Evolutionary dynamics of endogenous retroviruses (ERV-DC) in *Felis* lineage

ネコ科動物における内在性レトロウイルス（ERV-DC）の進化ダイナミクス

A DISSERTATION

Submitted by

NGO MINH HA

in Partial Fulfillment of the Requirement for the Degree of

DOCTOR OF PHILOSOPHY

(Virology, Epidemiology and Molecular Biology)

The United Graduate School of Veterinary Science

YAMAGUCHI UNIVERSITY

JAPAN

March, 2020

DECLARATION

I hereby declare that the work reported in this research project report has been carried out by the undersigned. I also declare that where reference has been made to the results of other workers, appropriate acknowledgment of the source of information has been made.

March, 2020 Yamaguchi University Author

NGO MINH HA

PREFACE

This dissertation discusses evolutionary dynamics of endogenous retroviruses in *Felis* lineage with reference to ERV-DC known as one of the youngest groups of endogenous retroviruses in domestic cat. Invasion of infectious endogenous retroviruses in mixed-breed and purebred cats is firstly discussed here. Furthermore, I investigated whether or not existence of ERV-DC in wildcats and their fate after invading different subspecies.

The first study on epidemiological survey of invasion of infectious endogenous retroviruses with reference to ERV-DC10 and ERV-DC14 in different breeding cats was conducted by polymerase chain reaction and viral infection assay in vitro. This study described prevalence of these two infectious proviruses not only in different breeding cat but also in different locations where cats live. This study suggested a possible direction of cat migration after domestication. This study also provided useful information for further study on the potentially associated disease induction in cats by these two viruses.

The second study was about tracking the fate of endogenous retrovirus segregation in wild and domestic cats. This is the first report found that a naturally occurring single mutation could inactivate an infectious retrovirus. A common mechanism found would be helpful for the strategy to produce antiviral therapies against viral infections in the future.

The third study found that existence of a new ERV-like ERV-DC in jungle cat. Jungle cat is known as the farthest wildcat specie in *Felis* lineage and is now threaten to be extinct in some countries. This study suggested that there potentially exists an infectious endogenous retrovirus like ERV-DC10 in jungle cat results in a potential disease induction in this wildcat specie.

This dissertation is organized in five main sections. The first section is general introduction whereby I described fundamental understanding about retroviruses and expressed rationale of

carrying out this research. The three following sections are corresponding with three studies respectively. The last section is general discussions and conclusions whereby the main research findings are discussed and concluded as well as restating significances of this research.

I believed this research would be interesting for all readers who are interested in virology, epidemiology, molecular biology, genetics as well as animal evolution.

ACKNOWLEDGMENT

This research was conducted in the Laboratory of Molecular Immunology and Infectious Diseases, Joint Faculty of Veterinary Medicine, Yamaguchi University, Japan since I was a research student until nearly finishing my PhD course between April 2015 and March 2020.

First of all, I would like to express my deep gratitude to my main supervisor Prof. Dr. Kazuo Nishigaki who wholeheartedly guided, advised and inspired me throughout my PhD student life. Initially, he gave me an opportunity to be his student even I did not own any good background. Furthermore, I am sure that I could not learn many knowledge and techniques regarding to virology, molecular biology, infectious diseases, bioinformatics as well as reviewing and writing a scientific paper without his supervision. After five-year in this laboratory, I am now confident with advancing some molecular techniques although I will never forget to enrich my knowledge more and more. Once again, I would like to say “Thank you very much” for your support during my study life in Japan.

I would like to express my sincere thanks to Assistant Prof. Dr. Ariko Miyake who patiently and tirelessly tutored me during my first time in the laboratory. Without her support, probably I cannot carry out the experiment as quickly and accurately as I can at the moment. I have leant a lot of knowledge from her not only techniques but also the way to organize experiments. Moreover, she is so kind to me and she is always willing to help me whenever I face with Japanese language difficulties. Japanese language is still a big trouble for me even I stayed here for 5 years so she always sympathizes and supports me as a professional translator.

I am deeply indebted to Prof. Dr. Yoshimi Yamamoto (Professor Emeritus of Yamaguchi University) who firstly introduced and advertised me to my main supervisor. Secondly, he always gives me thoughtful advice on not only my study life but also my real life in Japan.

I am deeply grateful to Assoc. Prof. Tetsuya Yanagida who carefully proofread my doctoral thesis and all my defending committee members who gave me valuable advice to improve my final PhD thesis.

I would like to acknowledge all present and past members in the Nishigaki sensei's laboratory, Yamaguchi University for their kindness, help and technical support throughout my study. Especially Dr. Junna Kawasaki, Dr. Isaac Makundi (PhD), Dr. Taiji Endo, and Dr. Ryosuke Sumi.

I am deeply thankful to the Ministry of Education, Culture, Sports, Science and Technology (Monbukagakusho) of Japan for financial support and giving me an opportunity to study and research in Yamaguchi University during my PhD's life.

I am deeply grateful to Hirakawa-no-Kazeno Kai members especially Ryoco Saito (Aunty), international student support office of Yamaguchi University especially Nohara-san for social life support.

I also would like to express my deep appreciation to my parents, my parents-in-law, all of my relatives and friends for their kindness and contributions.

Last but not least, I would like to express the deepest gratitude to my wife, my daughter and my son who tirelessly supported me and encouraged me to overcome difficulties to obtain the highest achievements throughout my study journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Retroviruses belonging to Retroviridae family are classified into exogenous retroviruses (exRVs) and endogenous retroviruses (ERVs). ERVs are known as DNA sequences comprising approximately 6–10% of cat, human, and mouse genome. ERVs are descendants of exRVs that integrated into the germ line of the ancestral host lineage and are now transmitted vertically from parents to offspring in a Mendelian inheritance pattern. ERVs are ubiquitously present in all vertebrates. Most ERVs are inactivated but some still retain replication capacity to produce infectious viruses in mice and cats. Interestingly, although most ERVs seem to represent junk DNA, a few ERVs have been co-opted by their hosts to gain physiological functions through "ERV domestication" such as antiviral factors, placenta formation ability, innate immunity regulation, and so on. Thus, ERV domestication has contributed dramatically to host evolution.

Feline endogenous retroviruses have been identified and grouped phylogenetically into different classes. Endogenous retroviruses of domestic cats (ERV-DCs) are one of the youngest feline ERV groups in domestic cats (*Felis silvestris catus*); some members are replication-competent (ERV-DC10, ERV-DC18, and ERV-DC14), produce the anti-retroviral soluble factor Refrex-1 (ERV-DC7 and ERV-DC16), or can generate recombinant feline leukemia virus (FeLV). However, the evolution of ERV-DC in *Felis* linage remains unclear.

In chapter one, I assessed the invasion by two distinct infectious ERV-DCs, ERV-DC10 and ERV-DC14, in domestic cats. Of a total sample of 1646 cats, 568 animals (34.5%) were positive for ERV-DC10 (heterozygous: 377; homozygous: 191), 68 animals (4.1%) were positive for ERV-DC14 (heterozygous: 67; homozygous: 1), and 10 animals (0.6%) were positive for both ERV-DC10 and ERV-DC14. ERV-DC10 and ERV-DC14 were detected in

domestic cats in Japan as well as in Tanzania, Sri Lanka, Vietnam, South Korea and Spain. Breeding cats including Singapura, Norwegian Forest and Ragdoll cats showed high frequencies of ERV-DC10 (60–100%). By contrast, ERV-DC14 was detected at low frequency in breeding cats. These results suggest that ERV-DC10 is widely distributed while ERV-DC14 is maintained in a minor population of cats. Thus, ERV-DC10 and ERV-DC14 have invaded cat populations independently.

In chapter two, I investigated ERV-DC in European wildcats (*Felis silvestris silvestris*) and detected four loci: ERV-DC6, ERV-DC7, ERV-DC14, and ERV-DC16. ERV-DC14 was detected at a high frequency in European wildcats; however, it was replication-defective due to single G→A nucleotide substitution resulting in an E148K substitution in ERV-DC14 envelope (Env). This mutation results in a cleavage-defective Env that is not incorporated into viral particles. Introduction of the same mutation into feline and murine infectious gammaretroviruses resulted in similar Env dysfunction. Interestingly, the same mutation was found in a FeLV from naturally-occurring thymic lymphoma and a mouse ERV, suggesting a common mechanism of virus inactivation. Refrex-1 was present in European wildcats; however, ERV-DC16, but not ERV-DC7, was unfixed in European wildcats. Thus, Refrex-1 has had an antiviral role throughout *Felis* evolution, pre-dating cat exposure to feline retroviruses. ERV-DC sequence diversity was present across wild and domestic cats but was locus-dependent. To sum up, ERVs have evolved species-specific phenotypes through the interplay between ERVs and their hosts. The mechanism of viral inactivation may be similar irrespective of the evolutionary history of retroviruses. The tracking of ancestral retroviruses can shed light on their roles in pathogenesis and host-virus evolution.

In chapter three, I conducted analysis of ERV-DC in jungle cats (*Felis chaus*) showing evolutionary lineage. Based on phylogenetic analysis, I found existence of ERV-DC/*F. chaus*

(endogenous retrovirus in jungle cat) like ERV-DC genotype III compared to three genotypes identified in domestic cats previously. The integration time of ERV-DC/*F. chaus* was estimated to be approximately 160,000 years ago. ERV-DC/*F. chaus* integration polymorphism was not fixed in jungle cat. The sites of integration were different between *Felis chaus* and domestic cats. ERV-DC/*F. chaus* integrated into *Felis chaus* have been inactivated due to gene mutations and deletions, and have lost autonomous growth and infectivity. However, existence of intact ERV-DC/*F. chaus env* was found. Using pseudotyped viruses, this ERV-DC/*F. chaus env* still retained infectious capacity suggesting that existence of infectious ERV-DC/*F. chaus* proviruses in jungle cats. These results suggested that ERV-DC and ERV-DC/*F. chaus* internalized in the *F.s.catus* and *F. chaus* have undergone evolutionary route independently in the process of co-evolution with each host. Determination and characterization of ERV-DC/*F. chaus* would be helpful for understanding virus evolution and host-virus interaction.

In conclusion, my studies firstly reported the invasion of infectious endogenous retroviruses in domestic cats, the inactivation of an infectious endogenous retrovirus via a single nucleotide polymorphism, and the potential existence of infectious endogenous proviruses in jungle cats. This information provided new insights into evolution of endogenous retroviruses in *Felis* lineage based on ERV-DCs.

GENERAL INTRODUCTION

Retroviridae family consists of two subfamilies including *Orthoretrovirinae* and *Spumaretrovirinae* based on the International Committee on Taxonomy of Viruses (1). The retrovirus subfamily *Orthoretrovirinae* included six genera which are *Alpharetroviruses, Betaretroviruses, Deltaretroviruses, Epsilonretroviruses, Gammaretroviruses and Lentiviruses* (1). This retrovirus subfamily (*orthoretrovirinae*) is distinct from the other subfamily due to an occurrence of reverse transcription within the viral particle (1). Retroviruses are classified as exogenous retroviruses (exRVs) and endogenous retroviruses (ERVs) based on their modes of transmission (2). Retroviruses exclusively infect vertebrates and invertebrates, and cause various pathologies such as lymphoma, myelodysplastic syndrome, anaplastic anemia, acute myelogenous leukemia, immune deficiency and cancer (3-9). RNA genome of retroviruses typically consists of 5'-capped, 3'-polyadenylated, *gag* genes coding for capsid protein, *pol* genes coding for enzymes, and *env* genes coding for envelope glycoproteins (10). A typical retrovirus virion is roughly spherical, approximately 100nm in diameter, icosahedral or conical capsid, packaging two identical copies of positive strand RNA and viral enzymes (reverses transcriptase, integrase and protease) in an enveloped particle (10). After reverse transcription occurs during infection, the retroviruses integrated into the host genome as proviruses which are along with two flanking noncoding long terminal repeats (LTRs) containing a transcriptional start site and various regulatory cis elements that determine the tropism of viral transcription (11-13). Retroviral envelope (Env) proteins are composed of a trimer of heterodimers formed between the surface subunit (SU) and the transmembrane subunit (TM). The SUs of gammaretroviruses are composed of two globular domains, the N-terminal and C-terminal domains, and mediate viral attachment to target cells through viral

receptor recognition and binding (14, 15). The TMs tether Env to membranes, and their fusion peptides mediate viral entry through fusion between the viral envelope and the cell membrane. Env is first synthesized as a SU–TM precursor polypeptide in the rough endoplasmic reticulum and then transported into the trans-Golgi network, where it is cleaved into the SU and TM at the cleavage motif (R-X-K/R-R2Y) by cellular proteases (16-21). Mutations in either the SU or TM region of retroviruses result in virus inactivation through the production of a cleavage-defective Env protein (17, 22-29). Thus, the Env maturation process is important for the formation of infectious viral particles, membrane fusion, and intracellular trafficking of proteins via the cellular secretory pathway (26).

Endogenous retroviruses (ERVs) are resident DNA copies that abound in host chromosomal DNA and comprise approximately 6–10% of cat, human, and mouse genome sequences (30-32). ERVs are descendants of exogenous retroviruses that integrated into the germ line of the ancestral host lineage and are now transmitted vertically from parents to offspring in a Mendelian inheritance pattern (33, 34). ERVs are ubiquitously present in all vertebrates (e.g., fish, amphibians, mammals, squamates, turtles, crocodilians, and birds) (35). Most ERVs are inactivated subsequent to the original retroviral infection through accumulation of mutations and deletions in their genes during viral and host genome replications (24, 36-39). ERVs are also controlled gene silencing mechanisms (40-42). However, some ERVs are known to retain replication capacity to produce infectious viruses; this phenomenon has been observed in several different species (43), including mice (36), koalas (44, 45), pigs (46-50), cats (42, 51), and mule deer (52, 53). Interestingly, although most ERVs seem to represent junk DNA, a few ERVs have been co-opted by their hosts to gain physiological functions through “ERV domestication”. Some ERVs that have been domesticated by their hosts eventually gain specific physiological functions; for example, some are antiviral factors (36, 54), have placenta

formation ability (55), have myoblast fusion ability (56), act as mRNA transporters in the nervous system (57, 58), regulate innate immunity (59), or play a physiological role in pregnancy (HEMO) (60). Together, these reports suggest that ERV domestication has contributed dramatically to host evolution.

ERV of domestic cats (ERV-DC) is an endogenous gammaretrovirus of the domestic cat (*F.s catus*), and is classified as an ERV1-3FCa-I by Repbase (51, 61, 62). ERV-DC has a simple genomic structure and encodes a Gag–Pol polyprotein and an Env protein in its 8.8-kbp genome. A unique feature of the ERV-DC family is that proviruses are phylogenetically classified into three genotypes: Genotype I, Genotype II, and Genotype III. Genotype-specific characteristics have been reported. For example, Genotype I and III proviruses can produce replication-competent viruses (ERV-DC10, ERV-DC18 and ERV-DC14), and these viruses use different receptors for infection (42, 51, 63). The Env gene from Genotype I proviruses has also been transduced into feline leukemia virus (FeLV), generating a novel interference subgroup: FeLV subgroup D (FeLV-D) (51). The Genotype II proviruses encode an antiviral factor, Refrex-1, that specifically inhibits Genotype I ERV-DC and FeLV-D infections (64).

Felis lineage is composed of eight small wildcats in the felis genus including Jungle cat (*Felis chaus, F. chaus*), Black-footed cat (*Felis nigripes*), Sand cat (*Felis margarita*), European wildcat (*Felis silvestris silvestris, F.s. silvestris*), Near Eastern wildcat (*Felis silvestris lybica*), central Asian wildcat (*Felis silvestris ornata*), southern African wildcat (*Felis silvestris cafra*), and Chinese desert cat (*Felis silvestris bieti*) as well as the domestic cat (*Felis silvestris catus, F.s. catus*) (65, 66). Previous study regarding to evolution of the Felidae cat family indicated that the *Felis* lineage (also known as the domestic cat lineage) separated from Leopard Cat lineage about 6.2 million years ago (MYA) in North America. Then, the *Felis* lineage subsequently split off at approximately 3.4MYA from Jungle cat and is the most recent lineage.

Through the Bering land bridge, some ancestral cats moved back to Asia via American migrant whilst some other spread in Europe and further to Africa (65, 66). The earliest archaeological evidence indicates that the European wildcat appeared in Europe ~230,000 years ago (65). A wildcat phalanx from the site of Klimonas shows that they were introduced to Cyprus 11000–10500 B.P. (all dates are reported in calibrated years before present), providing the earliest connection between humans and cats (67). The earliest cat to demonstrate a close association with humans is also from Cyprus, where a young wildcat was interred next to a human at the site of Shillourokambos ca. 9,500 y ago (68). Another report on the presence of cats directly dated between 5560–5280 cal B.P. in the early agricultural village of Quanhucun in Shaanxi, China (67).To date, the relationships between wildcats and domestic cats leading to the process of cat domestication remain controversial. For example, the question of whether domestic cats and wildcats evolved from a common ancestor, or whether domestic cats descended from wildcats, remains unresolved. Our previous studies have found that existence of ERV-DC in domestic cat but not in Tsushima leopard cat (51). These results indicated that probably invasion of ERV-DC in *Felis* lineage occurred since separation from Leopard cat lineage. In addition, the existence of ERV-DCs and their functionality in wildcats remains unclear, as does the overall evolution of ERV-DC in members of the *Felis* genus. Therefore, we undertook a study of ERVs, and ERV-DCs in particular, in the *Felis* genus to begin to elucidate these issues. On the other hand, domestic cats are presently widespread worldwide, so interbreeding between free-ranging domestic cats and wildcats occurs frequently both in captivity and in nature results in existence of hybrid cats (69). In addition, many domestic cat breeds were born by human activity results in diversity of cat populations which may associate with genetically infectious diseases and cat migration (70). Thus, the invasion of infectious endogenous

retroviruses such as ERV-DC10 and ERV-DC14 vertically transmitted among cat populations was investigated to figure out these unknown issues.

In this dissertation, I focused on the evolutionary dynamic of ERV-DC in *Felis* lineage which could contribute to both viral and animal evolution as well as cat domestication.

SCOPE OF THIS DISSERTATION

AIM 1.

To demonstrate the distribution of two infectious endogenous retroviruses (ERV-DC10 and ERV-DC14) in mixed-breed and purebred cats

AIM 2.

To demonstrate the fate of endogenous retrovirus segregation wild and domestic cats

AIM 3.

To demonstrate the evolutionary dynamic of ERV-DC in jungle cats

1. CHAPTER ONE

"Distribution of infectious endogenous retroviruses in mixed-breed and purebred cats"

This work has been published as follows:

Ngo MH, Soma T, Youn H-Y, Endo T, Makundi I, Kawasaki J, Miyake A, Nga BTT, Nguyen H, Arnal M, Fernández de Luco D, Deshapriya RMC, Hatoya S, Nishigaki K. 2019. Distribution of infectious endogenous retroviruses in mixed-breed and purebred cats. Archives of Virology. https://doi.org/10.1007/s00705-019-04454-z.

1.1. Abstract

Endogenous retroviruses of domestic cats (ERV-DCs) are members of the genus *Gammaretrovirus* that infect domestic cats (*Felis silvestris catus*). Uniquely, domestic cats harbor replication-competent proviruses such as ERV-DC10 (ERV-DC18) and ERV-DC14 (xenotropic and nonecotropic viruses, respectively). The purpose of this study was to assess invasion by two distinct infectious ERV-DCs, ERV-DC10 and ERV-DC14, in domestic cats. Of a total sample of 1646 cats, 568 animals (34.5%) were positive for ERV-DC10 (heterozygous: 377; homozygous: 191), 68 animals (4.1%) were positive for ERV-DC14 (heterozygous: 67; homozygous: 1), and 10 animals (0.6%) were positive for both ERV-DC10 and ERV-DC14. ERV-DC10 and ERV-DC14 were detected in domestic cats in Japan as well as in Tanzania, Sri Lanka, Vietnam, South Korea and Spain. Breeding cats including Singapura, Norwegian Forest and Ragdoll cats showed high frequencies of ERV-DC10 (60–100%). By contrast, ERV-DC14 was detected at low frequency in breeding cats. Our results suggest that ERV-DC10 is widely distributed while ERV-DC14 is maintained in a minor population of cats. Thus, ERV-DC10 and ERV-DC14 have invaded cat populations independently.

1.2. Introduction

Endogenous retroviruses (ERVs) are remnants of ancestral retroviral infections that are transmitted vertically from parents to offspring according to Mendelian fashion (2, 71, 72). ERVs are present in the genomes of all vertebrates, making up approximately 6%–10% of the cat, human and mouse genomes (30, 32, 73). Most ERVs are inactive, but some ERVs are replication-competent in several species (43) including mice (36), koalas (44, 45), pigs (46-50), and cats (42, 51).

Although many arguments on the time of cat domestication were researched based on archaeological evidences (65, 67, 68, 74-77), the domestic cat, *Felis silvestris catus (F.s. catus)*, is considered as one of the most recently evolved members of the Felidae. *F.s. catus* descended from *Felis silvestris lybica* in the Near East approximately 131,000 years ago (65). The development of domestic breeding cats was a consequence of artificial selection imposed by humans (78, 79). The International Cat Association currently recognizes 71 breeds of domestic cats (www.tica.org) while the Cat Fanciers' Association recognizes 42 pedigreed breeds for display in the Championship Class (www.cfa.org). Cat breed standards are defined by phenotypic characteristics and subspecies have diverged both morphologically and behaviorally (80). Moreover, several morphological characteristics of cat breed are determined by ERV insertions in KIT loci which alter coat pigmentation (81). Endogenous feline leukemia virus (enFeLV) sequences exist in the genomes of the domestic cat and wild species of the genus *Felis*, with an estimated frequency of 6–12 copies per haploid genome in the domestic cat. Approximately 9–16 distinct autosomal loci were detected per domestic cat examined (34, 82-87). A previous study analyzed insertional enFeLV polymorphisms among 79 domestic cats, including purebred and nonbreeding cats, and found that enFeLV-GGAG was present in 12 animals (15.2% of cats and 8.2% of chromosomes examined). The presence of enFeLVs in only these felid species suggested that enFeLVs entered the germline in a common ancestor of domestic cats before the lineage radiated (i.e., millions of years ago) (84). The Burmese cat had a higher proportion of homozygous enFeLV sites (77%) than other cats, perhaps due to limited outbreeding during development of the breed. Many enFeLVs are not fixed in different cat breeds due to the presence of heterozygous enFeLV insertional polymorphic sites, implying that associations between enFeLVs and disease may not affect all members of a breed (83). Associations between enFeLV copy number and the outcomes of exogenous feline leukemia

virus infections remain uncertain (83) although it was established that enFeLV was the counterpart of exogenous feline leukemia virus subgroup B (88, 89).

ERVs of domestic cats (ERV-DCs) are endogenous gammaretroviruses of the domestic cat (*F. s. catus*). ERV-DCs are classified into three genotypes: genotype I (ERV-DC1, -DC2, -DC3, -DC4, -DC8, -DC14, -DC17, and -DC19), genotype II (ERV-DC7 and -DC16), and genotype III (ERV-DC6, -DC10, and -DC18). ERV-DC10, -DC14 and -DC18 are infectious proviruses (42, 51). The *env* genes of genotype I proviruses were transduced into feline leukemia viruses, generating a novel interference subgroup called FeLV subgroup D (51). ERV-DC7 and ERV-DC16 were found to be homozygous in all Japanese domestic cats tested (51). These two loci encoded an anti-retroviral factor (Refrex-1) active against FeLV subgroup D (64).

Our previous work (51) showed that ERV-DC10 was broadly detected in Japanese domestic cats (N=244, 37.7% positive), while ERV-DC14 was only sporadically detected in this cat population (N=244, 2.5% positive). The ERV-DC18 provirus was identified in only one cat and its siblings. ERV-DC18 proviral sequences were nearly identical but distinct from ERV-DC10 sequences. Thus, ERV-DC18 was probably generated by mobilization of ERV-DC10 (51). ERV-DC14 and ERV-DC10 proviruses used different viral receptors, enabling these two ERVs to replicate in different types of cultured cells: while ERV-DC14 broadly infected many cell types, ERV-DC10 had a more limited tropism (42). A single nucleotide polymorphism (A280→T) in the ERV-DC 5' long terminal repeat (LTR) may represent the *cis* element influencing ERV-DC basal promoter activity. The ERV-DC A-type LTR (A280) was less prevalent in cat genomes compared with the T-type LTR (T280) conferring reduced promoter activity based on *in silico* analysis (42). However, the invasion of these two infectious proviruses in cat lineages and their pathogenesis remains unclear.

In this study, we assessed the invasion of these two infectious proviruses (ERV-DC10 and ERV-DC14) in domestic cats. Our results indicated that ERV-DC10 and ERV-DC14 independently invaded domestic cats and that their frequency distributions differed significantly. Furthermore, we investigated the relationships between the regions where domestic cats live and breed with the frequencies of ERV-DC10 and ERV-DC14.

1.3. Materials and Methods

1.3.1. Samples

Blood samples (N=955) from mixed breed cats (N=939) and purebred cats (N=16) were voluntarily submitted by veterinarians in Japan (90). Pure-breeding samples of Japanese domestic cats (N=516) were provided by the Veterinary Diagnostic Laboratory, Marupi Lifetech. Additionally, we collected blood, tissue and DNA samples from cats in different countries including South Korea (N=44), Vietnam (N=20), Sri Lanka (N=20), Tanzania (N=60) and Spain (N=31). Frequencies of ERV-DC14 in Spanish domestic cat samples were detected in our previous study (91). Details of these samples are shown in **Table 1.1**. DNA was extracted from blood and tissue using a DNeasy Blood and Tissue Kit (Qiagen, Osaka, Japan), by phenol/chloroform extraction (92), or using DNAzol (Life Technologies Japan, Tokyo, Japan).

1.3.2. Detection of ERV-DC10 and ERV-DC14

For genotyping of ERV-DC10 (51), we performed PCR using the primer pair Fe-122S (5′-TGAAGGAAGGAACTTTTCATGTAGG-3′) and Fe-38R (5′-CACACATGCTCTAGACACAATACCC-3′) to detect preintegration sites. We performed PCR using the internal primer Fe-36S (5′-AACCGCTTGGTACARTTCATAAGAG-3′) to detect ERV-DC10 insertional polymorphic sites. For ERV-DC14, we performed PCR using

the primer pair Fe-58S (5′-CATTCAGACTTGCAGTTAAGGGACT-3′) and Fe-42R (5′-CCATAGCAGCTGACTAGTTTGAATG-3′) to detect preintegration sites. We performed PCR using the primer Fe-102R (5′-GGATGAGATCCTCCCAGGTG-3′) to detect ERV-DC14 insertional polymorphisms. PCRs were performed using KOD Fx Neo (Toyobo, Osaka, Japan) and cycled as follows: 94ºC for 2 min (predenaturation); 30 cycles of 98ºC for 10 s (denaturation), 62ºC for 30 s (annealing), and 68ºC for 1.5 min (extension). PCR cycling conditions using Gotaq polymerase (Promega, Madison, WI, USA) were as follows: 95ºC for 2 min (predenaturation); 30 cycles of 95ºC for 30 s (denaturation), 57ºC for 30 s (annealing), and 72ºC for 1.5 min (extension); 72ºC for 5 min. To identify single nucleotide polymorphisms in ERV-DC14 *env*, the ERV-DC14 *env* gene was PCR-amplified with primers Fe-510S (5′-AAGGAATTGCCAAAGGAGTTCTAA-3′) and Fe-42R (5'-CCATAGCAGCTGACTAGTTTGAATG -3') using KOD Fx Neo polymerase. The detection and genotyping of proviruses by PCR is shown in **Figure 1.1**A, B. The pre-integration site spanned approximately 500 bp while the insertional polymorphic sites spanned approximately 800bp and 1.2kbp in ERV-DC10 and ERV-DC14, respectively. Heterozygosity was assessed in both pre-integration site and insertional polymorphic sites. However, homozygosity was assessed only in insertional polymorphic sites.

1.3.3. Statistical analyses

Associations between the frequencies of ERV-DC10 or ERV-DC14 and different geographic regions were evaluated by univariable analysis using Chi-square tests, Fisher's exact tests or one-way analysis of variances. P-values < 0.05 were considered statistically significant.

1.4. Results

1.4.1. Frequencies of ERV-DC10 and ERV-DC14 proviruses in Japanese domestic cats

We conducted a large-scale survey to assess the frequencies of ERV-DC10 and ERV-DC14 proviruses in mixed breed cats (N=939 including 244 domestic cats (51)) and purebred cats (N=532) in Japan. Of 1471 Japanese domestic cats, 482 animals (32.8%) were positive for ERV-DC10, of which 333 (22.6%) were heterozygous and 149 (10.2%) were homozygous. Only 57 Japanese domestic cats (4.0%) were positive for ERV-DC14, of which the majority (3.9%) were heterozygous and 1 (0.1%) was homozygous (**Table 1.2**). Only 6 of 1471 Japanese domestic cats (0.4%) were positive for both ERV-DC10 and ERV-DC14. The frequency of ERV-DC10 was 38.4% (N=361) in mixed breed cats while that was 22.7% in total purebred cats (N=121) (**Figure 1.2**A). By contrast, ERV-DC14 was detected more frequently in total purebred cats (N=36, 6.8%) while that was only 2.3% in mixed breed cats (N=21) (**Figure 1.2**B). Among Japanese domestic cats (N=6) positive for both ERV-DC10 and ERV-DC14, five animals were mixed breed cats and one animal (Scottish Fold, ERV-DC10 homozygous and ERV-DC14 heterozygous) was purebred (P = 0.248). These results indicated that ERV-DC10 was detected at significantly higher frequency than ERV-DC14 in Japanese domestic cats. ERV-DC10 was detected at significantly higher frequency in mixed breed compared with purebred cats. Conversely, the frequency of ERV-DC14 was significantly higher in purebred compared with mixed breed cats.

1.4.2. Frequencies of ERV-DC10 and ERV-DC14 proviruses in different cat breeds

To better understand the invasion of ERV-DC10 and ERV-DC14 in domestic cats, we further analyzed the frequencies of these two proviruses in purebred Japanese domestic cats. In total, 532 samples classified into 21 purebred cat breeds were analyzed. The frequencies of

ERV-DC10 in 16 different purebred cats are shown in **Figure 1.3** and were as follows: Singapura (N=10, positive=100%), Tonkinese (N=5, positive=40%), Himalayan (N=16, positive=12.5%), Norwegian Forest (N=22, positive=63.6%), Russian Blue (N=43, positive=4.6%), Scottish Fold (N=110, positive=25.5%), Abyssinian (N=33, positive=27.3%), Persian (N=12, positive=16.6%), American Curl (N=10, positive=30%), American Shorthair (N=132, positive=3.1%), Long-tailed Chinchilla (N=26, positive=11.6%), Maine Coon (N=16, positive=6.3%), Munchkin (N=20, positive=10%), Ragdoll (N=45, positive=82.2%),Somali (N=5, positive=40%) and British Shorthair (N=15, positive=0%). ERV-DC10 was not detected in any of the remaining five purebred cats: Birman (N=1), Siamese (N=2), Bengal (N=5), Exotic Shorthair (N=2), and Laperm (N=2). Only seven cat breeds (**Figure 1.4**) were positive for ERV-DC14 and all were heterozygous: Himalayan (N=16, positive=18.7%), Scottish Fold (N=110, positive=6.4%), British Shorthair (N=15, positive=6.7%), Persian (N=12, positive=16.6%), American Curl (N=10, positive=10.0%), American Shorthair (N=132, positive=15.3%), and Exotic Shorthair (N=2, positive=100%). Cat breeds could be classified into four breeding groups whose frequencies of ERV-D10 differed significantly from one another ($P < 0.05$) (**Figure 1.3**). Group I (82.2–100% positive) included Singapura and Ragdoll cats. Group II (40.0–63.6% positive) included Norwegian Forest, Tonkinese, and Somali cats. Group III (10–30% positive) included American Curl, Abyssinian, Scottish Fold, Persian, Himalayan, Long-tailed Chinchilla, and Munchkin cats. Group IV (0.0–6.3% positive) included Maine Coon, Russian Blue, American Shorthair, Birman, Siamese, British Shorthair, Bengal, Exotic Shorthair and Laperm cats. We also classified cat breeds into two groups based on their significantly different frequencies of ERV-DC14 provirus ($P < 0.05$) (**Figure 1.4**). Group I (10–100% positive) included Exotic Shorthair, Himalayan, Persian, American Shorthair, and American Curl cats. Group II (0.0–6.7% positive) included British Shorthair,

Scottish Fold, Tonkinese, Somali, Singapura, Siamese, Russian Blue, Ragdoll, Norwegian Forest, Munchkin, Maine Coon, Long-tailed Chinchilla, LaPerm, Birman, Bengal, and Abyssinian cats. Overall, the data indicated that ERV-DC10 was widely distributed while ERV-DC14 was present at lower frequencies in breeding cats in Japan.

Previous studies investigated the places and regions where cat breeds were originally established (93, 94). We analyzed associations between the origin of purebred cats and the presence of ERV-DC10 and ERV-DC14 proviruses. The purebred cats used in this study were classified into four origin groups: Asia, Europe, Middle East and North America (94). As shown in **Figure 1.5**, the average frequencies of ERV-DC10 in Asia, Europe, Middle East and North America were 35%, 22%, 21% and 18.3%, respectively. The average frequencies of ERV-DC14 in Asia, Europe, Middle East and North America were 0%, 6.4%, 8.3% and 12.5%, respectively. There were no statistically significant associations between the frequencies of the two proviruses and geographic origin. These results suggested that the distributions of ERV-DC10 and ERV-DC14 proviruses in breeding cats were similar regardless of geographic origin of the breed.

1.4.3. Frequencies of ERV-DC10 and ERV-DC14 proviruses in domestic cats in different countries

Next, we investigated the frequencies of ERV-DC10 and ERV-DC14 proviruses in domestic cats in different countries to determine whether or not their frequencies were similar to those of Japanese domestic cats. Domestic cats from South Korea (N=44), Sri Lanka (N=20), Vietnam (N=20), Tanzania (N=60) and Spain (N=31) were tested for ERV-DC10 and ERV-DC14. As shown in **Figure 1.6**A, frequencies of ERV-DC10 ranged from 19.4% to 66.7 % in domestic cats from five countries other than Japan. Frequencies of ERV-DC10 were highest in Tanzanian domestic cats (66.7%) and lowest in Spanish domestic cats (19.4%). Among

Asian domestic cats (Japan, South Korea, Vietnam, and Sri Lanka), frequencies of ERV-DC10 were highest in Sri Lankan cats (57.9%) and did not differ significantly in cats from Japan, South Korea and Vietnam (32.8%, 45.5% and 40.8%, respectively). ERV-DC14 was detected at similar frequencies in domestic cats from South Korea, Sri Lanka, Vietnam and Tanzania (4.5%, 5.0%, 6.7% and 5.0%, respectively) and with similar frequency to domestic cats in Japan (**Figure 1.6**B). Domestic cats in Spain showed somewhat higher frequencies of ERV-DC14 (9.7%) (91), although this difference was not statistically significant. Notably, one Spanish domestic cat and two South Korean domestic cats were positive for both ERV-DC10 and ERV-DC14. These results also indicated that ERV-DC10 and ERV-DC14 were broadly distributed in domestic cats from different countries, and that these two infectious proviruses independently invaded domestic cat populations.

