Name of PhD Student: POONAM GUPTA

Name of PhD Supervisor: Dr. MOHAMMAD HUSAIN

Name of PhD Co-Supervisor: Dr. ARVIND RAI

Department of Biotechnology, Jamia Millia Islamia

Title of PhD Thesis: Genotypic analysis of Nef gene from HIV-1 infected Rapid Progressors and Long-term Nonprogressors.

ABSTRACT

The HIV-1 *nef* protein is a 25- to 27-KDa regulatory protein known to perform multiple functions including CD4 and MHC-I down regulation, infectivity, actin remodeling, and viral spread leading to clinical progression to AIDS. These functions are accomplished through amino-acid motifs present at specific sites. Interactions between these motifs and associated host molecules have been suggested to be responsible for difference in disease progression resulting in rapid progression or delayed progression. The study was designed to perform quantitative analysis of HIV-1 proviral DNA/RNA using real-time PCR and sequence analysis to determine amino-acid substitutions along the whole length of *nef* including various functional motifs that are considered to be responsible for discernible difference in disease progression in patients presenting with rapid and delayed progression to AIDS. The HIV-1 seropositive patients were categorized as Rapid Progressors (RPs) and Long-term Nonprogressors (LTNPs) based on CD4 cell count and clinical data.

The HIV-1 proviral DNA load of LTNPs was higher than that of RPs but the difference between the loads was not statistically significant. A statistically significant negative correlation was found between HIV-1 proviral DNA load and CD4 count of RPs whereas in LTNPs no such correlation was observed. It was found that the majority of the strains characterized belonged to subtype C (40), followed by several BC (7) recombinants and subtype A1 (1). Complete *nef* subtype C sequences from 33 RPs and 7 LTNPs were compared and it was observed that in the majority of the sequences from both the groups, highly conserved functional motifs showed subtle changes. However, drastic changes were observed in two isolates of LTNPs where the arginine cluster was deleted while in one of them additionally, the acidic residues were substituted with basic residues (EEEEE→RK(R)KKE). The deletion of the arginine cluster and the mutation of acidic residues to basic residues are predicted to delay disease development by abolishing CD4 downmodulation and causing diminution of MHC-I downregulation, respectively.

The synonymous-to-nonsynonymous ratio was greater than one in both the groups, suggesting amino-acid conservation and functional robustness. The study also revealed a preference for certain amino-acids at specific sites such as M_{25}, A_{32}, Y_{53}, D_{66}, R_{76}, G_{103}, C_{190}, H_{199}, and F_{229} that occurred in both the groups. Additionally, frequency of certain residues such as F_{101} and H_{122} were significantly different in RPs and LTNPs ($p \leq 0.0326$ and 0.0016, respectively). The residues can be used as markers of different disease progression outcomes. Also, the entropy analysis showed that statistically significant variability of a residue of

acidic cluster was greater in RPs than in LTNPs whereas for $(Pxx)_4$ motif, the variability in LTNPs was greater than in RPs. Inter-patient nucleotide distance within the group and between the two groups showed very little variation, confirming genetic relatedness among isolates. Interestingly, phylogenetic analysis unexpectedly revealed no apparent separate clustering of LTNPs instead it was found to be interdigitated with RPs suggesting evolutionary relatedness with RPs. Deduced amino-acid sequences used to predict HLA binding epitopes for consensus *nef* gene sequences of RPs and LTNPs revealed three HLA subtype binding domains, GAFDLSFFL/GAVDLSFFL and LTFGWCFKL, in high frequency. The finding may encourage use of *nef* gene in designing a multiepitope HIV vaccine suitable for the Indian population.