2017

EVOLUTION OF EQUINE ARTERITIS VIRUS DURING PERSISTENT INFECTION IN THE REPRODUCTIVE TRACT OF THE STALLION AND THE MALE DONKEY

Bora Nam
University of Kentucky, bna222@g.uky.edu
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.499

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Bora Nam, Student

Dr. Udeni B. R. Balasuriya, Major Professor

Dr. Daniel Howe, Director of Graduate Studies

EVOLUTION OF EQUINE ARTERITIS VIRUS DURING PERSISTENT INFECTION IN THE REPRODUCTIVE TRACT OF THE STALLION AND THE MALE DONKEY

THESIS

A thesis submitted in partial fulfillment of the
Requirements for the degree of Master of Science in the
College of Agriculture, Food and Environment
at the University of Kentucky

By

Bora Nam

Lexington, Kentucky

Director: Dr. Udeni B. R. Balasuriya, Professor of Virology

Lexington, Kentucky

2017

ABSTRACT OF THESIS

EVOLUTION OF EQUINE ARTERITIS VIRUS DURING PERSISTENT INFECTION IN THE REPRODUCTIVE TRACT OF THE STALLION AND THE MALE DONKEY

Equine arteritis virus (EAV) establishes persistent infection in the stallion reproductive tract, and the carrier stallion continues to shed virus in semen for weeks to years or lifelong. The objective of this study was to elucidate the intra-host evolution of EAV during persistent infection in stallions. Seven EAV seronegative stallions were experimentally infected with EAV KY84 strain and followed for 726 days post-infection, and sequential clinical samples including semen were collected for virus isolation and next-generation sequencing (NGS). In addition, archived sequential semen samples from two stallions that were naturally infected with EAV KY84 for a long-period (up to 10 years) were also sequenced by NGS. The data demonstrated genetic bottleneck event and selection during acute infection followed by intra-host quasispecies diversification during persistent infection in the stallion reproductive tract.

Also, the full-length genome of a novel EAV donkey strain from Chile and a noncytopathic bovine viral diarrhea virus-1 (ncpBVDV-1) strain contaminating rabbit kidney-13 cells were also sequenced by NGS. The EAV donkey strain was genetically distinct but antigenically cross-reacted with EAV antisera, and it was phylogenetically closely related to the South African donkey strain of EAV. Genetic and phylogenetic analyses demonstrated that ncpBVDV-1 belongs to BVDV-1b group.

KEYWORDS: Equine arteritis virus, equine viral arteritis, persistent infection, intra-host evolution, quasispecies, bottleneck event

Bora Nam

December 8, 2017

EVOLUTION OF EQUINE ARTERITIS VIRUS DURING PERSISTENT INFECTION IN THE REPRODUCTIVE TRACT OF THE STALLION AND THE MALE DONKEY

By

Bora Nam

Udeni B.R. Balasuriya
Director of Thesis

Daniel Howe
Director of Graduate Studies

December 8, 2017

DEDICATION

To God for His unconditional love and mercy upon me.
To my parents and American host families for their support and belief in me.
To my friends for their encouragement.
To my cat Leo for always standing by me.
To the horses which sacrificed for my research.

ACKNOWLEDGEMENTS

When I arrived Kentucky in 2009 as a pre-vet student, I walked on the campus by myself and the final destination was the Maxwell H. Gluck Equine Research Center. At that time, I whispered in my heart saying "One day, I will work here." And the dream came true, except I did not expect to stay here for eight years. Mr. Gong Seoul, a previous undergraduate research assistant, introduced me to Dr. Udeni B. R. Balasuriya's molecular virology laboratory right before he left the laboratory after completing his undergraduate studies. Mr. Seoul turned over all the general laboratory maintenance and management work such as making cell culture medium. During eight years of working as an undergraduate research assistant, a senior laboratory technician, and a graduate research assistant, I had a great and happy time working with laboratory members involved in several research projects. I felt a value to me that I was able to support the laboratory. It is incredible how the relationship influenced my life and that was how I put my first step into Gluck. The completion of my Master's degree would not have been possible without the support, patience, belief, and encouragement of many people that I have met and worked with the past. It is my honor to acknowledge everyone to give my most profound sincere appreciation.

I give all my sincere heart and most profound appreciation to Dr. Udeni B. R. Balasuriya for his abundant support, supervision, and patience through my Master's degree program. He opened the door for me to join his laboratory and taught me classical virology and molecular biology techniques (virus isolation, virus titration, nucleic acid extraction, standard and real-time RT-PCR and sequencing). During the Masters program, I have obtained a fruitful understanding of equine arteritis virus, critical thinking skills

utilized in the research and the charm out of unlimited observations in the research field by looking over Dr. Balasuriya's ambition, passion and brilliant ideas on the research project. Not only is Dr. Balasuriya my advising professor, but he is also a great mentor, friend and sometimes like a father to me. He has shown patience, generosity and thoughtful consideration of myself and helped me to complete the degree and to move forward.