1.4.4. Frequencies of ERV-DC10 and ERV-DC14 proviruses in domestic cats used in this study

In summary, we assessed the presence of ERV-DC10 and ERV-DC14 in 1646 domestic cats from six different countries including Japan, Vietnam, Sri Lanka, South Korea, Tanzania and Spain. Of these animals, 568 cats were positive for ERV-DC10 (34.5%) and 68 cats were positive for ERV-DC14 (4.1%) (**Table 1.3**). Only 10 cats were double-positive for ERV-DC10 and ERV-DC14 (0.6%). Notably, ERV-DC14 homozygosity was observed in only one of 68 ERV-DC14-positive Japanese domestic cats. These results indicated that ERV-DC10 was more frequently detected in domestic cats compared with ERV-DC14. ERV-DC14 appeared to be maintained in a minor population. Thus, these two infectious proviruses may have independently invaded domestic cat populations.

1.5. Discussions

In this study, we assessed the frequencies of two infectious ERVs (ERV-DC10 and ERV-DC14) in domestic cats of different breeds and different geographic origins. The prevalence of ERV-DC10 in mixed breed cats was significantly higher than that in purebred cats in Japan, while the opposite was true for ERV-DC14 (**Figure 1.2**). This difference may relate to the different breeding strategies of mixed and purebred cats.

The frequencies of ERV-DC10 and ERV-DC14 were not associated with cat geographic origin. However, these frequencies appeared to be associated with specific breeds of cats as shown in **Figure 1.3** and **Figure 1.4**. Frequencies of ERV-DC14 were highest in Exotic Shorthair (100%), Himalayan (18.7%) and American Shorthair (15.3%) cats. Exotic Shorthair and American Shorthair cats originated in the US but now cluster in Western Europe (95). Moreover, Himalayan cat represent hybrids between UK and US cats (94). In the second chapter, we have detected the high prevalence of ERV-DC14 in European wildcats (96). Thus, ERV-DC14 may have originated from European wildcats. This suggestion will be supported and much clearer in the chapter two and three.

ERV-DC10 was detected at high frequencies in Singapura (100%), Ragdoll (82.2%), Norwegian Forest (63.6%), and Somali (40%) cats (**Figure 1.3**). Singapura cats are descended from a few populations in Asia from Asian wildcat ancestors, while Ragdoll cats are a cross breed between US and UK cats (94). Norwegian Forest cats are considered feral cats which were naturally selected (94). In addition, we only detected ERV-DC genotype III which was very close to ERV-DC10 in jungle cat (*F. chaus*) which is probably originated from Asian regions (97) so our results indicated that ERV-DC10 may have first originated in Asian and African wildcats. This suggestion will be supported in chapter two and three.

Our results indicated that ERV-DC10 and ERV-DC14 proviruses were similarly distributed among purebred cats of four origins (Asia, Europe, Middle East and North America) (**Figure 1.5**). This result was not unexpected, since after the establishment of cat breeds, they spread all over the world.

The frequencies of ERV-DC10 and ERV-DC14 in domestic cats in different countries differed within mixed breeding populations. While ERV-DC10 showed the highest and lowest frequencies in Tanzanian and Spanish domestic cats, respectively, frequencies of this provirus were broadly similar among Asian cats (Japan, South Korea and Vietnam). However, frequencies of ERV-DC10 were also high in Sri Lankan domestic cats, which resemble Tanzanian cats (**Figure 1.6**A). This result suggests that Asian and African domestic cats may harbor ERV-DC10 more frequently than domestic cats in Europe. In contrast, ERV-DC14 frequencies were similar among countries (3.9 to 9.7%) but were lowest in Asian and African countries (**Figure 1.6**B). We observed ERV-DC14 homozygosity in only one of 1646 domestic cats, while ERV-DC14 heterozygosity was observed in 67 domestic cats.

Endogenous Jaagsiekte sheep retrovirus (enJSRV) was highly active and abundant in female compared with male sheep because this enJSRV plays a role in trophoblast development (98). In our study, we found an equal distribution of the two infectious ERVs examined (ERV-DC10 and ERV-DC14) in male and female cats. Thus, gender does not appear to affect invasion by these two ERVs (data not shown), although it remains unknown whether ERV-DC activity may be gender-dependent.

Although most ERVs appear to represent junk DNA, a few ERVs have been co-opted by their hosts to gain a variety of physiological functions through “ERV domestication” (1). For example, some are antiviral factors (36, 54) or have placenta formation ability (99). enFeLVs are counterparts of exogenous FeLV subgroup B (88, 100) and produced recombinant

viruses with altered biological activity and pathogenicity (101, 102). enFeLVs were identified in wild species of the genus *Felis* closely related to domestic cats but were not detected in other lineages within the Felidae (82-84). Two endogenous retroviruses in domestic cat (ERV-DC) loci, including ERV-DC7 and ERV-DC16, are fixed in domestic cats and encode antiretroviral factors against FeLV subgroup D and ERV-DC genotype I. Our recent study showed that ERV-DC7 was fixed in European wildcats but that ERV-DC16 was unfixed. Thus, levels of antiretroviral activity against FeLV-D and ERV-DC14 differed slightly between these two cat groups (91). It remains unknown whether the presence of ERV-DC10 and ERV-DC14 in the cat genome is harmful to the host. However, the existence of any replication-competent retrovirus in the genome poses a potential risk. In this study, the infectivity of at least three full-length ERV-DC10 proviral clones and six full-length ERV-DC14 proviral clones from mixed breed and purebred cats was assessed. All ERV-DC10 and ERV-DC14 proviral clones assessed were infectious (data not shown). Thus, replication-competent ERV-DC10 and ERV-DC14 in the domestic cat genome may be mobile and interact with other exogenous retroviruses to generate new recombinant viruses. A similar pattern was observed in mice. The genomes of several mouse strains (e.g., AKR, C58, and HRS) carried endogenous ecotropic murine leukemia viruses (E-MLVs) called *Emvs*, most of which can produce infectious viruses during leukemogenesis when *Emv*-derived E-MLVs establish a chronic infection (36, 103, 104).

The process of cat domestication remains controversial. Domestic cats are thought to have originated from the Near Eastern wildcats (also known as *F.s. lybica*) (65). Ancestral retroviruses, including ERVs, can be valuable tools for understanding the domestication process. enFeLV insertional polymorphisms in different wildcats, breeding and non-breeding cats suggested a potential scenario for cat domestication (82-84), while a similar result using RD114 virus sequences suggested a map of cat migration (70). Although insertional

polymorphisms of the infectious proviruses ERV-DC10 and ERV-DC14 were not suggestive of any particular route of cat domestication, our study ruled out the hypothesis that the different wildcats mainly harbored these two ERV-DC loci (e.g., European wildcats may harbor ERV-DC14 (91) while Asian and African wildcats may harbor ERV-DC10 which will be supported in the chapter three). In addition, ERV-DC14 homozygosity was detected in only one of 68 ERV-DC14-positive domestic cats, while ERV-DC10 homozygosity and heterozygosity were detected in 191 and 377 domestic cats, respectively. Based on these results, we proposed the hypothesis that ERV-DC14 homozygosity (and potentially heterozygosity) may affect during embryogenesis and is deleterious in domestic cats. This hypothesis agrees with a previous study which found that ERV-DC14 broadly infects many species, while ERV-DC10 induced only limited infection in a subset of tested cells. In other words, the infectivity of ERV-DC14 is nonecotropic, whereas that of ERV-DC10 (ERV-DC18) appears to be xenotropic (42). Thus, screening for ERV-DC14 in purebred and mixed breed cats may have important implications for cat reproduction.

This issue needs to be considered further by future studies. In veterinary clinics, production of induced pluripotent stem cells, blood transfusion and bone marrow transplantation would be more safely accomplished using cat donors free of these two infectious ERVs. The relationships between these infectious ERVs (ERV-DC10 and ERV-DC14) and disease in domestic cats remain to be elucidated.

1.6. Conclusions

In summary, two infectious endogenous retroviruses (ERV-DC10 and ERV-DC14) were not fixed in cat populations, unlike ERV-DC7 or ERV-DC16 (51). However, these ERVs still showed high frequencies in domestic cats. These two infectious proviruses have potential to induce disease in cats and to recombine with other feline exogenous retroviruses to generate

new recombinant exogenous retroviruses. The existence of these two infectious ERVs in expanded purebred populations poses risks to the host, especially for treatment through transplantation methods (e.g., bone marrow transplantation). Moreover, investigations of ERV-DC10 and ERV-DC14 may help in understanding feline evolution. This study provides useful information toward improving our understanding of the pathogenicity of infectious ERVs as well as of cat evolution.

1.7. Tables and Figures in Chapter one

TABLES

Table 1.1 Characteristics of samples used in this study

Countries	Cat Breed	Sample No.	Source of DNA
Japan	Mix	939	Blood
	Purebred	532	Blood
Vietnam	Mix	20	Blood
South Korea	Mix	44	Blood
Sri Lanka	Mix	20	Tissues
Tanzania	Mix	60	Blood
Spain	Mix	31	Blood and tissues

Table 1.2 Prevalence of ERV-10 and ERV-DC14 in Japanese domestic cats[a]

Provirus / Cat	ERV-DC10				ERV-DC14				**Subtotal**	**Source**
	-/- (%)	+/- (%)	+/+ (%)	+ (%)	-/- (%)	+/- (%)	+/+ (%)	+ (%)		
Mix	142 (61.2)	70 (30.2)	20 (8.6)	90 (38.8)	227 (97.8)	5 (2.2)	0 (0)	5 (2.2)	232 (100)	(51)
Purebred	10 (83.3)	2 (16.7)	0 (0)	2 (16.7)	11 (91.7)	1 (8.3)	0 (0)	1 (8.3)	12 (100)	
Subtotal	152 (62.3)	72 (29.5)	20 (8.2)	92 (37.7)	238 (97.5)	6 (2.5)	0 (0)	6 (2.5)	244 (100)	
Mix	436 (61.7)	196 (27.7)	75 (10.6)	271 (38.3)	690 (97.6)	16 (2.3)	1 (0.1)	17 (2.4)	707 (100)	This study
Purebred	401 (77.1)	65 (12.5)	54 (10.4)	119 (22.9)	485 (93.3)	35 (6.7)	0 (0)	35 (6.7)	520 (100)	
Subtotal	837 (68.2)	261 (21.3)	129 (10.5)	390 (31.8)	1175 (95.7)	51 (4.2)	1 (0.1)	52 (4.3)	1227 (100)	
Mix	578 (61.6)	266 (28.3)	95 (10.1)	361 (38.4)	917 (97.7)	21 (2.2)	1 (0.1)	22 (2.3)	939 (100)	Combined
Purebred	411 (77.3)	67 (12.6)	54 (10.1)	121 (22.7)	496 (93.2)	36 (6.8)	0 (0)	36 (6.8)	532 (100)	
Total	989 (67.2)	333 (22.6)	149 (10.2)	482 (32.8)	1413 (96.0)	57 (3.9)	1 (0.1)	58 (4.0)	1471 (100)	

[a](+/+), copy present on both chromosomes (homozygous); (+/-), copy present on one of two chromosomes (heterozygous); (-/-), no copies present (null); (+), provirus detected.

Table 1.3 Frequencies of ERV-10 and ERV-DC14 in domestic cats[b]

Provirus	ERV-DC10 (+)	ERV-DC14 (+)	No. cat (positive with two proviruses)
-/-	1078 (65.5)	1578 (95.9%)	
+/-	377 (22.9%)	67 (4.0%)	
+/+	191 (11.6%)	1 (0.1%)	10 (0.6%)
Total (+)	568 (34.5%)	68 (4.1%)	
Total	1646 (100%)	1646 (100%)	

[b](+), provirus detected; (+/-), heterozygous (copy present on one of two chromosomes); (+/+), homozygous (copy present on both chromosomes).

FIGURES

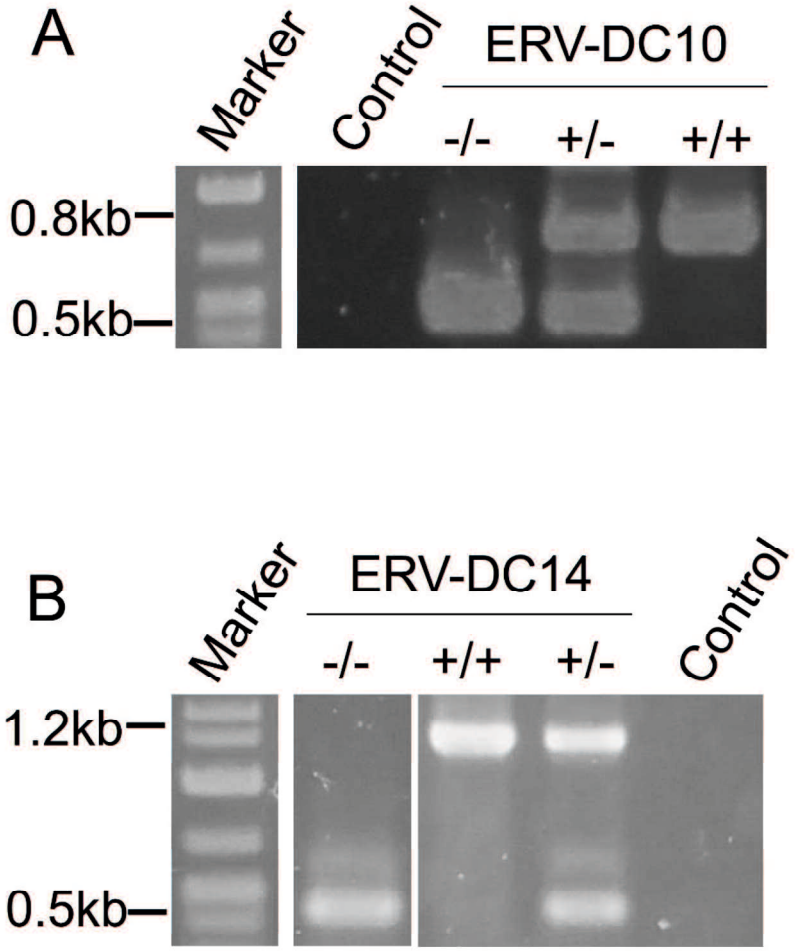

Figure 1.1 Genotyping of ERV-DC10 and ERV-DC14 by PCR. (A) ERV-DC10 detection and genotyping. Band sizes of 0.5kbp and 0.8kbp represent pre-integration sites and proviral insertional polymorphic sites, respectively. (B) ERV-DC14 detection and genotyping. Band sizes of 0.5kbp and 01.2kbp represent pre-integration sites and proviral insertional polymorphic sites, respectively. -/-, no copy of provirus present on either chromosome; +/+, proviral copy present on both chromosomes (homozygous); +/-, proviral copy present on one of two chromosomes (heterozygous).

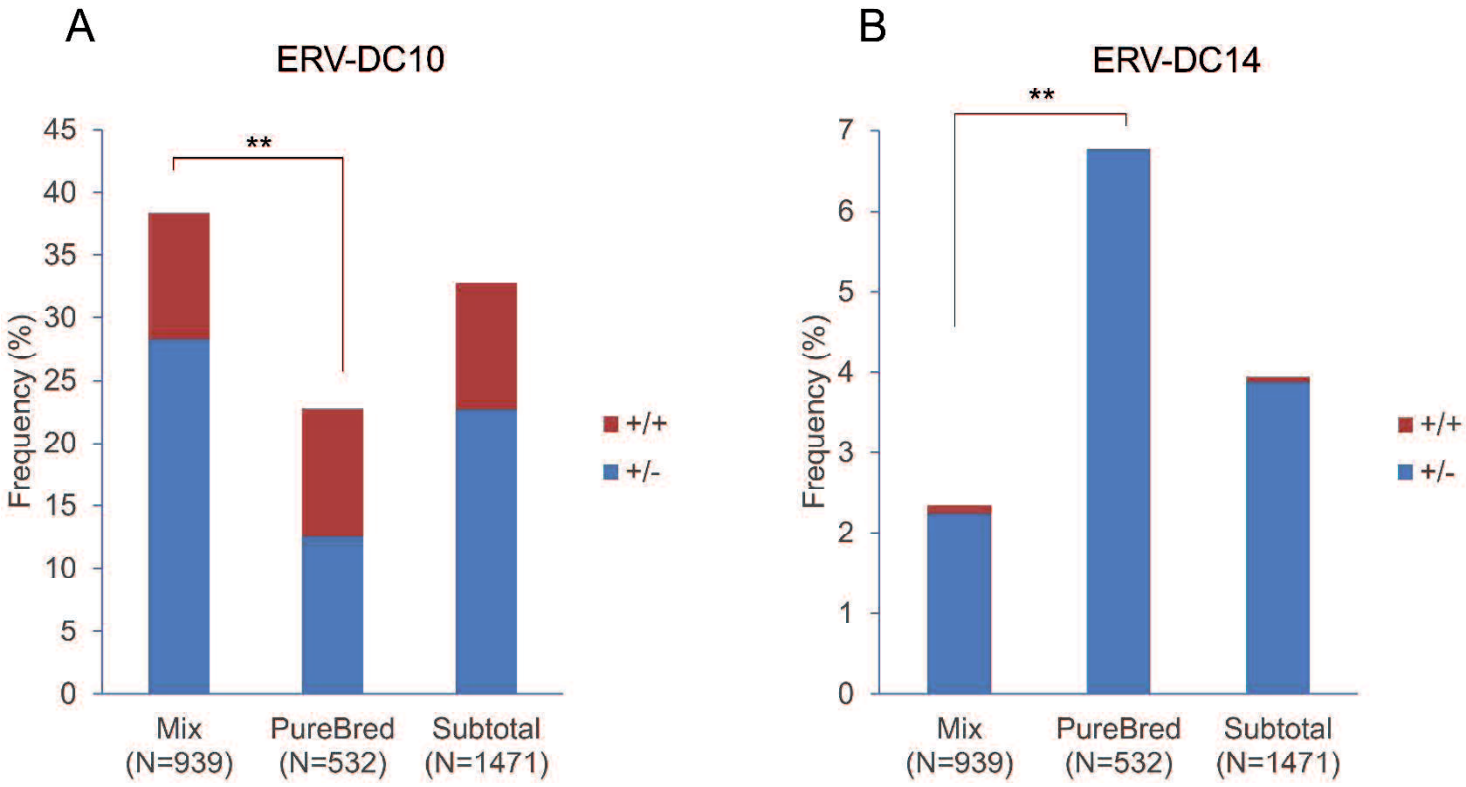

Figure 1.2 Prevalence of infectious ERV-DCs among domestic cats in Japan. (A) Prevalence of ERV-DC10 in Japanese domestic cats. (B) Prevalence of ERV-DC14 in Japanese domestic cats. Japanese domestic cats were divided into two groups including mixed (N=942) and purebred (N=532) animals. +/+, copy present on both chromosome (homozygous); +/-, copy present on one of two chromosomes (heterozygous). Differences between frequencies of ERV-DC10 and ERV-DC14 in different regions were analyzed using Chi-square tests or Fisher's exact tests; **, $P < 0.0001$.

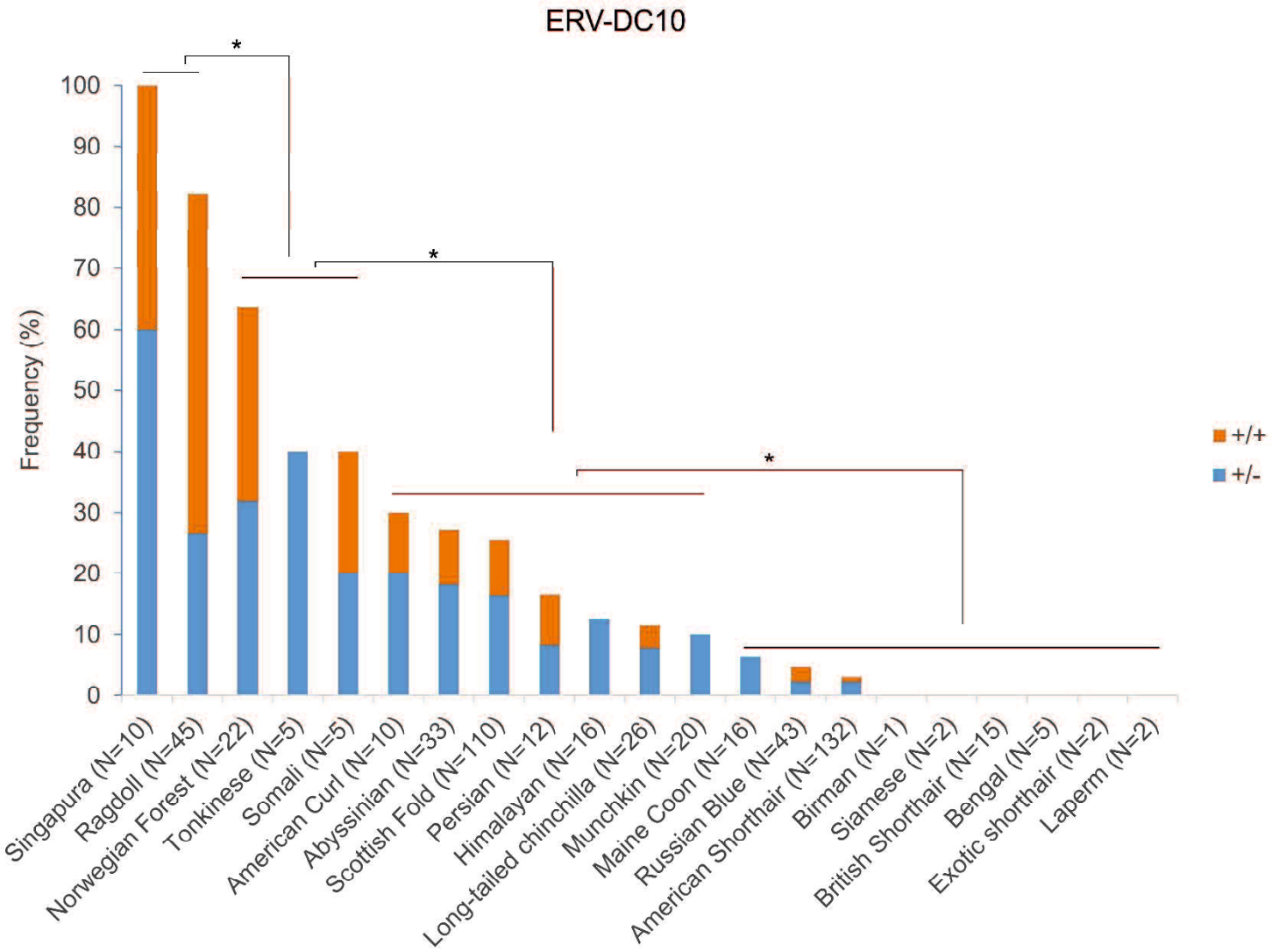

Figure 1.3 Comparison of ERV-DC10 prevalence among positive cat breeds. A total of 21 cat breeds investigated by ERV-DC10 genotyping are shown. +/+, proviral copy present on both chromosomes (homozygous); +/-, proviral copy present on one of two chromosomes (heterozygous). Numbers indicate sample sizes for each purebred cat on the X axis. *, $P < 0.05$ (Fisher's exact test).

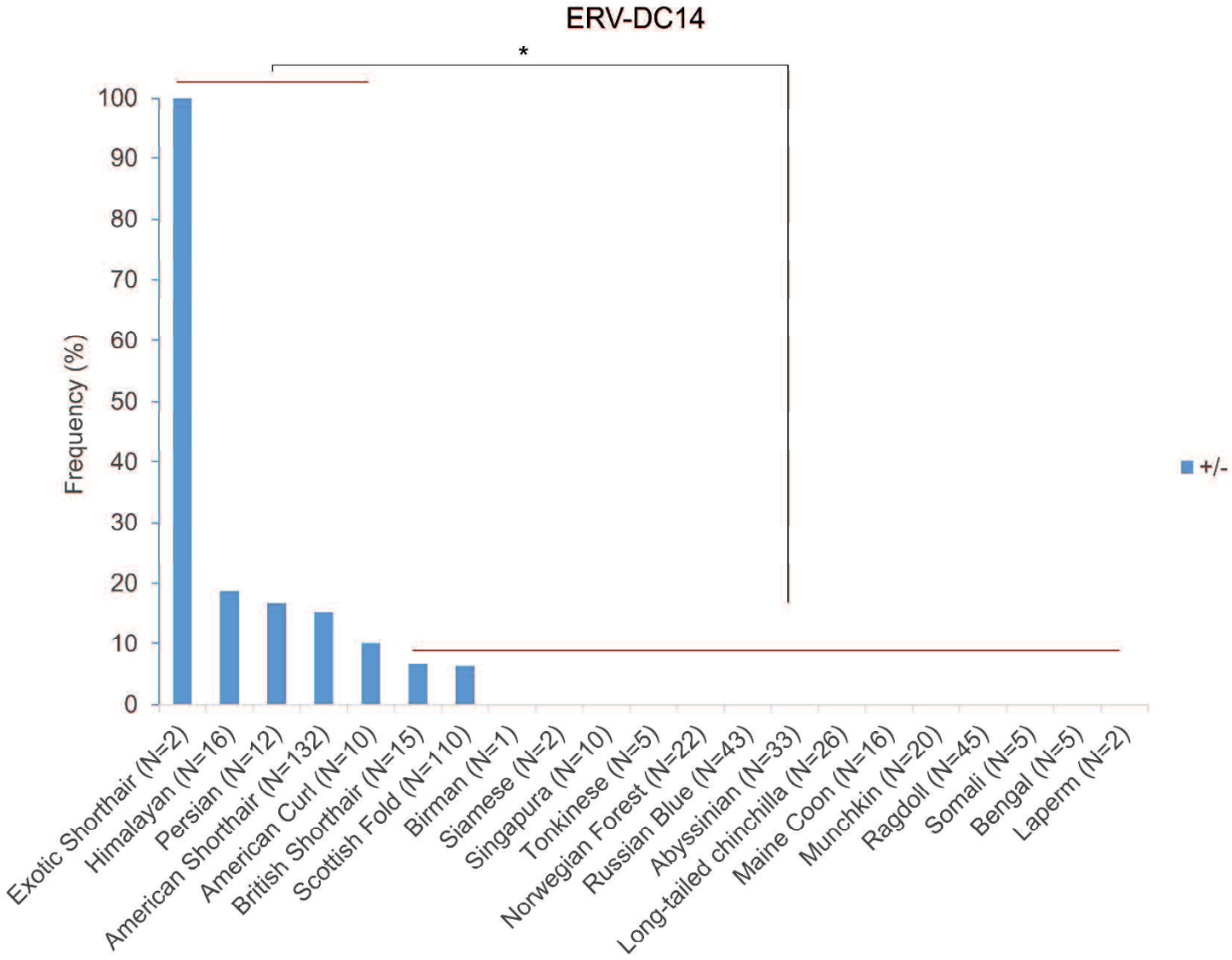

Figure 1.4 Comparison of ERV-DC14 prevalence among positive cat breeds. A total of 21 cat breeds investigated for ERV-DC14 were shown. +/+, copy present on both chromosomes (homozygous); +/-, copy present on one of two chromosomes (heterozygous). Numbers indicate sample sizes for each purebred cat on the X axis. *, $P < 0.05$ (Fisher's exact test).

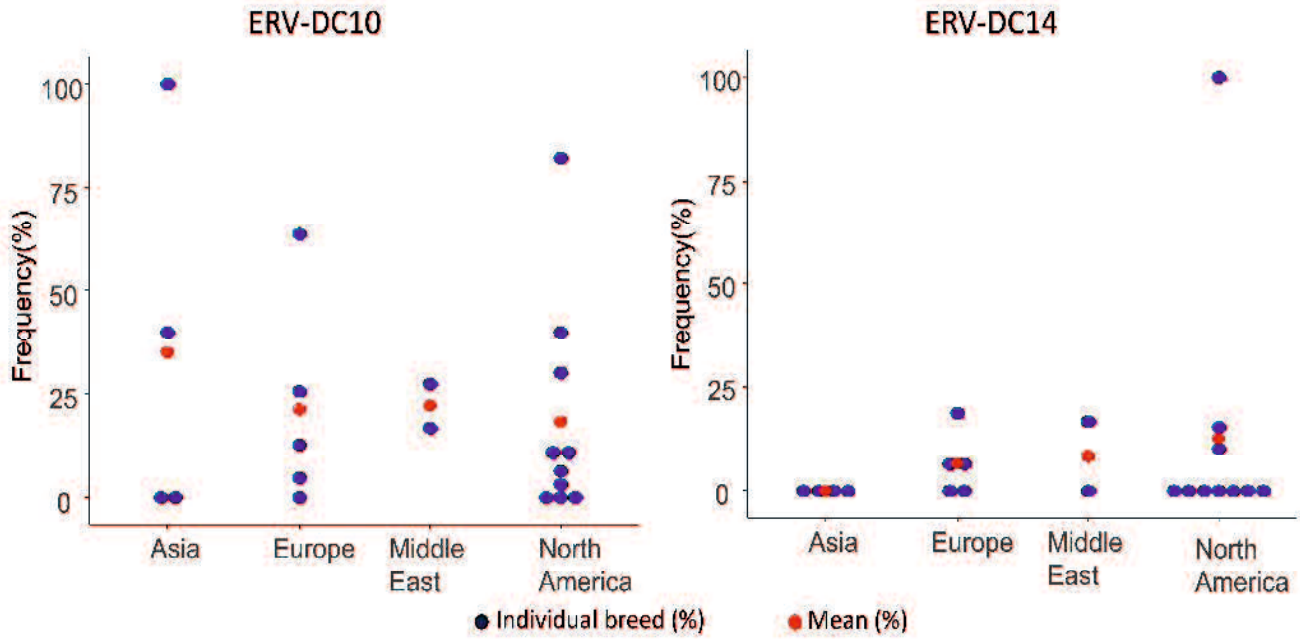

Figure 1.5 Prevalence of ERV-DC10 and ERV-DC14 among purebred cats of different origins. Based on the origins of different cat breeds previously described (94), we classified purebred cat samples into four regions. Four breeds were assigned to Asia (Birman, Siamese, Singapura, and Tonkinese). Five breeds were assigned to Europe (British Shorthair, Norwegian Forest, Russian Blue, Scottish Fold and Himalayan). Ten breeds were assigned to North America (Ragdoll, Munchkin, American Curl, American Shorthair, Bengal, Exotic Shorthair, Laperm, Long-tailed Chinchilla, and Maine Coon). Two breeds were assigned to the Middle East (Abyssinian and Persian). Red circles indicate the frequency of each pure breed. Black circles indicate the mean prevalence for each breed's origin with 95% confidence intervals. Differences between frequencies of ERV-DC10 and ERV-DC14 in different regions were analyzed using Fisher's exact test.

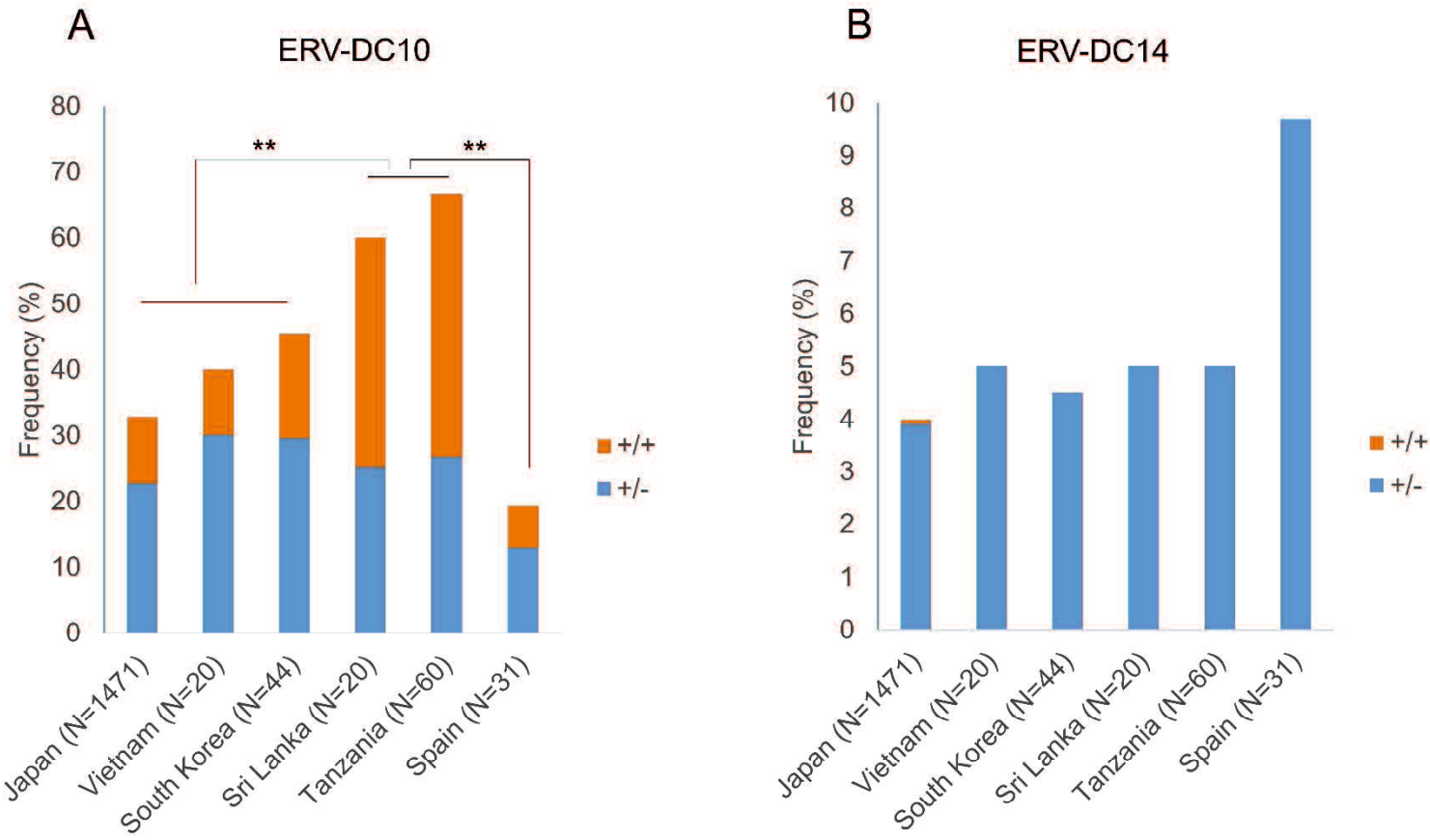

Figure 1.6 Prevalence of two infectious ERV-DCs in purebred cats of different countries. Frequencies of ERV-DC10 (A) and ERV-DC14 (B) in domestic cats in different countries. +/+, copy present on both chromosomes (homozygous); +/-, copy present on one of two chromosomes (heterozygous). Differences between frequencies of ERV-DC10 and ERV-DC14 in different regions were analyzed using Chi-square tests or Fisher's exact tests; **, $P < 0.0001$.

