Mapping of Genes and Identification of Retroviral Element in the *Nasalis larvatus* Genome Using Bioinformatic Approaches

Aswini Leela Loganathan, S. Vijay Kumar, Christopher Lok Yung Voo, Kenneth Francis Rodrigues

*Biotechnology Research Institute, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia*

*Email: kennethr@ums.edu.my*

**ABSTRACT**

*Nasalis larvatus*, also known locally as Proboscis Monkey (*Monyet Belanda* / *Bayau*), are endemic to the island of Borneo. These Old World Monkeys which parted from the human lineage 25 MYA are iconic to Sabah and serve as one of the major tourist attraction in Sabah. This current study focused on characterizing the genes linked to immunity and host defense which are the retroviral elements present in the *Nasalis larvatus* genome. Retroviruses are single stranded RNA viruses that reverse transcribed their RNA into DNA for integration into the host’s genome and have been posited to contribute to evolutionary processes. The aim of this study was to assess the genes related to immunity of different types and identification of retroviral elements in the *Nasalis larvatus* and gather novel data that can benefit in clinical studies and eventually enable effective conservation measures to be taken to protect the *Nasalis larvatus* in its natural habitats. Polymerase chain reaction (PCR) primers were designed from a total of 7 Retroviruses of non-human primate genes obtained from EST. The primers were used to independently amplify and sequence each DNA segment. The primers were designed using the online primer design tool (Primer 3). Amplification reactions were performed in a 96-well microtiter-plate thermal cycler, following which the PCR products were resolved using gel electrophoresis, extracted from the gel, purified, cloned and sequenced. The resulting sequences were analyzed. Alignment was done using ClustalW following which the phylogeny of the genes was resolved using MEGAS. This study highlighted the relatedness of the immunity related genes among the *Nasalis larvatus* and other non-primates and presumed ancestors and the descendants by analysis of homologous characters.

**Keywords:** Proboscis Monkey; Retrovirus; HERV; Evolution; Conserved
INTRODUCTION
The Proboscis monkey (*Nasalis larvatus*) is endemic to Borneo, occurring in Brunei, Indonesia (Kalimantan), Sabah and Sarawak (Groves, 2001). As one of the closest living Old World Monkey to the human, it may hold some unique genes that are beneficial to both human and other non-human primates. The *Nasalis larvatus* populate mangrove forests along river and estuaries (Bootratana, 2000). *Nasalis larvatus* are folivores and frugivores (Yeager, 1989). These primates are protected from hunting and confinement in Borneo. However, the deforestation of mangrove trees and both human and non-human predators have limited the population (Meijaard & Nijman, 2000a). This study will help in deciphering the phylogenetic order of the retroviruses of primates and identify the retroviruses that are highly conserved in *Nasalis larvatus*.

Retroviruses are single-stranded RNA viruses that reverse-transcribe their RNA into DNA for integration into the host’s genome. Generally retroviruses infect somatic cells. However, in some cases exogenous retroviruses include germ line cells infection, thus, allowing it to be passed on to subsequent progeny and thereby becoming ‘endogenous’. They are considered to play a key role in evolution (Miller, 2006). Endogenous retroviruses entered the germ line of their host during evolution and are now passed from parents to offspring like Mendelian genes (Boller et al., 2008). In this study, we examined human endogenous retroviruses (HERVs) which are present in the *Nasalis larvatus* and other primates. Prior to the evolutionary split of Hominoids and Old World Monkeys 25 million years ago (MYA), the HERV-K family is believed to have integrated into the germ line (Reus et al., 2001). HERV sequences are made of about 1% of the human genome. It is assumed that during the course of human evolution, exogenous progenitors of HERVs have inserted themselves into the cells of the germ line where they replicate along with the host’s cellular genes (Lwer et al., 1996).

MATERIALS AND METHODS
Targeted genes of various primates (human and non-human primates) were obtained from NCBI database. We used CYGWIN to convert all the FastA files to MultifastA. Following that we ran a stand-alone blast against the *Nasalis larvatus* using CLC Genomic Workbench. We chose the most relevant and most accurate hit and designed the primer using Primer3 software. Primers were designed to amplify conserved sequences of the primates from the NCBI database. Next, PCR amplification was performed in a BIORAD Thermal Cyclers and subsequently gel extraction was carried out using the commercial kit QIAquick (QIAGEN) following the manufacturer’s protocol. We then carried out cloning using the CloneJET
PCR cloning kit (Thermo Scientific). Plasmid miniprep was then performed using the Gene Jet Plasmid Mini Prep kit (Thermo Scientific).

DNA sequencing was done on an Applied Biosystem genetic analyzer system using the BigDye® Terminator v3.1 cycle sequencing kit chemistry. The DNA sequences were then assembled and analyzed using SeqMan (DNASTAR) software. Alignment was performed using ClustalW program in Mega 5.0. Phylogenetic tree was constructed using Mega 5.0 to study the pattern of relationship of the retroviral genes among the *Nasalis larvatus* and other primates and presumed ancestors and descendant are traced by analysis of homologous characters.

