Chapter 5

Phylogenetic studies of KFDV
CHAPTER – 5
PHYLOGENETIC STUDIES OF KFDV

1. Introduction
   a. Genome organization in flaviviruses
   b. Genes encoding structural proteins
      i. Capsid
      ii. PreM and M
      iii. Envelope
   c. Genes encoding non-structural proteins
      i. NS5 region
   d. Objectives of genotyping

2. Material and methods
   a. Selection of virus isolates proposed to be taken for genotyping
   b. Primer selection
   c. Virus propagation in-vivo
   d. Virus propagation in-vitro
   e. RNA extraction and amplification
   f. Sequencing of PCR products, assembly and curation
   g. Microevolution and macroevolution
   h. Test of selection
      i. Phylogenetic analysis
   j. Impact of low passage levels in tissue cultures

3. Results
   a. Virus propagation in-vivo
   b. Virus propagation in-vitro
   c. Phylogenetic studies of strains from patient’s serum
   d. Amplification of Envelope and NS5 gene fragments by RT PCR
   e. Sequencing results
   f. Multiple sequence alignment and phylogenetic analysis
   g. Microevolution and macroevolution
   h. Test of selection and calculation of divergence time
      i. Construction of phylogenetic tree based on colinearized data
   j. Impact of low passage levels in tissue culture
4. Discussion
   a. Phylogenetic analysis of strains from patient’s serum
   b. Phylogenetic studies based on two gene segments
   c. Probable factors for KFD spread to neighboring areas
      i. Construction of dams
      ii. Dynamics of forest cover
      iii. Climatic conditions
      iv. Vaccination programs
   d. Case report from Gulbarga district, Karnataka

1. INTRODUCTION
   a. Genome organization in Flaviviruses

      The flavivirus genome is around 11 kb and comprises of genes that encode a single polypeptide, which is cleaved into 10 different proteins. Mature flavivirus virions contain three structural and seven non-structural proteins. The structural proteins comprise of a nucleocapsid or core protein (C; 12 kDa), a non-glycosylated membrane protein (M; 8 kDa) and an envelope protein (E; 53 kDa), which is usually glycosylated.

   b. Genes encoding structural proteins
   b(i). Capsid

      The virion C protein is a small [predicted molecular mass (Mw) 11 kDa], highly basic protein that forms a structural component of the nucleocapsid. Sequence homology among C proteins from different flaviviruses is low, but regions of hydrophobic and hydrophilic amino acids are conserved. C protein participates in RNA and protein interaction which is important for nucleocapsid assembly, have not been defined.

   b(ii). PreM and M

      The prM protein is the glycosylated precursor (Mw, 26 kDa) of the structural protein M (Mw, 8 kDa). The prM undergoes a delayed cleavage to form M and the N-terminal pr segment, which is secreted into the extracellular medium. This cleavage occurs shortly before or coincident with virion release because prM and M are found on intracellular and extracellular viorn, respectively. The N-terminal pr segment is predominantly hydrophilic and contains one to three N-linked glycosylation sites and six conserved cysteine residues, all of which participate in intramolecular disulfide bridges.
The structural protein M, located in the C-terminal portion of prM is present in mature virions and contains a shortened ectodomain (41 amino acids) followed by two potential membrane-spanning domains. Antibodies to prM can mediate protective immunity perhaps by neutralization of released virions that contain some uncleaved prM.

b(iii). Envelope

The E protein is the major envelope protein of the virion. This protein is believed to play key roles in a number of important processes including virion assembly, receptor binding and membrane fusion and is the principal target for neutralizing antibodies. *Mutations in this protein can often have dramatic effects on viral pathogenesis.* Comparison of deduced E protein sequences shows areas of striking homology as well as divergence amongst the flaviviruses. All the twelve Cys residues in the E ectodomain are strictly conserved and involved in intramolecular disulfide bonds. The E protein is glycosylated in some, but not all the flaviviruses and the role of N-linked glycosylation in E function is very clear.

Virus interaction with the cell surface is important for explaining the pathogenesis of KFDV. The interaction between specific viral surface protein and the receptor on the target cell play an important role in viral tropism. Any constituent of cell membrane including carbohydrate, lipids and proteins as single or in combination may play a role as viral receptor depending on virus. For attachment and entry viruses use the cell membrane components serving normal function.

The viral envelope (E) glycoprotein embedded in lipid bilayer plays an important role in virus cell attachment; two sites of N linked glycosylation is supposed to play an important role in virus cell attachment. The antibodies against E protein have shown protection in mice proving its role in infection.

