Interspecies transmission of retroviruses in ancient and modern times

（古代および現代におけるレトロウイルスの種間伝播）

Joint Graduate School of Veterinary Medicine

Yamaguchi University

Loai Ghaleb Ahmed AbuEed

September 2026

DECLARATION

September 2026, Yamaguchi University, Japan.

I, Loai AbuEed, hereby declare that the work presented in this dissertation has been carried out independently by the undersigned. I confirm that all sources of information and data from the work of others have been duly acknowledged and cited within the text. This dissertation has not been submitted, in whole or in part, for the award of any other degree or for any other purpose.

ACKNOWLEDGMENT

I want to sincerely thank Prof. Kazuo Nishigaki, my supervisor, for his constant encouragement, guidance, and inspiration during my PhD studies. I sincerely appreciate him allowing me work in his lab and giving me the chance to learn and grow as a researcher under his guidance. His patience, dedication, and profound scientific knowledge have greatly influenced my development in the field of virology, as well as in scientific writing and critical thinking. My confidence and research abilities have grown during the years I have worked in his lab, and I will never forget the knowledge I have gained from him as I follow my goal of becoming a competent scientist.

I want to sincerely thank Assistant Professor Dr. Ariko Miyake for all of her help, support, and helpful guidance during my time in the lab. Especially in the early years of my studies, her commitment and careful guidance helped me to learn essential experimental and organizational abilities. I sincerely appreciate her constant encouragement, which allowed me to carry out my experiments effectively and confidently.

Additionally, I want to sincerely thank all of the current and past colleagues at Yamaguchi University's Laboratory of Molecular Immunology and Infectious Diseases for their generosity, collaboration, and technical support during my studies. I am extremely appreciative of Drs. Didik Pramono, Isaac Makundi, and Junna Kawasaki for their significant support, advice, and assistance during my research.

I am extremely thankful of the Japan International Cooperation Agency (JICA) for giving me the chance to complete my PhD studies and research at Yamaguchi University along with financial support.

Lastly, I want to sincerely thank my parents, parents-in-law, brother, sister, and all of my family members and friends for their unwavering support, encouragement, and prayers during my studies. Above all, I want to express my sincere thanks to my amazing sons, Kenan, Yaman,

and Mohammad, as well as my dear wife, Lamees Makhamre, for their unending love, patience, and sacrifices. I was able to overcome challenges and succeed in my PhD studies because of their continuous support and understanding

TABLE OF CONTENTS

DECLARATION I

ACKNOWLEDGMENT II

TABLE OF CONTENTS IV

GENERAL INTRODUCTION 4

SCOPE OF THIS DISSERTATION 12

1. CHAPTER ONE 8

1.1 ABSTRACT 9

1.2 INTRODUCTION 10

1.3 MATERIALS AND METHODS 12

1.3.1 Ethics Statemen 12

1.3.2 Study Location 12

1.3.3 Sample Collection 12

1.3.4 PCR Amplification and Sequencing 13

1.3.5 Multiple Sequence Alignment and Phylogenetic Analysis 14

1.3.6 Statistical Analyses 14

1.4 RESULTS 14

1.4.1 Prevalence and Demographics 14

1.4.2 FFV Co-Infection with FIV, FeLV, and FcaGHV1 15

1.4.3 Nucleotide Similarity with SU Region for FFV 15

1.4.4 Phylogenetic Analysis of FFV Based on SU Region 16

1.4.5 Nucleotide Sequence Accession Numbers 16

1.5 DISCUSSION 16

1.6 CONCLUSION 20

1.7 TABLES, FIGURES, SUPPLEMENTARY DATA IN CHAPTER ONE 21

2. CHAPTER TWO 31

2.1 ABSTRACT 32

2.2 INTRODUCTION 33

2.3 MATERIALS AND METHODS 34
2.3.1 Cell culture 34
2.3.2 Pseudo-typed virus preparation 34
2.3.3 Infection assay 35
2.3.4 Viral interference assay 36
2.3.5 Receptor screening 36
2.3.6 Establishment of cell lines expressing SLC52A2 and SLC52A1 37
2.3.7 Detection of SLC52A1 and SLC52A2 expression by reverse transcription quantitative PCR 38
2.3.8 Sequence analysis of SLC52A1 and SLC52A2 genes 38
2.3.9 Phylogenetic and sequence analyses 39
2.3.10 Estimation of the integration timing of CERV1 in primates 39
2.3.11 RNA-seq analysis 40
2.3.12 Statistical analysis 40
2.3.13 Ethics approval 40
2.4 RESULTS 41
2.4.1 Identification of CERV1 entry receptors 41
2.4.2 Viral receptor interference of CERV1 and PERV-A 42
2.4.3 CERV1 infectivity in human and primate cell lines 43
2.4.4 Expression of SLC52A1 and SLC52A2 in human and primate cell lines 43
2.4.5 Sequence analysis of SLC52A1 and SLC52A2 in human and primate cell lines 44
2.4.6 Expression of SLC52A1 and SLC52A2 in human and non-human primate tissues 44
2.4.7 Genetic diversity of CERV1 env in primates 45
2.4.8 CERV1 env expression in tissues 46
2.5 DISCUSSION 46
2.6 CONCLUSION 51
2.7 TABLES, FIGURES, SUPPLEMENTARY DATA IN CHAPTER TWO 52
GENERAL DISCUSSIONS AND CONCLUSIONS 138
REFERENCES 141

GENERAL ABSTRACT

Retroviruses are RNA viruses that infect a wide range of vertebrate hosts, including mammals. Their replication cycle is characterized by the reverse transcription of the single-stranded RNA genome into double-stranded DNA, which is subsequently integrated into the host cell's chromosomal DNA. Retroviruses are mainly classified into exogenous retroviruses (exRVs) and endogenous retroviruses (ERVs). exRVs are transmitted horizontally between susceptible hosts and are associated with neoplastic and degenerative diseases. In contrast, ERVs originate from ancestral infections of germ cells and are permanently inherited within host genomes.

Among exogenous retroviruses, foamy viruses (FVs), belonging to the subfamily *Spumaretrovirinae*, are widely distributed across mammals, including cattle, horses, cats, and primates, and have zoonotic potential. FVs are complex retroviruses characterized by terminal repeats and the major genes *gag*, *pol*, and *env*, sharing structural complexity with lentiviruses. Feline foamy virus (FFV) has been reported in domestic cats worldwide, including in Asia, Australia, Germany, Turkey, and the United States. Transmission mainly occurs through salivary exchange and close physical contact, although vertical transmission via lactation has been reported in felids. FFV exists as two genotypes (F17/951-type and FUV-type), which differ in the envelope surface (SU) protein and correspond to separate neutralization serotypes.

ERVs are gradually inactivated by mutations that arise during viral replication and through post-insertional mutations that occur during the replication of the host genome over time. Moreover, some ERVs undergo a process of domestication within their host organisms, thereby gradually acquiring physiological functions. ERV sequences serve as molecular fossils of past infections, enabling the study of paleovirulence and the evolutionary dynamics of host–virus interactions. With the exception of chimpanzee endogenous retrovirus 1 (CERV1)/Pan troglodytes endogenous retrovirus 1 and CERV2, all 40 ERV groups in chimpanzees are represented by lineal homologues in humans. However, although some studies have shed light on the CERV2 infection process, the detailed mechanisms of CERV1 infection remain poorly

understood. This dissertation investigates how retroviruses interact and co-evolve with their hosts, focusing on ancient viral endogenization events and ongoing cross-species transmission.

In Chapter 1, the potential transmission of FFV from domestic cats to the endangered Tsushima leopard cat (TLC), which is an endangered species found only at Tsushima Island in Japan were investigated. Screening of 89 TLC samples identified FFV infection in 7 (7.86%) of the cats, while 14.07% of 199 domestic cats tested positive. Phylogenetic analysis showed that FFV sequences from domestic cats and TLCs clustered within the same clade, indicating shared viral strains between domestic cats and TLCs and suggesting cross-species transmission. Statistical analysis revealed no significant association between FFV infection and sex (p = 0.28). However, in domestic cats, FFV infection was significantly associated with co-infection with feline immunodeficiency virus (FIV) (p = 0.002) and feline gammaherpesvirus 1 (FcaGHV1) (p = 0.0001), but not with feline leukemia virus (FeLV) (p = 0.21). These findings highlight the importance of monitoring FFV infection in both domestic cats and TLC populations as part of conservation and disease management strategies.

In Chapter 2, the transmission and endogenization pathways of CERV1 in primate genomes were examined. Using a human cDNA library, riboflavin transporter SLC52A2 was identified as a functional cellular receptor for CERV1 entry. The ectopic expression of human and chimpanzee SLC52A2, as well as the related SLC52A1, in heterologous cells conferred susceptibility to infection by both CERV1 and porcine endogenous retrovirus (PERV). Viral interference assays demonstrated reciprocal inhibition between CERV1 and PERV, indicating that they belong to the same interference group. Furthermore, CERV1 exhibited a wide host range, infecting various human and primate cell types. Genomic analyses revealed a high prevalence of CERV1 in great apes (*Pan troglodytes*, *Pan paniscus*, and *Gorilla gorilla*) and 15 Old World monkey species, but its absent in both humans and orangutans. These findings suggest that ERVs may contribute to host resistance against viral infections through convergent evolutionary mechanisms, with important implications for paleovirology and xenotransplantation safety.

In conclusion, this study provides comprehensive insights into retrovirus–host interactions, comprising both retroviral endogenization events and ongoing cross-species transmission. The findings on CERV1 enhance our understanding of the evolutionary integration and adaptation of retroviruses within primate genomes, while the FFV results demonstrate active viral exchange between domestic and wild felids. Collectively, this work contributes to a deeper understanding of retroviral evolution, host adaptation, and the ecological and evolutionary dynamics of host–virus relationships.

GENERAL INTRODUCTION

Retroviruses, classified within the family *Retroviridae*, represent a wide-ranging and varied group of viruses that infect all vertebrate organisms (1). They are characterized by a distinctive replication mechanism that includes the reverse transcription of genomic RNA into DNA, followed by its permanent integration into the host genome (2–4). A typical retroviral particle shows a spherical morphology, approximately 100 nm in diameter, and is surrounded within a lipid envelope derived from the host, which is coated with trimeric envelope glycoproteins that facilitate binding and fusion of the viral envelope with the host cell membrane (5). Underneath the envelope exists a matrix protein layer that encases an icosahedral or conical capsid, which contains two identical copies of positive-sense RNA along with the viral enzymes reverse transcriptase, integrase, and protease (6). Following infection, the genomic RNA is reverse transcribed, and the resultant DNA is integrated into the host chromosome as a provirus. This proviral form is flanked by long terminal repeats (LTRs) containing transcription initiation sites and regulatory cis-elements that control promoter activity, integration specificity, and viral transcriptional tropism (5). Historically, variations in capsid morphology and core structure were fundamental characteristics employed to differentiate retroviral genera (5).

The International Committee on Taxonomy of Viruses (ICTV) classifies retroviruses into various genera according to their genomic structure and biological characteristics (1). *Alpharetroviruses* and *gammaretroviruses* are categorized as simple retroviruses, while *deltaretroviruses*, *epsilonretroviruses*, *lentiviruses*, and *spumaviruses* are classified as complex retroviruses. *Betaretroviruses* exhibit features that are intermediate between these categories (6). Simple retroviruses encode only the canonical viral proteins—gag, pro, pol, and env. The gag polyprotein forms the internal structural proteins, containing matrix (MA), p12, capsid (CA), and nucleocapsid (NC) (7). The pro region encodes the viral protease, whereas pol encodes the reverse transcriptase and integrase enzymes. The env gene encodes the surface (SU) and transmembrane (TM) domains of the viral envelope glycoprotein (8). Conversely, complex retroviruses not only encode core proteins but also have several accessory and

regulatory genes that influence viral replication, immunological evasion, and host–virus interactions (5).

Foamy viruses (FVs), classified within the *Retroviridae* family and *Spumaretrovirinae* subfamily, are broadly found in cattle, horses, cats, and non-human primates (9). These viruses are classified as complex retroviruses and share certain structural and genomic features with lentiviruses, including the presence of LTRs and the canonical gag, pol, and env genes (10). Although humans are not natural hosts, zoonotic transmission from non-human primates has been documented, indicating a broad host range and cellular tropism (11). The Feline foamy virus (FFV), initially discovered in 1969 from a domestic cat in the United States, has since been identified in both healthy and afflicted cats across Asia, Australia, Europe, and North America (11–16). Experimental infections in domestic cats generally produce no apparent clinical signs, and FFV is therefore widely considered apathogenic (17). Notably, two distinct FFV serotypes FUV7-Type and 951-Type have been identified among Australian and American isolates, with each serotype forming a distinct env surface (SU) protein sequence cluster and corresponding neutralization group (16). Transmission of FVs generally occurs by salivary shedding and close contact, however vertical transmission via lactating cats has also been documented (17, 18). In domestic cats, virus transmission occurs via biting and social interactions like grooming (19), whereas in simian hosts, biting and predation are primary transmission pathways (20).

Endogenous retroviruses (ERVs) are inheritable remnants of earliest infections by exogenous retroviruses, found in the genomes of all vertebrate species (5). Through recurrent germline reinfection events, this vertical mode of transmission has contributed to the substantial accumulation of ERV-derived sequences within mammalian genomes (21). ERVs contain the common retroviral genomic structure, including of the basic genes gag, pro, pol, and env, located between two LTRs (22). However, the majority of ERVs have lost their ability to produce infectious virus due to mutations and deletions acquired during host genome replication (23). Despite this loss of infectivity, some ERVs have been functionally integrated

by the host and now serve beneficial biological functions (24). For instance, certain Env-derived proteins can inhibit infection by related retroviruses in species such as chickens, cats, mice, and humans (25–30). In addition to their functional roles, ERVs serve as genetic remnants of historical retroviral invasions and provide a basis for examining paleovirulence and the evolutionary patterns of ancient, and occasionally extinct, retroviruses (31). *Pan troglodytes* endogenous retroviruses (PtERVs) have a prevalent lineage of LTR retrotransposons within vertebrate genomes, categorized into three primary retroviral classes—class I (*gammaretroviruses*), class II (*betaretroviruses*), and class III (*spumaviruses*)—and further divided into 42 families, including two that are exclusive to chimpanzees (32, 33). Although most PtERV families are shared with humans and Asian apes, the CERV1 and CERV2 families are unique to African great apes and Old World monkeys (32). While the infection dynamics of CERV2 have been partially clarified (34, 35), the specific mechanisms of CERV1 infection remain inadequately understood.

This dissertation examines retroviral transmission dynamics by evaluating FFV prevalence in Tsushima leopard cats and domestic cats, and also investigating the historical transmission and endogenization of CERV1 in chimpanzees.

SCOPE OF THIS DISSERTATION

Chapter one:

To evaluate FFV transmission risk by examining infection rates in Tsushima leopard cats and domestic cats and assessing management systems designed to protect TLCs from pathogens carried by domestic cats.

Chapter two:

To demonstrate how CERV1, an extinct gammaretrovirus in chimpanzees, was originally transmitted and endogenized.

1. CHAPTER ONE

Feline Foamy Virus Transmission in Tsushima Leopard Cats *(Prionailurus bengalensis euptilurus)* on Tsushima Island, Japan.

This work has been published as follows:

AbuEed L, Makundi I, Miyake A, Kawasaki J, Minoura C, Koshida Y, Nishigaki K.2023. Feline Foamy Virus Transmission in Tsushima Leopard Cats (*Prionailurus bengalensis euptilurus*) on Tsushima Island, Japan. *Viruses* 15(4), 835. https://doi.org/ 10.3390/v15040835

1.1 ABSTRACT

Tsushima leopard cats (TLC; *Prionailurus bengalensis euptilurus*) only inhabit Tsushima Island, Nagasaki, Japan and are critically endangered and threatened by infectious diseases. The feline foamy virus (FFV) is widely endemic in domestic cats. Therefore, its transmission from domestic cats to TLCs may threaten the TLC population. Thus, this study aimed to assess the possibility that domestic cats could transmit FFV to TLCs. Eighty-nine TLC samples were screened, and FFV was identified in seven (7.86%). To assess the FFV infection status of domestic cats, 199 domestic cats were screened; 14.07% were infected. The phylogenetic analysis revealed that the FFV partial sequence from domestic cats and TLC sequences clustered in one clade, suggesting that the two populations share the same strain. The statistical data minimally supported the association between increased infection rate and sex ($p = 0.28$), indicating that FFV transmission is not sex dependent. In domestic cats, a significant difference was observed in FFV detection in feline immunodeficiency virus ($p = 0.002$) and gammaherpesvirus1 infection statuses ($p = 0.0001$) but not in feline leukemia virus infection status ($p = 0.21$). Monitoring FFV infection in domestic cats and TLC populations is highly recommended as part of TLC surveillance and management strategies.

1.2 INTRODUCTION

The Tsushima leopard cat (TLC), also known as *Prionailurus bengalensis euptilurus,* is one of the most endangered animal species in Japan (36). The TLC has a weight of 4.5 kg for males and 3.5 kg for females, making it a small felid. A typical cat has a long and fat tail, a long trunk, shorter-than-typical legs, and its color ranges from chestnut brown to cream with faint brown spots (37). It only inhabits Tsushima Island in the Nagasaki Prefecture (38). The TLC tends to be abundant in the upper region of Tsushima, known as Kamijima. Approximately 200 to 300 TLCs lived on Tsushima Island until the 1970s (39). According to the Japan Wildlife Research Center, the number of TLCs continues to decrease at approximately 10% every 10 years. Human activity in wildcat habitats is one of the biggest threats to their survival (40). Another important factor is infectious diseases. Disease transmission between domestic cats and TLCs have been previously documented (39, 41–43). Several cases of TLCs being infected with diseases commonly associated with domestic cats, including the feline immunodeficiency virus (FIV) and *Felis catus* gammaherpesvirus1 (FcaGHV1), have been reported (38, 39, 41). According to a study that utilized the Geographical Information System (GIS) data, FIV infection risk in TLCs is higher in areas densely populated with domestic cats than in areas sparsely populated with the species (36). A previous study reported the prevalence of feline leukemia virus (FeLV) infection in domestic cats but not in TLCs; however, it was demonstrated that the virus could replicate in their cells (38). A study reported the detection of the feline foamy virus (FFV) infection in Iriomote leopard cats (*Prionailurus bengalensis iriomotensis*) in Japan (44). A single FFV infection has been shown to be associated with potential pathological changes in domestic cats (17). A previous study reported histopathological changes in the kidney and lung after experimental FFV inoculation (45). Another study found that cats with FFV infection had a higher risk of

developing lymphoma compared to FFV-negative cats (46). However, coinfections with other viruses have been documented (47–49).

Foamy viruses (FVs) belong to the Retroviridae family, subfamily Spumaretrovirinae; they are present in cattle, horses, cats, and monkeys and are zoonotic (9, 50–53). FFVs have infected domestic cats globally, including in Asian countries (16, 54), Australia (15), Germany (14), Turkey (13), and the USA (12). FVs are complex retroviruses, similar to the lentiviral human immunodeficiency virus (HIV), and contain two terminal repeats and three main retroviral genes, *gag*, *pol*, and *env* (10, 55). All retroviruses, including FVs, replicate by first converting their RNA genome into a DNA intermediary that integrates into the host genome (19). FFV can be divided into two distinct sequence groups (F17/951-Type and FUV-Type) in the *env* surface (SU) protein, each belonging to one neutralization group (serotype) (16). Salivary shedding and ongoing contact between animals are speculated to be the primary routes of FV transmission, although vertical transmission through lactating cats has also been reported (17, 18). Cats may transmit the virus through bites and social contact, such as grooming (19). Furthermore, biting and predation are major sources of transmission among the simian populations (20, 56). Cats in the USA have been found to have FFV prevalence rates ranging from 10 to 75%, with age and male sex as the risk factors (57).

The biodiversity, evolutionary dynamics, and ecology of FFV are poorly understood. Given that controlling disease in wild populations is difficult, monitoring for infectious diseases among TLC populations is an essential part of surveillance and management (41). During the breeding season, TLCs usually have a large home range; therefore, they may contact free-roaming domestic cats. FIV prevalence among domestic cats is highest in Tsushima Island compared to other regions of Japan (36). Most FcaGHV1 and FIV infections spread through territorial aggression and fighting according to previous epidemiological data (41). Based on the correlation between FcaGHV1 and FIV and the identification of FIV in TLCs (36), there is a possibility that domestic cats could transmit FFV to TLCs.

To test this hypothesis, the FFV infection rate in TLCs and domestic cats was investigated, and the system to manage and protect TLCs from infectious diseases from domestic cats was evaluated. FFV infection was more frequent in domestic cats than in TLCs. The TLC sequences were identical to some domestic cat sequences over a 600-bp region of the virus. The phylogenetic analysis suggested that FFV transmission may occur between TLCs and domestic cats in Tsushima. Furthermore, monitoring infectious diseases in domestic cats will promote the conservative management of wild cats.

1.3 MATERIALS AND METHODS

1.3.1 Ethics Statemen

This study was approved by the Institutional Animal Care and Use Committee of Yamaguchi University (identification code 2017/315). Animal studies were conducted in accordance with the guidelines for the Care and Use of Laboratory Animals of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

1.3.2 Study Location

Tsushima (34°05′–34°42′ N, 129°10′–129°30′ E) is an island in the Japanese archipelago with a surface area of 708.6 km^2. It lies 50 km from Busan, the Korean peninsula, and 138 km from Kyushu Island, Japan. There are two main islands on Tsushima Island: the northern and southern Tsushima islands called Kamijima and Shimojima, respectively. There are four boroughs in Kamijima (Kamitsushima, Kamiagata, Mine, and Toyotama) and two in Shimojima (Mitsushima and Izuhara) (Figure 1.1).

1.3.3 Sample Collection

In total, 89 TLC (60 blood and 29 spleen) samples were obtained from Tsushima Island (Figure 1.1 B) between 1999 and 2020 (38). TLC samples were collected by the Kyushu Regional Environment Office (Tsushima Wildlife Conservation Center, Tsushima, Nagasaki, Japan), Ministry of Environment, Japan. In most cases, the cats were killed by

vehicles. The Tsushima Animal Medical Center, a non-profit animal hospital on the island, donated 199 domestic cat blood samples (38); the domestic cats brought to the center between 2009 and 2015 included both indoor-only and free-roaming cats. Blood samples and spleens were stored at −80 °C, and working samples were stored at −30 °C. Figure 1.1 A shows the sample collection site for domestic cats on Tsushima Island.

Blood samples were screened for FeLV and FIV infections using the SNAP FeLV/FIV Combo Kit (IDEXX Laboratories Inc., Westbrook, ME, USA). Total DNA was extracted from whole blood samples using the Dr. GenTLE System (Takara Bio Inc., Kyoto, Japan) and DNAzol reagent (Life Technologies Japan, Tokyo, Japan) according to the manufacturer's instructions. DNA was extracted from the spleen using a commercial kit (DNeasy® Blood & Tissue Kit; QIAGEN, Hilden, Germany). Conventional PCR for feline glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was conducted as previously described to confirm the presence of amplifiable template DNA (58).

1.3.4 PCR Amplification and Sequencing

The nested-PCR amplification method for FFV was conducted as described previously (47) using the primers listed in Table 1.1. Primers were synthesized by Fasmac Co., Ltd., Kanagawa, Japan and dissolved in ultra-pure water (PCR grade) to a concentration of 100.0 pmol/μL, from which working dilutions were made to a final concentration of 10.0 pmol/μL and stored at −30 °C until use.

The total volume of the reaction mixture used for each nested PCR was 24.0 μL. The nested PCR reaction mixture contained 12.5 μL of KOD One ™ PCR master Mix (TOYOBO Bio Inc., Osaka, Japan), 8.5 μL of ultra-pure water (PCR grade), and 1.5 μL of each of the forward and reverse primer working stock. In the first PCR round (nested PCR 1), 1.0 μL template DNA was added, spun quickly, and placed on a thermocycler with 30 cycles of denaturation at 98 °C for 10 s, annealing at 63 °C for 5 s, and extension at 68 °C for 8 s. For the second PCR round (nested PCR 2), 1 μL of the PCR 1 product was added, quickly spun, and placed in a thermocycler with 30 cycles of denaturation at 98 °C for 10 s,

annealing at 63 °C for 5 s, and extension at 68 °C for 1 s. PCR amplification products were further identified using electrophoresis with 1% agarose gels and purified using a FastGene Gel/PCR extraction kit (Nippon Genetics Co., Ltd., Tokyo, Japan). DNA sequencing by Fasmac Co., Ltd., Kanagawa, Japan, was performed on the second PCR product in both directions. UGENE 45.0 software was used to visualize and analyze the sequences, and GeneTyx (Software Development Co., Tokyo, Japan) and NCBI BLAST programs were used to compare the 600 bp SU region sequences with other FFV SU region sequences.

1.3.5 Multiple Sequence Alignment and Phylogenetic Analysis

The SU region nucleotide sequences for the samples and international FFV isolates, which were retrieved from GenBank, were aligned using ClustalW (59). Phylogenetic analyses were performed using two different software packages. The most suitable method for phylogenetic analysis at the nucleic acid level was determined using the jModel-test (60). MEGA6 was used for phylogenetic analysis. The phylogenetic tree was constructed using the maximum likelihood (Kimura 2-parameter) method for nucleic acids (59). A total of 1000 bootstrap replicates were used to construct the phylogenetic tree.

1.3.6 Statistical Analyses

All statistical analyses were performed using the Minitab Statistical program (Minitab version 18, Minitab Inc., Shanghai, China, 2018). The odds ratios (ORs), 95% confidence intervals (95% CIs), and *p*-values were calculated to analyze the risk factors for FFV and FFV co-infection; p values < 0.05 were considered statistically significant.

1.4 RESULTS

1.4.1 Prevalence and Demographics

The creation of amplicons relies on the primer that amplifies the sequence of the partial *env* gene (SU region). Therefore, nested PCR primers were obtained from a previous

study (47). The nested PCR method successfully amplified the SU region in seven TLC samples indicating an overall FFV detection frequency of 7.86% in TLCs (Table 1.2). These seven FFV-positive TLCs originated from Kamijima and comprised four males and three females.

Given that domestic cats are the primary natural hosts of FFV, their FFV infection status was investigated to better understand TLC management. In total, 28 of the 199 (14.07%) domestic cats were positive for FFV according to the nested PCR test (Table 1.2). The characteristics of the domestic cats tested in this study are presented in Table 1.3. No significant difference was observed in the FFV-positive sample (FFV infection rate) between Kamijima and Shimojima (Table 1.4). Of the 35 FFV-positive samples for domestic cats and TLCs, 29 (82.85%) were obtained from Kamitsushima and Kamiagata (Figure 1.1A,B).

1.4.2 FFV Co-Infection with FIV, FeLV, and FcaGHV1

To study the dynamics of retrovirus infection in TLCs, FFV co-infection with FeLV, FIV, and FcaGHV1 in the 199 domestic cats, which had previously been tested for FeLV, FIV, and FcaGHV1, was investigated (Table 1.3). FIV and FcaGHV1 infection statuses were significantly associated with FFV detection (Table 1.4). In contrast, FeLV infection status was not associated with FFV detection. Furthermore, no sex difference was observed for single infections of FFV and FeLV in the domestic cat samples (Table 1.4 and Table 1.5). FIV and FcaGHV1 had higher infection rates in males than females (Table 1.5). In contrast, the FFV and FeLV coinfection rates in males were not significantly different from those in females.

1.4.3 Nucleotide Similarity with SU Region for FFV

In total, 35 FFV-positive samples (7 TLCs and 28 domestic cats) were sequenced; however, two domestic cat samples were lost. Nucleotide sequences of the partial *env* gene (SU region) were determined. Two distinct *env* genotypes (F17/951-Type and FUV-Type) were

identified with a nucleotide similarity ranging from 98 to 100% and 82 to 85% within and between genotypes, respectively (Figure 1.2). A total of 12 mutations (5 non-synonymous and 7 synonymous) were identified in 15 samples (F17/951-Type) (Figure 1.2A), and 12 (4 non-synonymous and 8 synonymous) in 18 samples (FUV-Type) (Figure 1.2B). The FUV and F17/951 types are distributed along the Tsushima Island (Figure 1.3). The nucleotide differences may appear as a logical result because the amplification target in this study was the SU region, which is considered a hypervariable region.

1.4.4 Phylogenetic Analysis of FFV Based on SU Region

A total of 33 FFV sequences and 51 international FFV sequences retrieved from the GenBank database were used to construct a phylogenetic tree based on nucleic acids (Figure 1.4). Phylogenetic analysis demonstrated that 15 FFV-positive domestic cat samples belonged to the F17-Type and 18 samples (11 domestic cat and 7 TLC samples) belonged to the FUV-Type. Evidence of FFV transmission on Tsushima Island is strong, although some subclades indicate region-specific viral populations. FFV isolates from domestic cats were interspersed throughout the tree; in contrast, TLC FFV isolates were confined to one clade.

1.4.5 Nucleotide Sequence Accession Numbers

The partial FFV nucleotide sequences obtained in this study were deposited in the DDBJ, GenBank, and EMBL databases under accession numbers LC747016–LC747048.

1.5 DISCUSSION

This study is the first molecular survey and genetic analysis of FFV in TLCs (*Prionailurus bengalensis euptilurus*). FFV detection in felids other than domestic cats, which are its natural hosts, suggests that FFV can be transmitted to other felines. There is no “gold standard” diagnostic test for FFV, and most previous studies have used qPCR assays. FFV

may not be detected in spleen samples in some cases, even if other tissues or fluids (such as blood or saliva) test positive, depending on the stage of infection (17). According to a previous study, enzyme-linked immunoassay (ELISA) and PCR tests had similar sensitivities; however, ELISA had a higher specificity (9). As a result, it is possible that the assays in this study do not detect all FFV positive cases in either the domestic cats or TLC populations.

Previous studies have demonstrated FIV and FcaGHV1 transmission in TLC (36, 41). The phylogenetic analysis of retrovirus-positive individuals supports the transmission of FIV and FcaGHV1 from domestic cats to TLC on Tsushima Island. Therefore, FFV was included in this survey because TLCs only inhabit this region, and FFV transmission from domestic cats to TLCs may affect management strategies for protecting this endangered species.

The prevalence of FFV in domestic cats and TLCs in Tsushima was 21.96% (domestic cats, 14.07%; TLCs, 7.89%), similar to the previously reported prevalence, which ranged from 8% to 80%, based on geographic location (9, 14). Thus, domestic cats are more likely to be infected with FFV than TLCs, and FFV may be a cross-species transmission infection. Furthermore, FFV transmission may be possible in other cat species. No significant difference was observed in the FFV infection rate between Kamijima and Shimojima (ORs, 1.01; 95% CI, 0.38–2.67). The identification of FFV-positive samples from Kamijima and Shimojima indicated that FFV is widely spread across free-ranging populations. Although no FFV-positive TLC was detected in Shimojima, FFV infection of TLC may spread there in the future as it has been confirmed to be present in the area in recent years (Figure 1.1B). FFV has been suggested to be pathogenic through co-infection with retroviruses (18). FFV co-infection with FeLV or FIV has been reported to increase FeLV proviral load and detection rate (61, 62); previous studies have shown that non-regenerative anemia and T-cell lymphoma may occur (47). Therefore, the relationship between FFV infection and other infections (FeLV, FIV, and FcaGHV1) were analyzed. FIV- and FcaGHV-positive individuals had a higher FFV infection rate than the negative individuals. In contrast,

FeLV-positive individuals had a lower FFV infection rate than negative individuals (OR, 0.39; 95% CI, 0.08–1.74). Therefore, FFV may spread horizontally by bites, similar to FIV and FcaGHV1. Among the 199 domestic cat samples tested, 183 were strays or cats housed outdoors, whereas 16 were completely housed cats, suggesting that the population bias affected the study outcome. Most sampled cats died from traffic accidents and were discovered by chance; therefore, the relationship between the co-infection of these viruses and pathogenicity was unknown. In addition, these infectious diseases may have caused the death of individuals living in isolated environments.

Previous studies have demonstrated that the male sex is a significant risk factor for FcaGHV1 infection (58, 63, 64). In addition, the FIV infection rate in males has been reported to be higher than that in females (41). The statistical data from this study minimally supported the association between the increased single infection rate of FFV and FeLV and sex, indicating that FFV is readily transmitted between sexes. Additionally, previous reports have indicated that sex may not be a risk factor for FFV infection (19). FIV and FcaGHV1 had higher infection rates in males, consistent with previous findings (65). Regarding the sex difference in co-infection rates between FFV and other retroviruses, males were more likely to have a co-infection with FFV and FIV than females (OR, 3.30; 95% CI, 1.29–8.43). Furthermore, the co-infection rate of FFV and FcaGHV1 was higher in males (OR, 6.02; 95% CI, 1.62–22.34), suggesting that horizontal transmission by bite is the main route of infection in domestic male cats, similar to FIV and FcaGHV1, whereas FFV may be transmitted in domestic female cats via a different infection route (vertical transmission, etc.). No significant sex-related difference was observed in the co-infection rate between FFV and FeLV (OR, 1.26; 95% CI, 0.07–20.50). Therefore, it is unlikely that FFV would be transmitted through a prolonged contact route.

According to phylogenetic analysis, TLCs and domestic cats are hosts for the same virus. Phylogenetic analysis demonstrated that all TLC isolates belonged to one clade; however, domestic cat isolates clustered throughout the phylogenetic tree, basal to or within the TLC dominant clade (Figure 1.4). Additionally, F17/951-Type spreads between domestic cats,

while FUV-Type infects both domestic cats and TLCs. A previous study reported that domestic cats are the main host of FFV (61), and that there is a high possibility of FFV transmission from domestic cats to TLCs. Previous studies have demonstrated the viral transmission of FIV and FcaGHV1 in TLCs (36, 41). Phylogenetic analysis of these retrovirus-positive individuals supports the interspecific transmission of FIV and FcaGHV1 from domestic cats to TLCs (36). According to recent reports, wild cats are increasingly preying on domestic animals near populated areas, suggesting an increase in the viral spillover from domestic cats to wild cats (66, 67).

Phylogenetic analysis of the SU region indicated two distinct clades denoting the two SU *env* genotypes (F17-Type and FUV-Type; Figure 1.4). Nucleotide sequence similarity within and between genotypes ranged from 98 to 100% and 83 to 85%, respectively. Similar mutations between the isolates (LC747042 and LC747037) support our suggestion that the FFV virus is transmitted between the domestic cat and the TLC. In addition, random mutations have been found in some isolates (such as LC747041) that may occur after the infection of cats. Both SU *env* genotypes were isolated from domestic cats and obtained from Kamijima and Shimojima. This suggests a synergistic interaction between both genotypes, which are adapted to domestic cats and regularly co-circulate. Both genotypes were highly prevalent in domestic cats, suggesting that they are well-adapted despite the large quantity of amino acid variability in this SU region. Recent studies have reported two unique *env* SFV subtypes that are consistent with this observation in non-human primates (68, 69). In domestic cats, serotype-specific PCR, sequencing, and neutralizing assays have been used to identify these two genotypes (15). In addition, a previous study reported the existence of the F17-Type in Iriomote cats on the Yaeyama Islands of Japan (70). However, there are no available data about the existence of the FUV-Type in Japan; therefore, this study is the first to report the existence of the FUV-Type on Tsushima Island. Both SU *env* genotypes were isolated from domestic cats, and one genotype (FUV-Type) was isolated from TLCs, supporting the speculation that the virus was moved to Tsushima.

Additionally, further research on the transmission routes and co-infections is required for the revision of this management regime.

1.6 CONCLUSION

The study findings demonstrated that TLCs were infected with FFV, in addition to previously reported diseases, which may be originating from domestic cats. Furthermore, the infection rate of FFV in TLCs was approximately 7.86% indicating that it was more widespread than FcaGHV1. The regional distribution of positive samples, co-infection trends, and previous findings support the hypothesis that FFV may be transmitted interspecies between domestic cats and TLCs. Consequently, domestic cats and TLCs may have a closer relationship than previously speculated, posing a new challenge for the management and protection of endangered TLCs.

Considering the infectious diseases that originate from domestic cats, based on previous research and the FFV infection rate in TLCs in this study, various measures are necessary for TLC conservation. Specifically, to minimize the interaction between TLCs and domestic cats, stray domestic cats in Tsushima should be eliminated, and households should be advised to keep their cats indoors and ensure their thorough vaccination. In addition to regular testing for FeLV and FIV, testing for FFV infection should be added as a new index to investigate the spread of infectious diseases.

1.7 TABLES, FIGURES, SUPPLEMENTARY DATA IN CHAPTER ONE

Table

Table 1.1 Primer used in nested-PCR amplification. The expected amplicon size after the second PCR is 600 nucleotides.

Target Gene	Name	Sense	Primer Sequence (5′-3′)	Reference
envelope	env-f2	1st forword	GCTACTTCTACTAGAATAATGTTTTGGATA	(47)
	env-r2	1st reverse	AGCCACAGTAGTAATTGCATGGCCAGGCC	(47)
	env-f3	2nd forward	GCTTTCAAAAATATGGACATTGTTATGTTA	(47)
	env-r3	2nd reverse	GTTTCTCCAAAATCTGCAAGCATATGGATG	(47)

Table 1.2 FFV detected in feline DNA sample by nested PCR.

Host Species	No. of Samples	No. of FFV Positive	% Positive
Tsushima leopard cat	89	7	7.86
Domestic cat	199	28	14.07

Table 1.3 Prevalence of FFV among domestic cat samples

		FFV Status			
Variable	**Categories**	**Positive**	**Negative**	**Total**	**% Positive**
Sex	Male	15	73	88	17.04
	Female	13	98	111	11.71
Location	Kamijima	22	134	156	14.10
	Shimojima	6	37	43	13.95
FIV [1]	Positive	23	61	84	27.38
	Negative	12	103	115	10.43
FeLV [2]	Positive	2	22	24	8.33
	Negative	33	142	175	18.85
FcaGHV1 [1]	Positive	14	11	25	56
	Negative	21	153	174	12.06

[1] Cited from reference number (42); [2] Cited from reference number (37).

Table 1.4 Univariable logistic regression analysis to evaluate the association between variables and FFV detection in domestic cats

Variables	Categories	Z Statistic	Odds Ratio	95% CI [1]	*p* Value
Location	Kamijima vs. Shimojima	0.02	1.01	0.38–2.67	0.98
Sex	Male vs. female	1.06	1.54	0.69–3.45	0.284
FIV	Positive vs. negative	3	3.23	1.50–6.96	0.002
FeLV	Positive vs. negative	1.22	0.391	0.08–1.74	0.21
FcaGHV1	Positive vs. negative	4.78	9.27	3.72–23.08	0.0001

[1] Confidence intervals; CI, Confidence interval.

Table 1.5 Evaluation of the association between sex and another feline virus detection in domestic cats.

Variable	No. of Positive Samples	Male	Female	% Positive	Z Statistic	Odds Ratio	95% CI [1]	*p* Value
FIV	84	51	33	42.2	3.94	3.25	1.81–5.86	0.0001
FeLV	24	12	12	12	0.60	1.30	0.55–3.06	0.54
FcaGHV1	25	17	8	12.2	2.47	3.08	1.26–7.53	0.013
FeLV + FFV	2	1	1	1	0.16	1.26	0.07–20.50	0.86
FIV + FFV	23	16	7	11.6	2.49	3.30	1.29–8.43	0.012
FcaGHV1 + FFV	14	11	3	7.03	2.68	6.02	1.62–22.34	0.007

[1] Confidence intervals; CI, Confidence interval.

Figures

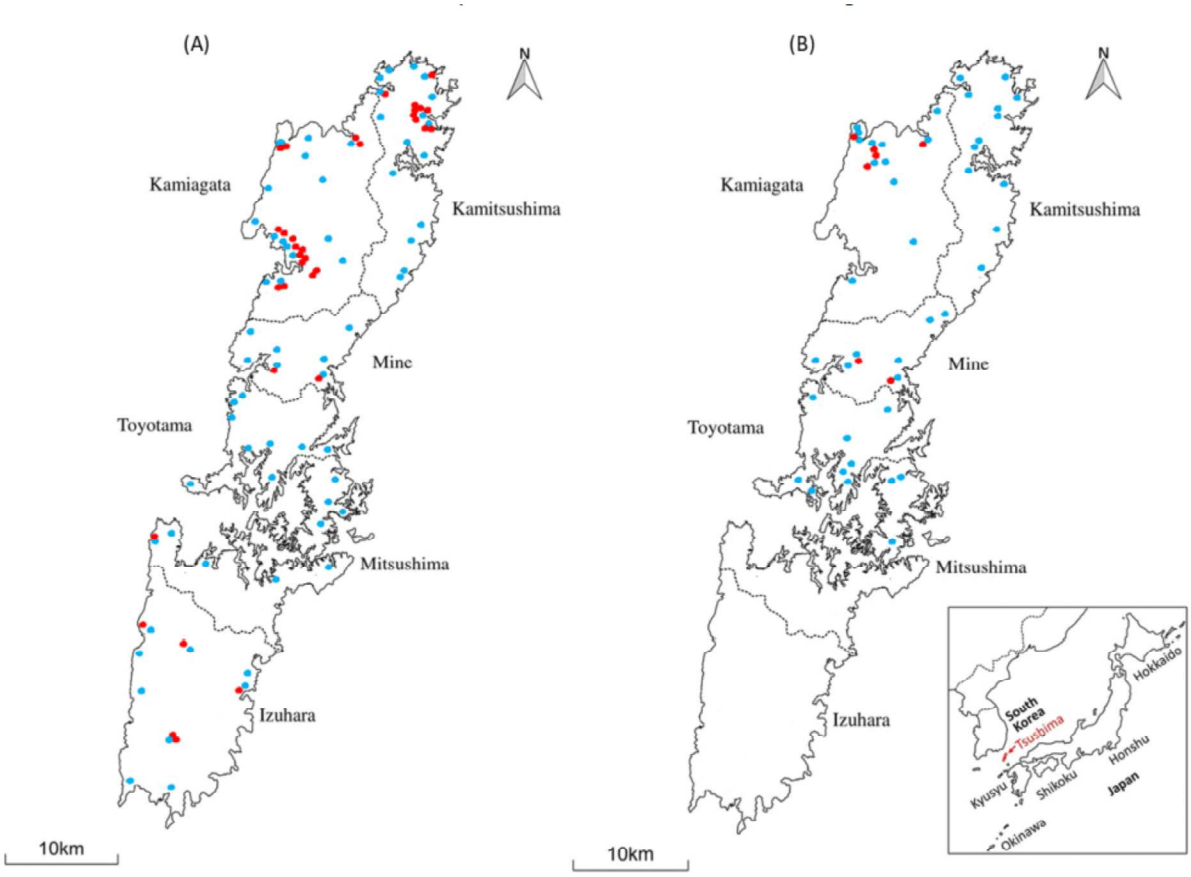


Figure 1.1 Maps of Tsushima Island and sample collection sites. (A) Map of sample collection sites for domestic cats. (B) Map of sample collection sites for wild cats. Blue dots indicate sample locations. Red dots represent positive samples.



Figure 1.2 Comparison of the nucleic acid sequence of the SU region. (A) Multiple sequence alignment of the SU region for F17/951-Type. (B) Multiple sequence alignment of the SU region for FUV-Type. Dots indicate identical consequences. Black and red color labels indicate domestic cat and TLC samples, respectively.

