# "Laboratory–based collaboration net work of infectious diseases in Asia"

**Funding Organization: National Institute of Infectious Diseases*, Tokyo, Japan***

**Indian Investigators :**

**Principal Investigators:** ***Dr. S. Chakrabarti, Director in Charge, National Institute of Cholera and Enteric Diseases, Kolkata***

**Co-Investigators:**

***Dr. Sekhar Chakrabarti, Scientist G***
***Dr. Banwarilal Sarkar, Scientist E***
***Dr. Asish K. Mukhopadhyay, Scientist D***
***Dr. Sandipan Ganguly, Scientist C***
***Dr. Hemanta Koley, Scientist C***

**Foreign Investigators :**

**Principal Investigators*:*** ***Dr. Haruo Watanabe, Director General, National Institute of Infectious Diseases, Japan & Ichiro Kurane, Deputy Director-General, National Institute of Infectious Diseases, Japan.***

**1. Title of the Research Project :**

**"**Laboratory –based collaboration net work of infectious diseases in Asia**"**

**2. Objectives**

The proposal has been developed with a broader objective to understand the evolution of pathogenesis of enteric pathogens and HIV that include phenotypic and genotypic characteristics of identified pathogens from diarrheal specimens as well as identification of novel factors for pathogenesis, to develop potential vaccine and to study mutations in HIV during infection which has an association with different genotypes and might be used as biomarkers; the study objectives will be covered though extensive work under the following areas and each of these areas will be covered by the Co-PIs of the project

*i) Retrospective analysis on the evolutionary aspects of Vibrio cholerae*

*ii) Differential pathogenesis of Giardia : Role of Giardia Virus*

*iii) Development of universal Shigella vaccine based on virulence gene expression.*

*iv) Analysis of HLA associated HIV-1 mutations in India and Japan.*

**3. Summary of the research for fiscal year 2012-13**

***1. Retrospective analysis on the evolutionary aspects of Vibrio cholera***

# Second Year Report

## *1. Study Title: Retrospective analysis on the evolutionary aspects of Vibrio cholerae*

## 2. Study facility:

National Institute of Cholera and Enteric Diseases
Division of Bacteriology
Kolkata, India

National Institute of Infectious Diseases
Tokyo, Japan

**3. PI from Indian Side for this report:** Asish K. Mukhopadhyay

**Summary:**

Our previous study on *Vibrio cholerae* strains isolated from Kolkata over 17 years from 1989 to 2005 depicted that in Kolkata, *V. cholerae* O1 strains with classical allele of *ctxB* have totally replaced seventh pandemic El Tor strains possessing El Tor allele of *ctxB* since 1995.

The recent devastating cholera outbreak in Haiti, for the first time in almost a century, placed this ancient scourge at the forefront of the global public health agenda. The causative agent of this outbreak *V. cholerae* contained a unique mutation at the 58$^{th}$ nucleotide of *ctxB* gene that has motivated us to investigate the emergence and dissemination of these new variants in Kolkata, India. Our newly developed double mismatch mutation assay and sequencing analysis showed that the Haitian *ctxB* first appeared in Kolkata during April, 2006 and 93.3% strains isolated in Kolkata during 2011 carried the new allele. This finding indicates that Haitian *ctxB* allele may have originated in Kolkata and then spread to the neighboring regions. Exploitation of this newly developed PCR assay will play important role in understanding the intercontinental spread of cholera. It has been hypothesized that the unique genetic composition of the new variants increases their relative fitness, perhaps as a consequence of increased pathogenicity.

Chronological analysis of one hundred twenty five *Vibrio cholerae* O139 strains isolated during 1993 - 2005 from Kolkata revealed the prevalence of two new genotypes of cholera toxin (CT) along with the different combinations of *ctxB* and *rstR* alleles resulting in variant CTX prophages. The prototype El Tor CTX phages gradually disappeared in O139 strains and the predominant CTX phages in O139 since 2002 are CTX Calcutta phages with genotype 4 and CTX El Tor phages with genotype 5. During the course of appearance and disappearance over a decade, five types of O139 strains have been detected based on CT genotypes. Very recently, we have found that few O139 strains contain CTX prophage in the small chromosome. These kinds of changes in O139 strains might be the probable reasons for their declining prevalence in human cases in Kolkata, India.

**Purpose:**

Genetic analysis of *Vibrio cholerae* O1 and O139 strains from Kolkata

**Materials and Methods:**

**Bacteriology and serology:** One hundred and forty two *V. cholerae* O1 strains were selected for this study. These strains were selected from the strain repository of the National Institute of Cholera and Enteric Diseases, Kolkata and were isolated from 2004 to 2011. All the strains were grown in Luria Bertani broth (Becton Dickinson, Sparks, Maryland, USA) for 18 hrs and then streaked on Luria agar (LA) plates. Identity of these strains was reconfirmed serologically by the slide agglutination with O1 specific polyclonal antiserum and serotype specific antisera (Becton Dickinson). *V. cholerae* O1 strains O395 (serotype Ogawa), N16961 (serotype Inaba) and EL-1786 (Ogawa, El tor) were used as standard strains for classical. El Tor and Haitian type, respectively.

Separately, one hundred and twenty five O139 strains were selected for this study. These strains were selected from the strain repository of National Institute of Cholera and Enteric Diseases, Kolkata and these were isolated in different time frames from 1993 to 2005.

**Preparation of template for PCR:** One loopful of an overnight culture from LA plate was suspended in 200 µL of Tris-EDTA buffer (pH 8.0) and then lysed by vigorous vortexing with mixture of phenol: chloroform: isoamyl alcohol (25:24:1) saturated with 10 mM Tris and 1mM EDTA.(Sigma-Aldrich, St Louis, USA) Supernatant was collected carefully following centrifugation at 12,000 rpm for 15 min and was extracted once with 100 µL of mixture of chloroform: isoamyl alcohol (24:1) and centrifuged for 15 min at 12,000 rpm. The supernatant containing the DNA was used as template for PCR analysis.

**PCR analysis:** Double mismatch amplification mutation assay (DMAMA) PCR have been developed by exploiting *ctxB* sequence polymorphisms between CT genotype 1 (classical type) and genotype 7 (Haitian type). Briefly, PCR mixture (20µl) was set up in two sets, each containing 10 ng of template DNA, 200 nM of allele specific forward primers (i.e. either *ctxB*-F3 or *ctxB*-F4), and a reverse primer in common, Rv-cla, 200 µM of each dNTP (Roche Diagnostics, Mannheim, Germany) with 1 unit of Taq DNA Polymerase (Bangalore Genei, Bangalore, India) and 1.5mM $MgCl_2$ (Bangalore Genei, Bangalore, India). The optimal annealing temperature was determined using Veriti Thermal Cycler (Applied Biosystems, Foster city, USA) which generates a thermal gradient of annealing temperatures (from 54°C to 64°C). Initial standardization of the PCR was done with the control strains and best result was obtained using the protocol depicted in **figure 1.** *V. cholera* O1 strains O395 and N16961 were used as standard strains for classical and El Tor biotypes, respectively. Haitian strain, EL-1786 was used as the reference strain for CT genotype 7. The amplified fragments underwent agarose gel electrophoresis, stained with ethidium bromide and digitally recorded.

**Nucleotide Sequenceing of *ctxB*:** To determine the nucleotide sequence of the *ctxB*, PCR amplification of *ctxB* locus of 12 *V. cholerae* O1 isolates was performed in a 25µL reaction mixture.

PCR primers and conditions used have been previously described. PCR amplicons were purified using the Qiaquick PCR purification kit (QIAGEN, Hilden, Germany) and both the strands were sequenced in an automated sequencer (ABI PRISM 3100 Genetic Analyser, Applied Biosystems, Foster city, USA). The entire coding sequences of the *ctxB* gene of three representative strains have been deposited in GenBank with accession numbers JN806157-JN806159. The deduced amino acid sequences of CTB from these three strains were aligned with corresponding sequences from N16961 (GenBank accession number NC-002505) and O395 (GenBank accession number CP001235) by using the online software Clustal W.

**Pulsed field gel electrophoresis (PFGE):**

Analysis of *Not*I digested PFGE patterns obtained with sixteen representative strains was carried out. The enzyme restricted the chromosomal DNA in to 15-27 fragments of different sizes ranging from 6 to 398 kb. Agarose-embedded genomic DNA from *V. cholerae* strains was treated with 50 U *Not*I (Takara Bio Inc, Otsu, Japan)) and the resulted DNA fragments were separated by electrophoresis under the contour-clamped homogeneous electric field (CHEF) on a CHEF Mapper system (Bio- Rad Laboratories, Hercules, CA, USA). Run conditions were generated as described earlier. The resulting fragments were typically followed an electrophoresis time of 40 hrs 24 min Following electrophoresis, the gels were stained for 30 min in Elix water (Millipore, Billerica, MA, USA) containing 1mg/mL ethidium bromide, destained in Elix water for 15 min and digitally recorded under UV light using the Gel Documentation System (Bio- Rad Laboratories). Dendogram analysis on the PFGE patterns was performed using Bionumeric software (Applied Maths, Sint-Martens-Latem, Belgium).

