J Biosci Vol. 43, No. 2, June 2018, pp. 239–242
DOI: 10.1007/s12038-018-9742-9

Brief Communication

Evolution of genome organization and epigenetic machineries

Tapas K Kundu* and Sweta Sikder

Transcription and Disease Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bengaluru 560 064, India

**Corresponding author (Email, tapas@jncasr.ac.in)*

MS received 30 January 2018; accepted 7 February 2018; published online 20 March 2018

The unit content of DNA in a living organism containing all of its genes is referred to as 'genome', which forms the basis of life and heredity. In the path of evolution from a single-cellular prokaryotic life to multi-cellular eukaryotic system, the genome has become more and more complex not only in the context of size but also in sequence and content. Although the size of the genome does not directly correlate with the complexity and hierarchy of a living organism, in the eukaryotic system with restricted nuclear size, disproportionately higher size of DNA is packed into a highly ordered manner (Kumari et al., in Kundu (eds) Epigenetics: development and disease. Subcellular biochemistry, Springer, Dordrecht, 2013). This packaging of genome with the assistance of protein and RNA is not a unique feature of the eukaryotic system alone. In fact, even in a tiny prokaryotic cell, to reduce the volume and to restrict the signal dependent availability of the genomic material (DNA), packing of genome is essential. In this brief article, we shall try to put forward a concept of genome organization and epigenetic machineries to assist the functional ability of the genomic material in the evolutionary perspective.

Keywords. Chromatin dynamics; epigenetic language; evolution of histones; genome organization; nucleoid structures

As the 1.6 mm long bacterial genome needs to be compressed into a 2 μm long, and 1 μm wide cell (Sherratt 2003), the DNA molecule in an *E. coli* cell needs to be condensed almost 400 times. The compact folded structure of bacterial genomic material is termed as nucleoid. The Nucleoid Associated Proteins (NAPs) assists in the folding of the super-coiled topological domains of bacterial genome thereby controlling its availability for different DNA templated phenomena including transcription, replication and repair. This compact structure of bacterial genome varies profoundly both in terms of structure, topology and composition of proteins (as described in the next section). Evolutionarily in this context, prokaryotic genome organization significantly differs from the orderliness of eukaryotic genome where it is organized as a universal 'beads on a string' structure (to be discussed in the following section). Counterintuitively, prokaryotic genomes carry structurally less-defined basic unit of genome packing, while the universal 'beads on a string' model strictly defines the basic unit in the genome compaction hierarchy of all eukaryotes. Forces of genome evolution might have tuned this unit of 'beads on a string' towards multiple states of regulation.

There are different classes of NAPs, which induce the genome compaction in prokaryotic system at different stages of their life cycle. These proteins are classified functionally as DNA bridging proteins (e.g. H-NS, SMC complex, Lrp), DNA bending proteins (e.g. IHF, HU, Fis) and nonspecific compactor (e.g. Dps). Interestingly, the content of these proteins significantly varies during the different phases of the bacterial life cycle. The concentration of Fis protein is found to be highest in the early exponential phase of the bacterial life cycle, whereas the HU protein takes over in the exponential phase. While the bacteria reaches the stationary phase the most dominant NAP to be found is IHF. Dps (the DNA binding protein from starved cells) specifically appears at the stationary phase when the genome should be organized into a compact structure to ensure that most of the genes are shut down, whereas the DNA bending proteins are predominant in the early and late exponential phase to facilitate the recruitment of transcription machineries to complete the life cycle within a span of 20 minutes. It is intriguing to know that this phenomenon of variable genome composition in accordance with the functional state of the genome is well conserved across several living organisms during the course of evolution. If we look into parasite (e.g. Malaria parasite, *Plasmodium sp.*) where the life cycle is completed in two different hosts, the composition of nucleosomal histones changes from one stage of life cycle to another stage. Even in the higher eukaryotic system, during the process of differentiation from pluripotent to the differentiated stage of a cell, there is a dramatic

alteration both in the terms of compactness and composition of the orderly organized nucleosomes.