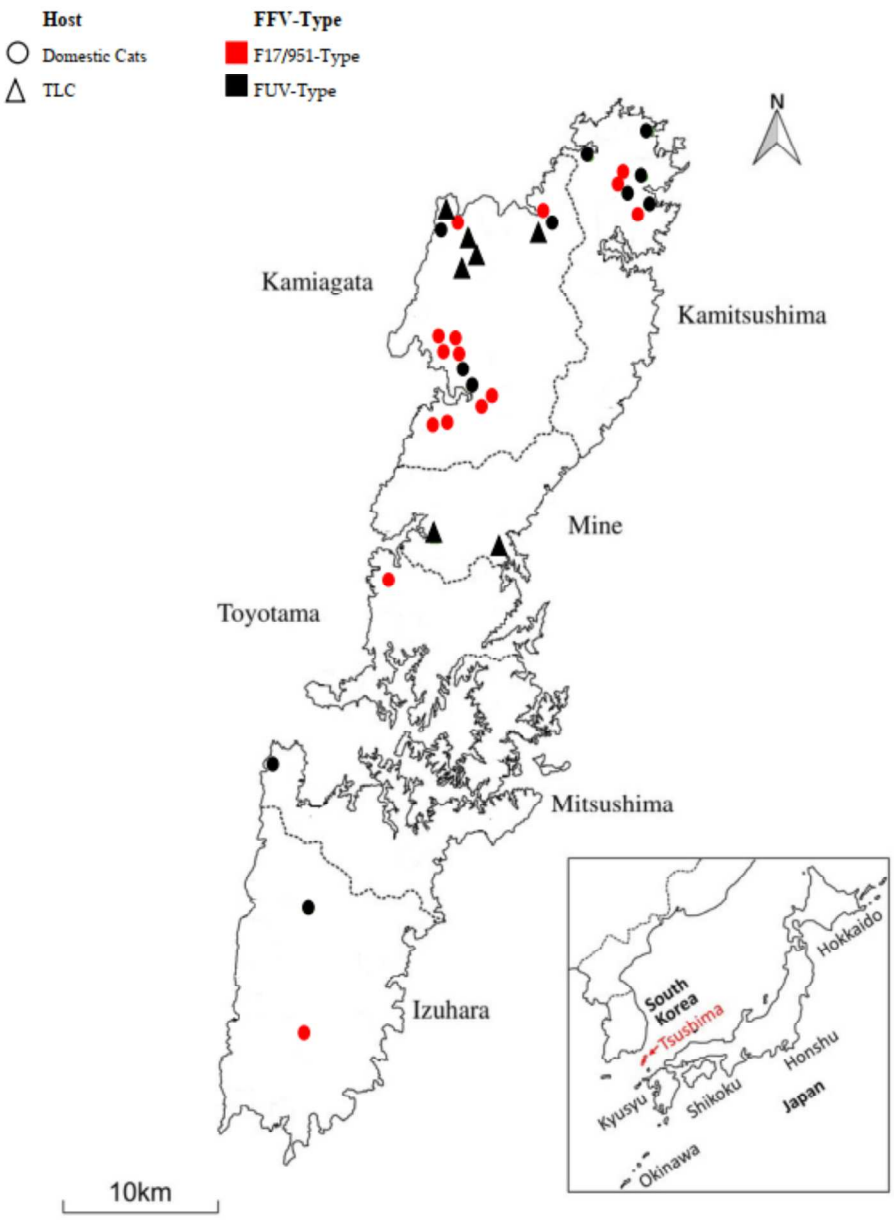


Figure 1.3 Maps of Tsushima Island and positive sample sites. Cycle indicate domestic cat samples. Triangle indicate TLC samples. Red color indicates F17/951-Type. Black color indicates FUV-Type

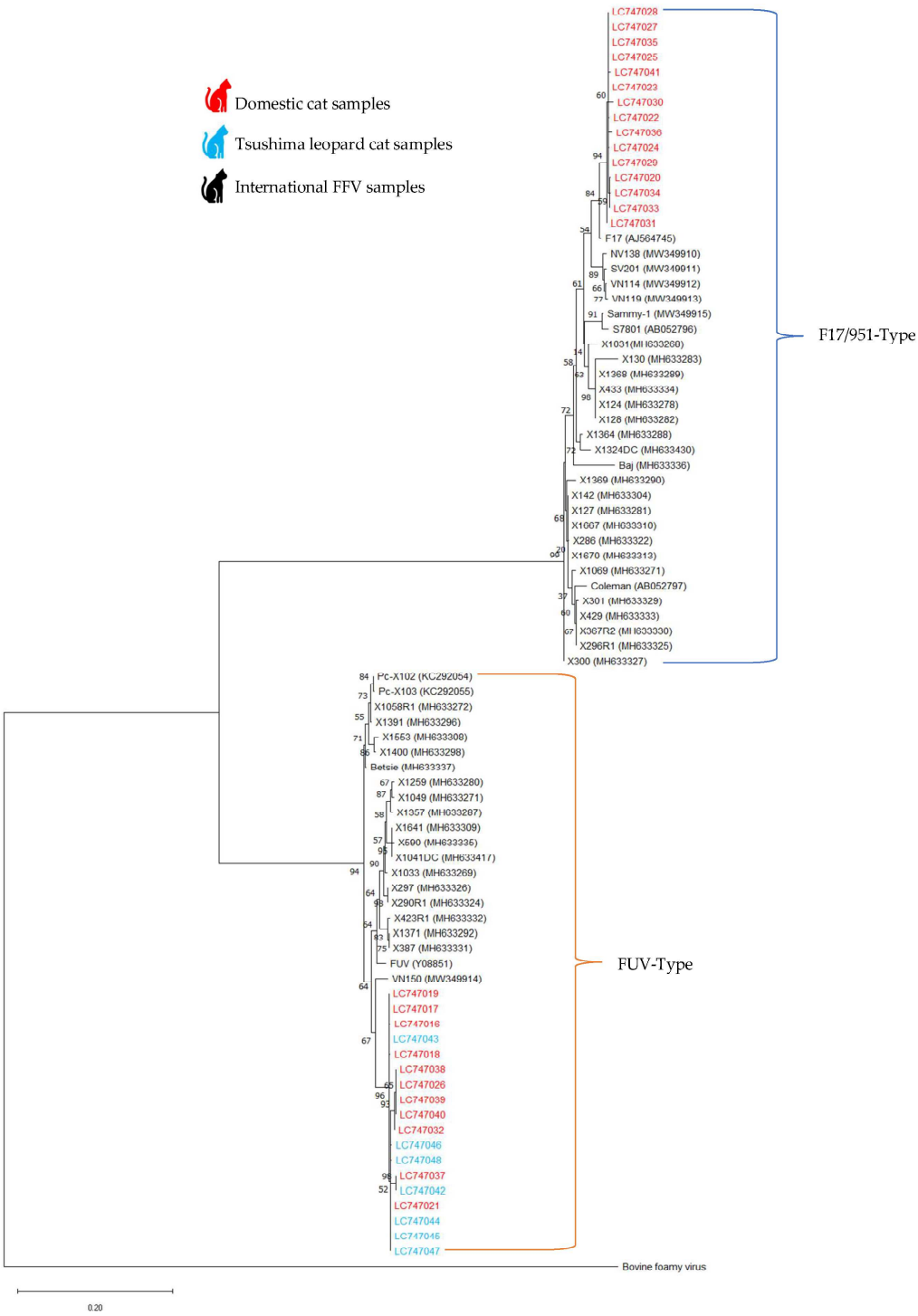


Figure 1.4 Maximum likelihood tree for SU region nucleic acid sequence. The percentage of replicate trees in which the associated taxa clustered in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary

distances were computed using the Kimura 2-parameter method and are in the units of the number of base differences per site. This analysis involved 84 nucleotide sequences.

2. CHAPTER TWO

Riboflavin transporter: evidence of a role as entry receptor for chimpanzee endogenous retrovirus.

This work has been published as follows:

AbuEed L, Miyake A, Wanjala N, Pramono D, Abdillah D, Imamura M, Shimojima M, Denner J, Kawasaki J, Nishigaki K. 2025. Riboflavin transporter: evidence of a role as entry receptor for chimpanzee endogenous retrovirus, *Virus Evolution* 11(1), veaf031. https://doi.org/10.1093/ve/veaf031.

2.1 ABSTRACT

Endogenous retroviruses (ERVs) are remnants of ancestral viral infections in germ cells that constitute a substantial proportion of the mammalian genome and are assumed to provide molecular fossil records of ancient infections. Analysis of these sequences may reveal the mechanisms of virus–host co-evolution, viral endogenization, and extinction. Chimpanzee endogenous retrovirus 1 (CERV1), a gamma retrovirus, is estimated to have circulated within primates for ~10 million years, although it is now apparently extinct. In this study, we aimed to gain an understanding of how the extinct CERV1 was transmitted and endogenized. On the basis of the identification of CERV1 fossils in the primate genome and using the expression-cloning method with the human cDNA library, we found that riboflavin transporter human SLC52A2 served as a receptor for CERV1 entry. The ectopic expression of human and chimpanzee SLC52A2 and its related SLC52A1 in heterogenic cells confers susceptibility to infection by CERV1 and porcine endogenous retrovirus (PERV). Virus interference experiments have shown that CERV1 inhibits infection by PERV and *vice versa*. This finding indicates that CERV1 and PERV belong to the same virus interference group. CERV1 shows infection in a wide range of human and primate cells. Notably, CERV1 infection is observed in human cell lines that express human SLC52A2 abundantly but hardly express human SLC52A1. Although CERV1 has been established to be present at high copy numbers in the great apes (*Pan troglodytes*, *Pan paniscus*, and *Gorilla gorilla*) and 15 Old World monkey species of the Cercopithecinae and Colobinae subfamilies, it is absent in humans and orangutans. CERV1 gene expression is observed in primates, including chimpanzees, suggesting that CERV1 has co-evolved with its hosts. Our results suggest that ERVs may have conferred resistance to viral infections in a convergent evolutionary manner. These findings are significant not only for advancing the field of paleovirology but also in terms of gaining an understanding of the potential risks of viral infection with respect to xenotransplantation, such as that from pigs to humans.

2.2 INTRODUCTION

Retroviruses can be classified as exogenous (exRVs) and endogenous (ERVs) retroviruses based on their mode of transmission. exRVs are transmitted horizontally between hosts and can induce malignant disease in susceptible individuals, whereas ERVs are ancient retroviral sequences that infect germ cells and are transmitted vertically to the offspring of different vertebrate hosts (5). In the latter case, this mode of transmission can lead to the accumulation of a substantial proportion of ERVs in the mammalian genome *via* repeated germline re-infections (21). Notably, ERVs are gradually attenuated or inactivated by mutations occurring during viral replication and *via* post-insertional mutation events occurring during host genome replication over time (23). Hosts frequently regulate the gene expression of ERVs *via* epigenetic mechanisms, including CpG methylation of nucleotides or histone modifications, and it is assumed that these events ensure that ERVs do not pose a threat to the host (71). Moreover, some ERVs undergo a process of domestication within their host organisms, thereby gradually acquiring physiological functions (24). In this regard, it has been demonstrated that ERVs impede the circulation of exogenous retroviruses; e.g. ERV envelope (Env)-derived proteins that restrict retrovirus infection have been well documented in several species, including chickens, cats, and mice, as well as in humans (25–30). Moreover, endogenous viral sequences represent a fossil record of past infections, thereby facilitating the study of paleovirulence (31), which seeks to determine how ancient, possibly extinct, retroviruses infected host species. With the exception of chimpanzee endogenous retrovirus 1 (CERV1)/*Pan troglodytes* endogenous retrovirus 1 and CERV2, all 40 ERV groups in chimpanzees are represented by lineal homologues in humans (32, 33). However, whereas studies have shed some light on the CERV2 infection process (34, 35), the detailed mechanisms of CERV1 infection have yet to be sufficiently determined. Accordingly, in this study, we aimed to gain an understanding of how CERV1, an extinct strain of gamma retrovirus found in chimpanzees, was originally transmitted and endogenized. We found that CERV1 began to endogenize within primate germlines by utilizing riboflavin transporters as receptors. Furthermore, our findings

provide evidence that ERVs may confer resistance to viral infections in a convergent evolutionary manner.

2.3 MATERIALS AND METHODS

2.3.1 Cell culture

The cells were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM; FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) supplemented with 10% foetal bovine serum. The cells were then incubated in a 5% CO_2 incubator at 37°C. In this study, we used the following cell lines: HEK293T (human embryonic kidney transformed with SV40 large T antigen) kindly provided by Dr Mari Kannagi (Tokyo Medical and Dental University); MDTF (*Mus dunni* tail fibroblast) kindly provided by Dr Yoshinao Kubo (Nagasaki University); HeLa (human), Cos7 (African green monkey), and Vero (African green monkey) obtained from the Cell Resource Center for Biomedical Research, Institute of Development, Aging and Cancer, Tohoku University; HT1080 (human), Hs68 (human), NCMDF (normal cynomolgus monkey dermal fibroblasts), and PK15 (pig) obtained from the Japanese Collection of Research Bioresources (JCRB) cell bank; and Chimp-0363M (chimpanzee fibroblasts) kindly provided by Dr Masanori Imamura (Kyoto University) (72) and transformed with SV40 small T antigen (Chimp-0363M-T cells). HEK293T cells persistently infected with PERV-A/C (73) were provided by Dr Joachim Denner (Free University Berlin). Additionally, GPLac cells harboring murine leukemia virus (MLV) Gag-Pol expression vector and pMXs retroviral vector carrying LacZ (74), and PLAT-GP cells (kindly provided by Dr. Toshio Kitamura, The University of Tokyo) harboring MLV Gag-Pol expression vector were cultured under the same conditions.

2.3.2 Pseudo-typed virus preparation

Approximately 5×10^5 GPLac cells were seeded in 6-well plates 1 day prior to transfection. The cells were transfected with 2.5 μg Env expression plasmid using TransIT®-293 reagent (Takara, Shiga, Japan) to produce LacZ-carrying Env-pseudotyped viruses. After 48 h, cell

supernatants were collected, filtered through a 0.22-μm filter (Merck, Darmstadt, Germany), and stored at −80°C. The following Env expression plasmids were used for pseudo-typed virus preparations: pFUΔss CERV1 (CERV1 env) (34, 35), pFUΔss CERV1 Env-10346 [gene accession number, NC_072421.1 (position 34511146–34509476)], pFUΔss CERV1 Env-10346 L241P (substitution of proline at position 241 with leucine in the Env-10346 protein), pFUΔss PERV-A (PERV-A env) (74), pFUΔss PERV-A/C (PERV-A 14/220 env) (accession number, GenBank: AAT77168.1), and pFUΔss 4070A (amphotropic MLV 4070A env) (74). pFUΔss CERV1 Env-10346 and pFUΔss PERV-A/C (PERV-A 14/220 env) were constructed *via* gene synthesis and were Myc-tagged at the 3′ end (Eurofins Genomics K.K., Tokyo, Japan). pFUΔss CERV1 Env-10346 L241P was constructed *via* site-directed mutagenesis using the following primers: Fe-896S, 5′GTACCAGTGGTGGGGCGGAGAGGAGCAGATTGATCCCAGA-3′, and Fe-920R, 5′-TCTGGGATCAATCTGCTCCTCTCCGCCCCACCACTGGTAC-3′. Mutagenesis was performed using the KOD-ONE Blue system (Toyobo, Osaka, Japan) in accordance with the manufacturer's protocol. The resulting mutants were validated through sequencing.

2.3.3 Infection assay

Cells were inoculated with the virus in 24-well plates containing 10 μg/ml polybrene (Nacalai Tesque, Inc., Kyoto, Japan) for 2 h. After the addition of fresh medium, the cells were cultured for 2 days post-infection. To conduct the LacZ assay, culture supernatants were discarded, and the cells were fixed with 250 μl of 2% glutaraldehyde for 15 min at room temperature and then stained with 250 μl of 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-Gal) solution. After incubation at 37°C for 2 h, the nuclei in LacZ-positive cells were counted using an Olympus CKX31 microscope (Olympus, Tokyo, Japan), and images were captured using a camera.

2.3.4 Viral interference assay

For the purpose of interference assays, HEK293T cells persistently infected with replication-competent PERV-A/C (73) and HEK293T cells seeded in 24-well plates were used as target cells. In addition, HEK293T cells were transfected with the CERV1 Env expression plasmid or empty vector (termed 293 T/CERV1 or 293 T/empty cells, respectively) using TransIT®-293 reagent (Takara), and as target cells, were thereafter cultured for 24 h prior to viral infection. Target cells were infected with each type of Env-pseudotyped virus in the presence of polybrene for 2 days and subsequently stained with X-Gal in accordance with the method described in the 'Infection assay' subsection. Single-cycle infectivity was assessed by counting the number of blue-stained nuclei.

2.3.5 Receptor screening

CERV1 Env-pseudotyped virus has previously been shown to infect HEK293T cells (34), and consequently, we screened a HEK293 cDNA library for genes encoding receptors potentially used by CERV1 for entry. A retroviral HEK293T cDNA library was constructed using a pMX retroviral vector. Poly(A)+ RNA was prepared from HEK293 cells using a FastTrack kit version 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol. cDNA was synthesized from the RNA using oligo dT12–18 primers, random hexamers, and Superscript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA). Double-stranded cDNA was synthesized using DNA ligase, DNA polymerase I, and RNase H. Blunt-end cDNA was adapted using a BstXI adaptor (Invitrogen) and ligated to the BstXI sites of the pMX vector. The ligated DNA was then amplified in *Escherichia coli* DH10B competent cells, and plasmid DNA was extracted from cells using a Qiagen Plasmid Midi Kit (QIAGEN, Tokyo, Japan).

MDTF cells, which are resistant to CERV1 infection (34), were transduced with the retroviral HEK293 cDNA library using a Retrovirus Packaging Kit Ampho (Takara, Kusatsu, Japan). PLAT-GP cells stably transduced with MLV Gag-Pol were transfected with CERV1 Env plasmid and either retroviral vector (pMSCV-puro) (Clontech, Mountain

View, CA, USA) or pMXs encoding green fluorescent protein (GFP) using TransIT®-293 reagent (Takara). The supernatants were collected 2 days later and filtered (0.22 μm) to prepare viral suspensions, which were stored at −80°C until used for analysis. The viral titre was > 1 × 105 IU/ml. The CERV1 Env-pseudotyped virus was also used to challenge MDTF cells transduced with the cDNA library in eight 10-cm-diameter culture dishes, with the cells being selected using 6 μg/ml puromycin (InvivoGen, San Diego, CA, USA). The puromycin-resistant colonies were plated onto culture plates, and these cells were subsequently challenged with CERV1-enveloped viruses carrying GFP. Sequence analysis revealed that the cDNAs encoded the riboflavin transporter human solute carrier 52A2 (huSLC52A2) in all GFP-positive colonies.

2.3.6 Establishment of cell lines expressing SLC52A2 and SLC52A1

Human SLC52A2 cDNA was polymerase chain reaction (PCR)-amplified using cDNA from HEK293T cells as a template with KOD One Master Mix (Toyobo) and the primer pair huPAR1-F1 (5′-CTAGAATTCGCCACCATGGCAGCACCCACGCCCGCCCG-3′) and huPAR1-R1 (5′GTGAAGATCTTCAAGCGTAATCTGGAACATCGTATGGGTAG GAGTCACAGGGGTCTGCACAGTCCTTTC-3′), with the resulting amplicons being inserted into a pMSCVneo expression vector (Clontech, Mountain View, CA, USA). Human solute carrier 52A1 (huSLC52A1) (accession number NP_001098047.1), chimpanzee SLC52A2 (accession number XP_016801823.1), and chimpanzee SLC52A1 (accession number XP_001164395.4) cDNAs were commercially synthesized by Eurofins Genomics K.K. (Tokyo, Japan). cDNA fused with a C-terminal hemagglutinin tag was also inserted into the pMSCVneo expression vector.

MDTF-chimpanzee SLC52A2 (MDTF-chSLC52A2), MDTF-human SLC52A2 (MDTF-huSLC52A2), and MDTF-human SLC52A1 (MDTF-huSLC52A1) cell lines were established. Briefly, PLAT-GP packaging cells expressing MLV Gag-Pol were co-transfected with expression vectors (pMSCVneo-chSLC52A2, pMSCVneo-chSLC52A1, pMSCVneo-huSLC52A2, or pMSCVneo-huSLC52A1) and a pFUΔss-Ampho 4070A

MuLV Env expression vector using TransIT®-293 reagent, and then filtered (0.22 μm) supernatants were used to infect MDTF cells. Cells were cultured in DMEM containing 600 μg/ml neomycin.

2.3.7 Detection of SLC52A1 and SLC52A2 expression by reverse transcription quantitative PCR

Total RNA was extracted from the cell lines HEK293T, HeLa, HT1080, HS68, NCMDF, Vero, Cos7, and chimpanzee cells using the RNeasy Mini kit (QIAGEN) in accordance with the manufacturer's instructions. Thereafter, cDNA was synthesized using the PrimeScript II first-strand cDNA synthesis kit (Takara) according to the manufacturer's instructions. Prior to reverse transcription, the RNA samples were treated with recombinant DNase I (Takara). The cDNA was amplified using SYBR Premix Ex Taq II (Tli RNase H Plus; Takara) in the CFX96 Touch real-time PCR detection system (Bio-Rad, Hercules, CA, USA). SLC52A2 was amplified using the primers Fe-995S (5`-TTCCTGCTTACCCTACGGGC-`3) and Fe-1012R (5`-GGAAGCAGGCCAGGGGATT-`3). SLC52A1 was amplified using the primers Fe-996S (5`-TTGCTGTTGCCATCACTACC-`3) and Fe-1013R (5`-CAAAGCCTCTTCTTCCTCCTTC-`3), and the internal control glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was amplified using the primers hGAPDH-F-Tsu (5′-CACCACCATGGAGAAGGCTG-3′) and hGAPDH-R-Tsu (5′-GCTGATGATCTTGAGGCTGTTGT-3′). All reactions were performed using 8-well white strips (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and run on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Inc., USA). Thermal cycling was performed according to the manufacturer's instructions.

2.3.8 Sequence analysis of SLC52A1 and SLC52A2 genes

PCR amplification was performed using the KOD One Master Mix (Toyobo) following the manufacturer's instructions. cDNA from cultured cells—including HEK293T, HeLa, HT1080, HS68, NCMDF, Vero, Cos7, and Chimp-0363M cells—was used as the template.

For SLC52A2 amplification, primers Fe-991S (5′-AGGTGGGAAAAGAACTGGCT-3′) and Fe-1010R (5′-CAGGGGCTTACCCTACAGC-3′) were used. For SLC52A1 amplification, primers Fe-993S (5′-GGGAAGGACCTGCCTGTGA-3′) and Fe-1011R (5′-GAGTGTGCAGGTGTCCATGA-3′) were used.

2.3.9 Phylogenetic and sequence analyses

To identify CERV1 *env* gene sequences in primates, the CERV1 env amino acid sequence (35) was used as a query for a tBLASTn search against 41 primate genome databases (Table 2.1) at the National Center for Biotechnology Information website (https://blast.ncbi.nlm.nih.gov/Blast.cgi, accessed 3 May 2024). The resulting sequences were filtered based on identity (≥80%) and sequence length (≥450 bp). All the identified CERV1 *env* genes are listed in Table 2.2. Alignment was performed using Muscle (75), and MEGA 11 (76) was used for phylogenetic analysis. A phylogenetic tree of the *env* genes was constructed based on the full-length open reading frame amino acid sequences (Table 2.3) and generated using the neighbour-joining method (77) with the *p*-distance model (78). Phylogenetic trees for the SLC52A1 and SLC52A2 genes (Table 2.4) were constructed using the maximum likelihood method (75). Substitution models were selected based on the lowest Bayesian information criterion score (JTT + G) (79). Bootstrap values from 1000 replicates were used as branch junction percentages to assess nodal support, and tree visualization was performed using iTOL (version 6.9.1) (80).

2.3.10 Estimation of the integration timing of CERV1 in primates

An orthologue-dating method was employed to estimate the integration timing of CERV1 in primates (81). Initially, CERV1 *env* genes were identified using a tBLASTn search of the NCBI genomic database. Subsequently, 10-kbp sequences flanking the CERV1 *env* genes, both upstream and downstream, were extracted, and sequence similarity was analysed using BLASTn software. The presence or absence of orthologous *env* loci was assessed based on the alignment coverage of the flanking sequences, with the

divergence times of the primate species being determined using the TimeTree database (82).

2.3.11 RNA-seq analysis

The expression levels of SLC52A1, SLC52A2, and CERV1 *env* were assessed using publicly available RNA-seq data obtained from the NCBI Sequence Read Archive (SRA) database (83). Data were pre-processed using fastp (version 0.20.1) (84) with specific parameters (−w 4, −y −3, and −x), and fasterq-dump (version 2.7.8a) was used to download the RNA-seq data. Genome index generation and mapping were performed using STAR (version 2.7.8a) (85) with default settings. Expression levels were quantified as transcripts per million (TPM), and reads per kilobase of exon per million mapped reads were calculated using featureCounts (version 2.0.1) (86). The following tissues were not examined owing to the absence of RNA-seq data: placenta, colon, and lymph nodes in *Macaca mulatta*; and placenta, testes, and ovaries in *Macaca fascicularis* and *P. troglodytes*. The SRA run accessions, genome files, and gene annotation files are listed in Table 2.5.

2.3.12 Statistical analysis

Data are presented as the averages with standard deviations. The results of assays were considered statistically significant at $P < 0.01$, as determined using Student's *t*-test.

2.3.13 Ethics approval

Experiments were conducted according to the Guidelines for the Care and Use of Laboratory Animals of the Ministry of Education, Culture, Sports, Science, and Technology, Japan. All the experiments were approved by the Genetic Modification Safety Committee of Yamaguchi University.

2.4 RESULTS

2.4.1 Identification of CERV1 entry receptors

It has been established that CERV1 Env-pseudotyped viruses can infect HEK293T cells (34), and consequently, in the present study, we screened an HEK293 cDNA library for genes encoding receptors potentially used by CERV1 for entry. Given that MDTF cells are resistant to CERV1 infection (34), MDTF cells transduced with the retroviral HEK293 cDNA library were challenged with CERV1-enveloped virus carrying a puromycin resistance gene, and 47 puromycin-resistant colonies were plated onto culture plates. To increase the specificity of the screening for receptor identification, these cells were subsequently challenged with CERV1-enveloped viruses carrying GFP, and among these 47 colonies, nine were identified as GFP-positive. Sequence analysis revealed that in all GFP-positive colonies, the cDNAs encoded the human riboflavin transporter (huSLC52A2).

Having thus identified SLC52A2, we investigated whether this protein and the related SLC52A1 (huSLC52A1) protein (SLC52A2 and SLC52A1, respectively) function as CERV1 entry receptors. The amino acid sequence identity between SLC52A2 and SLC52A1 was 86.2% in *Homo sapiens* and 86% in *P. troglodytes* (Figure 2.1). SLC52A1 and SLC52A2 showed extremely high evolutionary similarity among humans and non-human primates (Figure 2.2). Chimpanzee SLC52A2-, human SLC52A2-, chimpanzee SLC52A1-, and human SLC52A1-expressing MDTF cells, designated MDTF-chSLC52A2, MDTF-huSLC52A2, MDTF-chSLC52A1, and MDTF-huSLC52A1, respectively, were characterized by permissiveness to CERV-1 Env-pseudotyped virus infection, whereas MDTF cells were not (Figure 2.3A). These findings thus provided evidence to indicate that the expression of chimpanzee and human SLC52A2 and SLC52A1 in MDTF cells confers susceptibility to infection with CERV1-pseudotyped viruses. Furthermore, huSLC52A2 and huSLC52A1 have been established as entry receptors for porcine endogenous retrovirus subgroup A (PERV-A) and human tropic PERV-A/C (87). Viral infection using the PERV-A and PERV-A/C Env-pseudotyped viruses was confirmed

in MDTF-chSLC52A2, MDTF-huSLC52A2, MDTF-chSLC52A1, and MDTF-huSLC52A1 cells (Figure 2.3A), thereby indicating that chimpanzee and human SLC52A2 and SLC52A1 expression confers susceptibility to infection with pseudo-typed PERV.

To investigate the possibility that primates have functional CERV1 *env* genes in their genomes, we analysed publicly available whole-genome sequence data (Table 2.1). Using the *env* gene of CERV1 as a reference (35), we downloaded all available CERV1 *env* genes from primate genomes (Table 2.2) and extracted their full-length amino acid sequences (Table 2.3). We then determined the full-length amino acid sequence with the highest similarity to the reference CERV1. The chimpanzee sequence Env-10346 [NC_072421.1 (position 34511146–34509476)] showed the highest similarity (99.28%) to the reference CERV1 (Figure 2.3B). The CERV1 Env-pseudotyped virus using Env-10346 was assessed for infectivity; however, a loss of infectivity was detected. In the analysis of several amino acid substitutions in Env-10346 protein [substitution of glutamic acid at position 119 with lysine (E119K) and substitution of leucine at position 241 with proline (L241P)], infectivity was restored by substituting leucine at position 241 with proline in the Env protein (Figure 2.3C), but not by E119K. This result shows that CERV1 *env* acquires infectivity through at least one amino acid change, but although 31 chimpanzee CERV1 *env* genes have been identified (Table 2.3), it is currently unknown whether they retain their infectivity. Notably, a single amino acid change of Env in chimpanzees results in restoration of infectivity, suggesting that it has the potential to become a threat, although CERV1 has become extinct.

2.4.2 Viral receptor interference of CERV1 and PERV-A

The PK15 porcine kidney cell line, which produces PERVs, was found to be characterized by resistance to CERV1 infection, as well as infection with both PERV-A and PERV-A/C, by using pseudo-typed viruses (Figure 2.4A). Therefore, we performed viral interference assays to determine whether CERV1 and PERV can be classified within the same interference group. We accordingly observed that infection with CERV1, PERV-A, and PERV-A/C Env-pseudotyped viruses was inhibited in the presence of CERV1 Env (Figure

2.4B). Subsequently, HEK293T cells persistently infected with PERV-A/C (293T//PERV-A/C inf. cells) were assessed for their susceptibility to CERV1 infection and were found to be resistant to CERV1, PERV-A, and PERV-A/C infections (Figure 2.4C), thereby indicating that CERV1 interferes with the infection of both PERV-A and PERV-A/C *via* at least Env-mediated interference. These results indicate that CERV1, PERV-A, and PERV-A/C belong to the same interference group.

2.4.3 CERV1 infectivity in human and primate cell lines

As CERV1, PERV-A, and PERV-A/C utilize the same receptor, human and primate cell lines were assessed for susceptibility by Env-pseudotyped viruses (Figure 2.5). CERV1 was found to infect humans (293T, HeLa, HT1080, and Hs68 cells), chimpanzees (Chimp-0363M cells), and African green monkeys (Cos7 and Vero cells), although not NCMDF. Contrastingly, PERV-A was found to infect HEK293T and HeLa cells but not other lines. Moreover, with the exception of Vero cells, PERV-A/C was established to show infectivity in the human and primate cell lines assessed in this study. These findings indicate that despite using the same viral receptor, there are differences in susceptibility with respect to viral cell entry.

2.4.4 Expression of SLC52A1 and SLC52A2 in human and primate cell lines

SLC52A1 and SLC52A2 mRNA expression was evaluated by reverse transcription quantitative (RT-q)PCR in total RNA extracted from various human and non-human primate cell lines (Fig. 2.6A and B). SLC52A2 expression was observed in all non-human primates (*P. troglodytes*, *M. fascicularis*, and *Chlorocebus sabaeus*) and human cell lines tested. In contrast, SLC52A1 expression was very low or undetectable in the 293T, HeLa, HT1080, Vero, and Cos7 cells, but high in chimp-0363M, NCMDF, and Hs68 cells. Furthermore, the gene expression levels of SLC52A1 and SLC52A2 in Hs68, Vero, and Cos7 cells were analysed using publicly available RNA-sequencing data (Table 2.5). The expression data for SLC52A1 and SLC52A2 in HEK293, HeLa, and HT1080 cells were

obtained from the Human Protein Atlas Database (88) (accessed 3 March 2025). The expression level of SLC52A1 was 0 or close to 0 in all cells tested in this study, while SLC52A2 was expressed abundantly (Figure 2.6C). The results are consistent with those obtained by RT-qPCR. These results show that SLC52A2 is highly expressed across cell lines, whereas the expression level of SLC52A1 varies from highly expressed to barely expressed.

2.4.5 Sequence analysis of SLC52A1 and SLC52A2 in human and primate cell lines

We analysed the sequences of SLC52A1 and SLC52A2 in the cell lines used in this study. The amino acid sequences of SLC52A2 were identical to the reference genome in all cases (Figure 2.7). As a characteristic amino acid, the 109th amino acid of SLC52A2 was serine in NCMDF cells derived from crab-eating macaques, but not in other cells. In contrast, SLC52A1 showed several substitution mutations across different cell lines (Figure 2.8). In the chimpanzee cell line, SLC52A1 showed the substitution mutation M200L; in the NCMDF cell line, the F181S mutation was observed; and in the African green monkey cell line, an L73S substitution mutation was observed (Figure 2.8). In the Hs68 cell line, SLC52A1 showed two substitution mutations, Q71R and A271V. In contrast, in HT1080, 293T, and HeLa cell lines, the nucleotide sequences of SLC52A1 were not determined due to the lower or undetectable expression by using nested PCR.

2.4.6 Expression of SLC52A1 and SLC52A2 in human and non-human primate tissues

The gene expression levels of SLC52A1 and SLC52A2 in non-human primates were analysed using publicly available RNA-sequencing data from *P. troglodytes*, *M. fascicularis*, and *M. mulatta* (Table 2.5). For humans, expression data were obtained from the Human Protein Atlas (89) (accessed 15 December 2024). Overall, SLC52A2 exhibited higher expression than SLC52A1 in most tissues across all species, with particularly low

levels in the liver and skeletal muscle (Figure 2.9A–D). In contrast, the expression levels of SLC52A1 were apparently different between human and non-human primates. In humans, SLC52A1 was highly expressed in the placenta, but expression levels were very low in other tissues (Figure 2.9B). In non-human primates, SLC52A1 was expressed in all tissues and highly expressed in the thymus (Figure 2.9B–D).

2.4.7 Genetic diversity of CERV1 *env* in primates

A search of publicly available whole-genome data obtained from humans and primates revealed CERV1 *env* genes in great apes (*P. troglodytes, Pan paniscus,* and *Gorilla gorilla*) and 15 Old World monkey species from the Cercopithecinae and Colobinae subfamilies (Table 2.1).

Contrastingly, CERV1 was not detected in the genomes of either humans or orangutans (*Pongo*), and similarly, *Nomascus leucogenys* and *Symphalangus syndactylus*, in the family Hylobatidae, do not harbour CERV1. The phylogenetic trees constructed based on env amino acid sequences revealed four distinct genetic clusters (Clades 1, 2, 3, and 4) within the CERV1 *env* sequences (Figure 2.10), three of which (Clades 1, 2, and 4) were found in Old World monkeys, whereas Clade 3 was found exclusively in great apes. Furthermore, our phylogenetic analysis revealed that CERV1 *env* sequences from Old World monkeys and apes do not mix, which would tend to indicate an absence of any recent cross-species transmission of CERV1 between these species. However, we found that the copy number of the CERV1 *env* gene in the primate genome varied widely, ranging from 35 to 231 (Table 2.2).

The timing of CERV1-related virus integration, which we determined based on an analysis of orthologous *env* genes in sister species, revealed that CERV1 probably emerged subsequent to a significant divergence between ape and Old World monkey lineages, as certain ape species were found to lack CERV1 (Figure 2.11, Table 2.6). In contrast, ancestral Old World monkeys were infected with CERV1 and transmitted the infection within their lineages several million years ago. Notably, the findings of our TimeTree

phylogenetic tree analysis indicate that Old World monkeys may have been infected with CERV1 prior to apes.

2.4.8 CERV1 *env* expression in tissues

Tissue-specific expression of the CERV1 *env* gene was examined in *P. troglodytes*, *M. fascicularis*, and *M. mulatta*. CERV1 *env* expression was observed in most organs, with the highest levels of expression being detected in the lungs of chimpanzees, the thymus gland in *M. fascicularis*, and the brain frontal cortex in *M. mulatta* (Figure 2.12). Comparatively, we detected low levels of CERV1 *env* expression in the livers of all three species. These results indicate that CERV1 may have co-evolved with the host subsequent to its endogenization.

2.5 DISCUSSION

Our findings in the present study provide evidence to indicate that CERV1 began to endogenize in primate germlines by utilizing riboflavin transporters as receptors ~10 million years ago. Riboflavin transporters have been identified as receptors of both PERV and modified retroviruses (87, 90), thereby indicating that the interaction between the retroviral envelope and the entry receptor is convergent in nature.

In the present study, CERV1 was found to be a viral interference group of PERV-A and PERV-A/C. This conclusion is based on the results showing the resistance of CERV1 infection to HEK293T cells persistently infected with PERV-A/C (Figure 2.4C). Furthermore, it was suggested that this interference mechanism is mediated by the Env protein (Figure 2.4B and C). These results suggest that the receptor for CERV1 is SLC52A1/2. PERV infection has been reported to have a greater function for SLC52A1 as an entry receptor (87, 91). Non-permissive MDTF cells expressing chimpanzee and human SLC52A1/2 cDNA are highly susceptible to CERV1 infection. The infection with a pseudo-virus showed that MDTF cells expressing human SLC52A2 and chimpanzee SLC52A1 had a higher infection titre (Figure 2.3A). Species-specific preferences regarding

the usage of receptors may exist. However, the amount of receptor expression on the cell surface has not been measured, so at present, the preference for SLC52A1/2 is unknown for CERV1 infections. However, we do not intend to overemphasize this functional difference, but rather to show that there is a consistent significant difference in the experiments. Phylogenetic analysis of SLC52A1/2 showed high sequence similarity between species, suggesting that it is a riboflavin transport protein (Figure 2.2).

Analysis of the expression levels of SLC52A1/2 in human and non-human primate cell lines used in the present study showed that the expression levels of SLC52A2 were similar among them (Figure 2.6A–C). The expression levels of SLC52A2 were considerably higher than those of SLC52A1 in all cells, while the expression levels of SLC52A1 varied from undetectable to high in cell lines. In particular, HEK293T, HeLa, and HT1080 cells exhibited extremely low or undetectable levels of SLC52A1. Interestingly, it has been reported that the inactive form of human SLC52A1 cDNA is expressed in cell lines and tissues. This is due to a 118 bp insertion between exon 2 and exon 3, inserting a stop codon due to a frameshift, and therefore, the cDNA encodes a 167 amino acid protein (92). Therefore, SLC52A1 monitored by RT-qPCR is thought to be the sum of the normal form and inactive form, suggesting that the normal form of SLC52A1 was barely expressed (Figure 2.6B and C). Indeed, we failed to determine the sequence of SLC52A1 in these cells. Notably, these cells are permissive for infection with CERV1 and PERV-A/C. SLC52A2 is expressed at extremely high levels in these cultured cells, so it is possible that viral infection occurs using SLC52A2 in these cell lines. Alternatively, even if SLC52A1 is expressed in relatively small amounts, it may function as a receptor for viral infection.

The results of the infectivity of PERV in cell lines in the present study are generally consistent with the results of previous studies (87, 93). The differences in infectivity between PERV-A and PERV-A/C in cell lines are due to regions outside the receptor binding region within Env, and this enhanced the infectivity of PERV-A/C in cells by stabilization of the envelope glycoprotein and increased receptor binding (94).

It has been reported that the resistance of PERV to infection is related to the expression level and glycosylation of viral receptors (e.g. Vero and Cos7 cells) (95). Furthermore, in rhesus macaques, crab-eating macaques, and baboons, SLC52A2 is genetically deficient for PERV infection due to the substitution of leucine with serine at amino acid position 109 in SLC52A2 (91, 95, 96).

In the present study, NCMDF cells were the only cells that did not become infected with CERV1. NCMDF cell are normal dermal fibroblasts and are classified as primary cultured cells that have not been immortalized. Both SLC52A1 and SLC52A2 are highly expressed (Figure 2.6A and B), but the reason for the resistance to CERV1 infection is unknown. Since it was confirmed that NCMDF cells have a serine at position 109 of SLC52A2, it is thought that SLC52A1 functions as the entry receptor for PERV-A/C infection in NCMDF cells. On the other hand, it is unknown whether a serine at position 109 of SLC52A2 affects CERV1 infection.

As a result of experiments on the virus infection in various cell lines (Figure 2.5), it was found that there was a difference in infectivity between CERV1 and PERV, even though they belong to the same viral interference group, suggesting different properties of the virus-receptor interaction.

In the present study, the expression analysis of SLC52A1/2 was only based on quantitative analysis of RNA, so the evaluation of the amount and localization of these endogenous proteins is a future issue.

The utilization of transport proteins for cell entry is a common feature of gammaretroviruses. For example, ecotropic MuLV utilizes mCAT1, the cationic amino acid transporter (97). To date, all known receptors for feline and murine are multi-transmembrane receptors (98, 99). The discovery of SLC52A1/2 as an entry receptor for CERV1 follows this pattern of multipass membrane transport molecules acting as retroviral receptors (100). Some gammaretroviruses are known to share a receptor, such as GaLV, FeLV-B, KoRV-A, and 10A1-MuLV, which all use Pit1 (101–105); FeLV-D, ERV-DC, primate ERVs, and CERV2, which use CTR1 (34, 35, 106); or FeLV-E and murine ERV,

which use RFC1 (107, 108). As shown in these examples, the interactions between retrovirus envelopes and entry receptors can be shown to be convergent in nature. Similar convergent evolution has been reported for antiviral molecules (e.g. Refrex-1) derived from the retroviral envelope (34). It is extremely rare for a gammaretrovirus to utilize multiple entry receptors. There have been reports of FeLV-C mutant strains using molecules, such as THTR1, FLVCR1, and FLVCR2 (109). In addition, although the receptor for FeLV-B is feline Pit1, some FeLV-B mutants have been found to use feline Pit2 (25, 104), although their viral infectivity is low. On the other hand, CERV1, PERV, and CP virus (modified retrovirus) utilize both SLC52A1 and SLC52A2. One reason for this may be that the identities of SLC52A1 and SLC52A2 are very high (Figure 2.1). In the present study, the lack of evaluation of interactions of viral envelopes with endogenous receptors in the experimental system for identifying receptors is an issue for future research.

No remains of CERV1 infection are found in the human genome (35). This research has shown that CERV1 can infect human cell lines that express SLC52A2, regardless of the expression level of SLC52A1, suggesting that species-specific envelope-receptor incompatibility does not account for the absence of endogenous CERV1 copies in the human genome. Furthermore, SLC52A2 is widely expressed in human tissues (Figure 2.9B), including the testes and ovaries, and SLC52A1 is highly expressed in the placenta, suggesting that the lack of endogenization is unlikely to result from the absence of receptor expression in the germline.

There are clear differences in the tissue expression of SLC52A1 between humans and non-human primates. In humans, the expression of SLC52A1 is limited to the placenta; in contrast, it is expressed in a wide range of tissues in non-human primates. This observation may be hypothesized to be due to the different roles of SLC52A1 in different species. Furthermore, the reason why CERV1 is not endogenized in humans may be related to the expression pattern of SLC52A1/2.

It is not yet clear why humans, orangutans, and New World monkeys do not possess endogenous CERV1. In addition to the pattern of receptor expression, factors such as the

time and place of CERV1 prevalence may also be involved. Furthermore, the lack of CERV1 may be due to evolutionary resistance to viral infections, such as APOBEC3 or TRIM5α (110, 111). Furthermore, the results of the present study do not clarify the relationship between SLC52A1/2 gene polymorphisms and susceptibility to viral infection. However, it is known that there are more than 30 and 6 polymorphisms at the amino acid level in human SLC52A1 and SLC52A2, respectively, and the evolution of these genes may be expected. It is possible that there have been polymorphisms that confer resistance to CERV1 infection (100), but all of these are speculations, and this is a limitation of the current research.