I am most grateful to Dr. Peter J. Timoney for his continuous trust, belief, and encouragement. Whenever I fell, he grabbed my hand to stand up and still walk through the steps of the Master's program. He always believed in my talent, strength, and value and encouraged me to be motivated. He is one of the most beautiful gentlemen I have met; with humongous cheerful energy, he is the role model that I would like to be. His broad knowledge covering epidemiology, infectious disease and sincere care for equine health and well-being in the horse industry will continuously motivate me to seek and work related to it for the horse industry.

I would like to express my sincere appreciation to Dr. R. Frank Cook for accepting being part of my Masters degree committee. Within a limited period, he took a closer look at my research projects and provided me with helpful discussion and valuable input. I give my special thanks to Dr. Daniel K. Howe, the Director of Graduate Studies, for always providing me his time, thoughtful guidance and good notice while going through the Master's program. Without his consideration, I would not have been able to complete the degree by this time.

I would like to show my gratitude to Dr. Ganwu Li and Ms. Ying Zheng at the Department of Veterinary Diagnostic and Production Animal Medicine, College of

Veterinary Medicine, Iowa State University for their excellent collaboration to perform next-generation sequencing of equine arteritis virus samples and allow me to visit their laboratory to learn the complex bioinformatics analysis, to set up my computer and to perform the command line coding method.

I owe my most profound gratitude to Dr. Ernest Bailey for teaching me how much fun equine genetics is and reminding me that I can be a mad scientist. His joy, encouragement, and trust always cheered me up. I give my sincere gratitude to Dr. Eugene Lyons for a great friendship. He was the first person who took me over to necropsy to collect a parasite specimen. This is a great memory that I held my breath until the necropsy was over, but then screamed and cried after I saw a big cockroach in the laboratory. I express special gratitude to Dr. Thomas W. Swerczek for sharing his story of the research foundation on Woundade and in-depth consideration in equine health. Also, I highly acknowledge both Drs. Swerczek and William McCollum, who isolated the KY84 strain of equine arteritis virus for my primary research project.

My admiration to Dr. Yun Young Go, currently, a senior researcher at the Korea Research Institute of Chemical Technology, is enormous; she was my role model as the equine veterinarian and the brilliant scientist from South Korea. After watching her hard work, motivation, and enthusiasm over the research, I stepped into the Masters program. She was the person who taught me how to do cell culture and Dr. Jianqiang Zhang, currently an associate professor at Iowa State University, allowed me to pipet aliquots for the first time in the hood. That was how I gradually enter the bench works in the laboratory.

I must highly acknowledge Dr. Zelalem Mekuria that I owe my deepest gratitude to for his time, great consideration and abundant input on my primary research project (chapter 2). He is a bright budding scientist full of curiosity and highly motivated. With his significant investment, my central research project concluded wonderfully, and this research project should be continued to the next level.

I express my special appreciation to Dr. Wangisa Dunuwille for his dedicated help when I was very fortunate to take the crazy nightmare integrated biomedical science (IBS) courses. Without him, I would never have survived it. Although it was the most challenging time for me to learn the fundamental knowledge in cell biology, molecular biology, genetics, virology, immunology and biochemistry, it was a valuable time to learn in depth these sciences, which help me a better understand my research projects.

I gratefully thank Dr. Mariano Carossino for a great friendship. He is a brilliant, sharp, delightful and sophisticated person that I always enjoyed spending spare time with to make it easier during the intensive Masters study. I especially owe my gratitude for taking care of Leo, my cat, while I was out of town. Also, I thank him for teaching me the serum neutralization test and providing me with his thoughtful opinions on my primary research project related to determining the distinctive neutralization phenotype of the virus.

I give my heartfelt and deepest gratitude to my American grandmother Ms. Pam Henney for her unending support, love, trust, and encouragement. It was challenging for me coming all the way from South Korea to study abroad by myself. Without her love and support, I would never have been able to complete my education and move forward. I

promise myself that I will pass along all my help to the person who needs it from what I received. Also, I will continuously listen to what my heart is saying and follow it.

I am grateful to Ms. Kathy Shuck who patiently corrected and proofread my writing for improvement. Whenever I had questions, she always provided me with the right answers. Her kindness and cheerful encouragement helped me to grow more. I was very fortunate to be surrounded by good people especially Dr. Thomas Chambers, Stephanie Reedy, Michelle Yeargan and Sheila Cook on the fourth floor at the Gluck Center. I will miss the unique birthday parties on the fourth floor, and I appreciate everyone sharing their care and kindness. Many thanks to the previous and current laboratory members, Drs. Yanqiu Li, Lakshman Chelvarajan, Zhengchun Lu, Juliana Campos, Kathryn Smith, Sanjay Sarkar and Shankar Mondal; Ms. Kristin Chu, Ashley Skillman, Annet Kyomuhangi, Tessa Teardo, Cecilia Winfrey and Youjin Hwang; and Mr. Casey Edwards, Chase Tucker and Zach Willand. Also thanks to Drs. Sergey Artiushin, Karen McDowell and Fatai Oladunni; Ms. Karin Davies and Doris Gabbard and Mr. Brock Daugherty for good friendships. I would like to thank all of the administrative staff of the Department of Veterinary Science, especially Ms. Diane Furry, Gail Watkins, Debbie Mollett, Patsy Garrett and Jenny Evans. Not only was I being a scientist, but I was also being an accountant, scheduler, and shipper and played many other roles, and it was always a pleasure to work with them.