The E protein is a major component of the virion surface; it contains the important antigenic determinants sub serving haemagglutination-inhibition and neutralization and thus induce immunological responses in the infected host. Structural elements of the E protein are assumed to be involved in the binding of virions to cell receptors and in intraendosomal fusion at low pH (Monath, 1990).

c. Genes encoding non-structural proteins

c(i). NS5 gene

This is the last protein encoded in the long ORF, the largest (MW 103 - 104 kDa) and most highly conserved flavivirus protein. NS5 is a basic protein, lacking any long
hydrophobic stretches and is believed to be the flavivirus RNA-dependent RNA polymerase. This assumption is based on the presence of a highly conserved region, including the sequence motif GDD (YF NS5 residues 666 to 668), which is characteristic of known or putative RNA-dependent RNA polymerases of positive-strand RNA viruses.

The N-terminal domain of NS5 (between residues 60 and 45) is homologous to a region of methyltransferases implicated in S-adenosylmethionine binding. It has been suggested that this domain may be involved in methylation of the 5' cap structure. Although enzymological characterization of the protein is lacking, it seems likely that the NS5 protein is at least bifunctional, possessing both methyltransferases and RNA polymerase activities. Phylogeny was carried out of all the available strains from India using NS5 gene, which encodes polymerase and is conserved among the members of flaviviridae and E gene.

d. Objectives of genotyping

Genotyping of Indian strains was intended to know:

- What is the genetic diversity of KFDV involved in human infections?
- Can the relationship between human, tick and monkey infections be explained on the basis of a particular isolate or the source of infection?
- What is the antigenic similarity between human and tick isolates and the leap towards future of vaccine development?
- Can better candidate vaccine stocks be developed for an effective response in case multiple strains are circulating and which may be antigenically different from the current vaccine strain?
- What could be the possible cause of recurrent epidemics in spite of vaccination programs?

Genetic plasticity and adaptability to an environment is also one of the crucial factors to be looked upon. Partial regions of the two genes; structural (preM, M, Envelope) and non-structural gene (NS5) were the bottleneck of this study. Finally the complete structural genes of the selected viral strains were used based on the previous analysis of partial genes in order to understand the genetic diversity of KFDV.
2. MATERIAL AND METHODS

a. Selection of virus isolates for genotyping

Over 200 isolates were available at the National virus repository of which 50 isolates were selected to represent various host species, different geographical locations and different time periods (Table 12). These include 16 samples from human, 12 from monkey, 3 from rodent and 19 from ticks. 49 isolates are from year 1957 to 1972 with minimum of 2 isolates representative of each year whereas one isolate from human serum was obtained from year 2006.

b. Primer selection

Primers for Envelope and NS5 region of KFDV were designed from a nucleic acid sequence taken from NCBI Entrez database (NCBI accession # X74111, AF013385 respectively) (Table 13). The regions of 642 bp from NS5 gene and structural region were selected for amplification and further phylogenetic analysis.