Expression analysis of CERV1 in primates and the discovery of a viral receptor provide important insights into xenotransplantation from pigs to humans. Experimental primate models of pig organ transplantation have shown potential resistance to PERV infection (112, 113). This resistance may be due to the fact that the SLC52A2 gene encodes serine at amino acid position 109 in macaque and baboon species, but this substitution is not present in humans, chimpanzees, or African green monkeys. At least, this mutation of serine may explain the resistance to PERV infection. Notably, it is not known whether this amino acid mutation in SLC52A2 affects the infectivity of CERV1. However, considering the timing of CERV1 endogenization in non-human primates, the 109 Ser of SLC52A2 is thought to have occurred after CERV1 endogenization. The mechanism of CERV1 endogenization may not involve the genetic change in SLC52A2. In the present study, the finding that CERV1 and PERV-A belong to the same interference group and the fact that the CERV1 *env* gene is expressed in primates suggest that it is possible that primates exhibit resistance to PERV infection. At present, it is not yet clear whether the endogenization of CERV1 confers resistance to PERV infection. However, the physiological significance of endogenized CERV2-related sequences as a defence against feline and primate retroviral infection has been reported (34). Therefore, these gaps need to be tackled in future research.

Finally, since CERV1 belongs to the same interference group as PERV, the riboflavin transporter is the entry receptor for CERV1. However, it should be noted that the identification of the receptor is based on the results of experiments of ectopic expression of the receptors; it will be necessary to clarify this issue in the future using other approaches, such as receptor knockout experiments in cells.

2.6 CONCLUSION

Collectively, the findings in this study will contribute to enhancing our understanding of the transmission trajectories of ancient viruses, provide clues as to why these viruses have not been endogenized in humans, and offer insights for research on human and primate evolution. In particular, given that ancient retroviruses have integrated into the genomes of humans and animals, the question of how these viruses infected their hosts represents a fundamental issue in the field of life sciences, which ideally needs resolution.

2.7 TABLES, FIGURES, SUPPLEMENTARY DATA IN CHAPTER TWO

Table

Table 2.1 Primate genome reference used in this study.

No.	**Species**	**Assembly name**	**GenBank no.**
1	*Pan troglodytes*	NHGRI_mPanTro3-v2.0_pri	GCA_028858775.2
2	*Pan paniscus*	NHGRI_mPanPan1-v2.0_pri	GCA_029289425.2
3	*Gorilla gorilla*	NHGRI_mGorGor1-v2.0_pri	GCA_029281585.2
4	*Rhinopithecus roxellana*	ASM756505v1	GCA_007565055.1
5	*Piliocolobus tephrosceles*	ASM277652v5	GCA_002776525.5
6	*Macaca mulatta*	Mmul_10	GCA_003339765.3
7	*Macaca fascicularis*	T2T-MFA8v1.0	GCA_037993035.1
8	*Macaca thibetana*	ASM2454274v1	GCA_024542745.1
9	*Theropithecus gelada*	Tgel_1.0	GCA_003255815.1
10	*Papio anubis*	Panubis1.1	GCA_008728515.2
11	*Mandrillus leucophaeus*	Mleu.le_1.0	GCA_000951045.1
12	*Papio hamadryas*	ASM2378191v1	GCA_023781915.1
13	*Cercopithecus mona*	KIZ_CMon_1.0	GCA_014849445.1
14	*Erythrocebus patas*	ASM2378345v1	GCA_023783455.1
15	*Macaca silenus*	ASM2380736v1	GCA_023807365.1
16	*Macaca nigra*	mMacNir1.1	GCA_928851695.1
17	*Lophocebus aterrimus*	ASM2378323v1	GCA_023783235.1
18	*Rhinopithecus strykeri*	ASM2376470v1	GCA_023764705.1
19	*Nycticebus bengalensis*	ASM2389825v1	GCA_023898255.1
20	*Nycticebus coucang*	mNycCou1.pri	GCA_027406575.1
21	*Loris tardigradus*	ASM2378313v1	GCA_023783135.1
22	*Galago moholi*	ASM2378343v1	GCA_023783435.1
23	*Otolemur garnetti*	OtoGar3	GCA_000181295.3
24	*Daubentonia madagascariensis*	ASM2378347v1	GCA_023783475.1
25	*Microcebus murinus*	Mmur_3.0	GCA_000165445.3
26	*Lemur catta*	mLemCat1.pri	GCA_020740605.1
27	*Pithecia pithecia*	ASM2377967v1	GCA_023779675.1
28	*Aotus nancymaae*	86718_ANA_hifiasm-v0.15.2.pri	GCA_030222135.1
29	*Callithrix jacchus*	mCalJa1.2.pat.X	GCA_011100555.2
30	*Saguinus midas*	ASM2149847v1	GCA_021498475.1
31	*Cebus albifrons*	ASM2378357v1	GCA_023783575.1
32	*Sapajus apella*	GSC_monkey_1.0	GCA_009761245.1
33	*Ateles geoffroyi*	ASM2378355v1	GCA_023783555.1
34	*Hoolock leuconedys*	ASM2374817v1	GCA_023748175.1
35	*Symphalangus syndactylus*	NHGRI_mSymSyn1-v2.1_pri	GCA_028878055.3
36	*Hylobates pileatus*	ASM2149846v1	GCA_021498465.1
37	*Nomascus leucogenys*	Asia_NLE_v1	GCA_006542625.1
38	*Homo sapiens*	GRCh38.p14	GCA_000001405.29
39	*Pongo pygmaeus*	NHGRI_mPonPyg2-v2.0_pri	GCA_028885625.2
40	*Pongo abelii*	NHGRI_mPonAbe1-v2.0_pri	GCA_028885655.2
41	*Cephalopachus bancanus*	ASM2725705v1	GCA_027257055.1

Table 2.2 CERV1 env gene locations identified in the *Macaca mulatta* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_041755.1	96	57540227	57538557	0
2	NC_041755.1	95	135766213	135767883	0
3	NC_041755.1	90	1749040	1750710	0
4	NC_041755.1	89	172962530	172964200	0
5	NC_041755.1	90	46939078	46940685	0
6	NC_041755.1	92	147044530	147045738	0
7	NC_041755.1	94	30289	29318	0
8	NC_041755.1	92	1726083	1727219	0
9	NC_041755.1	94	53015936	53015067	0
10	NC_041755.1	97	53015067	53014261	0
11	NC_041755.1	98	29271	28624	0
12	NC_041755.1	86	1725563	1726087	0
13	NC_041755.1	93	65560783	65559923	6.30E-136
14	NC_041755.1	87	57546685	57545933	4.12E-109
15	NC_041755.1	82	57547140	57546583	1.32E-89
16	NC_041755.1	95	57545629	57545153	3.28E-76
17	NC_041760.1	97	168720624	168718954	0
18	NC_041760.1	91	145780424	145778754	0
19	NC_041760.1	97	61585384	61584062	0
20	NC_041760.1	94	2729461	2730777	0
21	NC_041760.1	94	31552258	31553379	0
22	NC_041760.1	84	31551710	31552255	0
23	NC_041760.1	93	82531890	82532711	3.31E-127
24	NC_041760.1	89	103651184	103651855	1.37E-122
25	NC_041764.1	96	95253351	95251681	0
26	NC_041764.1	91	69994228	69992606	0
27	NC_041764.1	89	54546184	54547809	0
28	NC_041764.1	87	119684573	119683299	0
29	NC_041764.1	88	104072236	104071034	0
30	NC_041764.1	89	58527726	58528877	0
31	NC_041764.1	95	62308969	62307785	0
32	NC_041764.1	94	59003518	59002337	0
33	NC_041764.1	94	81933354	81934286	0
34	NC_041764.1	94	58528907	58529395	0
35	NC_041764.1	84	62309454	62308972	0
36	NC_041764.1	96	32894076	32893618	5.13E-90
37	NC_041764.1	95	119683080	119682604	1.30E-76
38	NC_041757.1	96	44872379	44870709	0
39	NC_041757.1	92	13504084	13502414	0
40	NC_041757.1	91	10270260	10271930	0
41	NC_041757.1	91	127274297	127275967	0
42	NC_041757.1	90	11585521	11583851	0
43	NC_041757.1	86	58957049	58955487	0

44	NC_041757.1	96	124926723	124927559	1.11E-150
45	NC_041757.1	92	86797693	86798559	2.36E-136
46	NC_041756.1	96	153594191	153595855	0
47	NC_041756.1	91	51734746	51733076	0
48	NC_041756.1	92	31351556	31353226	0
49	NC_041756.1	91	151906851	151908512	0
50	NC_041756.1	90	39014840	39016300	0
51	NC_041756.1	90	41889098	41890417	0
52	NC_041756.1	96	164435582	164434905	0
53	NC_041756.1	96	164434405	164433917	0
54	NC_041756.1	97	164434908	164434435	0
55	NC_041756.1	92	176036776	176037231	3.46E-111
56	NC_027914.1	94	5491012	5489342	0
57	NC_027914.1	87	5824423	5825562	0
58	NC_027914.1	91	9632279	9631632	0
59	NC_027914.1	81	9633299	9632799	0
60	NC_027914.1	91	9632796	9632326	0
61	NC_027914.1	96	5825562	5826044	0
62	NC_041754.1	93	104508124	104509794	0
63	NC_041754.1	93	103832211	103830541	0
64	NC_041754.1	91	2370944	2372614	0
65	NC_041754.1	92	132928280	132929950	0
66	NC_041754.1	91	1954143	1952473	0
67	NC_041754.1	90	167262728	167261058	0
68	NC_041754.1	92	102896431	102898101	0
69	NC_041754.1	90	31061137	31059560	0
70	NC_041754.1	94	97483996	97485453	0
71	NC_041754.1	97	127592557	127593846	0
72	NC_041754.1	96	93653371	93652220	0
73	NC_041754.1	88	115120639	115121679	0
74	NC_041754.1	94	115121840	115122307	0
75	NC_041754.1	91	31008615	31009211	1.63E-179
76	NC_041754.1	95	45300182	45299322	4.94E-173
77	NC_041754.1	84	85272287	85273120	2.31E-146
78	NC_041754.1	93	2703528	2702641	2.40E-143
79	NC_041754.1	83	2704310	2703477	6.03E-143
80	NC_041754.1	92	28631551	28632417	3.99E-137
81	NC_041754.1	92	45299373	45298513	3.26E-129
82	NC_041754.1	93	173582262	173581789	2.53E-91
83	NC_041754.1	93	85273075	85273650	3.91E-84
84	NC_041766.1	92	74208538	74206868	0
85	NC_041766.1	92	78171365	78169695	0
86	NC_041766.1	92	46368511	46370181	0
87	NC_041766.1	92	1554085	1555755	0
88	NC_041766.1	91	95251885	95253555	0
89	NC_041766.1	89	11731817	11733487	0

90	NC_041766.1	88	104077288	104078958	0
91	NC_041766.1	89	29669980	29671131	0
92	NC_041766.1	91	29671163	29671651	0
93	NC_041768.1	92	107005390	107003720	0
94	NC_041768.1	91	77042731	77044398	0
95	NC_041768.1	93	105267932	105267204	7.09E-108
96	NC_041765.1	92	5614666	5612996	0
97	NC_041765.1	89	28450519	28452186	0
98	NC_041765.1	88	27215942	27217609	0
99	NC_041765.1	95	29946826	29945852	0
100	NC_041765.1	97	29945798	29945334	0
101	NC_041765.1	90	247886	246822	2.28E-174
102	NC_041765.1	86	248829	248194	3.05E-112
103	NC_041767.1	92	68406838	68405168	0
104	NC_041767.1	92	42557609	42555939	0
105	NC_041767.1	91	40463234	40461564	0
106	NC_041767.1	98	24256312	24257394	0
107	NC_041767.1	80	87128412	87127570	0
108	NC_041767.1	96	24255726	24256316	0
109	NC_041767.1	97	87127510	87126863	0
110	NC_041767.1	90	23445606	23444923	2.91E-115
111	NC_041774.1	92	56654836	56653166	0
112	NC_041774.1	92	100578184	100576514	0
113	NC_041774.1	91	29961957	29963627	0
114	NC_041774.1	90	91815781	91817388	0
115	NC_041763.1	92	57998518	58000188	0
116	NC_041763.1	91	91714340	91716010	0
117	NC_041763.1	86	61858493	61860163	0
118	NC_041763.1	90	21018076	21019191	0
119	NC_041763.1	85	26745447	26746475	5.39E-149
120	NC_041763.1	88	26744810	26745595	1.69E-145
121	NC_041759.1	92	154322916	154321246	0
122	NC_041759.1	91	14634062	14632392	0
123	NC_041759.1	88	50355992	50354322	0
124	NC_041759.1	90	30226087	30227484	0
125	NC_041759.1	98	160720307	160721422	0
126	NC_041759.1	90	105287120	105288196	0
127	NC_041759.1	95	17209468	17210133	0
128	NC_041759.1	92	160719753	160720304	0
129	NC_041759.1	96	105288199	105288786	0
130	NC_041759.1	96	17210163	17210651	0
131	NC_041759.1	84	17208983	17209465	0
132	NC_041759.1	92	71860408	71859587	1.76E-127
133	NC_041759.1	96	131751763	131752251	2.58E-117
134	NC_041759.1	94	149840757	149840269	6.56E-116
135	NC_041771.1	91	60859583	60857913	0

136	NC_041771.1	94	53335427	53336377	8.90E-157
137	NC_041771.1	89	45489172	45489897	1.05E-105
138	NC_041770.1	91	10989258	10987588	0
139	NC_041770.1	91	21558502	21560172	0
140	NC_041770.1	91	91002286	91003956	0
141	NC_041770.1	92	79223357	79222098	0
142	NC_041770.1	97	10136036	10134834	0
143	NC_041770.1	90	31892408	31892941	0
144	NC_041770.1	91	10136503	10136036	0
145	NC_041770.1	95	30516174	30516662	1.32E-119
146	NC_041761.1	92	30423754	30422084	0
147	NC_041761.1	91	57150927	57149257	0
148	NC_041761.1	90	43863624	43865294	0
149	NC_041761.1	95	114234986	114233835	0
150	NC_041761.1	88	48765031	48763880	0
151	NC_041761.1	94	48820982	48820335	0
152	NC_041761.1	93	48821548	48821030	0
153	NC_041761.1	96	114233803	114233315	0
154	NC_041761.1	95	48763848	48763360	0
155	NC_041761.1	92	94259348	94258527	3.20E-124
156	NC_041761.1	93	56344891	56345712	4.58E-122
157	NC_041761.1	96	56254271	56253783	9.22E-117
158	NC_041758.1	91	231340	233010	0
159	NC_041758.1	90	125827477	125829156	0
160	NC_041758.1	91	138706925	138705258	0
161	NC_041758.1	96	69040523	69039168	0
162	NC_041758.1	92	102370870	102372129	0
163	NC_041758.1	87	78301070	78301930	3.24E-158
164	NC_041758.1	89	78301867	78302739	7.93E-131
165	NC_041762.1	91	38268077	38266455	0
166	NC_041762.1	97	75658905	75657664	0
167	NC_041762.1	96	49830452	49831555	0
168	NC_041762.1	89	72303825	72304976	0
169	NC_041762.1	95	93659800	93660750	0
170	NC_041762.1	88	93659081	93659800	0
171	NC_041762.1	96	72305006	72305494	0
172	NC_041762.1	97	49831587	49832075	0
173	NC_041762.1	91	63058197	63057340	1.54E-135
174	NC_041772.1	92	20757432	20759030	0
175	NC_041772.1	87	22123430	22122738	4.57E-169
176	NC_041772.1	93	22122702	22122214	4.57E-169
177	NC_041772.1	84	22123885	22123373	3.39E-84

Table 2.2 CERV1 env gene locations identified in the *Gorilla gorilla* genome

No.	Acc.	Iden.	Start	End	E-Value
1	NC_073229.2	99	44988427	44986757	0
2	NC_073229.2	99	29340199	29338532	0
3	NC_073229.2	97	121562656	121560986	0
4	NC_073229.2	97	70715724	70717394	0
5	NC_073229.2	96	52741970	52740300	0
6	NC_073229.2	94	70692906	70691236	0
7	NC_073229.2	98	29459886	29461241	0
8	NC_073229.2	95	130502966	130504282	0
9	NC_073247.2	99	154819965	154821635	0
10	NC_073247.2	99	120978923	120980593	0
11	NC_073247.2	99	125864299	125865969	0
12	NC_073247.2	98	63587424	63585754	0
13	NC_073247.2	97	85178664	85180334	0
14	NC_073247.2	97	135803421	135801751	0
15	NC_073247.2	97	151083990	151082320	0
16	NC_073247.2	97	96752209	96750716	0
17	NC_073247.2	95	156170589	156171968	0
18	NC_073235.2	98	60397899	60396229	0
19	NC_073235.2	98	116154567	116152897	0
20	NC_073235.2	98	101571915	101573585	0
21	NC_073235.2	97	30244626	30242956	0
22	NC_073235.2	97	33834160	33832493	0
23	NC_073235.2	97	101594741	101596411	0
24	NC_073235.2	97	34143874	34142207	0
25	NC_073235.2	97	90203180	90201510	0
26	NC_073235.2	94	95848992	95850662	0
27	NC_073234.2	98	28818056	28819726	0
28	NC_073234.2	98	69166433	69164763	0
29	NC_073234.2	97	70067488	70069158	0
30	NC_073234.2	97	66970660	66972330	0
31	NC_073234.2	97	139529935	139528265	0
32	NC_073234.2	97	113236525	113234855	0
33	NC_073234.2	97	120408913	120410583	0
34	NC_073234.2	96	94865628	94867298	0
35	NC_073234.2	93	82736975	82738645	0
36	NC_073234.2	92	91027151	91028821	0
37	NC_073234.2	95	66889875	66888547	0
38	NC_073234.2	91	38560184	38561551	0
39	NC_073234.2	98	21411047	21412168	0
40	NC_073234.2	93	66908285	66907158	0
41	NC_073234.2	96	21410499	21411044	0
42	NC_073234.2	88	66908821	66908282	0
43	NC_073228.2	98	98492720	98494390	0

44	NC_073228.2	96	87046902	87048620	0
45	NC_073228.2	97	134854835	134856505	0
46	NC_073228.2	97	39259958	39258288	0
47	NC_073228.2	98	80339552	80337882	0
48	NC_073228.2	98	79418271	79419881	0
49	NC_073228.2	92	158040494	158039001	0
50	NC_073228.2	98	33766427	33767743	0
51	NC_073228.2	92	79424818	79423931	0
52	NC_073228.2	91	79423796	79423149	0
53	NC_073231.2	99	14883177	14884847	0
54	NC_073231.2	98	6517943	6516273	0
55	NC_073231.2	97	79869364	79867697	0
56	NC_073231.2	97	61890947	61892614	0
57	NC_073231.2	97	38431037	38432707	0
58	NC_073231.2	97	65671744	65673414	0
59	NC_073231.2	96	81756747	81755098	0
60	NC_073231.2	97	65033193	65031892	0
61	NC_073231.2	96	15010549	15011541	0
62	NC_073231.2	95	15009875	15010552	0
63	NC_073227.2	98	127833441	127835111	0
64	NC_073227.2	98	204370172	204368502	0
65	NC_073227.2	97	174968823	174967153	0
66	NC_073227.2	97	25189061	25190731	0
67	NC_073227.2	97	188792892	188794562	0
68	NC_073227.2	96	68959246	68957576	0
69	NC_073227.2	93	195605208	195603574	0
70	NC_073227.2	94	201719085	201717415	0
71	NC_073227.2	94	61682487	61681009	0
72	NC_073227.2	93	127800076	127801320	0
73	NC_073227.2	88	146655456	146656817	0
74	NC_073227.2	95	146316434	146315334	0
75	NC_073227.2	97	133975429	133976259	0
76	NC_073227.2	97	133976258	133977085	0
77	NC_086017.1	97	157425383	157427053	0
78	NC_086017.1	95	84022308	84020638	0
79	NC_086017.1	94	162688511	162686841	0
80	NC_086017.1	93	106095726	106097396	0
81	NC_086017.1	94	85287060	85285660	0
82	NC_086017.1	95	131480824	131479508	0
83	NC_086017.1	95	160333373	160334251	0
84	NC_073232.2	98	96501440	96503110	0
85	NC_073232.2	97	142655034	142653364	0
86	NC_073232.2	96	15890060	15891730	0
87	NC_073232.2	96	123651440	123649803	0
88	NC_073232.2	95	66536618	66534969	0

89	NC_073232.2	98	128220845	128219280	0
90	NC_073232.2	94	101688870	101687200	0
91	NC_073232.2	98	142944993	142946486	0
92	NC_073232.2	96	102431032	102429632	0
93	NC_073232.2	97	20575080	20573764	0
94	NC_073232.2	92	73040845	73039586	0
95	NC_073232.2	96	15793357	15792191	0
96	NC_073232.2	99	15793851	15793357	0
97	NC_073230.2	98	48286665	48284995	0
98	NC_073230.2	97	43943633	43941963	0
99	NC_073230.2	94	44096968	44098638	0
100	NC_073230.2	93	62459584	62457980	0
101	NC_073230.2	97	74209827	74211134	0
102	NC_073230.2	87	50997637	50999001	0
103	NC_073230.2	89	14387680	14388681	0
104	NC_073236.2	98	72673020	72671350	0
105	NC_073236.2	98	121599707	121598037	0
106	NC_073236.2	97	76586304	76584634	0
107	NC_073236.2	99	68672826	68674052	0
108	NC_073224.2	98	192826092	192824422	0
109	NC_073224.2	97	95857361	95859031	0
110	NC_073224.2	97	192842961	192841291	0
111	NC_073224.2	94	193524977	193526647	0
112	NC_073224.2	96	104756661	104754991	0
113	NC_073224.2	98	148932360	148933952	0
114	NC_073224.2	94	193604148	193602481	0
115	NC_073224.2	99	172933263	172934756	0
116	NC_073224.2	92	193514232	193512562	0
117	NC_073224.2	91	226254227	226255873	0
118	NC_073224.2	98	97823660	97824973	0
119	NC_073224.2	93	176409151	176407835	0
120	NC_073224.2	95	136014450	136015247	6.86E-125
121	NC_073237.2	97	68936248	68934578	0
122	NC_073237.2	93	49291654	49290035	0
123	NC_073237.2	94	14972861	14974240	0
124	NC_073237.2	94	14728159	14726780	0
125	NC_073237.2	96	47206208	47204907	0
126	NC_073237.2	94	12926189	12927568	0
127	NC_073237.2	93	16433747	16435126	0
128	NC_073246.2	97	37255372	37257042	0
129	NC_073246.2	97	29081537	29079870	0
130	NC_073246.2	97	16607637	16609307	0
131	NC_073246.2	95	2654545	2652875	0
132	NC_073233.2	97	70082795	70081125	0
133	NC_073233.2	97	63199815	63201485	0

134	NC_073233.2	96	66350743	66349073	0
135	NC_073233.2	97	75644777	75646093	0
136	NC_073233.2	90	75823761	75822445	0
137	NC_073233.2	93	99040740	99041894	0
138	NC_073233.2	91	99040228	99040740	0
139	NC_073244.2	97	28218291	28219961	0
140	NC_073244.2	98	30019020	30017731	0
141	NC_073244.2	95	26463806	26465134	0
142	NC_073244.2	98	29456307	29457281	0
143	NC_073244.2	97	30135756	30134764	0
144	NC_073244.2	98	26439383	26438586	0
145	NC_073244.2	97	30136430	30135753	0
146	NC_073244.2	97	29457328	29457975	0
147	NC_073244.2	96	26438556	26438068	0
148	NC_073239.2	97	35983895	35985565	0
149	NC_073239.2	97	35972257	35970587	0
150	NC_073239.2	97	104779353	104781023	0
151	NC_073239.2	96	13364170	13362500	0
152	NC_073241.2	96	48273314	48271644	0
153	NC_073241.2	94	27167554	27168933	0
154	NC_073241.2	94	27183020	27184399	0
155	NC_073241.2	94	27205963	27207342	0
156	NC_073241.2	94	25410117	25411496	0
157	NC_073241.2	94	29500956	29502335	0
158	NC_073241.2	94	25056035	25054656	0
159	NC_073241.2	94	29210078	29208699	0
160	NC_073241.2	94	25404261	25405640	0
161	NC_073241.2	94	30432959	30434338	0
162	NC_073241.2	94	29221792	29220413	0
163	NC_073241.2	94	29215935	29214556	0
164	NC_073241.2	97	45993916	45992600	0
165	NC_073241.2	92	61188930	61190252	0
166	NC_073241.2	96	83387744	83386794	0
167	NC_073241.2	98	83388463	83387744	0
168	NC_073243.2	96	69268288	69269955	0
169	NC_073243.2	96	67279148	67280818	0
170	NC_073243.2	91	99210871	99212538	0
171	NC_073243.2	98	68325441	68326571	0
172	NC_073243.2	97	57746971	57745847	0
173	NC_073243.2	99	57747516	57746971	0
174	NC_073243.2	94	68326549	68327112	0
175	NC_086018.1	95	2200896	2199226	0
176	NC_086018.1	96	8763056	8764336	0
177	NC_086018.1	90	6341581	6342414	1.13E-151
178	NC_086018.1	93	6342363	6343250	1.21E-139

179	NC_073238.2	94	63006969	63008639	0
180	NC_073238.2	88	55433695	55435347	0
181	NC_073238.2	94	6138053	6136674	0
182	NC_073238.2	94	1044565	1043186	0
183	NC_073238.2	95	95375212	95374064	0
184	NC_073238.2	98	95375756	95375211	0
185	NC_073238.2	86	26698837	26699670	2.45E-143
186	NC_073238.2	91	26699619	26700506	3.02E-138
187	NC_073248.2	95	10998207	10996537	0
188	NC_073248.2	92	29464674	29466344	0
189	NC_073248.2	90	28601249	28599579	0
190	NC_073248.2	90	28594974	28593304	0
191	NC_073248.2	92	32022316	32020706	0
192	NC_073248.2	94	53288366	53289745	0
193	NC_073248.2	82	10098329	10099999	0
194	NC_073248.2	93	44397782	44399083	0
195	NC_073248.2	95	31595964	31597151	0
196	NC_073248.2	92	32004764	32003610	0
197	NC_073248.2	94	32005276	32004764	0
198	NC_073248.2	94	31597154	31597633	0
199	NC_073248.2	90	31694761	31695603	2.78E-152
200	NC_073248.2	91	31695572	31696432	1.20E-131
201	NC_073245.2	93	46362168	46360498	0
202	NC_073245.2	92	62424026	62425342	0
203	NC_073245.2	93	46385814	46387130	0

Table 2.2 CERV1 env gene locations identified in the *Pan paniscus* genome

No.	Acc.	Iden.	Start	End	E-Value
1	NC_073259.2	99.282	65032142	65030472	0
2	NC_073259.2	98.564	20162288	20160618	0
3	NC_073259.2	97.691	55525596	55527284	0
4	NC_073259.2	97.846	11095466	11097136	0
5	NC_073259.2	94.973	35005036	35003366	0
6	NC_073259.2	94.255	84309137	84310792	0
7	NC_073259.2	92.998	55049629	55051299	0
8	NC_073259.2	97.59	56063291	56064784	0
9	NC_073259.2	92.101	64732260	64733930	0
10	NC_073259.2	85.278	25413518	25412028	0
11	NC_073259.2	98.25	81860017	81861216	0
12	NC_073259.2	98.701	81861216	81861677	0
13	NC_085927.1	98.923	58955161	58956831	0
14	NC_085927.1	97.846	27028800	27030470	0
15	NC_085927.1	90.647	2341854	2342687	2.93E-154
16	NC_085927.1	97.297	2342636	2343523	4E-149
17	NC_073269.2	99.102	43652227	43650557	0

18	NC_073269.2	98.205	49304038	49302368	0
19	NC_073269.2	97.307	33760265	33758598	0
20	NC_073269.2	96.768	49562070	49560400	0
21	NC_073269.2	93.537	26873238	26874908	0
22	NC_085926.1	98.743	92381223	92379553	0
23	NC_085926.1	98.743	44720755	44719085	0
24	NC_085926.1	97.666	39278326	39279996	0
25	NC_085926.1	94.973	160665968	160664298	0
26	NC_085926.1	95.332	19448235	19446565	0
27	NC_085926.1	89.228	128025538	128023880	0
28	NC_085926.1	96.903	88361657	88363012	0
29	NC_085926.1	97.897	87882167	87880884	0
30	NC_085926.1	91.667	127992286	127991027	0
31	NC_085926.1	81.818	127992693	127992199	3.15E-69
32	NC_073257.2	99.102	56070927	56069257	0
33	NC_073257.2	98.205	76345198	76343528	0
34	NC_073257.2	97.307	142306597	142304927	0
35	NC_073257.2	91.382	96096586	96094916	0
36	NC_073257.2	98.263	27809005	27810213	0
37	NC_073257.2	96.687	142575473	142576468	0
38	NC_073257.2	95.575	142560720	142561397	0
39	NC_073257.2	94.714	142574799	142575479	0
40	NC_073257.2	98.995	142561394	142561990	0
41	NC_073257.2	98.052	27810213	27810674	0
42	NC_073272.2	99.102	35734954	35733284	0
43	NC_073272.2	98.564	138609954	138611624	0
44	NC_073272.2	95.871	120543158	120544801	0
45	NC_073272.2	91.562	65927648	65925978	0
46	NC_073272.2	87.831	44461184	44459523	0
47	NC_073272.2	97.788	35853468	35852113	0
48	NC_073272.2	96.239	36147573	36148928	0
49	NC_073258.2	99.102	42290359	42292029	0
50	NC_073258.2	98.384	57214727	57216397	0
51	NC_073258.2	98.205	95506858	95508528	0
52	NC_073258.2	97.487	78498479	78500149	0
53	NC_073258.2	96.948	51402647	51404317	0
54	NC_073258.2	91.921	72838009	72839679	0
55	NC_073258.2	83.692	56461924	56460410	0
56	NC_073256.2	98.743	87624744	87623074	0
57	NC_073256.2	98.413	114406771	114405449	0
58	NC_073261.2	98.923	95199209	95197539	0
59	NC_073261.2	98.384	42398429	42400099	0
60	NC_073261.2	98.491	61115657	61114266	0
61	NC_073261.2	97.608	55973281	55973907	0
62	NC_073261.2	91.204	55973918	55974562	0

63	NC_073249.2	98.743	160388416	160386746	0
64	NC_073249.2	98.205	66412006	66410336	0
65	NC_073249.2	97.487	115679120	115680790	0
66	NC_073249.2	96.047	171933141	171931852	0
67	NC_073263.2	98.564	8522182	8520512	0
68	NC_073263.2	97.556	3565239	3563644	0
69	NC_073263.2	90.647	2846024	2846857	2.93E-154
70	NC_073263.2	97.635	2846806	2847693	2.84E-150
71	NC_073263.2	98.02	18343497	18342892	2.62E-97
72	NC_073267.2	98.384	10863599	10861929	0
73	NC_073267.2	97.666	11046542	11048212	0
74	NC_073255.2	98.025	48626608	48628278	0
75	NC_073255.2	98.384	53969058	53970728	0
76	NC_073255.2	97.846	41463382	41465052	0
77	NC_073255.2	95.699	75899916	75898249	0
78	NC_073255.2	94.614	60759837	60761507	0
79	NC_073255.2	95	27733088	27734707	0
80	NC_073255.2	94.075	60607865	60609535	0
81	NC_073255.2	97.135	62911227	62910076	0
82	NC_073255.2	98.16	62910046	62909558	0
83	NC_073273.2	97.666	38798134	38799804	0
84	NC_073273.2	97.487	41558383	41560053	0
85	NC_073273.2	97.307	44040920	44042590	0
86	NC_073273.2	96.409	39780893	39782563	0
87	NC_073273.2	97.127	44097042	44098706	0
88	NC_073273.2	95.332	36981202	36979532	0
89	NC_073273.2	96.05	40509822	40511492	0
90	NC_073273.2	94.794	38729397	38727727	0
91	NC_073273.2	93.178	40480782	40482452	0
92	NC_073273.2	96.296	38137773	38139149	0
93	NC_073273.2	93.498	40434527	40435495	0
94	NC_073273.2	93.377	40434075	40434527	0
95	NC_073273.2	90.909	43587032	43586445	1.68E-105
96	NC_073273.2	96.482	38207629	38207033	1.02E-95
97	NC_073273.2	96.465	38920372	38919779	1.86E-94
98	NC_073252.2	98.384	157474306	157472636	0
99	NC_073252.2	98.564	57130677	57129007	0
100	NC_073252.2	98.025	65156974	65158644	0
101	NC_073252.2	98.743	114071158	114069488	0
102	NC_073252.2	97.666	105512326	105513996	0
103	NC_073252.2	96.23	58449410	58451080	0
104	NC_073252.2	95.153	157462083	157460413	0
105	NC_073252.2	95.332	4417679	4416009	0
106	NC_073252.2	98.883	80522020	80523093	0
107	NC_073252.2	98.469	80521434	80522021	0

108	NC_073262.2	98.564	57540145	57541815	0
109	NC_073262.2	97.127	33595917	33594247	0
110	NC_073262.2	94.828	32057327	32058892	0
111	NC_073271.2	97.846	10363775	10365445	0
112	NC_073271.2	90.288	2396580	2397413	8.04E-154
113	NC_073271.2	97.635	2397362	2398249	2.84E-150
114	NC_073254.2	98.205	167458439	167456769	0
115	NC_073254.2	98.384	187489582	187487912	0
116	NC_073254.2	96.768	52907549	52909219	0
117	NC_073254.2	96.23	187450087	187448417	0
118	NC_073253.2	98.025	54203110	54201440	0
119	NC_073253.2	97.666	91587772	91586102	0
120	NC_073253.2	98.141	38371019	38372629	0
121	NC_073253.2	98.025	67427896	67429566	0
122	NC_073253.2	91.087	151994224	151995888	0
123	NC_073253.2	93.125	103831053	103832492	0
124	NC_073253.2	96.314	19848974	19847754	0
125	NC_073253.2	99.379	56079369	56078887	0
126	NC_073260.2	97.487	57323808	57325478	0
127	NC_073260.2	96.948	57298709	57300373	0
128	NC_073270.2	96.768	39828860	39827190	0
129	NC_073270.2	97.593	31909539	31908169	0
130	NC_073270.2	92.788	39846313	39847560	0
131	NC_073270.2	90.196	39884392	39885156	0
132	NC_073270.2	92.337	39885138	39885920	0
133	NC_073266.2	98.564	53243126	53241456	0
134	NC_073266.2	97.556	3539965	3538370	0
135	NC_073266.2	94.375	52045608	52044169	0
136	NC_073266.2	89.928	2860374	2861207	2.16E-152
137	NC_073266.2	95.946	2861156	2862043	2.37E-145
138	NC_073264.2	95.691	2774917	2776587	0
139	NC_073264.2	92.708	59727611	59729050	0
140	NC_073265.2	96.91	14531035	14532102	0

Table 2.2 CERV1 env gene locations identified in the *Pan troglodytes* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_072421.1	99	34511146	34509476	0
2	NC_072421.1	99	136625144	136626814	0
3	NC_072421.1	99	87564916	87566586	0
4	NC_072421.1	98	34921940	34923610	0
5	NC_072421.1	97	118598327	118599970	0
6	NC_072421.1	92	64049688	64048018	0
7	NC_072421.1	88	43258655	43256994	0

8	NC_072421.1	96	34629067	34627712	0
9	NC_072421.1	94	87358279	87357446	8.56E-167
10	NC_072421.1	98	87357497	87356610	1.82E-152
11	NC_072406.1	99	85550049	85548379	0
12	NC_072406.1	98	67212020	67213690	0
13	NC_072406.1	98	111877867	111876545	0
14	NC_072399.1	99	83396786	83398456	0
15	NC_072399.1	99	84089612	84087942	0
16	NC_072399.1	99	65412394	65413785	0
17	NC_072399.1	98	31311230	31312456	0
18	NC_072399.1	95	33693957	33692842	0
19	NC_072399.1	97	70544290	70543664	0
20	NC_072399.1	96	70543653	70543006	0
21	NC_072399.1	100	31312441	31312899	0
22	NC_072399.1	99	33692857	33692399	0
23	NC_072401.1	99	48771490	48769820	0
24	NC_072401.1	97	97651853	97650183	0
25	NC_072401.1	98	43334514	43336184	0
26	NC_072401.1	98	82574242	82572572	0
27	NC_072401.1	94	165820286	165818616	0
28	NC_072401.1	91	133209716	133208223	0
29	NC_072401.1	98	92422608	92423963	0
30	NC_072401.1	98	91948641	91947358	0
31	NC_072401.1	91	23470010	23469177	1.73E-154
32	NC_072401.1	95	23469228	23468341	9.49E-147
33	NC_072409.1	99	43903273	43904943	0
34	NC_072409.1	98	57410605	57412275	0
35	NC_072409.1	98	95987624	95989294	0
36	NC_072409.1	97	78688325	78689995	0
37	NC_072409.1	97	53048184	53049854	0
38	NC_072409.1	95	90641249	90642919	0
39	NC_072409.1	92	73026152	73027822	0
40	NC_072409.1	98	8315113	8314139	0
41	NC_072409.1	98	8246944	8245970	0
42	NC_072409.1	97	8393719	8392745	0
43	NC_072409.1	97	8392692	8392048	0
44	NC_072409.1	97	8314086	8313442	0
45	NC_072409.1	97	8245917	8245273	0
46	NC_072409.1	96	56925410	56926057	2.28E-106
47	NC_072410.1	99	62320673	62319003	0
48	NC_072410.1	98	20516742	20515072	0
49	NC_072410.1	98	55908714	55910384	0
50	NC_072410.1	98	90956543	90954873	0
51	NC_072410.1	98	11363240	11364910	0
52	NC_072410.1	97	52786256	52784586	0

53	NC_072410.1	95	35372574	35370904	0
54	NC_072410.1	93	55441462	55443132	0
55	NC_072410.1	98	56463739	56465232	0
56	NC_072410.1	92	62093411	62095081	0
57	NC_072410.1	98	79203087	79204286	0
58	NC_072410.1	100	79204286	79204747	0
59	NC_072410.1	96	81670464	81671264	1.65E-159
60	NC_072410.1	93	81671215	81672120	6.32E-144
61	NC_072405.1	99	14757977	14759647	0
62	NC_072405.1	98	47217844	47219514	0
63	NC_072405.1	99	40051726	40053396	0
64	NC_072405.1	98	59795771	59794101	0
65	NC_072405.1	98	52566604	52568274	0
66	NC_072405.1	95	64239113	64237449	0
67	NC_072405.1	95	26469868	26471538	0
68	NC_072405.1	94	59338487	59340157	0
69	NC_072405.1	94	147470859	147469189	0
70	NC_072405.1	95	59187641	59188624	0
71	NC_072405.1	95	59186955	59187638	0
72	NC_072408.1	99	63105751	63107421	0
73	NC_072408.1	99	46785874	46784204	0
74	NC_072408.1	97	133440567	133442237	0
75	NC_072408.1	94	105918372	105916933	0
76	NC_072408.1	93	133418290	133419726	0
77	NC_072408.1	92	86911296	86910190	0
78	NC_072408.1	89	86911859	86911284	0
79	NC_072402.1	98	117456636	117454966	0
80	NC_072402.1	98	83466343	83464673	0
81	NC_072402.1	98	83894154	83895824	0
82	NC_072402.1	98	161331036	161329366	0
83	NC_072402.1	98	9266758	9268428	0
84	NC_072402.1	97	58514534	58516204	0
85	NC_072402.1	97	108873554	108875224	0
86	NC_072402.1	95	4476291	4474621	0
87	NC_072402.1	95	161318813	161317143	0
88	NC_072402.1	97	57180096	57179257	0
89	NC_072402.1	99	57179284	57178445	0
90	NC_072412.1	99	12266338	12264662	0
91	NC_072412.1	92	49866936	49868375	0
92	NC_072414.1	98	35581254	35579584	0
93	NC_072414.1	99	12962041	12960374	0
94	NC_072414.1	98	13145955	13147472	0
95	NC_072398.1	98	165559922	165561592	0
96	NC_072398.1	98	66780725	66782395	0
97	NC_072398.1	95	55216927	55218213	0

98	NC_072398.1	93	113409120	113410172	0
99	NC_072398.1	97	111462541	111461717	2.48E-169
100	NC_072398.1	89	111461811	111460873	2.86E-145
101	NC_072417.1	98	24979533	24981203	0
102	NC_072417.1	98	40593315	40591645	0
103	NC_072417.1	98	46104957	46103287	0
104	NC_072417.1	97	46359775	46358105	0
105	NC_072417.1	97	30882653	30880986	0
106	NC_072417.1	94	24674018	24675688	0
107	NC_072404.1	98	44313383	44315053	0
108	NC_072404.1	98	159349997	159348327	0
109	NC_072404.1	97	165794230	165792911	0
110	NC_072404.1	91	179350423	179349590	1.48E-156
111	NC_072404.1	97	179349641	179348820	8.95E-136
112	NC_072419.1	98	12564231	12565901	0
113	NC_072419.1	96	69352	71022	0
114	NC_072419.1	97	799405	798056	0
115	NC_072403.1	98	55330309	55328639	0
116	NC_072403.1	98	94070909	94069239	0
117	NC_072403.1	91	154509228	154510892	0
118	NC_072403.1	94	106344356	106345795	0
119	NC_072403.1	97	21002991	21001771	0
120	NC_072403.1	99	57212328	57211846	0
121	NC_072403.1	97	70082406	70083374	4.95E-168
122	NC_072403.1	98	41176895	41177662	1.96E-158
123	NC_072403.1	97	41177631	41178536	1.06E-151
124	NC_072403.1	98	70081684	70082424	2.11E-149
125	NC_072400.1	98	110340042	110338372	0
126	NC_072400.1	98	112521035	112519365	0
127	NC_072400.1	97	43364253	43365923	0
128	NC_072400.1	96	44893761	44892091	0
129	NC_072400.1	99	56897110	56898261	0
130	NC_072400.1	98	56898293	56898781	0
131	NC_072422.1	97	9223282	9221612	0
132	NC_072422.1	98	3323276	3321606	0
133	NC_072422.1	97	8561179	8559509	0
134	NC_072422.1	98	3273323	3271656	0
135	NC_072422.1	97	5772894	5771224	0
136	NC_072422.1	96	7554186	7552516	0
137	NC_072422.1	96	7575594	7577264	0
138	NC_072422.1	95	10381329	10382999	0
139	NC_072422.1	95	6831881	6830211	0
140	NC_072422.1	93	6860950	6859280	0
141	NC_072422.1	95	8629905	8631395	0
142	NC_072422.1	93	6907776	6906355	0

143	NC_072422.1	95	3722130	3721105	0
144	NC_072422.1	89	35848181	35849014	0
145	NC_072422.1	89	35758329	35759162	0
146	NC_072422.1	89	35774509	35775342	0
147	NC_072422.1	89	35808399	35809232	0
148	NC_072422.1	89	35720660	35721493	0
149	NC_072422.1	97	10628493	10627621	0
150	NC_072422.1	95	10629212	10628490	0
151	NC_072422.1	93	35759137	35759997	0
152	NC_072422.1	93	35775317	35776177	0
153	NC_072422.1	93	35809207	35810067	0
154	NC_072422.1	93	35848989	35849849	0
155	NC_072422.1	92	35721468	35722328	0
156	NC_072422.1	90	3722708	3722175	0
157	NC_072422.1	94	35873297	35874205	6.44E-149
158	NC_072422.1	81	35872623	35873369	1.03E-129
159	NC_072416.1	98	7294560	7292890	0
160	NC_072416.1	99	44571662	44569992	0
161	NC_072416.1	96	86678	88348	0
162	NC_072416.1	95	43369153	43367714	0
163	NC_072416.1	97	823417	822068	0
164	NC_072418.1	98	40371297	40369627	0
165	NC_072418.1	97	35076840	35075170	0
166	NC_072418.1	94	26783503	26782064	0
167	NC_072418.1	93	40388465	40389712	0
168	NC_072418.1	89	40420867	40421631	0
169	NC_072418.1	93	40421613	40422395	0
170	NC_072420.1	98	12464109	12465779	0
171	NC_072420.1	98	44218046	44219716	0
172	NC_072411.1	96	65137	66807	0
173	NC_072411.1	98	1024814	1023465	0
174	NC_072411.1	98	4994818	4996167	0
175	NC_072411.1	98	12731511	12730906	4.32E-97
176	NC_072407.1	97	63611353	63610010	0
177	NC_072413.1	96	17183705	17184772	0