Moving up in the phylogenetic tree, the next evolved prokaryotes after bacteria are 'Archaea'. Significantly, the genome in archaea is more orderly organized with primitive histone-like molecules (Pereira *et al.* 1998). The archaeal histones are smaller in size lacking the extended N and C terminals. In the eukaryotic system, there are nucleosomes containing four different highly conserved histones: H3, H2B, H2A and H4. In archaea bacteria only homologues of histone H3 and H4 are found. In this form of life, the H2A and H2B are yet to appear. The primitive histones present in archaea bacteria are HMfA and HMfB, which combine with each other or other different proteins to generate the primitive type of nucleosome referred to as subnucleosomal particle. Keeping the evolutionary phenomenon in flow, the concentration of these HMf proteins also changes in different phases in the life of archaea bacteria (for example, the level of HMfB increases as the cell approaches to the stationary phase). However, some of the HMf interacting protein seems to possess ATP binding domain, raising the possibility of primitive chromatin remodeling system. Besides HMf protein, the archaea bacteria also contain other DNA compacting protein to orchestrate higher-order folding of the genome. These proteins include Alba, MC1 and HU. Therefore, it seems that from archaea bacteria onwards, the order compaction of genome is initiated.

In the evolutionary tree, the highly ordered genome compaction and organization is first observed in the parasites (e.g. Malaria parasite, *Plasmodium falciparum*) where all the four canonical core histones and even their variants are present. These histones already attain the conserved sequences as observed in the higher eukaryotes (Cui and Miao 2010; Henikoff and Smith 2015). There are only eight amino acid replacements between plasmodium histone H4 and human H4, five of which are conserved substitutions. However, unlike eukaryotic organisms where we observe multiple copies of histone genes per haploid genome, plasmodium genome has a single copy of all the core histones. The concentration of these core histones is not same all along the complex life cycle of the parasite. As is observed in bacteria or Archaea bacteria, different histones and their variants are highly variable during different stages of parasite life. The core histones are predominantly expressed in late trophozoite and schizont phase where the DNA synthesis occurs. However, throughout the intraerythrocytic asexual life cycle, marked deposition of H2A.Z is observed along with the histone modification marks which favors transcriptional activation (Petter *et al.* 2011). This brings out the aspect of the constantly evolving organization of the genome in terms of its composition, conserving the underlying phenomenon/forces which governs its functionality.

In order to regulate each biological phenomenon in a more precise manner and to respond to different environmental cues, the genome complexity and size both increased in the consequent evolutionary stages. The most remarkable signature observed is the highly ordered, positioned nucleosomal organization of genome in the eukaryotic system. Besides the centromeric region, in all the eukaryotes 170 bp DNA is wrapped around an octameric nucleosomal core comprising of histones H3, H2B H2A and H4 in two copies of each (Kumari *et al.* 2013).

To fine-tune the functional regulation of genome function, the canonical histones are replaced by the respective variants which seems to be essential in several phases of life. Although in higher eukaryotic systems, genes for canonical histones are present as gene clusters, genes for histone variants are present as a single gene similar to malaria parasite. The nucleosomes in eukaryotic system are regularly spaced and organized in such a way that it can form further higher-order chromatin domains comprising of loosely packed (euchromatin) and highly compact (heterochromatin) chromatin fiber (figure 1). The complexity in the genome organization of higher eukaryotes is brought out by the higher-order folding of the beads on a string structure mediated by the linker histone H1. These linker histones are absent in the parasite as well as in the other single cellular eukaryotes. where the higher-ordered compaction of the genome is not found. During the process of differentiation, chromatin structure is not compact unlike in the differentiated cells where the chromatin is present in a highly ordered compact conformation with minimum access to the genomic DNA.