c. Virus propagation in-vivo

For preparation of KFDV stock, 1-2 day old suckling Swiss albino mice were used for propagation. The infants were injected intracerebrally with KFDV procured from virus registry of National Institute of Virology. Following strains of the virus were used: W1930, 601011, G27678, 664518. The stored virus ampoules were reconstituted in 0.5 ml distilled water. Each mouse received intracerebrally about 20μl of virus suspension. These mice were observed daily for sickness. When mice showed sickness, the brains of these animals were dissected out with sterile scissors-forceps. To avoid any cross contamination one virus was handled at a time. The brains were collected in pre-weighed sterile vials. 10% w/v solution of brain was prepared in BAPS. After centrifugation at 10000 rpm for 60 min the solution was aliquoted. The vials were stored at -80°C until used.
<table>
<thead>
<tr>
<th>#</th>
<th>Isolate</th>
<th>Yr.</th>
<th>Locality</th>
<th>Original source</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>W379</td>
<td>1957</td>
<td>Baraghi</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>2</td>
<td>P9605</td>
<td>1957</td>
<td>Shigga</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>3</td>
<td>G11333</td>
<td>1957</td>
<td>Barasi</td>
<td>Haemaphysalis spinigera</td>
<td>T</td>
</tr>
<tr>
<td>4</td>
<td>W399</td>
<td>1957</td>
<td>Shigga</td>
<td>Macaca radiata</td>
<td>M</td>
</tr>
<tr>
<td>5</td>
<td>P16011</td>
<td>1958</td>
<td>Kaisodi</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>6</td>
<td>W399</td>
<td>1958</td>
<td>Hessari</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>7</td>
<td>W6043</td>
<td>1959</td>
<td>Belusiri</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>8</td>
<td>W6178</td>
<td>1959</td>
<td>Koppalgadde</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>9</td>
<td>G27667</td>
<td>1959</td>
<td>Kunvalhalli</td>
<td>Haemaphysalis spinigera from dead monkey</td>
<td>T</td>
</tr>
<tr>
<td>10</td>
<td>P20924</td>
<td>1959</td>
<td>Mullahre</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>11</td>
<td>P21092</td>
<td>1959</td>
<td>Halagpur</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>12</td>
<td>601203</td>
<td>1960</td>
<td>Tuddikoppa</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>13</td>
<td>611661</td>
<td>1961</td>
<td>Sagar station</td>
<td>H. turturis</td>
<td>T</td>
</tr>
<tr>
<td>14</td>
<td>612057</td>
<td>1961</td>
<td>Barur</td>
<td>Rattus rattus wrightoni</td>
<td>R</td>
</tr>
<tr>
<td>15</td>
<td>62844</td>
<td>1962</td>
<td>Hillensarur</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>16</td>
<td>62849</td>
<td>1962</td>
<td>Hillensarur</td>
<td>Rattus rattus wrightoni</td>
<td>R</td>
</tr>
<tr>
<td>17</td>
<td>62957</td>
<td>1962</td>
<td>Hillensarur</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>18</td>
<td>62969*</td>
<td>1962</td>
<td>VRC Poona</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>19</td>
<td>63661</td>
<td>1963</td>
<td>Malve</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>20</td>
<td>63696</td>
<td>1963</td>
<td>Sarnagadde</td>
<td>P. entellus</td>
<td>M</td>
</tr>
<tr>
<td>21</td>
<td>64244</td>
<td>1964</td>
<td>Balagodu</td>
<td>P. tarsius</td>
<td>T</td>
</tr>
<tr>
<td>22</td>
<td>64350</td>
<td>1964</td>
<td>Murasara</td>
<td>H. formosensis</td>
<td>T</td>
</tr>
<tr>
<td>23</td>
<td>642034</td>
<td>1964</td>
<td>Kangodou</td>
<td>H. turturis</td>
<td>T</td>
</tr>
<tr>
<td>24</td>
<td>642046</td>
<td>1964</td>
<td>Kangodou</td>
<td>H. papauana kinneari</td>
<td>T</td>
</tr>
<tr>
<td>25</td>
<td>652</td>
<td>1965</td>
<td>Kangodou</td>
<td>H. wellingtoni</td>
<td>T</td>
</tr>
<tr>
<td>26</td>
<td>651521*</td>
<td>1965</td>
<td>VRC Poona</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>27</td>
<td>652980</td>
<td>1965</td>
<td>Vadhala</td>
<td>Haemaphysalis spp.</td>
<td>T</td>
</tr>
<tr>
<td>28</td>
<td>6616</td>
<td>1966</td>
<td>Yelagadale</td>
<td>P. entellus</td>
<td>M</td>
</tr>
<tr>
<td>29</td>
<td>66364-1</td>
<td>1966</td>
<td>VRC staff Sagar</td>
<td>2 day acute human serum</td>
<td>H</td>
</tr>
<tr>
<td>30</td>
<td>66928-2</td>
<td>1966</td>
<td>Sagar</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>31</td>
<td>664518</td>
<td>1966</td>
<td>Konagadale</td>
<td>H. turturis</td>
<td>T</td>
</tr>
<tr>
<td>32</td>
<td>67965</td>
<td>1967</td>
<td>Sagar</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>33</td>
<td>671004</td>
<td>1967</td>
<td>Bhadrapura</td>
<td>P. entellus</td>
<td>M</td>
</tr>
<tr>
<td>34</td>
<td>673514</td>
<td>1967</td>
<td>Siraval</td>
<td>H. papauana kinneari</td>
<td>T</td>
</tr>
<tr>
<td>35</td>
<td>68142</td>
<td>1968</td>
<td>Hotagadale</td>
<td>P. entellus</td>
<td>M</td>
</tr>
<tr>
<td>36</td>
<td>68159</td>
<td>1968</td>
<td>Siraval</td>
<td>H. turturis</td>
<td>T</td>
</tr>
<tr>
<td>37</td>
<td>68484</td>
<td>1968</td>
<td>Halagalale</td>
<td>Rattus blanfordi</td>
<td>R</td>
</tr>
<tr>
<td>38</td>
<td>681960</td>
<td>1968</td>
<td>Barur</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>39</td>
<td>692156</td>
<td>1969</td>
<td>Chikkanallur</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>40</td>
<td>692163</td>
<td>1969</td>
<td>Thonagoda</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>41</td>
<td>712419</td>
<td>1971</td>
<td>Nodahalli</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>42</td>
<td>713038</td>
<td>1971</td>
<td>Jaidinakere</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>43</td>
<td>716810</td>
<td>1971</td>
<td>Gunjir</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>44</td>
<td>72166</td>
<td>1972</td>
<td>Gadgeri-sirsi</td>
<td>H. kyasanuensis</td>
<td>T</td>
</tr>
<tr>
<td>45</td>
<td>72827</td>
<td>1972</td>
<td>Holekoppa</td>
<td>P. entellus</td>
<td>M</td>
</tr>
<tr>
<td>46</td>
<td>A106</td>
<td>2006</td>
<td>Chikkaassulur</td>
<td>Human</td>
<td>H</td>
</tr>
<tr>
<td>47</td>
<td>W6204</td>
<td>1959</td>
<td>Kannahalli</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>48</td>
<td>G27678</td>
<td>1959</td>
<td>Kopalgadde</td>
<td>H. spinigera</td>
<td>T</td>
</tr>
<tr>
<td>49</td>
<td>W1930</td>
<td>1958</td>
<td>Chennnoor</td>
<td>Semnopithecus entellus</td>
<td>M</td>
</tr>
<tr>
<td>50</td>
<td>601201</td>
<td>1960</td>
<td>Chikkaassulur</td>
<td>Human</td>
<td>H</td>
</tr>
</tbody>
</table>

