

Genome wide snRNP motifs and regulatory sequences in HIV1 isolates

Paushali Roy*, Protip Basu, Sayak Ganguli and Abhijit Datta

DBT-Centre for Bioinformatics, Presidency College, Kolkata, India.

*corresponding author: paushali.06@gmail.com

Abstract- The pathogenesis of HIV-1 is complex and characterized by the interplay of both viral and host factors. Within HIV 1 genome there are several snRNP motifs responsible for pre mRNA splicing and stabilization. By locating these motifs within the genome and disturbing them may result in an impaired ability of the cells to sustain HIV-1 replication. One of such regulatory sequences is riboswitches that regulate the dimerization of HIV-1 RNA, which is an essential step during packaging. The current work was undertaken to identify possible regulatory RNA motifs in the HIV1 genome from different isolates. The current work has successfully identified multiple snRNP motifs in the genome sequences of different strains of HIV-1 isolates. The identification of the multiple snRNP motifs in the genomic sequences of the various isolates lead us to believe that future studies with artificially constructed snRNPs might have the potential to inhibit HIV1 replication. Apart from containing snRNP motifs they also possess regulatory riboswitch motifs. Riboswitches bind metabolites and control the dimerization and packaging of the genome. Thus the occurrence of such motifs further strengthens the idea that apart from serving as a regulatory domain for structural constraints such motifs may also regulate genome integration and production of the necessary products by using the host transcriptional machinery. It is however beyond doubt that such sequence motifs must have originated in the RNA world as they have the power to mediate RNA induced regulation of gene expression.

Keywords: Riboswitch, snRNP motifs, HIV-1, gene regulation, RNA processing

INTRODUCTION

The human immunodeficiency virus type 1 (HIV-1) is the primary cause of the acquired immunodeficiency syndrome (AIDS), which is a slow, progressive and degenerative disease of the human immune system. The pathogenesis of HIV-1 is complex and characterized by the interplay of both viral and host factors. Despite years of intensive research and some therapeutic success, AIDS, continues to be a major health problem worldwide. It is a type of lentivirus and widely recognized as the etiologic agent of acquired immunodeficiency syndrome (AIDS). It is characterized by its cytopathic effect and affinity for the T4-lymphocyte. The strains of HIV-1 can be classified into three groups: the "major" group M, the "outlier" group O and the "new" group N. These three groups may represent three separate introductions of simian immunodeficiency virus into humans [1]. The enzyme reverse transcriptase (RT) is used by HIV 1 (a retrovirus) to transcribe their single-stranded RNA genome into single-stranded DNA and to subsequently construct a complementary strand of DNA, providing a DNA double helix capable of integration into host cell chromosomes. Functional HIV-1-RT is a heterodimer containing subunits of 66 kDa (p66) and 51 kDa (p51) [2]. Many viral or cellular genes are involved in HIV-1 multiplication and therefore represent potential targets. Indeed, several strategies attempting to interfere with the production or function of such gene products are being tested at pre-clinical or clinical levels. Within HIV 1 genome there are several snRNPs (small nuclear ribonucleoproteins) motifs mainly responsible for pre mRNA splicing and stabilization. By locating these motifs within the genome and disturbing them may result in an impaired ability of the cells to sustain HIV-1 replication. HIV-1 pathogenesis is multifactorial and involves complex interactions between host

and viral genes [3]. Several regulatory sequences that play significant role in HIV-1 infection have been so far identified. One of such regulatory sequences is riboswitches that regulate the dimerization of HIV-1 RNA, which is an essential step during packaging. Riboswitches are complex folded RNA domains that serve as receptors for specific metabolites. These domains are found in the non-coding portions of various mRNAs, where they control gene expression by harnessing allosteric structural changes that are brought about by metabolite binding. New findings indicate that riboswitches are robust genetic elements that are involved in regulating fundamental metabolic processes in many organisms. The riboswitches are made up of the three-dimensional structure of RNA, in which RNA can undergo two mutually exclusive conformations in response to an environmental signal in the form of a metabolite. Riboswitches comprise two domains: an aptamer and an expression platform. The aptamer is highly conserved even in distantly related organisms, and serves as a precise sensor for its target metabolite. The expression platform is far more variable in sequence and in structure as it can function by assuming one of many structural forms to control gene expression [4]. For all these reasons riboswitches seem to be a significant form of genetic control. The advantage of this system is that they are highly specific to their substrates. But the lack of universality is because riboswitch-mediated translational regulation is limited to monocistronic m-RNA. Riboswitches have been shown to function in repressing gene expression (negative regulation) in response to metabolites. If riboswitch is a primitive mode of gene regulation then it suggests that in the RNA world negative regulation was predominant. There are also speculations that riboswitch-mediated control mechanism might