**Results and Discussion:**

**Development of a double mismatch mutation assay (DMAMA) PCR:** In this study, we focused on the PCR detection of *ctxB* alleles in *V. cholerae* strains. Initial purpose of our study was to establish a PCR based assay which can discriminate the *V. cholerae* strains carrying Haitian, classical and El Tor alleles of *ctxB* in a simple and rapid way. Current methods for differentiating the biotype specific CTB subunit of *V. cholerae* O1 requires MAMA PCR of the *ctxB* gene with biotype specific primers, nucleotide sequencing of the *ctxB* allele or performing an ELISA using a specific monoclonal antibody to the classical or El Tor CT. Among these the first one has been the method of choice as it is simple and less time consuming. But the reports of influx of new variant strains of *V. cholerae* O1 with a classical CTB additionally mutated at the 20$^{th}$ amino acid (58$^{th}$ nucleotide position) clearly point out its limitation as this allelic discrimination is based on two biotype specific reverse primers each bearing a mismatch at nucleotide position 203 and so it was unable to discriminate between the classical and Haitian type *ctxB* allele. Therefore, for discriminating the classical, El Tor and Haitian type *ctxB* alleles in a simple, rapid and accurate way, DMAMA was designed and validated (**Figure 1**). We designed two allele-specific polymorphism detection forward primers, *ctxB*-F3 and *ctxB*-F4 each bearing a mismatch at their 3' ends. These allele-specific primers each carry specific nucleotide, A and C, for Haitian and classical type, respectively, at the 3' end. Furthermore, we enhanced the 3' mismatch effect by introducing another nucleotide alternation G (rather than A) at the second nucleotide (i.e, the 57$^{th}$ nucleotide) from the 3'end of both the primers. We used the *ctxB* reverse primer specific for the classical biotype (Rv-cla) as described by Morita *et al* as the conserved reverse primer. As shown in **Figure 2**, the DMAMA successfully discriminated the three different allelic subtypes of *ctxB. V. cholerae* O1 strains having the *ctxB* allele of genotype 7 yielded a 191 bp fragment of DNA with the primer pair *ctxB*-F3/Rv-cla but not with *ctxB*-F4/Rv-cla. An exact reversal of the first result was found in case of the classical control strain (O395), and the El tor strain (N16961) did not produce any amplicon in both set PCR assays due to double mismatch in the forward and reverse primers. We have standardized the PCR to optimize both the specificity and sensitivity using the protocol depicted in **figure 1**.

**Sequencing analysis to evaluate the PCR based result:** To further reconfirm our PCR based result, 10 representative strains, which yielded positive bands for Haitian *ctxB* gene by DMAMA-PCR, were selected for DNA sequencing of *ctxB* gene. For sequencing, a separate pair of primer (*ctxB*-F: and *ctxB*-R) was used to provide the sequences of whole *ctxB* gene. Nucleotide sequence analysis of the *ctxB* genes of 10 representative strains of *V. cholerae* O1 revealed that the strains possessed DNA sequences identical to that of the classical type of *ctxB* with an additional mutation at the 58$^{th}$ position (C to A). The deduced amino acid sequences of all 10 representative strains were aligned with the CTB sequences of the reference strains N16961 (El Tor) and O395 (classical). The deduced amino acid sequences of all strains were found to be identical to the deduced amino acid sequence of the CTB of the O395 classical reference strain, with a histidine at position 39 and a threonine at position 68, except at 20$^{th}$ position with an asparagine instead of histidine (**Figure 3**). Thus, the result from DNA sequencing of *ctxB* gene confirmed the results of DMAMA-PCR and GenBank submission

ID is 1484711. Whole CTB is made up of 124 amino acids. A proteolytic cleavage between amino acids at position 21 and 22 results in generation of a 21 amino acids signal peptide and 103 amino acids mature CTB. New variant *V. cholerae* strains (with Haitian type of CTB) were mutated in signal peptide and histidine at position 20 was replaced by asparagine. We also sequenced *ctxB* gene from three representative strains those yielded amplicons with the classical specific primers (*ctxB*-F4/Rv-cla). The deduced amino acid sequences of all 3 strains were found to be identical to the classical reference strain, with a histidine at position 39 and a threonine at position 68 (data not shown). Thus, the result from DNA sequencing of *ctxB* gene confirmed the result of DMAMA-PCR.

**Screening of the Kolkata strains using the DMAMA PCR:** After standardizing the DMAMA, we extensively used it to investigate the emergence and dissemination of these new Haitian variant of *V. cholerae* strains in Kolkata, India. All the strains tested in 2004-05 were positive for classical type of *ctxB* indicating they are El Tor variant strains. Our analysis showed that first appearance of Haitian type *ctxB* was noted in Kolkata during April, 2006. There was a sudden decrease in the isolation profile of *V. cholerae* O1 strains with Haitian *ctxB* allele (CTB genotype 7) during 2007and 2008. The percentage of the O1 isolates with CT B genotype 7 started to increase from 2009 (**Figure 4**) and more than 93% Kolkata strains carried Haitian *ctxB* allele in 2011.

**Phylogenetic analysis based on PFGE:** Results of DMAMA and the sequencing data clearly indicated arrival of novel variant among the circulating strains of *V. cholerae* O1 around Kolkata and motivated us to take a closer look on the relatedness of these variants with the Haitian isolates. The PFGE profiles of *V. cholerae* strains from Kolkata were compared using Bio Numeric software (Applied Maths, Sint-Martens-Latem, Belgium) (**Figure 5**). The similarity between strains was determined using the Dice coefficient, and cluster analysis was carried out using the unweighted-pair group method using average linkages (UPGMA). All the tested *V. cholerae* strains with classical *ctxB* (genotype 1) clustered together (cluster A), with a similarity matrix of > 95 (**Figure 5**). All the tested strains with Haitian *ctxB* (genotype 7) were also found to be closely related to each other, with a similarity matrix of >92 (cluster B). Dendogram analysis showed two distinct PFGE pattern suggesting considerable diversities in genomic content between two groups (genotype 1 and genotype 7). Phylogenetic patterns indicate a close relationship between Haitian and Kolkata *V. cholerae* strains containing Haitian *ctxB* (genotype 7).

**Conclusions:** Our results not only signify a cryptic change in the circulating strains in Kolkata but also raise questions about the origin of these variants of *V. cholerae* O1 El Tor. This new type of *ctxB* (genotype 7) was first reported in *V. cholerae* O1 strains by Goel *et al* isolated from a cholera outbreak in Kalahandi, Orissa in 2007. But our results clearly show that in Kolkata genotype 7 prevailed since April 2006. This finding tempted us to speculate that Haitian type of *ctxB* may have originated from Kolkata and then disseminated to the neighboring regions like Orissa and other places, although conformation of this hypothesis requires several other epidemiological and experimental validations, and then may have spread incontinently from Nepal to Haiti as shown by recent evidences. It has been hypothesized that the unique genetic composition of the new variants increases their relative fitness, perhaps as a consequence of increased pathogenicity.

Recent reports by several research groups showed a putative link between the strains associated with cholera in Haiti and in Nepal underscoring the speed at which infectious diseases can be transported globally and this situation puts at risk other non-endemic countries also. Implementing a coordinated, integrated multisectoral approach is the only efficient way to prevent and contain outbreaks among vulnerable populations living in high-risk areas. Prevention, preparedness, and response all depend upon an effective and holistic surveillance system and are linked and interdependent. We strongly believe that the DMAMA will be an easy and accurate tool for tracking the emergence and dissemination of Haitian variant *ctxB* in *V. cholerae* O1 isolates and therefore will impart an integral role in understanding the cholera epidemiology around the globe.

**Genetic analysis of CTX prophages of *Vibrio cholerae* O139 strains from Kolkata:**

**Analysis of *ctxB* by MAMA PCR:** The results of MAMA-PCR showed that *V. cholerae* O139 strains isolated up to 1995 yielded amplicon only with El Tor allelic primer pair of *ctxB*. In case of strains isolated in 1996, along with El Tor *ctxB* among 54% of the strains, 18% of the strains yielded amplicon for classical *ctxB* primer pairs and 28% of the strains yielded amplicon with both classical and El Tor primer pairs of *ctxB*. This same trend continued in *V. cholerae* O139 strains isolated in 1997. In case of strains isolated in 1998, the result from MAMA PCR depicted that no strains yielded amplicon only with El Tor *ctxB* primer pair while 68% yielded amplicon with classical allele of *ctxB* primer pair and 32% yielded amplicons with both classical and El Tor allele of *ctxB* primer pairs. Since 1999 to 2005, the O139 strains yielded amplicon only with classical allelic primer pairs of *ctxB*.

**Nucleotide sequence analysis and nested PCR:** To confirm the results of MAMA-PCR, representative *V. cholerae* O139 strains isolated from 1993 to 2005 were selected for sequencing the *ctxB* loci. The results were depicted in **Figure 6.** Nested PCR results also depicted the schematic representation (**Figure 7**) of variable combination of CT genotypes and *rstR* alleles prevailed among O139 strains in Kolkata.

The prominent events in the changing profile of CTX prophases with respect to CT genotypes and *rstR* alleles among O139 strains over a decade were indicated in **Figure 8**, which demonstrated the graphical representation of isolation status of *V. cholerae* O139 strains from patients hospitalized due to acute secretary diarrhea at the Infectious Diseases Hospital, Kolkata from January 1993 to December 2005.

**Chromosomal localization of CTX prophage:** The schematic diagram displaying the chromosomal localization of CTX prophage among reemerged *V. cholerae* O139 strains during 1996 and 2003 are presented in **Figure 9**.

**Conclusion:** The diverse genotypes of *ctxB* as well as *rstR* alleles occurring among *V. cholerae* O139 strains along with the variations in other genetic segments of O139 strains which are not yet ascertained, perhaps consequences the temporal variation of incidence and prevalence of O139. The structural and functional aspects of these new CT genotypes will be the interesting areas to be explored in future, which may reveal vital information about the evolutionary phases of *V. cholerae* O139 strains. Frequent mutations thus acquired by *V. cholerae* O139 strains since its genesis may have an impact in their declining prevalence in cholera endemic areas like Kolkata.