2. CHAPTER TWO

"Tracking the fate of endogenous retrovirus segregation in wild and domestic cats"

This work has been published as follows:

Ngo MH, Arnal M, Sumi R, Kawasaki J, Miyake A, Grant CK, Otoi T, Fernández de Luco D, Nishigaki K. 2019. Tracking the fate of endogenous retrovirus segregation in wild and domestic cats. Journal of Virology. https://doi.org/10.1128/JVI.01324-19.

2.1. Abstract

Endogenous retroviruses of domestic cats (ERV-DCs) are one of the youngest feline ERV groups in domestic cats (*Felis silvestris catus*); some members are replication-competent (ERV-DC10, ERV-DC18, and ERV-DC14), produce the anti-retroviral soluble factor Refrex-1 (ERV-DC7 and ERV-DC16), or can generate recombinant feline leukemia virus (FeLV). Here, we investigated ERV-DC in European wildcats (*Felis silvestris silvestris*) and detected four loci: ERV-DC6, ERV-DC7, ERV-DC14, and ERV-DC16. ERV-DC14 was detected at a high frequency in European wildcats; however, it was replication-defective due to single G→A nucleotide substitution resulting in an E148K substitution in ERV-DC14 envelope (Env). This mutation results in a cleavage-defective Env that is not incorporated into viral particles. Introduction of the same mutation into feline and murine infectious gammaretroviruses resulted in similar Env dysfunction. Interestingly, the same mutation was found in a FeLV from naturally-occurring thymic lymphoma and a mouse ERV, suggesting a common mechanism of virus inactivation. Refrex-1 was present in European wildcats; however, ERV-DC16, but not ERV-DC7, was unfixed in European wildcats. Thus, Refrex-1 has had an antiviral role throughout *Felis* evolution, pre-dating cat exposure to feline retroviruses. ERV-DC sequence diversity was present across wild and domestic cats but was locus-dependent. In conclusion, ERVs have evolved species-specific phenotypes through the interplay between ERVs and their hosts. The mechanism of viral inactivation may be similar irrespective of the evolutionary history of retroviruses. The tracking of ancestral retroviruses can shed light on their roles in pathogenesis and host-virus evolution.

2.2. Introduction

Endogenous retroviruses (ERVs) are integrated retroviral elements that make up 6%–10% of cat, human, and mouse genome sequences (30, 32, 73). Most ERVs are disrupted subsequent to the original retroviral integration through accumulation of mutations, deletions or insertions in their genes during viral and host genome replication (24, 36-39). However, some infectious ERVs have been identified in mice (36) and cats (42, 51). Feline endogenous retroviruses have been identified and grouped phylogenetically into different classes (105-107). For example, endogenous feline leukemia viruses (enFeLV) became integrated within the genomes of members of the *Felis* genus more than 2 million years ago (34, 43, 82, 84). These ERVs are present at 6–12 copies per haploid genome in domestic cats (85-87), while fluorescent *in situ* hybridization detected 9–16 distinct autosomal enFeLV loci per domestic cat (83). enFeLVs can recombine with exogenous feline leukemia virus (FeLV) to yield recombinant FeLV subgroup B (102, 108). Additional feline endogenous ERVs have been characterized including RD-114 (109), MAC-1 (110, 111), and feline endogenous retrovirus gamma4 (112).

One of the youngest feline ERV groups, called ERVs in domestic cats (ERV-DCs), are estimated to have integrated within the cat genome approximately 2.8 million years ago. ERV-DCs are classified as endogenous gammaretroviruses (51, 61, 62, 105, 106). We previously identified and cloned 13 ERV-DC loci and estimated that there were 7–17 ERV-DC copies present in each domestic cat. ERV-DCs have a simple retroviral structure including *gag, pol,* and *env* genes enclosed between two non-coding long terminal repeats (LTRs) (51, 63). A unique feature of the ERV-DC family is that proviruses can be phylogenetically classified into three genotypes (**Figure 2.1A**): Genotype I (ERV-DC1, -DC2, -DC3, -DC4, -DC8, -DC14, -DC17, and -DC19), Genotype II (ERV-DC7 and -DC16), and Genotype III (ERV-DC6, -DC10 and -DC18). Among the ERV-DCs, ERV-DC10, -DC14 and -DC18 are infectious proviruses.

ERV-DC18 may have been generated by retrotransposition during ERV-DC10 reintegration or reinfection in different members of one cat family (51). ERV-DC14 showed low promoter activity in its 5'LTR due to a single A to T mutation. Reverting this mutation in ERV-DC14 (called ERV-DC14TA) enhanced its replication and enabled the ERV to persistently infect HEK-293T cells (42). A survey of insertional polymorphisms within ERV-DCs in Japanese domestic cats indicated that a low proportion (2.5%) of cats tested carried ERV-DC14 (51). Notably, FeLV-positive cells were transduced with the *env* gene from a Genotype I provirus, generating a novel interference subgroup called FeLV subgroup D (FeLV-D) (51). Genotype II proviruses were disrupted by mutations and deletions in the *pol* and *env* genes. However, the truncated Env protein of these proviruses (ERV-DC7 and -DC16) encoded an antiviral factor, called Refrex-1, that specifically inhibits ERV-DC genotype I and FeLV-D infections. Refrex-1 is efficiently secreted from feline cells as a soluble protein and may interfere with virus interaction with host cell receptors (64). ERV-DC6, -DC7 and -DC16 were apparently fixed in Japanese domestic cats, while the other ERV-DCs were polymorphic (51). Other examples of ERV *env* genes conferring resistance to viral infection have been demonstrated in the laboratory and in house and wild mice; these include *Fv-4, Rmcf, and Rmcf2*, which inhibit infection by ecotropic, polytropic, and xenotropic murine leukemia virus (MLV), respectively (36, 113-116). The *Mus musculus castaneus* endogenous virus (MLV/MmCN, located in the qE1 region of chromosome 8) was amplified from DNA of *M. musculus castaneus* (strain CAST/Ncr) trapped in Lake Casitas (117). The sequence of this Cas subtype Env resembled that of *Fv-4* (117), a defective endogenous MLV encoding a truncated Env that acts as a host restriction factor to block infection by ecotropic MLVs (114).

We previously reconstructed the full-length *env* genes of ERV-DC7 and ERV-DC16 (called ERV-DC7fl and ERV-DC16fl respectively) to assess the role of Refrex-1 in virus-host

coevolution. ERV-DC7fl and ERV-DC16fl were unable to produce infectious viral particles due to defects in Env cleavage. Defects in ERV-DC7fl Env resulted from three determinant residues (R407, I421 and T429). Reverse genetics methods were used to successfully reconstruct an infectious ERV-DC7fl *env* bearing the ERV-DC14 *env* consensus residues at these three positions (R407G, I427N, and T429A). Analyses of ERV-DC7 *env* sequence diversity in Japanese domestic cats indicated that the determinants of ERV-DC7fl dysfunction were not fixed in the population. Four variants were identified with different combinations of residues at these positions: 407G and 427N–429A (G-NA), 407R and 427N–429A (R-NA), 407G and 427I–429T (G-IT), and 407R and 427I–429T (R-IT). These variants have been present because the integration of ERV-DC7 into the host genome and the sequence of ERV-DC7 *env* noncoding regions underwent purifying selection between the time of its integration and its truncation. The T nucleotide at position 801 encodes a latent stop codon in ERV-DC7fl Env, and was predominant in most Japanese domestic cats examined. Two animals were exceptions (IDs ON-C and ON-R) and bore a C nucleotide at this position, which did not affect the function of ERV-DC7fl Env (23). Thus, characterization of ERV-DC7 in the *Felis* genus could help clarify the diversity of ERV-DCs in cat lineages.

The wildcat (*Felis silvestris*) population is dispersed throughout the Old World, but there has been little description of its subspecies (118). The coexistence of three subspecies including the European wildcat (*F.s. silvestris*), the African wildcat (*F.s. lybica*), and the domestic cat (*F.s. catus*), was reported in different regions across Europe (118-120). The earliest archaeological evidence indicates that the European wildcat appeared in Europe ~230,000 years ago (65). *F.s. catus* were domesticated from *F.s. lybica* approximately 131,000 years ago in the Near East (65, 121). The results of a mitochondrial DNA analysis suggested that a common ancestor of both European wildcats and sand cats (*F.s. margarita*) was genetically

distinct from the ancestor of domestic cats (*F.s. catus*) (65). ERV-DCs were demonstrated to be phylogenetically distinct from enFeLVs in both domestic cat (*F.s. catus*) and wildcats (51, 82, 105, 106). The domestic cat and its immediate progenitor, the wildcat (*F.s. silvestris/F.s. lybica*), showed a tremendous diversity of enFeLVs (82). To date, the relationships between European wildcats and domestic cats remain controversial. For example, the question of whether domestic cats and wildcats evolved from a common ancestor, or whether domestic cats descended from wildcats, remains unresolved. In addition, the existence of ERV-DCs and their functionality in wildcats remains unclear, as does the overall evolution of ERV-DC in members of the *Felis* genus. Therefore, we undertook a study of ERVs, and ERV-DCs in particular, in the *Felis* genus to begin to elucidate these issues.

In this study, we assessed the presence of ERV-DCs and analyzed their insertional polymorphisms in European wildcats and domestic cats. Ours is the first report showing that a species-specific inactivation of infectious endogenous retroviruses also contributes to a common mechanism of viral inactivation employed by the host against both endogenous and exogenous retroviruses. Additional analyses of ERV-DCs in wildcats and domestic cats revealed unexpected retroviral diversity and clarified several other issues regarding the fate of ERV endogenization, retroviral pathogenesis, and host-virus interactions.

2.3. Materials and Methods

2.3.1. Cell lines.

HEK293T (122) and *Mus dunni* tail *fibroblast* (MDTF) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 1× penicillin/streptomycin. GPLac cells (51), an *env*-negative packaging cell line containing a β-galactosidase (LacZ)-coding pMXs retroviral vector, and 293Lac cells (123) containing a

LacZ-coding pMXs retroviral vector, were cultured in DMEM supplemented with 10% FCS and 1× penicillin/streptomycin.

2.3.2. Samples.

This study on European wildcats was performed with the permission and support of local authorities (Navarra Government) to update information on the health status of this population. All samples were collected from carcasses of European wildcats during between 2000 and 2007 and stored frozen at −18 °C. The collection of this material did not require the approval of the Ethics Committee for Animal Experimentation because it was considered a routine veterinary practice without planned experimentation. Tissue samples from domestic cats were collected from animals that died and were brought to the Department of Animal Pathology in Faculty of Veterinary Medicine in University of Zaragoza for necropsy. Blood samples were provided by clinical veterinarians with the owner's permission. Muscle tissues from European wildcats and blood from domestic cats in Spain were used for DNA extraction with a DNeasy Blood and Tissue Kit (Qiagen, Osaka, Japan) or with phenol and chloroform extraction (92).

2.3.3. Cloning of ERV-DC proviruses from European wildcats.

ERV-DC7, ERV-DC14, and ERV-DC16 full-length proviral genomes were amplified from the splenic DNA of European wildcats using different primer pairs [58S-42R(51), 66S-53R(51), and 219S-44R (64) respectively]. Each full-length DNA fragment was cloned into a pCR4 blunt-TOPO (Invitrogen) vector and sequenced.

2.3.4. PCR.

We used KOD FX Neo (Toyobo, Osaka, Japan), KOD plus Neo (Toyobo, Osaka, Japan), and GoTaq (Promega, Madison, WI, USA) for genotyping insertional polymorphisms and various cloning procedures. The primers designed based on unique sequences both outside and inside

of each ERV-DC provirus to determine ERV-DC haplotypes are listed in a previous publication (51) and in **Table 2.1**.

2.3.5. Construction of chimera proviruses.

We constructed chimera proviruses between ERV-DC14/SO38 (51) and ERV-DC14/*F.s. silvestris* using both restriction enzymes and site-directed mutagenesis. First, two restriction enzymes, *Sal*I (Takara) and *Not*I (Takara), were used to make Chimera1 and Chimera2. Second, we used two restriction site pairs (*Sal*I–*Xho*I) and (*Xho*I–*Not*I) to produce Chimera3 and Chimera4. Finally, Mutant1 (A→G at nucleotide position 6735) and Mutant2 (T→C at nucleotide position 7110), resulting in K148E and S273P changes in ERV-DC14/*F.s. silvestris* Env, respectively, were constructed using a QuickChange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer's protocol. The primer pairs used for site-directed mutagenesis were 619S-597R and 620S-598R for Mutant1 and Mutant2, respectively. Chimeras and mutants were confirmed by direct sequencing.

2.3.6. Construction of expression vectors.

The expression vector pFUΔss (51) was used for constructing the expression plasmids pFUΔss DC7 Env/*F.s. silvestris* and DC14 Env/*F.s. silvestris*. *Env* fragments were amplified from ERV-DC7/*F.s. silvestris* and ERV-DC14/*F.s. silvestris* proviruses using the primer pairs Fe-650S/Fe-626R and Fe-610S/Fe-184R (64), respectively, and cloned into a pFUΔss vector between *Bam*HI and *Eco*RI restriction sites. The site-directed mutagenesis method described above was also used to construct expression plasmids for the Env point mutants pFUΔss ERV-DC14/*F.s. silvestris*/K148E, pFUΔss ERV-DC14/*F.s. silvestris*/S273P, pFUΔss Ampho-MLV/148K, pFUΔss Friend-MLV/148K, pFUΔss FeLV-A/148K, and pFUΔss FeLV-B/148K. The primers used for site-directed mutagenesis are shown in **Table 2.1**. The resulting mutants and Env expression plasmids were confirmed by sequencing.

2.3.7. Transfection.

HEK293T, 293Lac, and GPLac cells were transfected with plasmids using TransIT®-293 Transfection Reagent (Mirus) or Lipofectamine 3000 (Invitrogen) in accordance with each manufacturer's instructions.

2.3.8. Viruses.

To obtain infectious viruses, 293Lac cells containing a LacZ-coding pMXs retroviral vector were first seeded at a concentration of 1×10^6 cells in a six-well plate 1 day prior to transfection. The cells were then transfected with ERV-DC14TA (42), ERV-DC14/(clones: SO38 (51), GM21, IK19, FO16), ERV-DC14/*F.s. silvestris* (clones: wildcat54 and wildcat63), chimera1, chimera2, chimera3, chimera4, mutant1, and mutant2. Three days later, the corresponding supernatants were collected, filtered through a 0.22-μm filter (Merck Millipore, Burlington MA), and stored at −80 °C. To obtain LacZ-carrying Env pseudotyped viruses, GPLac cells were transfected with the corresponding Env expression plasmids of ERV-DC14, ERV-DC14/*F.s. silvestris*/WT, K148E (ERV-DC14/*F.s. silvestris*/K148E), S273P (ERV-DC14/*F.s. silvestris*/S273P), Ampho-MLV (4070A)/WT, Friend-MLV (clone57)/WT, FeLV-A (FeLV clone33)/WT, FeLV-B (Gardner-Arnstein)/WT, Ampho-MLV (4070A)/148K, Friend-MLV (clone57)/148K, FeLV-A (FeLV clone33)/148K, FeLV-B (Gardner-Arnstein)/148K, FeLV/KS16-1 and FeLV/KS16-2. The culture supernatants were also collected after 72 h, filtered through a 0.22-μm filter, and stored at −80 °C.

2.3.9. Infection assay.

The target cells, HEK293T or MDTF, were seeded at a concentration of 3×10^4 cells in 24-well plates one day prior to infection. HEK293T or MDTF cells (targeted for Friend-MLV) were separately incubated with 250 μL of each virus listed above in the presence of polybrene (8 μg/mL). After 48 h of incubation, the cells were stained with 5-bromo-4-chloro-3-indolyl-β-

D-galactopyranoside (X-Gal), and the single-cycle infectivity was assessed by counting blue-stained nuclei as visualized under a microscope. The viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations.

2.3.10. Viral infection assay in the presence of Refrex-1.

HEK293T cells were transfected with provirus clones (ERV-DC7, ERV-DC16, ERV-DC7/*F.s. silvestris*, and ERV-DC16/*F.s. silvestris*), Env-expression vectors (ERV-DC7/*F.s. silvestris* or ERV-DC7), or empty vector (mock). The resulting culture supernatants were collected, filtered, and stored at −80 °C. Fresh HEK293T cells were incubated with serial dilutions of the supernatants for 6 h at 37 °C under a humidified atmosphere containing 5% CO_2. After removing the supernatants, the cells were inoculated with a replication-competent virus, ERV-DC14TA (42), which produces a persistent infection in HEK293T cells. At 48 h post-infection, the infected cells were stained with X-Gal, and viral titers were calculated as the log of infectious units (IU) per mL with standard deviations.

2.3.11. Virus purification.

Culture supernatants (4 mL) were collected at approximately 72 h post-transfection, filtered through 0.45-μm filters, and ultracentrifuged for 90 min at 29,000 ×*g* at 4 °C in an Optima Max-XP ultracentrifuge (Beckman Coulter KK, Ariake, Tokyo, Japan). The resulting virions were resuspended in 20 μL of PBS, then used for western blotting.

2.3.12. Immunoblotting.

Immunoblotting was performed as previously described (5). The primary antibodies used in these assays were goat polyclonal anti-FeLV SU (gp70) (National Cancer Institute [NCI], Frederick, MD, USA), mouse monoclonal anti-FeLV SU (gp70) (C11D8) (Custom Monoclonals International, CA, USA), goat anti-Rauscher MLV SU (gp70) (NCI, Frederick, MD, USA), goat anti-Rauscher MLV CA (p30) (NCI), rat monoclonal anti-Ampho-MLV SU

(gp70) (83A25) (a gift of Dr. Leonard Evans from NIH/NIAD), mouse anti-FeLV TM (p15E) (PF6J2A) (Custom Monoclonals International, CA, USA), mouse anti-FeLV TM (p15E) (EC6-6B1) (Custom Monoclonals International, CA, USA), and mouse monoclonal anti-human β-actin (Santa Cruz Biotechnology, Dallas. Texas). The secondary antibodies used in these assays were horseradish peroxidase (HRP)-conjugated anti-mouse IgG (Cell Signaling Technology, Danvers MA) or HRP-conjugated donkey anti-goat IgG (Santa Cruz Biotechnology).

2.3.13. Flow cytometry analysis.

HEK293T cells were seeded in six-well plates and transfected with empty vector or Env expression plasmids for either ERV-DC14 *env* or ERV-DC14/*F.s. silvestris env*. After 48 h, the cells were harvested with PBS containing 5 mM ethylenediaminetetraacetic acid and washed with PBS containing 0.1% bovine serum albumin (BSA). Cells were fixed for 30 min at room temperature with PBS containing 1% FCS and 10% formaldehyde. Some cells were permeabilized with 0.2% Trion X-100 at room temperature for 15 min. Permeabilization was used to detect protein in cytoplasm (intracellular) while non-permeabilization was used to detect protein expressed at the cell membrane (cell surface). All cells were blocked with 1% BSA in PBS for 30 min at 4 °C. Goat anti-FeLV SU (gp70) (NCI) or rabbit anti-human AKT (serine/threonine-protein kinase) (Cell Signaling, Danvers MA) were used as the primary antibodies, and phycoerythrin (PE)-conjugated mouse anti-goat IgG (Santa Cruz Biotechnology) or fluorescein isothiocyanate (FITC)-conjugated goat F(ab')2 anti-rabbit IgG Fc (Abcam, Cambridge, UK) were used as the secondary antibodies. Cells were treated with each antibody for 30 min at 4 °C. Finally, cells were resuspended in 500 µL of wash buffer and analyzed using a BD Accuri C6 flow cytometer (Becton, Dickinson and Company, Franklin Lakes, NJ, USA).

2.3.14. Phylogenetic and sequencing analysis.

We retrieved 3'-LTR sequences of 13 loci of ERV-DC from NCBI. As a nucleotide substitution model, a Kimura 2-parameter model (124) with uniform rates was used because it had the lowest Bayesian information criterion (BIC) score. All positions containing gaps and missing data were eliminated from the analysis. Tree robustness was evaluated using the bootstrap method (1,000 replicates). All programs used for phylogenetic analyses were packaged in MEGA X (125). Accession numbers are ERV-DC1 (AB674439.1), ERV-DC2 (AB674449.1), ERV-DC3 (AB674440.1), ERV-DC4 (AB674441.1), ERV-DC6 (AB674450.1), ERV-DC7 (AB807599.1), ERV-DC8 (AB674443.1), ERV-DC10 (AB674444.1), ERV-DC14 (AB674445.1), ERV-DC16 (AB807600.1), ERV-DC17 (AB674446.1), ERV-DC18 (AB674447.1), and ERV-DC19 (AB674448.1).

The Env amino acid sequences of ERV-DC (ERV-DC10 and -DC14), gibbon ape leukemia virus (GALV), Friend murine leukemia virus (Friend-MLV), feline leukemia virus subgroup D/TY26peL (FeLV-D), feline leukemia virus TG35-2 (FeLV/TG35-2), moloney murine leukemia virus (MoMLV), koala retrovirus (KoRV-A and -B), endogenous FeLV AGTT locus (enFeLV/AGTT), feline leukemia virus subgroup B/Gardner-Arnstein (FeLV-B), ampho murine leukemia virus (Ampho-MLV), porcine endogenous retrovirus (PERV-A), feline leukemia virus subgroup A clone33 (FeLV-A), feline leukemia virus subgroup C/Sarma strain (FeLV-C), feline leukemia virus subgroup T (FeLV-T), xenotropic murine leukemia virus-related virus (XMRV), polytropic murine leukemia virus AKR13 (P-MLV), FeLV/KS16-1, MLV/MmCN (murine endogenous virus clone MmCN), HERV-T/Pongo (human endogenous retrovirus in *Pongo pygmaeus*), and HERV-T (human endogenous retrovirus T) were obtained from the NCBI database. Multiple alignments of the above amino acid sequences were generated using mafft (126).

We next obtained ERV-DC7 *env* sequences of 20 Japanese domestic cats (23, 64) from the National Center for Biotechnology Information (NCBI) databases. We also used PCR to amplify ERV-DC7 *env* from the chromosomal DNA of 11 European wildcats and 14 European domestic cats with the primer pair Fe-615S (5′-CCTCCAAGCCCTTTATCCTC-3′) and Fe-53R designed for the *pol* region and 3′-flanking region, respectively, and directly determined their *env* region sequences. The accession numbers for the nucleotide sequences of ERV-DC7 *env* are included in the following section. As a nucleotide substitution model, a Kimura 2-parameter model (124) with discrete gamma-distributed rate variation and inferred proportion of invariable sites (+ G, + I; $\alpha = 0.0500$, invar = 0.4934) was used. This model was selected because it had the lowest BIC score. All positions containing gaps and missing data were eliminated from the analysis. Tree robustness was evaluated using the bootstrap method (1,000 replicates). All programs used for phylogenetic analyses were packaged in MEGA X (125).

We next used PCR to amplify the 5′-LTR of ERV-DC14 from the chromosomal DNA of nine European wildcats, four European domestic cats, and six Japanese domestic cats with the primer pair Fe-603S and Fe-140R. These primers were designed to anneal to the 5′-flanking region and the *gag* region, respectively, and we directly determined the 5′-LTR region sequences. The accession numbers for the nucleotide sequences of the ERV-DC14 5′-LTR are included in the following section. As a nucleotide substitution model, a Kimura 2-parameter model (124) with uniform rates was used because it had the lowest BIC score. All positions containing gaps and missing data were eliminated from the analysis. Tree robustness was evaluated using the bootstrap method (1,000 replicates). All programs used for phylogenetic analyses were packaged in MEGA X (125).

We also used PCR with the primer pair Fe-609S (5’-ATATGCCCTCCCTAAGACTTCAAG-3’) and Fe-591R (5’-GATTCCATGGCCCTGAAGTAAGAA-3’) to amplify the gene

encoding cytochrome b, including a partial control region in mitochondrial DNA, to confirm European wildcat's samples. As a nucleotide substitution model, a Hasegawa-Kishino-Yano model (127) with uniform rates was used because it had the lowest BIC score. All positions containing gaps and missing data were eliminated from the analysis. Tree robustness was evaluated using the bootstrap method (1,000 replicates). All programs used for phylogenetic analyses were packaged in MEGA X (125).

2.3.15. Statistical analysis.

The results of infection assays were considered statistically significant if *p*-values were <0.05 by a Student's *t*-test and one-way analysis of variance (ANOVA).

2.3.16. Ethical approval.

Animal studies were conducted following the guidelines for the care and use of laboratory animals of the Ministry of Education, Culture, Sports, Science and Technology, Japan. All experiments were approved by the Genetic Modification Safety Committee of Yamaguchi University, Yamaguchi, Japan.

2.3.17. Accession numbers.

The nucleotide sequences reported in this study were deposited in the DDBJ, EMBLE, and GenBank databases under accession numbers LC485052 to LC485064, and LC485155.

2.4. Results

2.4.1. Insertional polymorphic distribution of ERV-DC in European wildcats and domestic cats.

The ERV-DC provirus insertions at 13 loci that were previously identified in domestic cats in Japan [**Figure 2.1**A and B, (51)] were investigated in 11 European wildcats in Spain. We detected ERV-DC proviral insertions of ERV-DC6, -DC7, -DC14, and -DC16 at frequencies

of 10%, 100%, 82%, and 18%, respectively, in European wildcats (**Figure 2.1**C). The ERV-DC provirus insertion was also investigated in 35 domestic cats in Spain. Eleven proviruses were detected at frequencies of 9.1–100%, and the insertional patterns were similar to those previously described for Japanese domestic cats (**Figure 2.1**B and D). ERV-DC6 was detected in only one European wildcat (<10%), but it was fixed in domestic cats in Japan and Spain. ERV-DC17 and ERV-DC18 were detected only in Japanese domestic cats. ERV-DC14 was detected at a frequency of 82% (9 cats) in European wildcats; this rate was significantly higher than those in domestic cats in Japan (2.4%) and Spain (11.4%). ERV-DC10 was not detected in European wildcats; it was detected only in domestic cats in Japan and Spain at frequencies of 38% and 24.5%, respectively. ERV-DC7 was fixed (100% frequency) in all cat populations. ERV-DC16 was detected at a frequency of 18% in European wildcats, whereas it was fixed in domestic cats (**Figure 2.1**E).

Next, we performed a PCR analysis to detect ERV-DC in a genotype-specific manner. As shown in **Figure 2.1**F, ERV-DCs of Genotypes I and II were detected in all European wildcats. However, ERV-DC6 (Genotype III) was detected in only one wildcat (No. 54), which was also positive for ERV-DC14 (Genotype I). These results demonstrated that ERV-DC was present in European wildcats and that the ERV-DC insertional polymorphic pattern was quite variable among European wildcats. ERV-DCs of Genotypes I and II were fixed in all wildcats, but ERV-DC of Genotype III had limited spread.

2.4.2. Cloning of ERV-DC14 from European wildcats and analysis of viral replication.

Of the replication-competent proviruses (ERV-DC10, -DC14, and -DC18), ERV-DC14 was the only one detected at a high frequency in European wildcats. Thus, we attempted to isolate the ERV-DC14 provirus to compare the properties of ERV-DC14 between domestic cats and

European wildcats. We successfully amplified two full-length ERV-DC14 proviruses from European wildcats (Nos. 54 and 63) via PCR, then cloned (ERV-DC14/*F.s. silvestris*/wildcat54 and ERV-DC14/*F.s. silvestris*/wildcat63) and determined their direct sequences. Two single nucleotide polymorphism (SNPs) were found between these two clones including G4367A and C4633T (data not shown). In this study, we used the terms ERV-DC14/*F.s. silvestris* and ERV-DC14 to refer to the clones amplified from European wildcat63 and the Japanese domestic cat SO38, respectively. Sequence analysis indicated that intact full-length open reading frames (ORFs) for all ERV-DC14/*F.s. silvestris* genes (*gag*, *pol*, and *env*) were present. We observed a difference of seven nucleotides compared with ERV-DC14 from domestic cats (**Figure 2.2**A). Next, we determined if ERV-DC14/*F.s. silvestris* could replicate in cultured cells by infection of fresh HEK293T cells. As shown in **Figure 2.2**B, all tested ERV-DC14 clones from Japanese domestic cats (SO38, GM21, IK19, and FO16) could infect HEK293T cells, and their viral titers were approximately 10^2 to 10^3 infectious units per mL. In contrast, ERV-DC14/*F.s. silvestris* clones (Nos. 54 and 63) could not infect HEK293T cells. Thus, unlike ERV-DC14 from domestic cats, ERV-DC14/*F.s. silvestris* from European wildcats was replication-incompetent.

2.4.3. Identification of the mutation responsible for replication incompetence of ERV-DC14/F.s. silvestris.

We next investigated potential reasons for the replication incompetence of ERV-DC14/*F.s. silvestris* provirus. Sequence analyses indicate that several nucleotide differences existed in the 5′LTR, *gag*, *pol*, and *env* genes between ERV-DC14 from domestic cats and ERV-DC14/*F.s. silvestris* (**Figure 2.2**A). We constructed four chimeric full-length proviruses (Chimera1, Chimera2, Chimera3, and Chimera4) consisting of ERV-DC14 and ERV-DC14/*F.s. silvestris*

sequences (**Figure 2.2**C). Next, we tested whether these four chimeric proviruses were infectious by infection of fresh HEK293T cells. As shown in **Figure 2.2**D, Chimera2 and Chimera4 exhibited viral infectivity, whereas the other chimeras were not infectious. Thus, the *env* gene that contained two nucleotide differences (nucleotide positions 6735 and 7110) could be responsible for the replication incompetence of ERV-DC14/*F.s. silvestris*.

Next, Mutant1 and Mutant2 were constructed (**Figure 2.2**C and **Figure 2.3**A) and their infectivity was tested in fresh HEK293T cells. Mutant1 showed viral infectivity, but Mutant2 did not (**Figure 2.2**D). This result demonstrated that 148K in ERV-DC14/*F.s. silvestris* Env was critical for infectivity of the proviruses. Notably, the viral titer of Mutant1 was significantly different from that of Chimera4 ($p < 0.0001$). To clarify the mechanism of Env dysfunction, we tested the infectivity of pseudotyped viruses produced from cells transfected with Env expression plasmids for ERV-DC14, ERV-DC14/*F.s. silvestris* (WT), Mutant1 (ERV-DC14/*F.s. silvestris*/K148E), and Mutant2 (ERV-DC14/*F.s. silvestris*/S273P) against fresh HEK293T cells. As shown in **Figure 2.3**B, K148E-Env-pseudotyped virus and ERV-DC14 Env-pseudotyped virus from domestic cats could both efficiently infect cells, and their titers ($10^{3.4}$ and 10^4, respectively) were significantly different ($p < 0.0001$). In contrast, neither ERV-DC14/*F.s. silvestris* nor S273P-Env-pseudotyped viruses were able to infect HEK293T cells. These findings confirmed that 148K in Env was responsible for the replication dysfunction of the ERV-DC14/*F.s. silvestris* provirus.

2.4.4. Mechanism of Env-dysfunction of ERV-DC14/F.s. silvestris.

Using a goat polyclonal anti-FeLV surface glycoprotein (SU) antibody that detects ERV-DC Env, we conducted a western blot analysis in GPLac cells that had been transfected with one of the Env expression vectors (ERV-DC14, ERV-DC14/*F.s. silvestris*, K148E, or S273P). As

shown in **Figure 2.3**C (top left panel), Env proteins were detected in cells transfected with either the ERV-DC14 or the K148E Env expression vectors as multiple bands of approximately 75 kDa and 70 kDa (representing precursor and mature SU protein, respectively). In contrast, cells transfected with ERV-DC14/*F.s. silvestris* and S273P Env expression vector produced only a single 75-kDa band corresponding to the Env protein. Both the ERV-DC14/*F.s. silvestris* and S273P Env expression proviruses were highly expressed in cells.

These findings suggested that the ERV-DC/*F.s. silvestris* Env protein, which consists of the SU and transmembrane (TM), had a cleavage dysfunction. Thus, we looked for antibodies that cross-reacted with the ERV-DC14 TM protein. Among seven monoclonal antibodies against the FeLV TM, we found two suitable monoclonal antibodies (PF6J-2A and EC6-6B1) which also cross-reacted with the FeLV TM (data not shown). Using the anti-FeLV TM antibody, a western blot analysis was conducted, and a TM protein of approximately 17 kDa was detected as a single band in cells transfected with either the ERV-DC14 or K148E Env expression vectors. By contrast, the TM protein was not detected in cells expressing either ERV-DC14/*F.s. silvestris* or S273P Env (**Figure 2.3**C, middle left panel). These results suggested that ERV-DC14/*F.s. silvestris* Env was not cleaved into the SU and TM. Therefore, we again used two antibodies, anti-FeLV SU and anti-FeLV TM, to detect Env proteins in the viral pellets prepared by ultracentrifugation of the supernatants (**Figure 2.3**C, top and middle right panel). Env SU protein (approximately 70 kDa) and Env TM protein (approximately 17 kDa) were detected in the viral pellets of ERV-DC14 and K148E pseudotyped viruses, whereas we failed to detect these proteins or even visible viral pellets of ERV-DC14/*F.s. silvestris* and S273P pseudotyped viruses. Gag proteins were detected using a goat anti-Raucher MLV CA antibody, and detected bands representing precursor Gag (Pr65, approximately 65 kDa) in cell lysates or Gag CA protein (p30, approximately 30 kDa) in cell supernatants (**Figure 2.3**C, bottom panel).

These Env proteins did not participate in the production of infectious viral particles and were not incorporated into virions, even though they were highly expressed in the cultured cells. These results indicate that the dysfunction of ERV-DC14/*F.s. silvestris* was caused by defects in cleavage of the Env protein and that infectious viral particles were not produced from cells exposed to ERV-DC14/*F.s. silvestris* due to the non-functionality of Env.

