethylomics: Studying of the molecular organization and functions of gethylomes and their contribution and application in synthetic biology (Ibrahim, 2011).

## **SYNTHETIC BIOLOGY AND GETYLOMICS CIRCUITS**

Synthetic biology is broadly defined as “the re-design and fabrication of existing biological systems” ([www.syntheticbiology.org](http://www.syntheticbiology.org)), this process includes engineering the networks and objects to understand and manipulate the regulatory mechanisms. To perform this task there is a necessity to build up genetic circuits and to uncover the design principles of natural biological systems through the rational design of gene and protein circuits (Mukherji and van Oudenaarden, 2009). Considering the role of DNA methylation in gene expression, the interaction with histones and the environment (Ibrahim, 2011; Zhang *et al.*, 2010), it is possible to imagine the importance of designing and redesigning gethylomic genetic circuits. These circuits will have great contribution and impact in improving the human’s healths and might find other applications.

A recent published article showed that DNA biomolecules are suited substrates for self-assembly and have proved to be a powerful scaffold with which it is possible to design and build molecular devices with multiple forms and function (Bhatia *et al.*, 2011). In this respect, there is another important point to add in this argument, that is the interaction between the chromatin [chromatin

is the complex of DNA and (histone) protein of which the chromosomes are composed] and methylated regions of DNA sequence. This interaction should be considered since active regions of the chromatin which enable gene expression, are associated with hypomethylated DNA whereas hypermethylated DNA is packaged in inactive chromatin (Razin and Cedar, 1977; Razin, 1998). These chromosomal regions might be good candidates for redesigning chromosomal circuits. Accordingly, I expect that studies in the field of “synthetic gethylomics” will have important applications in designing or redesigning of future gene molecular circuit devices.

## CONCLUSIONS

Gethylomes are specific regions on the chromosomes which contain methyl cytosine and refer to specific interactive part(s) of genome and methylome with histone and possibly other molecules. These regions play important role in gene expression and structural changes of the genome. On the other hand, gethylomics focus on studies of the molecular organization and functions of gethylomes and their contribution and application in synthetic biology. The interactive components of glethyomic systems and their crucial role in gene expression will facilitate possible designing or redesigning of gethylomic circuits for future applications in synthetic biology.

## REFERENCES

- Adams, R.L., 1990. DNA methylation-the effect of minor bases on DNA protein interactions. *Biochem. J.*, 265: 309-320.
- Baron, U., I. Turbachova, A. Hellwag, F. Eckhardt and K. Berlin *et al.*, 2006. DNA methylation analysis as a tool for cell typing. *Epigenetics*, 1: 55-60.
- Beck, S. and V.K. Rakyan, 2008. The methylome: Approaches for global DNA methylation profiling. *Trends. Genet.*, 24: 231-237.
- Bhatia, D., S. Sharma and Y. Krishnan, 2011. Synthetic, biofunctional nucleic acid-based molecular devices. *Curr. Opin. Biotechnol.*, 22: 475-484.
- Bibikova, M., B. Barnes, C. Tsan, V. Ho and B. Klotzle *et al.*, 2011. High density DNA methylation array with single CpG site resolution. *Genomics*, 98: 288-295.
- Brunet, A. and T.A. Rando, 2007. Ageing: From stem to stern. *Nature*, 449: 288-291.
- Ehrlich, M., M.A. Gama-Sosa, L.H. Huang, R.M. Midgett and K.C. Kuo *et al.*, 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.*, 10: 2709-2721.
- Feng, S., S.J. Cokus, X. Zhang, P.Y. Chen and M. Bostick *et al.*, 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. USA.*, 107: 8689-8694.
- Fraga, M.F. and M. Esteller, 2007. Epigenetics and aging: The targets and the marks. *Trends Genet.*, 23: 413-418.
- Guil, S. and M. Esteller, 2009. DNA methylomes, histone codes and miRNAs: Tying it all together. *Int. J. Biochem. Cell Biol.*, 41: 87-95.
- Ibrahim, M.A., N. Saleh, K.M. Mousawy, N. Al-Hmoud, E. Archoukieh, H.W. Al-Obaide and M.M. Al-Obaidi, 2009. Molecular analysis of RAPD-PCR genomic patterns in age related acute myeloid leukemia. *Trends Med. Res.*, 4: 35-41.
- Ibrahim, M.A., 2010a. Advances in genomic DNA methylation analysis. *Biotechnology*, 9: 459-468.
- Ibrahim, M.A., 2010b. Perspective of DNA methylation in cancer research. *Int. J. Cancer Res.*, 6: 188-201.