RESULTS AND DISCUSSION

We compiled 19 retrovirus genes from various primates from the NCBI and run a stand-alone blast against the *Nasalis larvatus* genome using CLC Genomic Workbench. Out of the 19 retrovirus genes, seven showed an e-value of 0.00 which is considered to be most relevant hits with high accuracy. Primers were designed from this seven genes using Primer3 software. PCR amplification was performed and it was proved that all seven genes were present in the *Nasalis larvatus* genome. The amplicons were sequenced and assembled. Out of the seven, only four gave the expected result. The alignment from the ClustalW showed that the four retroviruses were highly conserved.

Phylogenetic tree was constructed and analyzed using Neighbour-Joining (NJ) method in MEGA 5.0. Our data was also subjected to bootstrap analysis with 1000 replication to assess the strength of support for any particular clade. We used nine various primates i.e. [*Pan troglodytes* (Chimpanzee), *Pan paniscus* (Bonobo), *Gorilla gorilla* (Gorilla), *Homo sapiens* (Human), *Macaca mulatta* (Rhesus Monkey), *Pongo abelii* (Sumatran Orang Utan), *Pongo pygmaeus* (Bornean orang Utan), *Hylobates* (Gibbon) and *Macaca fuscata* (Japanese Macaque/Snow Monkey)] retroviruses nucleotide data set as a model to construct the tree to see if the retroviruses shared a common ancestry with the *Nasalis larvatus*. From the NJ trees created individually for the 4 retrovirus data, the result shows high consistency of conservation of the retroviruses (Figures 1-4).

Figure 1 and Figure 2, show that the tree topologies obtained from NJ analysis of the *Nasalis larvatus* retrovirus are highly congruent with respect to the phylogenetic position of the *Macaca fuscata*. While in Figure 3, the tree topologies suggests that the *Nasalis larvatus* retrovirus is highly congruent with respect to the phylogenetic position of humans. In Figure 4, The *Nasalis larvatus* retrovirus branches out from the model groups of primates. The Phylogenetic
branches suggest diversity of the retroviruses in *Nasalis larvatus* to human. The four detected retroviruses suggest multiple signs of long-time presence of the retroviruses in the genomes even though the *Nasalis larvatus* split from the human lineage 25 MYA.

**Figure 1** Bootstrapped Neighbour-Joining tree calculated from Kimura 2-parameter assuming that a certain fraction of sites are evolutionarily invariable (+I) based on 1000 replicates constructed using MEGA 5.0 with the data sets of primate nucleotides.

**Figure 2** Bootstrapped Neighbour-Joining tree calculated from Tamura 3-parameter assuming that a certain fraction of sites are evolutionarily invariable (+I) based on 1000 replicates constructed using MEGA 5.0 with the data sets of primate nucleotides.
Figure 3 Bootstrapped Neighbour-Joining tree calculated from Tamura 3-parameter based on 1000 replicates constructed using MEGA 5.0 with the data sets of primate nucleotides.

Figure 4 Bootstrapped Neighbour-Joining tree calculated from Tamura 3-parameter with non-uniformity of evolutionary rates among sites may be modeled by using a discrete Gamma distribution (+G) with 5 rate categories based on 1000 replicates constructed using MEGA 5.0 with the data sets of primate nucleotides.
Phylogenetic analyses of retroviral elements, as well as endogenous retroviruses, express the highly conserved nucleotides genes which are also endowed with conserved imperative functions. This domain is conserved among a large series of retroviral elements, and we have therefore attempted to generate phylogenetic links between retroviral elements identified from *Nasalis larvatus* which is obtained from the databases together with the other primates following the alignments of the nucleotide sequences. This allowed us to unravel a conserved organization among the domains and to identify a large number of human endogenous retroviruses (HERVs) from sequence databases. Benit *et al.* (2001) suggested that these elements were hallmarks of ancient infections of the germ line by retroviruses which have subsequently been “endogenized” and can be used as molecular markers of evolution.

**CONCLUSION**

Human endogenous retroviruses (HERVs) are transmitted perpendicularly through the germ line and are thus inherited by consecutive generations in a Mendelian manner. HERVs have been implicated in certain autoimmune diseases and cancers and might have a role in the aetiology and pathology of disease. However, many HERVs have been present in our genome for a considerable period of time so that their presence may also be of benefit to the human host. This study shows that the HERV is highly conserved in Old World Monkey, i.e. the *Nasalis larvatus*. Based on the data, the retroviruses were probably present in the genomes before the divergence of the Old World Monkeys and Hominoid lineages as suggested by Reus (2001). Overall, the accomplishment of the alignment in this study enables the identification of retrovirus in the *Nasalis larvatus*, an Old World Monkey. We also could identify the similarity and diversity of the retroviruses in the primates.

**ACKNOWLEDGEMENTS**

This project is funded by the Ministry of Science, Technology and Innovation (MOSTI), through the Genomics and Molecular Biology R&D Initiatives Grant under the Malaysia Genome Institute (MGI). It is a collaborative effort by the Malaysian Genome Resource Centre (MGRC), Universiti Malaysia Sabah (UMS) and also Universiti Malaysia Sarawak (UNIMAS).
REFERENCES