Table 2.2 CERV1 env gene locations identified in the *Papio anubis* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_044987.1	97	69751617	69753287	0
2	NC_044987.1	95	80750254	80748584	0
3	NC_044987.1	88	4554482	4553328	0
4	NC_044987.1	95	4553298	4552810	0
5	NC_044979.1	97	145520540	145522210	0

6	NC_044979.1	91	76132784	76131117	0
7	NC_044979.1	90	86685237	86686907	0
8	NC_044979.1	94	9187252	9186176	0
9	NC_044979.1	85	9187797	9187243	0
10	NC_044979.1	92	140276336	140275659	4.61E-118
11	NC_044978.1	97	70499061	70497391	0
12	NC_044978.1	90	61386402	61384732	0
13	NC_044978.1	91	184673093	184671720	0
14	NC_044978.1	93	13347241	13348005	3.42E-116
15	NW_022163318.1	96	1427	3097	0
16	NW_022165198.1	96	13842	12172	0
17	NC_044977.1	96	15573561	15575231	0
18	NC_044977.1	90	79480504	79478834	0
19	NC_044977.1	92	192097353	192095866	0
20	NC_044977.1	87	73457093	73455837	0
21	NC_044977.1	93	130560976	130562052	2.80E-180
22	NC_044977.1	92	51968830	51969651	3.66E-127
23	NC_044977.1	93	167030041	167029313	8.75E-107
24	NC_044980.1	95	156032021	156033691	0
25	NC_044980.1	95	155916348	155918018	0
26	NC_044980.1	91	134743758	134742088	0
27	NC_044980.1	90	155887793	155889463	0
28	NC_044980.1	84	169035375	169036334	3.89E-169
29	NC_044980.1	93	169036831	169037553	2.26E-102
30	NW_022167358.1	95	26539	24869	0
31	NW_022161486.1	92	11580	9910	0
32	NC_044996.1	92	96027437	96029107	0
33	NC_044996.1	92	9188082	9186412	0
34	NC_044996.1	91	65087665	65085995	0
35	NC_044996.1	91	63328246	63329916	0
36	NC_044996.1	90	33799215	33798064	0
37	NC_044996.1	94	77706079	77707251	0
38	NC_044996.1	95	33798034	33797546	0
39	NC_044982.1	92	14996222	14994552	0
40	NC_044982.1	90	19446046	19447716	0
41	NC_044982.1	89	131238891	131240561	0
42	NC_044982.1	83	81916073	81914544	0
43	NC_044982.1	91	95086525	95085653	8.36E-138
44	NC_044982.1	89	72059535	72058666	3.29E-131
45	NC_044982.1	92	41224460	41223639	6.17E-125
46	NC_044982.1	93	117755172	117755819	4.75E-123
47	NC_044982.1	94	139893716	139894204	3.02E-86
48	NC_044976.1	92	119624399	119626069	0
49	NC_044976.1	91	118026390	118028060	0
50	NC_044976.1	91	124623833	124625503	0

51	NC_044976.1	88	46697083	46698564	0
52	NC_044976.1	89	77316580	77315429	0
53	NC_044976.1	93	77315397	77314909	0
54	NC_044976.1	92	49327971	49328699	8.10E-107
55	NC_044986.1	92	108406298	108404628	0
56	NC_044986.1	91	86494607	86495428	4.04E-126
57	NC_044986.1	91	18288523	18289260	6.91E-124
58	NW_022161054.1	90	213215	211545	0
59	NC_044990.1	92	39324115	39322445	0
60	NC_044990.1	90	10779356	10777767	0
61	NC_044990.1	89	87668697	87669848	0
62	NC_044990.1	94	75933768	75932647	0
63	NC_044990.1	92	87669880	87670368	0
64	NW_022162457.1	90	10537	12159	0
65	NW_022162021.1	90	18040	19710	0
66	NC_044981.1	90	99494848	99496515	0
67	NC_044981.1	91	108896325	108895798	3.20E-98
68	NW_022160492.1	94	4783	6375	0
69	NC_044988.1	93	2950419	2948827	0
70	NC_044988.1	86	44580134	44580907	2.50E-148
71	NC_044988.1	92	44580871	44581743	4.16E-141
72	NC_044988.1	92	7629558	7630379	3.29E-127
73	NW_022159734.1	89	12031	10361	0
74	NC_044993.1	90	63466226	63464661	0
75	NC_044993.1	93	46008145	46009266	0
76	NC_044993.1	86	46007597	46008154	0
77	NC_044993.1	95	26430453	26429653	8.35E-127
78	NC_044989.1	89	90445254	90443584	0
79	NC_044989.1	88	51571633	51570617	0
80	NC_044989.1	91	51570630	51570154	0
81	NC_044989.1	92	13403772	13402885	6.20E-142
82	NC_044989.1	93	55616361	55616849	1.81E-117
83	NW_022165816.1	91	34	1566	0
84	NC_044997.1	87	6620579	6622249	0
85	NC_044997.1	89	5106317	5107186	1.07E-131
86	NW_022165809.1	91	34	1479	0
87	NW_022162557.1	90	3407	4852	0
88	NW_022162103.1	89	6579	7967	0
89	NW_022166797.1	90	1	1359	0
90	NW_022167311.1	95	65	1285	0
91	NW_022165504.1	89	1316	3	0
92	NW_022166709.1	92	1223	3	0
93	NW_022161704.1	89	10928	9651	0
94	NC_044985.1	92	22331082	22329856	0
95	NC_044985.1	94	117055062	117054535	4.04E-117

96	NC_044985.1	93	23596933	23597661	7.03E-105
97	NC_044983.1	88	43707824	43706673	0
98	NC_044983.1	96	43706641	43706153	0
99	NW_022163909.1	91	1474	2448	0
100	NW_022159724.1	88	11607	10393	0
101	NW_022159724.1	81	12062	11517	3.17E-86
102	NC_044995.1	87	7706373	7707470	0
103	NC_044995.1	91	7707446	7708039	0
104	NW_022162395.1	88	4241	5215	0
105	NW_022162395.1	94	5256	5903	0
106	NW_022164371.1	84	1191	2165	0
107	NW_022166015.1	91	1517	672	0
108	NW_022166015.1	90	640	152	0
109	NW_022166390.1	91	855	1	0
110	NW_022161442.1	82	5745	6425	0
111	NW_022161442.1	90	6475	7191	0
112	NW_022162630.1	80	2527	3207	0
113	NW_022162630.1	91	3707	4195	0
114	NW_022162630.1	86	3192	3677	0
115	NW_022164099.1	90	1014	1628	0
116	NW_022164099.1	90	1631	2251	0
117	NW_022161136.1	97	17995	16982	3.16E-176
118	NW_022161136.1	84	18635	17916	2.93E-122
119	NW_022167021.1	97	2754	1741	3.23E-176
120	NW_022167021.1	84	3394	2675	3.00E-122
121	NW_022167928.1	86	291	1064	4.51E-165
122	NC_044984.1	90	4537439	4538452	8.43E-163
123	NC_044984.1	84	4536799	4537467	2.67E-116
124	NC_044984.1	98	47743744	47743256	9.05E-81
125	NW_022169733.1	97	704	6	9.01E-159
126	NW_022167014.1	93	872	3	6.30E-154
127	NW_022167014.1	87	1326	832	8.58E-93
128	NW_022160125.1	88	1401	391	4.94E-153
129	NW_022165738.1	91	1586	747	2.93E-143
130	NW_022163357.1	91	1222	362	1.35E-135
131	NW_022169114.1	94	468	1085	3.35E-134
132	NW_022164381.1	92	1567	2160	1.22E-119
133	NW_022162003.1	91	6429	5941	1.17E-114
134	NW_022165539.1	85	1039	1641	2.97E-113
135	NW_022161247.1	89	737593	736910	5.99E-113
136	NW_022165717.1	92	1591	1031	7.18E-112
137	NW_022168286.1	82	543	1	3.61E-105
138	NW_022167789.1	94	1037	549	7.93E-91
139	NW_022166656.1	93	547	2	9.75E-88
140	NW_022167020.1	94	133781	133293	3.19E-86

141	NW_022165929.1	88	48109	48672	5.22E-84
142	NW_022167925.1	94	11082	11570	1.82E-76
143	NW_022161959.1	96	728	255	7.56E-76
144	NW_022160135.1	88	470	3	7.24E-71

Table 2.2 CERV1 env gene locations identified in the *Theropithecus gelada* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_037683.1	97	75051215	75049545	0
2	NC_037683.1	91	107463265	107464935	0
3	NC_037683.1	93	108957191	108956040	0
4	NC_037683.1	94	108956008	108955520	0
5	NC_037683.1	93	24191511	24190645	2.67E-139
6	NC_037683.1	82	90025131	90024265	2.26E-116
7	NW_020258762.1	97	3403	5073	0
8	NW_020258872.1	97	1837	167	0
9	NW_020258871.1	97	3875	5545	0
10	NC_037672.1	97	116115533	116117203	0
11	NC_037672.1	96	77974926	77976596	0
12	NC_037672.1	90	182897	184567	0
13	NC_037672.1	97	86211078	86212532	0
14	NC_037672.1	88	101484283	101485785	0
15	NC_037672.1	84	29211593	29210442	0
16	NC_037672.1	97	186458899	186459723	0
17	NC_037672.1	97	186460078	186460566	0
18	NC_037672.1	91	29210409	29209921	0
19	NC_037672.1	86	73021330	73022247	1.11E-132
20	NC_037672.1	92	173286049	173285285	6.36E-115
21	NC_037672.1	89	67206641	67205994	1.09E-99
22	NC_037668.1	97	13559466	13557796	0
23	NC_037668.1	96	191304398	191306068	0
24	NC_037668.1	92	98226204	98224534	0
25	NC_037668.1	92	14645875	14647545	0
26	NC_037668.1	91	141423731	141422061	0
27	NC_037668.1	96	26819253	26820746	0
28	NC_037668.1	90	22715532	22713862	0
29	NC_037668.1	88	76110107	76111591	0
30	NC_037668.1	93	55167128	55168393	0
31	NC_037668.1	89	106850907	106849756	0
32	NC_037668.1	87	116655095	116654124	0
33	NC_037668.1	95	19435078	19436022	0
34	NC_037668.1	96	116654077	116653430	0
35	NC_037668.1	88	19434534	19435079	0
36	NC_037668.1	92	106849724	106849236	0

37	NC_037668.1	92	22915921	22916649	9.52E-105
38	NW_020265518.1	97	6343	8013	0
39	NW_020265529.1	97	6342	8012	0
40	NW_020262659.1	97	1909	3579	0
41	NW_020262659.1	97	7662	9332	0
42	NC_037670.1	96	72847570	72849240	0
43	NC_037670.1	95	97992724	97991054	0
44	NC_037670.1	93	157991835	157990165	0
45	NC_037670.1	92	13490861	13489239	0
46	NC_037670.1	90	114004937	114003267	0
47	NC_037670.1	90	96302174	96300504	0
48	NC_037670.1	92	106802578	106804143	0
49	NC_037670.1	96	1786340	1785069	0
50	NC_037670.1	89	7385868	7387019	0
51	NC_037670.1	96	7387052	7387540	0
52	NC_037670.1	92	36741238	36742104	7.57E-138
53	NC_037670.1	94	150256005	150255277	3.77E-108
54	NC_037689.1	96	46933706	46932036	0
55	NC_037689.1	96	59175413	59173743	0
56	NC_037689.1	90	9650306	9648636	0
57	NC_037689.1	91	91700602	91702095	0
58	NC_037689.1	92	79157212	79155881	0
59	NC_037689.1	95	105809253	105809741	4.64E-117
60	NW_020259314.1	97	1775	105	0
61	NW_020264637.1	96	6516	8186	0
62	NC_037671.1	96	106205014	106206684	0
63	NC_037671.1	96	118526860	118528530	0
64	NC_037671.1	91	77101382	77099712	0
65	NC_037671.1	90	103330598	103332220	0
66	NC_037671.1	95	32943162	32942110	0
67	NC_037671.1	95	32934975	32933923	0
68	NC_037671.1	90	111712034	111710994	0
69	NC_037671.1	92	111711000	111710365	0
70	NC_037671.1	96	32942111	32941497	0
71	NC_037671.1	96	32933924	32933310	0
72	NC_037671.1	92	32691861	32690992	3.35E-133
73	NC_037671.1	91	108736168	108736827	3.35E-101
74	NC_037677.1	96	81363346	81365016	0
75	NC_037677.1	90	12271778	12273364	0
76	NC_037669.1	95	151984462	151986132	0
77	NC_037669.1	92	7641696	7643366	0
78	NC_037669.1	91	192568627	192566957	0
79	NC_037669.1	90	147553621	147555291	0
80	NC_037669.1	90	192592987	192591500	0
81	NC_037669.1	93	125867142	125868446	0

82	NC_037669.1	94	88627114	88628406	0
83	NC_037669.1	86	73300431	73299145	0
84	NC_037669.1	93	130852303	130853379	1.90E-179
85	NC_037669.1	97	113960796	113959936	3.07E-140
86	NC_037669.1	93	72606564	72605704	5.35E-136
87	NC_037669.1	93	128756000	128756860	9.30E-135
88	NC_037669.1	91	126399299	126398475	6.28E-124
89	NC_037669.1	86	28211887	28212519	1.54E-111
90	NC_037669.1	95	28212575	28213219	2.64E-107
91	NW_020252566.1	95	6016	7638	0
92	NW_020259513.1	96	1732	140	0
93	NW_020258771.1	93	3420	5090	0
94	NW_020258773.1	92	3419	5089	0
95	NW_020254939.1	93	4224	5894	0
96	NW_020254939.1	89	10212	11096	8.64E-166
97	NW_020259510.1	92	1761	3431	0
98	NW_020260032.1	92	2284	614	0
99	NW_020258587.1	92	3580	5250	0
100	NC_037673.1	91	138355522	138353852	0
101	NC_037673.1	90	93521326	93519656	0
102	NC_037673.1	90	159776082	159777752	0
103	NC_037673.1	96	166583637	166582318	0
104	NC_037673.1	85	172599242	172600630	0
105	NC_037673.1	95	120756435	120757118	1.19E-120
106	NC_037673.1	89	12015418	12016026	4.53E-113
107	NC_037676.1	92	54792694	54794364	0
108	NC_037676.1	92	59589118	59590788	0
109	NC_037676.1	91	114955558	114957180	0
110	NC_037676.1	91	59609618	59611288	0
111	NC_037676.1	90	42236784	42238454	0
112	NC_037676.1	90	2730301	2728622	0
113	NC_037676.1	90	41729136	41727514	0
114	NC_037676.1	89	50288938	50287838	0
115	NC_037676.1	95	50287806	50287318	0
116	NC_037676.1	93	133670857	133671678	2.56E-126
117	NW_020251718.1	92	6132	7802	0
118	NW_020251728.1	92	6131	7801	0
119	NW_020266101.1	92	5346	7016	0
120	NW_020256217.1	91	5455	7125	0
121	NW_020255614.1	91	1763	93	0
122	NC_037675.1	92	48857919	48859589	0
123	NC_037675.1	92	90114283	90115953	0
124	NC_037675.1	88	49126965	49128635	0
125	NC_037675.1	83	25908380	25909909	0
126	NC_037675.1	89	35499348	35500217	5.30E-131

127	NC_037675.1	93	66638878	66639606	1.29E-105
128	NW_020258583.1	92	3582	5252	0
129	NW_020259226.1	92	1941	271	0
130	NC_037687.1	90	20782197	20780527	0
131	NC_037680.1	92	20436839	20438509	0
132	NC_037680.1	92	17545605	17544391	0
133	NC_037680.1	90	18719474	18718434	0
134	NC_037680.1	90	18718437	18717805	0
135	NC_037680.1	92	18819071	18819799	6.30E-107
136	NW_020256117.1	91	5407	7029	0
137	NW_020259151.1	91	3609	5279	0
138	NW_020259213.1	94	3142	4764	0
139	NW_020257591.1	92	5612	7282	0
140	NC_037678.1	91	29460930	29459260	0
141	NC_037678.1	91	29455119	29453626	0
142	NC_037678.1	91	68575912	68577204	0
143	NC_037678.1	98	68583674	68582574	0
144	NC_037678.1	94	68584243	68583668	0
145	NC_037678.1	88	20002341	20002973	1.62E-108
146	NC_037678.1	93	20003295	20003792	2.32E-78
147	NW_020258710.1	91	1857	187	0
148	NW_020259170.1	91	3334	5004	0
149	NC_037679.1	90	100182585	100180918	0
150	NC_037679.1	90	109900422	109902092	0
151	NC_037679.1	91	109057345	109055675	0
152	NC_037679.1	91	126485911	126486558	2.56E-101
153	NW_020258327.1	90	2138	468	0
154	NC_037685.1	91	91416375	91414705	0
155	NC_037685.1	89	10471727	10470231	0
156	NC_037685.1	93	36613597	36612383	0
157	NW_020253500.1	90	1826	204	0
158	NW_020261833.1	91	4357	6027	0
159	NW_020258451.1	91	2508	838	0
160	NW_020259736.1	91	2729	4399	0
161	NW_020253039.1	91	1810	140	0
162	NW_020251647.1	90	1952	282	0
163	NC_037681.1	92	33980811	33979141	0
164	NC_037681.1	89	40345516	40346724	0
165	NC_037681.1	92	25151512	25150646	2.73E-138
166	NC_037681.1	85	100050401	100049751	3.39E-93
167	NC_037681.1	94	40347025	40347498	4.35E-73
168	NW_020258864.1	90	3961	5631	0
169	NW_020255603.1	89	5517	7187	0
170	NW_020258970.1	90	3829	5499	0
171	NW_020265361.1	90	2634	4304	0

172	NW_020258975.1	90	3825	5495	0
173	NC_037686.1	97	61474854	61473382	0
174	NC_037686.1	90	39726370	39726984	0
175	NC_037686.1	91	39726987	39727547	0
176	NW_020255074.1	97	1983	511	0
177	NW_020259486.1	91	1639	2	0
178	NW_020258992.1	89	1773	103	0
179	NW_020258095.1	88	2064	394	0
180	NC_037674.1	88	32670994	32669324	0
181	NC_037688.1	90	72222798	72221233	0
182	NC_037688.1	94	21325897	21326472	3.48E-108
183	NC_037688.1	95	3263616	3263128	4.60E-79
184	NW_020252996.1	90	5950	7620	0
185	NW_020261166.1	89	5803	7473	0
186	NW_020258825.1	88	3954	5624	0
187	NW_020254058.1	92	1902	457	0
188	NW_020259034.1	96	4112	5509	0
189	NW_020252103.1	95	6364	7695	0
190	NW_020252830.1	91	1760	495	0
191	NC_037682.1	93	121370868	121372136	0
192	NC_037682.1	94	65069414	65068176	0
193	NC_037682.1	89	23297208	23296525	5.67E-114
194	NW_020258301.1	92	5124	6392	0
195	NW_020258296.1	92	5128	6396	0

Table 2.2 CERV1 env gene locations identified in the *Mandrillus leucophaeus* genome

No.	Acc.	Iden.	Start	End
1	NW_012099765.1	96	88308	89978
2	NW_012101987.1	92	2461686	2463356
3	NW_012099410.1	88	59146	60801
4	NW_012107563.1	88	1103167	1101932
5	NW_012110264.1	89	264284	265600
6	NW_012105087.1	94	1003440	1004414
7	NW_012103843.1	91	6373	7347
8	NW_012103843.1	91	7445	8041
9	NW_012101598.1	93	90764	91882
10	NW_012101598.1	82	90219	90764
11	NW_012106764.1	91	1243649	1242762
12	NW_012105388.1	92	1530253	1529348
13	NW_012107731.1	87	1138733	1137807
14	NW_012109086.1	86	73330	72356
15	NW_012101365.1	91	150061	150612
16	NW_012100432.1	88	771125	770256

17	NW_012106386.1	92	1303226	1303774
18	NW_012098399.1	87	599759	598926
19	NW_012098399.1	90	598962	598090
20	NW_012104698.1	95	6715837	6716628
21	NW_012102809.1	89	3139235	3140011
22	NW_012102809.1	88	3142125	3142610
23	NW_012106286.1	86	1350391	1351179
24	NW_012103599.1	93	1084843	1085709
25	NW_012105043.1	91	1467725	1468591
26	NW_012106397.1	87	180598	179861
27	NW_012105864.1	92	754080	753502
28	NW_012107141.1	90	81119	81802
29	NW_012107552.1	92	174049	174822
30	NW_012103399.1	86	1716626	1716072
31	NW_012108864.1	92	999716	998982
32	NW_012105509.1	93	38	640
33	NW_012104354.1	94	2013447	2013935
34	NW_012108142.1	89	23210	23806
35	NW_012102821.1	89	520738	521226

Table 2.2 CERV1 env gene locations identified in the *Macaca thibetana* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_065588.1	97	91796391	91794721	0
2	NC_065588.1	96	56507554	56509224	0
3	NC_065588.1	90	30711446	30709776	0
4	NC_065588.1	91	66621233	66619866	0
5	NC_065588.1	94	78545993	78547369	0
6	NC_065588.1	91	47183420	47182083	0
7	NC_065588.1	93	116169477	116168158	0
8	NC_065588.1	95	34312185	34311001	0
9	NC_065588.1	94	116027877	116028965	0
10	NC_065588.1	90	31811341	31812189	0
11	NC_065588.1	83	116027293	116027874	0
12	NC_065588.1	83	34312672	34312190	0
13	NC_065588.1	96	31812228	31812707	0
14	NC_065580.1	96	145344934	145343264	0
15	NC_065580.1	92	127280308	127281975	0
16	NC_065580.1	91	125888003	125886333	0
17	NC_065580.1	91	131598437	131600107	0
18	NC_065580.1	93	127495832	127497430	0
19	NC_065580.1	90	76597949	76596279	0
20	NC_065580.1	90	172479773	172478103	0
21	NC_065580.1	93	17163198	17164583	0

22	NC_065580.1	95	126124500	126125705	0
23	NC_065580.1	97	20987371	20988363	0
24	NC_065580.1	97	20986697	20987374	0
25	NC_065580.1	92	16409003	16409623	0
26	NC_065580.1	98	16409674	16410144	0
27	NC_065580.1	84	126124033	126124509	0
28	NC_065580.1	91	9318375	9317509	1.46E-136
29	NC_065580.1	91	110395208	110394624	5.29E-84
30	NC_065582.1	96	11819860	11818190	0
31	NC_065582.1	92	40898524	40900194	0
32	NC_065582.1	92	103810014	103811684	0
33	NC_065582.1	92	103787482	103789152	0
34	NC_065582.1	90	61603811	61602141	0
35	NC_065582.1	91	135404172	135405845	0
36	NC_065582.1	91	109144479	109142809	0
37	NC_065582.1	91	48727858	48729387	0
38	NC_065582.1	92	37021857	37023203	0
39	NC_065582.1	89	122039123	122038542	1.58E-83
40	NC_065598.1	97	85009374	85011044	0
41	NC_065598.1	92	133409359	133411029	0
42	NC_065598.1	92	29827216	29828886	0
43	NC_065598.1	91	99449577	99447907	0
44	NC_065598.1	90	35885980	35884643	0
45	NC_065598.1	91	139944676	139944092	1.95E-85
46	NC_065578.1	96	175239498	175241171	0
47	NC_065578.1	94	119675082	119676752	0
48	NC_065578.1	91	174714702	174713032	0
49	NC_065578.1	90	218157344	218155674	0
50	NC_065578.1	94	218869987	218871003	0
51	NC_065578.1	92	119012793	119013653	5.66E-133
52	NC_065578.1	93	134362251	134362739	3.63E-115
53	NC_065578.1	94	50085055	50085528	1.03E-91
54	NC_065578.1	93	129887350	129886862	6.31E-77
55	NC_065584.1	96	107948283	107949953	0
56	NC_065584.1	92	33480152	33481822	0
57	NC_065584.1	92	42551289	42552959	0
58	NC_065584.1	95	116949462	116950766	0
59	NC_065584.1	92	110988177	110989424	0
60	NC_065584.1	89	139504430	139505581	0
61	NC_065584.1	94	139505611	139506099	0
62	NC_065584.1	90	66168672	66168001	1.07E-123
63	NC_065594.1	96	10214213	10212543	0
64	NC_065594.1	92	11425136	11423466	0
65	NC_065594.1	92	27701933	27700263	0
66	NC_065594.1	91	8342700	8341030	0

67	NC_065594.1	92	89124484	89126154	0
68	NC_065594.1	89	77339387	77338074	0
69	NC_065594.1	95	30589798	30590286	2.00E-118
70	NC_065594.1	94	48859714	48860202	5.43E-78
71	NC_065599.1	96	4061097	4059427	0
72	NC_065599.1	95	5776635	5778305	0
73	NC_065599.1	89	1694120	1695790	0
74	NC_065599.1	87	5460338	5459562	0
75	NC_065599.1	93	5459563	5458718	0
76	NC_065586.1	96	45570772	45572394	0
77	NC_065586.1	91	43336046	43337620	0
78	NC_065586.1	91	86549608	86548448	0
79	NC_065586.1	85	71337390	71336221	0
80	NC_065586.1	95	71337892	71337392	0
81	NC_065586.1	92	58732474	58731608	1.28E-139
82	NC_065591.1	93	65050041	65051711	0
83	NC_065591.1	92	42455571	42453901	0
84	NC_065591.1	91	85787093	85785600	0
85	NC_065591.1	91	69667522	69668451	1.06E-149
86	NC_065591.1	86	69666786	69667592	3.90E-143
87	NC_065591.1	90	23212052	23211369	2.36E-115
88	NC_065579.1	92	12679303	12680973	0
89	NC_065579.1	90	32136528	32138198	0
90	NC_065579.1	92	2121855	2123516	0
91	NC_065579.1	91	2145913	2147574	0
92	NC_065579.1	90	47304187	47305857	0
93	NC_065579.1	89	76496947	76498617	0
94	NC_065579.1	91	66319026	66320633	0
95	NC_065579.1	82	134000409	134002079	0
96	NC_065579.1	92	145308954	145310339	0
97	NC_065579.1	89	171184864	171186240	0
98	NC_065579.1	95	190280121	190281329	0
99	NC_065579.1	81	49371	50687	0
100	NC_065579.1	93	65954935	65954075	1.31E-136
101	NC_065579.1	86	48694	49329	9.61E-112
102	NC_065590.1	92	72525077	72523407	0
103	NC_065590.1	92	54044124	54042454	0
104	NC_065590.1	92	76482294	76480624	0
105	NC_065590.1	89	102421992	102423662	0
106	NC_065590.1	91	93574224	93575894	0
107	NC_065590.1	88	18910322	18911659	0
108	NC_065590.1	92	54695231	54693915	0
109	NC_065590.1	88	27931294	27932445	0
110	NC_065590.1	96	24159097	24158192	0
111	NC_065590.1	84	24159848	24159372	0

112	NC_065590.1	90	27932477	27932965	0
113	NC_065590.1	85	94009180	94010022	4.54E-151
114	NC_065590.1	96	94009991	94010851	2.77E-142
115	NC_065590.1	84	53352303	53351641	5.87E-117
116	NC_065585.1	91	88164993	88166663	0
117	NC_065585.1	91	92124942	92126615	0
118	NC_065585.1	92	9553726	9555396	0
119	NC_065585.1	90	100005515	100003845	0
120	NC_065585.1	90	96513213	96514835	0
121	NC_065585.1	89	87551937	87553607	0
122	NC_065585.1	92	2393390	2394835	0
123	NC_065585.1	91	22745253	22743940	0
124	NC_065585.1	89	96570645	96571448	0
125	NC_065585.1	95	96571682	96572170	0
126	NC_065585.1	88	89073423	89072557	2.46E-125
127	NC_065585.1	91	51204062	51204883	2.74E-123
128	NC_065585.1	91	88983983	88983162	6.28E-117
129	NC_065592.1	92	87148838	87150508	0
130	NC_065592.1	92	97463798	97465468	0
131	NC_065592.1	96	103615189	103615662	5.69E-106
132	NC_065592.1	93	102265215	102264568	1.04E-101
133	NC_065592.1	80	75452916	75452215	2.24E-78
134	NC_065597.1	92	4441732	4440062	0
135	NC_065597.1	86	45304180	45302690	0
136	NC_065597.1	93	1954594	1953266	0
137	NC_065597.1	93	5944387	5945253	2.16E-140
138	NC_065581.1	91	27853058	27854728	0
139	NC_065581.1	92	27027406	27025784	0
140	NC_065581.1	97	124019773	124020921	0
141	NC_065581.1	89	33418910	33419788	0
142	NC_065581.1	96	31667006	31667851	0
143	NC_065581.1	90	31666531	31667007	0
144	NC_065581.1	92	33420008	33420571	0
145	NC_065581.1	95	124020920	124021444	0
146	NC_065581.1	96	45354965	45354081	3.70E-147
147	NC_065581.1	83	45355750	45354917	5.73E-143
148	NC_065581.1	91	31412713	31411811	2.94E-138
149	NC_065581.1	90	31344141	31343239	3.54E-137
150	NC_065581.1	92	83597827	83596961	3.65E-137
151	NC_065581.1	93	20977324	20976515	2.25E-127
152	NC_065595.1	92	13395967	13394294	0
153	NC_065595.1	86	11360336	11361169	0
154	NC_065595.1	96	11361142	11361780	0
155	NC_065595.1	94	51545206	51546156	5.17E-157
156	NC_065595.1	93	43710554	43711282	4.95E-108

157	NC_065583.1	91	73835471	73837141	0
158	NC_065583.1	91	54945376	54943958	0
159	NC_065583.1	95	168126336	168127298	0
160	NC_065583.1	88	168125629	168126333	0
161	NC_065583.1	92	106486475	106487293	2.76E-125
162	NC_065583.1	95	47394703	47394215	4.69E-117
163	NC_065587.1	91	44808769	44810430	0
164	NC_065587.1	87	26106043	26107704	0
165	NC_065587.1	89	34718936	34717467	0
166	NC_065587.1	90	20367861	20368976	0
167	NW_026088941.1	91	40788	42398	0
168	NC_065589.1	89	101672740	101671073	0
169	NC_065589.1	89	102905578	102903911	0
170	NC_065589.1	87	100179771	100181267	0
171	NC_065589.1	96	29308981	29309469	1.21E-118
172	NC_065593.1	91	70418169	70416760	0
173	NC_065596.1	89	763730	762579	0
174	NC_065596.1	93	762543	762055	0
175	NW_026089201.1	95	1152	1	0

Table 2.2 CERV1 env gene locations identified in the *Macaca fascicularis* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_052268.1	97	101564604	101562934	0
2	NC_052268.1	92	59100091	59101761	0
3	NC_052268.1	91	83358953	83360644	0
4	NC_052268.1	91	93765368	93763698	0
5	NC_052268.1	89	102375911	102376594	1.08E-114
6	NC_052261.1	97	166613093	166611423	0
7	NC_052261.1	97	58857803	58856133	0
8	NC_052261.1	92	121455514	121453844	0
9	NC_052261.1	89	30561008	30562678	0
10	NC_052261.1	89	151385905	151384475	0
11	NC_052261.1	86	100733130	100733855	7.90E-124
12	NC_052259.1	97	174223931	174225601	0
13	NC_052259.1	91	78198723	78200393	0
14	NC_052259.1	91	239927	241597	0
15	NC_052259.1	91	148959242	148957572	0
16	NC_052259.1	90	76732699	76734369	0
17	NC_052259.1	90	137261226	137259559	0
18	NC_052259.1	89	124372981	124374348	0
19	NC_052259.1	92	100956893	100958152	0
20	NC_052255.1	97	92360046	92358376	0
21	NC_052255.1	96	45183803	45182133	0

22	NC_052255.1	94	102518793	102517123	0
23	NC_052255.1	92	103202366	103204036	0
24	NC_052255.1	92	101572257	101573927	0
25	NC_052255.1	90	2306352	2308022	0
26	NC_052255.1	90	131540200	131541693	0
27	NC_052255.1	97	126124475	126125764	0
28	NC_052255.1	89	131536989	131538116	0
29	NC_052255.1	93	2639916	2639029	2.26E-143
30	NC_052255.1	82	2640698	2639865	8.44E-141
31	NC_052255.1	94	103184808	103183948	2.61E-137
32	NC_052255.1	92	28568293	28569159	3.76E-137
33	NC_052255.1	93	86143846	86143358	1.54E-114
34	NC_052255.1	94	172161484	172161014	1.59E-91
35	NC_052265.1	96	92440839	92439169	0
36	NC_052265.1	94	79153654	79155324	0
37	NC_052265.1	91	67232727	67231105	0
38	NC_052265.1	86	116810107	116808914	0
39	NC_052265.1	89	31743480	31744328	0
40	NC_052265.1	95	31744367	31744846	0
41	NC_052265.1	95	31618702	31618244	1.47E-89
42	NC_052265.1	96	116808695	116808219	3.01E-77
43	NC_052263.1	96	45314277	45315947	0
44	NC_052263.1	93	64271764	64273434	0
45	NC_052263.1	91	52052087	52053757	0
46	NC_052263.1	98	70985750	70984509	0
47	NC_052263.1	92	58372149	58371283	8.56E-138
48	NC_052275.1	96	84217160	84218830	0
49	NC_052275.1	92	34952044	34950374	0
50	NC_052275.1	91	108283866	108285536	0
51	NC_052275.1	91	132667534	132668943	0
52	NC_052275.1	91	98706784	98705633	0
53	NC_052275.1	88	90026113	90027168	0
54	NC_052275.1	95	98705600	98705112	0
55	NC_052275.1	88	90027126	90027722	2.64E-84
56	NC_052258.1	96	123245383	123247053	0
57	NC_052258.1	91	158423631	158421961	0
58	NC_052258.1	90	44762308	44761157	0
59	NC_052258.1	94	44761125	44760637	0
60	NC_052258.1	91	31161892	31162794	2.92E-138
61	NC_052258.1	91	83048359	83047493	2.61E-135
62	NC_052256.1	95	60288396	60286726	0
63	NC_052256.1	92	14400448	14398778	0
64	NC_052256.1	91	136548653	136550323	0
65	NC_052256.1	91	192362515	192360845	0
66	NC_052256.1	90	147160444	147158774	0

67	NC_052256.1	91	192385970	192384312	0
68	NC_052256.1	89	23159184	23157514	0
69	NC_052256.1	91	49020906	49019521	0
70	NC_052256.1	90	117663294	117661957	0
71	NC_052256.1	96	136557402	136558616	0
72	NC_052256.1	94	136556947	136557399	0
73	NC_052256.1	93	128517837	128518697	6.25E-135
74	NC_052269.1	92	36313785	36312115	0
75	NC_052269.1	92	60796463	60794793	0
76	NC_052269.1	91	18494874	18496544	0
77	NC_052269.1	91	88172057	88170387	0
78	NC_052269.1	93	8052963	8053691	1.65E-105
79	NC_052271.1	92	50298789	50297119	0
80	NC_052271.1	90	77334169	77332547	0
81	NC_052271.1	91	89167615	89169285	0
82	NC_052271.1	90	38968275	38969945	0
83	NC_052271.1	89	30527081	30527602	4.19E-116
84	NC_052267.1	92	24549796	24551466	0
85	NC_052267.1	92	72397798	72396128	0
86	NC_052267.1	90	76367468	76365798	0
87	NC_052267.1	91	93481910	93483580	0
88	NC_052267.1	92	44480224	44481813	0
89	NC_052267.1	88	102338053	102339723	0
90	NC_052267.1	89	11651224	11652894	0
91	NC_052267.1	92	54540809	54539493	0
92	NC_052267.1	89	27804031	27805182	0
93	NC_052267.1	91	27805214	27805702	0
94	NC_052266.1	97	27598895	27597303	0
95	NC_052266.1	90	70736072	70737736	0
96	NC_052266.1	88	26093013	26094683	0
97	NC_052266.1	89	24854127	24855794	0
98	NC_052266.1	93	79812401	79810998	0
99	NC_052260.1	92	95034047	95035717	0
100	NC_052260.1	92	103149494	103151161	0
101	NC_052260.1	94	103014451	103013135	0
102	NC_052260.1	93	152050831	152049689	0
103	NC_052260.1	98	158438571	158439686	0
104	NC_052260.1	92	158438017	158438568	0
105	NC_052260.1	93	70375134	70374313	1.53E-127
106	NC_052260.1	96	129457625	129458113	8.98E-118
107	NC_052260.1	94	147554284	147553796	1.70E-117
108	NC_052273.1	92	20768092	20769762	0
109	NC_052273.1	87	22670772	22671464	1.67E-169
110	NC_052273.1	93	22671500	22671988	1.67E-169
111	NC_052273.1	84	22670317	22670829	2.81E-85

112	NC_052272.1	92	26177053	26175383	0
113	NC_052272.1	92	51657374	51658324	1.66E-152
114	NC_052272.1	89	43820833	43821558	1.70E-104
115	NC_052262.1	92	28964777	28963107	0
116	NC_052262.1	91	125400885	125399215	0
117	NC_052262.1	89	46231838	46230411	0
118	NC_052262.1	88	46289013	46287910	0
119	NC_052262.1	95	111631516	111630629	0
120	NC_052262.1	96	111630632	111629847	0
121	NC_052262.1	95	46287878	46287390	0
122	NC_052262.1	90	54622669	54621794	8.53E-171
123	NC_052262.1	92	91647003	91646182	3.02E-124
124	NC_052262.1	93	53811476	53812297	1.37E-122
125	NC_052262.1	96	53720391	53719903	8.70E-117
126	NC_052262.1	86	235141	234506	9.54E-112
127	NC_052262.1	83	54623124	54622609	1.43E-83
128	NC_052270.1	91	59247445	59249115	0
129	NC_052270.1	89	61842779	61843930	0
130	NC_052270.1	93	61843962	61844450	0
131	NW_025540891.1	91	19510	21180	0
132	NC_052257.1	91	76290616	76288946	0
133	NC_052257.1	91	176089605	176088145	0
134	NC_052257.1	90	166251100	166252419	0
135	NC_052257.1	95	183954653	183955900	0
136	NC_052257.1	97	20936908	20937900	3.16E-173
137	NC_052257.1	87	20936234	20937007	3.71E-138
138	NC_052257.1	91	9296707	9295841	6.11E-136
139	NC_052257.1	91	125667822	125666962	3.85E-131
140	NC_052274.1	86	33684961	33686451	0
141	NC_052274.1	91	49694331	49695206	1.99E-140
142	NC_052274.1	94	33008656	33009144	3.35E-78
143	NC_052264.1	89	80861676	80862791	0
144	NC_052264.1	84	86596219	86597247	2.00E-146

Table 2.2 CERV1 env gene locations identified in the *Rhinopithecus roxellana* genome

No.	Acc.	Iden.	Start	End	E-value
1	NC_044550.1	89	10341751	10340090	0
2	NC_044550.1	85	47191515	47189866	0
3	NC_044550.1	84	111025602	111023923	0
4	NC_044550.1	82	113502514	113504160	0
5	NC_044550.1	83	150072581	150074230	0
6	NC_044550.1	80	152542012	152540417	0
7	NC_044550.1	85	91860762	91862441	0

8	NC_044550.1	81	61986959	61985304	0
9	NC_044550.1	88	188155418	188154633	0
10	NC_044550.1	88	143324086	143323217	1.57E-122
11	NC_044550.1	86	178809281	178808511	2.59E-107
12	NC_044551.1	82	124907730	124906081	0
13	NC_044551.1	84	78594191	78595840	0
14	NC_044551.1	82	38912237	38913886	0
15	NC_044551.1	82	125421216	125422865	0
16	NC_044551.1	83	70030018	70029128	0
17	NC_044551.1	85	85814403	85813489	1.73E-121
18	NC_044551.1	87	16368809	16368183	5.19E-89
19	NC_044554.1	82	1466457	1464808	0
20	NC_044554.1	82	37641296	37639647	0
21	NC_044554.1	81	52146340	52144793	0
22	NC_044554.1	83	55124067	55122595	0
23	NC_044554.1	89	2030097	2029381	0
24	NC_044554.1	86	79372648	79373199	0
25	NC_044560.1	81	129737540	129739132	0
26	NC_044560.1	82	30712362	30714011	0
27	NC_044560.1	82	128982305	128980656	0
28	NC_044560.1	83	125584773	125586422	0
29	NC_044560.1	86	129086189	129087160	0
30	NC_044556.1	83	88895931	88897580	0
31	NC_044556.1	83	54990038	54988389	0
32	NC_044556.1	83	87740905	87739328	0
33	NC_044556.1	87	50856961	50857674	5.54E-108
34	NC_044563.1	82	102279519	102281168	0
35	NC_044563.1	83	79115006	79116655	0
36	NC_044563.1	87	119653819	119652614	0
37	NC_044563.1	87	20612208	20611237	0
38	NC_044563.1	88	58688643	58687753	1.89E-135
39	NC_044563.1	86	79683062	79682211	7.21E-123
40	NC_044555.1	84	117499695	117501344	0
41	NC_044555.1	84	82278722	82280371	0
42	NC_044555.1	83	22396452	22394803	0
43	NC_044555.1	81	96191216	96190020	0
44	NC_044555.1	91	96190021	96189563	0
45	NC_044555.1	89	27394120	27394944	3.32E-122
46	NC_044555.1	89	115447513	115446776	1.97E-97
47	NC_044555.1	90	30170242	30170817	2.24E-93
48	NC_044559.1	82	22482086	22483735	0
49	NC_044559.1	83	7875117	7873438	0
50	NC_044559.1	87	32529083	32527851	0
51	NC_044559.1	91	106371236	106371961	0
52	NC_044549.1	82	5623363	5625012	0

53	NC_044549.1	82	147190152	147191801	0
54	NC_044549.1	82	111645131	111643482	0
55	NC_044549.1	82	149150298	149151830	0
56	NC_044549.1	80	20469566	20471074	0
57	NC_044549.1	85	198137050	198138324	0
58	NC_044549.1	84	173173854	173172499	0
59	NC_044549.1	80	65896922	65895792	0
60	NC_044549.1	86	5768179	5769045	0
61	NC_044549.1	88	43584342	43583695	0
62	NC_044549.1	90	65895760	65895272	0
63	NC_044562.1	83	15169181	15167532	0
64	NC_044562.1	82	26804766	26806361	0
65	NC_044562.1	84	73060347	73058698	0
66	NC_044562.1	83	68616486	68614837	0
67	NC_044562.1	87	3535303	3534533	2.04E-105
68	NC_044561.1	83	106457655	106456006	0
69	NC_044561.1	81	30092926	30093486	0
70	NC_044561.1	87	30093708	30094196	0
71	NC_044561.1	84	31787420	31786686	1.88E-95
72	NC_044565.1	84	18117499	18119148	0
73	NC_044565.1	84	75820550	75818901	0
74	NC_044565.1	82	16442340	16440691	0
75	NC_044565.1	82	16459805	16458156	0
76	NC_044565.1	80	75779653	75778007	0
77	NC_044565.1	81	17623146	17624657	0
78	NC_044565.1	85	29402090	29400804	0
79	NC_044565.1	88	52577672	52578577	9.88E-137
80	NC_044565.1	88	32226264	32227151	1.66E-128
81	NC_044565.1	87	1157204	1156434	1.37E-104
82	NC_044565.1	86	106028768	106029526	1.22E-102
83	NC_044567.1	82	67559356	67557761	0
84	NC_044553.1	83	32527303	32525654	0
85	NC_044553.1	85	28618509	28620188	0
86	NC_044553.1	84	121224425	121226104	0
87	NC_044553.1	86	52715340	52716746	0
88	NC_044553.1	86	28947230	28945977	0
89	NC_044553.1	87	95417497	95416697	0
90	NC_044564.1	83	47422472	47420823	0
91	NC_044564.1	84	30979552	30977897	0
92	NC_044564.1	84	45457276	45455627	0
93	NC_044564.1	87	63039645	63040880	0
94	NC_044564.1	86	104952909	104951965	0
95	NC_044564.1	88	93336743	93336096	0
96	NC_044564.1	83	67151314	67152327	1.83E-144
97	NC_044564.1	80	67150692	67151384	5.98E-100