2.4.5. ERV-DC14/F.s. silvestris Env localizes on the cell surface.

To better understand the mechanism of Env dysfunction, the subcellular localization of ERV-DC14/*F.s. silvestris* Env was investigated by flow cytometry. As shown in **Figure 2.4**A and B, ERV-DC14 and ERV-DC14/*F.s. silvestris* Env-expressing cells had higher signals compared with the mock-transfected cells in both permeabilized (detecting protein in the cytoplasm) and nonpermeabilized (detecting protein on the cell surface) samples. These results suggested that ERV-DC14/*F.s. silvestris* Env is transported to and expressed on the cell surface. To confirm the results of this assay, we detected the serine/threonine-protein kinase, AKT present in the cytoplasm in both permeabilized and nonpermeabilized samples from mock-transfected cells. The permeabilized samples had higher fluorescent signals compared with control (without primary antibody) samples, whereas the nonpermeabilized samples had fluorescent signals similar to those of control samples (**Figure 2.4**C).

2.4.6. Mutations in the ERV-DC14 env gene among European wildcats.

The results above indicated that ERV-DC14 was highly prevalent in European wildcats but not in domestic cats from Japan or Spain (**Figure 2.1**). We next investigated whether or not the ERV-DC14 *env* gene mutations were evolutionally conserved in European wildcats and domestic cats. Sequence analyses showed that all domestic cats in Japan that were positive for

ERV-DC14 (N = 6) displayed the 6735G (148E in Env) and 7110C (273P in Env) polymorphisms in the *env* gene. In contrast, all European wildcats that were positive for ERV-DC14 (N = 9) displayed the G6735A (E148K in Env) and C7110T (P273S in Env) polymorphisms in ERV-DC14 (**Table 2.2**). Three of the four ERV-DC14-positive domestic cats from Spain had the same sequences as Japanese domestic cats in the ERV-DC14 *env* gene, whereas the other ERV-DC14-positive domestic cat in Spain (cat ID. 317) had the same ERV-DC14 *env* mutations as European wildcats. These results suggested that both ERV-DC14 phenotypes are present in cat populations (i.e., an active ERV-DC14 encoding 148E and 273P in Env that is mainly present in the domestic cat population and an inactive ERV-DC14 encoding 148K and 273S in Env that is abundantly distributed in the European wildcat population).

2.4.7. Identification of specific mutations in FeLV and murine ERV corresponding to the ERV-DC14 Env 148K mutation.

To identify specific mutations corresponding to the 148K mutation in the SU N-terminal domain of ERV-DC14 Env, we next analyzed the sequences of gammaretroviruses in different species, including our previous data on the major FeLV strains circulating in Japan (90). As shown in **Figure 2.5**A, we identified two virus sequences bearing this specific mutation. One is the *Mus musculus castaneus* endogenous virus (MLV/MmCN). The other was the FeLV from cat ID KS16, a 5-year-old, neutered male with no history of FeLV vaccination; this animal presented with dyspnea and was diagnosed with thymic lymphoma. Two FeLV *env* variants isolated from peripheral blood mononuclear cell DNA of the Japanese cat KS16 (90) showed single nucleotide changes at position 148 resulting in an E residue in FeLV/KS16-1 Env and a K residue in FeLV/KS16-2 Env (**Figure 2.5**A).

To determine the infectivity of the FeLV/KS16-2 variant, we tested infection of fresh HEK293T cells by pseudotyped viruses produced by transfection of cells with Env expression plasmids for FeLV/KS16-1 (as a positive control), FeLV/KS16-2, and empty vector (mock). The FeLV-A/KS16-2 variant was unable to infect cells, whereas the FeLV/KS16-1 variant successfully infected cells (**Figure 2.5**B), consistent with a previous report (90). The transfected cells were analyzed by western blotting with specific anti-FeLV SU and TM antibodies. Both the mature SU protein (approximately 70 kDa) and TM protein (approximately 17 kDa) were detected in the FeLV/KS16-1 variant. However, only the precursor SU protein was detected in the FeLV/KS16-2 variant (the TM protein was not detected) (**Figure 2.5**C). The precursor Gag (Pr65) protein was detected in all samples. These results revealed that the specific mutation causing Env dysfunction and viral inactivation occurred not only in ERVs but also in exogenous retroviruses.

2.4.8. Mutational analysis of 148E within the SU N-terminal domain of Env conserved among gammaretroviruses.

The 148E residue in ERV-DC14 Env is mainly conserved within gammaretroviruses (**Figure 2.5**A). To determine if this mutation causes Env dysfunction in other gammaretroviruses, we constructed Env expression plasmids for Ampho-MLV (4070A), Friend-MLV (clone57), FeLV-A (FeLV clone33), and FeLV-B (Gardner-Arnstein) bearing the E148K mutation and tested their infectivity. Ampho-MLV, FeLV-A, FeLV-B, or Friend-MLV Env-pseudotyped viruses bearing the E148K mutation were unable to infect fresh HEK293T cells or MDTF cells, whereas wide-type (WT) Ampho-MLV, FeLV-A, FeLV-B, or Friend-MLV Env-pseudotyped viruses successfully infected HEK293T cells and MDTF cells (**Figure 2.6**A). Western blotting with a specific anti-SU antibody detected two bands or a broad band of Env proteins from

GPLac cells transfected with Ampho-MLV, Friend-MLV, FeLV-A, or FeLV-B Env (WT) expression plasmids, whereas only a single band was detected in cells transfected with any of the Env-pseudotyped viruses bearing the E148K mutation (**Figure 2.6**B) even though these Env-pseudotyped viruses were highly expressed in cells. The precursor Gag (Pr65) protein was detected using an anti-MuLV CA antibody in cell lysates as a control.

Virus pellets prepared by ultracentrifugation of the supernatants of transfected cells were analyzed by western blotting with specific anti-SU antibody. Gag CA protein was detected using goat anti-MLV CA antibody in cell supernatants from all samples. Env SU proteins were detected in the viral pellets of Ampho-MLV, Friend-MLV, FeLV-A, and FeLV-B Env-pseudotyped viruses, whereas these proteins were not detected in viral pellets of FeLV-A and FeLV-B pseudotyped viruses bearing the E148K mutation or were detected as only a faint band in viral pellets of Ampho-MLV and Friend-MLV pseudotyped viruses bearing the E148K mutation (**Figure 2.6**B). The mutant Env proteins did not participate in the production of infectious viral particles and were not incorporated into virions, even though they were highly expressed in cultured cells. These results indicated that the E148K mutation caused the defect in Env cleavage in Ampho-MLV, Friend-MLV, FeLV-A, and FeLV-B and that infectious viral particles were not produced from cells transfected with pseudotyped viruses bearing the E148K mutation. Additionally, these findings suggest that the critical amino acid substitution of E148K within the SU N-terminal domain caused the same dysfunctions in other gammaretroviruses as those observed in ERV-DC14/*F.s. silvestris*.

2.4.9. Analysis of Refrex-1 in European wildcats

ERV-DC7 is fixed in both European wildcats and domestic cats, whereas ERV-DC16 is fixed in only domestic cats (**Figure 2.1**E). To determine whether Refrex-1 was evolutionally

conserved in European wildcats, we isolated the full-length ERV-DC7 and ERV-DC16 proviruses, termed ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris*, respectively, from European wildcats. We found that these two proviruses had defective ORFs encoding *gag, pol,* and *env* similar to the ERV-DC7 and ERV-DC16 from domestic cats (64). However, SNPs, deletions or insertions existed comparing the ERV-DC7 and ERV-DC16 sequences of European wildcats and domestic cats (**Figure 2.7**A). Next, the Refrex-1 activities of ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris* were analyzed. The inhibition assay indicated that both ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris* proviruses specifically inhibited the infection of ERV-DC14TA in a dose-dependent manner, showing Refrex-1 activity similar to that of ERV-DC from domestic cats (**Figure 2.7**B).

To confirm if the truncated Env proteins conferred Refrex-1 activity, as reflected by the absence of differences in the Refrex-1 coding region between ERV-DC16 from domestic cat and ERV-DC16/*F.s. silvestris* (**Figure 2.7**A), we performed experiments with the Env-expression vector ERV-DC7/*F.s. silvestris* Env. This expression vector encodes an Env bearing a difference of two amino acids in comparison with ERV-DC7 from domestic cats. The supernatants of cells transfected with the Env-expression plasmids of either ERV-DC7 or ERV-DC7/*F.s. silvestris* also specifically inhibited the infection of ERV-DC14TA (**Figure 2.7**C). Western blotting of those cell lysates was conducted using anti-FeLV SU antibody. Refrex-1 protein was detected as bands of ~28-kDa in size for ERV-DC7 and of ~32-kDa in size for ERV-DC16 from both domestic cat and wildcat sources. Interestingly, the amount of Refrex-1 from ERV-DC7/*F.s. silvestris* was slightly lower than from ERV-DC7 (**Figure 2.7**D). In addition, we also investigated whether Refrex-1 from ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris* proviruses could inhibit FeLV-D infection; the inhibition assay results showed that Refrex-1 expressed from European wildcats can also inhibit the pseudotyped

FeLV-D/TY2.0 virus, and the viral titers decreased from $10^{4.5}$ (Mock) to 10^{2} (data not shown). These results are consistent with those for Refrex-1 from domestic cats (64). Our data suggest that truncated Env from ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris* both encode Refrex-1.

2.4.10. Sequence diversity of ERV-DC between domestic cats and wildcats.

We previously ascertained the sequence diversity of ERV-DC7 *env* in Japanese domestic cats (23). In the present study, we investigated the sequence diversity of ERV-DC7 *env* in both European wildcats and domestic cats. The phylogenetic tree constructed from ERV-DC7 *env* sequences shows the evolutionary diversity of this provirus in each cat population (**Figure 2.8**A). We identified a total of 16 alleles of ERV-DC7 *env*, including nine alleles that were newly identified in this study. All of the alleles had the same stop codon mutation in the middle of *env* and encoded a truncated *env* as Refrex-1. Among the 30 SNPs observed in European wildcats and Spanish domestic cats, two new SNPs were identified. These two SNPs (nucleotide positions 1555 and 1782) caused nonsynonymous substitutions in ERV-DC7 *env* via deletion of a stop codon mutation. The nucleotide sequences of the ERV-DC7/*F.s. silvestris* *env* gene were mostly conserved among European wildcats, with the exception of one European wildcat (ID: European wildcat55) which showed four nucleotide differences compared with the other European wildcats. The genetic diversity of the ERV-DC7 *env* sequences suggested that wildcats form a genetically different clade from that of domestic cats. Additionally, based on analyses of the amino acid sequence of defective ERV-DC7fl *env* at positions 407, 427 and 429, we found that the combination (407G, 427I and 429T) was conserved in European wildcats. In contrast, the six combination variants (R–IT, G–IT, G–IA, R–IA, G–NA and R–NA) only existed in domestic cats (**Figure 2.8**A).

Because LTR sequences may be involved in the integration and transcription of viral DNA, we were interested in whether ERV-DC14 LTRs had evolved in different cat populations. The 5′-LTRs of ERV-DC14 from different cat populations were amplified and determined their sequences. **Figure 2.8**B shows the phylogenetic tree constructed based on the ERV-DC14 5′-LTRs among different cat populations. These data showed that, based on 5′-LTR sequences of ERV-DC14 proviruses, two cat populations exist: wildcats and domestic cats. Sequence analyses indicated that the 5′-LTR of ERV-DC14 is mainly conserved between these two cat populations. Only one nucleotide difference was identified between wildcats (T) and domestic cats (C) (**Figure 2.8**B). Interestingly, one wild cat (cat 58) displayed a mutation in the TATA box (TATA to CATA), and this mutation was also observed in the replication-competent ERV-DC10 and ERV-DC18 proviruses (**Figure 2.8**B). In addition, we also analyzed sequence diversity of the 3′-LTRs of ERV-DC14 in European wildcats (N=2) and Japanese domestic cats (N=4). The 3′-LTR of ERV-DC14 was conserved between these two cat populations (data not shown).

2.5. Discussions

Here, we identified ERV-DC/*F. s. silvestris* in the European wildcat (*F.s. silvestris*), which is evolutionally close to the domestic cat. Our study detected four ERV-DC loci (ERV-DC14, -DC6, -DC7, and -DC16) in wildcats and 11 ERV-DC loci in domestic cats in Spain (**Figure 2.1**). Several ERV-DC loci were less prevalent in European wildcats than in domestic cats (**Figure 2.1**). Although two European wildcats (Nos. 52 and 53) were negative for ERV-DC14 (Genotype I) and showed no evidence of any other known Genotype I ERV-DC, genotype-specific PCR still amplified a fragment of Genotype I ERV-DC. This finding may suggest the existence of an unknown Genotype I ERV-DC locus in European wildcats (**Figure 2.1**F).

Furthermore, genotype-specific PCR also suggested that Genotype I and Genotype II ERV-DCs have invaded European wildcats, whereas Genotype III ERV-DCs have not. As our previous results demonstrated, Genotype I and Genotype II ERV-DCs use the same receptor for viral entry, while Genotype III ERV-DC uses a distinct receptor (23). Genotype I and Genotype II ERV-DCs might have infected the ancestors of wild and domestic cats, whereas Genotype III ERV-DC more recently infected domestic cats. The phenomenon of retrotransposition of ERV-DC18, which belong to Genotype III, was only observed in Japanese domestic cats and not in European wildcats or European domestic cats (**Figure 2.1**C and D). This result suggests that this phenomenon may be specific to one cat family. ERV-DC14 was found to have significantly invaded in European wildcats, where it was detected at higher prevalence (80%) compared with domestic cats (2.4% and 11.4% in domestic cats from Japan and Spain, respectively) (**Figure 2.1**). This result may be explained by the existence of two different ERV-DC14 phenotypes.

The identification and characterization of replication-defective ERV-DC14/*F.s. silvestris* indicates that a feline ancestor with an E148K mutation in ERV-DC14 Env probably gave rise to European wildcats, whereas a cat ancestor with a replication-competent version of ERV-DC14 probably gave rise to domestic cats. **Figure 2.1** shows that the integration pattern of ERV-DC in wildcats was distinct from those of European and Japanese domestic cats. One of four European domestic cats (cat ID 317) had a replication-defective ERV-DC14 (**Table 2.2**); the ERV-DC14 locus found in this animal could be derived from European wildcats due to the high potential for interbreeding (69, 128, 129). The 5′-LTR sequence of ERV-DC14 from this cat showed a single nucleotide polymorphism (C→T), which was observed in ERV-DC14/*F.s. silvestris*, further supporting this hypothesis (**Figure 2.8**B). These results also suggest that the

ERV-DC14 genotype is conserved among all cats, but its phenotype differs between wildcats and domestic cats.

The replication-defective ERV-DC14/*F.s. silvestris* provirus in European wildcats was found to result from a single G6735A mutation in *env* that resulted in an E148K substitution in ERV-DC14 Env. The other mutation (C7110T) in *env*, resulting in a P273S amino acid substitution (**Table 2.2**), also might cooperatively contribute to inactivation of the ERV-DC14/*F.s. silvestris* provirus (**Figure 2.2**D). This lethal mutation may have been caused by apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3 proteins, which are mammalian-specific cellular deaminases (130, 131) that play important roles as viral restriction factors (132). Interestingly, the same mutation observed in FeLV/KS16-2 and MLV/MmCN is also due to a G→A substitution, indicating that this phenomenon occurred in ERV-DC14/*F.s silvestris*, MLV/MmCN and FeLV/KS16-2, despite two of these retroviruses being endogenous and the other being exogenous. A reverse genetics technique revealed that a common mutation within the SU domain (E148K) resulted in defective Env cleavage and Env incorporation into virions. The 148E at Env was found to be relatively conserved among gammaretroviruses (**Figure 2.6**), and the substitution of this residue (E148K) in FeLV-A, FeLV-B, Ampho-MLV, and Friend-MLV caused the same dysfunctions of Env. This result suggests that the three-dimensional structure and/or folding pathway of the N-terminal SU is generally conserved among gammaretroviruses. Furthermore, our findings also indicate that any exogenous retrovirus, like FeLV-D, that is newly generated in wildcats would potentially be inactivated by this mechanism, although specific ERV-DC loci for FeLV-D generation have not yet been identified. In this study, the same mutation occurring in the SU region of MLV/MmCN resulted in a defective endogenous MLV encoding a truncated Env which functioned as a known restriction factor (*Fv-4*) in *Mus musculus castaneus* (117). The *Fv-4 env*-encoded protein is

processed normally and can be incorporated into viral particles but is unable to produce infectious virus (28, 29). Conversely, Refrex-1 is a soluble anti-retroviral factor and is not incorporated into viral particles (64).

The Env E148K substitution is close to the N-terminus of the SU and is distant from the SU–TM cleavage motif recognized by furin. One possible mechanism through which the E148K mutation causes Env dysfunction is by altering the protein's structure in a way that prevents furin from accessing Env. Flow cytometry results indicated that ERV-DC14/*F.s. silvestris* provirus affected by this conformational change was unable to be incorporated into viral particles, even though Env was transported to the cell surface (**Figure 2.4**); this mechanism is similar to that of Env dysfunction in HERV-K (27), ERV-DC7fl, and ERV-DC16fl (23).

The ERV-DC7 and ERV-DC16 loci are fixed in domestic cats, and the presence of Refrex-1 is evidence that ERVs are involved in coevolution between host and virus (64). Our study found that only ERV-DC7/*F.s. silvestris* was fixed in all European wildcats, while ERV-DC16/*F.s. silvestris* was polymorphic in European wildcats (**Figure 2.1**C). We demonstrated that ERV-DC7/*F.s. silvestris* and ERV-DC16/*F.s. silvestris* from European wildcats also encoded the anti-retroviral factor Refrex-1 (**Figure 2.7**B). The results of our work suggest that Refrex-1 has had an antiviral role throughout *Felis* evolution, pre-dating the exposure of cats to feline retroviruses. European wildcats still maintain Refrex-1 in the population to avoid the threat of infection by Genotype I ERV-DCs and FeLV subgroup D. One interesting observation was that the amount of Refrex-1 in ERV-DC7/*F.s. silvestris* was slightly lower than in ERV-DC7 (**Figure 2.7**D), which may be due to genetic diversity of ERV-DC7s between domestic cats and wildcats.

Many intact ERVs are targeted for transcriptional silencing by modification or mutation of their LTRs, and species-specific ERV LTRs play important roles in regulatory effects during

development by modulating the transcriptome (40, 42, 133-135). Here, ERV-DC14/*F.s. silvestris* probably still had a low level of promoter activity due to a mutation (A→T) in the 5'-LTR, similar to that of ERV-DC14 in domestic cats (42). As shown in **Figure 2.8**B, ERV-DC14/*F.s. silvestris* from cat No.58 had a mutation in the TATA box (TATA→CATA), a sequence which was observed in the replication-competent ERV-DC10 and ERV-DC18 proviruses, suggesting that the LTR was probably not inactivated in this animal.

ERVs are known to have different fates in the host germline following endogenization (1). ERVs are mainly inactive, acting much like fossil records in both vertebrate and invertebrate animals. Thus, there are very few infectious endogenous retroviruses in mammals (36, 45, 50, 136-138). Replication-competent ERV-DC14 in the domestic cat genome may be mobile and interact with other exogenous retroviruses to generate new recombinant viruses. Any replication-competent retrovirus in the host genome poses a potential risk to the host. Thus, the host will exert negative selection to eliminate deleterious endogenous elements during host–virus coevolution. Our data indicate that inactivated ERV-DC14/*F.s. silvestris* is potentially being selected during the evolutionary process of European wildcats, whereas this EV-DC locus has not yet been inactivated in domestic cats. Although ERV-DC14/*F.s. silvestris* is inactivated, infectious ERV could be generated by recombination events as reported in mice (103). Moreover, the promoter activity of ERV-DC14 in domestic cats is also low due to a mutation in the 5′-LTR (42). In contrast, the frequency of inactive ERV-DC14/*F.s. silvestris* in European wildcats was high (82%) (**Figure 2.1**C), so the presence of ERV-DC14 in the wildcat genome may play an important role for the host.

Here, a broad sequence analysis of ERV-DC7 from European wildcat and domestic cat populations showed different sequence combinations, and the determinants of ERV-DC7fl *env* dysfunction were found to consist of six combination variants (G-IT, R-IT, G-IA, R-IA, G-NA,

and R-NA), represented in **Figure 2.8**A by different panel patterns. Interestingly, we found that ERV-DC7 *env* was highly conserved in the wildcat population, with only a single combination (G-IT) observed in these animals. In contrast, six combination variants of ERV-DC7 were observed in domestic cats. This result suggests that ERVs in general, and ERV-DC in particular, genetically diverged during the process of domestication. Sequence diversity of ERV-DCs depends on the locus and the cat subspecies; in particular, ancestral ERV-DC7 may be conserved in European wildcat populations (**Figure 2.8**). As previously described, a reconstructed ERV-DC7/*F.s. silvestris* full-length *env* did not contain a latent stop codon (23). Nucleotide position 801 was a conserved C in all European wildcats, whereas in domestic cats, a T nucleotide existed in all cats examined but one (European *F.s. catus* 45) (**Figure 2.8**). This result concurred with our previous study (23). This latent stop codon suggests that ERV-DC7fl may evolve to a different prototype in domestic cat populations during endogenization.

2.6. Conclusions

In conclusion, this study revealed the existence of ERV-DC in the European wildcat population. Notably, this is the first report showing that an infectious endogenous retrovirus is inactivated in a feline species-specific manner. The strategy of tracking infectious ERVs that have invaded cat lineages can reveal the different fates of ERVs and uncover new properties of retroviruses. Our study may contribute to the understanding of evolution and domestication of cat lineages. Furthermore, our findings provide insights into retroviral pathogenesis and virus-host interactions.

2.7. Tables and Figures in Chapter two

TABLES

Table 2.1 Sequences of primers used in this study

Target	Primers		Sequence (5'-3')
ERV-DC14/*F.s. silvestris* provirus	Forward	Fe-603S	AGTTAAGGGACTGTGGACTT
	Reverse	Fe-587R	GCTGGGCATTGTTCTCCTTT
pFUΔss DC14 env	Forward	Fe-610S	GGATCCGGATCCCACCATGAAACCCCCAGCGGGAAT
	Reverse	Fe-184R	GAATTCGAATTCTATTCGATTGTATCTGGCCTTT
ERV-DC14/*F.s. silvestris*/K148E	Forward	Fe-619S	GCATCGCCAGTAGTCTCGCAGCCCCATGCCG
	Reverse	Fe-597R	CGGCATGGGGCTGCGAGACTACTGGCGATGC
ERV-DC14/*F.s. silvestris*/S273P	Forward	Fe-620S	GGTGTCTGAGTTGGGGCCCTTGGTGGC
	Reverse	Fe-598R	GCCACCAAGGGCCCCAACTCAGACACC
pFUΔss Ampho-MLV env mutant	Forward	Fe-636S	GGTAAATGGGGGTGTAAAACCACCGGACAGG
	Reverse	Fe-614R	CCTGTCCGGTGGTTTTACACCCCCATTTACC
pFUΔss Friend-MLV env mutant	Forward	Fe-637S	CCTCTTGGGGCTGCAAGACAACCGGTAGA
	Reverse	Fe-615R	TCTACCGGTTGTCTTGCAGCCCCAAGAGG
pFUΔss FeLV-A env mutant	Forward	Fe-638S	GCTGCATGGGGTTGCAAAACTACGGGAGAAG
	Reverse	Fe-616R	CTTCTCCCGTAGTTTTGCAACCCCATGCAGC
pFUΔss FeLV-B env mutant	Forward	Fe-639S	CTGTATGGGGTTGCAAGACCACCGGGGAA
	Reverse	Fe-617R	TTCCCCGGTGGTCTTGCAACCCCATACAG
ERV-DC14/pol-3'fla	Forward	Fe-501S	AAGGAATTGCCAAAGGAGTTCTAA
	Reverse	Fe-587R	GCTGGGCATTGTTCTCCTTT
pFUΔss DC7 env	Forward	Fe-650S	GGATCCGGATCCCACCATGAAACCCCCAACAGGAAT
	Reverse	Fe-626R	GCGCGAATTCGAATTCTCATTCGATTGTATCTGGCC
ERV-DC14 gag	Reverse	Fe-140R	ACAAAACATAGAACACAATACC
ERV-DC7/*F.s. silvestris* provirus	Forward	Fe-66S	CCGAAAAMTTCCTGACTGTTTAAGA
	Reverse	Fe-53R	AGAGGAAATAAACCGGGTAGTGTGT

ERV-DC16/*F.s. silvestris* provirus	Forward	Fe-219S	GCCACGGTCATGAAAATAAAAA
	Reverse	Fe-44R	TGCAGACAGAACATACTGTGACAAA

Table 2.2 Properties of the ERV-DC14 provirus in different cat populations[a]

Cat	Sample size	ERV-DC14 (+)	ERV-DC14 env gene	
			6735A/148K and 7110T/273S	6735G/148E and 7110C/273P
F.s. silvestris	11	9	9	0
European *F.s. catus*	35	4	1	3
Japanese *F.s. catus*	247	6	0	6

[a](+), ERV-DC14 genotyping positive. 6735 and 7710, nucleotide position of ERV-DC14. 148 and 273, amino acid position of Env. A, G, T, and C are nucleotides. K, E, S, and P are corresponding amino acids.

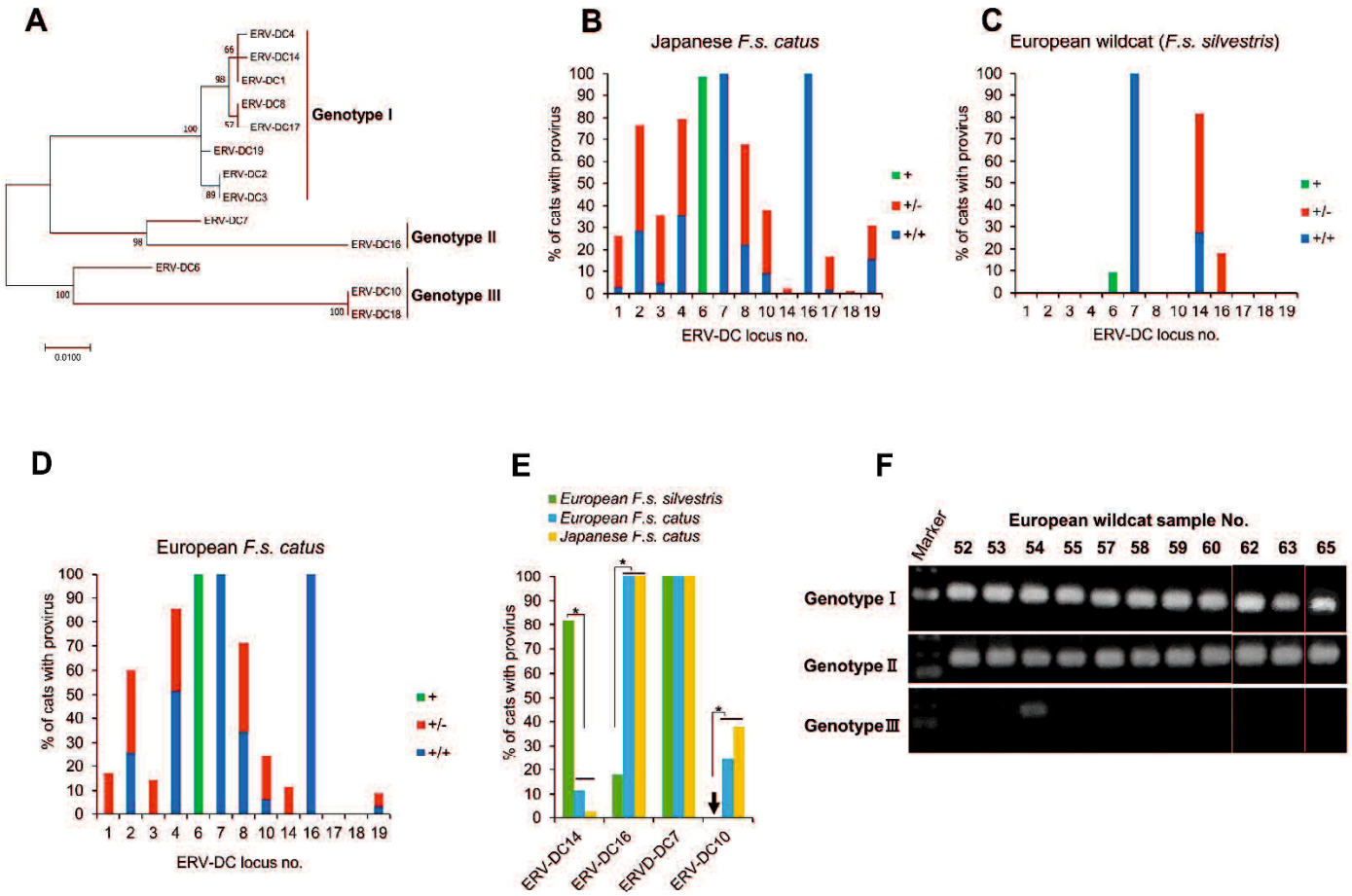

Figure 2.1 Detection of ERV-DC proviruses in domestic cat and wildcat genomes. (A) A phylogenetic tree of ERV-DC 3'LTR was constructed using maximum-likelihood methods. The percentages at branch junctions indicate bootstrap values (1,000 replicates). Thirteen ERV-DC loci were classified into three genotypes: Genotype I (ERV-DC1, DC2, -DC3, -DC4, -DC8, -DC14, -DC17, and -DC19), Genotype II (ERV-DC7, and -DC16), and Genotype III (ERV-DC6, -DC10 and -DC18). (B–D) Insertional polymorphisms of 13 ERV-DCs in Japanese domestic cats (B), European wildcats (C), and European domestic cats (D). Green and +, provirus detected; red and +/-, heterozygous (copy present on one of two chromosomes); blue and +/+, homozygous (copy present on both chromosomes). (E) Comparison of genotype frequencies for three loci (ERV-DC14, -DC16, and -DC7) among different cat populations. (F)

PCR detection of ERV-DC genotypes in European wildcats (*F.s. silvestris*). Statistical analyses were conducted using Student's *t*-tests and one-way ANOVAs. *$p < 0.0001$.

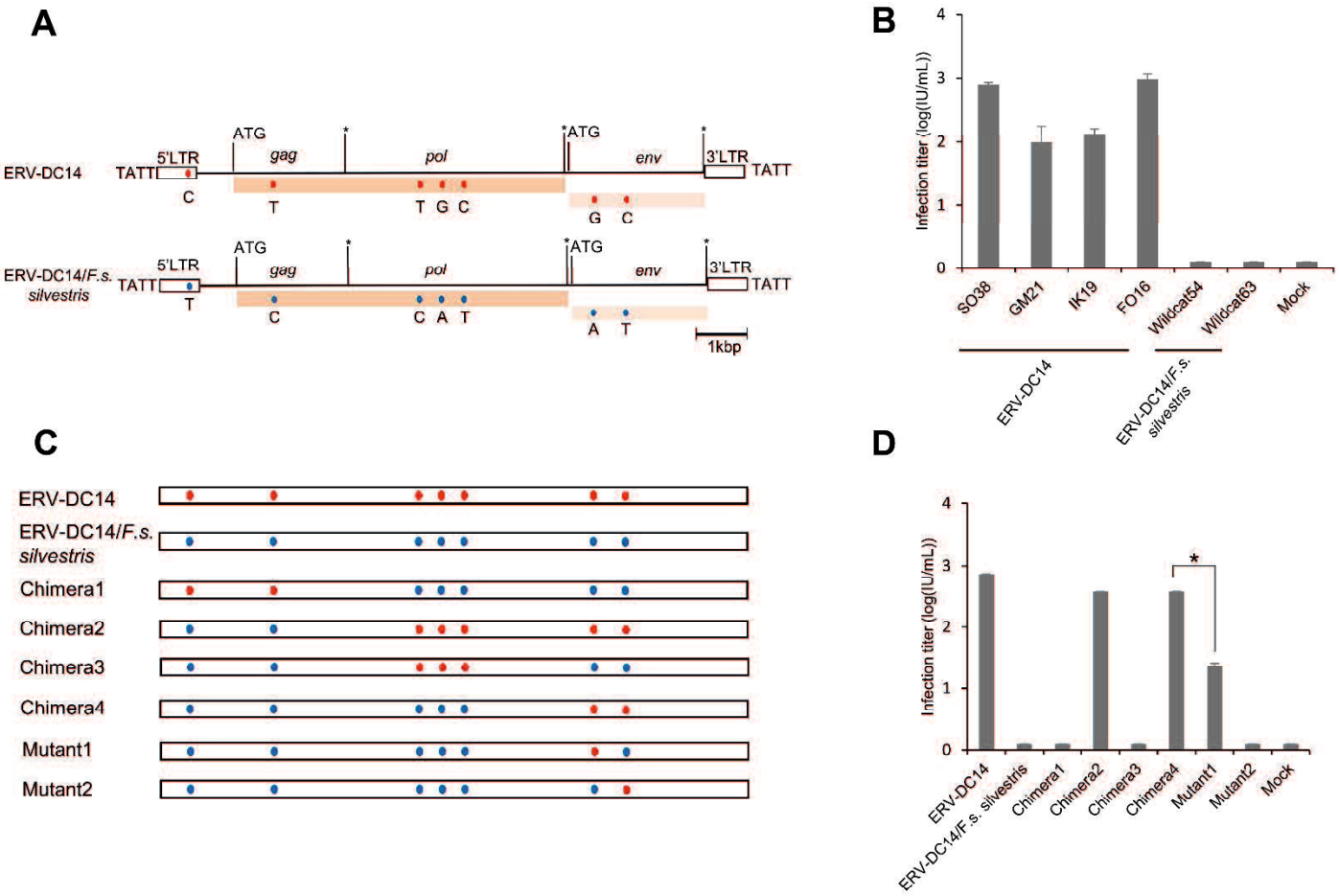

Figure 2.2 Characterization and assessment of ERV-DC14/*F.s. silvestris*. (A) Schematic image of the full-length ERV-DC14/*F.s. silvestris* provirus. The ERV-DC14 clone SO38 strain was used as the ERV-DC14 reference genome. The *gag*, *pol*, and *env* genes are illustrated together with the 5′- and 3′-LTRs and positions of the *gag* and *env* translational initiation codons (ATG). Asterisks, stop codons; dark pink box, open reading frame (ORF) of the Gag–Pol polyprotein; light pink box, Env protein; blue and red round circles indicate single nucleotide polymorphisms (SNPs) between the two proviruses ERV-DC14/SO38 and ERVDC14/*F.s. silvestris*. Nucleotide substitutions are shown. Flanking 4-bp target duplicate site (TSD) sequences are shown for each provirus. (B) Assessment of replication competent activity of ERV-DC14 in European wildcats (*F.s. silvestris*) and Japan domestic cats. All tested proviral clones including ERV-DC14 from different Japanese domestic cats (SO38, GM21, IK19,

FO16), ERV-DC14/*F.s. silvestris* (wildcat54, and wildcat63), or the empty vector (mock) were transfected with 293Lac cells and then tested their infectivity with fresh HEK293T cells. The viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations. (C) Schematic representation of the chimeric structures of the two proviruses. Chimeras1–4 were constructed via recombination between the two proviruses using restriction enzyme digestion, and the two mutants were constructed by site-directed mutagenesis. Blue and red round circles indicate single nucleotide polymorphisms (SNPs) between the two proviruses ERV-DC14/SO38 and ERVDC14/*F.s. silvestris*. (D) Assessment of the replication competence of chimeric ERV-DC14. 293Lac cells were transfected with plasmids containing different chimeric ERV-DC14 or mock (empty vector) transfected, and the resulting supernatants were collected and used to infect fresh HEK293T cells. The viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations. *$p < 0.0001$ (one-way ANOVA).