I express my special gratitude to Ms. Kimberly A. Bennett at the UK Student Healthcare. She was an amazing nurse practitioner to me and took very good care of my health and mind whenever I had problems going through this stressful time. I gratefully thank my former advisors, Drs. Mary Rossano, Roberta Dwyer and Larry Grabau and Ms.

Susan Skees from the UK College of Agriculture. Without their support, belief, and encouragement, I would not be here right now. Also, I give my special thanks to Dr. Michael Goodin at the Department of Plant Pathology for a good friendship with excellent coffee time.

I gratefully acknowledge the funding sources for my study and personal financial stipend, Hildegard Rose Shapiro Endowment and the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA) grant no. 2013-68004-20360.

I owe my deepest sincere gratitude for faithful friendship especially to Ms. Yujung Lee, Hyemi Lee, Hyeyun Min, Nayon Kang, Heejoo Kim and Yoonie Choi whom I got to know in Lexington and became a part of my life. Without their friendship, support, and care, I would never be able to go through all the hardship during eight years of studying abroad in Kentucky. At each step, I grew stronger by grabbing their hands. I express a special thank you to my friends of more than 10 years, Ms. Myeongeun Lee (elementary school friend), Halim Michelle Song (National Research Council Canada; Yang Young Middle School friend) and Minkyung Bae (Lehigh Valley Christian High School friend). Whenever I had a hard time, they were always on the phone to me giving excellent life advice (known as seven-up sparkling [TOK-SSO-NUN-CIDER] advice), everlasting cheers and support so that I was able to stand up and continuously move forward. My time in Lexington, Kentucky was enjoyable with UK 1990 girls, Ms. Heejin Yang, Yeseul Kim and Jiyoung Lee. I gratefully thank friends from the Jeja Church, especially Dr. Jun Ho Seok, Mr. and Mrs. Moon and Ms. Na-Ra Lee for their support and friendship in God.

Lastly, I would like to thank my parents, Min Woo Nam and Jin Myung Park for their unconditional love, trust, and encouragement. My uncle, Dr. Seung Rib Park, deserves special thanks for showing me the joy of intellectual pursuit and caring. A special thank you to Leo, my cat, who was always next to me during the past eight years. Also, I wholeheartedly thank my American host families, the Sterling and Carolyn Buss family (Christy, Jeremy and Lauren) and the Debra Tirrell-Yeager family (Paul James, Jonathan, Timothy and Katie Lynn Hutchison) who always taught me family love in God and made sure I felt comfortable during my stay in Pennsylvania. I truly believe that I was very fortunate and God blessed me to have everyone as part of my life. Thank you so much and I love you all!

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES xiv

LIST OF FIGURES xvi

CHAPTER ONE - Literature Review 1

1. Introduction 1

1.1. Equine arteritis virus (EAV) 3

1.1.1. Classification and genome organization 3

1.1.2. Nonstructural proteins of EAV 7

1.1.3. Structural proteins of EAV 8

1.1.3.1. Minor envelope proteins 8

1.1.3.2. Major envelope proteins 10

1.2. The replication cycle of EAV 12

1.2.1. Virus attachment and entry 13

1.2.2. Post-translation in processing of replicase polyproteins 16

1.2.3. Genome replication 19

1.2.4. Viral sg mRNA transcription 21

1.2.5. Virus assembly, budding and release 24

1.3. Equine viral arteritis 26

1.3.1. Epidemiology 26

1.3.1.1. Transmission 27

1.3.1.2. Immune response to EAV infection 28

1.3.2. Clinical signs 29

1.3.3. Pathogenesis 31
1.3.4. Persistent EAV infection 35
1.4. Diagnosis of EAV infection 36
1.4.1. Appropriate specimen collection for examination 36
1.4.2. Virus isolation in cell culture 37
1.4.3. Molecular diagnosis (nucleic acid detection by RT-PCR) 38
1.4.4. Antigen visualization by immunohistochemistry 39
1.4.5. Serological diagnosis 39
1.4.6. Differential diagnosis 41
1.5. Treatment of EVA 41
1.5.1. Antiviral therapy 41
1.5.2. Anti-GnRH vaccination and GnRH antagonists 42
1.6. Disease prevention and control 42
1.6.1. Vaccine 45
2. Evolution of RNA viruses 48
2.1. Mechanisms of virus evolution 50
2.1.1. Accumulation of mutations 50
2.1.1.1. Concept of viral quasispecies 50
2.1.1.2. Major characteristics of viral quasispecies 51
2.1.1.3. Biological outcomes of quasispecies dynamics 55
2.1.2. Reassortment 60
2.1.3. Recombination 61
3. Research objectives 61