In order to gain access to genomic DNA in a signal dependent manner, the compact and organized genomes need to be constantly folded and unfolded. This chromatin dynamics fine-tune the gene expression as well as most of DNA templated phenomenon. The dynamic state of the chromatin is regulated by DNA methylation, post translational modifications of histones and non-coding RNA. Since these elements function above the so-called gene sequence, the phenomenon is termed as 'epigenetics'. The fundamental modification of histone-like proteins and histones are observed from prokaryotic system to the most evolved higher, multi-cellular eukaryotes (Baetson *et al.* 2014). For example, acetylation of lysine on these proteins is highly predominant and functionally conserved in a wide array of living organisms. Recent evidences suggest that, like eukaryotic system, there might be large number of lysine methyltransferases in the prokaryotic world as well (Zilberman *et al.* 2008). The acetyltransferases are found in prokaryotes, archaea, parasite, unicellular eukaryotes and also multicellular eukaryotes. However, the complex regulatory mechanisms of these enzymes have gradually got fine tuned during the course of evolution. The existence of non-coding RNA could also be noticed even in the prokaryotic system. Among all the epigenetic machineries, the complexity of the evolution of DNA methylation is remarkable. Although DNA methylation is observed in prokaryotic system, it is yet to be documented in the unicellular

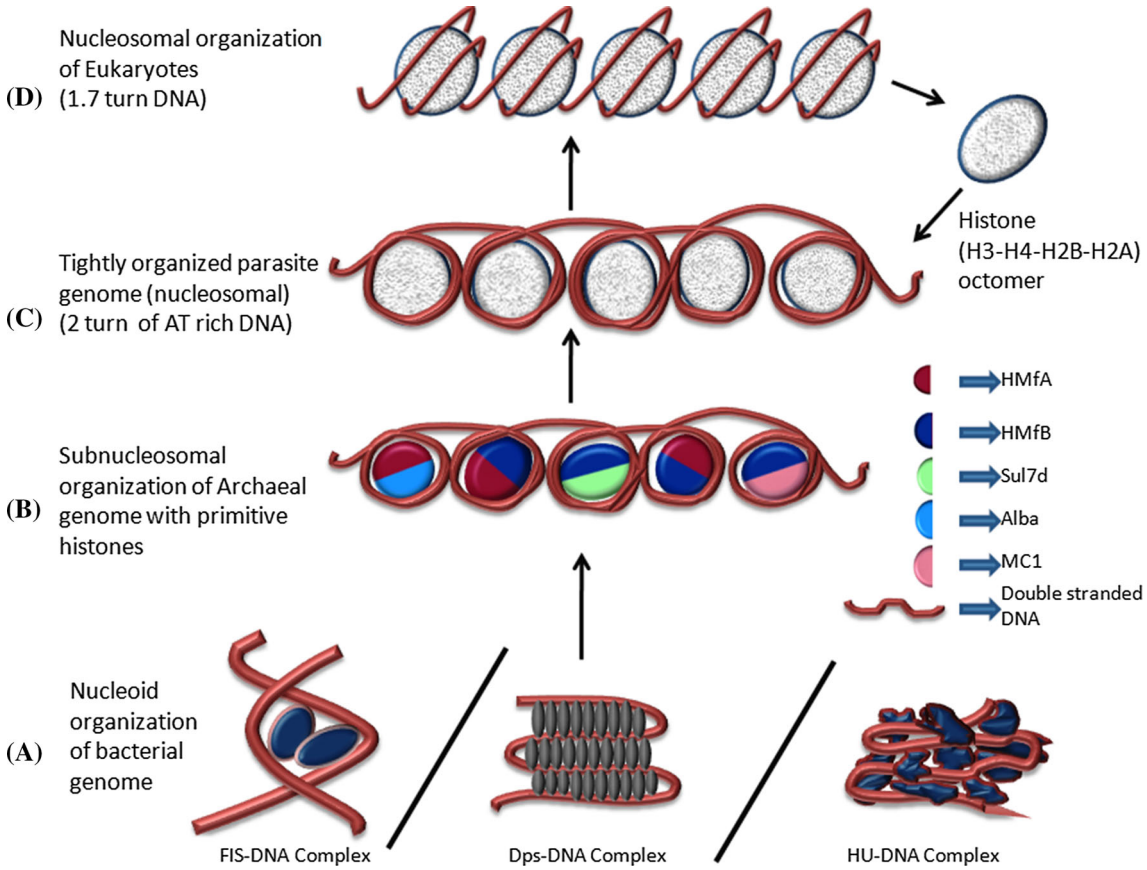

Figure 1. Evolution of the nucleoprotein organization of the genome: (**A**) In bacteria, genome compaction or organization is mediated by several DNA binding proteins (Fis, DPS, HU, etc.) in an irregular fashion. (**B**) In Archaea, primitive histone-like proteins, along with some other DNA binding proteins, form the subnucleosomal particle and organize the genome in a relatively ordered fashion. (**C**) In parasites, all the four core histones assemble in an ordered nucleosomal array, where two turns of AT rich DNA wrap around the octameric core. (**D**) In Eukaryotes, histone octamers are assembled orderly as a beads on string, where 1.7 turns of DNA wraps around the nucleosome to form the primary level organization, which further folds and form a higher-ordered chromatin organization.

