4.4 Post-doctoral Fellowship and Research Associate-ship studies:

4.4.1 Determination of Sero Type Specific Gene Sequences of Dengue Viruses and their Cross Infections from Human Serum and Mosquito Samples for Risk of Dengue Hemorrhagic Fever in Western Rajasthan (PDF Fellowship Project)

PDF Fellow: *Dr. Annette Angel*
Guide: *Dr. Vinod Joshi, Scientist-G, DMRC, Jodhpur*

Commencement: March, 2014 **Duration:** One & half years **Status:** Concluded
Funding Agency: Indian Council of Medical Research, New Delhi

OBJECTIVES

1. Determination of type specific whole genome sequences of dengue viruses (DEN-1, 2, 3 & 4) isolated from field caught mosquitoes and human serum samples.
2. Study of heterogeneity of cross infections of circulating extrinsic and intrinsic virus strains as possible cause of Dengue Hemorrhagic Fever in disease endemic settings.
3. Application of viral genomics for demarcation of risk areas for DHF.

PROGRESS

RT-PCR was performed using the 17 primer sets of DENV-3. Since the sample (87sBalotra) was also passaged in C6/36 cell lines as shown in the table, hence viral RNA was extracted from both the passaged and unpassaged (pure serum) samples. After amplification, 5µl of amplicons was run in 2% Agarose gel and bands were obtained for each of the 17 sets in both the sample types.

Reverse Transcriptase -Polymerase Chain Reaction (RT-PCR) and Gene Sequencing to test the whole genome primers

Dengue virus primers for types 1, 2, 3 and 4 were designed keeping in view the spatio-temporal diversity within the circulating dengue genomes, specificity, sensitivity, amplicon size, properties of primer annealing, primer-dimer possibility, G-C lock at 3' end and the Melting temperature (Tm) of primers. For checking the working of the whole genome sequencing primers, the positive control courteously provided by CDC, USA in the RT-PCR kit i.e. was used. RNA was extracted from the lyophilized PC and RT-PCR was performed using the newly designed primer sets. After amplification, 5µl of amplicons was run in 2% Agarose gel and bands were obtained for each of the four Dengue virus types. The expected MW for each amplified product was approximately 800-1000 bp while those amplified with primer sets covering the start and stop points had MW of approximately 450-550 bp.The amplicons obtained for each of the four dengue virus types were then were subjected to sequencing in the 3130xl Genetic Analyzer (m/s ABI, USA).

Gene Sequence Analysis of the Positive Dengue sample

The fragmented sequences obtained for the four dengue serotypes were then assembled together to get the whole genome with respect to each of the four serotypes. Using the BLAST software, the four assembled sequences were then searched for its possible alignment and matching.

Reverse Transcriptase -Polymerase Chain Reaction (RT-PCR) and Gene Sequencing of the dengue samples using the whole genome primers

Once the primers sets designed were checked using the positive sample, few serum and mosquitoes samples were taken for analysis from the virus repository in the Lab keeping in view the clinical profile of the sample. A total of 3 serum and 3 mosquitoes were attempted for whole genome sequencing as shown in Table 1.

Table 1. Details of the serum and mosquito from infected household processed for sequence analysis

S. No.	Patient's code	Dengue virus related clinical profile	Serum Passaged in C6/36 or not	Dengue virus type identified using RT-PCR	Mosquito from patient's house (code)	Detection of dengue virus in mosquito	Dengue virus type identified using RT-PCR
1.	87 S Balotra	NS1 +ve; IgM +ve	Yes	DENV-3, 4	Female (7=3f)	IFAT +ve	DENV-3,4
2.	87 S Balotra	NS1 +ve; IgM +ve	No	DENV-3, 4	-	-	-
3.	101 Jodhpur	NS1 +ve; IgM –ve	No	DENV-2	Female (101=5f)	IFAT +ve	DENV-3
4.	102 Jodhpur	NS1 +ve; IgM +ve	No	DENV-2	-	-	-

Each of these 17 products were purified using the PCR purification kit (M/s Qiagen, CA) and then were subjected to sequencing using the Big Dye Terminator v3.1 ready reaction kit (m/s ABI, USA). Subsequently, any unincorporated dNTP's were removed using the Dye Ex 2.0 Spin kit (m/s Qiagen, CA). The purified products were then sequenced using the 16 capillary Genetic Analyzer (M/s ABI, USA). The rest of the samples as shown in Table 1 are in the process of sequencing.

Genome sequence and Phylogenetic analysis

For further analysis of the whole DENV-3 genome obtained for the sample code 87S Balotra, the SeqScape v2.6 (M/s ABI, USA) software was used. Reference based assembly was performed using NCBI reference genome for DENV-3 (NC_001475.2). Comparative genome analysis was also performed with nearest genome in phylogenetic analysis. To perform the phylogenetic analysis, assembled genome sequence was subjected to BLAST (Megablast algorithm) on NCBI Non Redundant database. All hits obtained were of DENV-3 genome only. Fasta format of the genome sequences of the top 100 hits were downloaded. These sequences alongwith the obtained assembled genome sequence were then subjected to local pair-wise Multiple Sequence Alignment (MSA) using MAFFT v7.221. Phylogenetic analysis and distance calculation were performed using the MEGA v.6 software with the Neighbour-Joining method of the Maximum Composite Likelihood model, uniform rates among sites with 1000 bootstrap replicates.