**Studies in progress:**

**We are presently perusing for expression study along with some animal model assay to make this work into a publication format.**
**Studying of O139 strains having CTX prophage in the small chromosome**

## Standardized PCR protocol used for this study

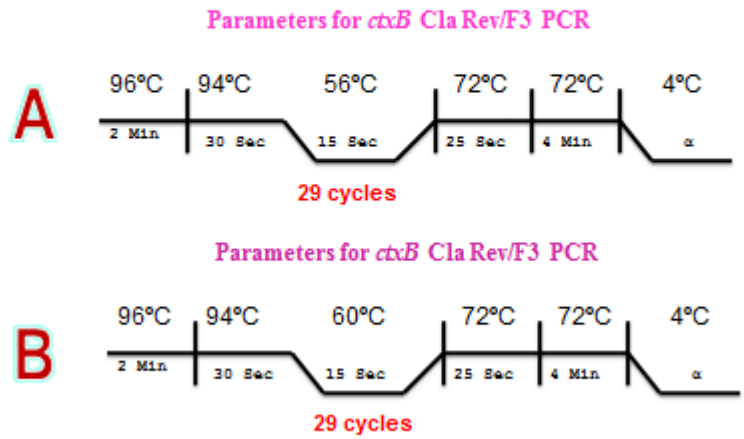
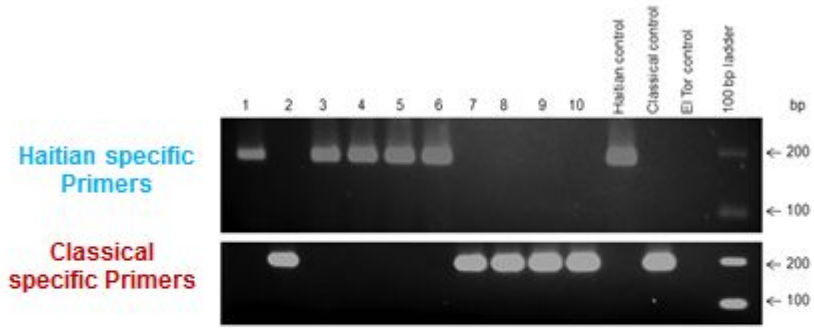


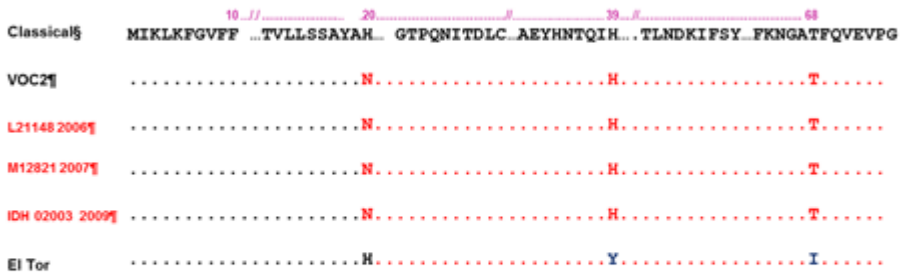

Figure 1




Double Mismatch amplification Mutation PCR assay to discriminate the classical, El Tor and Haitian type *ctxB* alleles

Figure 2

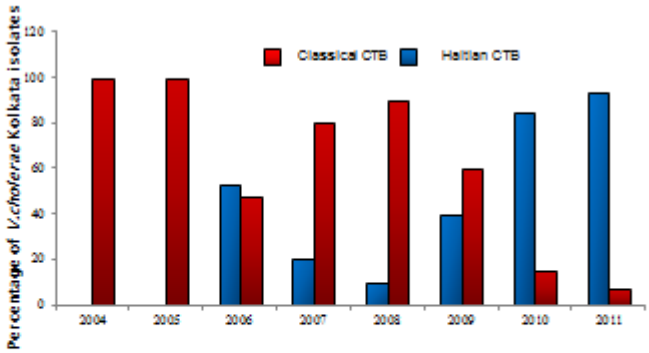
# Sequence analysis of *ctxB*




The deduced amino acid sequences of *ctxB* of representative IDH isolates were found to be identical to CT of the Haitian strain with a asparagine at position 20, histidine at position 39 and a threonine at position 68 confirming its identity with the Haitian type CT.

Figure 3

Isolation profile of *Vibrio cholerae* O1 strains with classical and Haitian type of *ctxB* in Kolkata




*Naha et al*; JCM, May 2012

**_V. cholerae_ O1 strain with Haitian type of _ctxB_ was first time isolated in Kolkata during April 2006**

Figure 4

**Dendogram Analysis of the PFGE pattern using Bionumeric software (Applied Maths, Belgium)**

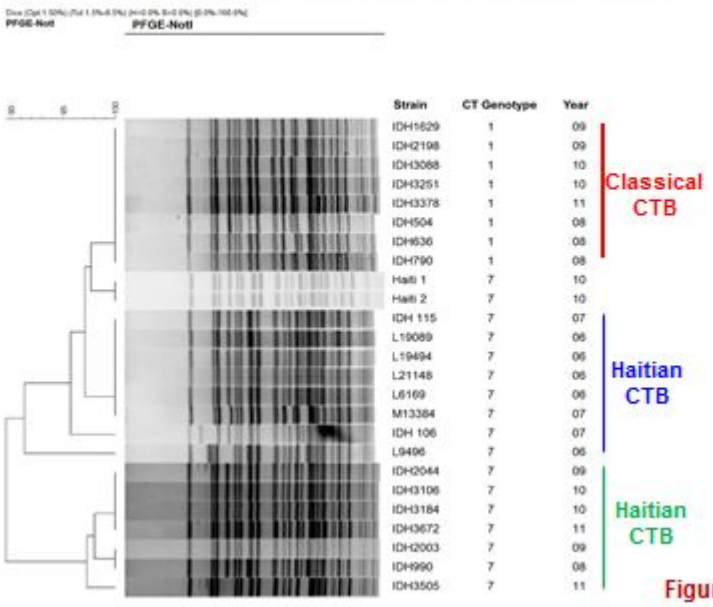



Figure 5

**Analysis of *ctxB* and *rstR of* O139 strains isolated from Kolkata over a decade**

| Year | Nucleotide positions | | | Amino acid positions | | | CT genotype | *rstR* type |
|---|---|---|---|---|---|---|---|---|
| | 83 | 115 | 203 | 28 | 39 | 68 | | |
| 1993-1995 | A | T | T | Asp | Tyr | Ile | Genotype 3 | El Tor |
| 1996-1998 | A<br>C | T<br>C | T<br>C | Asp<br>Ala | Tyr His | Ile<br>Thr | Genotype 3<br>Genotype 4 | El Tor<br>Calcutta |
| 1996-1997 | C | C | C | Ala | His | Thr | Genotype 4 | El Tor |
| 1998 | C | C | C | Ala | His | Thr | Genotype 4 | Calcutta |
| 1998-2001 | A | T | C | Asp | Tyr | Thr | Genotype 5 | El Tor |
| 2000-2005 | C<br>A | C<br>T | C<br>C | Ala<br>Asp | His Tyr | Thr<br>Thr | Genotype 4<br>Genotype 5 | Calcutta<br>El Tor |

Color coded year indicates that single strain contains two allelic type of *ctxB*

Figure 6

**Different combinations of *rstR* alleles and *ctxB* genotypes prevailing among O139 strains isolated from Kolkata**

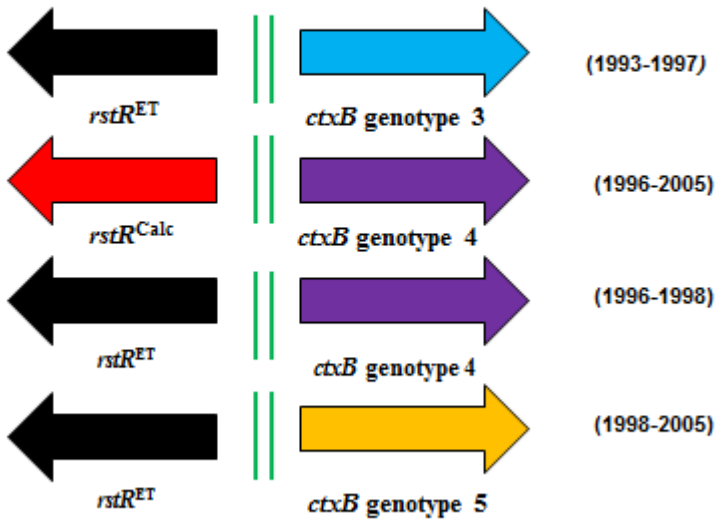



Figure 7

## Changing profile of CTX prophage among O139 strains over a decade in Kolkata

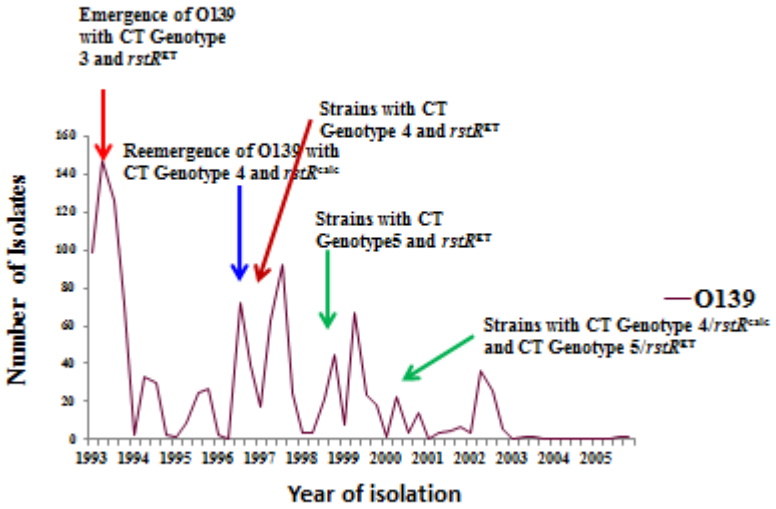



Figure 8

**Schematic diagram showing copy number of CTX prophage and allelic combinations of *rstR* and *ctxB* in reemerged O139 of 1996 (A) and recent O139 (B) of Kolkata**