CHAPTER TWO - Genetic Bottleneck and Selection of Equine Arteritis Virus During Acute Infection and Intra-host Quasispecies Diversification During Long-term Persistent Infection in the Reproductive Tract of the Stallion 65

2.1. Summary 65

2.2. Introduction 66

2.3. Materials and methods 70

2.4. Results 81

2.5. Discussion 123

CHAPTER THREE - Genomic, Phylogenetic and Antigenic Characterization of a Novel Field Strain of Equine Arteritis Virus Isolated from a Feral Donkey in Chile 130

3.1. Summary 130

3.2. Introduction 130

3.3. Materials and methods 132

3.4. Results 138

3.5. Discussion 153

CHAPTER FOUR - Complete Genome Sequence of a Noncytopathic Strain of Bovine Viral Diarrhea Virus 1 (BVDV-1), a Contaminant of the High Passage RK-13 Cell Line 158

4.1. Summary 158

4.2. Introduction 159

4.3. Materials and methods 160

4.4. Results and discussion 162

CHAPTER FIVE - Summary of Research Findings 167

APPENDICES

APPENDIX 1 - List of Abbreviations 173

APPENDIX 2 - Glossary Relevant to the Virus Evolution 177

APPENDIX 3 - Supplementary Tables..... 180

APPENDIX 4 - Experimental Methods..... 190

REFERENCES 234

VITA..... 262

LIST OF TABLES

Table 1.1. The new classification of the Order *Nidovirales* 3

Table 2.1. Nonsynonymous and synonymous substitution site estimate for ORF3, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions 108

Table 2.2. Nonsynonymous and synonymous substitution site estimate for ORF5, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions 109

Table 2.3. Nonsynonymous and synonymous substitution site estimate for ORF3 was identified in virus isolates from both experimentally and naturally infected horses, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions 115

Table 2.4. Analysis of positively selected codon sites in ORF3 and ORF5 of sequential viruses..... 122

Table 3.1. Comparative nucleotide and amino acid sequence analysis of donkey VD7634 strain of EAV and published EAV sequences (EAVUtr [NC002532] and EAV VBS [DQ846750] 140

Table 3.2. Insertion and deletion positions in full-length genome alignments of EAV isolates from North America and Europe 141

Table 3.3. Serum neutralization test titers of mAbs and polyclonal equine antisera against EAV donkey VD7634 strain and other EAV isolates..... 151

Table 4.1. Cleavage sites between structural and nonstructural proteins of reference BVDV strains and HP-KY-RK13 strain 164

Table 4.2. Nonstructural and structural proteins of BVDV strains..... 164

Table 4.3. Insertion and deletion positions in full-length genome alignments of BVDV-1b strains 165

LIST OF FIGURES

Figure 1.1. Schematic representations of EAV genome organization and gene expression and of the EAV virion particle.. 6

Figure 1.2. Predicted membrane topology of the minor envelope proteins (E, GP2, GP3, GP4 and ORF5a protein).. 9

Figure 1.3. Predicted alternative topology models of the GP2/GP3/GP4 heterodimer complex.. 10

Figure 1.4. Schematic membrane topology of GP5 and M heterodimer structure. 11

Figure 1.5. Schematic diagram of EAV replication cycle. 13

Figure 1.6. Schematic representation of post-translational processing of EAV replicase polyproteins..... 16

Figure 1.7. Two alternative proteolytic processing pathways of EAV replicase polyprotein pp1a.. 18

Figure 1.8. Illustration of the replication and transcription model in nidoviruses.. 22

Figure 2.1. Schematic presentation of the experimental design for this study. 83

Figure 2.2. SNV population in the inoculum used for experimental infection.. 85

Figure 2.3. Mapping of SNV in nasal secretions (NS), buffy coat (BC) cells and semen (S) in seven experimentally infected stallions during acute infection (0 to 9 dpi).. 87

Figure 2.4. Maximum likelihood phylogenetic trees showing a star-like network, which demonstrates the presence a major founder population in the experimental stallions similar to the stock virus. 89

Figure 2.5. SNV population and structure during the acute phase of infection (0 to 9 dpi).. 90

Figure 2.6. Sliding window analyses showing evidence for positive evolutionary selections in the sequential viruses during acute infection.. 92

Figure 2.7. Quasispecies diversity and population size of sequential viruses in nasal secretions, buffy coat and semen during acute and persistent infections.. 94

Figure 2.8. Heat map showing the evolution of 49 sequential viruses from seven experimentally infected stallions.. 98

Figure 2.9. Quasispecies population structure during persistent infection (107 to 726 dpi).. 100

Figure 2.10. Quasispecies populations and positive selection in semen viruses. 101

Figure 2.11. Genetic divergence and evolution of EAV KY84 sequential viruses in experimental stallions..... 103

Figure 2.12. Rate of synonymous and nonsynonymous substitutions across genomic regions and ORFs of the sequential viruses. 105

Figure 2.13. Comparison of nonsynonymous to synonymous nucleotide diversity in the experimental horses.. 106

Figure 2.14. Sequence alignments and comparison of sequential viruses in naturally infected stallions D and E. 112