98	NC_044564.1	91	53524079	53524657	6.30E-85
99	NC_044552.1	81	2805497	2803905	0
100	NC_044552.1	84	152675440	152673791	0
101	NC_044552.1	80	73521217	73519595	0
102	NC_044552.1	84	54136413	54137837	0
103	NC_044552.1	89	1174813	1173926	0
104	NC_044552.1	89	61986261	61987031	0
105	NC_044552.1	91	1173920	1173369	0
106	NC_044552.1	87	14967453	14966461	1.14E-157
107	NC_044566.1	80	13167698	13169347	0
108	NC_044566.1	87	29011342	29010056	0
109	NC_044569.1	83	43888135	43889784	0
110	NC_044557.1	81	13546266	13547864	0
111	NC_044557.1	82	37666321	37667970	0
112	NC_044557.1	83	70080400	70081119	0
113	NC_044557.1	87	112541835	112541191	0
114	NC_044557.1	91	70081149	70081637	0
115	NW_022142205.1	82	314556	313204	0
116	NC_044570.1	82	3464996	3463965	2.21E-137
117	NC_044570.1	86	3465633	3464935	1.25E-126
118	NC_044558.1	81	74231994	74233124	0
119	NC_044558.1	91	74233156	74233644	0
120	NC_044558.1	86	114541405	114541947	3.61E-73
121	NC_044568.1	80	9674538	9673435	0
122	NC_044568.1	80	36264626	36265687	0
123	NC_044568.1	92	36265686	36266276	0
124	NC_044568.1	90	9673435	9672890	0

Table 2.2 CERV1 env gene locations identified in the *Piliocolobus tephrosceles* genome

No.	Acc.	Iden.	Start	End	E-value
1	NW_022335374.1	88	1032	2702	0
2	NW_022319399.1	85	5666	7345	0
3	NW_022319400.1	85	1784	105	0
4	NW_022319738.1	85	5963	7642	0
5	NW_022319739.1	85	1820	141	0
6	NW_022323402.1	84	1813	158	0
7	NW_022323403.1	84	4470	6125	0
8	NC_045437.1	85	139717580	139719259	0
9	NC_045437.1	81	41293399	41294913	0
10	NC_045437.1	84	98917087	98916299	8.96E-105
11	NC_045439.1	83	136200535	136202184	0
12	NC_045439.1	83	144910730	144909081	0
13	NW_022307976.1	82	8425	6776	0

14	NC_045443.1	84	35049671	35048028	0
15	NC_045443.1	88	18240136	18241116	0
16	NC_045443.1	88	53987524	53986736	0
17	NC_045443.1	91	125425198	125425686	0
18	NW_022324670.1	82	1778	129	0
19	NW_022324671.1	82	5828	7477	0
20	NC_045440.1	84	2832672	2834315	0
21	NC_045438.1	81	79427977	79426403	0
22	NC_045445.1	83	21561049	21559400	0
23	NC_045450.1	89	30512522	30513931	0
24	NC_045450.1	86	47732443	47731541	2.15E-129
25	NC_045455.1	83	75154154	75155803	0
26	NC_045455.1	89	4723769	4724302	2.43E-170
27	NC_045436.1	83	118856460	118857971	0
28	NW_022321856.1	82	4032	5504	0
29	NW_022321857.1	82	3681	2209	0
30	NC_045446.1	83	14790319	14788889	0
31	NC_045446.1	81	50944552	50943449	0
32	NC_045446.1	86	80748986	80747937	0
33	NC_045446.1	86	49877644	49878291	0
34	NC_045446.1	84	100634383	100633706	0
35	NC_045446.1	89	100633706	100633110	0
36	NC_045446.1	88	50943432	50942944	0
37	NW_022317456.1	83	3620	5044	0
38	NW_022317457.1	83	4504	3080	0
39	NW_022318574.1	83	3622	5046	0
40	NW_022322277.1	84	6172	7503	0
41	NW_022322278.1	84	1718	387	0
42	NW_022310249.1	81	6476	5031	0
43	NW_022332403.1	88	1393	104	0
44	NC_045453.1	87	38883970	38882876	0
45	NW_022321936.1	89	5601	6488	0
46	NW_022321936.1	90	6485	7270	0
47	NW_022321937.1	89	1747	860	0
48	NW_022321937.1	90	863	78	0
49	NC_045435.1	90	99719775	99719128	0
50	NC_045444.1	88	7347436	7348146	0
51	NW_022302957.1	85	3118	2486	0
52	NC_045447.1	86	75488168	75486960	3.76E-177
53	NC_045441.1	87	94706674	94705895	0
54	NC_045441.1	83	44112633	44113766	1.39E-158
55	NC_045441.1	87	104883605	104882958	4.95E-137
56	NC_045441.1	85	93239498	93238902	5.44E-80
57	NW_022323647.1	87	3011	2076	9.69E-139
58	NW_022323648.1	87	1	936	9.69E-139

59	NW_022302993.1	83	2558	3535	2.78E-137
60	NW_022318386.1	89	6790	7677	2.17E-136
61	NW_022318184.1	89	1031	144	2.17E-136
62	NW_022318185.1	89	6792	7679	2.17E-136

Table 2.2 CERV1 env gene locations identified in the *Macaca silenus* genome

No.	Acc.	Ident.	Start	End	E-value
1	JAKFHO010001111.1	97	17061132	17059462	0
2	JAKFHO010001111.1	89	1770790	1771626	2.16E-164
3	JAKFHO010001111.1	94	2363840	2363190	5.48E-146
4	JAKFHO010001111.1	88	1771550	1772410	3.30E-123
5	JAKFHO010001111.1	95	7150913	7150143	7.86E-120
6	JAKFHO010001111.1	91	17881315	17881926	2.03E-109
7	JAKFHO010000446.1	97	34994696	34993026	0
8	JAKFHO010000446.1	92	24848935	24847265	0
9	JAKFHO010000446.1	89	1511270	1509831	0
10	JAKFHO010000446.1	89	24854779	24853805	0
11	JAKFHO010000446.1	95	4113460	4114431	0
12	JAKFHO010000446.1	96	22687151	22686504	0
13	JAKFHO010000446.1	92	22687741	22687199	0
14	JAKFHO010000446.1	96	24853599	24853111	0
15	JAKFHO010000446.1	94	30855297	30856307	4.39E-168
16	JAKFHO010001189.1	97	11196462	11194792	0
17	JAKFHO010001189.1	91	14684657	14686279	0
18	JAKFHO010001744.1	95	3858669	3856999	0
19	JAKFHO010000113.1	96	24994612	24992942	0
20	JAKFHO010000113.1	92	8286384	8287250	7.53E-137
21	JAKFHO010000934.1	95	11282357	11280687	0
22	JAKFHO010000934.1	93	21927944	21926277	0
23	JAKFHO010000934.1	90	15797386	15798894	0
24	JAKFHO010000934.1	93	34501401	34500541	3.61E-136
25	JAKFHO010000922.1	93	39392453	39390783	0
26	JAKFHO010000922.1	89	69310638	69312308	0
27	JAKFHO010000922.1	90	11354446	11355708	0
28	JAKFHO010000922.1	96	43356211	43355099	0
29	JAKFHO010000922.1	84	43356765	43356283	0
30	JAKFHO010000922.1	88	60466373	60467302	1.89E-142
31	JAKFHO010000922.1	88	60465637	60466422	3.59E-139
32	JAKFHO010000922.1	89	24337447	24336881	5.41E-132
33	JAKFHO010000158.1	92	270666	272336	0
34	JAKFHO010001078.1	93	11354687	11356357	0
35	JAKFHO010001244.1	92	1498625	1496955	0
36	JAKFHO010001244.1	94	13353466	13352993	9.69E-92

37	JAKFHO010001766.1	92	7580926	7582596	0
38	JAKFHO010001766.1	90	1809069	1810184	0
39	JAKFHO010001236.1	92	16420	18090	0
40	JAKFHO010001034.1	91	20017700	20016030	0
41	JAKFHO010001034.1	95	13100469	13101359	2.34E-147
42	JAKFHO010001034.1	96	6213532	6214020	9.66E-119
43	JAKFHO010001034.1	81	13099694	13100248	1.42E-111
44	JAKFHO010001499.1	92	43231	44901	0
45	JAKFHO010000905.1	92	1451860	1453530	0
46	JAKFHO010001669.1	91	53593	55263	0
47	JAKFHO010000002.1	91	31628110	31626440	0
48	JAKFHO010000002.1	88	43801220	43802890	0
49	JAKFHO010000002.1	83	45493348	45494217	1.33E-156
50	JAKFHO010000002.1	97	45474690	45475310	1.96E-155
51	JAKFHO010001267.1	92	12452356	12454026	0
52	JAKFHO010001267.1	88	623001	621466	0
53	JAKFHO010000001.1	90	41555448	41553778	0
54	JAKFHO010000001.1	92	28624805	28626472	0
55	JAKFHO010000001.1	92	16850152	16851474	0
56	JAKFHO010000001.1	92	30725945	30724935	4.77E-164
57	JAKFHO010001355.1	91	408855	407185	0
58	JAKFHO010001000.1	92	4318563	4316944	0
59	JAKFHO010001000.1	91	16317727	16319349	0
60	JAKFHO010001000.1	97	26572547	26571855	0
61	JAKFHO010001000.1	94	26571879	26571022	0
62	JAKFHO010001133.1	90	1727865	1729535	0
63	JAKFHO010001133.1	89	2065835	2064789	0
64	JAKFHO010001133.1	94	2064724	2064167	0
65	JAKFHO010001133.1	92	9848215	9847508	7.73E-141
66	JAKFHO010001133.1	89	15893065	15892193	6.10E-124
67	JAKFHO010001133.1	93	19830260	19829772	6.89E-116
68	JAKFHO010001133.1	86	9849109	9848519	1.10E-97
69	JAKFHO010000208.1	97	10164	11699	0
70	JAKFHO010001145.1	89	18929	17259	0
71	JAKFHO010001367.1	91	11841341	11839671	0
72	JAKFHO010001367.1	91	11817074	11815965	4.36E-177
73	JAKFHO010001367.1	85	11817633	11817043	3.28E-98
74	JAKFHO010000979.1	91	196576	194909	0
75	JAKFHO010000557.1	90	18633257	18634924	0
76	JAKFHO010000557.1	98	18601595	18600645	0
77	JAKFHO010000557.1	97	18602173	18601592	0
78	JAKFHO010000557.1	91	30080837	30081703	1.08E-135
79	JAKFHO010000447.1	91	1331107	1332726	0
80	JAKFHO010000458.1	90	1616446	1614776	0
81	JAKFHO010001222.1	90	5661931	5663601	0

82	JAKFHO010001222.1	91	13219242	13220774	0
83	JAKFHO010001444.1	90	4470900	4472567	0
84	JAKFHO010001255.1	90	3675861	3677531	0
85	JAKFHO010001255.1	92	2435099	2436370	0
86	JAKFHO010001255.1	96	56704113	56703226	0
87	JAKFHO010001255.1	95	5171272	5170757	0
88	JAKFHO010001255.1	84	56704669	56704187	0
89	JAKFHO010001255.1	96	16998080	16998940	2.68E-140
90	JAKFHO010001255.1	90	29279028	29278510	7.97E-73
91	JAKFHO010001167.1	89	16152394	16154064	0
92	JAKFHO010001167.1	88	16178934	16180037	0
93	JAKFHO010001167.1	91	16487789	16487193	0
94	JAKFHO010001167.1	91	16180072	16180560	0
95	JAKFHO010001036.1	90	399814	401484	0
96	JAKFHO010001045.1	89	2477456	2475786	0
97	JAKFHO010001045.1	83	23902822	23904471	0
98	JAKFHO010000535.1	91	2415620	2414046	0
99	JAKFHO010000535.1	96	156508	155357	0
100	JAKFHO010001699.1	92	13517184	13518737	0
101	JAKFHO010001699.1	93	38009529	38010332	0
102	JAKFHO010001699.1	97	38010310	38011194	0
103	JAKFHO010001699.1	83	15316181	15315225	0
104	JAKFHO010000224.1	89	20633947	20635617	0
105	JAKFHO010000224.1	93	27196169	27197179	2.74E-167
106	JAKFHO010000890.1	88	9892015	9890348	0
107	JAKFHO010000890.1	88	9905144	9906118	0
108	JAKFHO010000890.1	87	1473274	1472360	0
109	JAKFHO010000890.1	85	10518324	10517518	0
110	JAKFHO010000890.1	96	7363270	7363758	2.98E-142
111	JAKFHO010000890.1	91	7362747	7363238	2.98E-142
112	JAKFHO010000890.1	91	9078723	9079544	5.51E-121
113	JAKFHO010000890.1	83	8987834	8987280	9.20E-111
114	JAKFHO010001003.1	88	46619	44952	0
115	JAKFHO010001397.1	88	21927	20263	0
116	JAKFHO010000998.1	83	524767	523100	0
117	JAKFHO010000982.1	85	300066	298399	0
118	JAKFHO010000258.1	85	3031261	3029591	0
119	JAKFHO010001422.1	86	5959382	5961043	0
120	JAKFHO010000080.1	88	38452	39882	0
121	JAKFHO010000412.1	93	14152	12773	0
122	JAKFHO010001033.1	95	10190047	10188716	0
123	JAKFHO010001033.1	90	48269822	48270796	0
124	JAKFHO010001033.1	93	30835570	30836022	2.00E-118
125	JAKFHO010000569.1	90	1410143	1411462	0
126	JAKFHO010000569.1	93	7943	7365	0

127	JAKFHO010000945.1	92	22177226	22176153	0
128	JAKFHO010000945.1	85	23698859	23697924	0
129	JAKFHO010000945.1	86	23699539	23698871	0
130	JAKFHO010000945.1	84	22177824	22177201	0
131	JAKFHO010001144.1	80	3518615	3520054	0
132	JAKFHO010001144.1	93	51335998	51336819	3.66E-129
133	JAKFHO010001722.1	91	225595	224321	0
134	JAKFHO010001322.1	95	4791430	4792695	0
135	JAKFHO010001322.1	84	4650244	4649066	0
136	JAKFHO010001322.1	84	4650729	4650247	0
137	JAKFHO010000037.1	96	35952	34714	0
138	JAKFHO010001178.1	91	16531143	16532459	0
139	JAKFHO010001178.1	95	16522769	16523878	0
140	JAKFHO010001178.1	95	15105884	15104796	0
141	JAKFHO010001178.1	81	15106438	15105884	0
142	JAKFHO010001035.1	96	524196	522970	0
143	JAKFHO010001016.1	91	43224	44375	0
144	JAKFHO010001016.1	97	44407	44895	0
145	JAKFHO010001777.1	85	3979664	3978405	0
146	JAKFHO010000069.1	91	4045743	4044592	0
147	JAKFHO010000069.1	98	4044563	4044075	0
148	JAKFHO010000069.1	93	2793088	2794098	5.15E-165
149	JAKFHO010001103.1	83	229192	230643	0
150	JAKFHO010001555.1	88	2131242	2132438	0
151	JAKFHO010001555.1	95	2132435	2132911	0
152	JAKFHO010001478.1	89	6232172	6233323	0
153	JAKFHO010001478.1	90	6233355	6233843	0
154	JAKFHO010001478.1	93	4102351	4101341	2.66E-167
155	JAKFHO010000891.1	97	1271663	1272787	0
156	JAKFHO010000891.1	91	1271178	1271663	0
157	JAKFHO010001456.1	89	34241	33090	0
158	JAKFHO010001456.1	94	33058	32570	0
159	JAKFHO010001098.1	98	69886	68939	0
160	JAKFHO010001098.1	87	68756	68232	0
161	JAKFHO010000335.1	98	6637773	6636658	0
162	JAKFHO010000335.1	93	6638327	6637776	0
163	JAKFHO010000335.1	95	14680666	14679656	3.10E-169
164	JAKFHO010000335.1	92	35721376	35720516	7.77E-134
165	JAKFHO010000335.1	85	1492138	1491575	4.94E-133
166	JAKFHO010000335.1	95	35727299	35726811	2.15E-116
167	JAKFHO010000335.1	94	17521140	17521628	2.71E-116
168	JAKFHO010001533.1	97	3320451	3319618	0
169	JAKFHO010001533.1	94	3319642	3318782	0
170	JAKFHO010000779.1	93	29113842	29113075	1.29E-114
171	JAKFHO010000989.1	86	9270842	9271990	0

172	JAKFHO010000989.1	97	9272020	9272508	0
173	JAKFHO010000925.1	89	37080	35929	0
174	JAKFHO010000925.1	91	35893	35405	0
175	JAKFHO010001293.1	81	184316	185632	0
176	JAKFHO010001588.1	96	29251289	29251777	5.85E-120
177	JAKFHO010001256.1	95	13406907	13408001	0
178	JAKFHO010001256.1	84	13406332	13406910	0
179	JAKFHO010001256.1	94	5159886	5160896	1.02E-171
180	JAKFHO010001477.1	90	28895529	28894858	1.24E-123
181	JAKFHO010001022.1	89	10086470	10087399	1.05E-143
182	JAKFHO010001022.1	87	10086019	10086519	6.48E-118
183	JAKFHO010001400.1	89	4490209	4489136	0
184	JAKFHO010001366.1	83	36216847	36217959	0
185	JAKFHO010001366.1	95	36218022	36218510	0
186	JAKFHO010001366.1	92	28812770	28813627	4.52E-135
187	JAKFHO010001366.1	93	567074	566253	8.34E-130
188	JAKFHO010001366.1	96	39051518	39052111	7.56E-95
189	JAKFHO010001455.1	90	1436794	1435871	0
190	JAKFHO010001455.1	93	1435820	1435173	0
191	JAKFHO010001117.1	89	375972	374857	0
192	JAKFHO010001311.1	96	4915406	4914471	0
193	JAKFHO010001311.1	83	4915963	4915406	0
194	JAKFHO010000283.1	83	14013	12904	0
195	JAKFHO010000283.1	84	14569	14024	0
196	JAKFHO010000136.1	87	28173	27292	0
197	JAKFHO010001666.1	95	4280235	4281116	0
198	JAKFHO010000923.1	88	19101566	19102309	0
199	JAKFHO010000923.1	96	19102744	19103232	0
200	JAKFHO010000724.1	94	202617	203186	0
201	JAKFHO010001209.1	94	194241	193618	0
202	JAKFHO010001209.1	93	194735	194241	0

Table 2.2 CERV1 env gene locations identified in the *Erythrocebus patas* genome

No.	Acc.	Iden.	Start	End	E-value
1	JAKFHT010004611.1	91	2980400	2978730	0
2	JAKFHT010001557.1	90	462563	464233	0
3	JAKFHT010001557.1	80	6149798	6150766	9.33E-137
4	JAKFHT010001668.1	91	24906579	24904957	0
5	JAKFHT010001668.1	80	21368389	21370056	0
6	JAKFHT010001668.1	87	11433615	11432464	0
7	JAKFHT010001113.1	91	435386	433716	0
8	JAKFHT010001113.1	92	6259580	6258306	0
9	JAKFHT010000001.1	91	1315956	1317578	0

10	JAKFHT010000001.1	94	2600897	2600115	0
11	JAKFHT010000001.1	94	20822696	20822208	0
12	JAKFHT010000001.1	94	2598903	2598415	5.25E-77
13	JAKFHT010004456.1	89	12223009	12224628	0
14	JAKFHT010004456.1	87	201773	202888	0
15	JAKFHT010004456.1	85	194147	195202	0
16	JAKFHT010004456.1	88	202846	203442	5.60E-83
17	JAKFHT010000491.1	90	1350981	1349311	0
18	JAKFHT010000491.1	93	2355238	2356008	0
19	JAKFHT010000491.1	84	2354690	2355241	0
20	JAKFHT010001765.1	89	9498	7828	0
21	JAKFHT010001002.1	90	1383968	1382298	0
22	JAKFHT010001002.1	93	1990913	1990425	2.33E-159
23	JAKFHT010001002.1	83	1993789	1993313	1.53E-151
24	JAKFHT010001002.1	81	1991650	1991201	5.00E-65
25	JAKFHT010000047.1	89	3751557	3749890	0
26	JAKFHT010000335.1	89	31801565	31799970	0
27	JAKFHT010000335.1	95	25921307	25921795	1.42E-78
28	JAKFHT010004389.1	85	11874992	11873331	0
29	JAKFHT010000557.1	87	32680581	32678917	0
30	JAKFHT010000668.1	91	22561336	22562829	0
31	JAKFHT010000668.1	84	967378	966134	0
32	JAKFHT010000668.1	87	32471830	32472924	0
33	JAKFHT010004415.1	83	53529	55196	0
34	JAKFHT010004415.1	90	56527	57519	2.65E-159
35	JAKFHT010004352.1	88	222588	221059	0
36	JAKFHT010004455.1	81	46635324	46636991	0
37	JAKFHT010004455.1	95	38266941	38267429	0
38	JAKFHT010004455.1	81	46633972	46634769	3.47E-168
39	JAKFHT010000158.1	87	426916	428439	0
40	JAKFHT010001379.1	82	8881	10545	0
41	JAKFHT010000646.1	86	907586	906042	0
42	JAKFHT010000113.1	89	7122935	7121589	0
43	JAKFHT010000113.1	84	4322165	4320687	0
44	JAKFHT010000535.1	88	4042883	4041528	0
45	JAKFHT010003001.1	83	5227221	5228747	0
46	JAKFHT010003001.1	93	13967837	13967010	4.03E-130
47	JAKFHT010002112.1	85	19002101	19000770	0
48	JAKFHT010002112.1	92	19008925	19008359	2.00E-86
49	JAKFHT010004356.1	92	9685775	9687091	0
50	JAKFHT010004356.1	86	13288126	13286915	0
51	JAKFHT010004356.1	81	7491599	7492708	0
52	JAKFHT010004356.1	86	7499825	7500706	0
53	JAKFHT010004356.1	95	7500997	7501485	0
54	JAKFHT010004356.1	92	14867867	14867007	2.76E-134

55	JAKFHT010000480.1	90	86113	87435	0
56	JAKFHT010000480.1	85	4094922	4094272	0
57	JAKFHT010004344.1	87	8072820	8071501	0
58	JAKFHT010004344.1	80	38844689	38845384	4.65E-107
59	JAKFHT010003491.1	89	268756	270081	0
60	JAKFHT010000369.1	89	2494235	2495449	0
61	JAKFHT010000513.1	92	2085485	2084247	0
62	JAKFHT010000735.1	82	1146306	1145053	0
63	JAKFHT010003668.1	80	1816042	1817010	0
64	JAKFHT010003668.1	96	1817058	1817630	0
65	JAKFHT010000169.1	88	5453424	5452300	0
66	JAKFHT010000169.1	94	5452249	5451761	0
67	JAKFHT010002335.1	92	3201136	3200027	0
68	JAKFHT010004666.1	84	509917	508772	0
69	JAKFHT010004666.1	93	508743	508255	0
70	JAKFHT010004567.1	84	8220570	8219431	0
71	JAKFHT010004567.1	96	8219395	8218907	0
72	JAKFHT010004677.1	94	25822734	25821622	0
73	JAKFHT010004677.1	85	25823291	25822737	0
74	JAKFHT010004677.1	93	29370867	29371355	0
75	JAKFHT010002291.1	84	782397	781165	0
76	JAKFHT010001591.1	83	77603	78712	0
77	JAKFHT010001446.1	91	6184469	6185080	0
78	JAKFHT010001446.1	88	1149337	1149918	0
79	JAKFHT010001446.1	90	1150366	1150998	0
80	JAKFHT010001446.1	90	1149899	1150375	0
81	JAKFHT010003335.1	84	7240999	7241973	0
82	JAKFHT010003335.1	91	20145804	20146340	0
83	JAKFHT010003335.1	93	7242176	7242664	0
84	JAKFHT010003335.1	88	20146340	20146825	0
85	JAKFHT010001112.1	83	45305770	45304886	0
86	JAKFHT010001112.1	93	35884218	35883739	6.91E-117
87	JAKFHT010002557.1	88	9835286	9834402	0
88	JAKFHT010002557.1	93	1033654	1033007	0
89	JAKFHT010002557.1	95	9834110	9833622	0
90	JAKFHT010000290.1	88	36554	37402	0
91	JAKFHT010003779.1	85	12189760	12190596	0
92	JAKFHT010002402.1	94	77501	78379	0
93	JAKFHT010003223.1	89	12697607	12698374	0
94	JAKFHT010003223.1	93	12698365	12698958	0
95	JAKFHT010000391.1	84	4199278	4200111	0
96	JAKFHT010000391.1	94	4328109	4328957	0
97	JAKFHT010000391.1	94	4200131	4200817	0
98	JAKFHT010001435.1	86	1469489	1470265	0
99	JAKFHT010001435.1	90	1470264	1471109	0

100	JAKFHT010002513.1	95	202327	203097	0
101	JAKFHT010002513.1	82	91773	91153	1.19E-102
102	JAKFHT010003334.1	93	53873125	53873841	0
103	JAKFHT010000446.1	92	30246512	30245754	0
104	JAKFHT010000446.1	93	30246984	30246532	0
105	JAKFHT010004789.1	94	6595774	6595181	0
106	JAKFHT010004789.1	84	6596330	6595848	0
107	JAKFHT010004789.1	96	6595151	6594663	0
108	JAKFHT010002446.1	88	5079244	5078672	0
109	JAKFHT010001524.1	93	1227779	1228258	1.81E-180
110	JAKFHT010001524.1	94	1228475	1229011	1.81E-180
111	JAKFHT010001890.1	94	19487776	19488264	2.85E-105
112	JAKFHT010001224.1	94	10159906	10160388	0
113	JAKFHT010001224.1	92	6258905	6259765	3.65E-136
114	JAKFHT010001224.1	91	30200232	30201038	4.94E-125
115	JAKFHT010002447.1	85	221529	220642	7.83E-175
116	JAKFHT010002447.1	83	460882	461367	4.03E-113
117	JAKFHT010001302.1	85	1716132	1716932	0
118	JAKFHT010001302.1	92	1716932	1717396	0
119	JAKFHT010000114.1	83	5818296	5817067	4.00E-174
120	JAKFHT010002224.1	94	15167055	15166408	1.86E-168
121	JAKFHT010002224.1	81	15167638	15167189	1.86E-168
122	JAKFHT010004766.1	88	5374460	5373630	5.73E-163
123	JAKFHT010002779.1	90	5037115	5036507	3.94E-107
124	JAKFHT010002779.1	89	17607062	17606538	2.96E-89
125	JAKFHT010004668.1	88	58607	59251	6.85E-160
126	JAKFHT010004333.1	89	3129490	3128468	4.40E-155
127	JAKFHT010000293.1	94	13385	14086	4.75E-151
128	JAKFHT010001779.1	92	11979850	11980362	3.05E-165
129	JAKFHT010001779.1	92	11981994	11982569	7.88E-162
130	JAKFHT010001779.1	93	11982560	11983105	7.88E-162
131	JAKFHT010001779.1	90	17399132	17399587	4.96E-113
132	JAKFHT010000002.1	93	5057487	5058209	0
133	JAKFHT010000002.1	91	5056972	5057502	0
134	JAKFHT010000002.1	93	20705510	20706433	4.41E-146
135	JAKFHT010002223.1	94	5720351	5719491	3.29E-138
136	JAKFHT010000976.1	87	10439	11308	1.13E-129
137	JAKFHT010000291.1	94	4906721	4906233	5.18E-111
138	JAKFHT010004467.1	93	10166879	10166076	5.41E-122
139	JAKFHT010000779.1	93	26303262	26303945	1.17E-119
140	JAKFHT010004522.1	92	4252021	4251338	2.08E-119
141	JAKFHT010004367.1	88	15187062	15186481	7.51E-143
142	JAKFHT010004622.1	91	8516343	8516855	1.23E-116
143	JAKFHT010001158.1	80	652053	652811	4.01E-99
144	JAKFHT010001158.1	93	650284	650736	1.03E-82

145	JAKFHT010004345.1	81	10054893	10055432	6.83E-107

Table 2.2 CERV1 env gene locations identified in the *Rhinopithecus Strykeri* genome

No.	Acc.	Iden.	Start	End	E-value
1	JAKFHF010003989.1	88	10409848	10408190	0
2	JAKFHF010001512.1	84	891931	890282	0
3	JAKFHF010000101.1	85	10121745	10123397	0
4	JAKFHF010000101.1	89	8475040	8475504	0
5	JAKFHF010000101.1	89	9616831	9616199	2.08E-105
6	JAKFHF010000101.1	90	6765541	6764966	2.49E-94
7	JAKFHF010001412.1	84	4248405	4250054	0
8	JAKFHF010001412.1	80	15550838	15549246	0
9	JAKFHF010001412.1	87	7355255	7355863	0
10	JAKFHF010004305.1	84	4197735	4196086	0
11	JAKFHF010002731.1	83	591038	589389	0
12	JAKFHF010004197.1	84	854800	856479	0
13	JAKFHF010002617.1	84	2494525	2496180	0
14	JAKFHF010002617.1	88	10699832	10699110	0
15	JAKFHF010002617.1	89	12691570	12690815	0
16	JAKFHF010002617.1	91	18846942	18847520	3.05E-86
17	JAKFHF010000984.1	84	895711	897360	0
18	JAKFHF010000520.1	82	49175520	49173928	0
19	JAKFHF010000520.1	83	9377	11026	0
20	JAKFHF010000520.1	83	4495923	4497323	0
21	JAKFHF010000520.1	88	58134266	58133481	0
22	JAKFHF010001883.1	83	4022641	4020992	0
23	JAKFHF010002192.1	83	13171288	13172937	0
24	JAKFHF010004320.1	82	534705	536354	0
25	JAKFHF010000356.1	83	14946392	14944743	0
26	JAKFHF010000281.1	82	43603994	43602348	0
27	JAKFHF010000281.1	85	24499761	24501440	0
28	JAKFHF010000281.1	81	18947577	18945922	0
29	JAKFHF010001739.1	82	15198992	15200641	0
30	JAKFHF010001739.1	85	4762075	4761161	1.62E-120
31	JAKFHF010001739.1	88	11652147	11651521	9.48E-94
32	JAKFHF010000464.1	82	2630689	2632338	0
33	JAKFHF010001075.1	81	5885688	5887280	0
34	JAKFHF010001075.1	84	43808878	43807307	0
35	JAKFHF010001075.1	87	20531064	20531828	0
36	JAKFHF010001429.1	84	27462834	27464483	0
37	JAKFHF010001779.1	83	18175600	18177249	0
38	JAKFHF010001779.1	85	21281247	21281879	2.03E-176
39	JAKFHF010004362.1	83	10745596	10747245	0

40	JAKFHF010004362.1	90	8257570	8257061	8.04E-108
41	JAKFHF010001904.1	81	24989	26638	0
42	JAKFHF010003966.1	83	1943275	1941626	0
43	JAKFHF010001888.1	80	18941	17289	0
44	JAKFHF010003104.1	81	11856257	11854641	0
45	JAKFHF010000133.1	81	3054798	3053380	0
46	JAKFHF010001292.1	81	43231	44655	0
47	JAKFHF010004511.1	85	3981090	3979816	0
48	JAKFHF010001442.1	81	10740279	10738948	0
49	JAKFHF010000626.1	89	25694496	25695221	0
50	JAKFHF010000626.1	87	26289236	26288385	2.19E-125
51	JAKFHF010003863.1	86	3279277	3280563	0
52	JAKFHF010000845.1	81	2311	1052	0
53	JAKFHF010004406.1	86	713811	712708	6.98E-164
54	JAKFHF010001476.1	81	1107921	1106743	0
55	JAKFHF010001476.1	92	1211819	1211331	0
56	JAKFHF010002163.1	83	18388	17102	0
57	JAKFHF010003047.1	81	578000	576762	0
58	JAKFHF010003441.1	83	850187	851383	0
59	JAKFHF010003441.1	86	851383	851865	0
60	JAKFHF010002807.1	87	5409606	5410058	0
61	JAKFHF010004098.1	85	28995730	28996824	0
62	JAKFHF010003771.1	86	501899	500958	0
63	JAKFHF010003852.1	88	3507054	3507701	0
64	JAKFHF010003852.1	88	17291467	17290820	0
65	JAKFHF010003852.1	81	15592596	15593156	1.90E-83
66	JAKFHF010002577.1	80	19048690	19049739	0
67	JAKFHF010002577.1	88	19049736	19050338	0
68	JAKFHF010004020.1	83	21160280	21159249	0
69	JAKFHF010002693.1	90	12552327	12553100	0
70	JAKFHF010001246.1	88	17287649	17288404	0
71	JAKFHF010001246.1	89	23658638	23659300	0
72	JAKFHF010001246.1	91	18984138	18984689	8.78E-159
73	JAKFHF010001246.1	87	23666176	23667027	4.94E-115
74	JAKFHF010001266.1	85	28447	29277	0
75	JAKFHF010000442.1	90	1171844	1172560	0
76	JAKFHF010001117.1	87	600469	599669	0
77	JAKFHF010001314.1	83	16223	16921	0
78	JAKFHF010004602.1	89	27658	27017	0
79	JAKFHF010004602.1	84	28269	27712	0
80	JAKFHF010000838.1	83	2680170	2680628	0
81	JAKFHF010002187.1	87	4029	3037	3.85E-158
82	JAKFHF010002152.1	83	304924	303944	2.40E-151
83	JAKFHF010002152.1	86	303728	303246	8.91E-68
84	JAKFHF010001491.1	82	1734317	1735222	0

85	JAKFHF010000617.1	88	543048	543533	0
86	JAKFHF010000617.1	83	542307	542774	0
87	JAKFHF010000617.1	83	500472	501440	1.67E-133
88	JAKFHF010002273.1	84	12958173	12958772	3.36E-142
89	JAKFHF010002273.1	81	9396441	9397469	3.98E-132
90	JAKFHF010003601.1	87	4908781	4907894	8.26E-132
91	JAKFHF010003601.1	89	4794119	4794856	3.72E-94
92	JAKFHF010001533.1	87	4739226	4739723	1.77E-65
93	JAKFHF010003179.1	85	95283	95963	3.99E-88
94	JAKFHF010001438.1	81	7347952	7347344	1.02E-82
95	JAKFHF010000331.1	87	2828551	2829108	2.67E-98
96	JAKFHF010003647.1	87	35472148	35471606	4.15E-74
97	JAKFHF010001043.1	89	211188	210451	6.76E-97

Table 2.2 CERV1 env gene locations identified in the *Macaca nigra* genome

No.	Acc.	Iden.	Start	End	E-value
1	CAKMXP010000015.1	97	36545260	36543590	0
2	CAKMXP010000015.1	91	21269298	21270920	0
3	CAKMXP010000015.1	90	18102632	18103747	0
4	CAKMXP010000015.1	90	20813420	20812305	0
5	CAKMXP010000015.1	95	20812312	20811749	0
6	CAKMXP010000015.1	94	18103740	18104303	0
7	CAKMXP010000015.1	90	37370908	37371591	8.79E-116
8	CAKMXP010000006.1	97	37974950	37976620	0
9	CAKMXP010000006.1	93	51307225	51305555	0
10	CAKMXP010000006.1	91	63285206	63286825	0
11	CAKMXP010000006.1	90	78829531	78828215	0
12	CAKMXP010000006.1	88	78803337	78802234	0
13	CAKMXP010000006.1	93	78802198	78801710	0
14	CAKMXP010000014.1	96	37321520	37323190	0
15	CAKMXP010000014.1	91	69856343	69858013	0
16	CAKMXP010000014.1	90	22424009	22422858	0
17	CAKMXP010000014.1	88	58957591	58958742	0
18	CAKMXP010000014.1	88	9421928	9423043	0
19	CAKMXP010000014.1	94	34227863	34227012	0
20	CAKMXP010000014.1	93	9423036	9423599	0
21	CAKMXP010000014.1	86	34228350	34227868	0
22	CAKMXP010000014.1	92	22422826	22422341	0
23	CAKMXP010000005.1	96	67896297	67894627	0
24	CAKMXP010000005.1	92	42457755	42456085	0
25	CAKMXP010000005.1	91	99118400	99120070	0
26	CAKMXP010000005.1	90	2029193	2027571	0
27	CAKMXP010000005.1	89	1970931	1969261	0

28	CAKMXP010000005.1	84	16889902	16888334	0
29	CAKMXP010000005.1	89	10397626	10396475	0
30	CAKMXP010000005.1	88	11021904	11020753	0
31	CAKMXP010000005.1	89	7903738	7904853	0
32	CAKMXP010000005.1	89	92373422	92374537	0
33	CAKMXP010000005.1	96	7904846	7905409	0
34	CAKMXP010000005.1	92	92374530	92375093	0
35	CAKMXP010000005.1	94	10396445	10395957	0
36	CAKMXP010000005.1	90	11020721	11020233	0
37	CAKMXP010000005.1	94	59957831	59956821	1.46E-168
38	CAKMXP010000005.1	96	7858075	7858563	1.03E-138
39	CAKMXP010000005.1	92	7857551	7858006	1.03E-138
40	CAKMXP010000005.1	89	47765072	47764251	6.98E-120
41	CAKMXP010000005.1	94	9498470	9497982	3.40E-110
42	CAKMXP010000005.1	96	9591096	9591584	1.27E-109
43	CAKMXP010000002.1	95	60732005	60730335	0
44	CAKMXP010000002.1	91	64229909	64231531	0
45	CAKMXP010000002.1	91	26339859	26338189	0
46	CAKMXP010000002.1	88	23367667	23365997	0
47	CAKMXP010000002.1	92	49407457	49406072	0
48	CAKMXP010000002.1	90	121362869	121361553	0
49	CAKMXP010000002.1	93	6226056	6224872	0
50	CAKMXP010000002.1	90	30237336	30236221	0
51	CAKMXP010000002.1	93	8882508	8883518	7.54E-164
52	CAKMXP010000002.1	82	107547616	107548188	2.08E-117
53	CAKMXP010000019.1	96	45127734	45126064	0
54	CAKMXP010000019.1	90	11610650	11608980	0
55	CAKMXP010000019.1	95	22728618	22729781	0
56	CAKMXP010000019.1	84	22728110	22728592	0
57	CAKMXP010000012.1	96	62825964	62824342	0
58	CAKMXP010000012.1	92	48040111	48041781	0
59	CAKMXP010000012.1	92	38355588	38353918	0
60	CAKMXP010000012.1	92	77355952	77357622	0
61	CAKMXP010000012.1	92	66922371	66924041	0
62	CAKMXP010000012.1	90	63368321	63366651	0
63	CAKMXP010000012.1	96	62668844	62667723	0
64	CAKMXP010000012.1	81	62669392	62668835	0
65	CAKMXP010000011.1	93	58541839	58540169	0
66	CAKMXP010000011.1	90	9845518	9844403	0
67	CAKMXP010000011.1	89	66203557	66204624	0
68	CAKMXP010000011.1	92	66204617	66205180	0
69	CAKMXP010000011.1	94	9844334	9843846	0
70	CAKMXP010000011.1	92	48370891	48370328	2.58E-119
71	CAKMXP010000011.1	88	75600561	75599893	1.44E-96
72	CAKMXP010000016.1	93	55621126	55622793	0

73	CAKMXP010000016.1	86	37944241	37942571	0
74	CAKMXP010000016.1	95	33316983	33317591	0
75	CAKMXP010000016.1	95	33317584	33318147	0
76	CAKMXP010000016.1	84	33316475	33316957	0
77	CAKMXP010000034.1	93	16915861	16917531	0
78	CAKMXP010000034.1	93	16241445	16239775	0
79	CAKMXP010000034.1	95	9794323	9795438	0
80	CAKMXP010000035.1	92	12002575	12000905	0
81	CAKMXP010000035.1	96	20909588	20909130	1.77E-91
82	CAKMXP010000024.1	92	35699792	35701462	0
83	CAKMXP010000024.1	89	3996291	3994672	0
84	CAKMXP010000022.1	92	42717178	42715508	0
85	CAKMXP010000022.1	91	29201098	29200541	5.92E-93
86	CAKMXP010000013.1	92	17911521	17909851	0
87	CAKMXP010000029.1	91	30972548	30974218	0
88	CAKMXP010000029.1	90	8242026	8240356	0
89	CAKMXP010000038.1	91	13545513	13543843	0
90	CAKMXP010000038.1	92	11503361	11504854	0
91	CAKMXP010000018.1	92	100573	102243	0
92	CAKMXP010000007.1	92	13213909	13215579	0
93	CAKMXP010000007.1	91	7593675	7592056	0
94	CAKMXP010000007.1	91	36448837	36447452	0
95	CAKMXP010000007.1	96	6710112	6709624	1.56E-109
96	CAKMXP010000025.1	92	3228497	3226827	0
97	CAKMXP010000025.1	95	23490398	23491903	0
98	CAKMXP010000025.1	88	31131951	31130887	0
99	CAKMXP010000025.1	87	18977897	18976992	0
100	CAKMXP010000025.1	93	18976788	18976225	0
101	CAKMXP010000025.1	94	31130818	31130330	0
102	CAKMXP010000025.1	90	227588	228145	4.72E-113
103	CAKMXP010000046.1	91	14030186	14028516	0
104	CAKMXP010000046.1	88	1196035	1196949	0
105	CAKMXP010000046.1	94	1196942	1197505	0
106	CAKMXP010000046.1	84	4089804	4090286	0
107	CAKMXP010000001.1	91	163313218	163311548	0
108	CAKMXP010000001.1	91	136751528	136749858	0
109	CAKMXP010000001.1	88	129035794	129037461	0
110	CAKMXP010000001.1	89	151613108	151611957	0
111	CAKMXP010000001.1	94	116790023	116788839	0
112	CAKMXP010000001.1	90	21098971	21097820	0
113	CAKMXP010000001.1	98	18479066	18477951	0
114	CAKMXP010000001.1	90	156166634	156165519	0
115	CAKMXP010000001.1	94	74171297	74170530	0
116	CAKMXP010000001.1	95	74306487	74307236	0
117	CAKMXP010000001.1	93	74170537	74169977	0