* - Samples obtained from humans after lab infection. M = Monkey, T = Tick, H = Human, R = Rodent.
Table 13: List of the primers used for amplification and the internal primers for sequencing

<table>
<thead>
<tr>
<th>Region</th>
<th>Primer Name</th>
<th>Genome location</th>
<th>Primer Sequence (5' → 3')</th>
<th>Tm</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>preM-Env</td>
<td>EF2</td>
<td>459</td>
<td>TGGTGTTCTCTGCGACAGTT</td>
<td>56.9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>ER2</td>
<td>1258-1238</td>
<td>TCCTGCACCTCGTCTCGCTT</td>
<td>57.9</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>EF3*</td>
<td>606</td>
<td>TCATCGAGTGTGTTGACACCATT</td>
<td>56.8</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>ER1*</td>
<td>701-678</td>
<td>TTCCGTATCCCAGTGACATCGCT</td>
<td>59.9</td>
<td>24</td>
</tr>
<tr>
<td>NS5</td>
<td>F3</td>
<td>9422-9441</td>
<td>GGCCTGATGCATGGACATCAT</td>
<td>58.45</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>R4</td>
<td>10046-10063</td>
<td>TCCATTGTTGGTATGCT</td>
<td>58.67</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>F4*</td>
<td>9660-9680</td>
<td>TGAGACCTTCTGACGGCGCT</td>
<td>60.2</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>R3*</td>
<td>9801-9819</td>
<td>TTCTTCATGGAAACTCAT</td>
<td>51.5</td>
<td>19</td>
</tr>
</tbody>
</table>

* - Internal primers used only for sequencing for some of the problematic sequences

Figure 37: Schematic representation of the relative genomic positions of the primers used for amplification from structural and NS5 gene. Black arrows – primers used for amplification of PCR products. White arrows – Internal primers used for sequencing some of the isolates.

d. Virus propagation *in-vitro*

The lyophilized stock was reconstituted in distilled water and 0.6 ml was used for infection of Vero and PS cell lines in 25cm² bottles along with normal control. The suspension was kept for viral adsorption for 1 hr, followed by MEM supplemented with 2 % FCS was added to the tissue culture bottles. By the 6th day of infection the infected Vero cells showed typical rounding of cells, clumping and complete disruption of the monolayer.

c. RNA extraction and amplification

Total RNA was extracted from 250 μl of either 10% mouse brain suspension or infected Vero cell lysate and added to 750 μl Trizol (Gibco-BRL). The aqueous phase was separated at 14,500 rpm / 10°/ 4°C. The upper aqueous phase was removed and 500 ul Isopropanol added to it. It was mixed and incubated at room temp for 15° and centrifuged at 14500 rpm /10°/4°C. The pellet was washed with chilled 75% ethanol followed by chilled absolute ethanol and centrifuged at 14,500 rpm / 10°/ 4°C. The supernatant was
removed and the pellet air-dried. The RNA was dissolved in 10 μl of diethylpyrocarbonate (DEPC)-treated water.

cDNA were prepared using Avian Molony Virus Reverse Transcriptase (Promega) according to the manufacturer’s protocol, separately for both NS5 and structural genes; 2 μl Reverse primers were used for each reaction respectively. They were incubated at 65°C for 10 min in a thermal cycler and snapped on ice. Master Mix was prepared by adding 25 mM of MgCl₂, 5x buffer and 10U of AMV and further incubated at 42°C for 45 min and 85°C for 5 min in thermocycler (Mastercycler gradient™, Eppendorf). PCR was performed using 2.5 μl of 10x PCR buffer (Invitrogen), 1.5 μl of 50 mM MgCl₂, 0.5 μl of 10 mM dNTP, 1.25 μl of 10 μM of each primer, 5U of Taq polymerase (Invitrogen), 1.0 μl of cDNA and distilled water up to 25 μl reaction volume. The reaction was kept at denaturation at 94°C for 5 min; followed by 35 cycles of 1 min steps for 94°C, 55°C and 72°C each; and a final 5 min extension at 72°C.

f. Sequencing of PCR products, assembly and curation

The sequencing was done as described earlier using the standard protocol given by ABI Applied Biosystems. The quality of each sequence was monitored using sequence analysis software version 5.1. All the good quality sequences were assembled automatically using Kodon software version 2.1 (Applied Maths, Inc., Austin, Texas, U.S.A.) and curated manually.