Figure 2.15. Quasispecies population structure in naturally infected stallions D and E.. 113

Figure 2.16. Comparison of nonsynonymous to synonymous nucleotide diversity of viruses from naturally infected stallions.. 114

Figure 2.17. Estimation of evolutionary rate of the sequential viruses. 117

Figure 2.18. Comparative trend in nucleotide diversity and selection between viruses from the natural and experimental infections.. 118

Figure 2.19. Fast-unconstrained Bayesian approximations for inferring selection (FUBAR) analysis of ORF3.. 120

Figure 2.20. Fast-unconstrained Bayesian approximations for inferring selection (FUBAR) analysis of ORF5. 121

Figure 3. 1. Based on percentage homology among EAV 29 EAV strains; nucleotide sequence identity (%) and amino acid sequence identity (%) analyses were used to create bar graphs.. 145

Figure 3.2. Phylogenetic tree based on the partial ORF5 of EAV donkey VD7634 strain and nine EAV donkey reference strains..... 146

Figure 3.3. Phylogenetic analyses of EAV were inferred based on primary nucleotide alignment..... 148

Figure 3.4. Comparative amino acid substitution of ORF5 of EAV donkey VD7634 to 29 full-length EAV strains according to the serum neutralization test (SNT) result by mAb..... 152

Figure 4. 1. Phylogenetic analyses of whole genome of 34 reference BVDV strains obtained from GenBank and the ncpBVDV HP-KY-RK13 strain identified with a black dot..... 166

CHAPTER ONE

Literature Review

1. Introduction

Equine arteritis virus (EAV) is a small, enveloped, positive-sense, single-stranded RNA virus that belongs to the family *Arteriviridae* (genus: *Equartevirus,* order: *Nidovirales*)[1]. Viruses in the order *Nidovirales* (*Arteriviridae, Coronaviridae* and *Roniviridae*) have a similar genome organization and viral replication strategy, mainly characterized by the generation of a nested set of subgenomic viral messenger RNAs (sg mRNAs) [1–6]. Recently, the International Committee on Taxonomy of Viruses (ICTV) reorganized the genus designation and proposed expansion of the family *Arteriviridae,* a taxonomic group that includes EAV, porcine reproductive and respiratory syndrome virus (PRRSV), lactate dehydrogenase-elevating virus (LDV) of mice, simian hemorrhagic fever virus (SHFV), wobbly possum disease virus (WPDV) and African pouched rat arterivirus (APRAV-1)[6–8].

EAV is the causative agent of equine viral arteritis (EVA), which is a reproductive and respiratory disease of the family *Equidae* that occurs worldwide[9–13]. In 1953, EAV was first isolated from the lung of an aborted fetus during an extensive outbreak of respiratory disease and abortion on a Standardbred breeding farm near Bucyrus, OH, United States of America (USA)[9,14]. While there is only one known serotype of EAV, field strains of the virus can vary in virulence and neutralization capabilities[15–21]. The clinical signs of EVA are variable depending on the age and physical condition of the horse, route of exposure, virus strain, challenge dose and environmental factors[22,23]. Most importantly, EAV can cause abortion in pregnant mares,

neonatal death in young foals, and establishment of persistent infection in 10-70% of infected stallions (carrier state), with significant economic consequences for the equine industry[11].

This chapter provides a review of EAV (replication cycle, viral structure and neutralization determinants) and of the disease it causes (EVA) including epidemiology, modes of transmission, immune response, clinical signs, pathogenesis, persistent infection in the stallion, diagnosis, treatment, prevention and control. Lastly, an overview will be presented on RNA virus quasispecies structure and evolutionary mechanisms.