118	CAKMXP010000001.1	98	135262025	135261537	0
119	CAKMXP010000001.1	96	21097788	21097300	0
120	CAKMXP010000001.1	94	156165450	156164962	0
121	CAKMXP010000001.1	82	116790510	116790028	0
122	CAKMXP010000001.1	94	74307268	74307756	0
123	CAKMXP010000001.1	95	151611925	151611437	0
124	CAKMXP010000001.1	92	149841349	149840531	2.68E-126
125	CAKMXP010000001.1	94	144593538	144594026	3.13E-118
126	CAKMXP010000001.1	88	144592682	144593200	1.50E-117
127	CAKMXP010000001.1	93	29416019	29416507	1.23E-116
128	CAKMXP010000001.1	94	47642942	47642454	1.36E-116
129	CAKMXP010000001.1	93	107419342	107419986	2.78E-108
130	CAKMXP010000044.1	91	8071274	8069655	0
131	CAKMXP010000044.1	94	17827034	17825850	0
132	CAKMXP010000044.1	92	1533855	1534688	8.57E-143
133	CAKMXP010000021.1	91	43429150	43430820	0
134	CAKMXP010000021.1	89	31106031	31107146	0
135	CAKMXP010000021.1	93	31107139	31107702	0
136	CAKMXP010000030.1	91	8972621	8974291	0
137	CAKMXP010000030.1	88	21854687	21855751	0
138	CAKMXP010000030.1	94	21855744	21856307	0
139	CAKMXP010000030.1	97	10338070	10337537	5.54E-159
140	CAKMXP010000030.1	98	10337505	10337017	5.54E-159
141	CAKMXP010000047.1	90	9877776	9879446	0
142	CAKMXP010000032.1	90	24798955	24800625	0
143	CAKMXP010000067.1	91	782719	781049	0
144	CAKMXP010000067.1	90	32206	33876	0
145	CAKMXP010000076.1	90	488121	486451	0
146	CAKMXP010000003.1	91	76627535	76625916	0
147	CAKMXP010000003.1	96	20939678	20940670	0
148	CAKMXP010000003.1	96	20939004	20939681	0
149	CAKMXP010000003.1	86	5392384	5391788	0
150	CAKMXP010000003.1	94	5391791	5391234	0
151	CAKMXP010000003.1	91	9309564	9309001	1.79E-117
152	CAKMXP010000010.1	91	13519201	13517531	0
153	CAKMXP010000010.1	93	30694172	30695494	0
154	CAKMXP010000010.1	92	52568686	52569969	0
155	CAKMXP010000010.1	90	62714273	62713122	0
156	CAKMXP010000010.1	94	33586578	33585394	0
157	CAKMXP010000010.1	90	34668358	34669485	0
158	CAKMXP010000010.1	87	56761672	56762823	0
159	CAKMXP010000010.1	91	4667964	4667335	0
160	CAKMXP010000010.1	92	4667342	4666779	0
161	CAKMXP010000010.1	96	34669542	34670030	0
162	CAKMXP010000010.1	84	33587065	33586583	0

163	CAKMXP010000010.1	83	4668451	4667969	0
164	CAKMXP010000010.1	96	62713090	62712602	0
165	CAKMXP010000010.1	92	56762855	56763343	0
166	CAKMXP010000010.1	91	20243331	20242876	1.69E-131
167	CAKMXP010000010.1	81	4285553	4284990	5.38E-127
168	CAKMXP010000010.1	91	4286001	4285546	5.38E-127
169	CAKMXP010000119.1	90	212262	213881	0
170	CAKMXP010000008.1	89	77641394	77643064	0
171	CAKMXP010000008.1	95	84546536	84545289	0
172	CAKMXP010000008.1	97	58414237	58415211	0
173	CAKMXP010000008.1	94	58415261	58415908	0
174	CAKMXP010000008.1	92	3062206	3061499	1.10E-141
175	CAKMXP010000008.1	92	86791632	86790745	3.79E-139
176	CAKMXP010000008.1	91	71132687	71133553	2.94E-138
177	CAKMXP010000008.1	86	86792318	86791581	2.51E-137
178	CAKMXP010000040.1	89	5214395	5216068	0
179	CAKMXP010000009.1	90	9950240	9948621	0
180	CAKMXP010000009.1	89	1830058	1831185	0
181	CAKMXP010000009.1	87	2168023	2167049	0
182	CAKMXP010000009.1	94	2167001	2166354	0
183	CAKMXP010000009.1	91	1831166	1831729	0
184	CAKMXP010000009.1	93	14829257	14830267	1.00E-167
185	CAKMXP010000009.1	92	28287370	28288236	5.54E-136
186	CAKMXP010000009.1	89	16013360	16012488	2.10E-123
187	CAKMXP010000009.1	91	86107310	86106822	3.49E-112
188	CAKMXP010000059.1	95	649520	648144	0
189	CAKMXP010000020.1	88	46376190	46377728	0
190	CAKMXP010000020.1	95	24836286	24835135	0
191	CAKMXP010000020.1	90	4444190	4443039	0
192	CAKMXP010000020.1	93	16277597	16278712	0
193	CAKMXP010000020.1	91	4443007	4442519	0
194	CAKMXP010000042.1	97	5200065	5198776	0
195	CAKMXP010000079.1	93	1393437	1394759	0
196	CAKMXP010000017.1	96	1810131	1808980	0
197	CAKMXP010000017.1	88	306379	307503	0
198	CAKMXP010000017.1	90	307484	308047	0
199	CAKMXP010000017.1	96	1808950	1808462	0
200	CAKMXP010000017.1	92	42691598	42690891	2.40E-141
201	CAKMXP010000115.1	94	23290	24549	0
202	CAKMXP010000051.1	90	12097365	12096079	0
203	CAKMXP010000051.1	89	12074535	12073408	0
204	CAKMXP010000051.1	93	12073427	12072864	0
205	CAKMXP010000051.1	93	6718544	6719035	7.89E-143
206	CAKMXP010000051.1	93	6718021	6718512	7.89E-143
207	CAKMXP010000142.1	89	104559	103408	0

208	CAKMXP010000142.1	96	103376	102888	0
209	CAKMXP010000004.1	90	75609382	75608231	0
210	CAKMXP010000004.1	96	75608201	75607713	0
211	CAKMXP010000004.1	92	12115158	12115979	4.28E-127
212	CAKMXP010000053.1	88	6109	7224	0
213	CAKMXP010000053.1	84	5744852	5745541	0
214	CAKMXP010000053.1	94	5745898	5746386	0
215	CAKMXP010000027.1	96	26425401	26426375	0
216	CAKMXP010000027.1	92	22273170	22272097	0
217	CAKMXP010000027.1	96	26426425	26427072	0
218	CAKMXP010000027.1	86	22273768	22273172	0
219	CAKMXP010000050.1	95	284928	285938	1.05E-168
220	CAKMXP010000160.1	92	57180	56383	0
221	CAKMXP010000160.1	96	56351	55863	0
222	CAKMXP010000081.1	96	1330161	1330724	2.08E-146
223	CAKMXP010000081.1	93	1329713	1330168	2.08E-146
224	CAKMXP010000023.1	93	18213917	18214783	1.48E-140
225	CAKMXP010000023.1	92	777574	777086	3.88E-74
226	CAKMXP010000036.1	93	20531159	20531647	2.41E-140
227	CAKMXP010000036.1	89	20530621	20531127	2.41E-140
228	CAKMXP010000093.1	92	1329281	1329769	3.06E-131
229	CAKMXP010000093.1	81	1328101	1328871	5.89E-129
230	CAKMXP010000104.1	88	414081	414899	9.26E-123
231	CAKMXP010000092.1	95	1289744	1290529	1.96E-122

Table 2.2 CERV1 env gene locations identified in the *Papio hamadryas* genome

No.	Acc.	Iden.	Start	End	E-value
1	JAKFHJ010002405.1	97	19402235	19403902	0
2	JAKFHJ010002405.1	89	13329059	13330729	0
3	JAKFHJ010002405.1	91	3344023	3342530	0
4	JAKFHJ010000002.1	97	37185584	37187254	0
5	JAKFHJ010002838.1	96	5741314	5742981	0
6	JAKFHJ010002283.1	95	16137227	16135557	0
7	JAKFHJ010002283.1	91	14267157	14265487	0
8	JAKFHJ010002604.1	94	57279800	57278130	0
9	JAKFHJ010002627.1	92	10121583	10119913	0
10	JAKFHJ010002627.1	87	13525311	13526195	0
11	JAKFHJ010002627.1	96	13526492	13526980	0
12	JAKFHJ010002505.1	91	17125055	17123385	0
13	JAKFHJ010002505.1	87	12118037	12116370	0
14	JAKFHJ010000358.1	92	3727036	3725366	0
15	JAKFHJ010000358.1	89	182157	183827	0
16	JAKFHJ010002693.1	91	9247768	9246098	0

17	JAKFHJ010002693.1	89	12101496	12099826	0
18	JAKFHJ010002693.1	98	12902136	12903338	0
19	JAKFHJ010001335.1	90	19227886	19229556	0
20	JAKFHJ010001335.1	89	14622633	14621782	2.21E-126
21	JAKFHJ010000213.1	92	206088	207758	0
22	JAKFHJ010000113.1	90	31087133	31088803	0
23	JAKFHJ010002327.1	90	964474	962804	0
24	JAKFHJ010002527.1	90	11926517	11924847	0
25	JAKFHJ010002527.1	90	12806941	12807702	0
26	JAKFHJ010002527.1	90	12807961	12808605	0
27	JAKFHJ010000546.1	90	3137268	3135598	0
28	JAKFHJ010000546.1	92	3130834	3131742	6.45E-141
29	JAKFHJ010000546.1	92	3122704	3123183	3.50E-114
30	JAKFHJ010000224.1	90	3614121	3615791	0
31	JAKFHJ010000224.1	92	3622025	3623422	0
32	JAKFHJ010000224.1	83	40433321	40432176	0
33	JAKFHJ010000224.1	92	52183950	52183078	4.43E-141
34	JAKFHJ010000224.1	93	6345862	6345101	2.11E-116
35	JAKFHJ010002001.1	89	19844966	19843437	0
36	JAKFHJ010002001.1	81	1244288	1245811	0
37	JAKFHJ010002001.1	91	1247894	1248877	3.82E-159
38	JAKFHJ010002001.1	85	19847255	19846761	3.94E-94
39	JAKFHJ010002660.1	91	12553681	12552194	0
40	JAKFHJ010002660.1	87	12529996	12529163	5.07E-146
41	JAKFHJ010002660.1	94	12528822	12528334	4.10E-111
42	JAKFHJ010000458.1	92	3075639	3074242	0
43	JAKFHJ010000458.1	89	3658641	3660005	0
44	JAKFHJ010000458.1	84	1249720	1251174	0
45	JAKFHJ010002460.1	89	1270630	1272111	0
46	JAKFHJ010002460.1	92	4142573	4143301	8.62E-107
47	JAKFHJ010002571.1	80	2163066	2164730	0
48	JAKFHJ010002571.1	94	13305238	13305723	1.46E-87
49	JAKFHJ010002223.1	93	61648060	61649379	0
50	JAKFHJ010002223.1	88	21452973	21454181	0
51	JAKFHJ010002223.1	90	39334120	39333212	0
52	JAKFHJ010002223.1	92	39333236	39332664	0
53	JAKFHJ010002223.1	89	21454168	21454644	0
54	JAKFHJ010002223.1	92	70716738	70717604	2.38E-140
55	JAKFHJ010002223.1	94	16999056	16998568	1.73E-117
56	JAKFHJ010002223.1	92	71249238	71249966	2.25E-106
57	JAKFHJ010001112.1	93	74473760	74475064	0
58	JAKFHJ010001112.1	87	18401985	18400729	0
59	JAKFHJ010001112.1	94	34852068	34853303	0
60	JAKFHJ010001112.1	90	24828267	24827383	0
61	JAKFHJ010001112.1	91	24827091	24826603	0

62	JAKFHJ010001112.1	98	42256299	42256787	4.83E-161
63	JAKFHJ010001112.1	93	18721422	18720601	8.16E-127
64	JAKFHJ010001112.1	91	75104824	75104003	7.35E-126
65	JAKFHJ010002493.1	86	55615244	55616593	0
66	JAKFHJ010002493.1	89	55587051	55587935	0
67	JAKFHJ010002493.1	90	69590907	69591626	0
68	JAKFHJ010002493.1	92	55587955	55588539	0
69	JAKFHJ010002493.1	84	69591655	69592140	0
70	JAKFHJ010002493.1	83	32068563	32067616	4.78E-169
71	JAKFHJ010002960.1	94	3585467	3584220	0
72	JAKFHJ010000125.1	98	2022938	2021736	0
73	JAKFHJ010002394.1	95	16403219	16401984	0
74	JAKFHJ010002394.1	94	16404221	16403733	0
75	JAKFHJ010002394.1	87	11130570	11131193	1.75E-145
76	JAKFHJ010002738.1	90	11305627	11306946	0
77	JAKFHJ010002471.1	94	2097266	2096028	0
78	JAKFHJ010001435.1	88	247504	246188	0
79	JAKFHJ010002224.1	87	6470018	6468906	0
80	JAKFHJ010002224.1	85	26246627	26245872	0
81	JAKFHJ010002224.1	96	6468843	6468355	0
82	JAKFHJ010002224.1	94	6467222	6466686	1.18E-79
83	JAKFHJ010002349.1	89	17269555	17270706	0
84	JAKFHJ010002349.1	93	5118918	5117881	0
85	JAKFHJ010002349.1	92	17270738	17271226	0
86	JAKFHJ010000335.1	89	6028567	6027464	0
87	JAKFHJ010000335.1	97	6027435	6026947	0
88	JAKFHJ010000001.1	93	27372303	27373031	6.76E-106
89	JAKFHJ010002937.1	88	57172616	57173764	0
90	JAKFHJ010002937.1	95	57169244	57170218	0
91	JAKFHJ010002937.1	98	57173796	57174284	0
92	JAKFHJ010002937.1	90	4689704	4688871	1.29E-158
93	JAKFHJ010002937.1	94	4688937	4688038	4.02E-148
94	JAKFHJ010002937.1	96	57165835	57166380	1.75E-127
95	JAKFHJ010002937.1	96	18385916	18386389	9.49E-76
96	JAKFHJ010002704.1	90	5973055	5974170	0
97	JAKFHJ010002605.1	86	4764411	4763566	3.99E-152
98	JAKFHJ010002605.1	86	6221782	6221132	1.73E-92
99	JAKFHJ010001890.1	88	194359	195468	0
100	JAKFHJ010001890.1	96	19967030	19967953	0
101	JAKFHJ010001890.1	81	19966287	19967027	0
102	JAKFHJ010002494.1	80	16657954	16658925	0
103	JAKFHJ010002494.1	86	4564282	4565115	0
104	JAKFHJ010002494.1	89	4565165	4565947	0
105	JAKFHJ010002826.1	89	47469687	47470730	0
106	JAKFHJ010002826.1	94	47470730	47471353	0

107	JAKFHJ010002360.1	98	4301884	4300913	0
108	JAKFHJ010002406.1	91	155462	156439	0
109	JAKFHJ010002406.1	94	156502	156990	0
110	JAKFHJ010002406.1	95	162333	162821	2.68E-166
111	JAKFHJ010002771.1	90	4445836	4446612	0
112	JAKFHJ010002771.1	98	4446597	4447319	0
113	JAKFHJ010000407.1	81	25275	26387	0
114	JAKFHJ010000111.1	98	16894	15944	0
115	JAKFHJ010000247.1	84	827839	828945	0
116	JAKFHJ010000047.1	80	479883	480926	0
117	JAKFHJ010000047.1	89	480929	481549	0
118	JAKFHJ010000447.1	87	684479	683649	0
119	JAKFHJ010002779.1	89	57019	56015	0
120	JAKFHJ010002779.1	80	57684	57004	0
121	JAKFHJ010000668.1	90	17233839	17233063	0
122	JAKFHJ010000668.1	98	17232895	17232359	0
123	JAKFHJ010002294.1	95	5359397	5358675	0
124	JAKFHJ010002294.1	90	5358678	5357734	0
125	JAKFHJ010002294.1	90	6193705	6193022	7.76E-114
126	JAKFHJ010002294.1	84	6191689	6191219	4.02E-56
127	JAKFHJ010001558.1	92	499801	500688	0
128	JAKFHJ010001558.1	82	499020	499826	0
129	JAKFHJ010002208.1	83	3310	3990	0
130	JAKFHJ010002208.1	91	4490	4978	0
131	JAKFHJ010002208.1	86	3975	4460	0
132	JAKFHJ010001002.1	97	128582	129118	5.70E-179
133	JAKFHJ010001002.1	84	121311	120670	1.58E-110
134	JAKFHJ010001002.1	95	124669	124211	2.52E-74
135	JAKFHJ010002682.1	89	8262519	8262016	3.42E-84
136	JAKFHJ010000985.1	81	10736	9510	1.13E-168
137	JAKFHJ010000985.1	81	11167	10694	7.66E-74
138	JAKFHJ010001224.1	87	16058895	16059791	1.04E-164
139	JAKFHJ010000269.1	86	2926860	2927915	0
140	JAKFHJ010001068.1	97	931637	931149	6.26E-158
141	JAKFHJ010000533.1	98	7762	8229	2.58E-157
142	JAKFHJ010000891.1	92	1322666	1321779	6.60E-142
143	JAKFHJ010002593.1	80	4967048	4967743	3.14E-140
144	JAKFHJ010002314.1	92	90297	91163	5.35E-139
145	JAKFHJ010001668.1	91	16047157	16048023	6.44E-135
146	JAKFHJ010002383.1	97	2634547	2635035	1.59E-139
147	JAKFHJ010002383.1	91	124713	125573	6.49E-130
148	JAKFHJ010001957.1	88	528384	527686	4.73E-129
149	JAKFHJ010001957.1	84	527747	527271	3.20E-61
150	JAKFHJ010002904.1	92	5667219	5668040	3.89E-127
151	JAKFHJ010001557.1	83	12497064	12497930	3.75E-118

152	JAKFHJ010002715.1	85	21391638	21392195	1.13E-158
153	JAKFHJ010002715.1	91	13137872	13137051	2.51E-125
154	JAKFHJ010002715.1	95	48294619	48295086	2.62E-115
155	JAKFHJ010002715.1	96	13139744	13139256	1.77E-107
156	JAKFHJ010002649.1	86	12971400	12970537	1.87E-122
157	JAKFHJ010000291.1	91	3305270	3306091	6.53E-122
158	JAKFHJ010001113.1	95	13290796	13291566	2.15E-120
159	JAKFHJ010001113.1	82	13287888	13288514	1.60E-80
160	JAKFHJ010002760.1	81	9207540	9208436	4.77E-118
161	JAKFHJ010000091.1	92	5539930	5539253	4.97E-118
162	JAKFHJ010003015.1	92	4472751	4472098	7.59E-118
163	JAKFHJ010003015.1	88	719564	719001	5.72E-84
164	JAKFHJ010002893.1	91	3775372	3774836	5.16E-137
165	JAKFHJ010002893.1	91	3775847	3775389	5.16E-137
166	JAKFHJ010002893.1	84	3776487	3775819	2.84E-116
167	JAKFHJ010000557.1	86	16869246	16868614	1.66E-111
168	JAKFHJ010000557.1	95	16868558	16867914	7.76E-107
169	JAKFHJ010001728.1	81	3293	2370	1.53E-115

Table 2.2 CERV1 env gene locations identified in the *Cercopithecus mona* genome

No.	Acc.	Iden.	Start	End	E-value
1	JABBCR010000770.1	90	11127948	11126278	0
2	JABBCR010000989.1	91	9785573	9787195	0
3	JABBCR010000641.1	90	48985934	48984264	0
4	JABBCR010001471.1	90	7553603	7555273	0
5	JABBCR010001471.1	89	7547181	7548332	0
6	JABBCR010001471.1	85	7548362	7548850	0
7	JABBCR010001471.1	92	1110667	1111545	1.58E-139
8	JABBCR010000478.1	89	733518	735188	0
9	JABBCR010000330.1	89	2063520	2065190	0
10	JABBCR010001785.1	89	5712257	5710587	0
11	JABBCR010001785.1	89	49301558	49303228	0
12	JABBCR010001785.1	93	49305736	49306383	5.03E-103
13	JABBCR010001269.1	90	32748328	32746706	0
14	JABBCR010000917.1	89	141260	139590	0
15	JABBCR010000917.1	90	132101	130779	0
16	JABBCR010001597.1	89	29161349	29159679	0
17	JABBCR010001722.1	89	102590	104260	0
18	JABBCR010001722.1	89	276462	277766	0
19	JABBCR010001722.1	88	61165	62277	9.16E-180
20	JABBCR010001722.1	80	60615	61214	8.91E-86
21	JABBCR010001485.1	89	41470	43140	0
22	JABBCR010001416.1	90	2791	1181	0

23	JABBCR010001325.1	90	31679415	31681082	0
24	JABBCR010001325.1	88	26498407	26499621	0
25	JABBCR010001325.1	94	31682000	31682998	2.43E-167
26	JABBCR010001325.1	91	12068980	12069537	1.69E-140
27	JABBCR010001325.1	91	12209187	12209909	9.66E-121
28	JABBCR010001799.1	89	5561	7216	0
29	JABBCR010001478.1	88	12858666	12860333	0
30	JABBCR010001478.1	87	4730270	4731316	0
31	JABBCR010001478.1	92	4731319	4731942	0
32	JABBCR010001603.1	88	705979	707649	0
33	JABBCR010000769.1	88	17924966	17926636	0
34	JABBCR010000263.1	89	43763083	43764753	0
35	JABBCR010000263.1	86	64600177	64599131	0
36	JABBCR010000263.1	92	64599131	64598508	0
37	JABBCR010000263.1	91	32195880	32196503	2.79E-131
38	JABBCR010001117.1	87	3673644	3671974	0
39	JABBCR010001117.1	90	3445956	3446738	0
40	JABBCR010001117.1	85	3445414	3445956	0
41	JABBCR010000990.1	87	7263292	7264959	0
42	JABBCR010001754.1	87	3069194	3067524	0
43	JABBCR010001755.1	87	8100661	8098991	0
44	JABBCR010001755.1	82	4094028	4095698	0
45	JABBCR010001755.1	90	44861041	44860229	1.75E-122
46	JABBCR010001755.1	93	20404487	20403999	3.85E-113
47	JABBCR010001755.1	83	44858972	44858382	5.26E-77
48	JABBCR010000587.1	87	313580	311913	0
49	JABBCR010000262.1	87	22330817	22332484	0
50	JABBCR010001649.1	89	1164456	1166051	0
51	JABBCR010001653.1	85	10382	12052	0
52	JABBCR010001658.1	85	355411	353741	0
53	JABBCR010001120.1	89	21676998	21675397	0
54	JABBCR010001596.1	85	4914086	4915756	0
55	JABBCR010000184.1	85	12366201	12367868	0
56	JABBCR010000184.1	86	11385590	11384580	4.38E-152
57	JABBCR010000184.1	90	20058786	20057884	3.59E-140
58	JABBCR010001889.1	83	11098876	11100546	0
59	JABBCR010001118.1	89	14633351	14632029	0
60	JABBCR010001118.1	88	12212151	12213383	0
61	JABBCR010001118.1	82	8605715	8605239	0
62	JABBCR010001010.1	89	1324062	1322701	0
63	JABBCR010001010.1	89	1429860	1430924	0
64	JABBCR010001010.1	85	1429244	1429867	0
65	JABBCR010000163.1	89	1083105	1084295	0
66	JABBCR010001119.1	94	17196683	17195523	0
67	JABBCR010001119.1	83	17197192	17196713	0

68	JABBCR010001032.1	93	3662362	3663474	0
69	JABBCR010001032.1	85	3661797	3662411	3.51E-106
70	JABBCR010001635.1	81	1131498	1130224	0
71	JABBCR010001635.1	82	183549	184412	2.73E-149
72	JABBCR010000137.1	87	5400	4474	0
73	JABBCR010000137.1	85	4400	3783	0
74	JABBCR010000570.1	80	34176051	34177202	0
75	JABBCR010000570.1	87	31560816	31561649	0
76	JABBCR010000570.1	91	34177238	34177720	0
77	JABBCR010000570.1	89	31561649	31562104	0
78	JABBCR010000260.1	83	316693	315731	0
79	JABBCR010001065.1	83	9649048	9649809	0
80	JABBCR010000961.1	83	9716911	9716087	0
81	JABBCR010000289.1	84	17484344	17483862	0
82	JABBCR010000507.1	89	8815570	8814437	9.19E-178
83	JABBCR010000507.1	86	8817225	8816152	1.63E-162
84	JABBCR010000507.1	86	7608581	7608075	9.26E-125
85	JABBCR010000507.1	85	7608123	7607254	1.92E-123
86	JABBCR010001796.1	92	24966	25454	4.84E-174
87	JABBCR010000853.1	91	32464796	32464173	0
88	JABBCR010000853.1	87	32465445	32464879	0
89	JABBCR010000367.1	90	3307021	3307644	1.42E-138
90	JABBCR010001408.1	83	7698806	7699648	1.69E-141
91	JABBCR010001408.1	84	7699600	7700472	1.13E-121
92	JABBCR010001323.1	90	19929010	19929912	1.41E-140
93	JABBCR010001323.1	85	19927557	19928177	7.59E-88
94	JABBCR010001564.1	92	1123090	1122212	1.56E-140
95	JABBCR010001564.1	91	1119956	1119078	6.64E-136
96	JABBCR010001064.1	86	5072237	5071416	3.56E-117

Table 2.2 CERV1 env gene locations identified in the *Lophocebus aterrimus* genome

No.	Acc.	Iden.	Start	End	E-value
1	JAKFHR010003956.1	97	162644	160974	0
2	JAKFHR010000857.1	97	2931739	2930069	0
3	JAKFHR010000857.1	90	6669145	6667475	0
4	JAKFHR010003935.1	96	2914458	2916128	0
5	JAKFHR010003935.1	98	2913121	2914065	1.85E-163
6	JAKFHR010000613.1	96	471584	469914	0
7	JAKFHR010000613.1	83	579474	581003	0
8	JAKFHR010004069.1	97	13466	15136	0
9	JAKFHR010001090.1	97	2879229	2877559	0
10	JAKFHR010000383.1	97	11349	9727	0
11	JAKFHR010000768.1	97	3639515	3641185	0

12	JAKFHR010003934.1	96	672030	670360	0
13	JAKFHR010003934.1	89	755362	756396	4.32E-159
14	JAKFHR010003934.1	87	754725	755426	4.41E-125
15	JAKFHR010004257.1	96	2327299	2328969	0
16	JAKFHR010000780.1	96	5385950	5387620	0
17	JAKFHR010001079.1	96	2863910	2862240	0
18	JAKFHR010004213.1	96	1411898	1410228	0
19	JAKFHR010001347.1	96	4711	3041	0
20	JAKFHR010001347.1	90	1674102	1672432	0
21	JAKFHR010001347.1	91	1879205	1879933	2.13E-105
22	JAKFHR010003976.1	96	332801	334471	0
23	JAKFHR010000298.1	96	32347	33987	0
24	JAKFHR010003769.1	95	20278143	20276473	0
25	JAKFHR010003769.1	92	15988471	15989199	1.30E-106
26	JAKFHR010003002.1	95	175771	174101	0
27	JAKFHR010003002.1	85	1999405	2000055	4.36E-92
28	JAKFHR010002335.1	94	8221363	8223033	0
29	JAKFHR010004095.1	96	350033	348363	0
30	JAKFHR010004132.1	96	3922	2291	0
31	JAKFHR010000520.1	93	16018	17688	0
32	JAKFHR010000469.1	92	8354	6684	0
33	JAKFHR010000924.1	92	1227395	1229065	0
34	JAKFHR010000924.1	90	9835	8684	0
35	JAKFHR010000924.1	96	8654	8166	0
36	JAKFHR010000080.1	92	7954	6284	0
37	JAKFHR010000806.1	92	10222	11892	0
38	JAKFHR010001513.1	92	2426894	2428564	0
39	JAKFHR010001513.1	94	2440684	2442354	0
40	JAKFHR010001960.1	92	4776	6446	0
41	JAKFHR010001413.1	93	990117	991787	0
42	JAKFHR010000506.1	92	6024	7694	0
43	JAKFHR010001169.1	92	23292	21622	0
44	JAKFHR010000058.1	92	348993	350663	0
45	JAKFHR010000746.1	91	7387108	7385438	0
46	JAKFHR010001028.1	91	14638	12968	0
47	JAKFHR010001028.1	95	6055	5411	0
48	JAKFHR010001028.1	95	6506	6048	0
49	JAKFHR010000391.1	92	2106494	2104824	0
50	JAKFHR010000010.1	91	36899	38569	0
51	JAKFHR010004368.1	92	6069145	6067475	0
52	JAKFHR010000680.1	91	1365918	1367588	0
53	JAKFHR010000680.1	83	1364776	1365474	1.06E-121
54	JAKFHR010000136.1	91	1151617	1149947	0
55	JAKFHR010002236.1	91	2157088	2155418	0
56	JAKFHR010002236.1	91	1869188	1870009	8.32E-125

57	JAKFHR010004324.1	92	4955789	4957459	0
58	JAKFHR010004324.1	96	4962316	4963197	1.21E-147
59	JAKFHR010004324.1	84	4961528	4962361	1.53E-145
60	JAKFHR010001213.1	91	3228624	3226954	0
61	JAKFHR010001213.1	95	3233745	3233047	7.25E-167
62	JAKFHR010001001.1	91	760843	759173	0
63	JAKFHR010000502.1	91	84767	86437	0
64	JAKFHR010004041.1	91	298705	297035	0
65	JAKFHR010004002.1	91	2025113	2023443	0
66	JAKFHR010003853.1	91	6247	4577	0
67	JAKFHR010003980.1	91	7913414	7915036	0
68	JAKFHR010000256.1	90	14050	12428	0
69	JAKFHR010003847.1	91	1737637	1735967	0
70	JAKFHR010003847.1	89	1848280	1848963	8.46E-115
71	JAKFHR010003446.1	96	8079052	8077502	0
72	JAKFHR010000291.1	91	7797173	7795506	0
73	JAKFHR010004094.1	90	8851	10521	0
74	JAKFHR010000957.1	91	1251174	1252844	0
75	JAKFHR010000957.1	94	680264	679539	8.57E-120
76	JAKFHR010000369.1	91	3705027	3703405	0
77	JAKFHR010001291.1	92	14029	15699	0
78	JAKFHR010004362.1	90	49035	47365	0
79	JAKFHR010003819.1	91	607019	605349	0
80	JAKFHR010000754.1	91	8563	10233	0
81	JAKFHR010002491.1	92	1693701	1692115	0
82	JAKFHR010001112.1	90	20107120	20105456	0
83	JAKFHR010002068.1	91	2790346	2788676	0
84	JAKFHR010002068.1	89	3628926	3630596	0
85	JAKFHR010000231.1	90	19307	17637	0
86	JAKFHR010003113.1	89	574513	576183	0
87	JAKFHR010001913.1	95	2115726	2117219	0
88	JAKFHR010003223.1	92	3771661	3773268	0
89	JAKFHR010002035.1	90	467575	465905	0
90	JAKFHR010003886.1	89	25067	26737	0
91	JAKFHR010003886.1	90	18725	20290	0
92	JAKFHR010002313.1	90	1505994	1504375	0
93	JAKFHR010002313.1	97	2164153	2163206	0
94	JAKFHR010002313.1	97	2163206	2162511	0
95	JAKFHR010004102.1	91	509774	511327	0
96	JAKFHR010004083.1	89	369279	370949	0
97	JAKFHR010001668.1	84	7179273	7177606	0
98	JAKFHR010001491.1	92	546658	548163	0
99	JAKFHR010001491.1	87	20878	21510	3.25E-112
100	JAKFHR010001491.1	95	21566	22210	7.38E-107
101	JAKFHR010001125.1	97	4058117	4059457	0

102	JAKFHR010002713.1	88	1196759	1198243	0
103	JAKFHR010003112.1	88	13165290	13163788	0
104	JAKFHR010004357.1	91	4672956	4674398	0
105	JAKFHR010000779.1	82	17726899	17728524	0
106	JAKFHR010000779.1	89	13430402	13429611	1.07E-116
107	JAKFHR010004074.1	82	14444	16108	0
108	JAKFHR010004239.1	91	30938	32290	0
109	JAKFHR010000669.1	94	12500749	12502077	0
110	JAKFHR010001691.1	93	1347469	1346165	0
111	JAKFHR010001691.1	91	819539	820360	4.41E-126
112	JAKFHR010001691.1	91	820799	821446	4.89E-103
113	JAKFHR010002558.1	90	1142339	1140993	0
114	JAKFHR010003953.1	91	14464	15780	0
115	JAKFHR010004057.1	95	13664973	13663705	0
116	JAKFHR010000191.1	87	6396163	6394877	0
117	JAKFHR010000191.1	86	6398225	6397176	0
118	JAKFHR010000191.1	92	6710027	6709206	9.43E-125
119	JAKFHR010000191.1	90	5692619	5691804	6.82E-123
120	JAKFHR010004209.1	94	53402	52110	0
121	JAKFHR010000791.1	93	720872	722137	0
122	JAKFHR010001779.1	89	8306307	8305132	0
123	JAKFHR010002112.1	94	3247	2036	0
124	JAKFHR010001335.1	93	11825217	11826455	0
125	JAKFHR010001335.1	93	10553307	10552579	2.90E-106
126	JAKFHR010004267.1	90	441369	442517	0
127	JAKFHR010004267.1	96	442547	443035	0
128	JAKFHR010000818.1	90	8226	7075	0
129	JAKFHR010000818.1	96	7045	6557	0
130	JAKFHR010001013.1	89	1927448	1926297	0
131	JAKFHR010001013.1	90	1926259	1925771	0
132	JAKFHR010002791.1	94	2266563	2267747	0
133	JAKFHR010002791.1	94	2266076	2266603	5.58E-101
134	JAKFHR010004162.1	90	195937	197127	0
135	JAKFHR010003146.1	89	1412743	1411640	0
136	JAKFHR010003146.1	96	1411608	1411120	0

Table 2.3 Full-length ORF of CERV1 env genes in primates.

No.	**Name**	**Accession number**	**Env position**
1	*Pan_troglodytes_1*	NC_072421.1	34511146..34509476
2	*Pan_troglodytes_2*	NC_072421.1	87564916..87566586
3	*Pan_troglodytes_3*	NC_072406.1	85550049..85548379
4	*Pan_troglodytes_4*	NC_072399.1	83396786..83398456
5	*Pan_troglodytes_5*	NC_072399.1	84089612..84087942
6	*Pan_troglodytes_6*	NC_072401.1	48771490..48769820
7	*Pan_troglodytes_7*	NC_072410.1	62320673..62319003
8	*Pan_troglodytes_8*	NC_072410.1	20516742..20515072
9	*Pan_troglodytes_9*	NC_072410.1	90956543..90954873
10	*Pan_troglodytes_10*	NC_072410.1	11363240..11364910
11	*Pan_troglodytes_11*	NC_072405.1	40051726..40053396
12	*Pan_troglodytes_12*	NC_072405.1	59795771..59794101
13	*Pan_troglodytes_13*	NC_072408.1	46785874..46784204
14	*Pan_troglodytes_14*	NC_072408.1	133440567..133442237
15	*Pan_troglodytes_15*	NC_072402.1	117456636..117454966
16	*Pan_troglodytes_16*	NC_072402.1	83466343..83464673
17	*Pan_troglodytes_17*	NC_072402.1	83894154..83895824
18	*Pan_troglodytes_18*	NC_072412.1	12266338..12264662
19	*Pan_troglodytes_19*	NC_072414.1	35581254..35579584
20	*Pan_troglodytes_20*	NC_072398.1	165559922..165561592
21	*Pan_troglodytes_21*	NC_072417.1	24979533..24981203
22	*Pan_troglodytes_22*	NC_072417.1	40593315..40591645
23	*Pan_troglodytes_23*	NC_072404.1	44313383..44315053
24	*Pan_troglodytes_24*	NC_072419.1	12564231..12565901
25	*Pan_troglodytes_25*	NC_072403.1	55330309..55328639
26	*Pan_troglodytes_26*	NC_072403.1	94070909..94069239
27	*Pan_troglodytes_27*	NC_072422.1	9223282..9221612
28	*Pan_troglodytes_28*	NC_072422.1	6831881..6830211
29	*Pan_troglodytes_29*	NC_072416.1	7294560..7292890
30	*Pan_troglodytes_30*	NC_072416.1	44571662..44569992
31	*Pan_troglodytes_31*	NC_072420.1	44218046..44219716
32	*Rhinopithecus_strykeri_1*	JAKFHF010004197.1	854800..856479
33	*Piliocolobus_tephrosceles_1*	NW_022319399.1	5666..7345
34	*Piliocolobus_tephrosceles_2*	NW_022319400.1	1784..105
35	*Piliocolobus_tephrosceles_3*	NW_022319738.1	5963..7642
36	*Piliocolobus_tephrosceles_4*	NW_022319739.1	1820..141
37	*Piliocolobus_tephrosceles_5*	NC_045437.1	139717580..139719259
38	*Rhinopithecus_roxellana_1*	NC_044550.1	111025602..111023923
39	*Erythrocebus_patas_1*	JAKFHT010004611.1	2980400..2978730
40	*Erythrocebus_patas_2*	JAKFHT010001557.1	462563..464233
41	*Erythrocebus_patas_3*	JAKFHT010001113.1	435386..433716
42	*Cercopithecus_mona_1*	JABBCR010000641.1	48985934..48984264
43	*Cercopithecus_mona_2*	JABBCR010000330.1	2063520..2065190

44	*Cercopithecus_mona_3*	JABBCR010001785.1	5712257..5710587
45	*Cercopithecus_mona_4*	JABBCR010001785.1	49301558..49303228
46	*Cercopithecus_mona_5*	JABBCR010000917.1	141260..139590
47	*Cercopithecus_mona_6*	JABBCR010001117.1	3673644..3671974
48	*Cercopithecus_mona_7*	JABBCR010001754.1	3069194..3067524
49	*Mandrillus_leucophaeus_1*	NW_012101987.1	2461686..2463356
50	*Theropithecus_gelada_1*	NC_037683.1	75051215..75049545
51	*Theropithecus_gelada_2*	NW_020258762.1	3403..5073
52	*Theropithecus_gelada_3*	NW_020258872.1	1837..167
53	*Theropithecus_gelada_4*	NW_020258871.1	3875..5545
54	*Theropithecus_gelada_5*	NC_037672.1	116115533..116117203
55	*Theropithecus_gelada_6*	NC_037672.1	77974926..77976596
56	*Theropithecus_gelada_7*	NC_037668.1	13559466..13557796
57	*Theropithecus_gelada_8*	NC_037668.1	191304398..191306068
58	*Theropithecus_gelada_9*	NC_037668.1	98226204..98224534
59	*Theropithecus_gelada_10*	NC_037668.1	141423731..141422061
60	*Theropithecus_gelada_11*	NW_020262659.1	1909..3579
61	*Theropithecus_gelada_12*	NW_020262659.1	7662..9332
62	*Theropithecus_gelada_13*	NC_037670.1	72847570..72849240
63	*Theropithecus_gelada_14*	NW_020264637.1	6516..8186
64	*Theropithecus_gelada_15*	NW_020264628.1	6516..8186
65	*Theropithecus_gelada_16*	NC_037671.1	106205014..106206684
66	*Theropithecus_gelada_17*	NC_037671.1	118526860..118528530
67	*Theropithecus_gelada_18*	NW_020258771.1	3420..5090
68	*Theropithecus_gelada_19*	NW_020258773.1	3419..5089
69	*Theropithecus_gelada_20*	NW_020259510.1	1761..3431
70	*Theropithecus_gelada_21*	NW_020258587.1	3580..5250
71	*Theropithecus_gelada_22*	NC_037676.1	54792694..54794364
72	*Theropithecus_gelada_23*	NC_037676.1	59609618..59611288
73	*Theropithecus_gelada_24*	NW_020251718.1	6132..7802
74	*Theropithecus_gelada_25*	NW_020251728.1	6131..7801
75	*Theropithecus_gelada_26*	NW_020266101.1	5346..7016
76	*Theropithecus_gelada_27*	NW_020266091.1	5346..7016
77	*Theropithecus_gelada_28*	NC_037675.1	48857919..48859589
78	*Theropithecus_gelada_29*	NW_020258583.1	3582..5252
79	*Theropithecus_gelada_30*	NW_020259226.1	1941..271
80	*Theropithecus_gelada_31*	NW_020259151.1	3609..5279
81	*Theropithecus_gelada_32*	NC_037685.1	91416375..91414705
82	*Theropithecus_gelada_33*	NW_020253039.1	1810..140
83	*Theropithecus_gelada_34*	NW_020258864.1	3961..5631
84	*Theropithecus_gelada_35*	NW_020253050.1	1810..140
85	*Theropithecus_gelada_36*	NW_020258992.1	1773..103
86	*Lophocebus_aterrimus_1*	JAKFHR010003956.1	162644..160974
87	*Lophocebus_aterrimus_2*	JAKFHR010000857.1	2931739..2930069
88	*Lophocebus_aterrimus_3*	JAKFHR010003935.1	2914458..2916128
89	*Lophocebus_aterrimus_4*	JAKFHR010000613.1	471584..469914

90	*Lophocebus_aterrimus_5*	JAKFHR010004069.1	13466..15136
91	*Lophocebus_aterrimus_6*	JAKFHR010003934.1	672030..670360
92	*Lophocebus_aterrimus_7*	JAKFHR010001079.1	2863910..2862240
93	*Lophocebus_aterrimus_8*	JAKFHR010001347.1	4711..3041
94	*Lophocebus_aterrimus_9*	JAKFHR010000520.1	16018..17688
95	*Lophocebus_aterrimus_10*	JAKFHR010000469.1	8354..6684
96	*Lophocebus_aterrimus_11*	JAKFHR010000924.1	1227395..1229065
97	*Lophocebus_aterrimus_12*	JAKFHR010000080.1	7954..6284
98	*Lophocebus_aterrimus_13*	JAKFHR010000806.1	10222..11892
99	*Lophocebus_aterrimus_14*	JAKFHR010001513.1	2426894..2428564
100	*Lophocebus_aterrimus_15*	JAKFHR010001960.1	4776..6446
101	*Lophocebus_aterrimus_16*	JAKFHR010000058.1	348993..350663
102	*Lophocebus_aterrimus_17*	JAKFHR010000746.1	7387108..7385438
103	*Lophocebus_aterrimus_18*	JAKFHR010001028.1	14638..12968
104	*Lophocebus_aterrimus_19*	JAKFHR010000010.1	36899..38569
105	*Lophocebus_aterrimus_20*	JAKFHR010000136.1	1151617..1149947
106	*Lophocebus_aterrimus_21*	JAKFHR010004041.1	298705..297035
107	*Lophocebus_aterrimus_22*	JAKFHR010003853.1	6247..4577
108	*Lophocebus_aterrimus_23*	JAKFHR010000754.1	8563..10233
109	*Lophocebus_aterrimus_24*	JAKFHR010002068.1	2790346..2788676
110	*Papio_anubis_1*	NC_044987.1	69751617..69753287
111	*Papio_anubis_2*	NW_022163318.1	1427..3097
112	*Papio_anubis_3*	NW_022163319.1	1427..3097
113	*Papio_anubis_4*	NC_044980.1	155887793..155889463
114	*Papio_anubis_5*	NW_022161486.1	11580..9910
115	*Papio_anubis_6*	NC_044996.1	96027437..96029107
116	*Papio_anubis_7*	NC_044996.1	9188082..9186412
117	*Papio_anubis_8*	NC_044996.1	65087665..65085995
118	*Papio_anubis_9*	NC_044982.1	14996222..14994552
119	*Papio_anubis_10*	NC_044976.1	119624399..119626069
120	*Papio_anubis_11*	NC_044986.1	108406298..108404628
121	*Papio_anubis_12*	NC_044990.1	39324115..39322445
122	*Papio_hamadryas_1*	JAKFHJ010000002.1	37185584..37187254
123	*Papio_hamadryas_2*	JAKFHJ010002627.1	10121583..10119913
124	*Papio_hamadryas_3*	JAKFHJ010000358.1	3727036..3725366
125	*Papio_hamadryas_4*	JAKFHJ010002693.1	9247768..9246098
126	*Macaca_nigra_1*	CAKMXP010000015.1	36545260..36543590
127	*Macaca_nigra_2*	CAKMXP010000006.1	37974950..37976620
128	*Macaca_nigra_3*	CAKMXP010000014.1	69856343..69858013
129	*Macaca_nigra_4*	CAKMXP010000005.1	42457755..42456085
130	*Macaca_nigra_5*	CAKMXP010000005.1	99118400..99120070
131	*Macaca_nigra_6*	CAKMXP010000002.1	26339859..26338189
132	*Macaca_nigra_7*	CAKMXP010000012.1	48040111..48041781
133	*Macaca_nigra_8*	CAKMXP010000012.1	38355588..38353918
134	*Macaca_nigra_9*	CAKMXP010000012.1	77355952..77357622
135	*Macaca_nigra_10*	CAKMXP010000011.1	58541839..58540169