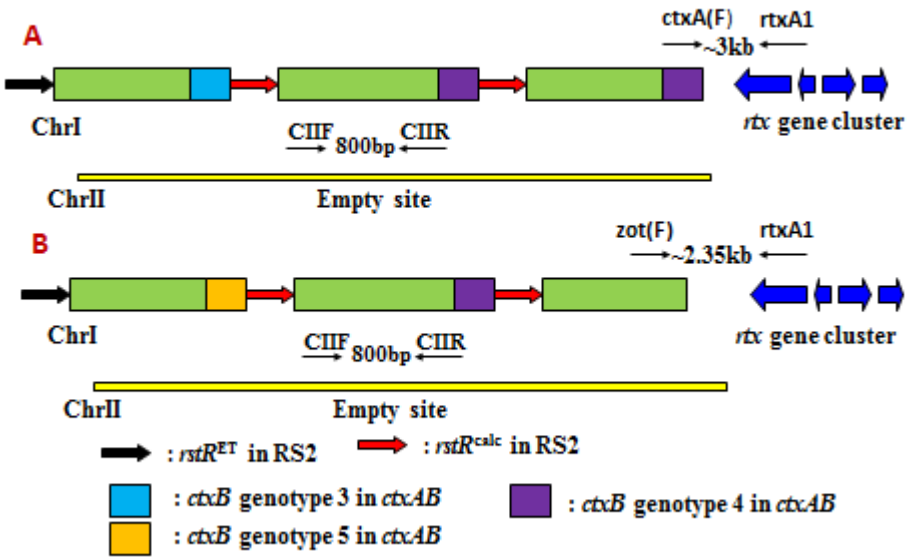



Figure 9

## 2. *Differential pathogenesis of Giardia : Role of Giardia Virus*

**i) Name of PI:** Dr. Sandipan Ganguly

**ii) Name of the Co-**PI: Prof. Tomoyoshi Nozaki

**iii) Title of the Project:** ***Differential pathogenesis of Giardia : Role of Giardia Virus***

**iv) Objectives of the Project:**

- To identify the infection of different GLV among different isolates of *Giardia lamblia* in patients with differential infection, i.e. with and without symptomatic *Giardia lamblia* infection and with multiple infections along with *Giardia lamblia* as a co-infection and also with asymptomatic *Giardia lamblia* infection.
- To understand if there is any genetic variability in housekeeping as well as pathogenic genes (*e.g.* the excretory secretory protein genes of *Giardia lamblia*) before and after infection of GLV or the outcome of differential pathogenesis is only a result of differential transcriptomic expression by PCR, RT PCR and Microarray hybridization.

**v) Plan of work for fiscal year 2012-2013**

- **Accessing the genetic variability among local Giardia strain**
- **Standardization of GLV detection procedure**

**Pilot survey for assessment of genetic variations among local Giardia isolates**

**Significance of study area**

Kolkata (N 22º577242 E 88º398743) is a densely populated city of India, where the giardiasis in human caused by infection of *G. duodenalis* is a serious problem of public health. It is one of the major metro cities of eastern India and Southeast Asia with plenty of immigration and emigration of multi cultured people with varied socio economic condition. It has a favorable climatic condition for the growth of enteric pathogens. So, it is very much likely to find plenty of opportunistic enteric parasites in this city and also a possibility of genetic variations among them.

**Procedure**

68 Giardia positive stool samples were randomly taken from the surveillance program of IDBG hospital and were subjected to multi-locus genotyping. The DNA was extracted directly from the positive stools using StoolDNAMiniKit (QIAGEN, USA) according to the manufacturer's protocol. A portion of *ß*-giardin (*ßg*) [511 bp] on 90 kb long contig ctg02-35, *Glutamate dehydrogenase* (*gdh*) [434 bp] on 231 kb long contig ctg02-15 and *Triose phosphate isomerase* (*tpi*) [530 bp] on 200 kb long contig ctg02-19 (www.*Giardia*db.org), were individually amplified according to the previously described nested PCR protocols [10-12]. The nested PCR products were separated in 1.5% (w/v) agarose gel and purified by gel cut purification process using High Pure PCR purification Kit (Roche, Germany) as per the manufacturer's protocol. Bi-directional sequencing was performed with the respective purified products and nested PCR primers on an ABI 3100 automated sequencer by using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystem, USA). The

sequences obtained from all three loci (*ßg, gdh* and *tpi*) were validated using the database BLAST search (i.e. NCBI and *Giardia*db) and were submitted to NCBI GenBank (ACC no. JF918436 – JF918523 & JN647526 – JN647641). The sequences from each locus were separately aligned by using 'MEGA Version 4' software [13,14] and were manually checked and edited. Previously reported sequences of the respective loci representing different *G. duodenalis* assemblages were included in the analysis to get a better resolution of the assemblage distribution. The extent of sequence diversity among the wild isolates based on the target loci was determined using the 'Maximum Composite Likelihood' method through the MEGA4 software.

**Findings**

Based on the cumulative sequence data of all three loci, 41 samples could be assigned as assemblage 'B' (60.2%) and 13 as assemblage 'A' (19.1%), while 14 (20.5%) isolates showed multiple assemblages depending on the marker loci.

Sequence analysis showed higher degree of diversity in the '*tpi*' loci in compare to other two (**Table 1).**

**Table 1. Estimation of average sequence diversity**

| **Target locus** | | | | | |
|---|---|---|---|---|---|
| **ß giardin (*ßg*)** | | **Glutamate dehydrogenase (*gdh*)** | | **Triose phosphate isomerase (*tpi*)** | |
| ***N*** | ***D* ± *SE*** | ***N*** | ***D* ± *SE*** | ***N*** | ***D* ± *SE*** |
| 68 | 0.001 ±0.004 | 67 | 0.053 ±0.011 | 64 | 0.106 ±0.019 |

*The number of base substitutions per site from averaging over all sequence pairs is shown. All results are based on the pairwise analysis of 'N' number of sequences. Standard error (SE) estimate(s) are shown besides the diversity (D) estimates and were obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Maximum Composite Likelihood method in MEGA4. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). The tpi loci showed highest diversity in compare to the other two.*

Detection of Assemblage 'B' and assemblage 'A' overlaps the results obtained from previous studies in southern India using single locus PCR-RFLP method [18]. However, non-specific sequence heterogeneity among same loci of identical assemblages led to the difficulty in assigning sub-assemblages [19]. In the previous reports from the same country, presence of mixed assemblages may be evident but, those were usually found through multiple assemblage specific bands in RFLP analysis of a particular locus [18]. This may occur due to presence of overlapping DNA sequence derived from two or more different isolates, although it was not proved. The unique finding of our study was, 14 (20.5%) isolates showed multiple assemblages depending on the marker loci. Clustal distribution with these 14 isolates in the NJ tree revealed a better picture where isolates were placed in distinct clusters of assemblage A and B (marked with red and blue bars respectively) supported by high boostrap values (Fig. 1) but, the cluster selection (i.e. A or B) of the isolates were different for each loci. For example the isolate number GLI11 is positioned in assemblage A cluster in '*ßg*' and '*gdh'* tree (Fig. 1A & B) but falls in the other cluster in '*tpi'* tree (Fig. 1C). Again isolate GLI23 is positioned in assemblage A cluster in *ßg* and '*tpi*' tree (Fig. 1A & C) but is present in the assemblage B cluster in *'gdh'* tree (Fig. 1B).

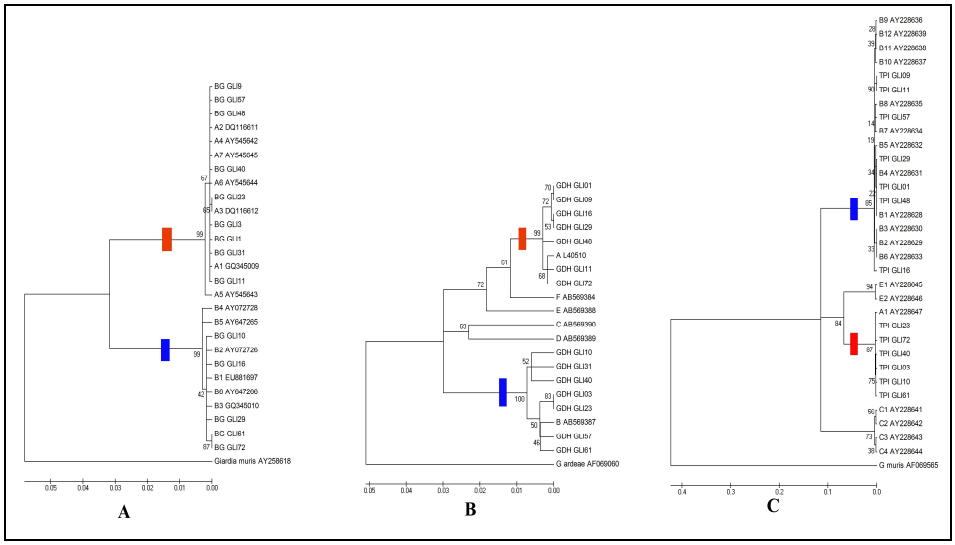



**Fig. 1. Phylogenetic analysis of *ß-giardin, Glutamate dehydrogenase and Triose phosphate isomerase loci* using the 'MEGA version 4 program'**