1.1. Equine arteritis virus (EAV)

1.1.1. Classification and genome organization

Table 1.1. The new classification of the Order *Nidovirales*

Order	Family	Subfamily	Genus	Species
Nidovirales	***Arteriviridae***		*Dipartevirus*	Wobbly possum disease virus
			Equartevirus	**Equine arteritis virus**
			Nesartevirus	African pouched rat arterivirus
			Porartevirus	Lactate dehydrogenase-elevating virus
				Porcine reproductive and respiratory syndrome virus 1
				Porcine reproductive and respiratory syndrome virus 2
				Rat arterivirus 1
			Simartevirus	DeBrazza's monkey arterivirus
				Free State vervet virus
				Kafue kinda chacma baboon virus
				Kibale red colobus virus 1
				Kibale red colobus virus 2
				Kibale red-tailed guenon virus 1
				Mikumi yellow baboon virus 1
				Pebjah virus
				Simian hemorrhagic encephalitis virus
				Simian hemorrhagic fever virus
	Coronaviridae	*Coronavirinae*	*Alphacoronavirus*	Alphacoronavirus 1
				Bat coronavirus CDPHE15
				Bat coronavirus HKU10
				Human coronavirus 229E
				Human coronavirus NL63
				Miniopterus bat coronavirus 1
				Miniopterus bat coronavirus HKU8
				Mink coronavirus 1
				Porcine epidemic diarrhea virus
				Rhinolophus bat coronavirus HKU2
				Scotophilus bat coronavirus 512
			Betacoronavirus	Betacoronavirus 1
				Hedgehog coronavirus 1
				Human coronavirus HKU1
				Middle East respiratory syndrome-related coronavirus
				Murine coronavirus
				Pipistrellus bat coronavirus HKU5
				Rousettus bat coronavirus HKU9
				Severe acute respiratory syndrome-related coronavirus
				Tylonycteris bat coronavirus HKU4
			Deltacoronavirus	Bulbul coronavirus HKU11
				Common moorhen coronavirus HKU21
				Coronavirus HKU15
				Munia coronavirus HKU13
				Night heron coronavirus HKU19
				Thrush coronavirus HKU12
				White-eye coronavirus HKU16
				Wigeon coronavirus HKU20
			Gammacoronavirus	Avian coronavirus
				Beluga whale coronavirus SW1
		Torovirinae	*Bafinivirus*	Fathead minnow nidovirus 1
				White bream virus
			Torovirus	Bovine torovirus
				Equine torovirus
				Human torovirus
				Porcine torovirus
			Unassigned	Ball python nidovirus 1
				Bovine dinovirus 1
				Chinook salmon nidovirus 1
	Mesoniviridae		*Alphamesonivirus*	Alphamesonivirus 1
				Alphamesonivirus 2
				Alphamesonivirus 3
				Alphamesonivirus 4
				Alphamesonivirus 5
			Unassigned	Mesonivirus 1
				Mesonivirus 2
	Roniviridae		*Okavirus*	Gill-associated virus

EAV was initially classified as a non-arthropod-borne togavirus along with LDV and SHFV because of the similarity of their virion morphology and protein compositions[24–26]. In 1984, the genus *Arterivirus* was established within the family *Togaviridae*[27,28]. However, since neither togaviruses nor flaviviruses synthesize a 5′-3′-coterminal set of nested sg mRNAs during viral replication, the replication strategy of EAV was more closely related to that of coronaviruses, LDV, SHFV and PRRSV[3,29–33]. Thus, the 6th ICTV report (1995) assigned all four viruses (EAV, LDV, SHFV and PRRSV) to the genus *Arterivirus*, following which the 7th ICTV report (2000) announced that the genus *Arterivirus* belongs to a new family (*Arteriviridae)* under a new order *Nidovirales*[34,35]. Recently, the ICTV proposed a reorganization and expansion of the family *Arteriviridae*. Under the new proposed classification, the genus designation "*Arterivirus*" is now specifically divided into *Dipartevirus*, *Equartevirus*, *Nesartevirus*, *Porartevirus and Simartevirus*[27]. Currently, there are four families (*Arteriviridae*, *Coronaviridae*, *Mesoniviridae* and *Roniviridae*) within the order *Nidovirales*, and the family *Arteriviridae* is formalized by five genuses: *Dipartevirus*, *Equartevirus*, *Nesartevirus*, *Porartevir*us and *Simartevirus* (Table 1.1.).

The EAV virion is an enveloped, spherical, 50 to 65 nm particle with an isometric core that contains a single-stranded, positive-sense ribonucleic acid (RNA) molecule (Figure 1.1.B.). The EAV genome length varies between 12,704 to 12,731 bp among different virus strains and includes a 5′ leader sequence (224 nucleotides) and at least ten open reading frames (ORFs; Figure 1.1.A.). The two large ORFs 1a and 1b occupy the 5′-proximal three-quarters of the EAV genome and encode two replicase polyproteins (pp1a and pp1ab) directly translated from the genomic RNA[3]. These two large replicase

polyproteins are processed into at least 13 nonstructural proteins (nsp1-12, including nsp7α/β) by three viral proteases (nsp1, nsp2 and nsp4)[36,37]. The 3′-distal quarter of the genome consists of eight ORFs (ORF 2a, 2b, 3-4, 5a, 5b, 6 and 7), which encode for the structural proteins E, GP2, GP3, GP4, GP5a protein, GP4, M and N proteins, respectively[5,38–41]. The seven envelope structural proteins (E, GP2, GP3, GP4, GP5a protein, GP5 and M) are classified into two major envelope proteins (GP5 and M) and five minor envelope proteins (E, GP2, GP3, GP4 and ORF5a protein)[1]. The structural proteins are expressed from a 3′-coterminal nested set of six sg mRNAs (2-7) that have a common 5′ leader sequence and transcription regulating sequence (TRS) at the 5′-end of each sg mRNA (Figure 1.1.A.)[29,42,43]. The sg mRNAs function monocistronically, except for the bicistronic sg mRNAs 2 and 5. The protein translation of sg mRNAs 2 (E and GP2 proteins) and 5 (ORF5a protein and GP5) is expressed by leaky scanning of the 5′-proximal end of the same sg mRNAs[39,40].