These sequences were trimmed in order to ensure better quality results. Following trimming 720 bp of structural region starting from position 500 nt to 1220 nt and 620 bp of NS5 region was used for genotyping (Table 14). Together the sequence of structural and the non-structural region genes (720 bp, 620 bp) comprises of around 12% of the genome. Multiple sequence alignment was performed using clustalX software version 1.8.

Table 14: The sequence information of 3' end of structural genes (encoding preM, membrane and envelope proteins) and non-structural gene (encoding NS5 protein) of KFDV used for the genotyping.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Location</th>
<th>Size</th>
<th>Initial nt</th>
<th>Amino acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreM</td>
<td>500-736</td>
<td>237 bp</td>
<td>ATG GTG ATC AGG GCA GAA</td>
<td>MVIRAE</td>
</tr>
<tr>
<td>Membrane</td>
<td>737-961</td>
<td>225 bp</td>
<td>TCA GTG TCG ATT CCG</td>
<td>SVSIP</td>
</tr>
<tr>
<td>Envelope</td>
<td>962-1220</td>
<td>240 bp</td>
<td>ACA CGA TGC ACA CAC</td>
<td>TRCTHL</td>
</tr>
<tr>
<td>NS5</td>
<td>9422-10063</td>
<td>579 bp</td>
<td>ATA ACC AGA AGA</td>
<td>ITRR</td>
</tr>
</tbody>
</table>

Microevolution and macroevolution
The mean of p distance / nucleotide substitutions was calculated for each year for all the sequences under study. Rate of mutation was calculated after plotting a graph with year of isolation at x axis versus the average nucleotide p distance (NPD) to the y axis; as calculated from MEGA software. Number of synonymous and non-synonymous substitution sites was calculated for the colinearized nucleic acid data for KFDV isolates using model of Nei-Gojobori method (number of differences) implementation in MEGA 3.0 software. Each number of substitutions was plotted with respect to the year of isolation of the virus strain of KFDV.

h. Test of selection

The data showed smaller substitutions therefore the larger sample Z test is too liberal in rejecting the null hypothesis. Fisher exact test was conducted as a test of selection to examine the null hypothesis of the neutral evolution wherein dN = dS was assumed to be Ho. Testing the hypothesis for positive selection (dN > dS)

i. Phylogenetic analysis

A total of 50 isolates were used for colinearized tree construction using both partial NS5 and partial structural region. The data included 1281 sites which are approximately 12% of the whole genome comprising of inframe coding regions. After multiple sequence alignment the tree construction was done using Minimum Evolution method with 1000 bootstraps pseudo replicates at 1 CNI level with initial NJ tree. Kimura 2-parameter was used as substitution model wherein substitutions included both transitions and transversions and the pattern among lineages were homogenous with uniform sites.

j. Impact of low passage levels in tissue culture

In order to study the impact of passages in tissue culture; on viral genome, a pilot study was conducted. KFDV isolate W1930 was propagated in Vero E6 cell line upto 15th passage level. Subsequently the complete envelope gene was sequenced for 0 and 15 passage tissue culture lysates. The sequence obtained were compared and analyzed for mutations upon repeated passages.

3. RESULTS

a. Virus propagation in-vivo

The virus caused demonstrable CPE in Vero E6 cell line between 4 and 7 days depending upon the virus titer (Figure 38).
b. Virus propagation in-vitro

Inoculated mice showed typical symptoms of KFDV infection in 60-72 hrs. This included severe hind limb paralysis varying from hemiplegia to paraplegia with staggering gait. Hemorrhagic manifestations at multiple focal points were prominent on the ventral side of abdomen especially at the site of liver showing enlargement (Figure 39). There was urinary bladder enlargement because of urine retention. Tremors occurred in all mice and gradually they became moribund by 65 to 70 hrs and 100% mortality was seen within 72 to 80 hrs.

Figure 39: Hemorrhage at multiple focal points in KFD infected mice (left). Bladder enlargement because of urine retention (right) was prominent in dissected mice.

c. Phylogenetic studies of strains from patient's serum

Using the nested RT-PCR system on the recent serum specimens from KFD patients, the genotype could be screened successfully. Maximum parsimony method was used as model for construction of phylogenetic tree with 1000 bootstrap pseudo replicates. Phylogenetic data obtained so far shows that at least two genotypes or quasispecies (Q) are circulating in the area (Figure 40). The nucleotide difference between these isolates ranges from 0.6% to 2.8% relative to the KFDV vaccine strain AY323490. However at the level
of amino acid there is no change within all these isolates showing the phenotypic conservation among the isolates. Isolate 72166 shows maximum of 2.6% nucleotide divergence. The NPD data is relative to the KFDV vaccine strain (NCBI Ac #AY323490) (Figure 41)

Figure 40: Phylogenetic analysis of KFD virus isolates based on the sequence obtained from patient’s serum samples using diagnostic sets of primers.