*The evolutionary history was inferred using the Neighbour-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap tests (1000 replicates) are shown next to the branches. The evolutionary distances have been computed using the Maximum Composite Likelihood method and are shown in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Assemblage A clusters are marked with RED bars and Assemblage B clusters are marked with BLUE bars. (1. A) Phylogenetic analysis of ß-giardin locus. (1. B) Phylogenetic analysis of Glutamate dehydrogenase locus. (1. C) Phylogenetic analysis of Triose phosphate isomerase locus.*

Homology analysis showed that the sequences within each assemblage A and B clusters for all the loci are mostly homologous (data not shown) taking all the isolates in account. Also in the combined phylogenetic analysis, these 14 isolates doesn't make any separate cluster and moreover they are evenly distributed in the respective cluster either assemblage A or B similar to Fig. 1. Hence, in this case the outcome of assemblage was solely depended on the marker loci. Association of genotype outcome with other physical factors such as Age, Sex and co-infection status was also checked but no particular association was observed. From this finding it can be said that the differential taxon position of this 14 isolates could not be due to chance and it is a true example of mixed assemblage. Two major reasons can be placed in favor of this outcome i.e. presence of mixed infection or co-infection of two different strain & occurrence of genetic recombination through sexual reproduction. Previous reports of mixed assemblages relied on the presence of multiple peaks in a particular position in the chromatograms which lead to the ambiguous taxon positioning of the isolates in the phylogenetic tree which can be due to mixed infection but, in this case all the isolates with dual genotype are positioned perfectly in a particular cluster for a specific loci. Although, recent reports suggest towards considering this type of observations as mixed assemblage infection, in spite of a thin probability of inter-assemblage recombination, still detailed molecular epidemiological study is required to find out the exact reason behind this unique finding. However, detection of high percentage of mixed genotype is evident, whether it is due to mixed infection or genetic recombination.

Other than sexual reproduction, genetic recombination can also occur through some extra-chromosomal genetic material which is capable of incorporating the host genome. Here comes the concept of GLV.

**GLV**

In 1985, the abundant presence of a 7-kb linear double-stranded RNA (ds RNA) species in the trophozoites of a *G. lamblia* strain, Portland I, which turned out to be the genome of a non-segmented, ds RNA virus was named as *Giardia lamblia virus* (GLV) (Ching C.Wang *et. al*). *Giardiavirus* is a small, icosahedron, nonenveloped virus comprising a monopartite double-stranded RNA genome, a major protein of 100 kDa, and a less abundant polypeptide of 190 kDa. It can be Isolated from the culture supernatant of *Giardia lamblia,* a parasitic flagellate in human and other mammals, and efficiently infects other virus-free *Giardia lamblia. (*Miller,R.L., *et. al*). Unlike the uninfectious *yeast killer virus* (ScV), *Leishmania RNA virus* (LRV), or other fungal viruses, GLV is shed into the culture supernatant and is highly infectious in its purified form. (Furfine, E.S., *et. al*). Virus particles, released by the transfected cells into the culture medium, were capable of infecting the virus-sensitive *Giardia lamblia* WB strain. Two aspects of GLV distinguish it from the rest of the totiviruses: i) its ability to infect *Giardia* trophozoite. ii)The fact that a single-stranded transcript of the GLV dsRNA genome can be isolated and introduced into *Giardia* trophozoites *by* electroporation to initiate *GLV* infection and replication. Moreover, GLV can possibly incorporate into host genome.

**GLV Detection procedure and outcomes:**

**RNA isolation and PCR amplification:** Viral RNA was isolated by using Viral RNA Minikit, Qiagen. cDNA preparation and PCR amplification was performed by using Superscript-III One Step RT-PCR kit, Invitrogen.

**Protocol**

- Various amounts (1–2 ng) of eluted RNA were subjected to RT-PCR using the SuperScriptIII One-Step RT-PCR kit (Invitrogen Cat no. 12574-030) and primers directed to the GLV-capsid protein sequence (GenBank L13218). [The forward primer sequence (GLV-CF) was 5'-GCCAGGATCTGGTAATTGCT-3' corresponding to nt 1251–1270; the reverse primer sequence (GLV-CR) was 5'-CTAGCGTCCTTTGAATACA-3' corresponding nt 1541–1569.]
- Viral RNA was denatured in the presence of GLV-CF and GLV-CR primers (0.4 mM) by heating at 94 C for 3 min, quick-chilled in wet ice, and subjected to RT-PCR following procedures provided by the manufacturer (Invitrogen).
- [RT-PCR consisted of 1 hr incubation at 53°C, followed by 3 min denaturation at 94°C, and then 40 cycles of 94°C for 30 sec, 53°C for 30 sec, and 68°C for 1 min, followed by a final extension at 68°C for 5 min]
- RT-PCR products were analysed by polyacrylamide gel electrophoresis followed by ethidium bromide staining.

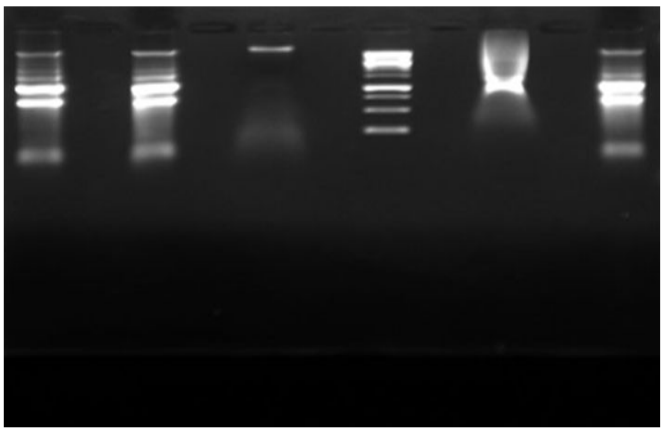



Fig: Isolated RNA with different treatment

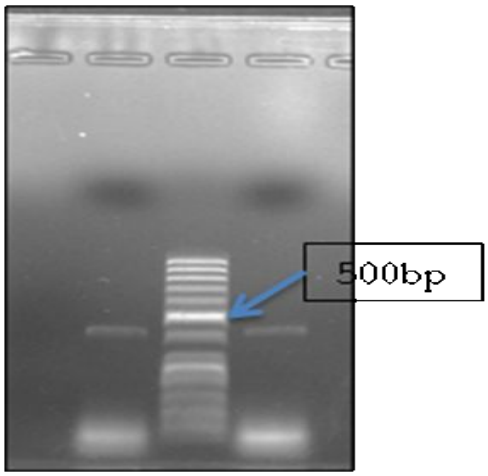



Fig: cDNA and PCR amplification using Superscript III one step RT PCR kit, Invitrogen (using Ref primer for GLV capsid protein.

**Sequencing Result:** The sequence obtained has no significant identity with reported cds of GLV capsid sequence. (alignment provided as FASTA file named GLV wg_cap_4F)
New PCR primers were designed targeting the conserved region of GLV capsid protein and PCR was done according to the reference protocol.

| Name of genes | sequence |
|---|---|
| Capv1F | 5'-CTGGTAATTGCTCACTTTCATC-3' |
| Capv1R | 5'-AACATATCCTTGTAAGCAGACC-3' |
| Capv2F | 5'-GCTCACTTTCATCGTCTATCTT-3' |
| Capv2R | 5'-CGGTGGAAACGTCGAGTG-3' |
| Capv3F | 5'-CTTCGAGAGCTCAATTCCACA-3' |
| Capv3R | 5'-GAAACGTCGAGTGAGGTGG-3' |
| Capv4F | 5'-CATGTCGAATAGAACGAGGTACT-3' |
| Capv4R | 5'-GTAACCATGGAAACATAGGG-3' |
| Capv5F | 5'-GGTCTGCTTACAAGGATATG-3' |
| Capv5R | 5'-AAGAACTGTGGGCCGCTCG-3' |

**PCR products**:

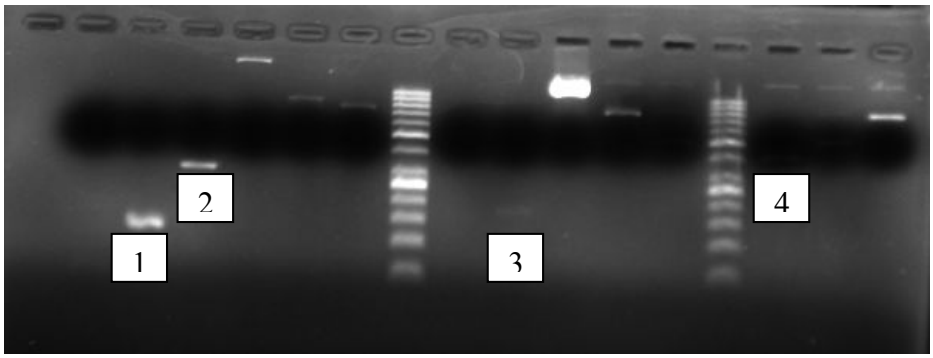



Fig: cDNA and PCR amplification using Superscript III one step RT PCR kit, Invitrogen (using new designed primer for GLV capsid protein.