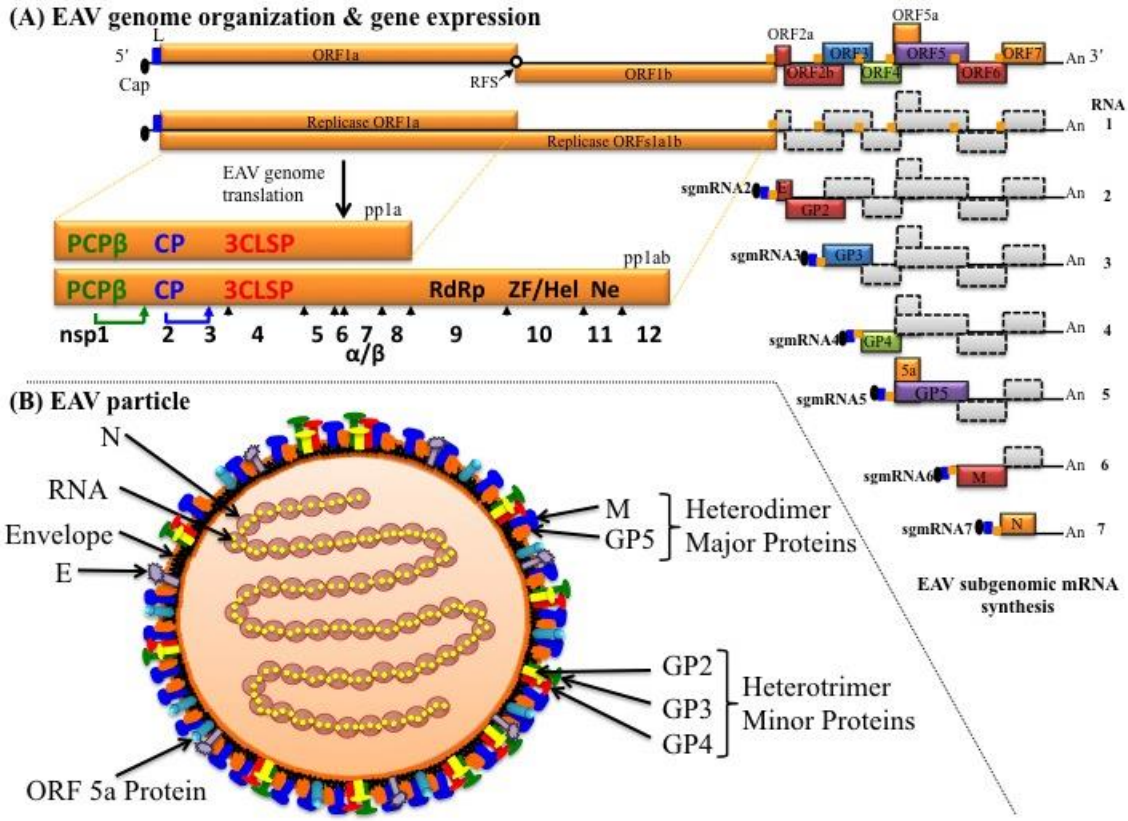


Figure 1.1. Schematic representations of EAV genome organization and gene expression and of the EAV virion particle. (A) The ORFs 1a/b are located at the 5′-end of the genome and are translated into two replicase polyproteins (pp1a and pp1ab) and further processed into at least 13 nsps by three viral proteases (PCPβ, CP and 3CLSP)[3,36,37] (B) The EAV virion particle contains seven envelope structural proteins: E, GP2, GP3, GP4, ORF5a protein, GP5 and M. The N protein encapsulates the linear, positive-sense, single-stranded RNA genome. This schematic representation was modified from Balasuriya *et al*. (2013) from Elsevier with permission[1].

1.1.2. Nonstructural proteins of EAV

The replicases pp1a and pp1ab encoded by ORF1a/1b undergo proteolytic processing at 11 cleavage sites by viral proteases (two papain-like cysteine protease domains: "accessory protease" nsp1, nsp2 and a chymotrypsin-like serine: "main protease" nsp4 which is encoded by ORF1a) and results in 13 nsps (nsp1-12, including nsp7α/β)[1,36,37]. The ORF1a protein is comprised of eight cleaved end products, nsp1-8 including a number of processing intermediates[44–46].

The nsp1 is an "accessory protease" that functions in virion biogenesis and processes the replicase polyprotein to produce sg RNA[47,48]. The nsp2 is another "accessory protease" responsible for the cleavage of the nsp2/3 site through auto-protease activity and the cofactor for processing of the nsp4/5 site[44,48,49]. The nsp3 is a transmembrane protein involved in double-membrane vesicle formation[48,50]. The nsp4 is the "main protease" that plays a significant role during the cleavage of nsp3-8 and nsp3-12[44,48]. The nsp5 participates in the formation of the EAV replication complex[44,48]. The nsp7 is further processed into nsp7α/β; the functions of nsp6-8 are undetermined[44,48,51].