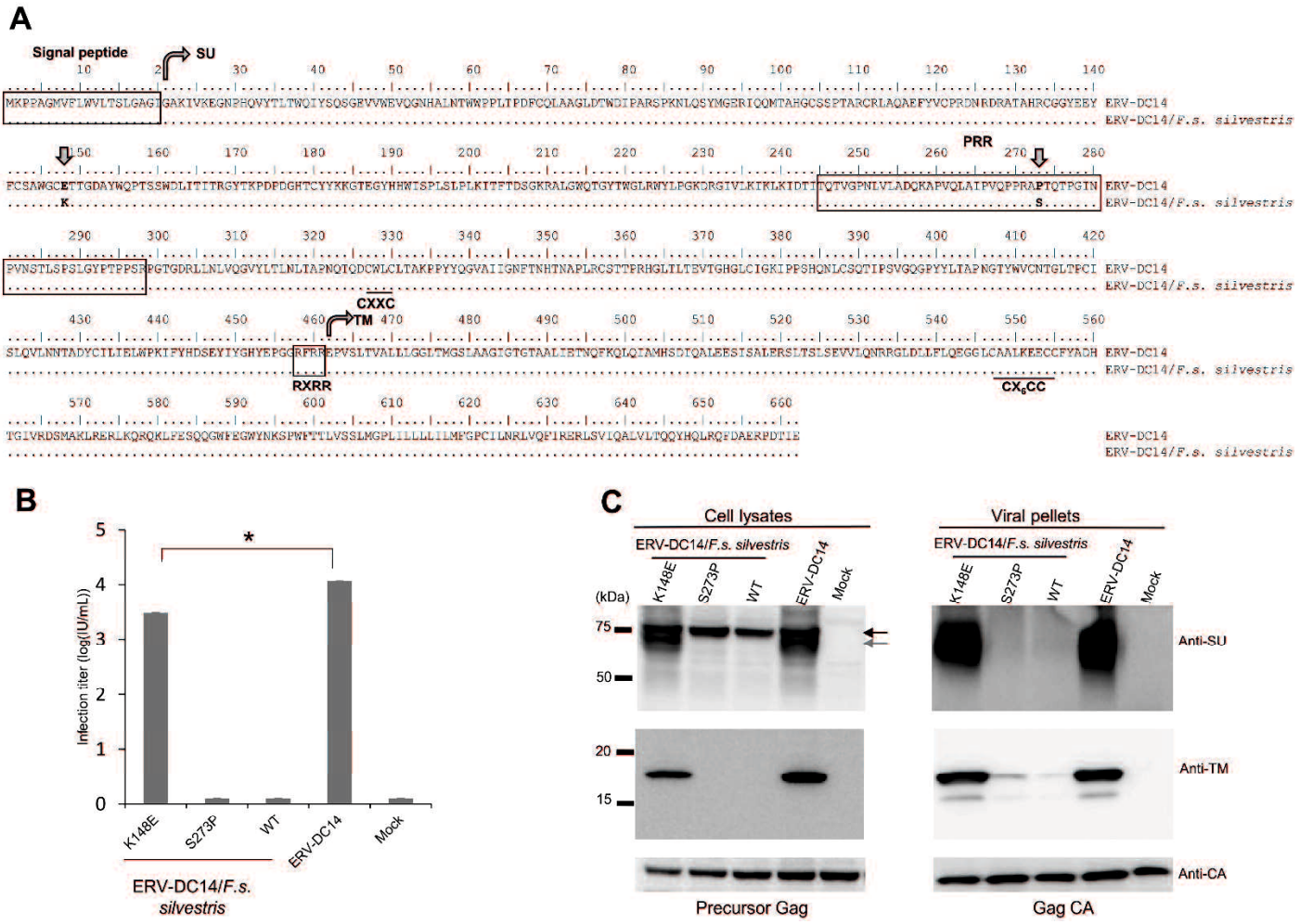

Figure 2.3 Determination of the mutation responsible for ERV-DC14/*F.s. silvestris* Env dysfunction. (A) Amino acid sequence alignment of Env proteins of ERV-DC14 and ERV-DC4/*F.s. silvestris.* SU, surface subunit; PRR, proline-rich region; TM, transmembrane subunit. RXRR is the cleavage motif. CXXC and CX_6CC are sites of covalent interaction. Arrows indicate the positions of amino acids 148 and 273, which differ between these two ERV-DC14 proviruses. (B) Assessment of Env-pseudotyped viruses based on ERV-DC14/*F.s. silvestris* wildtype (WT), mutant1 (K148E), and mutant2 (S273P) or on ERV-DC14. GPLac cells were transfected with indicated Env expression plasmids. The corresponding Env-pseudotyped viruses were used to infect fresh HEK293T cells. The viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations. *$p < 0.0001$ (one-way ANOVA). (C)

Western blotting analysis of GPLac cells expressing ERV-DC14/*F.s. silvestris* Env (K148E, S273P, or WT) or ERV-DC14 Env. The cell lysates and viral pellets from culture supernatants were analyzed. A goat polyclonal anti-FeLV SU (gp70) antibody was used to detect ERV-DC14 SU, and a mouse monoclonal anti-FeLV TM (p15E) was used to detect ERV-DC14 TM. Black arrow indicates immature SU; gray arrow indicates mature SU. Precursor Gag (Pr65) and Gag CA (p30) were both detected with a goat anti-Raucher MLV CA antibody.

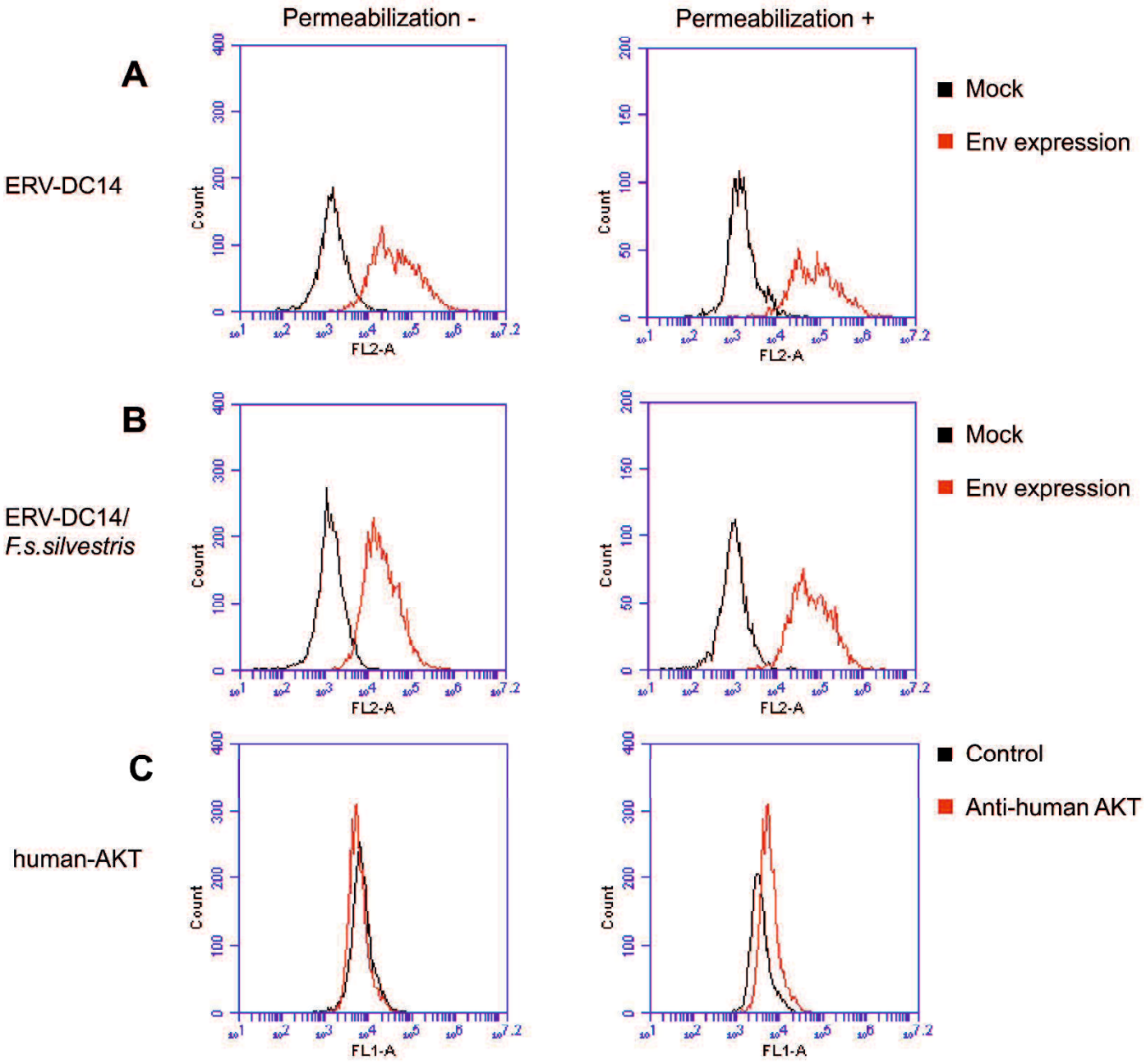

Figure 2.4 Flow cytometry analysis of ERV-DC14/*F.s. silvestris* cell surface expression. (A–B) Detection of Env on the interior and exterior of cells. HEK293T cells expressing ERV-DC14 Env (A) or ERV-DC14/*F.s. silvestris* Env (B) were permeabilized with 0.2% Triton X-100 in PBS (right panel) or were not permeabilized (left panel), and intracellular (right panel) and cell surface (left panel) Env proteins were stained with goat anti-FeLV SU (gp70) and phycoerythrin (PE)-conjugated anti-goat IgG antibody. Fluorescent signals were detected using

the FL-2 channel of a flow cytometer. Histograms of Env-expressing cells (red) and mock-transfected cells (black) are overlaid in each graph. The x-axis shows the signal intensity in FL-2; the y-axis shows the cell counts. (C) Staining of human AKT in permeabilized (right panel) or non-permeabilized (left panel) samples of mock-transfected cells using rabbit anti-human AKT and fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit IgG antibodies. Histograms of cells treated with anti-human-AKT antibody (red) and without antibody (black) are overlaid in each graph. The x-axis shows the signal intensity in FL-1; the y-axis shows the cell counts.

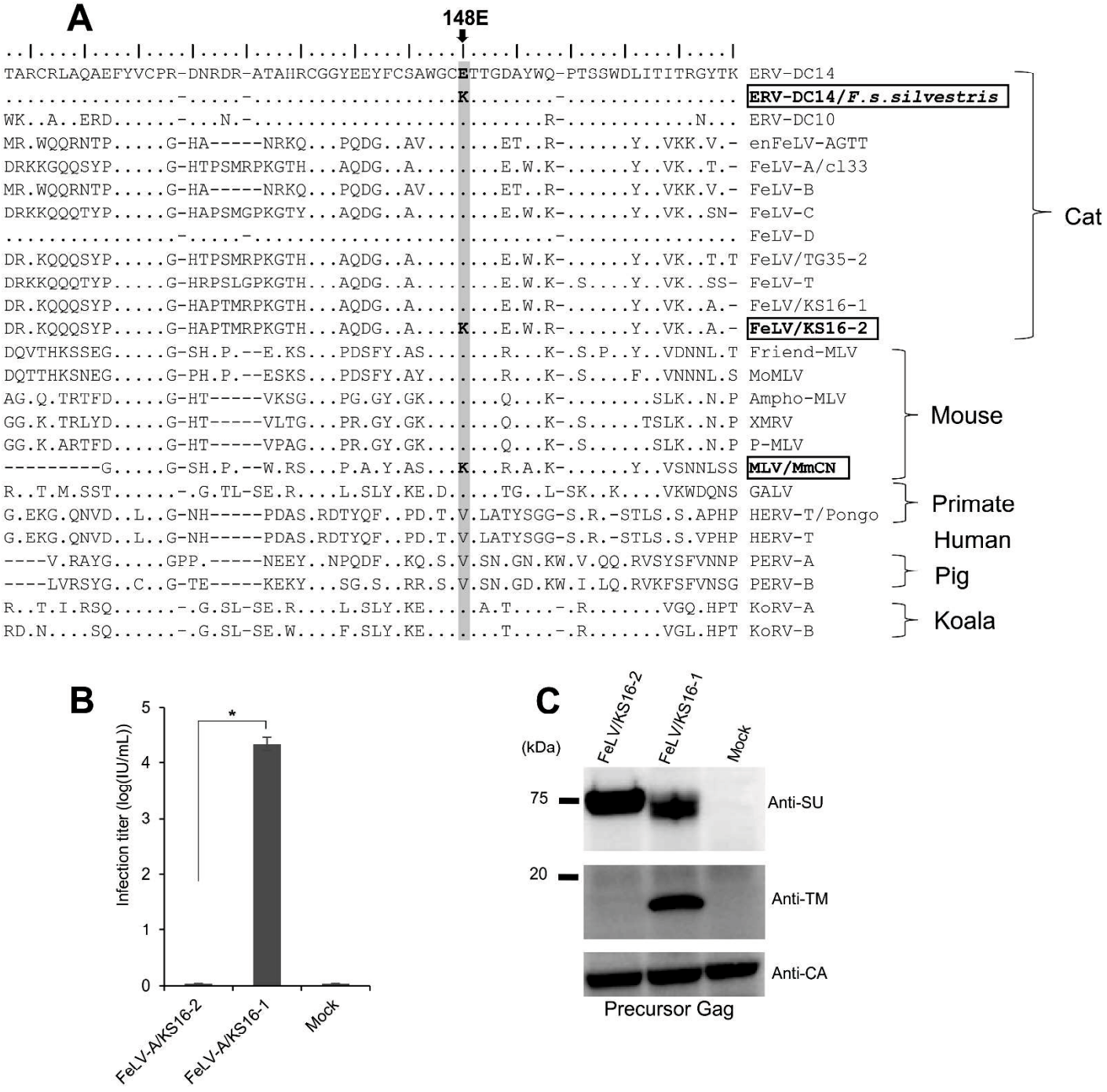

Figure 2.5 Sequence analysis of gammaretrovirus Env proteins and dysfunction of a FeLV variant induced by a single mutation in the SU N-terminal domain. (A) A sequence alignment of Env in gammaretroviruses, constructed by mafft (126). The amino acid sequences of the gammaretrovirus SU N-terminal regions are also presented. The critical amino acid position 148E is shaded in light gray. GenBank accession numbers of reference sequences are as

follows: ERV-DC14 (BAM33599.1), ERV-DC10 (BAM33597.1), enFeLV-AGTT (AY364318.1), FeLV-A (BAB63924.2), FeLV-B (AAA43052.1), FeLV-C (AAA43049.1), FeLV-D (BAM33588.1), FeLV/TG35-2 (BAU61794.1), FeLV-T (AAA43050.1), FeLV/KS16-1 (BAK41670.2), Friend-MLV (AAA46480.1), MoMLV (NP_057935.1), Ampho-MLV (AAA46515.1), XMRV (ADU55755.1), P-MLV (ARB03464.1), MLV/MmCN (AMK06448.1), GALV (AAA46811.1), HERV-T/Pongo (CAI15393.1), HERV-T (XP_011526770.1), PERV-A (AAQ83899.1), KoRV-A (BAM67147.1), KoRV-B (AGO86848.1), ERV-DC14/*F.s. silvestris* (BBL19108.1), and FeLV/KS16-2 (BBL19109.1) . (B) Infection assay using pseudotyped viruses of two FeLV variants (KS16-1 and KS16-2). GPLac cells were transfected with Env expression plasmids for FeLV/KS16-1 and FeLV/KS16-2. The filtered viral supernatants were used to infect fresh HEK293T cells. The viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations; **p* < 0.0001 (one-way ANOVA). (C) Immunoblotting analysis using cell lysates from GPLac cells transfected with the Env expression plasmids which are presented in panel B. Env proteins were detected by mouse anti-FeLV SU (gp70) antibody and anti-FeLV TM (p15E) antibody. Precursor Gag (Pr65) was detected with a goat anti-Raucher MLV CA antibody. The filter exposure time differed for each antibody.

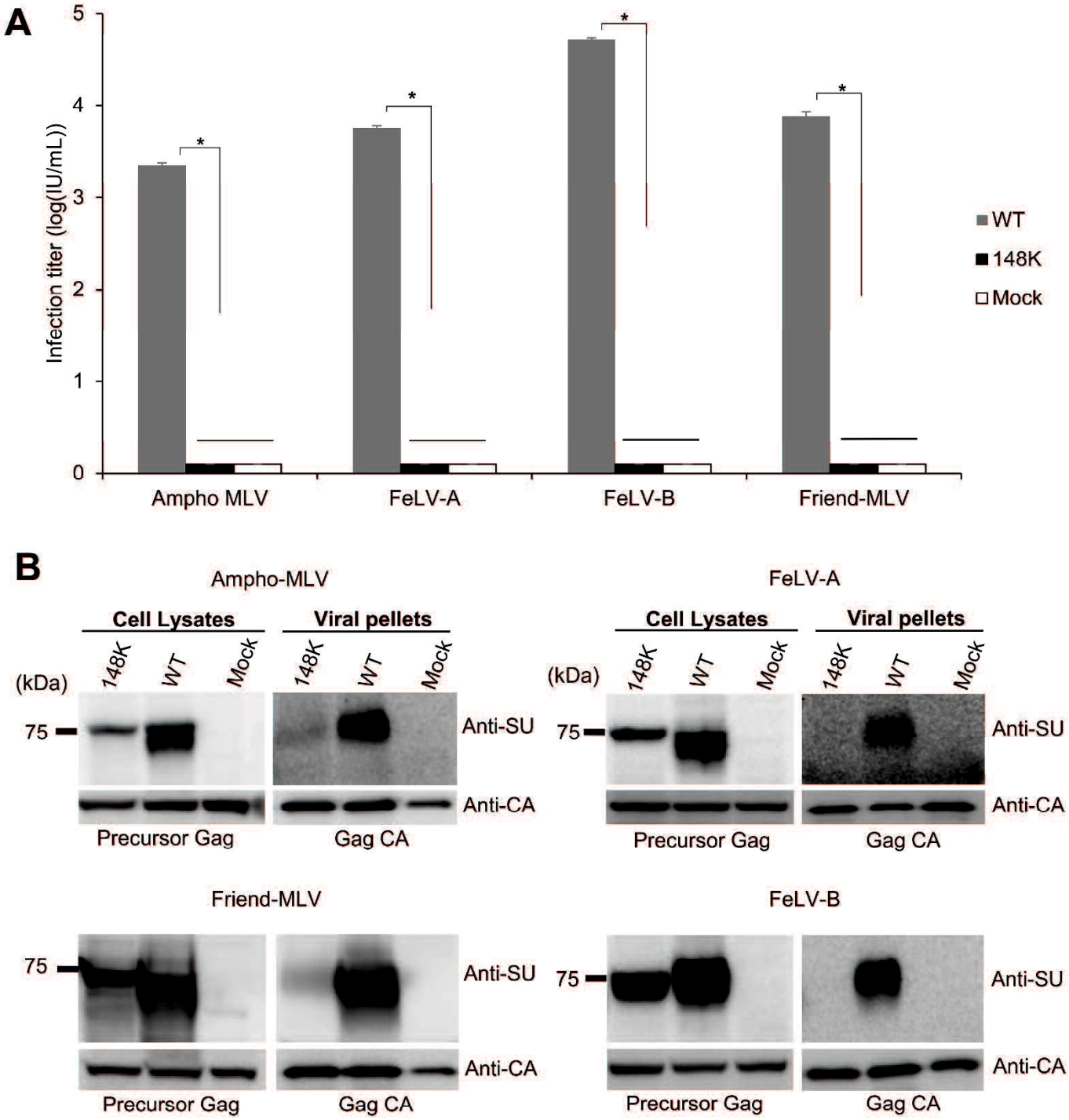

Figure 2.6 Dysfunction caused by mutations within the SU N-terminal domain in gammaretroviruses. (A) Infection assay using pseudotyped viruses of WT or mutant Env (E148K) of FeLV-A/clone33, FeLV-B/GA, Ampho-MLV/4070A, and Friend-MLV/clone 57. Expression plasmids were constructed for Env mutants and expressed in GPLac cells. After 72

h, the cell lysates and viral pellets were harvested from the culture supernatants. Fresh MDTF cells were inoculated with viral supernatants of Friend-MLV and fresh HEK293T cells were inoculated with the viral supernatants of the other pseudotyped viruses. After 48 h, the X-Gal-positive cells were counted and the viral titers are illustrated as the log of infectious units (IU) per mL with standard deviations; $*p < 0.0001$ (one-way ANOVA). (B) Immunoblotting analysis of cell lysates (left) and viral pellets (right) from GPLac cells expressing wildtype (WT) Env or Env mutants. FeLV-A and FeLV-B Env were detected with an anti-FeLV SU (gp70) antibody, Ampho-MLV Env was detected with an anti-Ampho-MLV SU (gp70) antibody, and Friend-MLV Env was detected with an anti-MLV SU (gp70) antibody. Precursor Gag (Pr65) and Gag CA (p30) were both detected with a goat anti-Raucher MLV CA antibody. The filter exposure times differed between the cell lysates and viral pellets.

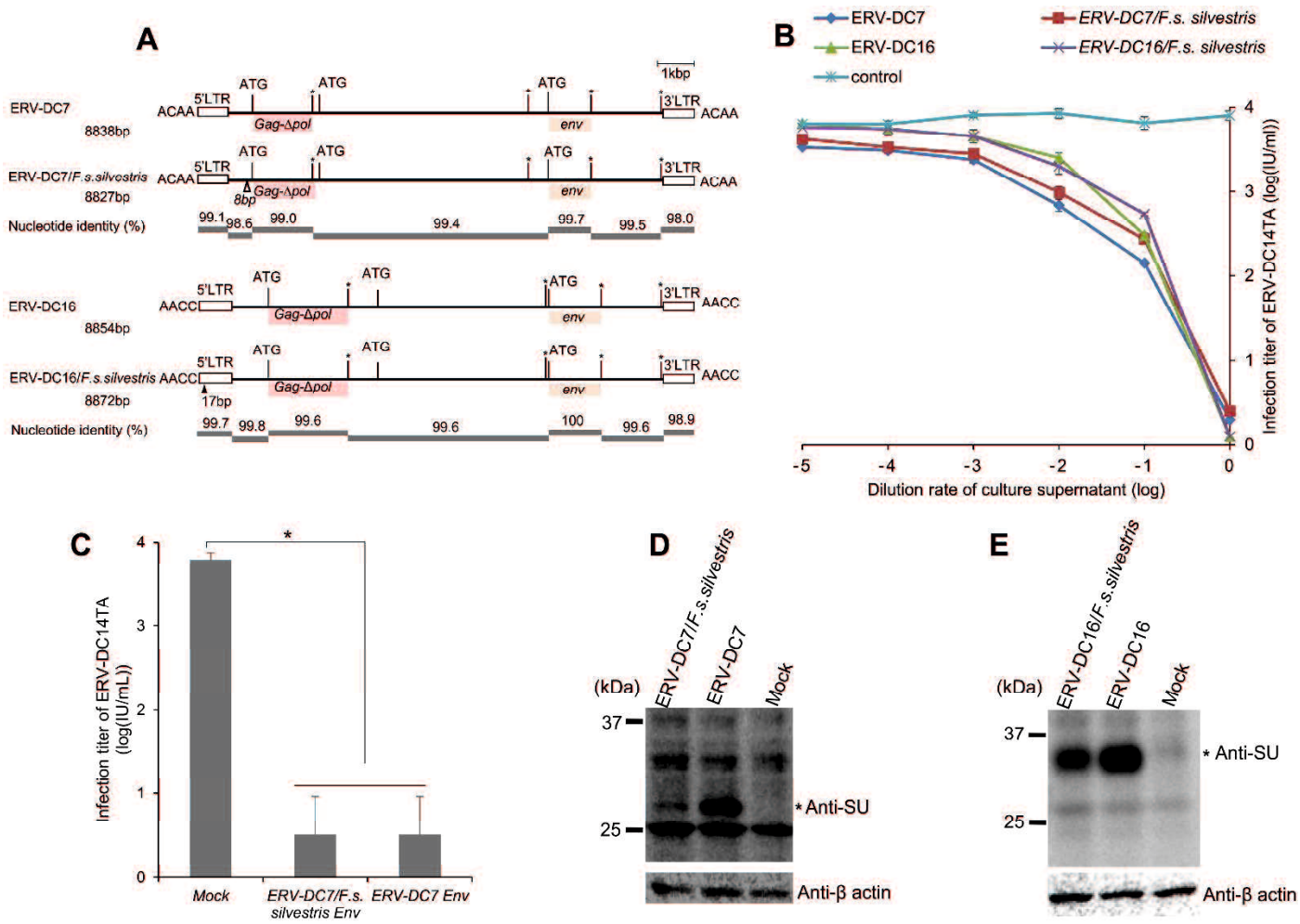

Figure 2.7 Analysis of Refrex-1 in European wildcats. (A) Schematic structure of ERV-DC7 and ERV-DC16 in domestic cat and European wildcats. The *gag*, deletion of *pol* (Δ*pol*), and *env* genes are illustrated together with the 5′- and 3′-LTRs and the positions of the *gag* and *env* translational initiation codons (ATG). Asterisks, stop codons; dark pink boxes indicate predicted Gag and light pink boxes indicate truncated Env protein; black triangles indicate insertions; and white triangles indicate deletions. Flanking 4-bp target duplicate site (TSD) sequences are shown for each provirus. (B) Dose-dependent inhibitory effect of Refrex-1 on viral infection. The supernatants of HEK293T cells transfected with each provirus [ERV-DC7, ERV-DC7/*F.s. silvestris*, ERV-DC16, ERV-DC16/*F.s. silvestris*, or empty vector (Mock)] were diluted and added to fresh HEK293T cells. After removing supernatants, those cells were

infected with the replication-competent ERV-DC14TA virus. X-Gal-positive cells were counted and viral titers were calculated as the log of infectious units (I.U.) per mL with standard deviations. (C) Inhibitory effect on viral infection of the Refrex-1 encoded by truncated Env from ERV-DC7 and ERV-DC7/*F.s. silvestris*. The supernatants of HEK293T cells transfected with the indicated expression vectors encoding ERV-DC7 Env or ERV-DC7/*F.s. silvestris* Env or with empty vector (Mock) were added to fresh HEK293T cells. After removing supernatants, cells were challenged with replication-competent ERV-DC14TA virus. X-Gal-positive cells were counted and viral titers were calculated as the log of infectious units (I.U.) per mL with standard deviations. (D–E) Detection of Refrex-1 expression in transfected cells. HEK293T cells were transfected with the provirus of ERV-DC7 or ERV-DC7/*F.s. silvestris* or with an empty vector (Mock). At 72 h post-transfection, the lysates were prepared for western blotting with a polyclonal goat anti-FeLV SU (gp70) antibody. Refrex-1 protein was detected as bands of ~28-kDa for ERV-DC7 and of ~32-kDa for ERV-DC16. Asterisk indicates the Refrex-1 protein. Human anti-β actin antibody was used as an internal control.

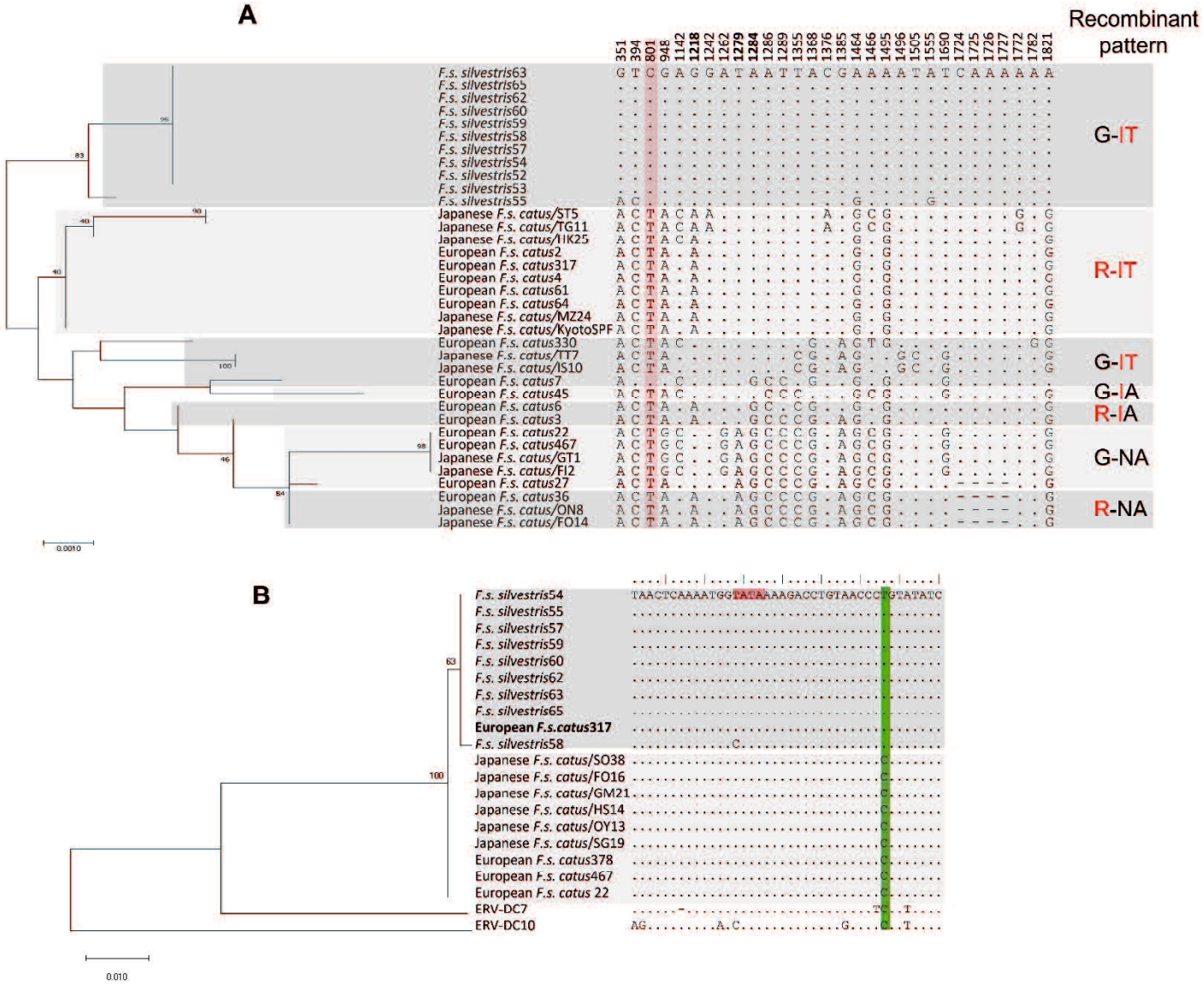

Figure 2.8 Sequence diversity of ERV-DC in wildcats and domestic cats. (A) Sequence diversity of ERV-DC7 *env* in wildcats and domestic cats. ERV-DC7 *env* sequences from 11 European wildcats were determined. A phylogenetic tree of the ERV-DC7 *env* sequences was constructed using maximum-likelihood methods (left). The percentages at branch junctions indicate their bootstrap values (1,000 replicates). Excerpts of polymorphic sites from the ERV-DC7 *env* sequence are shown in the middle. Numbers indicate nucleotide positions and bold numbers indicate nucleotide positions causing amino acid substitutions at positions 407, 427, and 429 (corresponding to R, I, and T, respectively) which cause defective Env cleavage in

ERV-DC7fl. The nucleotide change generating a latent stop codon in domestic cats is shaded in pink. On the right, sequence variations at positions 407 and 427–429 are shown, respectively. Amino acids suppressing Env cleavage are in red. (B) Phylogenetic and sequence analysis of ERV-DC14 based on the 5′-LTR. A phylogenetic tree of the ERV-DC14 5′-LTR in three different cat populations was constructed using maximum-likelihood methods (left). The percentages at branch junctions indicate their bootstrap values (1,000 replicates). Nucleotide sequences of the 5′-LTR are also presented (right). The position of the TATA box is shaded in pink, and the green shading indicates nucleotide determinants distinguishing wildcats and domestic cats.

3. CHAPTER THREE

"Evolutionary dynamics of ERV-DC in jungle cat (*Felis chaus*)"

This work will be published as follows:

Ngo MH, Baba T, Nishigaki K. Evolutionary dynamics of ERV-DC in jungle cats (*Felis chaus*).

3.1. Abstract

Our previous studies identified one of endogenous retrovirus groups in domestic cats termed as ERV-DC (endogenous retrovirus in domestic cat). We also determined existence of ERV-DC in domestic cat (*Felis silvestris catus*) and European wildcat (*Felis silvestris silvestris*). Since ERV-DC did not exist in the genome of the Tsushima leopard cat (*Prionailurus bengalensis euptilurus*), it is thought that ERV-DC internalization occurs only in the cat genome after the bifurcation of the domestic cat lineage and the Leopard cat lineage. We conducted analysis of ERV-DC in jungle cats (*Felis chaus*) showing evolutionary lineage. Based on phylogenetic analysis, existence of ERV-DC/*F. chaus* (endogenous retrovirus in jungle cat) like ERV-DC genotype III compared to three genotypes identified in domestic cats previously. The integration time of ERV-DC/*F. chaus* was estimated to be approximately 160,000 years ago. ERV-DC/*F. chaus* integration polymorphism was not fixed in jungle cat. The sites of integration were different between *F. chaus* and domestic cats. ERV-DCs integrated into *F. chaus* have been inactivated due to gene mutations and deletions, and have lost autonomous growth and infectivity. However, existence of intact ERV-DC/*F. chaus env* was found. Using pseudotyped viruses, this ERV-DC/*F. chaus env* still retained infectious capacity suggesting that existence of infectious ERV-DC/*F. chaus* proviruses. These results suggested that ERV-DC and ERV-DC/*F. chaus* internalized in the *F.s. catus* and *Felis chaus* have undergone evolutionary route independently in the process of co-evolution with each host. Determination and characterization of ERV-DC/*F. chaus* would be helpful for understanding virus evolution and host-virus interaction.