**1 represents: >L21F (primer set 1F/1R)**
2 represents: >L31F (primer set
3 represents: >L45R
4 represents: >L51F

**Findings:**
Although the primers were designed against the GLV capsid protein but most of the PCR products with variety of PCR conditions were non-specific in nature. Few DNA bands from the desired base pair were purified and sequenced with the specific primers but the results obtained were not desirable (Table)

| Primer Set | Identity with (Accession no.) | Score | E value | Query coverage |
|---|---|---|---|---|
| L21F | XM_001705748.1 | 141 | 9e-31 | 95% |
| L31F | XM_001707957.1 | 326 | 7e-86 | 91% |
| L45R | XM_001706447.1 | 463 | 5.5e-16 | 90% |
| L51F | XM_001706605.1 | 582 | 7e-163 | 99% |
| 4F | XM_001706802.1 | 129 | 0.78 | 59% |

**Possible reasons for negative results:**

- The lab Portland I lacks GLV
- We should look for some better protocol

**vi) Future goals and Plan for fiscal year 2013-2014**

1. Establishing better protocol for GLV identification.
2. Identification of infection of GLV among different *G. l.* isolates in patients with differential infection, e.g. with sole *G. l.* infection and with multiple infections.

3. Identification of infection of GLV among non symptomatic *G. l.* isolates (if any) with sole and multiple infections.

**vii) Publications:**

Avik Kumar Mukherjee, Sumallya Karmakar, Dibyendu Raj and Sandipan Ganguly*. Multi-locus Genotyping Reveals High Occurrence of Mixed Assemblages in *Giardia duodenalis* within a Limited Geographical Boundary. 2013 ***British Microbiology Research Journal***. (In press).

## *3. Development of universal Shigella vaccine based on virulence gene expression.*

1. Title of Project:

**Development of a universal Shigella vaccine based on virulence gene expression.**

2. Name of investigators: **Hemanta Koley, Jiro Mitobe, GB Nair.**

3. Name of Student: **Ritam Sinha**

5 .Objective of Our Work:

Current vaccines for bacterial diseases have a serotype direction that limits the effect of vaccination to a narrow range of bacteria within the same species. An attempt to develop vaccine against broad serotype is worthwhile but it is really difficult. This may result from powerful immunogenicity of serotype specific polysaccharide antigen that could camouflage potential antigenicity of common virulence proteins. We have developed a candidate of broad range of Shigella vaccine based on molecular mechanism of virulence gene expression.

Keeping above ideas we have started work with these salient objectives:

a) To understand protective efficacy and immunogenicity of live genetically manipulated shigella vaccine strain against homologous as well as heterologous Shigella strains in guinea pig model.
b) To study the duration of protection offered by live genetically manipulated shigella vaccine strain.

6) Brief Report:

Bacillary dysentery caused by Shigella species, is a major cause of infant morbidity and mortality in developed as well as in developing countries. At present, only antibiotic therapy is available for treatment of shigellosis. Unfortunately, due to the global emergence of multidrug resistance, the choice of antimicrobial agents for treating shigellosis is very limited and we are approaching where the shigellosis can become an untreatable disease because of lake of an effective antibiotic. Therefore, the possibilities of other preventive measures such as anti-

dysentery vaccines have attracted increasing attention in this field. Various trials of several candidates' vaccine are being done in different parts of the world, but till date no suitable Shigella vaccine is available for public health use. There are different serotypes of Shigella species and their distribution varies between endemic geographical regions. *Shigella sonnei*, the most frequent serogroup in developed countries, accounted for 40–57% of shigellosis infections reported between 1994 and 2002 . In India and Bangladesh, *S. dysenteriae* 1 epidemics spread to adjoining areas, had high attack rates and showed extremely high fatality rates among children and adults. Sanitation, and improved hygiene and water would help to control the disease but progress is slow in the poor communities where the disease is most prevalent. Therefore, the only hope to control *S. dysenteriae* 1 and *Shigella sonnei* infection is to develop a vaccine. The immune response against Shigella species are serotype-specific, so current immunization strategies have required the administration of live vaccine strains to provide protection against multiple serotypes. In our study, we evaluated the protective efficacy and immune response live attenuated shigella in guinea pig model. Constriction and preliminary protection work done by our Japanese scientist.

The protective efficacy after eye immunization with four doses (0, 1$^{st}$, 14$^{th}$, 15$^{th}$ Day) of Shigella strain was examined. The protection following challenge was 100% protection (against *Shigella dysenteriae*) in the immunized group whereas the non immunized group of animals developed keratoconjunctivitis mean zero protection ( Fig 1).

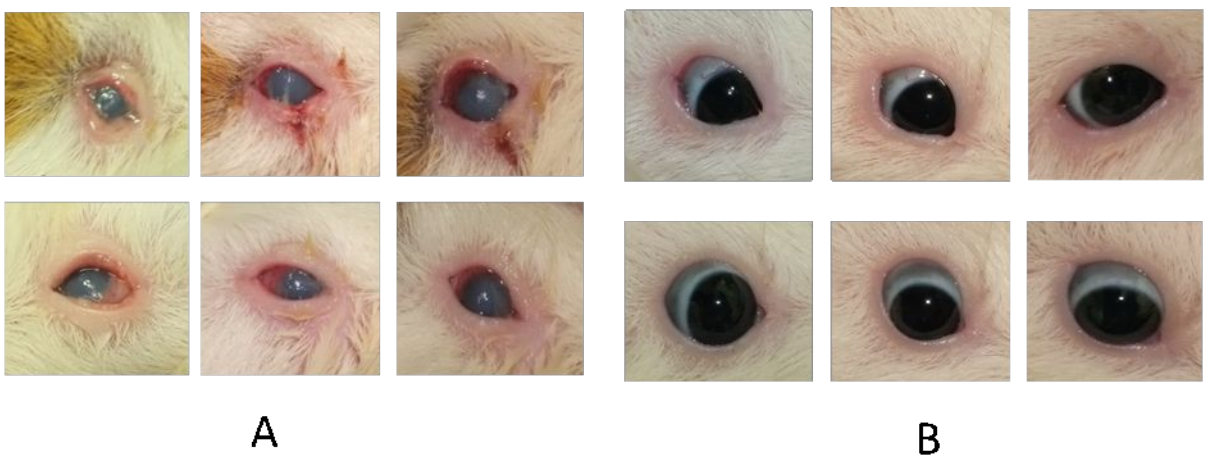



**Fig 1: Keratoconjunctisvitis developed after challenged with *Shigella dysenteriae* 1 to the Non - immunized guinea pigs ( A). No effect to the immunized animals ( B)**

Nature of colonization ability after challenged with *S. sonnei* from the distal colon of animals of both the immunized and control groups was observed. In animals challenged on day 42 after the first day of the immunization, at 72 hrs after the challenge, at most $5\pm2.8\times10^{3}$cfu per gram of tissue were recovered

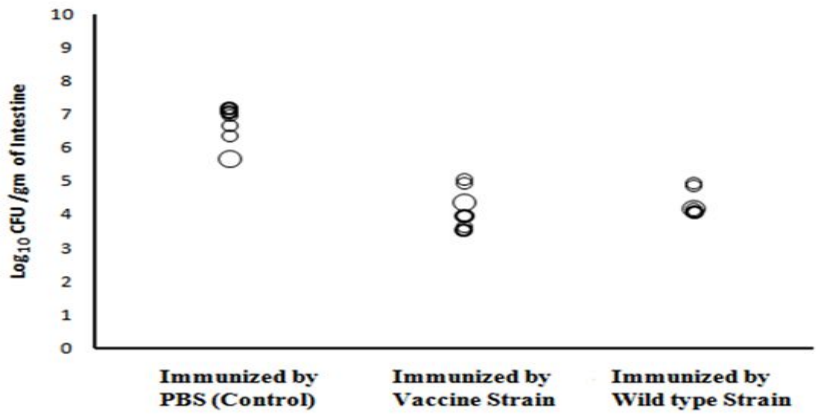



**Fig 2: . Nature of colonization ability of challenge strain *S. sonnei* in the distal colon among challenged immunized and non-immunized animals.**

from animals of the non-immunized group, whereas $1.8\pm 2.8 \times10^{9}$ cfu per gram of tissue were recovered from animals of immunized group of animal (Fig: 2 ).

**Table 1: Protective efficacy of the immunization with Vaccine strain MF 4831 (hfq mutant) and wild type strain.**

| Experimental animal | Immunogen used in Eye. | Challenged Strain used in intestine | Number of animal used | Disease symptoms | % of death with Shigellosis | % of protection against Shigellosis |
|---|---|---|---|---|---|---|
| PBS Control Group | Nil | Wild type S. sonnei | 6 | Shigellosis 100% ( 6/6) | 100% (6/6) | 0% ( 0/6) |
| Immunized Group | VACCINE STRAIN MF4831 | Wild type S. sonnei | 6 | 33.3% ( 2/6) | 0 % (0/6) | 100% |
| | Wild Type Strain ( S. flexneri 2a 2457T ) | Wild type S. sonnei | 5 | 40% ( 2/5) | 0 % (0/5) | 100% |

After 42th day of first immunization, luminal challenge was performed with *Shigella sonnei* to observed ectopic effect. We have observed that immunized groups of animals showed 100% protection against *Shigella sonnei* challenged. Among four animals of immunized group excreted semisolid stools within 24 hours after the challenge but recovered spontaneously within 48 hours. On the other hand, all animals of the control group challenged developed symptoms of bacillary dysentery, such as tenesmus and mucoidal and bloody diarrhea and died within 72 hrs. (Table1).