- Ibrahim, M.A., 2010c. Prospects for DNA methylation research in psychiatric disorders. *Trends Med. Res.*, 5: 1-15.
- Ibrahim, M.A., 2011. Implication of environmental health hazards in aberration of DNA methylation and prevalence of diseases. *Proceedings of the Biosafety and Biosecurity International Conference, September 13-15, 2011, Jordan.*
- Jeltsch, A., 2010. Molecular biology. Phylogeny of methylomes, *Science*, 328: 837-838.
- Lederberg, J. and A. T. McCray, 2001. Ome sweet omics: A genealogical treasury of words. *The Scientist*, 154: 8.
- Lister, L. and J.R. Ecker, 2009. Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Res.*, 19: 959-966.
- Lutsenko, E. and A.S. Bhagwat, 1999. Principal causes of hot spots for cytosine to thymine mutations at sites of cytosine methylation in growing cells A model, its experimental support and implications. *Mutat. Res.*, 437: 11-20.
- Maher, E.R. and W. Reik, 2000. Beckwith-Weideman syndrome: Imprinting in cluster revisited. *J. Clin. Invest.*, 105: 247-252.
- McKusick, V.A. and F.H. Ruddle, 1987a. A new discipline, a new name: A new journal. *Genomics*, 1: 1-2.
- McKusick, V.A. and F.H. Ruddle, 1987b. Toward a complete map of the human genome. *Genomics*, 1: 103-106.
- Mill, J., T. Tang, Z. Kaminsky, T. Khare and S. Yazdanpanah *et al.*, 2008. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am. J. Human Genet.*, 82: 696-711.
- Mukherji, S. and A. van Oudenaarden, 2009. Synthetic biology: Understanding biological design from synthetic circuits. *Nat. Rev. Genet.*, 10: 859-871.
- Nil, K.R., 2000. *Glossary of Biotechnology Terms*. 3rd Edn., CRC Press, Florida, USA., ISBN 1-58716-122-2.
- Ogoshi, K, S. Hashimoto, Y. Nakatani, W. Qu and K. Oshima *et al.*, 2011. Genome-wide profiling of DNA methylation in human cancer cells. *Genomics*, 98: 280-287.
- Oostra, B.A. and R. Willemsen, 2002. The X chromosome and fragile X mental retardation. *Cytogenet. Gen. Res.*, 99: 257-264.
- Razin, A. and H. Cedar, 1977. Distribution of 5-methylcytosine in chromatin. *Proc. Natl. Acad. Sci. USA*, 74: 2725-2728.
- Razin, A., 1998. CpG methylation, chromatin structure and gene silencing: A three-way connection. *EMBO J.*, 17: 4905-4908.
- Robertson, K.D., 2005. DNA methylation and human disease. *Nat. Rev. Genet.*, 6: 597-610.
- Strathdee, G. and R. Brown, 2002. Aberrant DNA methylation in cancer: Potential clinical interventions. *Expert. Rev. Mol. Med.*, 4: 1-17.
- Szyf, M., 2009. The early life environment and the epigenome. *Biochim. Biophys. Acta*, 1790: 878-885.
- Tucker, K.L., C. Beard, J. Dausman, L. Jackson-Grusby and P.W. Laird *et al.*, 1996. Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of non imprinted genes. *Genes Dev.*, 10: 1008-1020.
- Vallero, D.A., 2010. *Environmental Biotechnology: A Biosystems Approach*. Elsevier Inc., Oxford, UK., ISBN: 978-0-12-375089-1.

- Weissbach, A., 1993. A Chronicle of DNA Methylation (1948-1975). In: DNA Methylation: Molecular Biology and Biological Significance, Jost, J.P. and H. Saluz (Eds.). Birkhauser Verlag, Berlin, ISBN: 3-7643-2778-2, pp: 1-10.
- Zemach, A., I.E. McDaniel, P. Silva and D. Zilberman, 2010. Genome-wide evolutionary analysis of eukaryotic DNA. *Methylation Science*, 328: 916-919.
- Zhang, L., C.M. McHale, N. Rothman, G. Li and Z. Ji *et al.*, 2010. Systems biology of human benzene exposure. *Chem. Biol. Interact.*, 184: 86-93.