Figure 41: Nucleotide P Distance (NPD) for individual isolates calculated using MEGA 3.1 software.
d. Amplification of Envelope and NS5 gene fragments by RT-PCR

The pair of primers worked well giving bands of desired size when seen on 1.5% agarose gel (Figure 42). Primers pair EF2 and ER2 gave a specific product size of 780 bp of structural region covering partial preM, M and partial Envelope region and 642 bp of partial NS5 gene from location as depicted in Table 15. Both primer pairs were standardized to work on either tissue culture lysate or mice brain suspensions. For amplifications the negative controls were included to negate any cross contamination. These included either tissue culture supernatants or 10% of normal mouse brain suspensions depending upon the propagation of sample in either of the system.

![Figure 42: Amplification of two gene fragments of four representative isolates of KFDV. S- Partial structural region of 780 bp. NS5- partial 642 bp region of NS5. N- Negative control. L-100 bp step Ladder (Promega); intense band is of 500 bp](image)

e. Sequencing results

Well separated peaks were obtained with good quality values (QV) as specified by Applied Biosystems; blue peaks as a measure of quality control (Figure 43) for the sequences used in studies. Sequences with bad QV were discarded and the sequencing was repeated in few cases.
f. Multiple sequence alignment and Phylogenetic analysis

The optimum alignment was achieved without introducing any gap in the alignment (Figure 44).
Figure 44 (contd)

AY323490 ACA GAC ATT GGG ATT TGG GAT GAT TCA ATC ATG TAC GAG TGT GTC ACC ATT GAC TCG GGA GAA GAA CCA GCT GGC GGT

94
Figure 44 (contd)

```
AY323490 AGA AAC AGG AGG TCA GTG TCG ATT CCG GTG CAT OCT CAT AGT GAT CTC ACC GGA AGA GGG CAC AAG TGG CTT AAA

95
```
Figure 44 (contd)
Figure 44 (contd)

AY323490 GTC ACT GGG GAT GAT TGC GTT GTG AGA CCA ATT GAC GAC CGC TTC GGG AAG GCC CTC TAC TTC CTA AAT GAC ATG

601011

601203

611661

612057

623969

62844 T

62849 T

62957 T

63661 T

63696 T

642034 T

642046 T

64244 T

64350 T

651521

652 T

652980 T

6616 T

66364 T

664518 T

66928 T

671004 T

673514 T

67965 T

68142 T

68159 T

681960 T

68484 T

692156 T

692163 T

716810 T

72166

72827 T

A106

G11333

G27667

G27678

P16011

P20924

P21092

P9605

W1930

W3399

W379

W6043 T

W6178

VI6204

712419 T

X74111

AY323490 GCC AAG GTG AGA AAG GAC ATT GGA GAG TGG GAG CCC TCA GAA GGT TAC TCC AGC TGG GAG GAG GTC CCC TTT TCT
Figure 44 (contd)
Figure 44: Multiple sequence alignment of colinearized data of 50 KFDV isolates; using clustalW implementation in MEGA 3.0 software

g. Microevolution and macroevolution

The distance within all isolates was 0.04 with the standard error of 0.01. Substitution pattern homogeneity was also checked. Overall mean % nucleotide similarity is 99.9% within various KFDV isolates. There was 98.8 to 100% nucleotide similarity between the KFDV isolates as compared to the standard vaccine strain P9605. Recent isolates A106 from year 2006 shows 98.8% of nt similarity with the primitive isolates while primitive isolates shows similarity over 99%. The data shows % nucleotide change varies from 0-2% within KFDV isolates (Figure 45). The overall amino acid similarity varies from 99.5 to 100% within the isolates. The primitive isolates show almost 99.8 to 100% amino acid similarity with each other. The recent isolate shows 99.5% amino acid identity as compared to the vaccine strain and 99.3 to 99.5% amino acid similarities with the other isolates.

Figure 45: Nucleotide and amino acid P distance was calculated using MEGA software and plotted for different KFDV isolates used for phylogenetic studies.
The data obtained from isolates from cases of lab infection (P20924, 623969 and 652116) were not the true representative of time of isolation and the recent isolate (72166) because of discontinuity in the time; were not included for time series analysis. The trend line obtained from data is based on the method of least squares shows a straight line trend ($Y = a + bX$). The continuous mutations are increasing over a period of time with an NPD changing at the function of $0.0003x + 0.0015$ (Figure 46). The slope of this trend line (0.00033) represents the rate of change in NPD with respect to first isolate in 1957.