The results of the current study might suggest us that **this MF 4831 (hfq mutant)** vaccine strain could be a promising future vaccine candidate, to prove, we need to work more detailed in animal models.

7) Reference:


1. Mitobe, J., Morita-Ishihara, T., Ishihama, A., and Watanabe, H. Involvement of RNA-binding Protein Hfq in the Post-transcriptional Regulation of *invE* Gene Expression in *Shigella sonnei*[*](2008) *J BiolChem* 283, 5738-5747
2. Mitobe, J., Morita-Ishihara, T., Ishihama, A., and Watanabe, H. Involvement of RNA-binding protein Hfq in the osmotic-response regulation of *invE* gene expression in *Shigella sonnei* (2009) *BMC Microbiol* 9, 110
3. Sereny, B. Experimental keratoconjunctivitis shigellosa. 1957. *ActaMicrobiol. Acad. Sci. Hung.*(4:367-376.
4. Sack RB, Kline RL &SpiraWM Oral immunization of rabbits with enterotoxigenic Escherichia coli protects against intraintestinal challenge (1988). *Infect Immun* 56: 387-394.
5. Barman S, Saha DR, Ramamurthy T, Koley H. Development of a new guinea-pig model of shigellosis. (2011), *FEMS Immunol Med Microbiol*.. 62(3):304-14.

## *4. Analysis of HLA associated HIV-1 mutations in India and Japan.*

Initial few months were utilized for recruitment and training of field staffs, collaborating with NGOs of Manipur and West Bengal etc. Sample collection started from August 2012. Eligible HIV positive patients with and without receiving ART were approached for the study by our social/field workers to explain the purpose of this study and to seek their voluntary participation. A team from NICED visits these states every month to monitor and supervise sample and data collection procedure. So, far **42** blood samples have been collected from eligible candidates by trained laboratory technicians along with data related to their socio-demography, risk behavior and history of ART compliance. Written informed consent has been obtained from all eligible and willing candidates. Blood samples were transferred to NICED laboratory on same day of the collection. Data editing and entry are being carried out at present.

---

**Viral Load estimation**

Viral load is estimated using the Cobas Amplicor HIV-1 monitor test version 1.5 (Roche) by quantitaing HIV-1 RNA in human plasma on the cobas amplicor analyser .

1. 200µl of blood plasma is taken for RNA extraction and finally eluted in 400 µl of diluent medium.
2. 50 µl of RNA is mixed with 50 µl of provided master mix for the reverse transcription of target RNA to generate cDNA.
3. It is followed by PCR amplification of target cDNA by specific complementary primers.
4. The amplified products are then hybridized to target specific oligonucleotide probes.
5. Finally, the probe bound amplified product is detected by colorimetric method indicating the viral load estimated by RNA copies / ml of plasma.

The viral load results of the 42 samples have been shown in **Table 1**.

| SERIAL NO. | SAMPLE I.D. | CD4 COUNT | PLASMA VIRAL LOAD (HIV-1 RNA copies/ml) |
|---|---|---|---|
| 1 | JK-RI-000-4615 | 36 | 16800 |
| 2 | RPC-DEC-K3747/12 | 80 | 575 |
| 3 | RG-MCH-000-517 | 58 | 18800 |
| 4 | DURG/CG/0684 | 40 | 35700 |
| 5 | WB/MR/000/170 | 140 | 585 |
| 6 | WB/TM/00/2964 | 37 | 49 |
| 7 | JK/MG/00881 | 96 | 41000 |
| 8 | WB-RPC-40 | 373 | 90 |
| 9 | WB-MLD-73 | 118 | 35 |
| 10 | WB/STM/00/1047 | 273 | 26600 |
| 11 | WB-TM-00-6720 | 48 | 156000 |
| 12 | WB-RPC-592 | 89 | 83 |
| 13 | RGKAR-MCH-000-415 | 176 | 27200 |
| 14 | RGKAR-MCH-001066 | 136 | 165000 |
| 15 | JK-DHN-017 | 96 | 43000 |
| 16 | MDH-000-638 | 1120 | 4700 |
| 17 | DURG/0653 | 189 | 165000 |
| 18 | WB/MR/000/107 | 198 | 142000 |
| 19 | AS/GMC/1175 | 137 | 53100 |
| 20 | JK/RI/00/130 | 32 | 31500 |
| 21 | WB/STM/0000/77 | 12 | 14000 |
| 22 | RAIPUR/CG/08/1058 | 64 | 133000 |
| 23 | WB/MMCH/000/106 | 511 | 239 |
| 24 | JR/RI/000/2431 | 36 | 15300 |
| 25 | BMCH/0000/53/2010 | 46 | 143000 |
| 26 | JK/DHN/009 | 279 | 165 |
| 27 | WB/MR/000/723 | 357 | 105 |
| 28 | WB/TM/2660 | 51 | 1640000 |
| 29 | WB/NB/32 | 226 | 399000 |
| 30 | WB/RPC/001214 | 82 | 162000 |
| 31 | DURG/CG/1145 | 135 | 862000 |
| 32 | WB/TM/2883 | 84 | 4420 |
| 33 | WB/STM/00/5598 | 401 | 72900 |
| 34 | WB/TM/005011 | 171 | 10600 |
| 35 | WB/TM/9604 | 264 | 84500 |

| | | | |
|---|---|---|---|
| 36 | AS/GMC/1228 | 55 | 199000 |
| 37 | RPC/K-677 | 307 | 11000 |
| 38 | WB/RPC/K-192 | 135 | 61700 |
| 39 | OR/CTC/87/07 | 80 | 5590 |
| 40 | WB/TM/00/3674 | 96 | 76000 |
| 41 | WB/BMCH/000/016 | 230 | 92 |
| 42 | WB/TM/00/5299 | 162 | 48 |

**HIV-1 sub-typing and genetic characterization of the HIV-1 seropositive samples**

In continuation to the study of the molecular characterization of HIV-1 strains, till date **42** samples were analysed. Multi Region Hybridization Assay were performed with these samples. The blood samples were collected in Na-citrate solution after pre- and post-test counseling. HIV-1 sero-positivity was determined by rapid spot test (Immunocomb HIV-1/2 Bi-spot, Orgenics, Israel), followed by ELISA (Immunogenetics, Belgium) and line immunoassay (Inno-LIA, HIV-1/HIV-2). Peripheral blood mononuclear cells (PBMCs) were separated from whole blood by Ficoll-Hypaque gradient centrifugation.

*A) Multi-region Hybridization Assay* (*MHA) :*

An advanced genotyping method multi-region hybridization assay (MHA) was commenced for sub-typing analysis and determination of the recombination break points, among the HIV-1 positive samples. MHA was performed on the basis of 8 different genomic regions of HIV-1. The principle of the MHA was to amplify multiple short fragments throughout the HIV-1 genome and to assess the hybridization of clade specific fluorescent probes in real time PCR. Almost all the genomic regions of 40 samples were tested each with probes for subtype B, C and AE. The amplified regions were positioned in areas of the genome that would tend to maximize discrimination among different strains of interest. Viral RNA was extracted by Viral RNA Mini Kit (QIAGEN, Germany). The RT-PCR and first round PCR was done following proper PCR conditions. All the genomic regions of HIV-1 genome viz. *p17*, pro, reverse transcriptase (*rt*), *int*, *tat*, gp120, gp41 and *nef* were amplified in separate 1$^{st}$ round PCRs. The second round real time PCR was done using TaqMan 2x universal PCR master mix (ABI), 400 μM of each inner primer, 250 μM of probe and 1 μl of 1$^{st}$ round PCR product. Real time PCR amplification was performed in a 96 well ABI PRISM 7900HT sequence detection system (Applied Biosystem) with following routine: hold at 95°C-10 mins and 95°C-10s, 60°C-1 min and in case of gp120 region, the extension step was performed at

60°C-2 mins. Fluorescence intensity was measured during the reaction and was analysed by using the SDS v2.1 software (Applied Biosystem).

MHA analysis showed that in case of 35 samples, all the genomic regions (*p17*, *pro*, *rt*, *int*, *tat*, gp120 and *nef*) of HIV-1 reacted with the probes belonged to subtype C whereas 5 of the remaining samples showed probe reactivity for subtype C (*p17*, *pro*, *int*, *tat*, gp120 and *nef)* genes but the *rt* region of the *pol* gene showed dual probe reactivity.

**Table 3**

**(a) Results for multiregion hybridization assay (MHA):**

| **Subtype** | **Total no. of samples** | **No. of samples** | **Genomic regions** |
|---|---|---|---|
| C | 35 | 35 | p17, pro, rt, int, tat, gp120, gp41, nef |
| B | | - | p17, pro, rt, int, tat, gp120, gp41, nef |
| Dual reactivity | | 5 | pol |
| Multigenomic variants | | - | - |

The AIDS epidemic in Manipur, India, manifests unique features, having co-circulation of B and C HIV-1 subtypes along with recombinant forms. Manipur has the highest incidence of HIV-1 infection compared to the other states of India, but limited information is available regarding the full-length sequence of HIV-1 recombinants.

Viral RNA, extracted from the plasma of a male injecting drug user diagnosed with HIV-1 infection. Near full-length genome was amplified by polymerase chain reaction using primer walking approach. Phylogenetic relationships were determined with neighbor-joining trees. The recombination break points were detected using boot scan and Simplot analyses.