3.2. Introduction

Endogenous retroviruses in domestic cats (ERV-DC) has been identified 13 loci in domestic cat (*F.s. catus*) and 4 loci in European wildcat (*F.s. silvestris*) (51, 96). In domestic cat, almost all ERV-DCs were non-infectious except for ERV-DC14, ERV-DC10 and ERV-DC18 (42, 51). In contrast, all four ERV-DC loci in European wildcat were replication competent defective. Interestingly, we have found that ERV-DC14 in European wildcat were defective due a single nucleotide substitution in env gene which caused env cleavage defective mechanism. Our previous results also suggested that integration time of ERV-DC in European wildcat and domestic cat were almost similar. Thus, this results suggested that ERV-DC probably concurrently infected these two cat lineages after separation from ancestor cat lineage (96).

Mitochondrial DNA analysis indicated *Felis chaus* known as jungle cat is the most distant species belonging to the same *Felis* lineage as *F.s. catus* and is thought to have branched about 3.4 million years ago (66). *F. chaus* has widely distributed in many countries however some countries such as Egypt, Vietnam and some other parts of Asian countries showed significant decline in number due to habitat loss and put them into endangered species (97). Since ERV-DC has been confirmed to exist in *F.s. catus* and was not present in the Tsushima Leopard cat (*Prionailurus bengalensis euptilurus*), these results suggested that ERV-DC may have been internalized only in the domestic cat lineage after separation from Leopard cat lineage (51).

Our preliminary results showed that we have detected 11 proviral integration sites and isolated 9 proviral loci from jungle cat genome using genomic library and plague hybridization techniques. We also successfully isolated nine proviruses and two solo-LTRs in the *F. chaus* genome. The presence of ERV-DC in the *F. chaus* genome was verified by Cat BLAT Search (NCBI) for these 11 loci, and no previously known ERV-DC integration site was found in all loci. In other words, novel ERV-like ERV-DC integration sites that were different from that of

the ERV-DC in domestic cat were revealed. Further, when the nucleotide sequences of nine proviral clones were determined, all proviral clones were disrupted by gene deletions, and the env gene, which would be involved in infectivity, was largely deleted (**Figure 3.1**A). These results suggested that ERV-like ERV-DC in *F. chaus* had been inactivated by the host and lost their infectivity. In addition, target site duplication was identical in each clone (**Figure 3.1**B) so this result may suggest that integration of ERV-like ERV-DC in jungle cat due to infection (139).

In order to investigate whether or not ERV-DC exists in *F. chaus*, PCR was used to detect ERV-DC. As a result, the presence of another ERV-like ERV-DC genotype III was confirmed in *F. chaus*. In contrast, genotype I and II were not detected using specific primers for each genotype (**Figure 3.2**). In addition, ERV-DC integration sites were not detected in all *F. chaus* which were previously identified in *F. s. catus*. These results revealed that all known ERV-DC were not present in *F. chaus*, and only ERV-like ERV-DC genotype III was integrated. However, the integration site of ERV-like ERV-DC genotype III termed as ERV-DC/*F. chaus* was different from that of *F. s. catus*, suggesting that it was internalized by a unique ERV-DC infection in *F. chaus* (139).

Next, we investigated 8 chromosomal DNA of other *F. chaus* individuals. These chromosomal DNAs were not used for analysis of genomic library to analyze insertional polymorphism of ERV-DC/*F. chaus* previously. The presence or absence of integration sites and their genotypes of ERV-DC in these *F. chaus* were verified by PCR method. As a result, ERV-DC/*F. chaus* loci were insertionally polymorphic in jungle cats, and no ERV-DC/*F. chaus* locus was fixed (**Figure 3.3**). This result suggested that the integration of ERV-DC/*F. chaus* into these loci has

occurred very recently (**Table 3.1**) and the integration time calculation was described in the materials and methods section. In contrast, analyzing insertional polymorphism of ERV-DC/*F. chaus* in *F.s. catus*, we used PCR to detect the presence of 11 ERV-DC/*F. chaus* loci and the genotype using 65 *F.s. catus* individuals. As a result, ERV-DC/*F. chaus* integration into these 11 loci was not observed in all 65 domestic cat individuals. This result revealed that ERV-DC and ERV-DC/*F. chaus* were separately infected and internalized in domestic cats and *F. chaus* (139).

The phylogenetic tree analysis was also done in order to investigate which genotypes of nine ERV-DC/*F. chaus* proviruses isolated from the genomic library. We conducted phylogenetic trees based on their sequences of *LTR*, *gag*, *pol*, and truncated *env* genes. As a result, ERV-DC/*F. chaus* is classified as genotype III of ERV-DC in domestic cat based on any parts of proviruses (**Figure 3.4**, **Figure 3.5**, **Figure 3.6**, and **Figure 3.7**). However, ERV-DC/*F. chaus* was classified in another clade independent of genotype III of ERV-DC (139).

In addition, to estimate integration time of ERV-DC/*F. chaus* proviruses isolated in *F. chaus* genome and examine when the proviral integration occurred, the sequences of LTR(s) were analyzed. The number of mutations in the 5 'LTR and the 3' LTR was compared to estimate the integration time. As a result, in 5 clones out of 9 clones, the nucleotide sequences of both LTRs were completely identical, and the integration was supposed to be "Recently". Even in the most mutated clones, the integration was estimated to be 160,000 years ago. This result suggested that integration at each locus was very recent (**Table 3.1**) (139).

As a result of structural analysis of ERV-DC/*F. chaus* identified from the genomic library, all clones were found to have a defect in the *env* gene (**Figure 3.1**). However, due to estimated integration time of ERV-DC/*F. chaus*, these ERV-DC/*F. chaus* proviruses were considered to be very recently integrated so the presence of infectious ERV-DC/*F. chaus* was inferred. In this study, we characterized ERV-DC/*F. chaus* full-length *env* and we have isolated intact and infectious ERV-DC/*F. chaus* full-length *env* gene. The phylogenetic tree analysis of full length *env* gene sequences also showed that ERV-DC/*F. chaus* was distinct clade from ERV-DC genotype III. This result revealed that ERV-DC/*F. chaus* is supposed to be very recently internalized separately from ERV-DC in domestic cat. In addition, characterization of ERV-DC/*F. chaus* showed virus evolution and understanding host-virus interaction. This result also suggests the potential risks of recombination virus occurs in this endangered species.

3.3. Materials and Methods

3.3.1. Sample information

This experiment was conducted using blood of 9 *F. chaus* were collected from Bangladesh. Purification of chromosomal DNA was performed from 300 µl of blood using DNAzol® BD Reagent (Invitrogen). Next, we used a primer pair (the forward primer was Fe-289S and the reverse primer was Fe-247R) to amplify the partial control region of mitochondrial DNA in order to confirm that all blood samples were from *F. chaus*. The phylogenetic analysis indicated that all samples were identical to *F. chaus* (139).

3.3.2. PCR

The PCR reaction was performed using Takara PCR Thermal Cycler Dice (Takara, Shiga, Japan). Polymerases used in this study were GoTaqR Colorless Master Mix (Promega, Madison, WI, USA), KOD-Plus-Ver. 2 (Toyobo, Osaka, Japan), KOD FX Neo (Toyobo, Osaka,

Japan), KOD One^{TM} PCR Master Mix -Blue (Toyobo, Osaka, Japan). The PCR conditions were in accordance with manual instructions.

3.3.3. Amplification of ERV-DC/F. chaus proviruses

According to genomic library and plague hybridization experiment (139), we previously designed eleven primer pairs located 5' and 3' flanking region of 11 identified proviral integration sites. Each primer pair was in accordance with each ERV-DC/*F. chaus* locus including ERV-DC/*F. chaus*1 (Fe-315S and Fe-254R), ERV-DC/*F. chaus*2 (Fe-316S and Fe-255R), ERV-DC/*F. chaus*3 (Fe-317S and Fe-256R), ERV-DC/*F. chaus*4 (Fe-318S and Fe-257R), ERV-DC/*F. chaus*6 (Fe-319S and Fe-258R), ERV-DC/*F. chaus*7 (Fe-320S and Fe-259R), ERV-DC/*F. chaus*8 (Fe-321S and Fe-260R), ERV-DC/*F. chaus*9 (Fe-322S and Fe-261R), ERV-DC/*F. chaus*10 (Fe-323S and Fe-262R), ERV-DC/*F. chaus*11 (Fe-324S and Fe-263R), and ERV-DC/*F. chaus*13 (Fe-325S and Fe-264R). We used PCR to amplify full-length provirus in jungle cat no.4. The amplified PCR product were ligated to pCR-Blunt vector (Invitrogen) or pCR4Blunt-TOPO vector (Invitrogen), then were transformed into *E. coli* HST08 Premium Competent Cells (Takara, Shiga, Japan), and cultured in LB agar medium supplemented with kanamycin. The grown colonies were arbitrarily selected and cultured for 16 hours at 37°C in a kanamycin-supplemented LB broth. A part of which was purified by the alkaline method, the cleaved with specific restriction enzymes, and the inserted fragments were confirmed by electrophoresis. Thereafter, the plasmid was purified using High Pure Plasmid Isolation Kit (Roche). Primer sequences were listed in **Table 3.2**.

3.3.4. Sequences.

The purified plasmid and the purified PCR product were subjected to cycle sequence reaction and then sequenced using Applied Biosystems 3730xl DNA Analyzer (Applied Biosystems Inc.).

3.3.5. Detection of ERV-DC genotypes in F. chaus

We used specific primer pairs for different ERV-DC genotypes including all ERV-DC genotype (Fe-231S and Fe-208R), ERV-DC genotype I (Fe-228S and Fe-205R), ERV-DC genotype II (Fe-232S and Fe-206R), and ERV-DC genotype III (Fe-230S and Fe207R) (42, 51) to detect ERV-DC in 10 samples of jungle cat. We also used a specific primer pair to detect β-actin to validate the PCR experiment. After amplified by PCR, the PCR product was purified using FastGene Gel / PCR Extraction Kit (FastGene® Nippon Genetics), and the nucleotide sequence was determined by sequencing. Primer sequences were listed in **Table 3.2**.

3.3.6. Insertional polymorphism of distinct 11 ERV-DC in F. chaus and F.s. catus

Based on 11 proviral integration sites detected in jungle cat genomes previously, we designed a set of primer pairs located in 5' and 3' flanking region in each integration site. We also used a common forward primer (Fe-36S) which was specific for ERV-DC env gene to detect insertional polymorphism of all ERV-DC genotypes. We conducted PCR to determine presence or absence of ERV-WC proviral loci in all jungle cat samples. Similarly, we also used these primer sets to detect ERV-WC proviral loci in domestic cat samples (N=65). Primer sequences were listed in **Table 3.2**.

3.3.7. Detection of full-length env of ERV-DC/*F. chaus*

We used a primer pair [Fe-627S (forward) and Fe-168R (reverse)] which located in 3'-end of *pol* and U3 regions of 3'-LTR which are conserved in all 9 identified proviral loci. This primer pair was used to PCR amplify full-length env of ERV-DC/*F. chaus*. The PCR product was purified and then were cloned into pCR4Blunt-TOPO vector (Invitrogen) and sequenced. ERV-DC/*F. chaus*2 full length env fused with Myc-tag was amplified by a primer pair (Fe-575S and Fe-565R) and cloned to expression vector pFUΔss (51) using two specific restriction enzymes

(*BamHI* and *EcoRI*). The sequences of ERV-DC/*F. chaus* full-length *env* in expression vector was determined by sequencing. A similar strategy to construct an expression vector carrying ERV-DC10 full-length env fused with Myc tag was conducted for positive control.

3.3.8. Cells and viruses.

Human embryonic kidney 293 cells (HEK293T) and GPLac cells (51) were all maintained in Dulbecco's minimal essential medium with low glucose and supplemented with 10% fetal bovine serum (FBS). GPLac cell were separately transfected with expression plasmids of ERVDC-10, WC3env, and empty vector (mock) to generate corresponding pseudotyped viruses. Transfectants were collected and stored at -80ºC. Infection of target cells with those pseudotyped viruses above were carried out as described below.

3.3.9. Infection assay

Target cells (HEK293T) were seeded in 24-well plates (1.0 x 10^4 cells/well) and incubated overnight at 37°C. The following day the cells were incubated with 1 ml of those pseudotyped viral supernatants above for 2 h in the presence of polybrene (8 µg/ml). The virus supernatant was then replaced with fresh growth medium, and cells were allowed to incubate for a further 2 days before X-Gal (5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside) (Sigma-Aldrich, Canada) staining. Virus titers were determined by counting the number of blue CFU, and titers were calculated as the number of CFU obtained per milliliter of virus supernatant.

3.3.10. Protein analysis

Approximately 0.3 x10^6 HEK-293T cells grown in a 9-cm^2 tissue 6-well culture plate were lysed using 100µl of cell lysis buffer (20 mM Tris-HCl, pH 7.5, 1% Triton X-100, 0.05% sodium dodecyl sulfate, 5 mg/ml sodium deoxycholate, 150 mM NaCl, 1 mM

phenylmethylsulfonyl fluoride) at 4°C for 20 min. Cell genomic DNA was pelleted by centrifugation at 13,000 x g for 20 min at 4°C. Cell lysate supernatant was either stored at -80°C or used for protein analysis. Approximately 30µg of total protein was run on a 10% SDS-polyacrylamide gel, and proteins were subsequently transferred to a nitrocellulose membrane (Pall, Pensacola, FL). Myc-tagged ERV-DC10 and ERV-DC/*F. chaus* full-length env proteins were detected by incubation of nitrocellulose membranes with anti-myc monoclonal antibody (9B11) (Cell Signaling) diluted 1 in 1000 in phosphate-buffered saline (PBS) containing 0.1% Tween 20. This was followed by incubation with goat anti-mouse antibody conjugated to horseradish peroxidase (Sigma) diluted at 1:3000 ratio in PBS–0.1% Tween 20. Signals were detected using chemiluminescence reagent (Perkin Elmer, Boston, MA), followed by exposure to Kodak Biomax MR film.

3.3.11. Phylogenetic analysis

We retrieved sequence of ERV–DC1, –DC2, DC–3, –DC4, –DC6, –DC7, –DC8, –DC10, –DC14, –DC16, –DC17, –DC18, and –DC19 from NCBI. Multiple alignments were conducted using muscle (140). The phylogenetic trees of LTRs, *gag*, truncated *env* and full-length *env* genes were inferred using the Neighbor-Joining method based on Kimura 2-parameter model (124) while Maximum Likelihood method based on the Hasegawa-Kishino-Yano model (127) was used for *pol* gene. The best substitution pattern was selected based on the lowest Bayesian Information Criterion. The percentage of replicate trees in which the associated taxa clustered together 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 are in the units of the number of base

substitutions per site. All positions containing gaps and missing data were eliminated. All evolutionary analyses were packaged in MEGA 7 (141).

3.3.12. Integration time Estimation of ERV-DC in F. chaus

In order to estimate the time of ERV-DC incorporation in *F. chaus*, the number of mutations in the 5 'LTR and 3' LTR of each clone was compared, and the average branching ratio (1.2 x 10^{-8} substitutions / site / year) (142) was substituted into the formula of age estimation using. The 5 'LTR and the 3' LTR of each clone were multiply aligned, and the genetic distance (D) was determined at Mega 7 using the Kimura two-parameter model (124). The embedded age was calculated using the formula (embedded age) = (D / average branch ratio) × 1/2. The characteristics of LTR of each clone are described in **Table 3.1**.

3.3.13. Detection of full-length ERV-DC/F. chaus provirus

We used a different approach to amplify the ERV-DC/*F. chaus* provirus from *F. chaus* no.2 and no.3. For detection, a primer pair specific to the ERV-DC10 3'LTR region including Fe-148S (forward) and Fe-243R (reverse) were used. The full-length proviruses expected to be detected was cloned and sequenced.

3.3.14. GenBank accession numbers of reference sequences

ERV-DC1 (AB674439.1), ERV-DC2 (AB674449.1), ERV-DC3 (AB674440.1), ERV-DC4 (AB674441.1), ERV-DC6 (AB674450.1), ERV-DC7 (AB807599.1), ERV-DC8 (AB674443.1), ERV-DC10 (AB674444.1), ERV-DC14 (AB674445.1), ERV-DC16 (AB807600.1), ERV-DC17 (AB674446.1), ERV-DC18 (AB674447.1), ERV-DC19 (AB674448.1), and ECE1 (X51929.1)

3.4. Results

3.4.1. Detection of ERV-DC/F. chaus full-length env

In order to figure out whether or not existence of intact ERV-DC/*F. chaus env* in jungle cat, we used PCR to amplify full-length env of ERV-DC/*F. chaus* based on ERV-DC10 full-length env in domestic cat. As a result, we could detect 1.9kb fragment in three jungle cats (cat ID: JC2, JC3 and JC4) as well as 0.7kb fragment were known as truncated env form of ERV-DC/*F. chaus* (**Figure 3.8**). The three fragments were cloned and sequenced. As a result of genetic analysis, the nucleotide sequence of the full-length env gene of ERV-DC/*F. chaus* detected from *F. chaus* JC2 and JC3 were very close to that of ERV-DC10, and only 4-amino acid substitutions were observed while full-length env gene of ERV-DC/*F. chaus* detected from *F. chaus* JC4 was disrupted by a deletion (**Figure 3.9**). Phylogenetic analysis of full-length *env* gene also indicated that ERV-DC/*F. chaus* belonged to genotype III of ERV-DC but is distinct clade from that in domestic cat (**Figure 3.10**). These results suggested that there existed intact ERV-DC genotype III full-length *env* in jungle cats.

3.4.2. Examination of infectivity of full-length env detected in F. chaus

In order to investigate whether this env full-length clones retained infectivity similar to ERV-DC10, Lac-Z pseudotyped viruses were prepared and infection experiments were conducted. As the result, the env gene detected from *F. chaus* (JC2 and JC3) showed infectious (**Figure 3.11**) similar to ERV-DC10 env (positive control). This result indicated that even ERV-DC proviruses were disrupted by gene deletions, there still existed intact full-length ERV-DC env in two *F. chaus* which still conferred for their infectivity.

3.4.3. Examination of expression of ERV-DC/F. chaus full-length env in culturing cells

In order to assess the expression of ERV-DC/*F. chaus* full length env in culturing cells, we carried out western blotting using cell lysates. As a result, expression of ERV-DC/*F. chaus*

full-length env in culturing cells was similar to ERV-DC10 (positive control) (data not shown). This result also conferred the functionality of ERV-DC/*F. chaus* full-length env which could incorporate into a virion.

3.4.4. Examination of intact ERV-DC/F. chaus full-length proviruses

Based on these findings, we are wondering whether or not existence of infectious ERV-DC/*F. chaus* provirus. We used a primer pair (Fe-148S and Fe-243R) located in 3'LTR of ERV-DC/*F. chaus* full-length form to amplify the full-length proviruses. As a result, we could detect approximately 9.0kb fragment which can resemble to full-length proviruses in *F. chaus* JC3 and JC4 (**Figure 3.12**). Due to disrupted ERV-DC/*F. chaus*4 full-length *env*, we will next continue to isolate this full-length provirus from JC3 and determine its infectivity as a future plan.

3.5. Discussions

In this study, we found existence of endogenous retrovirus in jungle cat termed as ERV-DC/*F. chaus* like ERV-DC in *F. chaus*. All ERV-DC/*F. chaus* proviruses were noninfectious and disrupted by deletions in different region throughout proviral genome (**Figure 3.1**). These structures were similar patterns to ERV-DC. Besides, only ERV-DC genotype III was identified in *F. chaus* compared to existence of the three genotypes in *F.s. catus* (**Figure 3.2**). It was suggested that the integration of ERV-DC genotype III occurred to each species after *F. chaus* diverged from other domestic cat genera. In addition, the ERV-DC/*F. chaus* integration site identified this time was different from that of ERV-DC in domestic cat because ERV-DC was not detected in *F. chaus* at the previous ERV-DC integration site. The integration of ERV-DC/*F. chaus* into *F. chaus* was considered to have occurred independently. The question is why only genotype III of ERV-DC transmitted between species but not genotypes I and II, the

reason may be due to ERV-DC genotype III still retains as infectious viruses. In addition, genotype I is suppressed by Refrex-1 and shows no infectivity, and genotype II functions as Refrex-1. Thus, it is possible that ERV-DC genotype III could transmit among *Felis* lineages.

ERV-DC is considered to integrate into *F. chaus* genome recently. There is no fixed loci of ERV-DC in *F. chaus* compared to domestic cats and European wildcats. This result suggested that there is no existence of anti-retroviral molecule like refrex-1 in jungle cats. Examination of these newly identified loci in *F. s. catus* samples indicated that there is no existence of similar loci in domestic cats. These results suggested that integration of ERV-DC was probably independent between *F. s. catus* and *F. chaus*.

Phylogenetic tree revealed that all ERV-DCs isolated from *F. chaus* were classified as ERV-DC genotype III. This is consistent with the results of PCR detection of ERV-DC genotypes, suggesting that the integration of ERV-DC in *F. chaus* is only genotype III. In addition, ERV-DC/*F. chaus* detected from *F. chaus* form an independent clade compared to ERV-DC genotype III in domestic cat. This result suggests that ERV-DC/*F. chaus* has their own unique integration. These results suggested that ancestral ERV-DC independently infected *F. s. catus* and *F. chaus*.

Based on estimated integration time of ERV-DC in *F. chaus*, ERV-DC integrated into *F. chaus* genome recently about 160,000 years ago. However, *F. chaus* is considered to have diverged from the cat lineages about 3.4 million years ago. This result suggested that ERV-DC infected *F. chaus* after separation from cat lineages.

In this study, we found existence of intact full-length ERV-DC/*F. chaus env* which was also classified into ERV-DC genotype III and highly identical to ERV-DC10 besides identifying non-infectious ERV-DC/*F. chaus* proviruses in *F. chaus*. The phylogenetic analyses suggested that this ERV-DC/*F. chaus* full length *env* was distinct from ERV-DC full length *env* genotype

III in domestic cats. This full-length *env* was still infectious when forming pseudotyped viruses. These results suggest there is possibility to recombine with exogenous retroviruses during infections step to make a new recombinant pathogenic virus similar to FeLV-B and FeLV-D. In addition, it is possible that there existed infectious full-length ERV-DC/*F. chaus* proviruses (**Figure 3.12**) in jungle cat. More studies are needed to confirm this possibility.

3.6. Conclusions

In this study, we found ERV-DC/*F. chaus*. The integration time of ERV-DC/*F. chaus* was considered as recently. ERV-DC/*F. chaus* integrated into *F. chaus* independently from ERV-DC in domestic cat. ERV-DC/*F. chaus* was classified into ERV-DC genotype III but different clade from ERV-DC genotype III in domestic cat. There is no existence of ERV-DC genotype I and II in *F. chaus* genome. This result suggests that there is no existence of refrex-1 against FeLV-D and ERV-DC genotype I. We also found the intact full-length ERV-DC/*F. chaus env* showing infectious functionality. ERV-DCs internalized in domestic cat and *F. chaus* have evolved in the process of co-evolution with their respective hosts. It is possible that ERV-DC in *F. chaus* is expected to be destroyed as junk DNA, but it also has the possibility of acquiring a function as a host gene like Refrex-1 of the domestic cat. However, a new recombinant pathogenic virus could be generated due to a new infection into *F. chaus* existing intact full-length infectious ERV-DC/*F. chaus env*.

3.7. Tables and Figures in Chapter three

TABLES

Table 3.1 Genetic diversity of LTRs in ERV-DC/*F. chaus* and integration time estimation.

Proviral locus	No. of differences between 5' and 3' LTRs	Homology between 5' and 3' LTRs (%)	Length of LTR (bp)		Genetic distances between 5’ and 3’ LTR	Calculated integration age (mya)
			5' LTR	3' LTR		
Locus No.1	2	99.5	551	551	0.004	166666.7
Locus No.2	0	100	551	551	0	0
Locus No.3	1	99.8	551	551	0.002	83333.33
Locus No.4	0	100	551	551	0	0
Locus No.6	0	100	551	551	0	0
Locus No.7	1	99.8	551	551	0.002	83333.33
Locus No.8	0	100	551	551	0	0
Locus No.9	1	99.8	551	551	0.002	83333.33
Locus No.13	0	100	551	551	0	0

Table 3.2 Sequences of primers were used in this study.

Name	Sequence	Specificity
Fe-36S	5'-AACCGCTTGGTACARTTCATAAGAG-3'	ERV-DC env
Fe-148S	5'-TGGTYTAGYTTAYTAAAA-3'	ERV-DC10 3'LTR
Fe-228S	5'-GCTTGCACTTCCACCAGTTG-3'	ERV-DC Gl 3'LTR
Fe-230S	5'-GCCTCCCTACCCGACTTCC-3'	ERV-DC Glll env
Fe-231S	5'-TCCACCCTCACACCAGAATC-3'	ERV-DC env
Fe-232S	5'-GCCAGATACAATCGAATGAAAGG-3'	ERV-DC Gll 3'LTR
Fe-289S	5'-CATTTCAACGTGGGGGTTTC-3'	mtDNA control region
Fe-315S	5'-TATCCAAGCACACTTTCCAGCA-3'	ERV-DC/*F. chaus*1 5'flanking
Fe-316S	5'-TCTCAGCTCTTCCCAGGACTTT-3'	ERV-DC/*F. chaus*2 5'flanking
Fe-317S	5'-CTGTGTCTCCACACCCTAGCC-3'	ERV-DC/*F. chaus*3 5'flanking
Fe-318S	5'-GATGATAAGCTTTGCATTTGAGA-3	ERV-DC/*F. chaus*4 5'flanking
Fe-319S	5'-TGATAAGAAAGCACAAGTGGAAC-3'	ERV-DC/*F. chaus*6 5'flanking
Fe-320S	5'-TCCTAAGGAAGGGAGAAAAGG-3'	ERV-DC/*F. chaus*7 5'flanking
Fe-321S	5'-GATGTAACGTATCACCCAAGAGTAG-3'	ERV-DC/*F. chaus*8 5'flanking
Fe-322S	5'-GTCAGGTAATTGCCCAACCTTAC-3'	ERV-DC/*F. chaus*9 5'flanking
Fe-323S	5'-CTAAAACACAAAACAAAACAAAGACT-3'	ERV-DC/*F. chaus*10 5'flanking
Fe-324S	5'-ACCAGGCCTACCTATGTTCAC-3'	ERV-DC/*F. chaus*11 5'flanking
Fe-325S	5'-GTCACTCTTAGGCCCATTCTGT-3'	ERV-DC/*F. chaus*13 5'flanking
Fe-627S	5'-ACCAGAAACTCCCAAAACCTG-3'	ERV-DC pol
Hub-b-actin(DNA)S	5'-ATCATGTTTGAGACCTTCAA-3'	Human β-actin
Fe-168R	5'-GAAGRTAGGGTGGGGGTGTKTTAGTAAGCTA-3'	ERV-DC10 3'LTR
Fe-205R	5'-ACCTGTTCCTGTCTTGCGTAG-3'	ERV-DC Gl 3'LTR
Fe-206R	5'-TGCCAACTGGTTTTGTTACTTATG-3'	ERV-DC Gll 3'LTR
Fe-207R	5'-AGGGGGTTTAGCCGTTAGG-3'	ERV-DC Glll env
Fe-208R	5'-TGAGTCATGGTAGAAGATTTTTGG-3'	ERV-DC env
Fe-243R	5'-GCTCTCCCGCTTTCTAACACTG-3'	ERV-DC 3'LTR
Fe-247R	5'-CCATTGACTGAATAGCACCTTGA-3'	mtDNA control region
Fe-254R	5'-GTGGTGGGAAGTAATGAGCTAC-3'	ERV-DC/*F. chaus*1 3'flanking
Fe-255R	5'-ACACGATGAGCCTTGTTTGAG-3'	ERV-DC/*F. chaus*2 3'flanking
Fe-256R	5'-ATACTGCTATCCCCTCCTTCTG-3'	ERV-DC/*F. chaus*3 3'flanking
Fe-257R	5'-AAGAATTGGGATCCAAGGAATG-3'	ERV-DC/*F. chaus*4 3'flanking
Fe-258R	5'-GCATTTATCATTACTCGGTGTTACC-3'	ERV-DC/*F. chaus*6 3'flanking
Fe-259R	5'-GTGACTATACTCAGGGGGAAGTTA-3'	ERV-DC/*F. chaus*7 3'flanking
Fe-260R	5'-GCCCTTTGCCTTCAACTTACCT-3'	ERV-DC/*F. chaus*8 3'flanking
Fe-261R	5'-TGTCTGTCTGTCTTGGGGAGAC-3'	ERV-DC/*F. chaus*9 3'flanking
Fe-262R	5'-GGAACAGACTTTGAATGGTACAGA-3'	ERV-DC/*F. chaus*10 3'flanking
Fe-263R	5'-TAATCCGCACACCGTACTCC-3'	ERV-DC/*F. chaus*11 3'flanking
Fe-264R	5'-TGGCCACTCCTCTTTCCTACC-3'	ERV-DC/*F. chaus*13 3'flanking
Fe-510R	5'-ccggatccctacaggtcttcttcagagatcagtttctgttcTTCAATTGTATCTGGCCT-3'	ERV-DC10 env myc BamHl
Hub-b-actin(DNA)R	5'-AGATGGGCACAGTGTGGGT-3'	Human β-actin

FIGURES

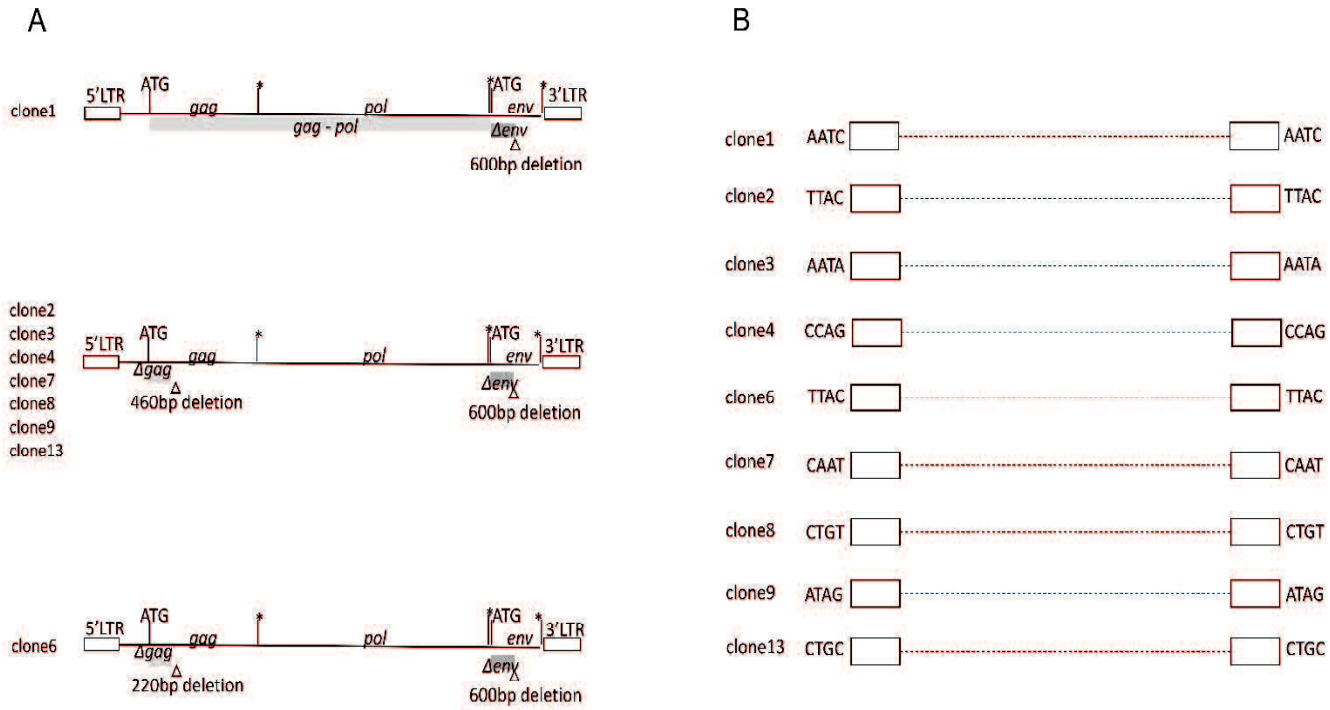

Figure 3.1 Characterization of ERV-DC in jungle cats. (A) Structures of the genomes of 11 full-length ERV-like ERV-DC proviruses. The gag, pol, and env genes are illustrated together with the 5' and 3' LTRs and positions of the *gag* and *env* translational initiation codons (ATG). Asterisks indicate stop codons. Gag and Pol proteins may be synthesized as a large single polypeptide precursor via termination suppression (143). An open triangle indicates a deletion of nucleotides. (B) Base sequence of TSD (Target Site Duplication) at each integration site. Flanking 4-bp TSD sequences are shown for each provirus.

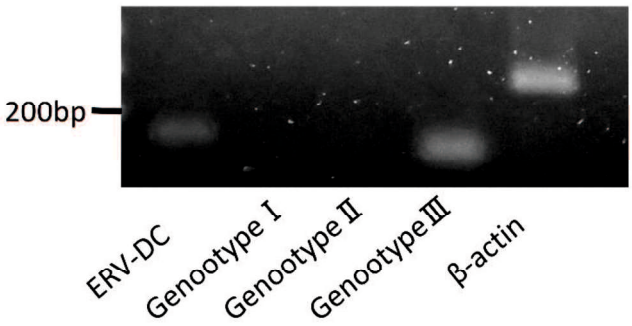

Figure 3.2 Detection of ERV-DC genotype in jungle cats. Light bands indicate detection. β-action was used to validate the PCR experiment.