The proportion of synonymous substitutions is over two fold than that of non-synonymous substitutions (Figure 47). Synonymous substitutions ($dS$) are increasingly accumulating over a period of time from 0 to 9 per 427 sites; however the non-synonymous substitutions ($dN$) are almost stable with respect to time ranging from 0 to 2 per 427 sites out of total 1200 bp. This clearly defines the purifying selection over a period of time where $dN < dS$. The rate of nucleotide substitutions was calculated as $8.16 \times 10^{-4}$ substitutions per year per site (including both synonymous and non-synonymous sites).
h. Test of selection and calculation of divergence time

Fisher exact test probability for all the sequences were greater than 0.05 (from 0.50 to 1.00) hence we accept the null hypothesis of positive selection (dN < dS) at 95% confidence intervals.

With the assumption that the rate of nucleotide substitutions within the genome in tick-borne viruses occurs at constant rate, the time of divergence is calculated based on KFDV NPD data. This is in accordance with the analysis done in previous chapter. If, T = Time of divergence, R = Rate of change of NPD, N = total number of bp used for analysis, so as 

\[ R = \frac{0.00033}{1200} \]

\[ = 2.75 \times 10^{-7} \]

Therefore 

\[ T = \frac{(NPD/N)}{R} \]

\[ T (AHFV/KFDV) = \frac{0.075}{0.00033}, T (POWV/KFDV) = \frac{0.316}{0.00033} \]

Hence KFDV must have diverged from AHFV around 250 years ago whereas the whole lineage of KFD and AHFV must have diverged from POWV around 1150 years ago.

i. Construction of phylogenetic tree based on colinearized data

Analyzing data on isolates from various localities of Shimoga district collected till 1973, it is seen that the taxa are clustered together, showing nucleotide divergence of 0 to 0.6% only, within the isolates.

Overall comparison revealed that two isolates - 72166 and A106 - are different quasispecies as compared to the earlier isolates form Shimoga district. The nucleotide divergence was around 1.2%. It is interesting to note that isolate 72166 was recovered from a tick in village Gadgeri in Sirsi in 1972 whereas isolate A106 was from a human case in Mangalore in 2006. This would suggest that the same virus has migrated from Shimoga in 1972 to south Kannada in 2006 (Figure 48).
Figure 48: Tree generated from colinearized data of both partial structural and non-structural region genes of 1281 bp from MEGA analysis using Maximum parsimony. Analysis supported by 1000 bootstrap replicates.
j. Impact of low passage levels in tissue cultures
Cumulative mutations were expected to occur on repeated passages in tissue culture. However, on comparison of KFDV W1930 isolate at 0 and 15 passage levels no difference was detected in the envelope gene.

4. DISCUSSION
a. Phylogenetic analysis of strains from patient’s serum
The strains A834 (Mangalore, 2005), A008 (Chikmagalur, 2006), A1130 and A1132 (Shimoga, 2005) showed more similarity with primitive isolates W1930 (isolated from monkeys in 1959), W6204 and G27678 (isolated from ticks in the same year). Therefore, it is presumed that the primitive strains are still circulating in the same area and probably infecting human beings (Figure 40). Same strains probably survived in multiple hosts. Like the other tick-borne viruses this virus species is able to infect multiple hosts taking advantage of its sustenance and existence over a span of 50 years since its discovery.

b. Phylogenetic studies based on two gene segments
In the present work majority of data has been generated from isolates which were obtained during 1957 to 1971 from Shimoga district. Later on work on this virus was discontinued due to lack of a BSL-3 facility. The localities listed in Table 13 were small hamlets of migrating people. It is evident that a number of mutations have accumulated over a period of time with respect to the first isolate P9605 (Figure 48). It is evident that KFDV evolved slowly as compared to the other flaviviruses. It is however not clear if this low rate of mutation is a function of sustenance into the same host such as ticks which have a long lifespan or it is a function of higher GC content that makes it less vulnerable to mutations or it is attributable to some proof reading function of its replication complex (polymerase and helicase proteins).

KFD epizootic was initially characterized by the spread of the disease to areas contiguous with the original focus of infection. Since the first record, epidemics of KFD have occurred repeatedly in Shimoga district of Karnataka and in its adjoining areas. During the epizootic period of 1964–1965, monkey deaths occurred only within the previously known infected area. During 1965–1966 the epizootic extended towards a
contiguous forest South-East of Sagar town, involving an area covering approximately 30 km². The epizootics, which appeared within the original infected area between October 1966 and September 1969 showed a tendency to move north-west of Sorab town (Figure 49).

During the epizootic season of 1969–1970 and 1970–1971, monkey deaths occurred in Yakshi and Gudavi forests respectively which are located further north-west of Sorab town. During 1971–1972, epizootics continued to occur in the old focus of Sagar and Sorab taluk wherein isolate 72827 (1972) was isolated in Holekoppa, which is similar to the primitive isolates. A new focus appeared in Gadgeri and a nearby valley and KFDV
isolate (72166) was from the same locality which has high similarity with the new isolate A106 isolated from Chikknallur further south to the Gadgeri (Figure 48).