eukaryotes (e.g. budding yeast). Apart from the architectural proteins of the genome, there are several DNA elements that have significantly contributed in the organization of the genome. Furthermore, in the course of evolution there have been enormous duplication of genes by means of recombination and enhancement of ploidy, which has definitely provided an added advantage to the constantly evolving living organisms but also created a huge pressure to organize extra-genomic material in the tiny cell nucleus and decode the encrypted message whenever it is necessary. Presumably the highly coordinated epigenetic language has been utilized for this purpose in higher organisms. In this short article, however, we have focused only on the structural components of the genome, their dynamics and modification and their evolution across time.

This perspective helps us to look in retrospection the genome composition and dynamics from simple unicellular bacteria to the complex chromatin fibers of the eukaryotes (figure 1). Besides the natural evolutionary pressure, the artificial stress created by human beings in terms of therapeutics as well as environmental stress presumably influences the rapid alteration in the genome dynamics. This article may be helpful to align our understanding of disease biology in terms of genetics and epigenetics attributing to the field of diagnostics and drug discovery.

Acknowledgements

We acknowledge Ms Kruthi HT for designing and compilation of the figure. The authors' laboratory is supported by funding from Jawaharlal Nehru Centre for Advanced Scientific Research, Department of Science and Technology (DST), India, and Department of Biotechnology (DBT), India. TKK is a recipient of the Sir JC Bose National Fellowship, Department of Science and Technology, India. SS is a Senior Research Fellow of Council of Scientific and Industrial Research (CSIR), India.

References

Bateson P 2014 Evolution, epigenetics and cooperation. *J. Biosci*. **39** 191–200

Cui L and Miao J 2010 Chromatin-mediated epigenetic regulation in the malaria parasite *Plasmodium falciparum*. *Eukaryot. Cell* **9** 1138–1149

Henikoff S and Smith MM 2015 Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* **7** a019364

Kumari S, Swaminathan A, Chatterjee S, Senapati P, Boopathi R and Kundu TK 2013 Chromatin organization, epigenetics and differentiation: An evolutionary perspective; in *Epigenetics: Development and disease. Subcellular biochemistry*, T Kundu (ed) (Springer, Dordrecht)

Pereira SL and Reeve JN 1998 Histones and nucleosomes in Archaea and Eukarya: a comparative analysis. *Extremophiles* **2** 141–148

Petter M, Lee CC, Byrne TJ, Boysen KE, Volz J, Ralph SA, Cowman AF, Brown GV and Duffy MF 2011 Expression of P. falciparum var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog.* **7** e1001292

Sherratt DJ 2003 Bacterial chromosome dynamics. *Science* **301** 780–785

Zilberman D 2008 The evolving functions of DNA methylation. *Curr. Opin. Plant Biol*. **11** 554–559

Corresponding editor: B Jagadeeshwar Rao