136	*Macaca_nigra_11*	CAKMXP010000034.1	16915861..16917531
137	*Macaca_nigra_12*	CAKMXP010000035.1	12002575..12000905
138	*Macaca_nigra_13*	CAKMXP010000024.1	35699792..35701462
139	*Macaca_nigra_14*	CAKMXP010000022.1	42717178..42715508
140	*Macaca_nigra_15*	CAKMXP010000013.1	17911521..17909851
141	*Macaca_nigra_16*	CAKMXP010000029.1	30972548..30974218
142	*Macaca_nigra_17*	CAKMXP010000018.1	100573..102243
143	*Macaca_nigra_18*	CAKMXP010000007.1	13213909..13215579
144	*Macaca_nigra_19*	CAKMXP010000025.1	3228497..3226827
145	*Macaca_silenus_1*	JAKFHO010001111.1	17061132..17059462
146	*Macaca_silenus_2*	JAKFHO010000446.1	24848935..24847265
147	*Macaca_silenus_3*	JAKFHO010001189.1	11196462..11194792
148	*Macaca_silenus_4*	JAKFHO010000934.1	11282357..11280687
149	*Macaca_silenus_5*	JAKFHO010000922.1	39392453..39390783
150	*Macaca_silenus_6*	JAKFHO010000922.1	69310638..69312308
151	*Macaca_silenus_7*	JAKFHO010000158.1	270666..272336
152	*Macaca_silenus_8*	JAKFHO010001078.1	11354687..11356357
153	*Macaca_silenus_9*	JAKFHO010001244.1	1498625..1496955
154	*Macaca_silenus_10*	JAKFHO010001766.1	7580926..7582596
155	*Macaca_silenus_11*	JAKFHO010001267.1	12452356..12454026
156	*Macaca_silenus_12*	JAKFHO010000001.1	41555448..41553778
157	*Macaca_silenus_13*	JAKFHO010001367.1	11841341..11839671
158	*Macaca_fascicularis_1*	NC_052268.1	101564604..101562934
159	*Macaca_fascicularis_2*	NC_052268.1	59100091..59101761
160	*Macaca_fascicularis_3*	NC_052268.1	83358953..83360644
161	*Macaca_fascicularis_4*	NC_052261.1	166613093..166611423
162	*Macaca_fascicularis_5*	NC_052261.1	121455514..121453844
163	*Macaca_fascicularis_6*	NC_052261.1	30561008..30562678
164	*Macaca_fascicularis_7*	NC_052255.1	2306352..2308022
165	*Macaca_fascicularis_8*	NC_052265.1	92440839..92439169
166	*Macaca_fascicularis_9*	NC_052263.1	64271764..64273434
167	*Macaca_fascicularis_10*	NC_052258.1	158423631..158421961
168	*Macaca_fascicularis_11*	NC_052256.1	14400448..14398778
169	*Macaca_fascicularis_12*	NC_052269.1	60796463..60794793
170	*Macaca_fascicularis_13*	NC_052269.1	18494874..18496544
171	*Macaca_fascicularis_14*	NC_052271.1	50298789..50297119
172	*Macaca_fascicularis_15*	NC_052267.1	24549796..24551466
173	*Macaca_fascicularis_16*	NC_052260.1	95034047..95035717
174	*Macaca_fascicularis_17*	NC_052272.1	26177053..26175383
175	*Macaca_fascicularis_18*	NC_052270.1	59247445..59249115
176	*Macaca_fascicularis_19*	NW_025540891.1	19510..21180
177	*Macaca_fascicularis_20*	NC_052257.1	76290616..76288946
178	*Macaca_thibetana_1*	NC_065588.1	91796391..91794721
179	*Macaca_thibetana_2*	NC_065588.1	30711446..30709776
180	*Macaca_thibetana_3*	NC_065580.1	131598437..131600107
181	*Macaca_thibetana_4*	NC_065580.1	172479773..172478103

182	*Macaca_thibetana_5*	NC_065582.1	11819860..11818190
183	*Macaca_thibetana_6*	NC_065582.1	40898524..40900194
184	*Macaca_thibetana_7*	NC_065582.1	103810014..103811684
185	*Macaca_thibetana_8*	NC_065582.1	103787482..103789152
186	*Macaca_thibetana_9*	NC_065582.1	61603811..61602141
187	*Macaca_thibetana_10*	NC_065582.1	135404172..135405845
188	*Macaca_thibetana_11*	NC_065598.1	85009374..85011044
189	*Macaca_thibetana_12*	NC_065584.1	33480152..33481822
190	*Macaca_thibetana_13*	NC_065584.1	42551289..42552959
191	*Macaca_thibetana_14*	NC_065599.1	4061097..4059427
192	*Macaca_thibetana_15*	NC_065591.1	65050041..65051711
193	*Macaca_thibetana_16*	NC_065591.1	42455571..42453901
194	*Macaca_thibetana_17*	NC_065579.1	12679303..12680973
195	*Macaca_thibetana_18*	NC_065590.1	72525077..72523407
196	*Macaca_thibetana_19*	NC_065590.1	76482294..76480624
197	*Macaca_thibetana_20*	NC_065585.1	92124942..92126615
198	*Macaca_thibetana_21*	NC_065581.1	27853058..27854728
199	*Macaca_mulatta_1*	NC_041755.1	57540227..57538557
200	*Macaca_mulatta_2*	NC_041760.1	168720624..168718954
201	*Macaca_mulatta_3*	NC_041757.1	13504084..13502414
202	*Macaca_mulatta_4*	NC_041756.1	51734746..51733076
203	*Macaca_mulatta_5*	NC_041754.1	104508124..104509794
204	*Macaca_mulatta_6*	NC_041754.1	2370944..2372614
205	*Macaca_mulatta_7*	NC_041754.1	167262728..167261058
206	*Macaca_mulatta_8*	NC_041766.1	78171365..78169695
207	*Macaca_mulatta_9*	NC_041766.1	1554085..1555755
208	*Macaca_mulatta_10*	NC_041768.1	107005390..107003720
209	*Macaca_mulatta_11*	NC_041765.1	5614666..5612996
210	*Macaca_mulatta_12*	NC_041767.1	68406838..68405168
211	*Macaca_mulatta_13*	NC_041767.1	42557609..42555939
212	*Macaca_mulatta_14*	NC_041774.1	56654836..56653166
213	*Macaca_mulatta_15*	NC_041774.1	100578184..100576514
214	*Macaca_mulatta_16*	NC_041763.1	57998518..58000188
215	*Macaca_mulatta_17*	NC_041763.1	91714340..91716010
216	*Macaca_mulatta_18*	NC_041759.1	14634062..14632392
217	*Macaca_mulatta_19*	NW_021161888.1	2526..4196
218	*Macaca_mulatta_20*	NC_041771.1	60859583..60857913
219	*Macaca_mulatta_21*	NC_041770.1	10989258..10987588
220	*Macaca_mulatta_22*	NC_041758.1	231340..233010
221	*Pan_paniscus_1*	NC_073259.2	65032142..65030472
222	*Pan_paniscus_2*	NC_073259.2	20162288..20160618
223	*Pan_paniscus_3*	NC_073259.2	11095466..11097136
224	*Pan_paniscus_4*	NC_085927.1	58955161..58956831
225	*Pan_paniscus_5*	NC_073269.2	43652227..43650557
226	*Pan_paniscus_6*	NC_085926.1	44720755..44719085
227	*Pan_paniscus_7*	NC_073257.2	56070927..56069257

228	*Pan_paniscus_8*	NC_073257.2	76345198..76343528
229	*Pan_paniscus_9*	NC_073258.2	78498479..78500149
230	*Pan_paniscus_10*	NC_073256.2	87624744..87623074
231	*Pan_paniscus_11*	NC_073261.2	95199209..95197539
232	*Pan_paniscus_12*	NC_073261.2	42398429..42400099
233	*Pan_paniscus_13*	NC_073267.2	10863599..10861929
234	*Pan_paniscus_14*	NC_073255.2	53969058..53970728
235	*Pan_paniscus_15*	NC_073273.2	38798134..38799804
236	*Pan_paniscus_16*	NC_073273.2	40509822..40511492
237	*Pan_paniscus_17*	NC_073252.2	157474306..157472636
238	*Pan_paniscus_18*	NC_073252.2	57130677..57129007
239	*Pan_paniscus_19*	NC_073252.2	114071158..114069488
240	*Pan_paniscus_20*	NC_073253.2	54203110..54201440
241	*Pan_paniscus_21*	NC_073253.2	91587772..91586102
242	*Pan_paniscus_22*	NC_073260.2	57323808..57325478
243	*Pan_paniscus_23*	NC_073266.2	53243126..53241456
244	*Gorilla_gorilla_1*	NC_073247.2	154819965..154821635
245	*Gorilla_gorilla_2*	NC_073247.2	120978923..120980593
246	*Gorilla_gorilla_3*	NC_073247.2	125864299..125865969
247	*Gorilla_gorilla_4*	NC_073247.2	63587424..63585754
248	*Gorilla_gorilla_5*	NC_073247.2	85178664..85180334
249	*Gorilla_gorilla_6*	NC_073235.2	60397899..60396229
250	*Gorilla_gorilla_7*	NC_073235.2	116154567..116152897
251	*Gorilla_gorilla_8*	NC_073235.2	101594741..101596411
252	*Gorilla_gorilla_9*	NC_073234.2	66970660..66972330
253	*Gorilla_gorilla_10*	NC_073234.2	139529935..139528265
254	*Gorilla_gorilla_11*	NC_073228.2	98492720..98494390
255	*Gorilla_gorilla_12*	NC_073228.2	87046902..87048620
256	*Gorilla_gorilla_13*	NC_073231.2	6517943..6516273
257	*Gorilla_gorilla_14*	NC_073231.2	65671744..65673414
258	*Gorilla_gorilla_15*	NC_073227.2	127833441..127835111
259	*Gorilla_gorilla_16*	NC_073227.2	174968823..174967153
260	*Gorilla_gorilla_17*	NC_086017.1	157425383..157427053
261	*Gorilla_gorilla_18*	NC_073232.2	142655034..142653364
262	*Gorilla_gorilla_19*	NC_073236.2	72673020..72671350
263	*Gorilla_gorilla_20*	NC_073236.2	121599707..121598037

Table 2.4 Reference sequence for the SLC52A1 and SLC52A2 genes.

Accession no.	Gene name	Species
XP_012315969.1	SLC52A1	*Aotus nancymaae*
XP_012328721.1	SLC52A2	*Aotus nancymaae*
XP_008064142.2	SLC52A1	*Carlito syrichta*
XP_008068176.1	SLC52A2	*Carlito syrichta*
XP_017352210.1	SLC52A1	*Cebus imitator*
XP_017367150.1	SLC52A2	*Cebus imitator*
XP_004058414.4	SLC52A1	*Gorilla gorilla*
XP_018888082.1	SLC52A2	*Gorilla gorilla*
XP_031993785.1	SLC52A1	*Hylobates moloch*
XP_032616695.1	SLC52A2	*Hylobates moloch*
XP_045383100.1	SLC52A1	*Lemur catta*
XP_045417156.1	SLC52A2	*Lemur catta*
XP_050621558.1	SLC52A1	*Macaca thibetana*
XP_050656898.1	SLC52A2	*Macaca thibetana*
XP_011850087.1	SLC52A1	*Mandrillus leucophaeus*
XP_011831087.1	SLC52A2	*Mandrillus leucophaeus*
XP_030655562.1	SLC52A1	*Nomascus leucogenys*
XP_003281204.1	SLC52A2	*Nomascus leucogenys*
XP_053424185.1	SLC52A1	*Nycticebus coucang*
XP_053414314.1	SLC52A2	*Nycticebus coucang*
XP_003791274.1	SLC52A1	*Otolemur garnettii*
XP_003802391.1	SLC52A2	*Otolemur garnettii*
XP_023040459.1	SLC52A1	*Piliocolobus tephrosceles*
XP_023056633.1	SLC52A2	*Piliocolobus tephrosceles*
XP_002826937.1	SLC52A1	*Pongo abelii*
XP_009242460.2	SLC52A2	*Pongo abelii*
XP_054315395.1	SLC52A1	*Pongo pygmaeus*
XP_054355269.1	SLC52A2	*Pongo pygmaeus*
XP_012515363.1	SLC52A1	*Propithecus coquereli*
XP_012503601.1	SLC52A2	*Propithecus coquereli*
XP_032124620.1	SLC52A1	*Sapajus apella*
XP_032112066.1	SLC52A2	*Sapajus apella*
XP_025219655.1	SLC52A1	*Theropithecus gelada*
XP_025250752.1	SLC52A2	*Theropithecus gelada*
XP_010361888.1	SLC52A1	*Rhinopithecus roxellana*
XP_010373525.1	SLC52A2	*Rhinopithecus roxellana*
XP_005582692.2	SLC52A1	*Macaca fascicularis*
XP_005564405.2	SLC52A2	*Macaca fascicularis*
XP_001099620.2	SLC52A1	*Macaca mulatta*
XP_001090939.2	SLC52A2	*Macaca mulatta*
XP_003810223.1	SLC52A1	*Pan paniscus*
XP_003819471.1	SLC52A2	*Pan paniscus*
XP_523560.4	SLC52A1	*Pan troglodytes*

XP_016801823.1	SLC52A2	*Pan troglodytes*
XP_003912217.1	SLC52A1	*Papio anubis*
XP_003903322.2	SLC52A2	*Papio anubis*
NP_060456.3	SLC52A1	*Homo sapiens*
NP_001350047.1	SLC52A2	*Homo sapiens*
XP_017721360.1	SLC52A1	*Rhinopithecus bieti*
XP_017738983.1	SLC52A2	*Rhinopithecus bieti*
XP_055143637.1	SLC52A2	*Symphalangus syndactylus*
XP_003942096.1	SLC52A2	*Saimiri boliviensis boliviensis*
XP_012615726.1	SLC52A2	*Microcebus murinus*
NP_212134.3	SLC52A3	*Homo sapiens*
XP_008008193.1	SLC52A1	*Chlorocebus sabaeus*
XP_037855901.1	SLC52A2	*Chlorocebus sabaeus*

Table 2.5 Sequence Read Archive (SRA) accession numbers of the references used in this study.

Run	Species	gtf	Organism	Tissue	Bioproject
SRR1758915	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Bone Marrow	PRJNA271912
SRR1758916	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Brain Cerebellum	PRJNA271912
SRR1758917	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Brain Frontal Cortex	PRJNA271912
SRR1758918	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Brain Pituitary	PRJNA271912
SRR1758919	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Brain Temporal Lobe	PRJNA271912
SRR1758920	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Colon	PRJNA271912
SRR1758921	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Heart	PRJNA271912
SRR1758922	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Kidney	PRJNA271912
SRR1758924	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Liver	PRJNA271912
SRR1758926	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Lung	PRJNA271912
SRR1758928	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Lymph Node	PRJNA271912
SRR1758929	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Skeletal Muscle	PRJNA271912
SRR1758931	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Spleen	PRJNA271912
SRR1758933	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Thymus	PRJNA271912
SRR1758934	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Bone Marrow	PRJNA271912
SRR1758935	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Brain Cerebellum	PRJNA271912
SRR1758936	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Brain Frontal Cortex	PRJNA271912
SRR1758937	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Brain Pituitary	PRJNA271912
SRR1758938	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Brain Temporal Lobe	PRJNA271912
SRR1758939	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Colon	PRJNA271912
SRR1758940	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Heart	PRJNA271912
SRR1758941	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Kidney	PRJNA271912
SRR1758942	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Liver	PRJNA271912
SRR1758943	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Lung	PRJNA271912

SRR1758944	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Lymph Node	PRJNA271912
SRR1758945	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Skeletal Muscle	PRJNA271912
SRR1758946	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Spleen	PRJNA271912
SRR1758947	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Thymus	PRJNA271912
SRR1602559	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Bone Marrow	PRJNA261940
SRR1602561	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Brain Cerebellum	PRJNA261940
SRR1602563	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Brain Frontal Cortex	PRJNA261940
SRR1602564	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Brain Pituitary	PRJNA261940
SRR1602566	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Heart	PRJNA261940
SRR1602568	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Liver	PRJNA261940
SRR1602569	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Lung	PRJNA261940
SRR1602572	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Thymus	PRJNA261940
SRR1602567	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Kidney	PRJNA261940
SRR1602570	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Skeletal Muscle	PRJNA261940
SRR1602571	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Spleen	PRJNA261940
SRR1602573	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Testis	PRJNA261940
SRR1602574	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Ovaries	PRJNA261940
SRR1602576	chimpanzee	panTro6.ncbiRefSeq.gtf	*Pan troglodytes*	Whole blood	PRJNA261948
SRR1602577	crab_eating_macaque	macFas5.ncbiRefSeq.gtf	*Macaca fascicularis*	Whole blood	PRJNA261948
SRR1602582	rheusus_macaque	rheMac10.ncbiRefSeq.gtf	*Macaca mulatta*	Whole blood	PRJNA261948
SRR8062405	African_green_monkey	Vero_WHO_p1.0_genomic.gtf	Chlorocebus sabaeus	Cos7 cell line	PRJNA496616
SRR23978815	African_green_monkey	Vero_WHO_p1.0_genomic.gtf	Chlorocebus sabaeus	Vero cell line	PRJNA949173
SRR27812394	Human	GRCh38.p14_genomic.gtf	Homo sapiens	Hs68 cell line	PRJNA1071594

Table 2.6 Integration timings and orthologs of the CERV1 env gene in primates

Host	Accession no.	Position	Integrated age MYA	Ortholog (Species)
P. paniscus	NC_073262.2	57532816..57543193	2.40–6.40	*P. troglodytes*
P. troglodytes	NC_072411.2	56889829..56900169	2.40–6.40	-
G. gorilla	-	-	unknown	N.D
E. patas	-	-	unknown	N.D
C. mona	-	-	unknown	N.D
P. hamadryas	JAKFHJ010002493.1	55581515..55590272	4.57–4.52	*P. anubis*
P. anubis	CM018184.2	155882257..155891076	4.57–4.52	-
T. gelada	NC_037673.1	159770546..159779372	4.57–4.52	*P. anubis*
L. aterrimus	-	-	unknown	N.D
M. leucophaeus	-	-	unknown	N.D
M. thibetana	-	-	unknown	N.D
M. mulatta	NC_041767.1	24248568..24258635	4.50–10.45	*M. nigra*
M. fascicularis	NC_088388.1	24364398..24374465	4.50–10.45	*M. nigra*
M. silenus	JAKFHO010001111.1	17066737..17058220	4.50–10.45	*M. nigra*
M. nigra	CAKMXP010000015.1	36552487..36542349	4.50–10.45	-
R. roxellana	NC_044550.1	111032812..111022679	2.19–12.79	*R. strykeri*
R. strykeri	JAKFHF010004197.1	848205..857643	2.19–12.79	-
P. tephrosceles	-	-	unknown	N.D

MYA, million years ago; N.D., orthologs of the CERV1-env gene were not detected; unknown, integration

Figure

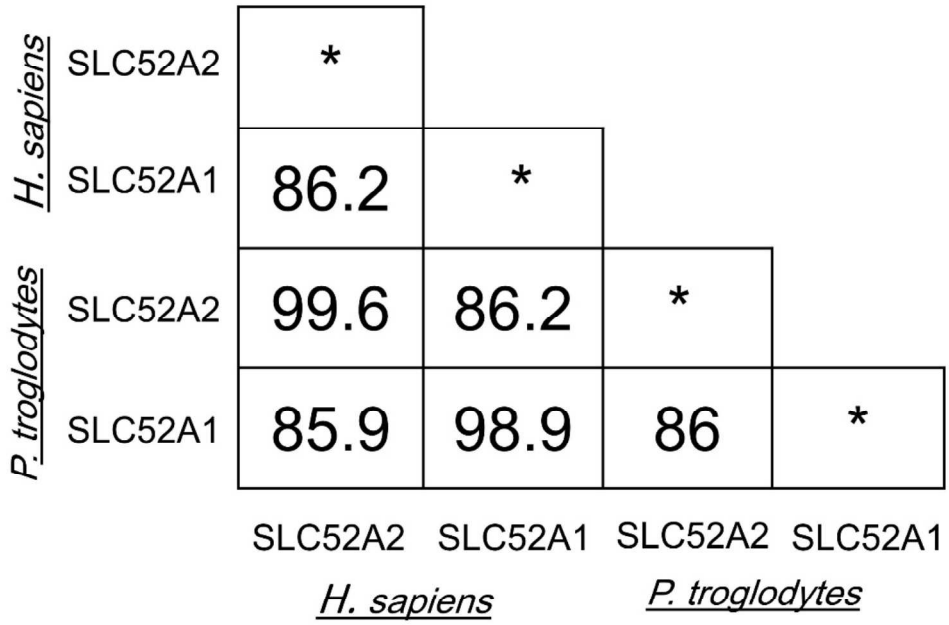


Figure 2.1 Amino acid sequence similarity of SLC52A1 and SLC52A2 between *Homo sapiens* and *Pan troglodytes*. Pairwise alignments were performed using MUSCLE, and the percentage of amino acid sequence similarity between the two species is shown in the intersecting cells.

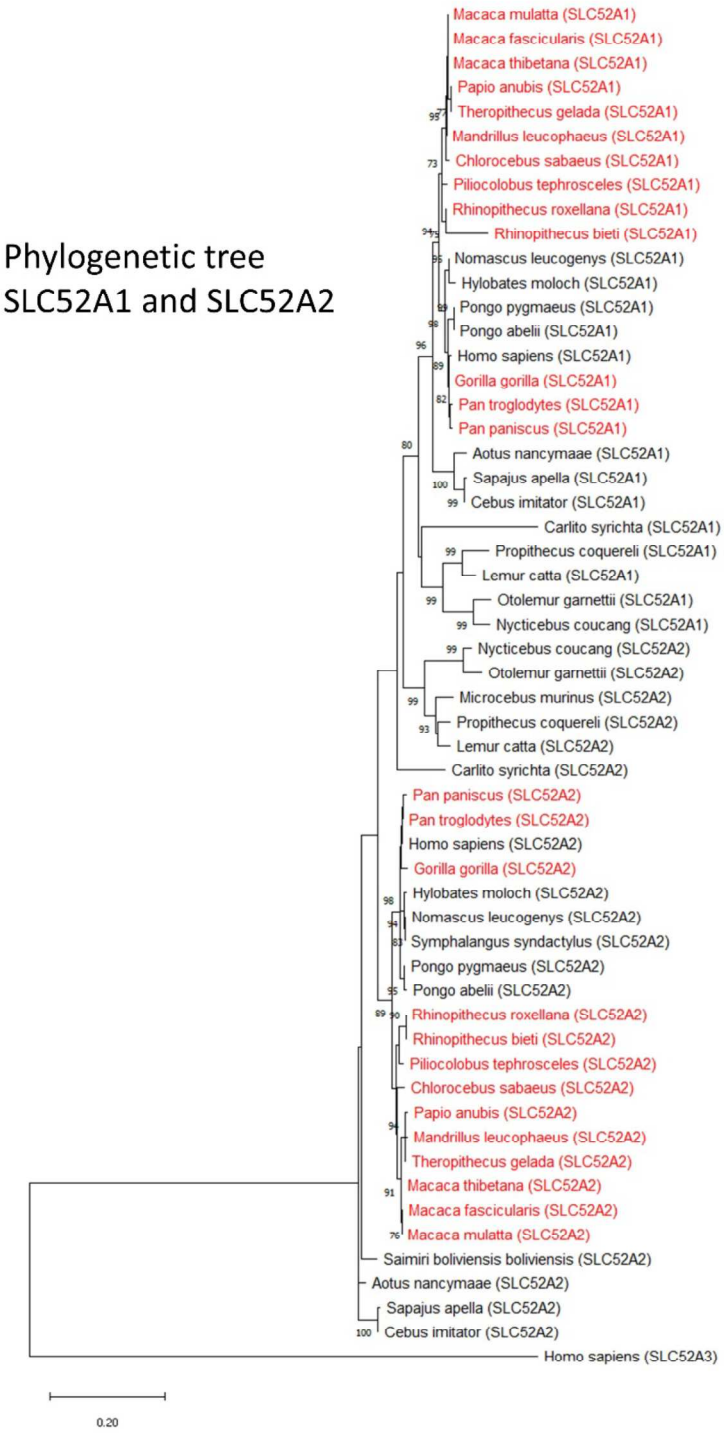


Figure 2.2 Phylogenetic analysis of the SLC52A2 and SLC52A1 amino acid sequences in primates. Phylogenetic tree for the SLC52A2 and SLC52A1. Human SLC52A3 was used as an outgroup. The evolutionary history was inferred using the maximum likelihood method and the JTT matrix-based model. Bootstrap values >70% are shown adjacent to the respective branches. The tree is drawn to

scale, with branch lengths representing the number of substitutions per site. This analysis involved 56 amino acid sequences. The red color indicates the species infected with CERV1.

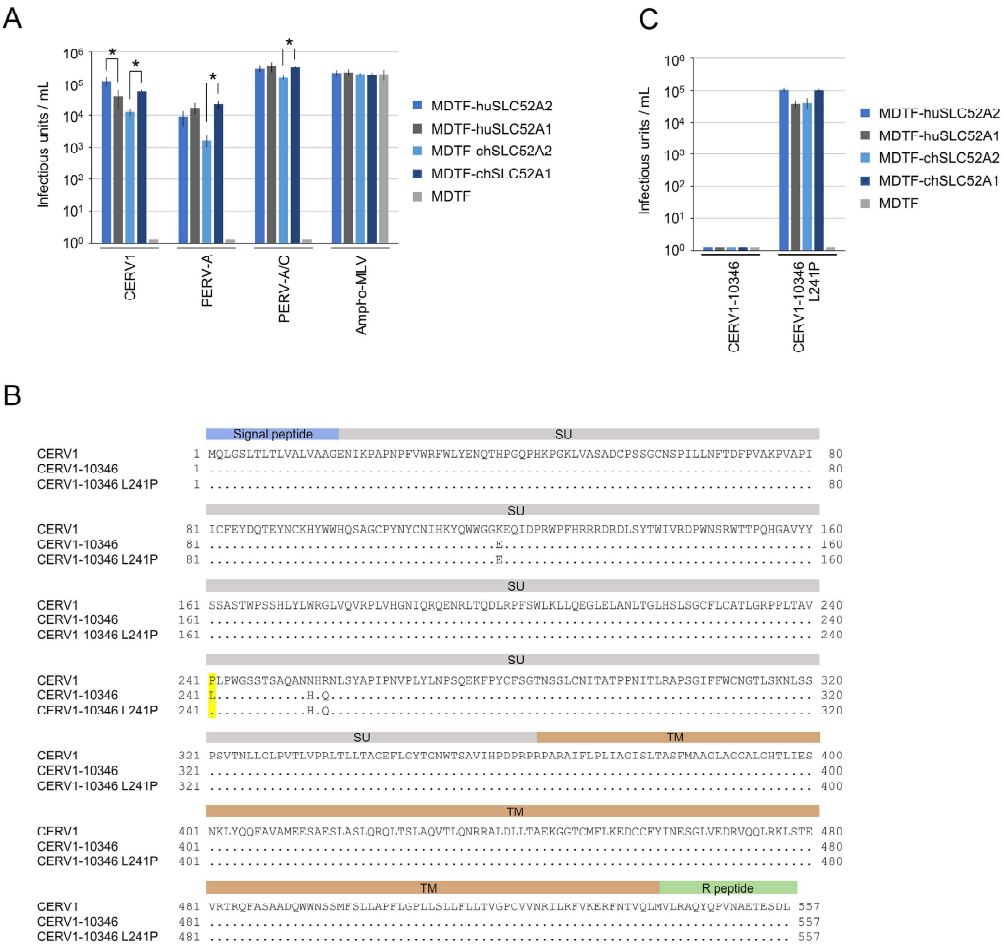


Figure 2.3 Infection of CERV1, PERV-A, and PERV-A/C. (A) Infection of LacZ-coding CERV1, PERV-A, and PERV-A/C Env-pseudotyped viruses in MDTF-huSLC52A2, MDTF-huSLC52A1, MDTF-chSLC52A2, MDTF-chSLC52A1, and MDTF cells. (B) Infection of LacZ-coding CERV1 10346 and CERV1-10346 L241P Env-pseudotyped viruses in MDTF-huSLC52A2, MDTF-huSLC52A1, MDTF-chSLC52A2, MDTF-chSLC52A1, and MDTF cells. (C) Alignment of the amino acid sequences of Envs of CERV1, CERV1-10346, and CERV1-10346 L241P. Envs are shown using single-letter amino acid codes. Leucine (L) and proline (P) at position 241 in CERV1 Env are indicated for comparison. Dots, conserved amino acid residues; SU, surface unit; TM, transmembrane domain. The data represent the averages of three independent experiments, with the standard deviations shown (A and B).

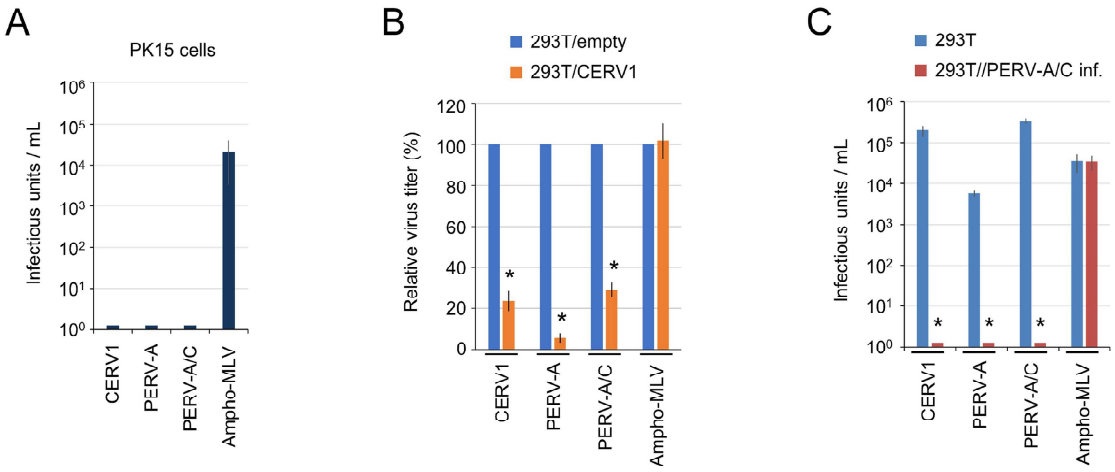


Figure 2.4 Interference group comparison of CERV1, PERV-A, and PERV-A/C. Infection of CERV1, PERV-A, and PERV-A/C Env-pseudotyped viruses in (A) PK15 cells, (B) HEK293T cells transiently expressing CERV1 Env (293/CERV1 cells) or empty plasmid (293/empty cells), and (C) HEK293T cells persistently infected with PERV-A/C (293T//PERV-A/C inf. cells) or 293T cells. Comparisons with HEK293T/empty or control cells were performed using Student's *t*-test ($^{*}P < .01$) in (B, C). The y-axis indicates infectious units per millilitre of virus (A, C) or relative virus titre (%) (B). The infectious titres for 293T/empty cells (B) are as follows. CERV1: $(1.38 \pm 0.62) \times 10^5$ IU/ml PERV-A: $(8.64 \pm 4.36) \times 10^3$ IU/ml, PERV-A/C: $(2.09 \pm 0.12) \times 10^5$ IU/ml, Ampho-MLV: $(9.16 \pm 0.52) \times 10^4$ IU/ml. The data represent the averages of three independent experiments, with the standard deviations shown.

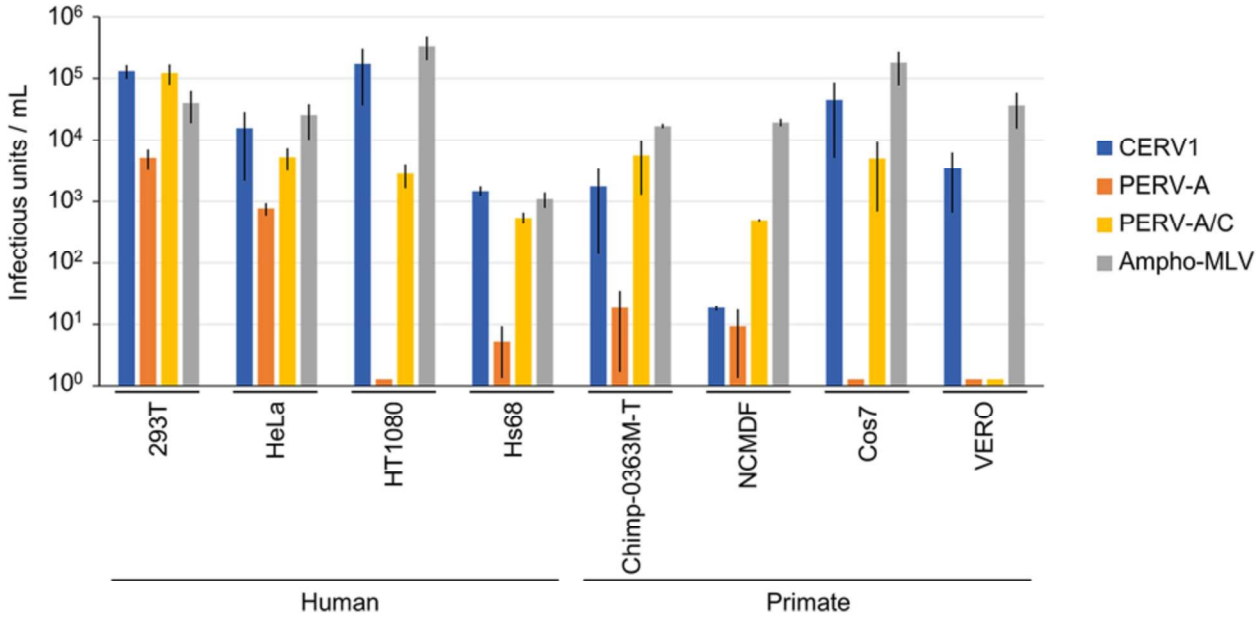


Figure 2.5 Comparison of the infectivity of CERV1, PERV-A, and PERV-A/C in human and primate cell lines. The infectivity of CERV1, PERV-A, and PERV-A/C Env-pseudotyped viruses was assessed in human and primate cell lines. The data represent the averages of three independent experiments, with the standard deviations shown.

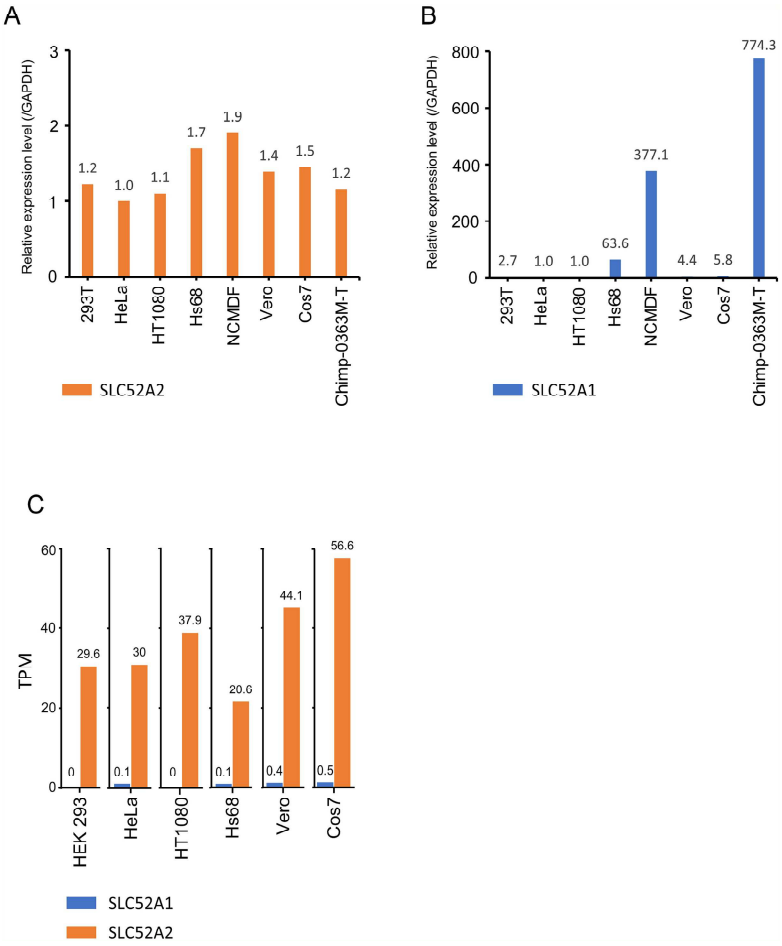


Figure 2.9 Expression of SLC52A1 and SLC52A2 in human and non-human primate cell lines. Quantification of SLC52A2 (A) and SLC52A1 (B) transcripts by quantitative RT-PCR in cell lines. The x-axis indicates the analysed samples. The y-axis indicates the expression level normalized to the expression of GAPDH. In each experiment, the normalized expression in HeLa cells is set to 1 for comparison across cell lines. (C) Expression of SLC52A1 and SLC52A2 in cell line. The x-axis indicates the analysed samples. The y-axis indicates TPM.

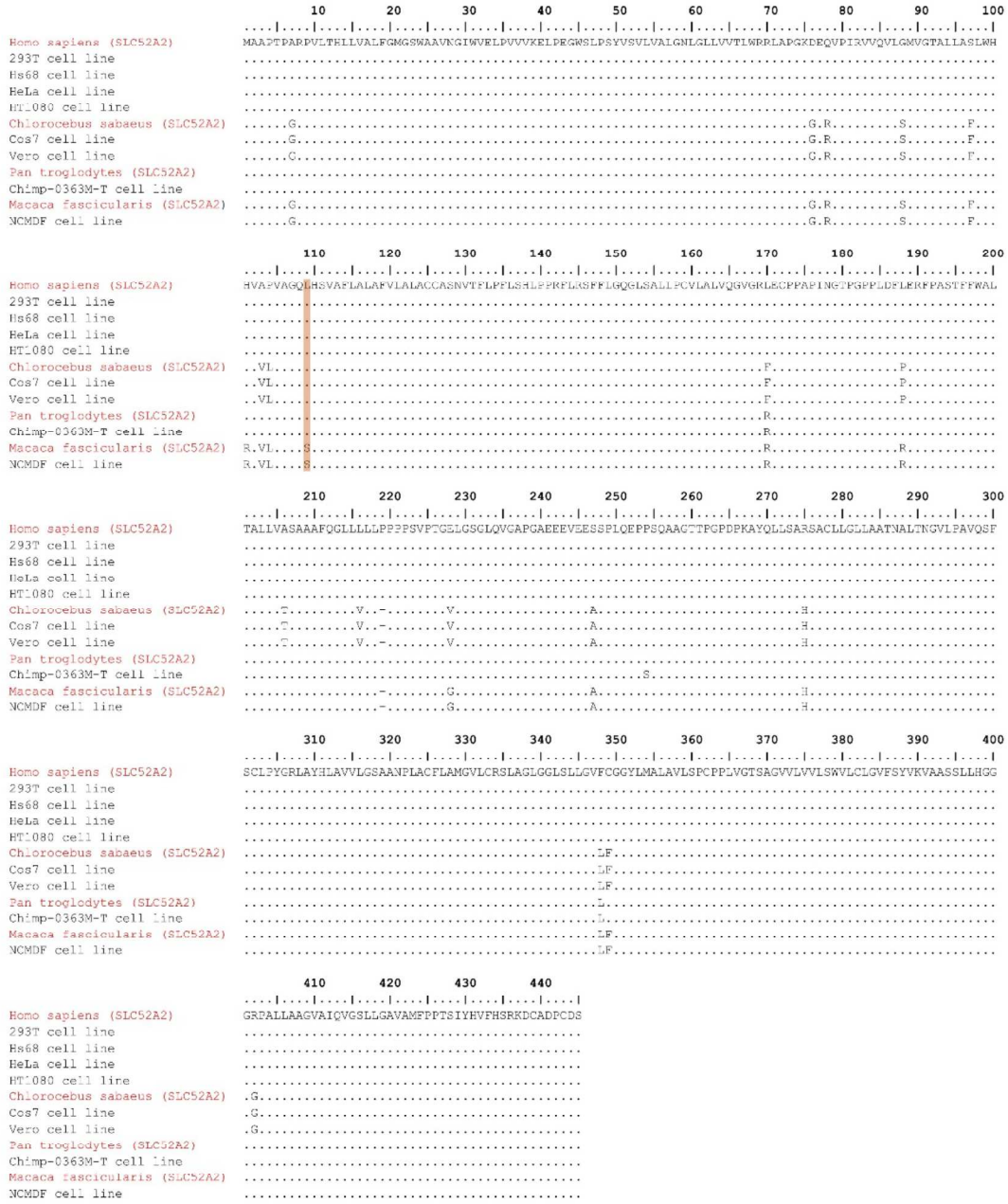


Figure 2.7 Comparison of the SLC52A2 gene sequence in the human cell line and primate cell line. Amino acid sequence identity of the SLC52A2 gene is shown. Hyphens indicate gaps in the amino acid sequence, while dots represent conserved amino acid residues. The orange color indicates the positions of leucine and serine amino acids. The red label indicates the reference amino acid sequences.