By the end of year 1973, epizootics and epidemics were recognized in several new foci, distant from the original focus (Figure 50). The new foci were Aramanekoppa in Hosanagar taluk of Shimoga district and Kodani in Honnavar taluk of North Kanara district. Epizootics and epidemics were also observed in certain localities of original focus (Barur, Jambai, Maisavi and Padavagodu) where the KFDV activity has been recorded over several years (Sreenivasan et. al., 1986). KFD was restricted to Shimoga districts until 1974, when the disease appeared in a new focal area in Uttar Kannad. During 1975, KFDV spread to Mandagadde area in Thirthahalli taluk, Shimoga district approximately 50 km South-West of the periphery of Aramanekoppa focus.

In the year 1980, another focus appeared in Chikmagalur district and Dakshin Kannada district (Figure 50). Later during 1982 the disease appeared at Patrane area in Beltangady taluk, South Kanara district, approximately 80 km south of the periphery of Mandagadde focus (Sreenivasan et. al., 1986) with a notable outbreak in 1982 reporting 326 cases. In the year 2000, the worst affected areas were Honnavar, Jaida and Six Chapur.
in Uttar Kannad, it was probably the same isolate which is subject to further phylogenetic analysis with representative isolate from that area at the time of epidemic. From the year 2001 to 2006 the same strain established itself at Mangalore as depicted by the phylogenetic analysis, which probably caused subsequent disease spread. Mangalore being a seaport may be the cause of concern in relation to the dispersal of KFDV to other geographical locations.

c. Probable factors for KFD spread to neighboring areas

c(i). Construction of dams

Building up of various dams has been implicated in the spread of disease. During 1962 the Saravathi hydroelectric project decreased the mobility of various infected ticks carrying animals to the river valleys (Banerjee, 1990). This might be one of the reasons of sustenance of KFDV in Shimoga only until 1973. Whereas in the south and southeastern edge, the human and monkey cases increased in 1962 and 1963.

c(ii). Dynamics of forest cover

KFD infection is confined to areas with thick forest covers where the tick activity is prevalent. However, the recent reports have been found outside this cover.
In 1983, the infection occurred in previously undisturbed forests where some 400 ha were clear cut to establish a cashew tree plantation. Most of the human patients were immigrant laborers employed to clear the forest (Varma, 2001).

C(iii). Climatic conditions

Nearly 70 villages in Chikmagalur, Shimoga, Uttar Kannada and Dakshin Kannada were affected with the disease, which are known to thrive under long period of dryness. Spell of dry season from December to June coincides with the nymphal activity in these regions.

C(iv). Vaccination programs

There was no notable outbreak after 1983; therefore the vaccine production in Shimoga was stopped in 1996. However, in 1999 it was started again. In 2000, 2002 and 2003 outbreaks occurred in other parts bypassing the focus area of Shimoga.

Number of villages affected by KFDV and the human cases consequently were maximum in 1974-79. Subsequently, there was a decrease in cases in the following years. This also correlated with a decrease in monkey deaths. This might be attributed to the migration of monkeys or cattle or other domestic animals carrying infected ticks to other places than Shimoga which is considered to be the focus area for infection. Implementation of vaccination programs could also be the probable reason for decrease in cases in subsequent years 2000-2003.

From the overall analysis, it seems that the two strains have diverged geographically i.e. as stated, the primitive isolates have spread towards the northern parts of its centre of origin whereas the recent strains are prevalent in the districts south-west of Shimoga. Finally, it is speculated that KFDV may show an altered degree of virulence due
to many changing factors such as changes in social behavior of humans due to changes in geographical conditions, large scale changes in ecology due to deforestation, or building of dams or canals and changing agricultural practices etc.

It is still not known and is a subject of further studies that so far new quasispecies emerged towards the northern parts which has altered host range or different antigenic profiles. Although there are lesser changes, since the difference at amino acid level is very low but a conclusive statement can not be made based on this data.

However, the probable reason of confinement of disease to this Shimoga for nearly three decades since its discovery could be that area in the outskirts of Shimoga is comparatively more populated by languor monkeys rather than bonnet monkeys. Bonnet monkeys show less mortality as compared to the languor monkeys thus can survive with infection and act as a source of virus.

d. Case report from Gulbarga district, Karnataka

Gulbarga district is situated in north-east of Karnataka state (physiographic region of northern Karnataka plateau) that has quite divergent climatic conditions, flora and fauna. The Northern Karnataka Plateau, at an elevation of 300 to 600 meters, is largely composed of the Deccan Trap and comprises of Belgaum, Bidar and Bijapur districts besides Gulbarga district. It represents a monotonous, treeless extensive plateau landscape rich in black cotton soil.