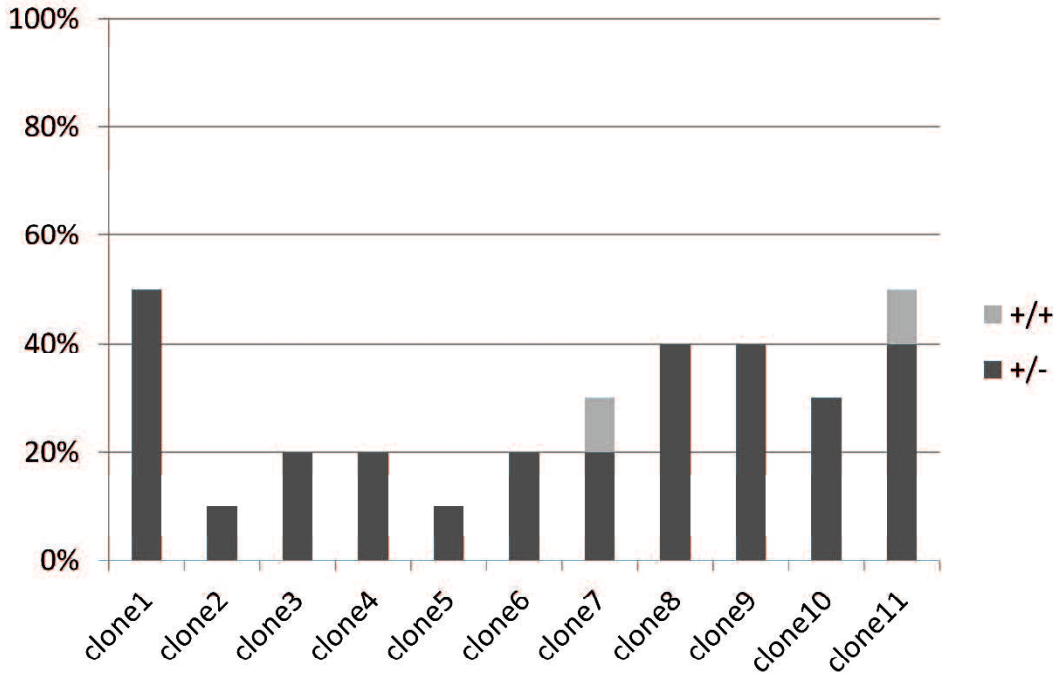

Figure 3.3 Insertional polymorphism of 11 ERV-DC/*F. chaus* loci in 10 *F. chaus* individuals. +/-, provirus was detected on heterozygous loci; +/+, provirus was detected on homozygous loci.

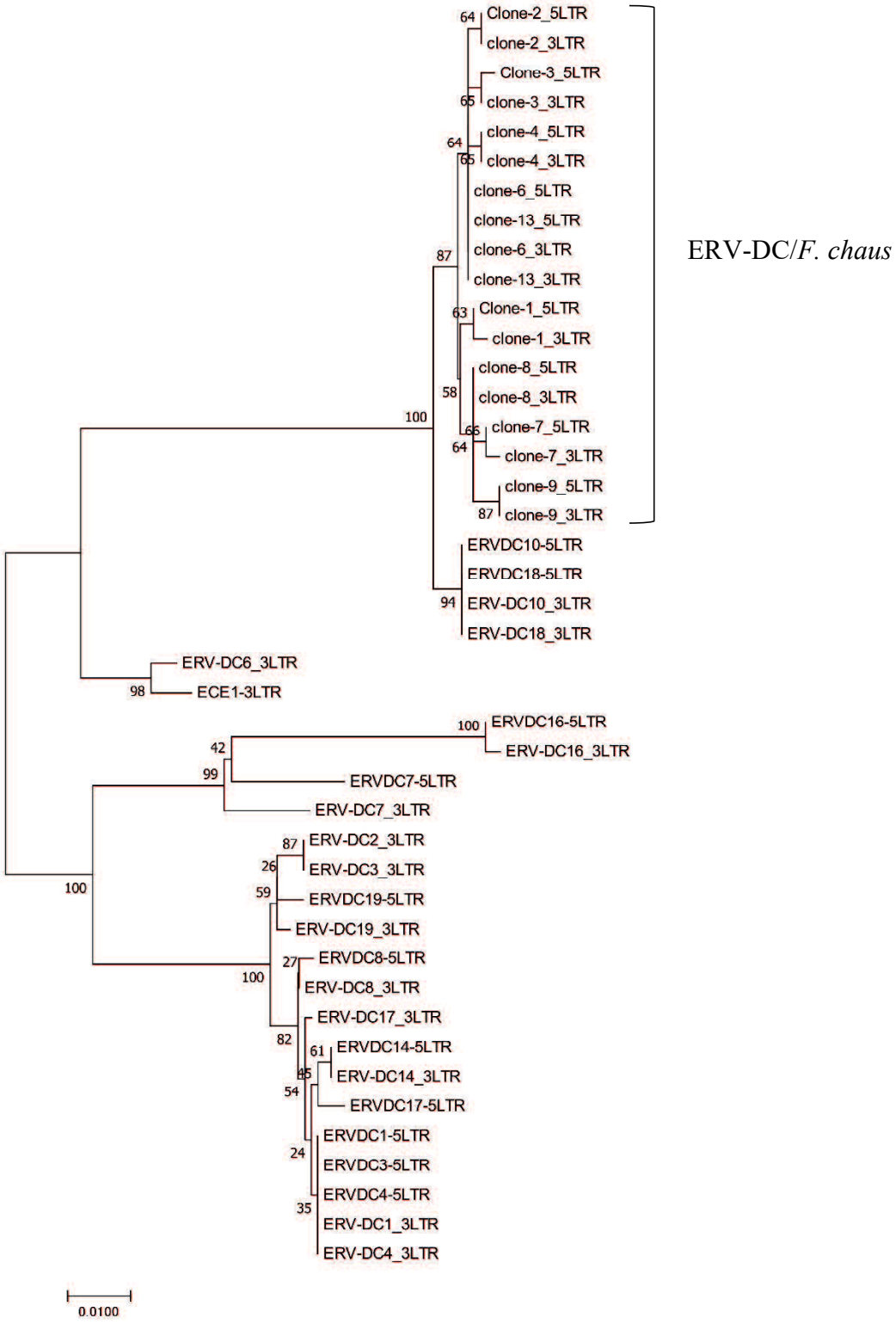

Figure 3.4 The phylogenetic analysis of ERV-DC/*F. chaus* and ERV-DCs based on *LTRs*. The phylogenetic tree was constructed based on neighbor-joining method. The percentages at the branch junctions indicated their bootstrap values (1,000 replicates).

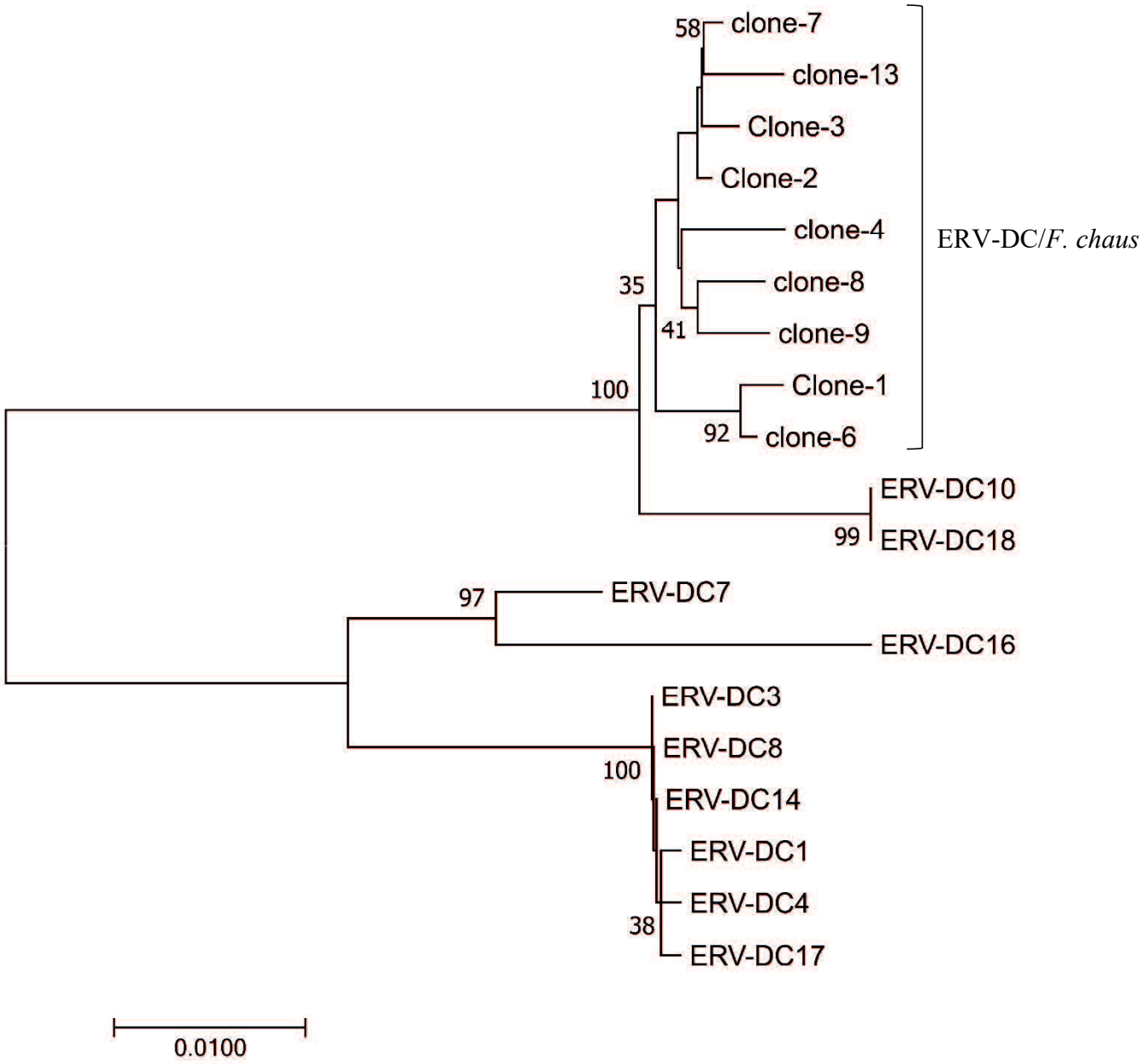

Figure 3.5 The phylogenetic analysis of ERV-DC/*F. chaus* and ERV-DCs based on *gag* gene. The phylogenetic tree was constructed based on neighbor-joining method. The percentages at the branch junctions indicated their bootstrap values (1,000 replicates).

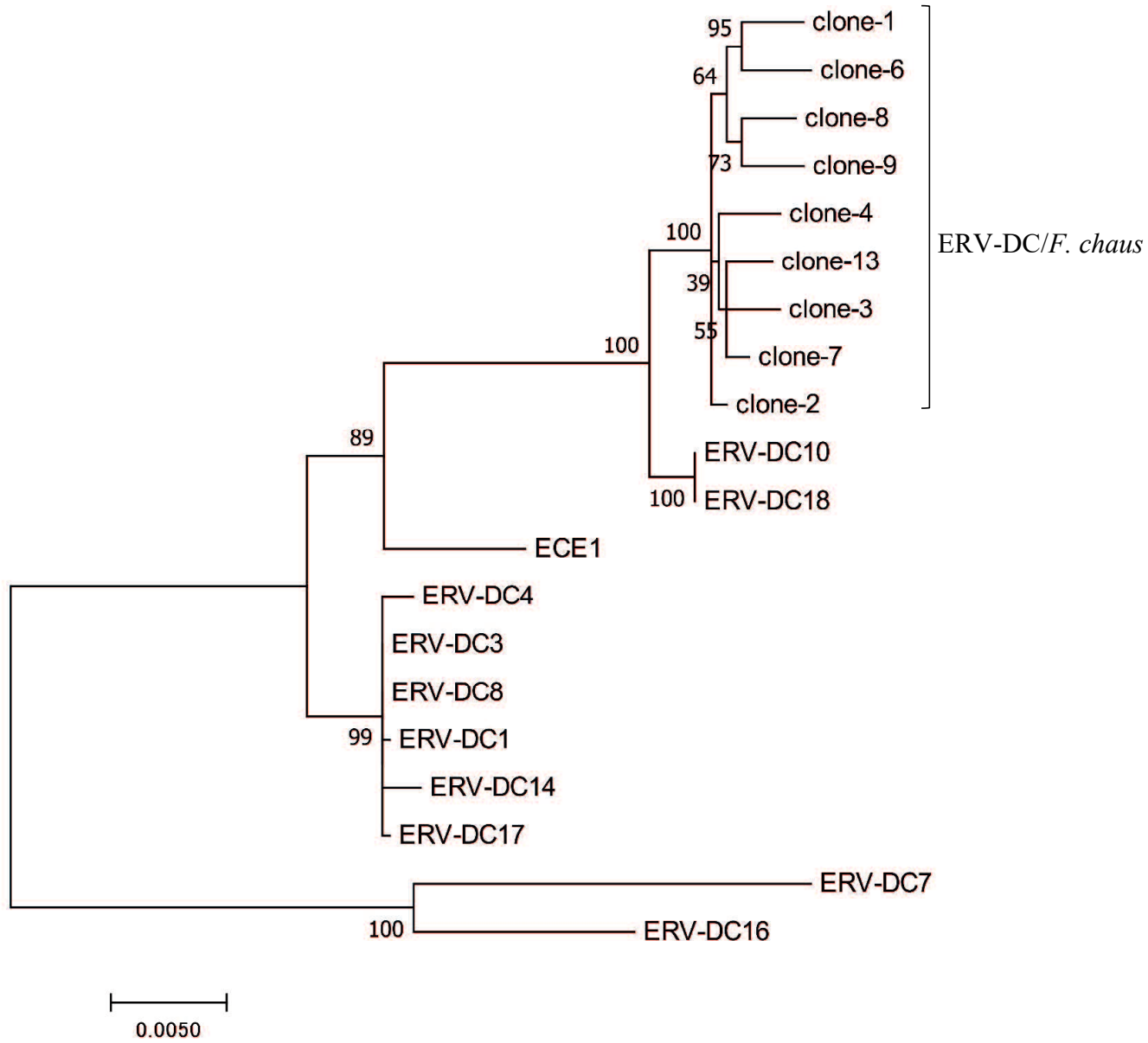

Figure 3.6 The phylogenetic analysis of ERV-DC/*F. chaus* and ERV-DCs based on *pol* gene. The phylogenetic tree was constructed based on Maximum Likelihood method. The percentages at the branch junctions indicated their bootstrap values (1,000 replicates).

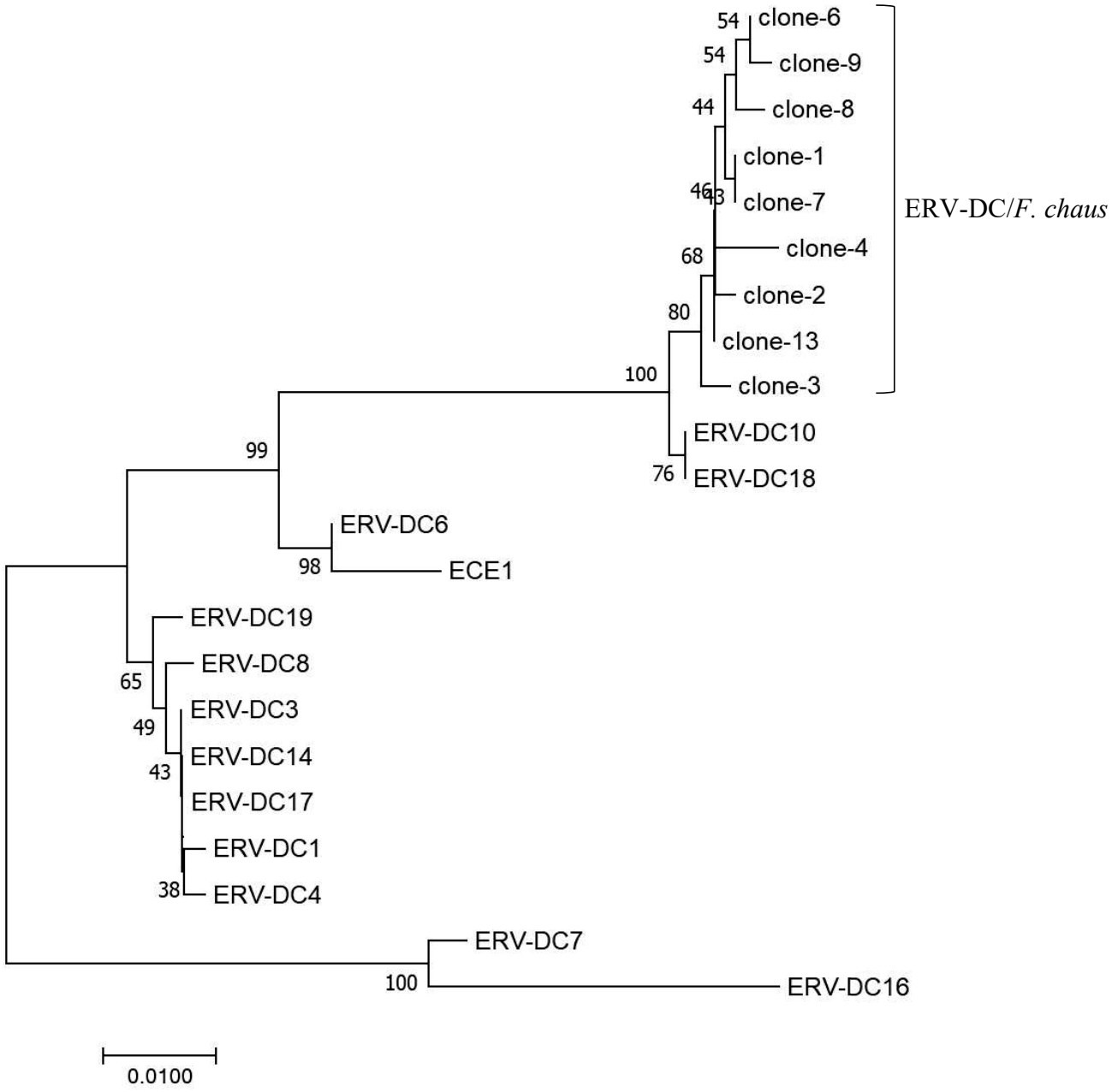

Figure 3.7 The phylogenetic analysis of ERV-DC/*F. chaus* and ERV-DCs based on *env* gene. The phylogenetic tree was constructed based on neighbor-joining method. The percentages at the branch junctions indicated their bootstrap values (1,000 replicates).

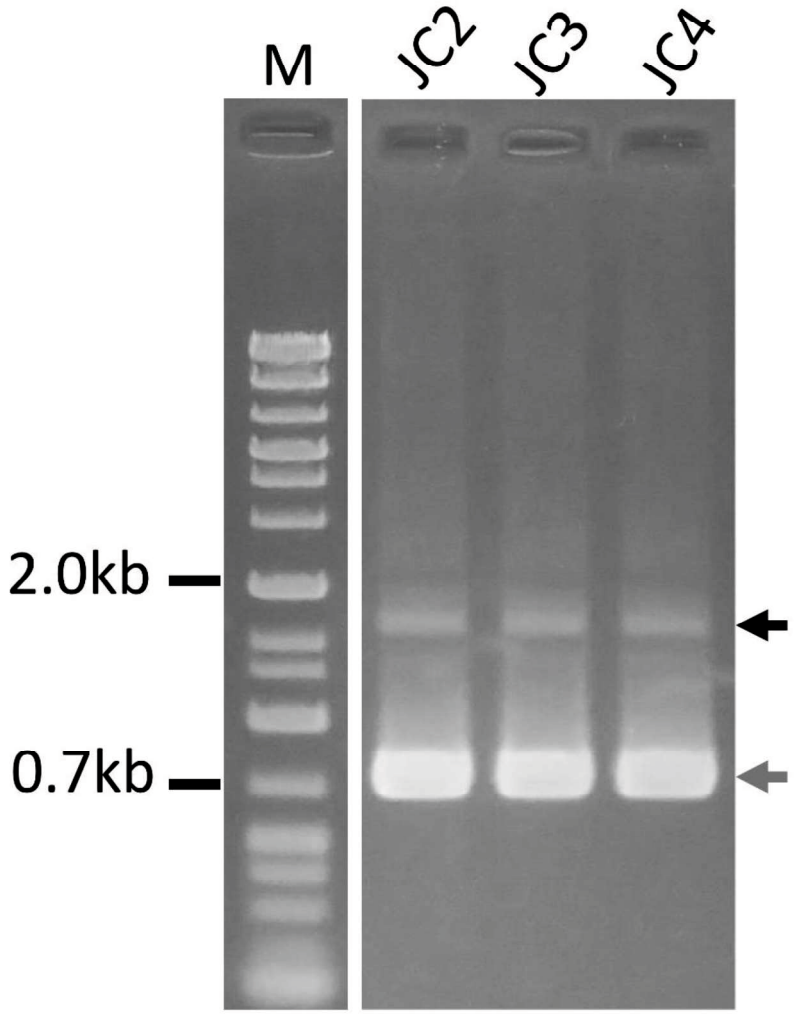

Figure 3.8 PCR detection of full-length *env* of ERV-DC/*F. chaus*. Black arrow indicates full-length *env*. Gray arrow indicates truncated *env*. Primer pairs were Fe-627S and Fe-168R which located in 3'-end of pol and U3 regions of 3'-LTR. JC1, JC2, JC3 were jungle cat samples no. 1,2 and 3.

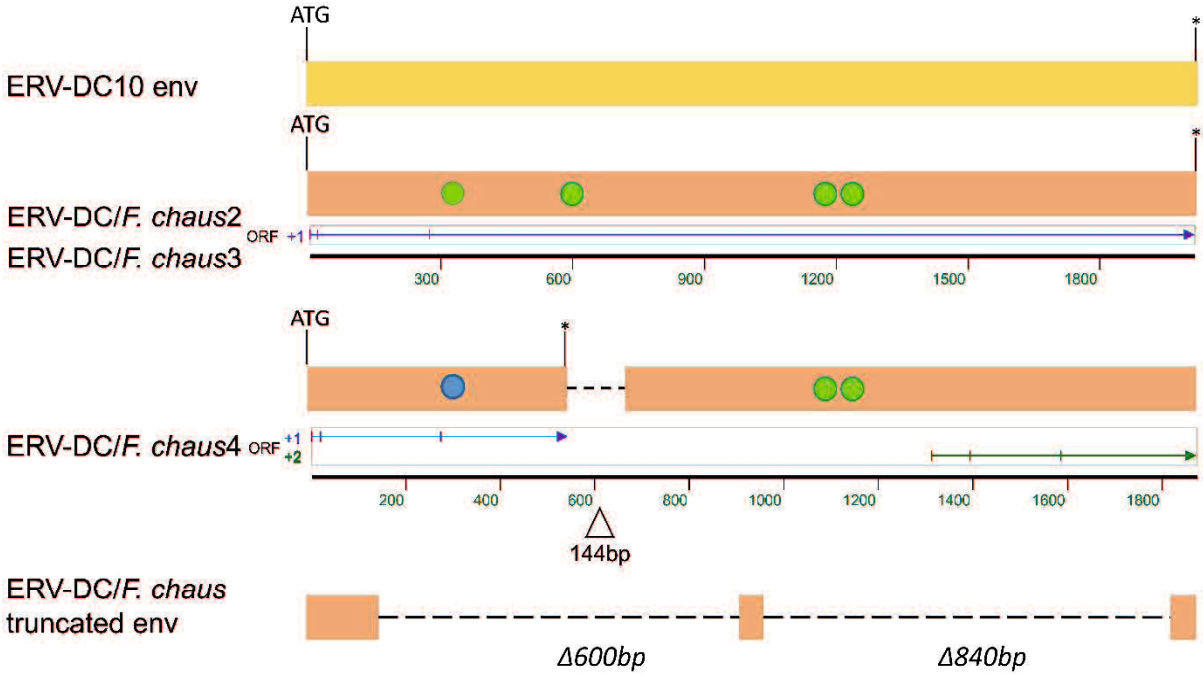

Figure 3.9 Sequence diversity of ERV-DC10 and ERV-DC/*F. chaus* based on env gene. ATG, start codon; asterisks indicate stop codon. Green and Blue circles indicate SNPs among three ERV-DC/*F. chaus* full length *env* compared with ERV-DC10. Light triangles and dash lines indicated sequence deletions.

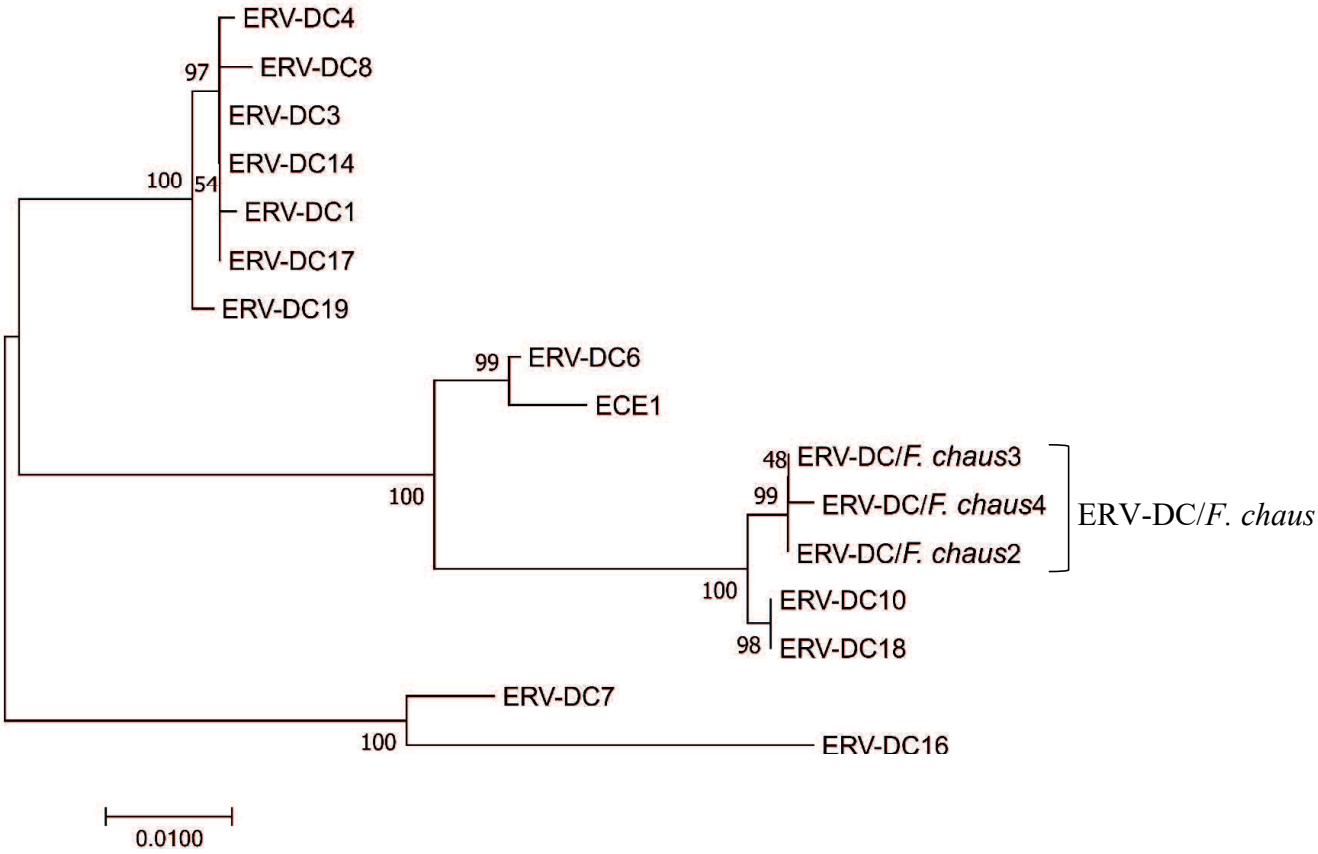

Figure 3.10 Analysis of ERV-DC/*F. chaus* full-length *env* gene. The phylogenetic tree was constructed based on neighbor-joining method. The percentages at the branch junctions indicated their bootstrap values (1,000 replicates).

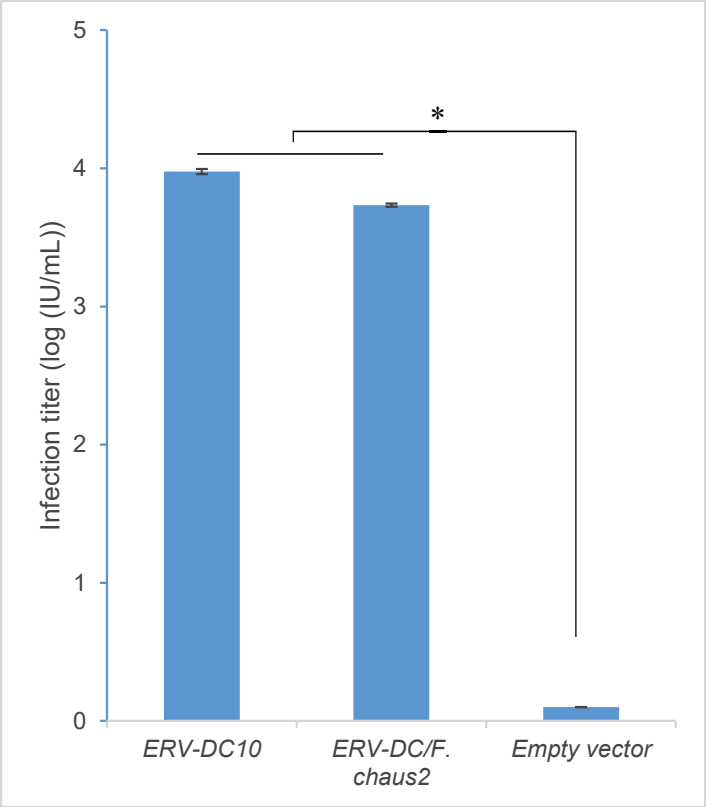

Figure 3.11 Infectivity of ERV-DC/*F. chaus* env pseudotyped virus. GPLac cells were transfected with env expression vector of ERV-DC/*F. chaus*2, ERV-DC10 (positive control) and empty vector (mock). Viral supernatants were collected after 72 h and used for infection assay with fresh HEK-293T cells. Viral titers are illustrated as the log number of infectious units (IU) per milliliter with standard deviations. *, P-value < 0.0001 (one-way ANOVA).

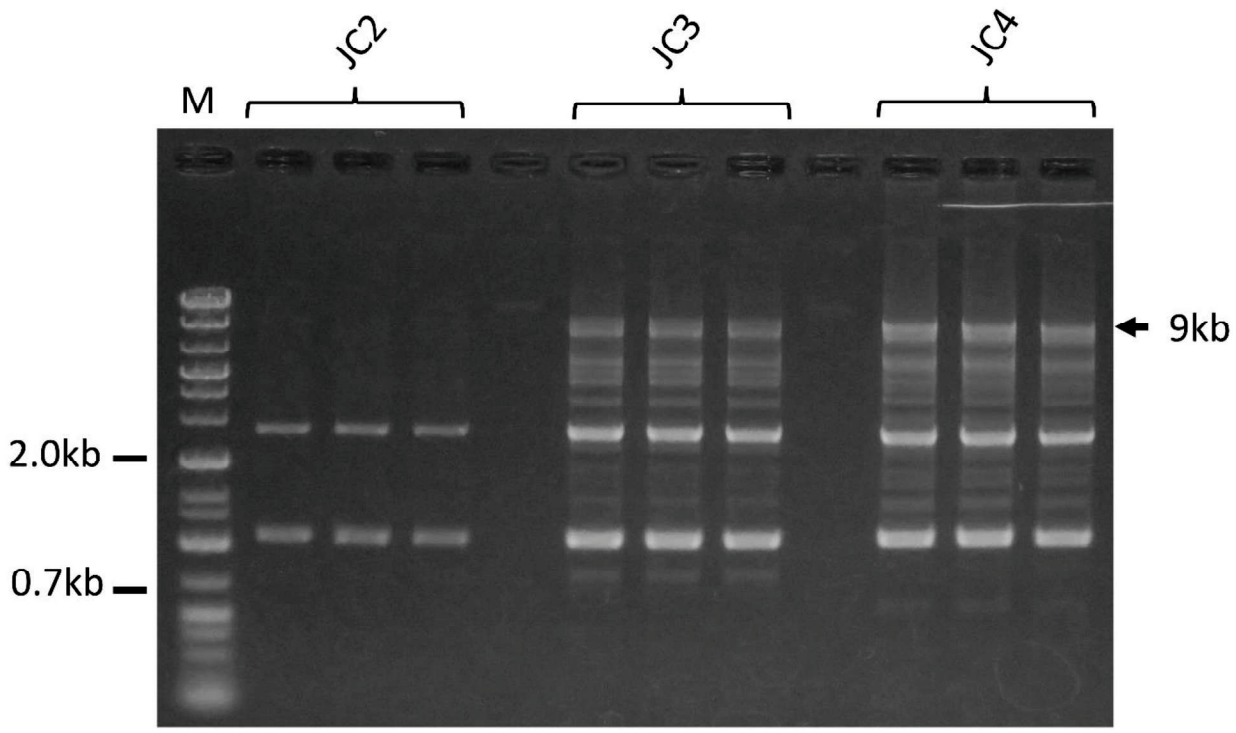

Figure 3.12 PCR detection of ERV-DC/*F. chaus* proviruses. Three chromosomal DNA sample from different three jungle cats were used. Black arrow indicated full proviruses. Primer pairs were Fe-148S and Fe-243R which located in 3' LTR. JC1, JC2 and JC3 are jungle cat sample No. 1, 2 and 3.

GENERAL DISCUSSIONS AND CONCLUSIONS

Endogenous retroviruses are known as remnants of ancestral germlines exogenous retroviral infection in the host genome. At present, molecular studies on biological cellular functions of ERV mainly focuses on mammalian genomes such as mice, human, chickens, livestock and pets gaining new insights into potential disease induction as well as animal and viral evolution. However, only inbred or domesticated species investigated could not reflect the real overview of ERV endogenization. Thus, broader comparative analyses of ERV in wild species probably is a golden key to understand biological properties of ERV, the reasons for predisposing ERV endogenization as well as their impacts on host biology (144).

Now, identifying and characterizing existence and prevalence of ERV in wildcats and domestic cats, determining the biological functions of those in different cat populations, and elucidating ERV evolution as well as animal evolution are the main objectives in this dissertation.

My first chapter described prevalence of two infectious endogenous retroviruses in mixed-breed and purebred cats. The frequency of ERV-DC10 (34.5%) was significantly higher than ERV-DC14 (4.1%) in domestic cats. Interestingly, prevalence of ERV-DC10 was higher in mix-breed cat population compared with purebred cat population but the opposite comparison was true with prevalence of ERV-DC14. In term of each purebred cat, invasion of ERV-DC10 is broader than that of ERV-DC14. Furthermore, existence of ERV-DC10 homozygous (N=191, 11.6%) is significantly different compared to only 0.1% of ERV-DC14 which was found only one mixed-breed Japanese domestic cat. These results suggested that ERV-DC10 may tend to expand in the domestic cat population while ERV-DC14 is probably to be deleterious in the domestic cat genome. One hypothesis is that existence of infectious ERV-DC14 provirus homozygous in domestic cat genome may predispose to affect the embryogenesis process. In addition, this research could rule out a second hypothesis that distinct cat ancestors harbored

these two ERV-DC loci (European wildcats may harbor ERV-DC14 while Asian and African wildcat may harbor ERV-DC10). The frequencies of these two infectious ERV-DC10 and ERV-DC14 in domestic cats in different countries also supported this second hypothesis.