Results & Observations

Partial Sequencing of mosquito and serum samples with earlier referred primers:

During the process of type identification of the DENV present in the mosquito as well as serum samples, few of the samples were carried forward for gene sequencing (using primers referred earlier (Lanciott et al, 1996 & Seah et al, 1995)). The partial/ short length sequences obtained were then submitted to the NCBI databank. The details of the sequence submitted are appended in the publication list.

Reverse Transcriptase -Polymerase Chain Reaction (RT-PCR) and Gene Sequencing to test the whole genome primers

Agarose gel electrophoresis was carried out using 2% agarose gel and 5µl of the cDNA amplified was run for 30 minutes at 100v. This was done for all the four DENV.

Complete genome characterization and phylogenetic analysis of the serum sample 87s Balotra

Agarose gel electrophoresis was also carried out for the serum (code 87S Balotra, positive for DENV-3). 5µl of the cDNA amplified with 17 sets of primers. Since the sample was also passaged in C6/36 cell line, hence both the pure/direct serum as well as the passaged serum was taken for sequencing.

Genomic and Phylogenetic analysis of the DENV-3 sample

Complete genome of Dengue-3 virus of the 87S Balotra sample was sequenced. In total 10, 674 base pairs (bp) in un-passaged sample and 10, 672 bp in passaged sample were analysed for their nucleotide sequence. The polyprotein gene range in case of unpassaged genome was from 88 to 10260 while that of passaged genome was from 86 to 10258. The sequence has been submitted to NCBI, USA. When pair-wise nucleotide alignment using BLAST showed that in first 5 best matches, observed genome resembled the complete genome sequences of India, Pakistan, China and Sri Lanka.

The observed sequence which is first complete genome of dengue- 3 virus from Rajasthan, India showed 98-99% identity with reported Indian genomes. The phylogenetic analysis of the genome is shown in Figure 1.

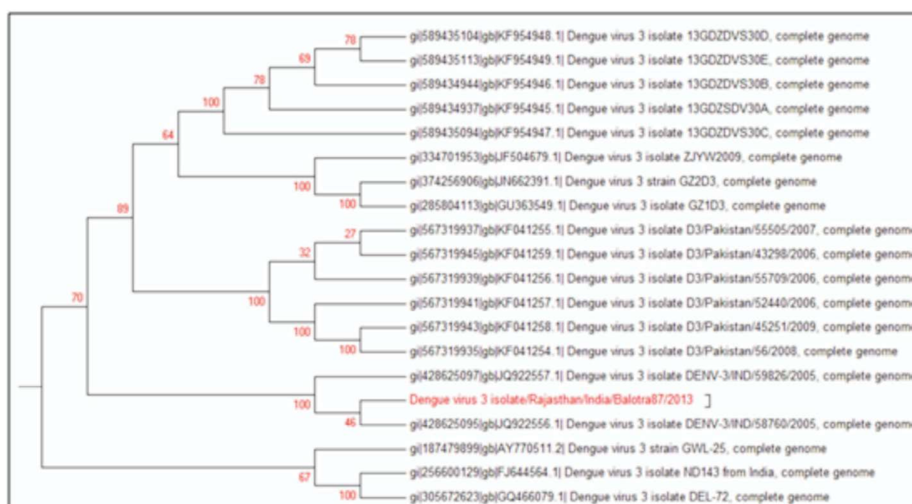

Fig. 1. Phylogenetic analysis of whole genome. (Fig shows a sub tree depicting location of Rajasthan genome in proximity to genomes from other parts of India followed by Pakistan and China).

Variations in nucleotide sequence of observed genome with reference to global and regional genomes

The whole genomes of the DENV-3 virus were subjected to mutation analysis. These genomes were compared taking two reference whole genomes sequences into consideration: one with NCBI reference genome (Global genome, Accession no.: NC_ 001475.2) and another with genome of nearest phylogeny (Regional genome, Accession no.: JQ922556.1). As many as 388 variations were observed in the genome sequence when compared with the NCBI reference genome (global genome) and 109 variations were observed when compared with the regional genome (genome of nearest phylogeny) (Table 2).

Table 2. Comparative analysis of the observed genomes with global and regional genomes

Strain	Variations observed at nucleotide level		Variation observed at amino acid level	
	With Global genome NC_001475.2	With Regional genome JQ922556.1	With Global genome NC_001475.2	With Regional genome JQ922556.1
DENV-3 isolate/ Rajasthan, India (Accession no. : KU216208)	388	109	34	25
DENV-3 Rajasthan, India (Accession no. : KU216209)	388	109	34	25

Nucleotide sequence variations leading to Amino acid alterations

Of the 388 variations observed in the nucleotide sequence of reported genome with reference to global genome, 34 variations led to Amino acid (AA) variations. Four variations were seen in envelope protein, 1 in Anchored capsid region where as remaining 29 showed AA variations in non structural (NS) proteins of dengue-3 virus. When the observed genome of Rajasthan, India (Accession no.: KU216208) was compared with the genome of nearest phylogeny, it showed 5 AA variation in the genes comprising the structural protein region (1 in Anchored capsid C and 4 in Envelope protein region) while remaining 20 variations were observed in the viral gene region coding non structural proteins. The NS4A region did not show any AA variation. Details of the nucleotide and amino acid variation were published in virology reports journal (http://dx.doi.org/10.1016/j.virep.2016.05.003).

Highlights of the Study

1. First study on the complete genome of Dengue-3 virus has been made in Rajasthan, India
2. The complete genome of 10672 base pairs was studied. The analysis showed 388 variations in the nucleotide sequences including 34 variations in the Amino Acids.
3. Phylogenetic studies showed similarity of genome with reported genomes from India, Pakistan and China