*B) Near full length genome sequencing:*

This recombinant predominantly had subtype C genome and exhibited mosaic structures with subtype B insertions at three different positions of HIV-1 genome. Simplot analysis of near full-length genome sequence from the recombinant HIV-1 strain, MAN86 exhibited similarity with the sequence of C.IN.93.93IN905 in its subtype C backbone, while the subtype B insertions showed resemblance with the sequence of B.TH.99.99$^{TH}$_C1416. This study confirms the presence of a unique recombinant HIV-1 strain, emerging as a result of

recombination between HIV-1 strains from India and Thailand (Fig.1). The Simplot analysis was shown in (Fig. 2) for sample no. MAN-86 (HIV-1 strain).

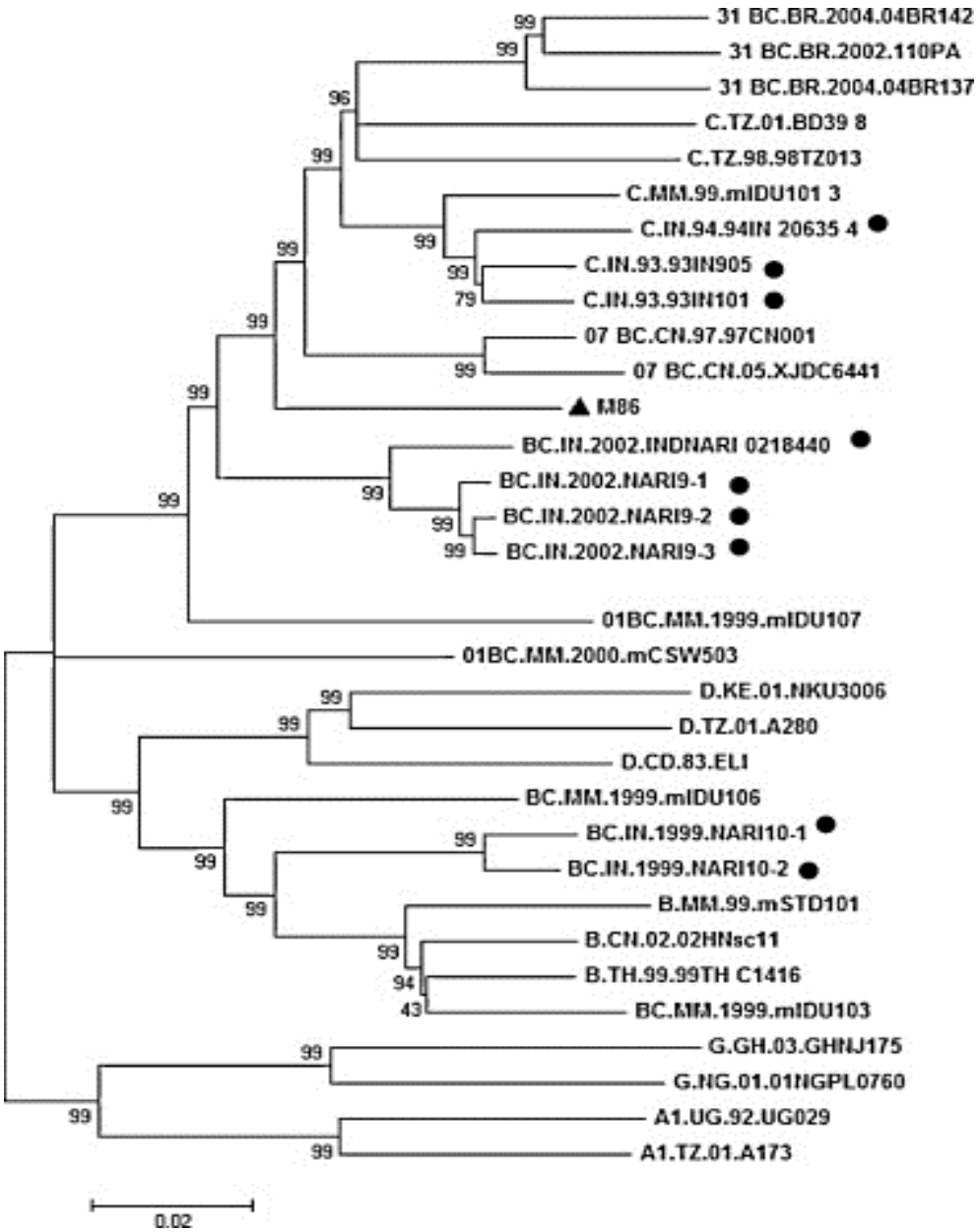



Fig. 1. Phylogenetic analysis of virtually full-length genome sequences. The tree was constructed by the neighbor-joining method. The Manipur sequence (MAN86) is shown as '▲', Indian strains are shown as '●' and the unmarked ones represent other global strains. The scale bar represents 10% difference. Interpatient HIV-1 diversity across the entire genome or individual genes among Manipur isolate from IDU compared with HIV-1 sequences from other global HIV-1 strains. The interpatient genetic distances were calculated using the Kimura two-parameter distance measurement.

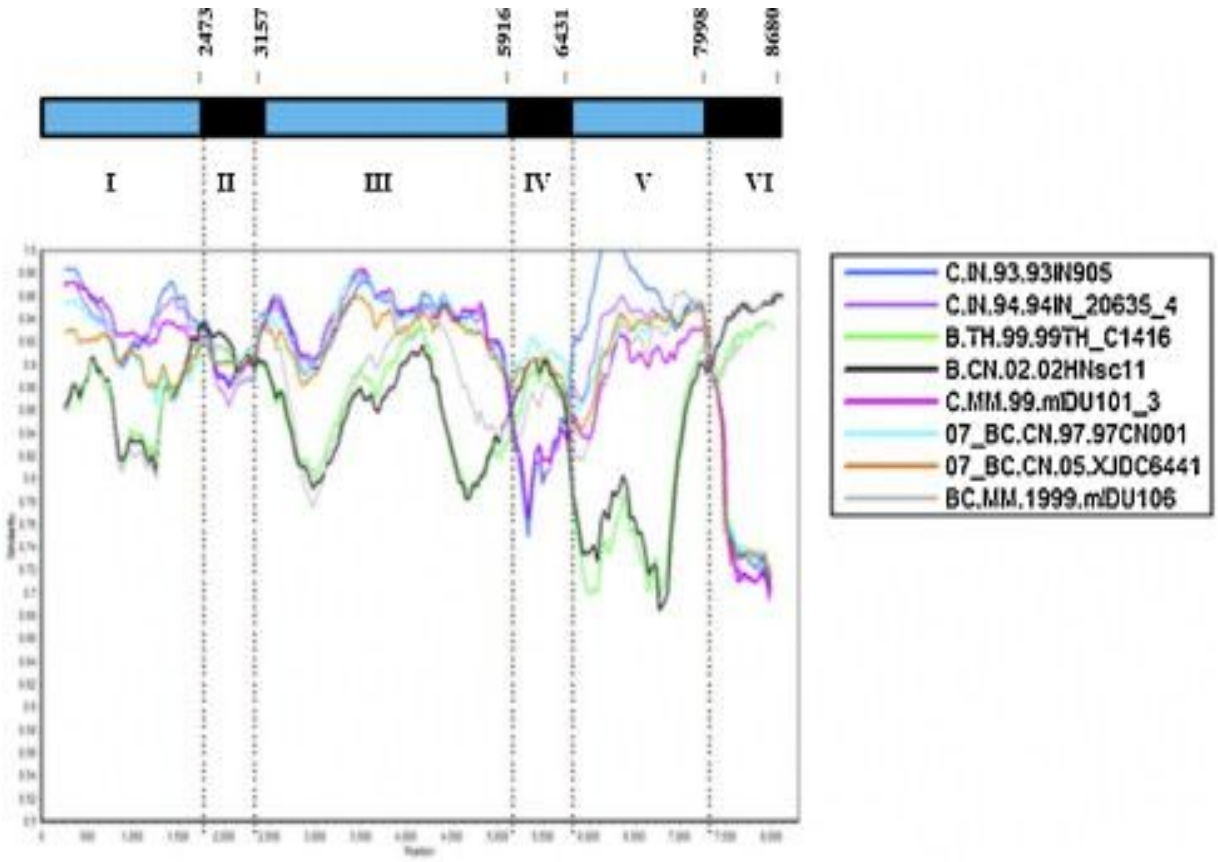



Fig. 2. Simplot analysis of a HIV-1 strain (MAN86) isolated from an IDU of Manipur. The HIV-1 strains included are C.IN.93.93IN905, C.IN.94.94IN_20635_4, B.TH.99.99$^{TH}$_C1416, B.CN.02.02.HNsc11, C.MM.99.mIDU101_3, 07_BC.CN.97.97.CN001, 07_BC.CN.05XJDC6441, BC.MM.1999.mIDU106. Recombinant viruses are indicated by the subtype designation followed by the name of the isolate. Bootstrap values "70%", based on 100 replicates, of some key nodes are shown. The Breakpoint positions were obtained using Simplot 3.5.1 and numbered according to HXB2 reference. The schematic drawing was performed to divide the HIV-1 genome of MAN86 into six different fragments based on mosaic recombination pattern. Subtype C and subtype B regions were shown in blue and black colors.

Recombinant viruses are indicated by the subtype designation followed by the name of the isolate. Bootstrap values 70 %, based on 100 replicates, of some key nodes are shown. The Breakpoint positions were obtained using Simplot 3.5.1 and numbered according to HXB2 reference.

## Publications:

1. **Roni Sarkar**, Kamalesh Sarkar, N. Brajachand Singh, Y. Manihar Singh, Debashis Mitra, Sekhar Chakrabarti. **Emergence of a unique recombinant form of HIV-1 from Manipur, India.** ***Journal of Clinical Virology***, **(2012)**. **55: 274- 277**.