My second chapter explained about tracking the fate of endogenous retrovirus with regard to ERV-DC segregation in wild and domestic cats. To better understanding of the evolution of ERV-DC after integration into the *Felis* linage, I investigated the possible integration of ERV-DC in European wildcat which is considered as one of cat ancestor based on phylogenetic analysis of cat genome (65). This study firstly revealed that existence of ERV-DC integration in European wildcat genome but different pattern compared with that in domestic cat genome (51). Next, I successfully cloned ERV-DC14 in European wildcat (ERV-DC14/*F.s. silvestris*) which showed the same integration time with ERV-DC14 in domestic cat (51). However, this ERV-DC14/*F.s. silvestris* is highly prevalent in European wildcat and is inactivated through a single nucleotide mutation G to A results in an E148K residue mutation in *env* gene. This mutation resulted in Env cleavage dysfunction which was observed in all infected European wildcats. This is the first report indicated that an infectious endogenous retrovirus was inactivated by a single mutation through a common mechanism about failure in viral incorporation into a virion. This result suggested that ancestral ERV-DC14 may infect domestic cats and European wildcat independently at the same integration time. While ERV-DC14 is still infectious in domestic cats, ERV-DC14/*F.s. silvestris* is inactivated in all tested European wildcats. This results also conferred the above hypothesis that infectious ERV-DC14 is probably deleterious in domestic cat genome. More interestingly, this mutation was also found in a Feline leukemia virus isolated from a cat showed naturally occurring thymic lymphoma. Introduction of the same mutation in other gammaretroviruses showed the similar dysfunctional results. These results suggested a common mechanism of virus inactivation

during the interaction between virus and the host. Moreover, this interesting finding may also contribute to a strategy to produce gene therapy against viral infections.

Another interesting point this second study is that Refrex-1 was also found in European wildcat. However, Refrex-1 level in European wildcat seems to be lower than that in domestic cat due to the different integration of ERV-DC7 and ERV-DC17 among European wildcat and domestic cat. Even so, Refrex-1 still retains their antiviral activity against FeLV subgroup D and ERV-DC genotype I. Diversity of ERV-DC between European wildcat and domestic cat based on ERV-DC7 could show different integration time of ERV-DC in these two cat populations. This result suggested that Refrex-1 maintained its antiviral role before and after cat domestication. Probably, activity of refrex-1 in modern cat is higher than that in ancestor cat due to emergence of newly recombinant pathogenic viruses invading modern cat population.

My third chapter was discussed about evolutionary dynamic of ERV-DC in jungle cat (*Felis chaus*). Based on phylogenetic analysis of cat genome, *Felis chaus* is considered as the farthest ancestor of domestic cat lineage after separation from leopard lineage (66). Therefore, I conducted a research into ERV-DC in this wildcat specie to show evolution of ERV-DC in *Felis* lineage because there was no integration of ERV-DC in Tsushima Leopard cat (51). Thus, I hypothesized that ERV-DC integration in to *Felis* lineage after separation from Leopard lineage. The initial result indicated that only ERV-DC genotype III was integrated into jungle cat but in a distinct clade based on phylogenetic analysis of *LTRs, gag*, *pol* and *env* genes. The integration time of ERV-DC in jungle cat was very recent suggested that invasion of ERV-DC in *Felis* genus was recently integrated compared to other endogenous retroviruses in domestic cat like endogenous feline leukemia virus.

More interestingly, I found that existence of intact full-length ERV-DC genotype III env gene in jungle cat which still retains infectious capacity. This result suggested that there may exist

infectious intact full-length ERV-DC provirus in jungle cat genome. Based on my research, I can continue to figure out the ancestral exogenous retrovirus infect jungle cat before integration of ERV-DC in this wildcat specie.

Lastly, the tracking of ancestral retroviruses provided insights into their roles in pathogenesis and host-virus evolution. Further study can continue to describe the overview of ERV-DC thought *Felis* genus in different wildcat species (66) which may contribute to both animal and viral evolution.

REFERENCES

1. **Greenwood AD, Ishida Y, O'Brien SP, Roca AL, Eiden MV.** 2018. Transmission, Evolution, and Endogenization: Lessons Learned from Recent Retroviral Invasions. Microbiology and Molecular Biology Reviews **82:**e00044-00017.
2. **Boeke JD, Stoye JP.** 1997. Retrotransponsons, endogenous retroviruses and the evolution of retroelements. *In* Coffin JM, Hughes SH, Varmus HE (ed), Retroviruses. New York, Cold Spring Harbor Laboratory Press.
3. **Hartmann K.** 2012. Clinical Aspects of Feline Retroviruses: A Review. Viruses **4:**2684-2710.
4. **Hartmann K.** 2011. Clinical aspects of feline immunodeficiency and feline leukemia virus infection. Veterinary Immunology and Immunopathology **143:**190-201.
5. **Umehara D, Watanabe S, Ochi H, Anai Y, Ahmed N, Kannagi M, Hanson C, Ruscetti S, Nishigaki K.** 2010. Role of Phosphatidylinositol 3-Kinase in Friend Spleen Focus-Forming Virus-Induced Erythroid Disease. Journal of Virology **84:**7675-7682.
6. **Hisasue M, Nagashima N, Nishigaki K, Fukuzawa I, Ura S, Katae H, Tsuchiya R, Yamada T, Hasegawa A, Tsujimoto H.** 2009. Myelodysplastic syndromes and acute myeloid leukemia in cats infected with feline leukemia virus clone33 containing a unique long terminal repeat. International Journal of Cancer **124:**1133-1141.
7. **Fujino Y, Satoh H, Hisasue M, Masuda K, Ohno K, Tsujimoto H.** 2003. Detection of the integrated feline leukemia viruses in a cat lymphoid tumor cell line by fluorescence in situ hybridization. J Hered **94:**251-255.
8. **Hisasue M, Nishigaki K, Katae H, Yuri K, Mizuno T, Fujino Y, Setoguchi A, Hasegawa A, Watari T, Masuda K, Ohno K, Tsujimoto H.** 2000. Clonality analysis of various hematopoietic disorders in cats naturally infected with feline leukemia virus. The Journal of veterinary medical science **62:**1059-1065.
9. **Kassiotis G.** 2014. Endogenous retroviruses and the development of cancer. Journal of immunology (Baltimore, Md : 1950) **192:**1343-1349.
10. **Coffin JM.** 1992. Structure and Classification of Retroviruses, p 19-49. *In* Levy JA (ed), The Retroviridae doi:10.1007/978-1-4615-3372-6_2. Springer US, Boston, MA.
11. **Varmus H.** 1988. Retroviruses. Science **240:**1427-1435.
12. **Guntaka RV.** 1993. Transcription termination and polyadenylation in retroviruses. Microbiological Reviews **57:**511-521.
13. **Hess J, Clements J, Narayan O.** 1985. cis- and trans-acting transcriptional regulation of visna virus. Science **229:**482-485.
14. **Fass D, Davey RA, Hamson CA, Kim PS, Cunningham JM, Berger JM.** 1997. Structure of a Murine Leukemia Virus Receptor-Binding Glycoprotein at 2.0 Angstrom Resolution. Science **277:**1662-1666.
15. **Barnett AL, Davey RA, Cunningham JM.** 2001. Modular organization of the Friend murine leukemia virus envelope protein underlies the mechanism of infection. Proceedings of the National Academy of Sciences **98:**4113-4118.
16. **Murakami T.** 2012. Retroviral Env Glycoprotein Trafficking and Incorporation into Virions. Molecular Biology International **2012:**11.
17. **Dubay JW, Dubay SR, Shin HJ, Hunter E.** 1995. Analysis of the cleavage site of the human immunodeficiency virus type 1 glycoprotein: requirement of precursor cleavage for glycoprotein incorporation. Journal of Virology **69:**4675-4682.

18. **Goodman LJ, Kain SR, Firestone GL.** 1993. Trafficking of wild-type and an endoproteolytic-site mutant of the mouse mammary tumor virus glycoprotein. Journal of Biological Chemistry **268:**2329-2336.
19. **Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk H-D, Garten W.** 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gpl60. Nature **360:**358-361.
20. **Dong JY, Dubay JW, Perez LG, Hunter E.** 1992. Mutations within the proteolytic cleavage site of the Rous sarcoma virus glycoprotein define a requirement for dibasic residues for intracellular cleavage. Journal of Virology **66:**865-874.
21. **Freed EO, Risser R.** 1987. The role of envelope glycoprotein processing in murine leukemia virus infection. Journal of Virology **61:**2852-2856.
22. **Zavorotinskaya T, Albritton LM.** 1999. Failure To Cleave Murine Leukemia Virus Envelope Protein Does Not Preclude Its Incorporation in Virions and Productive Virus-Receptor Interaction. Journal of Virology **73:**5621-5629.
23. **Ito J, Baba T, Kawasaki J, Nishigaki K.** 2016. Ancestral Mutations Acquired in Refrex-1, a Restriction Factor against Feline Retroviruses, during its Cooption and Domestication. Journal of Virology **90:**1470-1485.
24. **Nakaya Y, Miyazawa T.** 2014. Dysfunction of Bovine Endogenous Retrovirus K2 Envelope Glycoprotein Is Related to Unsuccessful Intracellular Trafficking. Journal of Virology **88:**6896-6905.
25. **Li Z, Pinter A, Kayman SC.** 1997. The critical N-linked glycan of murine leukemia virus envelope protein promotes both folding of the C-terminal domains of the precursor polyprotein and stability of the postcleavage envelope complex. Journal of Virology **71:**7012-7019.
26. **Apte S, Sanders DA.** 2010. Effects of retroviral envelope-protein cleavage upon trafficking, incorporation, and membrane fusion. Virology **405:**214-224.
27. **Dewannieux M, Blaise S, Heidmann T.** 2005. Identification of a Functional Envelope Protein from the HERV-K Family of Human Endogenous Retroviruses. Journal of Virology **79:**15573-15577.
28. **Taylor GM, Gao Y, Sanders DA.** 2001. Fv-4: Identification of the Defect in Env and the Mechanism of Resistance to Ecotropic Murine Leukemia Virus. Journal of Virology **75:**11244-11248.
29. **Takeda A, Matano T.** 2007. Inhibition of infectious murine leukemia virus production by Fv-4 env gene products exerting dominant negative effect on viral envelope glycoprotein. Microbes and Infection **9:**1590-1596.
30. **Pontius JU, Mullikin JC, Smith DR, Team AS, Lindblad-Toh K, Gnerre S, Clamp M, Chang J, Stephens R, Neelam B, Volfovsky N, Schäffer AA, Agarwala R, Narfström K, Murphy WJ, Giger U, Roca AL, Antunes A, Menotti-Raymond M, Yuhki N, Pecon-Slattery J, Johnson WE, Bourque G, Tesler G, Program NCS, O'Brien SJ.** 2007. Initial sequence and comparative analysis of the cat genome. Genome Research **17:**1675-1689.
31. **Mouse Genome Sequencing C.** 2002. Initial sequencing and comparative analysis of the mouse genome. Nature **420:**520.
32. **Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D,**

Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, et al. 2001. Initial sequencing and analysis of the human genome. Nature **409:**860-921.

33. **Stoye JP.** 2012. Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nature Reviews Microbiology **10:**395.
34. **Coffin JM.** 2004. Evolution of retroviruses: fossils in our DNA. Proc Am Philos Soc **148:**264-280.
35. **Xu X, Zhao H, Gong Z, Han G-Z.** 2018. Endogenous retroviruses of non-avian/mammalian vertebrates illuminate diversity and deep history of retroviruses. PLOS Pathogens **14:**e1007072.
36. **Kozak CA.** 2014. Origins of the endogenous and infectious laboratory mouse gammaretroviruses. Viruses **7:**1-26.
37. **Oliveira NM, Satija H, Kouwenhoven IA, Eiden MV.** 2007. Changes in viral protein function that accompany retroviral endogenization. Proceedings of the National Academy of Sciences **104:**17506-17511.
38. **Lee YN, Bieniasz PD.** 2007. Reconstitution of an Infectious Human Endogenous Retrovirus. PLOS Pathogens **3:**e10.
39. **Ghosh AK, Roy-Burman P.** 1989. Characterization of enhancer elements and their mutations in the long terminal repeat of feline endogenous RD-114 proviruses. Journal of Virology **63:**4234-4241.
40. **Robbez-Masson L, Rowe HM.** 2015. Retrotransposons shape species-specific embryonic stem cell gene expression. Retrovirology **12:**45.
41. **Slotkin RK, Martienssen R.** 2007. Transposable elements and the epigenetic regulation of the genome. Nature Reviews Genetics **8:**272.
42. **Kuse K, Ito J, Miyake A, Kawasaki J, Watanabe S, Makundi I, Ngo MH, Otoi T, Nishigaki K.** 2016. Existence of Two Distinct Infectious Endogenous Retroviruses in Domestic Cats and Their Different Strategies for Adaptation to Transcriptional Regulation. Journal of Virology **90:**9029-9045.
43. **Weiss RA.** 2006. The discovery of endogenous retroviruses. Retrovirology **3:**67.
44. **Tarlinton RE, Meers J, Young PR.** 2006. Retroviral invasion of the koala genome. Nature **442:**79-81.
45. **Hanger JJ, Bromham LD, McKee JJ, O'Brien TM, Robinson WF.** 2000. The Nucleotide Sequence of Koala (Phascolarctos cinereus) Retrovirus: a Novel Type C Endogenous Virus Related to Gibbon Ape Leukemia Virus. Journal of Virology **74:**4264-4272.
46. **Wilson CA, Wong S, Muller J, Davidson CE, Rose TM, Burd P.** 1998. Type C Retrovirus Released from Porcine Primary Peripheral Blood Mononuclear Cells Infects Human Cells. Journal of Virology **72:**3082-3087.
47. **Le Tissier P, Stoye JP, Takeuchi Y, Patience C, Weiss RA.** 1997. Two sets of human-tropic pig retrovirus. Nature **389:**681-682.
48. **Todaro GJ, Benveniste RE, Lieber MM, Sherr CJ.** 1974. Characterization of a type C virus released from the porcine cell line PK(15). Virology **58:**65-74.
49. **Moennig V, Frank H, Hunsmann G, Ohms P, Schwarz H, Schäfer W, Strandström H.** 1974. C-type particles produced by a permanent cell line from a leukemic pig: II. Physical, chemical, and serological characterization of the particles. Virology **57:**179-188.

50. **Armstrong JA, Porterfield JS, de Madrid AT.** 1971. C-type Virus Particles in Pig Kidney Cell Lines. Journal of General Virology **10:**195-198.
51. **Anai Y, Ochi H, Watanabe S, Nakagawa S, Kawamura M, Gojobori T, Nishigaki K.** 2012. Infectious Endogenous Retroviruses in Cats and Emergence of Recombinant Viruses. Journal of Virology **86:**8634-8644.
52. **Fábryová H, Hron T, Kabíčková H, Poss M, Elleder D.** 2015. Induction and characterization of a replication competent cervid endogenous gammaretrovirus (CrERV) from mule deer cells. Virology **485:**96-103.
53. **Aaronson SA, Tronick SR, Stephenson JR.** 1976. Endogenous type C RNA virus of odocoileus hemionus, a mammalian species of New World origin. Cell **9:**489-494.
54. **Malfavon-Borja R, Feschotte C.** 2015. Fighting Fire with Fire: Endogenous Retrovirus Envelopes as Restriction Factors. Journal of Virology **89:**4047-4050.
55. **Christian L, Guillaume C, Anne D, Cécile E, Odile H, Cécile V, Thierry H.** 2013. Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. Philosophical Transactions of the Royal Society B: Biological Sciences **368:**20120507.
56. **Redelsperger F, Raddi N, Bacquin A, Vernochet C, Mariot V, Gache V, Blanchard-Gutton N, Charrin S, Tiret L, Dumonceaux J, Dupressoir A, Heidmann T.** 2016. Genetic Evidence That Captured Retroviral Envelope syncytins Contribute to Myoblast Fusion and Muscle Sexual Dimorphism in Mice. PLOS Genetics **12:**e1006289.
57. **Pastuzyn ED, Day CE, Kearns RB, Kyrke-Smith M, Taibi AV, McCormick J, Yoder N, Belnap DM, Erlendsson S, Morado DR, Briggs JAG, Feschotte C, Shepherd JD.** 2018. The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. Cell **172:**275-288.e218.
58. **Ashley J, Cordy B, Lucia D, Fradkin LG, Budnik V, Thomson T.** 2018. Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. Cell **172:**262-274.e211.
59. **Chuong EB, Elde NC, Feschotte C.** 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science **351:**1083.
60. **Heidmann O, Béguin A, Paternina J, Berthier R, Deloger M, Bawa O, Heidmann T.** 2017. HEMO, an ancestral endogenous retroviral envelope protein shed in the blood of pregnant women and expressed in pluripotent stem cells and tumors. Proceedings of the National Academy of Sciences doi:10.1073/pnas.1702204114**:**201702204.
61. **Beyer W, Mohring R, Drescher B, Nötzel U, Rosenthal S.** 1987. Molecular cloning of an endogenous cat retroviral element (ECE1)—a recombinant between RD-114 and FeLV-related sequences. Archives of Virology **96:**297-301.
62. **van der Kuyl AC, Dekker JT, Goudsmit J.** 1999. Discovery of a new endogenous type C retrovirus (FcEV) in cats: evidence for RD-114 being an FcEV(Gag-Pol)/baboon endogenous virus BaEV(Env) recombinant. J Virol **73**.
63. **Kawasaki J, Nishigaki K.** 2018. Tracking the Continuous Evolutionary Processes of an Endogenous Retrovirus of the Domestic Cat: ERV-DC. Viruses **10:**179.
64. **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. Journal of Virology **87:**12029-12040.
65. **Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ, Macdonald DW.** 2007. The Near Eastern Origin of Cat Domestication. Science **317:**519-523.

66. **Johnson WE, Eizirik E, Pecon-Slattery J, Murphy WJ, Antunes A, Teeling E, O'Brien SJ.** 2006. The Late Miocene Radiation of Modern Felidae: A Genetic Assessment. Science **311:**73-77.
67. **Vigne J-D, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J.** 2012. First wave of cultivators spread to Cyprus at least 10,600 y ago. Proceedings of the National Academy of Sciences **109:**8445-8449.
68. **Vigne J-D, Guilaine J, Debue K, Haye L, Gérard P.** 2004. Early Taming of the Cat in Cyprus. Science **304:**259-259.
69. **Hubbard AL, McOris S, Jones TW, Boid R, Scott R, Easterbee N.** 1992. Is survival of European wildcats Felis silvestris in Britain threatened by interbreeding with domestic cats? Biological Conservation **61:**203-208.
70. **Shimode S, Nakagawa S, Miyazawa T.** 2015. Multiple invasions of an infectious retrovirus in cat genomes. Scientific Reports **5:**8164.
71. **Dimmic MW, Rest JS, Mindell DP, Goldstein RA.** 2002. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. J Mol Evol **55:**65-73.
72. **Jarrett WF, Crawford EM, Martin WB, Davie F.** 1964. A virus-like particle associated with leukemia (lymphosarcoma). Nature **202:**567-569.
73. **Chinwalla AT, Cook LL, Delehaunty KD, Fewell GA, Fulton LA, Fulton RS, Graves TA, Hillier LW, Mardis ER, McPherson JD, Miner TL, Nash WE, Nelson JO, Nhan MN, Pepin KH, Pohl CS, Ponce TC, Schultz B, Thompson J, Trevaskis E, Waterston RH, Wendl MC, Wilson RK, Yang S-P, An P, Berry E, Birren B, Bloom T, Brown DG, Butler J, Daly M, David R, Deri J, Dodge S, Foley K, Gage D, Gnerre S, Holzer T, Jaffe DB, Kamal M, Karlsson EK, Kells C, Kirby A, Kulbokas EJ, Lander ES, Landers T, Leger JP, Levine R, Lindblad-Toh K, Mauceli E, et al.** 2002. Initial sequencing and comparative analysis of the mouse genome. Nature **420:**520-562.
74. **Vigne J-D, Evin A, Cucchi T, Dai L, Yu C, Hu S, Soulages N, Wang W, Sun Z, Gao J, Dobney K, Yuan J.** 2016. Earliest "Domestic" Cats in China Identified as Leopard Cat (Prionailurus bengalensis). PLOS ONE **11:**e0147295.
75. **Hu Y, Hu S, Wang W, Wu X, Marshall FB, Chen X, Hou L, Wang C.** 2014. Earliest evidence for commensal processes of cat domestication. Proceedings of the National Academy of Sciences **111:**116-120.
76. **Bar-Oz G, Weissbrod L, Tsahar E.** 2014. Cats in recent Chinese study on cat domestication are commensal, not domesticated. Proceedings of the National Academy of Sciences **111:**E876-E876.
77. **Linseele V, Van Neer W, Hendrickx S.** 2007. Evidence for early cat taming in Egypt. Journal of Archaeological Science **34:**2081-2090.
78. **Driscoll CA, Clutton-Brock J, Kitchener AC, O'Brien SJ.** 2009. The taming of the cat. Scientific American **300:**68-75.
79. **Vella CM, Shelton LM, McGonagle JJ, Stanglein TW.** 1999. Robinson's genetics for cat breeders and veterinarians. Butterworth-Heinemann Ltd.
80. **Salonen M, Vapalahti K, Tiira K, Mäki-Tanila A, Lohi H.** 2019. Breed differences of heritable behaviour traits in cats. Scientific Reports **9:**7949.
81. **David VA, Menotti-Raymond M, Wallace AC, Roelke M, Kehler J, Leighty R, Eizirik E, Hannah SS, Nelson G, Schaffer AA, Connelly CJ, O'Brien SJ, Ryugo**

DK. 2014. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. G3 (Bethesda) **4:**1881-1891.
82. **Polani S, Roca AL, Rosensteel BB, Kolokotronis SO, Bar-Gal GK.** 2010. Evolutionary dynamics of endogenous feline leukemia virus proliferation among species of the domestic cat lineage. Virology **405:**397-407.
83. **Roca AL, Nash WG, Menninger JC, Murphy WJ, O'Brien SJ.** 2005. Insertional Polymorphisms of Endogenous Feline Leukemia Viruses. Journal of Virology **79:**3979-3986.
84. **Roca AL, Pecon-Slattery J, O'Brien SJ.** 2004. Genomically Intact Endogenous Feline Leukemia Viruses of Recent Origin. Journal of Virology **78:**4370-4375.
85. **Niman HL, Akhavi M, Gardner MB, Stephenson JR, Roy-Burman P.** 1980. Differential expression of two distinct endogenous retrovisus genomes in developing tissues of the domestic cat. J Natl Cancer Inst **64:**587-594.
86. **Koshy R, Gallo RC, Wong-Staal F.** 1980. Characterization of the endogenous feline leukemia virus-related DNA sequences in cats and attempts to identify exogenous viral sequences in tissues of virus-negative leukemic animals. Virology **103:**434-445.
87. **Benveniste RE, Sherr CJ, Todaro GJ.** 1975. Evolution of type C viral genes: origin of feline leukemia virus. Science **190:**886-888.
88. **McDougall AS, Terry A, Tzavaras T, Cheney C, Rojko J, Neil JC.** 1994. Defective endogenous proviruses are expressed in feline lymphoid cells: evidence for a role in natural resistance to subgroup B feline leukemia viruses. J Virol **68:**2151-2160.
89. **Stewart MA, Warnock M, Wheeler A, Wilkie N, Mullins JI, Onions DE, Neil JC.** 1986. Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. Journal of Virology **58:**825-834.
90. **Watanabe S, Kawamura M, Odahara Y, Anai Y, Ochi H, Nakagawa S, Endo Y, Tsujimoto H, Nishigaki K.** 2013. Phylogenetic and Structural Diversity in the Feline Leukemia Virus Env Gene. PLOS ONE **8:**e61009.
91. **Ngo MH, Arnal M, Sumi R, Kawasaki J, Miyake A, Grant CK, Otoi T, Luco DFD, Nishigaki K.** 2019. Tracking the fate of endogenous retrovirus segregation in wild and domestic cats. Journal of Virology **In press**.
92. **Sambrook J, Russell DW.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
93. **Kurushima JD, Lipinski MJ, Gandolfi B, Froenicke L, Grahn JC, Grahn RA, Lyons LA.** 2013. Variation of cats under domestication: genetic assignment of domestic cats to breeds and worldwide random-bred populations. Animal Genetics **44:**311-324.
94. **Menotti-Raymond M, David VA, Pflueger SM, Lindblad-Toh K, Wade CM, O'Brien SJ, Johnson WE.** 2008. Patterns of molecular genetic variation among cat breeds. Genomics **91:**1-11.
95. **Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, Pedersen NC, Lyons LA.** 2008. The ascent of cat breeds: Genetic evaluations of breeds and worldwide random-bred populations. Genomics **91:**12-21.
96. **Ngo MH, Arnal M, Sumi R, Kawasaki J, Miyake A, Grant CK, Otoi T, Fernández de Luco D, Nishigaki K.** 2019. Tracking the fate of endogenous retrovirus segregation in wild and domestic cats. Journal of Virology doi:10.1128/jvi.01324-19**:**JVI.01324-01319.

97. **Duckworth JW, Poole CM, Tizard RJ, Walston JL, Timmins RJ.** 2005. The Jungle Cat Felis chaus in Indochina: a threatened population of a widespread and adaptable species. Biodiversity & Conservation **14:**1263-1280.
98. **Palmarini M, Mura M, Spencer TE.** 2004. Endogenous betaretroviruses of sheep: teaching new lessons in retroviral interference and adaptation. J Gen Virol **85**.
99. **Cornelis G, Heidmann O, Bernard-Stoecklin S, Reynaud K, Véron G, Mulot B, Dupressoir A, Heidmann T.** 2012. Ancestral capture of syncytin-Car1, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. Proceedings of the National Academy of Sciences **109:**E432–E441.
100. **Donahue PR, Hoover EA, Beltz GA, Riedel N, Hirsch VM, Overbaugh J, Mullins JI.** 1988. Strong sequence conservation among horizontally transmissible, minimally pathogenic feline leukemia viruses. J Virol **62:**722-731.
101. **Pandey R, Ghosh AK, Kumar DV, Bachman BA, Shibata D, Roy-Burman P.** 1991. Recombination between feline leukemia virus subgroup B or C and endogenous env elements alters the in vitro biological activities of the viruses. Journal of Virology **65:**6495-6508.
102. **Overbaugh J, Riedel N, Hoover EA, Mullins JI.** 1988. Transduction of endogenous envelope genes by feline leukaemia virus in vitro. Nature **332:**731-734.
103. **Young GR, Eksmond U, Salcedo R, Alexopoulou L, Stoye JP, Kassiotis G.** 2012. Resurrection of endogenous retroviruses in antibody-deficient mice. Nature **491:**774-778.
104. **Stoye JP, Moroni C, Coffin JM.** 1991. Virological events leading to spontaneous AKR thymomas. Journal of Virology **65:**1273-1285.
105. **Song N, Jo H, Choi M, Kim JH, Seo HG, Cha SY, Seo K, Park C.** 2013. Identification and classification of feline endogenous retroviruses in the cat genome using degenerate PCR and in silico data analysis. J Gen Virol **94:**1587-1596.
106. **Mata H, Gongora J, Eizirik E, Alves BM, Soares MA, Ravazzolo AP.** 2015. Identification and characterization of diverse groups of endogenous retroviruses in felids. Retrovirology **12:**26.
107. **Weiss RA.** 2013. On the concept and elucidation of endogenous retroviruses. Philosophical Transactions of the Royal Society B: Biological Sciences **368:**20120494.
108. **Boomer S, Eiden M, Burns CC, Overbaugh J.** 1997. Three distinct envelope domains, variably present in subgroup B feline leukemia virus recombinants, mediate Pit1 and Pit2 receptor recognition. Journal of Virology **71:**8116-8123.
109. **McAllister RM, Nicolson M, Gardner MB, Rongey RW, Rasheed S, Sarma PS, Huebner RJ, Hatanaka M, Oroszlan S, Gilden RV, Kabigting A, Vernon L.** 1972. C-type virus released from cultured human rhabdomyosarcoma cells. Nat New Biol **235**.
110. **Bonner TI, Todaro GJ.** 1979. Carnivores have sequences in their cellular DNA distantly related to the primate endogenous virus, MAC-1. Virology **94:**224-227.
111. **Todaro GJ, Benveniste RE, Sherwin SA, Sherr CJ.** 1978. MAC-1, a new genetically transmitted type C virus of primates: "low frequency" activation from stumptail monkey cell cultures. Cell **13:**775-782.
112. **Kawasaki J, Kawamura M, Ohsato Y, Ito J, Nishigaki K.** 2017. Presence of a Shared 5′ -Leader Sequence in Ancestral Human and Mammalian Retroviruses and Its Transduction into Feline Leukemia Virus. Journal of Virology **91:**e00829-00817.
113. **Ikeda H, Sugimura H.** 1989. Fv-4 resistance gene: a truncated endogenous murine leukemia virus with ecotropic interference properties. Journal of Virology **63:**5405-5412.

114. **Ikeda H, Laigret F, Martin MA, Repaske R.** 1985. Characterization of a molecularly cloned retroviral sequence associated with Fv-4 resistance. Journal of Virology **55:**768-777.
115. **Wu T, Yan Y, Kozak CA.** 2005. Rmcf2, a Xenotropic Provirus in the Asian Mouse Species Mus castaneus, Blocks Infection by Polytropic Mouse Gammaretroviruses. Journal of Virology **79:**9677-9684.
116. **Jung YT, Lyu MS, Buckler-White A, Kozak CA.** 2002. Characterization of a Polytropic Murine Leukemia Virus Proviral Sequence Associated with the Virus Resistance Gene Rmcf of DBA/2 Mice. Journal of Virology **76:**8218-8224.
117. **Bamunusinghe D, Naghashfar Z, Buckler-White A, Plishka R, Baliji S, Liu Q, Kassner J, Oler AJ, Hartley J, Kozak CA.** 2016. Sequence Diversity, Intersubgroup Relationships, and Origins of the Mouse Leukemia Gammaretroviruses of Laboratory and Wild Mice. Journal of Virology **90:**4186-4198.
118. **Nowell K, Jackson P.** 1996. European wildcat, Felis silvestris, silvestris group Schreber, 1775. Wild cats: status survey and conservation action plan IUCN, Gland, Switzerland**:**110-113.
119. **Hemmer H.** 1978. The evolutionary systematics of living Felidae: present status and current problems. Carnivora **1:**71-79.
120. **Guggisberg CAW.** 1975. Wild cats of the world. David & Charles.
121. **Driscoll CA, Macdonald DW, O'Brien SJ.** 2009. From wild animals to domestic pets, an evolutionary view of domestication. Proceedings of the National Academy of Sciences **106:**9971-9978.
122. **Graham FL, Smiley J, Russell WC, Nairn R.** 1977. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. Journal of General Virology **36:**59-72.
123. **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. Journal of Virology **90:**4832-4837.
124. **Kimura M.** 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution **16:**111-120.
125. **Kumar S, Stecher G, Li M, Knyaz C, Tamura K.** 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol **35:**1547-1549.
126. **Katoh K, Standley DM.** 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution **30:**772-780.
127. **Hasegawa M, Kishino H, Yano T.** 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution **22:**160-174.
128. **Mattucci F, Oliveira R, Bizzarri L, Vercillo F, Anile S, Ragni B, Lapini L, Sforzi A, Alves PC, Lyons LA, Randi E.** 2013. Genetic structure of wildcat (Felis silvestris) populations in Italy. Ecology and Evolution **3:**2443-2458.
129. **Hertwig S, Schweizer M, Stepanow S, Jungnickel A, Böhle UR, Fischer M.** 2009. Regionally high rates of hybridization and introgression in German wildcat populations (Felis silvestris, Carnivora, Felidae). Journal of Zoological Systematics and Evolutionary Research **47:**283-297.
130. **Koito A, Ikeda T.** 2011. Intrinsic restriction activity by AID/APOBEC family of enzymes against the mobility of retroelements. Mobile Genetic Elements **1:**197-202.

131. **LaRue RS, Andrésdóttir V, Blanchard Y, Conticello SG, Derse D, Emerman M, Greene WC, Jónsson SR, Landau NR, Löchelt M, Malik HS, Malim MH, Münk C, O'Brien SJ, Pathak VK, Strebel K, Wain-Hobson S, Yu X-F, Yuhki N, Harris RS.** 2009. Guidelines for Naming Nonprimate APOBEC3 Genes and Proteins. Journal of Virology **83:**494-497.
132. **Harris RS, Liddament MT.** 2004. Retroviral restriction by APOBEC proteins. Nature Reviews Immunology **4:**868.
133. **Gifford WD, Pfaff SL, Macfarlan TS.** 2013. Transposable elements as genetic regulatory substrates in early development. Trends in Cell Biology **23:**218-226.
134. **Cowley M, Oakey RJ.** 2013. Transposable Elements Re-Wire and Fine-Tune the Transcriptome. PLOS Genetics **9:**e1003234.
135. **Isbel L, Whitelaw E.** 2012. Endogenous retroviruses in mammals: An emerging picture of how ERVs modify expression of adjacent genes. BioEssays **34:**734-738.
136. **Fiebig U, Hartmann MG, Bannert N, Kurth R, Denner J.** 2006. Transspecies Transmission of the Endogenous Koala Retrovirus. Journal of Virology **80:**5651-5654.
137. **Tarlinton R, Meers J, Hanger J, Young P.** 2005. Real-time reverse transcriptase PCR for the endogenous koala retrovirus reveals an association between plasma viral load and neoplastic disease in koalas. Journal of General Virology **86:**783-787.
138. **Tacke SJ, Kurth R, Denner J.** 2000. Porcine Endogenous Retroviruses Inhibit Human Immune Cell Function: Risk for Xenotransplantation? Virology **268:**87-93.
139. **Baba T, Ngo MH, Nishigaki K.** 2016. Analysis of ERV-DC in *Felis Chaus*. (unpublished thesis). Doctor of Veterinary Medicine. Yamaguchi, Yamaguchi University.
140. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research **32:**1792-1797.
141. **Kumar S, Stecher G, Tamura K.** 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution **33:**1870-1874.
142. **Lopez JV, Culver M, Stephens JC, Johnson WE, O'Brien SJ.** 1997. Rates of nuclear and cytoplasmic mitochondrial DNA sequence divergence in mammals. Molecular Biology and Evolution **14:**277-286.
143. **Yoshinaka Y, Katoh I, Copeland TD, Oroszlan S.** 1985. Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. Proceedings of the National Academy of Sciences **82:**1618-1622.
144. **Johnson WE.** 2019. Origins and evolutionary consequences of ancient endogenous retroviruses. Nature Reviews Microbiology **17:**355-370.