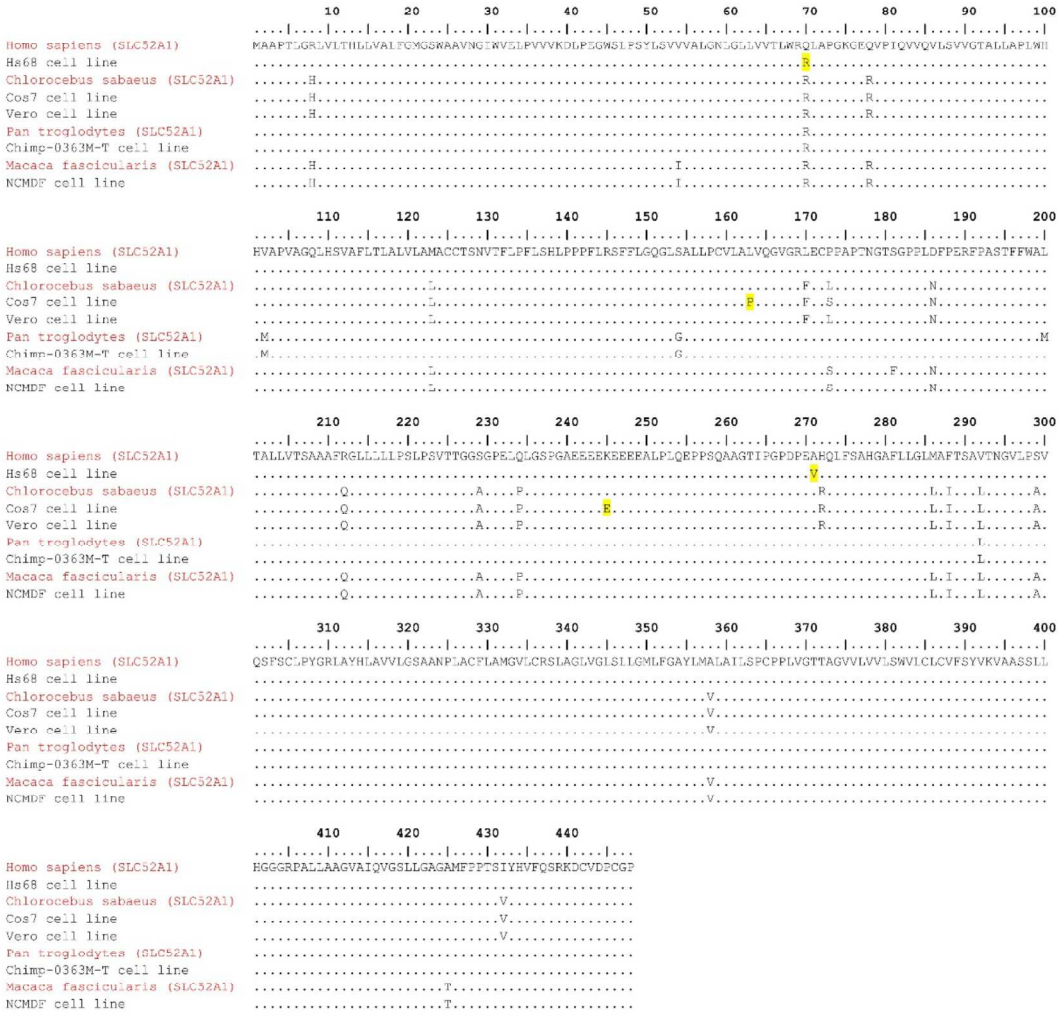


Figure 2.8 Comparison of the SLC52A1 gene sequence in human and primate cell lines. Amino acid positions are indicated above the sequences. Dots indicate residues identical to the reference sequence, while letters represent amino acid substitutions. Hyphens indicate gaps introduced into the alignment. Yellow highlights indicate amino acids that differ from the reference sequence in the corresponding cell lines.

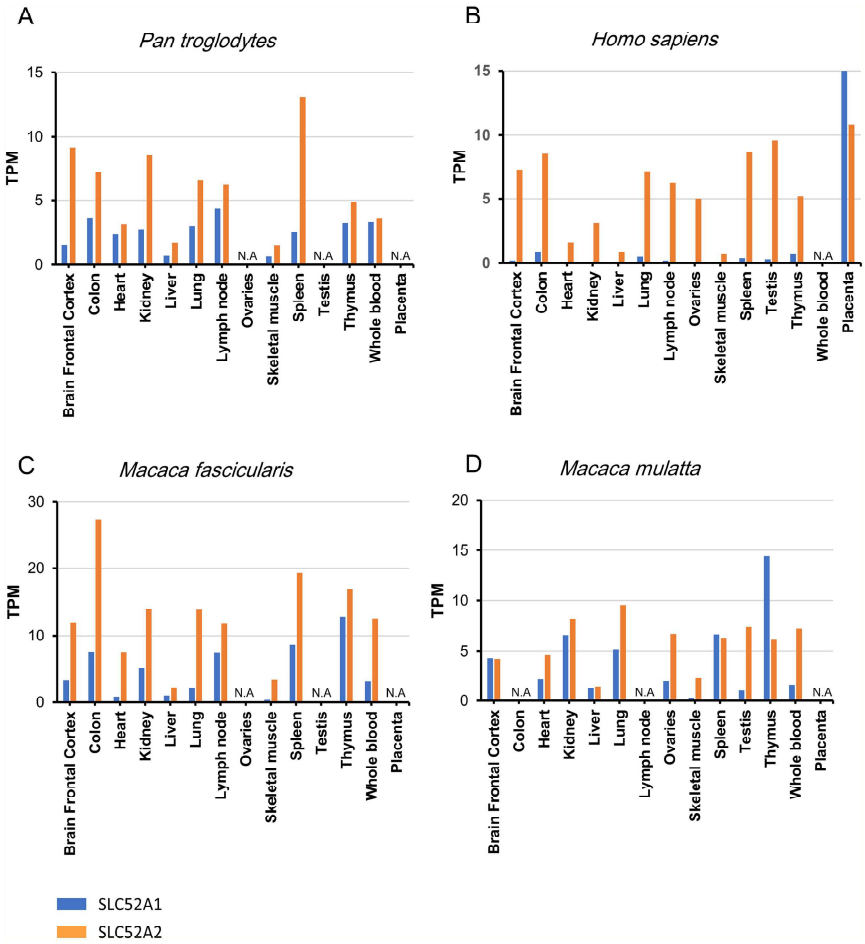


Figure 2.9 Expression of SLC52A1 and SLC52A2 genes in human and non-human primate tissues. (A–D) Bar graphs showing the expression levels of the SLC52A1 and SLC52A2 genes. The x-axis represents the tissues investigated, and the y-axis shows the normalized expression levels of SLC52A1 and SLC52A2 in TPM. N.A. indicates that expression levels in the specified tissue were not investigated due to the unavailability of RNA-seq data.

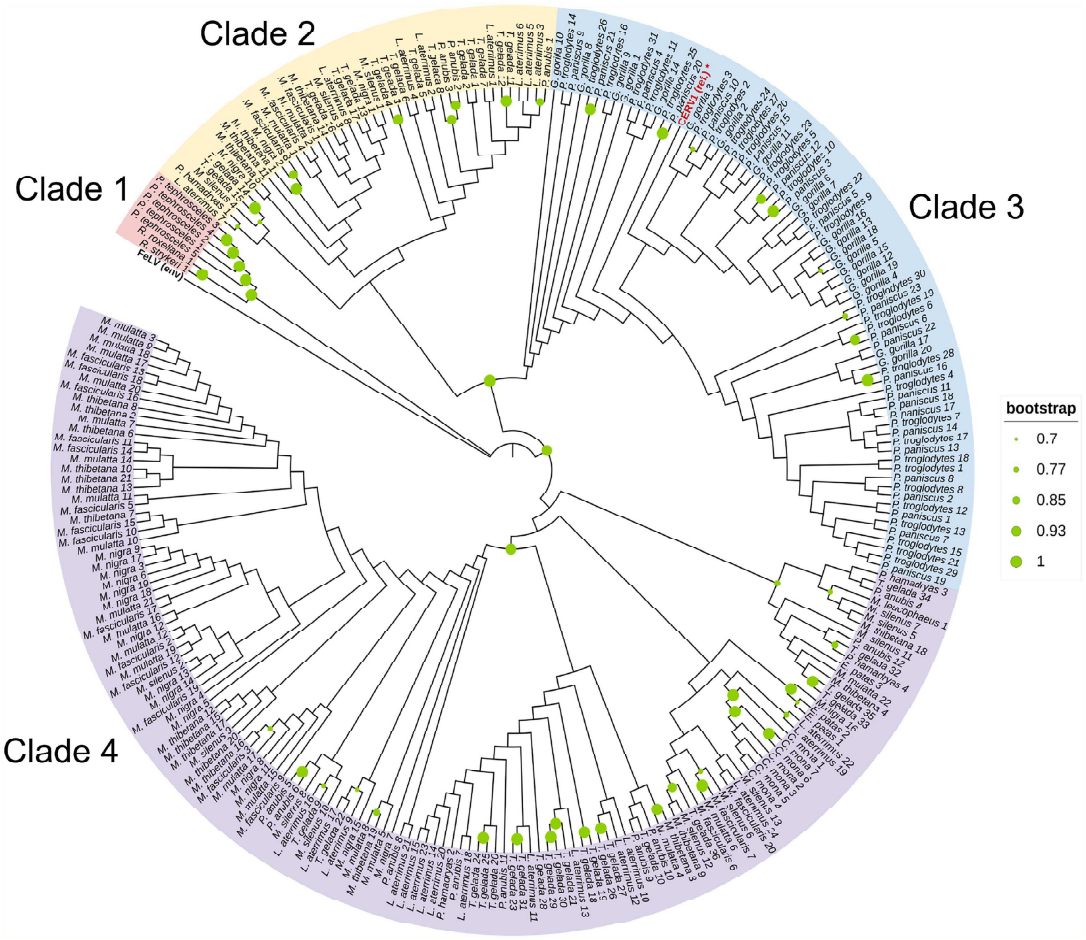


Figure 2.10 Phylogenetic analysis using the CERV1 Env amino acid sequences. The phylogenetic tree for the CERV1 Env amino acid sequences was constructed based on the neighbour-joining method. Bootstrap values >70% are shown adjacent to the respective branches. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. The CERV1 env type is categorized as Clade 1, Clade 2, Clade 3, and Clade 4. The asterisk indicates the CERV1 Env reference sequence used in this study (35).

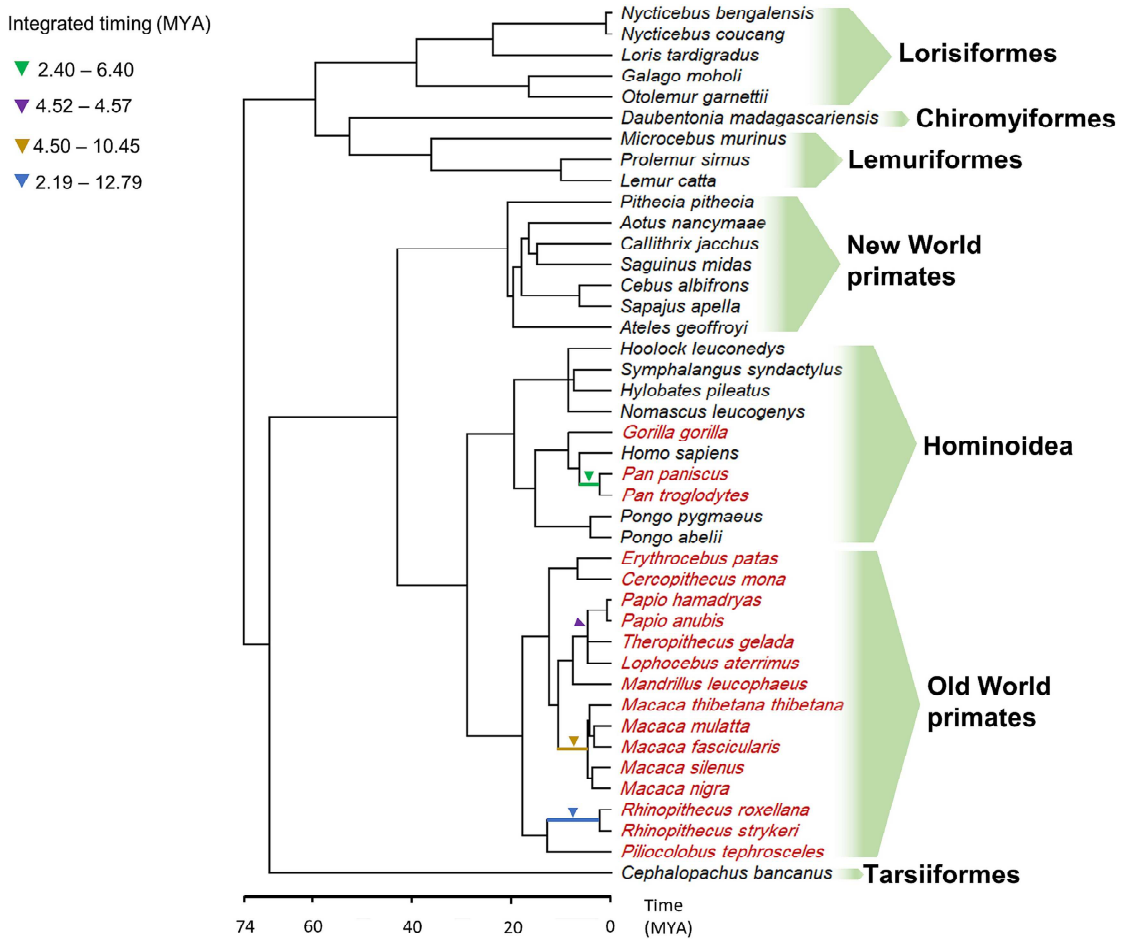


Figure 2.11 Evolutionary history of primate species with the integration timings of CERV1. The arrowheads indicate integration timings of the ancestor with CERV1. The red colour indicates the species infected with CERV1.

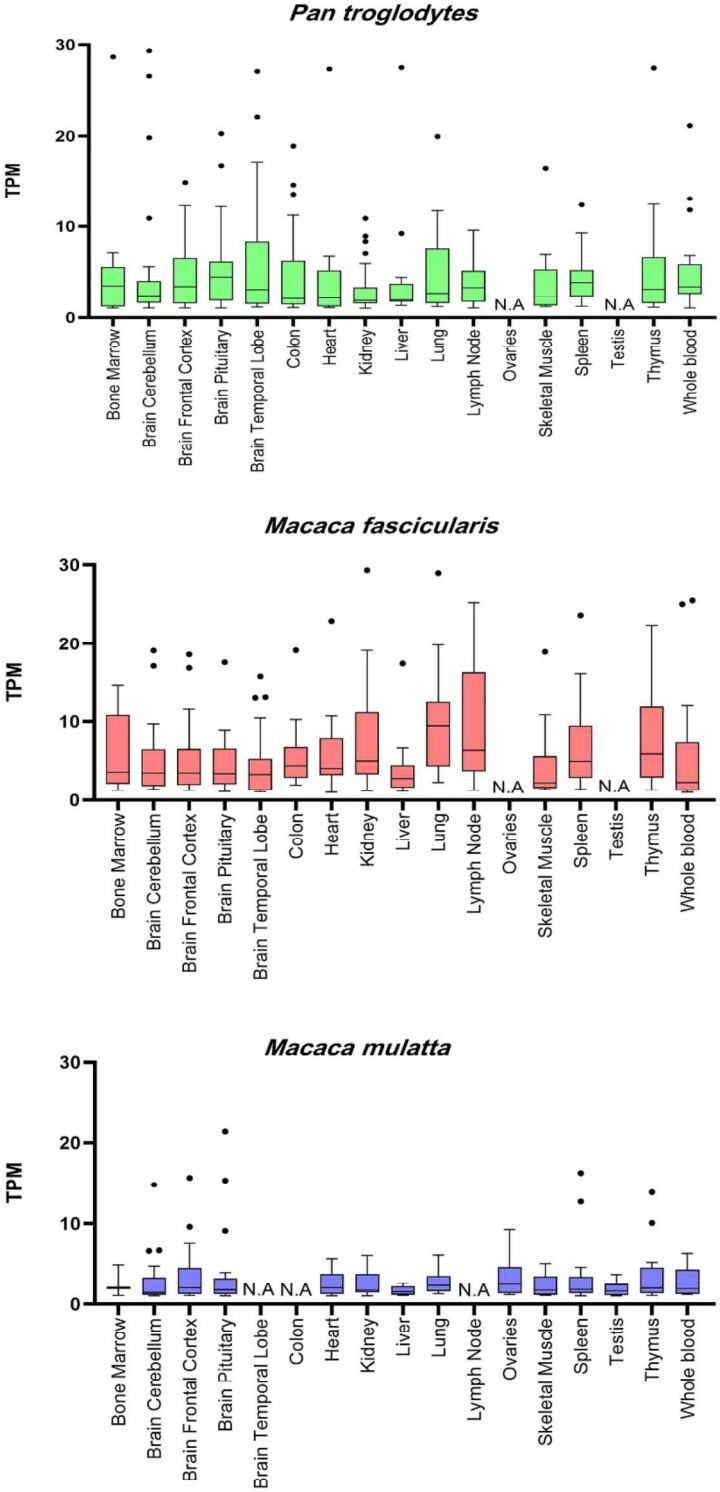


Figure 2.12 Expression of CERV1 *Env* in primate tissues. Box plot showing the CERV1 *Env* gene expression levels, analysed using publicly available RNA-seq data from *P. troglodytes*, *M. fascicularis*, and *M. mulatta*. Box-and-whisker plots depict the median as a line in the centre of a box bounded by the first and third quartiles, with whiskers to 1.5 times the interquartile range and outlier values shown as individual dots.

GENERAL DISCUSSIONS AND CONCLUSIONS

Retroviruses infect a wide variety of vertebrate hosts, including amphibians, reptiles, fish, birds, and mammals (24). Once infecting germline cells, these viruses can permanently integrate into the host genome, leading to the formation of ERVs that are transmitted across generations (114). This dual nature allows retroviruses to act both as active infectious agents and as long-term drivers of genome evolution. The transmission across species of retroviruses may represent a significant risk to animal conservation. The purpose of this study was to investigate these phenomena on two different levels: first, by determining the prevalence and transmission dynamics of FFV in endangered TLC and domestic cats to assess ongoing cross-species viral spillover; and second, by determining the receptor used by CERV1 in primates to understand the mechanisms of ancient viral endogenization.

In Chapter One, I performed the first molecular survey and genetic investigation of FFV in TLCs. I found that the prevalence of FFV in domestic cats and TLCs in Tsushima was 21.96% (domestic cats, 14.07%; TLCs, 7.89%), consistent with the frequencies reported in previous studies (9, 14). My results demonstrated that domestic cats have a higher susceptibility to FFV infection compared to TLCs. Consequently, these findings suggest that FFV may be a cross-species transmission infection, and that transmission may also be possible in other cat species.

I also examined the risk factors correlated with infection. My results showed that male cats had a significantly higher rate of co-infection with FFV and other retroviruses (FIV, FcaGHV1) compared to females. This indicates that horizontal transmission by bites is the main way of infection for males. While single FFV infections show no sex difference, this suggests that females may contract the virus vertically transmission. These findings are critical for conservation, since they highlight that management strategies must concentrate on the prevention of contact between domestic and wild populations to protect the endangered TLC.

Phylogenetic analyses of the SU region showed two genotypes: F17-Type, found in domestic cats, and FUV-Type, affecting both domestic cats and TLCs. The domestic cat isolates cluster with the TLC isolate in a single clade, suggesting directional transmission from domestic reservoirs. This study presents the first report of the FUV-Type genotype on Tsushima Island, where a shared mutation among

species support interspecific spillover. This is apparently due to ecological interactions such as predation, similar to transmission patterns seen in FIV and FcaGHV1. Both genotypes were found from Kamijima and Shimojima, indicating a synergistic correlation in which they are well-adapted to domestic cats and frequently co-circulate. While F17-Type was previously identified in Iriomote cats (70), this is the first confirmation of the FUV-Type in Japan. Therefore, more study into these transmission routes is necessary for the improvement of management strategies.

Chapter Two investigates the molecular evolution of endogenous retroviruses in primates. I demonstrated that CERV1 started endogenization in primate germlines approximately 10 million years ago by using riboflavin transporters (SLC52A1/2) as entrance receptors. This finding categorizes CERV1 within the same viral interference group as PERV, as CERV1 infection was inhibited in cells persistently infected with PERV-A/C. Significantly, while the majority of gammaretroviruses use a single receptor, as shown by GaLV and FeLV-B using Pit1 or FeLV-D using CTR1—CERV1 and PERV are unique in their use of both SLC52A1 and SLC52A2, likely due to the high sequence identity between these transporters.

My research highlights a pattern of convergent evolution, where distinct viruses independently evolved to utilize the same multi-transmembrane transport proteins for cellular entry. Expression analysis found that SLC52A2 is widely expressed in human tissue, such as the testes and ovaries, whereas CERV1 is notably absent from the human genome. This indicates that species-specific receptor incompatibility is not the cause of the absence of CERV1 endogenization in humans. Alternatively, other factors may be responsible, including the timing of viral prevalence, specific expression patterns of SLC52A1 in the placenta compared to other tissue, or evolutionary resistance mechanisms such as APOBEC3 or TRIM5α.

These findings provide significant insights for xenotransplantation. Resistance to PERV in specific primates, such as macaques and baboons, is associated with a specific serine-to-leucine substitution at position 109 in SLC52A2; however, this mutation is not present in humans and chimpanzees. However, because CERV1 and PERV share an interference group, the expression of endogenous CERV1 Env proteins in non-human primates may act as an alternative resistance response to PERV infection. This

possible "molecular shield" effect highlights the functional role that endogenized viral genes could be used in host defense, similar to the physiological significance found for CERV2-related sequences.

In conclusion, I have demonstrated that FFV is prevalent in domestic cats and may be a cross-species transmission infection affecting endangered felids, while simultaneously showing that ancient retroviruses like CERV1 have shaped primate genomes through convergent receptor usage. Together, these studies contribute to a broader understanding of how host-pathogen interactions drive both disease emergence and evolutionary adaptation.

REFERENCES

1. Greenwood AD, Ishida Y, O'Brien SP, Roca AL, Eiden M V. 2018. Transmission, evolution, and endogenization: lessons learned from recent retroviral invasions. Microbiology and molecular biology reviews 82:10–1128.

2. Hess JL, Clements JE, Narayan O. 1985. Cis-and trans-acting transcriptional regulation of visna virus. Science (1979) 229:482–485.

3. Guntaka R V. 1993. Transcription termination and polyadenylation in retroviruses. Microbiol Rev 57:511–521.

4. Varmus H. 1988. Retroviruses. Science (1979) 240:1427–1435.

5. Coffin JM, Hughes SH, Varmus HE. 1997. The interactions of retroviruses and their hosts. In: Coffin JM, Hughes SH, Varmus HE (eds.), Retroviruses, pp. 335-42. New York: Cold Spring Harbour Press.

6. Coffin JM. 1992. Structure and classification of retroviruses, p. 19–49. *In* The retroviridae. Springer.

7. Leis J, Baltimore D, Bishop JM, Coffin J, Fleissner E, Goff SP, Oroszlan S, Robinson H, Skalka AM, Temin HM and. 1988. Standardized and simplified nomenclature for proteins common to all retroviruses. J Virol 62:1808–1809.

8. Nelson PN, Carnegie PR, Martin J, Ejtehadi HD, Hooley P, Roden D, Rowland-Jones S, Warren P, Astley J, Murray PG. 2003. Demystified... Human endogenous retroviruses. Molecular Pathology 56:11.

9. Dannemiller NG, Kechejian S, Kraberger S, Logan K, Alldredge M, Crooks KR, VandeWoude S, Carver S. 2020. Diagnostic uncertainty and the epidemiology of feline foamy virus in pumas (Puma concolor). Sci Rep 10:1587.

10. Kehl T, Tan J, Materniak M. 2013. Non-simian foamy viruses: molecular virology, tropism and prevalence and zoonotic/interspecies transmission. Viruses 5:2169–209.

11. Pinto-Santini DM, Stenbak CR, Linial ML. 2017. Foamy virus zoonotic infections. Retrovirology 14:55.

12. Kechejian SR, Dannemiller N, Kraberger S, Ledesma-Feliciano C, Malmberg J, Roelke Parker M, Cunningham M, McBride R, Riley SPD, Vickers WT. 2019. Feline foamy virus is highly prevalent in free-ranging Puma concolor from Colorado, Florida and southern California. Viruses 11:359.

13. Koc BT, Oğuzoğlu TÇ. 2019. First report on the prevalence and genetic relatedness of Feline Foamy Virus (FFV) from Turkish domestic cats. Virus Res 274:197768.

14. Bleiholder A, Mühle M, Hechler T, Bevins S, Denner J, Löchelt M. 2011. Pattern of seroreactivity against feline foamy virus proteins in domestic cats from Germany. Vet Immunol Immunopathol 143:292–300.

15. Winkler IG, Flügel RM, Löchelt M, Flower RLP. 1998. Detection and molecular characterisation of feline foamy virus serotypes in naturally infected cats. Virology 247:144–51.

16. Phung HTT, Ikeda Y, Miyazawa T, Nakamura K, Mochizuki M, Izumiya Y, Sato E, Nishimura Y, Tohya Y, Takahashi E. 2001. Genetic analyses of feline foamy virus isolates from domestic and wild feline species in geographically distinct areas. Virus Res 76:171–81.

17. Ledesma-Feliciano C, Troyer RM, Zheng X, Miller C, Cianciolo R, Bordicchia M, Dannemiller N, Gagne R, Beatty J, Quimby J. 2019. Feline foamy virus infection: characterization of experimental infection and prevalence of natural infection in domestic cats with and without chronic kidney disease. Viruses 11:662.

18. Romen F, Backes P, Materniak M, Sting R, Vahlenkamp TW, Riebe R, Pawlita M, Kuzmak J, Löchelt M. 2007. Serological detection systems for identification of cows shedding bovine foamy virus via milk. Virology 364:123–31.

19. Kraberger S, Fountain-Jones NM, Gagne RB, Malmberg J, Dannemiller NG, Logan K, Alldredge M, Varsani A, Crooks KR, Craft M. 2020. Frequent cross-species transmissions of foamy virus between domestic and wild felids. Virus Evol 6:vez058.

20. Mouinga-Ondémé A, Caron M, Nkoghé D, Telfer P, Marx P, Saïb A, Leroy E, Gonzalez J-P, Gessain A, Kazanji M. 2012. Cross-species transmission of simian foamy virus to humans in rural Gabon, Central Africa. J Virol 86:1255–60.

21. Belshaw R, Katzourakis A, Paces J, Burt A, Tristem M. 2005. High copy number in human endogenous retrovirus families is associated with copying mechanisms in addition to reinfection. Mol Biol Evol 22:814–7.

22. Garcia-Etxebarria K, Jugo BM. 2010. Genome-wide detection and characterization of endogenous retroviruses in Bos taurus. J Virol 84:10852–10862.

23. Gifford R, Tristem M. 2003. The evolution, distribution and diversity of endogenous retroviruses. Virus Genes 26:291–315.

24. Johnson WE. 2019. Origins and evolutionary consequences of ancient endogenous retroviruses. Nat Rev Microbiol 17:355–70.

25. Pramono D, Takeuchi D, Katsuki M, AbuEed L, Abdillah D, Kimura T, Kawasaki J, Miyake A, Nishigaki K. 2024. FeLIX is a restriction factor for mammalian retrovirus infection. J Virol 98:e01771-23.

26. Frank JA, Singh M, Cullen HB, Kirou RA, Benkaddour-Boumzaouad M, Cortes JL, Garcia Pérez J, Coyne CB, Feschotte C. 2022. Evolution and antiviral activity of a human protein of retroviral origin. Science 378:422–8.

27. Blanco-Melo D, Gifford RJ, Bieniasz PD. 2017. Co-option of an endogenous retrovirus envelope for host defense in hominid ancestors. Elife 6:e22519.

28. Ito J, Watanabe S, Hiratsuka T, Kuse K, Odahara Y, Ochi H, Kawamura M, Nishigaki K. 2013. Refrex-1, a soluble restriction factor against feline endogenous and exogenous retroviruses. J Virol 87:12029–40.

29. Ikeda H, Laigret F, Martin MA, Repaske R. 1985. Characterization of a molecularly cloned retroviral sequence associated with Fv-4 resistance. J Virol 55:768–77.

30. Robinson HL, Astrin SM, Senior AM, Salazar FH. 1981. Host susceptibility to endogenous viruses: defective, glycoprotein-expressing proviruses interfere with infections. J Virol 40:745–51.

31. Emerman M, Malik HS. 2010. Paleovirology—modern consequences of ancient viruses. PLoS Biol 8:e1000301.

32. Yohn CT, Jiang Z, McGrath SD, Hayden KE, Khaitovich P, Johnson ME, Eichler MY, McPherson JD, Zhao S, Pääbo S. 2005. Lineage-specific expansions of retroviral insertions within the genomes of African great apes but not humans and orangutans. PLoS Biol 3:e110.

33. Polavarapu N, Bowen NJ, McDonald JF. 2006. Identification, characterization and comparative genomics of chimpanzee endogenous retroviruses. Genome Biol 7:R51.

34. Miyake A, Ngo MH, Wulandari S, Shimojima M, Nakagawa S, Kawasaki J, Nishigaki K. 2022. Convergent evolution of antiviral machinery derived from endogenous retrovirus truncated envelope genes in multiple species. Proc Natl Acad Sci 119:e2114441119.

35. Soll SJ, Neil SJD, Bieniasz PD. 2010. Identification of a receptor for an extinct virus. Proc Natl Acad Sci USA 107:19496–501.

36. Makundi I, Koshida Y, Endo Y, Nishigaki K. 2018. Identification of felis catus gammaherpesvirus 1 in Tsushima leopard cats (Prionailurus bengalensis euptilurus) on Tsushima Island, Japan. Viruses 10:378.

37. Saitoh T, Kaji K, Izawa M, Yamada F. 2015. Conservation and management of terrestrial mammals in Japan: its organizational system and practices. Therya 6:139–53.

38. Makundi I, Koshida Y, Kuse K, Hiratsuka T, Ito J, Baba T, Watanabe S, Kawamura M, Odahara Y, Miyake A, Yamamoto H, Kuniyoshi S, Onuma M, Nishigaki K. 2017. Epidemiologic survey of feline leukemia virus in domestic cats on Tsushima Island, Japan: Management strategy for Tsushima leopard cats. J. Vet. Diagn. Investig. 29:889–95.

39. Tateno M, Nishio T, Matsuo T, Sakuma M, Nakanishi N, Izawa M, Asari Y, Okamura M, Miyama TS, Setoguchi A. 2013. Epidemiological survey of tick-borne protozoal infection in iriomote cats and tsushima leopard cats in Japan. J. Vet. Med. Sci. 75:13–15.

40. Mitani N, Mihara S, Ishii N, Koike H. 2009. Clues to the cause of the Tsushima leopard cat (*Prionailurus bengalensis euptilura*) decline from isotopic measurements in three species of Carnivora. Ecol Res 24:897–908.

41. Hayama S, Yamamoto H, Nakanishi S, Hiyama T, Murayama A, Mori H, Sugitani A, Fujiwara S. 2010. Risk analysis of feline immunodeficiency virus infection in Tsushima leopard cats (*Prionailurus bengalensis euptilurus*) and domestic cats using a geographic information system. J. Vet. Med. Sci. 72:1003300203.

42. Nishimura Y, Goto Y, Yoneda K, Endo Y, Mizuno T, Hamachi M, Maruyama H, Kinoshita H, Koga S, Komori M. 1999. Interspecies transmission of feline immunodeficiency virus from the domestic cat to the Tsushima cat (*Felis bengalensis euptilura*) in the wild. J Virol 73:7916–21.

43. Tateno M, Nishio T, Sakuma M, Nakanishi N, Izawa M, Asari Y, Okamura M, Maruyama S, Miyama TS, Setoguchi A. 2013. Molecular epidemiologic survey of *Bartonella, Ehrlichia*, and *Anaplasma* infections in Japanese Iriomote and Tsushima leopard cats. J Wildl Dis 49:646–52.

44. Sumiyoshi A, Kitao K, Miyazawa T. 2022. Genetic and biological characterization of feline foamy virus isolated from a leopard cat (*Prionailurus bengalensis*) in Vietnam. J Vet Med Sci 84:157–65.

45. Alke A, Schwantes A, Zemba M, Flügel RM, Löchelt M. 2000. Characterization of the humoral immune response and virus replication in cats experimentally infected with feline foamy virus. Virology 275:170–76.

46. Court EA, Watson ADJ, Peaston AE. 1997. Retrospective study of 60 cases of feline lymphosarcoma. Aust Vet J 75:424–27.

47. Cavalcante LTF, Muniz CP, Jia H, Augusto AM, Troccoli F, Medeiros S de O, Dias CGA, Switzer WM, Soares MA, Santos AF. 2018. Clinical and molecular features of feline foamy virus and feline leukemia virus co-infection in naturally-infected cats. Viruses 10:702.

48. Switzer WM, Garcia AD, Yang C, Wright A, Kalish ML, Folks TM, Heneine W. 2008. Coinfection with HIV-1 and simian foamy virus in West Central Africans. J Infect Dis 197:1389–93.

49. Alais S, Pasquier A, Jegado B, Journo C, Rua R, Gessain A, Tobaly-Tapiero J, Lacoste R, Turpin J, Mahieux R. 2018. STLV-1 co-infection is correlated with an increased SFV proviral load in the peripheral blood of SFV/STLV-1 naturally infected non-human primates. PLoS Negl Trop Dis 12:e0006812.

50. Mekata H, Okagawa T, Konnai S, Miyazawa T. 2021. Molecular Epidemiology and Whole-Genome Analysis of Bovine Foamy Virus in Japan. Viruses 13:1017.

51. Hooks JJ, Gibbs CJ, Jr. 1975. The foamy viruses. Bacteriol Rev 39:169–85.

52. Khan AS, Bodem J, Buseyne F, Gessain A, Johnson W, Kuhn JH, Kuzmak J, Lindemann D, Linial ML, Löchelt M. 2018. Spumaretroviruses: Updated taxonomy and nomenclature. Virology 516:158–64.

53. Tobaly-Tapiero J, Bittoun P, Neves M, Guillemin M-C, Lecellier C-H, Puvion-Dutilleul F, Gicquel B, Zientara S, Giron M-L, de Thé H. 2000. Isolation and characterization of an equine foamy virus. J Virol 74:4064–73.

54. Pinto-Santini DM, Stenbak CR, Linial ML. 2017. Foamy virus zoonotic infections. Retrovirology 14:1-14.

55. Roy J, Rudolph W, Juretzek T, Gärtner K, Bock M, Herchenröder O, Lindemann D, Heinkelein M, Rethwilm A. 2003. Feline foamy virus genome and replication strategy. J Virol 77:11324–331.

56. Buseyne F, Betsem E, Montange T, Njouom R, Bilounga Ndongo C, Hermine O, Gessain A. 2018. Clinical signs and blood test results among humans infected with zoonotic simian foamy virus: a case-control study. J Infect Dis 218:144–51.

57. Kechejian S, Dannemiller N, Kraberger S, Ledesma Feliciano C, Löchelt M, Carver S, VandeWoude S. 2019. Feline foamy virus seroprevalence and demographic risk factors in stray domestic cat populations in Colorado, Southern California and Florida, USA. J Feline Med Surg Open Rep 5:2055116919873736.

58. Beatty JA, Troyer RM, Carver S, Barrs VR, Espinasse F, Conradi O, Stutzman-Rodriguez K, Chan CC, Tasker S, Lappin MR. 2014. Felis catus gammaherpesvirus 1; a widely endemic potential pathogen of domestic cats. Virology 460:100–7.

59. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–29.

60. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772.

61. Winkler IG, Lochelt M, Flower RLP. 1999. Epidemiology of feline foamy virus and feline immunodeficiency virus infections in domestic and feral cats: a seroepidemiological study. J Clin Microbiol 37:2848–51.

62. Miranda LHM de, Meli M, Conceição-Silva F, Novacco M, Menezes RC, Pereira SA, Sugiarto S, Reis ÉG dos, Gremião IDF, Hofmann-Lehmann R. 2018. Co-infection with feline retrovirus is related to changes in immunological parameters of cats with sporotrichosis. PLoS One 13:e0207644.

63. McLuckie A, Tasker S, Dhand NK, Spencer S, Beatty JA. 2016. High prevalence of Felis catus gammaherpesvirus 1 infection in haemoplasma-infected cats supports co-transmission. Vet J 214:117–21.

64. Ertl R, Korb M, Langbein-Detsch I, Klein D. 2015. Prevalence and risk factors of gammaherpesvirus infection in domestic cats in Central Europe. Virol J 12:146.

65. Stutzman-Rodriguez K, Rovnak J, VandeWoude S, Troyer RM. 2016. Domestic cats seropositive for Felis catus gammaherpesvirus 1 are often qPCR negative. Virology 498:23–30.

66. Chiu ES, Kraberger S, Cunningham M, Cusack L, Roelke M, VandeWoude S. 2019. Multiple introductions of domestic cat feline leukemia virus in endangered Florida panthers. Emerg Infect Dis 25:92.

67. Moss WE, Alldredge MW, Logan KA, Pauli JN. 2016. Human expansion precipitates niche expansion for an opportunistic apex predator (*Puma concolor*). Sci Rep 6:39639.

68. Lambert C, Couteaudier M, Gouzil J, Richard L, Montange T, Betsem E, Rua R, Tobaly-Tapiero J, Lindemann D, Njouom R. 2018. Potent neutralizing antibodies in humans infected with zoonotic simian foamy viruses target conserved epitopes located in the dimorphic domain of the surface envelope protein. PLoS Pathog 14:e1007293.

69. Richard L, Rua R, Betsem E, Mouinga-Ondémé A, Kazanji M, Leroy E, Njouom R, Buseyne F, Afonso P V, Gessain A. 2015. Cocirculation of two env molecular variants, of possible

recombinant origin, in gorilla and chimpanzee simian foamy virus strains from Central Africa. J Virol 89:12480–91.

70. Mochizuki M, Akuzawa M, Nagatomo H. 1990. Serological survey of the Iriomote cat (*Felis iriomotensis*) in Japan. J Wildl Dis 26:236–45.

71. Kassiotis G. 2023. The immunological conundrum of endogenous retroelements. Annu Rev Immunol 41:99–125.

72. Kitajima R, Nakai R, Imamura T, Kameda T, Kozuka D, Hirai H, Ito H, Imai H, Imamura M. 2020. Modeling of early neural development in vitro by direct neurosphere formation culture of chimpanzee induced pluripotent stem cells. Stem Cell Res 44:101749.

73. Denner J, Specke V, Thiesen U, Karlas A, Kurth R. 2003. Genetic alterations of the long terminal repeat of an ecotropic porcine endogenous retrovirus during passage in human cells. Virology 314:125–33.

74. Miyake A, Watanabe S, Hiratsuka T, Ito J, Ngo MH, Makundi I, Kawasaki J, Endo Y, Tsujimoto H, Nishigaki K. 2016. Novel feline leukemia virus interference group based on the env gene. J Virol 90:4832–7.

75. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–7.

76. Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–27.

77. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–25.

78. Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. Oxford university press.

79. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–82.

80. Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 52:W78–82.

81. Aiewsakun P, Katzourakis A. 2015. Endogenous viruses: Connecting recent and ancient viral evolution. Virology 479-480:26–37.

82. Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol 34:1812–9.

83. Leinonen R, Sugawara H, Shumway M, Collaboration INSD. 2010. The sequence read archive. Nucleic Acids Res 39:D19–21.

84. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–90.

85. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

86. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–30.

87. Ericsson TA, Takeuchi Y, Templin C, Quinn G, Farhadian SF, Wood JC, Oldmixon BA, Suling KM, Ishii JK, Kitagawa Y. 2003. Identification of receptors for pig endogenous retrovirus. Proc Natl Acad Sci USA 100:6759–64.

88. Jin H, Zhang C, Zwahlen M, Von Feilitzen K, Karlsson M, Shi M, Yuan M, Song X, Li X, Yang H. 2023. Systematic transcriptional analysis of human cell lines for gene expression landscape and tumor representation. Nat Commun 14:5417.

89. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A. 2015. Tissue-based map of the human proteome. Science 347:1260419.

90. Mazari PM, Linder-Basso D, Sarangi A, Chang Y, Roth MJ. 2009. Single-round selection yields a unique retroviral envelope utilizing GPR172A as its host receptor. Proc Natl Acad Sci 106:5848–53.

91. Marcucci KT, Argaw T, Wilson CA, Salomon DR. 2009. Identification of two distinct structural regions in a human porcine endogenous retrovirus receptor, HuPAR2, contributing to function for viral entry. Retrovirology 6:6-3.

92. Yonezawa A, Masuda S, Katsura T, Inui K. 2008. Identification and functional characterization of a novel human and rat riboflavin transporter, RFT1. Am J Physiol Cell Physiol 295:C632–41.

93. Takeuchi Y, Patience C, Magre S, Weiss RA, Banerjee PT, Le Tissier P, Stoye JP. 1998. Host range and interference studies of three classes of pig endogenous retrovirus. J Virol 72:9986–91.

94. Harrison I, Takeuchi Y, Bartosch B, Stoye JP. 2004. Determinants of high titer in recombinant porcine endogenous retroviruses. J Virol 78:13871–9.

95. Mattiuzzo G, Takeuchi Y. 2010. Suboptimal porcine endogenous retrovirus infection in non-human primate cells: implication for preclinical xenotransplantation. PLoS One 5:e13203.

96. Mattiuzzo G, Matouskova M, Takeuchi Y. 2007. Differential resistance to cell entry by porcine endogenous retrovirus subgroup A in rodent species. Retrovirology 4:1-13.

97. Albritton LM, Tseng L, Scadden D, Cunningham JM. 1989. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. Cell 57:659–666.

98. Kozak CA. 2014. Origins of the endogenous and infectious laboratory mouse gammaretroviruses. Viruses 7:1–26.

99. Chiu ES, Hoover EA, VandeWoude S. 2018. A retrospective examination of feline leukemia subgroup characterization: viral interference assays to deep sequencing. Viruses 10:29.

100. Yonezawa A, Inui K. 2013. Novel riboflavin transporter family RFVT/SLC52: identification, nomenclature, functional characterization and genetic diseases of RFVT/SLC52. Mol Asp Med 34:693–701.

101. Johann S V, Gibbons JJ, O'hara B. 1992. GLVR1, a receptor for gibbon ape leukemia virus, is homologous to a phosphate permease of *Neurospora crassa* and is expressed at high levels in the brain and thymus. J Virol 66:1635–40.

102. Takeuchi Y, Vile RG, Simpson GUY, O'Hara B, Collins MK, Weiss RA. 1992. Feline leukemia virus subgroup B uses the same cell surface receptor as gibbon ape leukemia virus. J Virol 66:1219–22.

103. Miller AD, Wolgamot G. 1997. Murine retroviruses use at least six different receptors for entry into Mus dunni cells. J Virol 71:4531–5.

104. Anderson MM, Lauring AS, Robertson S, Dirks C, Overbaugh J. 2001. Feline Pit2 functions as a receptor for subgroup B feline leukemia viruses. J Virol 75:10563–72.

105. Oliveira NM, Farrell KB, Eiden M V. 2006. In vitro characterization of a koala retrovirus. J Virol 80:3104–7.

106. Tury S, Giovannini D, Ivanova S, Touhami J, Courgnaud V, Battini J-L. 2022. Identification of copper transporter 1 as a receptor for feline endogenous retrovirus ERV-DC14. J Virol 96:e0022922.

107. Tsang J, Ribet D, Heidmann T, Dewannieux M. 2019. Identification of the receptor used by the ecotropic mouse GLN endogenous retrovirus. J Virol 93:e01125-18.

108. Miyake A, Kawasaki J, Ngo H, Makundi I, Muto Y, Khan AH, Smith DJ, Nishigaki K. 2019. Reduced folate carrier: an entry receptor for a novel feline leukemia virus variant. J Virol 93:e00269-19.

109. Shalev Z, Duffy SP, Adema KW, Prasad R, Hussain N, Willett BJ, Tailor CS. 2009. Identification of a feline leukemia virus variant that can use THTR1, FLVCR1, and FLVCR2 for infection. J Virol 83:6706–16.

110. Kaiser SM, Malik HS, Emerman M. 2007. Restriction of an extinct retrovirus by the human TRIM5α antiviral protein. Science 316:1756–8.

111. Perez-Caballero D, Soll SJ, Bieniasz PD. 2008. Evidence for restriction of ancient primate gammaretroviruses by APOBEC3 but not TRIM5α proteins. PLoS Pathog 4:e1000181.

112. Specke V, Schuurman H-J, Plesker R, Coulibaly C, Özel M, Langford G, Kurth R, Denner J. 2002. Virus safety in xenotransplantation: first exploratory in vivo studies in small laboratory animals and non-human primates. Transpl Immunol 9:281–8.

113. Specke V, Plesker R, Wood J, Coulibaly C, Suling K, Patience C, Kurth R, Schuurman H, Denner J. 2009. No in vivo infection of triple immunosuppressed non-human primates after inoculation with high titers of porcine endogenous retroviruses. Xenotransplantation 16:34–44.

114. Kawasaki J, Nishigaki K. 2018. Tracking the continuous evolutionary processes of an endogenous retrovirus of the domestic cat: ERV-DC. Viruses 10: 179.