The ORF1b translation (expression of pp1ab) occurs through a -1 ribosomal frameshift mechanism by a "slippery" sequence localized in the overlapping region of ORF1a/1b[3,36,37,44,45,49,52–62]. The ORF1b generates four major protein products after processing mediated by nsp4 that include the RNA-dependent RNA polymerase (RdRp, nsp9) and the NTPase/RNA helicase domains (nsp10), which are necessary for viral RNA replication and mRNA transcription[3,44,48,63,64]. The nsp11 has endoribonuclease activity important for viral RNA synthesis [48,65]. Lastly, the function of nsp12 has not yet been determined[1]. The mature nsps are also involved in the constitution of the membrane-

anchored replication/transcription complex (RTC) that promotes viral genome replication and synthesis of the nested set of sg mRNAs for the expression of the viral structural proteins[66,67]. Previous immunofluorescence studies have demonstrated that the ORF1b-encoded replicase subunits are observed in the perinuclear region of EAV infected cells; this suggested that the ORF1b-encoded replicase subunits might be associated with intracellular membranous compartments in the endoplasmic reticulum (ER)[61].

1.1.3. Structural proteins of EAV

1.1.3.1. Minor envelope proteins

The EAV virion contains five minor envelope proteins: E, GP2, GP3, GP4 and ORF5a protein encoded by ORF2a, ORF2b, ORF3, ORF4 and ORF5a, respectively (Figure 1.2.)[1]. The unglycosylated envelope protein E is highly hydrophobic and predicted to be a type III integral membrane protein and is intermediately abundant in the virion[40]. The E protein is essential for the production of infectious progeny [40,41]. The GP2 is a type I transmembrane minor envelope protein that is highly conserved on sequence comparison analysis[68]. Although GP2 is not essential for viral RNA replication and transcription[40,69], it is essential for the production of infectious progeny[40,48]. According to Wieringa *et al*. (2002), suppression of ORF3 expression restrains the production of infectious virus particles, which implies that GP3 membrane-associated envelope glycoprotein is essential for viral replication[48,70]. The GP3 is predicted to be either a type II protein or a class IV protein[70,71]. Although the function of GP4 is not fully understood, it is predicted to be as essential for viral replication as GP3[48,70]. The ORF5a protein is a novel structural protein encoded by an overlapping region in the 5′-end of ORF5[39,72]. The

ORF5a protein is a type III membrane protein that is expressed from the same sg mRNA as GP5 and is possibly involved in the leaky ribosomal scanning[1]. The function of ORF5a protein is not fully understood, but it is predicted to be essential for viral replication and infectivity [1,39].

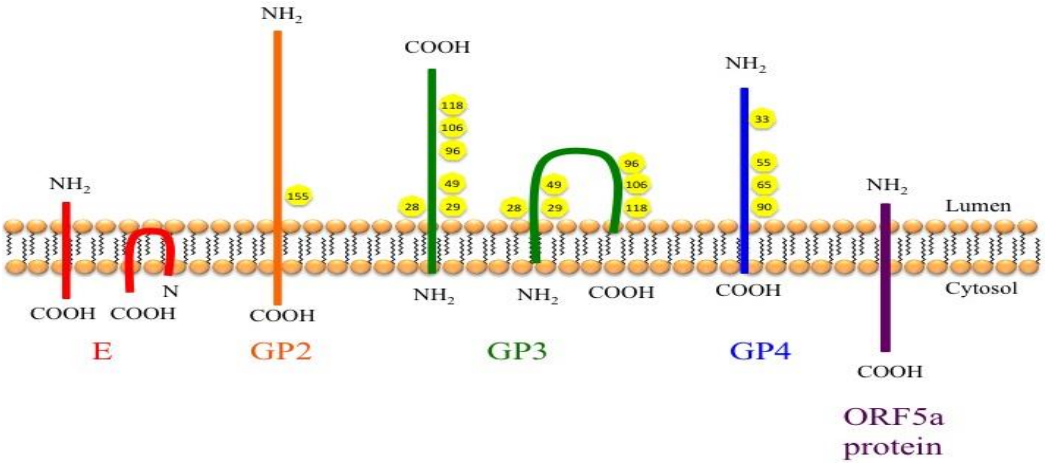


Figure 1.2. Predicted membrane topology of the minor envelope proteins (E, GP2, GP3, GP4 and ORF5a protein). Two alternative topology predictions for E and GP3 enveloped membrane proteins are demonstrated above. Predicted N-glycosylation sites in these proteins are shown in yellow circles with corresponding amino acid positions. Modified from Balasuriya *et al*. (2013) from Elsevier with permission[1].

The minor envelope proteins GP2, GP3 and GP4 are abundantly expressed in EAV infected cells and covalently form a GP2/GP3/GP4 heterotrimeric complex[38,41,70,73,74]. The GP2 and GP4 proteins are both type I integral membrane proteins that contain one and four functional N-glycosylation sites, respectively (Figure 1.3.)[1,70,75]. The intermolecular cysteine bridges with intra-chain disulfide bonds play a significant role in the conformation and stabilization of the GP2/GP3/GP4 heterotrimeric

complex (Figure 1.3.)[74]. The E protein may act as an intermediate receptor and is associated with the GP2/GP3/GP4 heterotrimeric complex during the process of viral attachment/entry and production of infectious viral progeny[1,76].

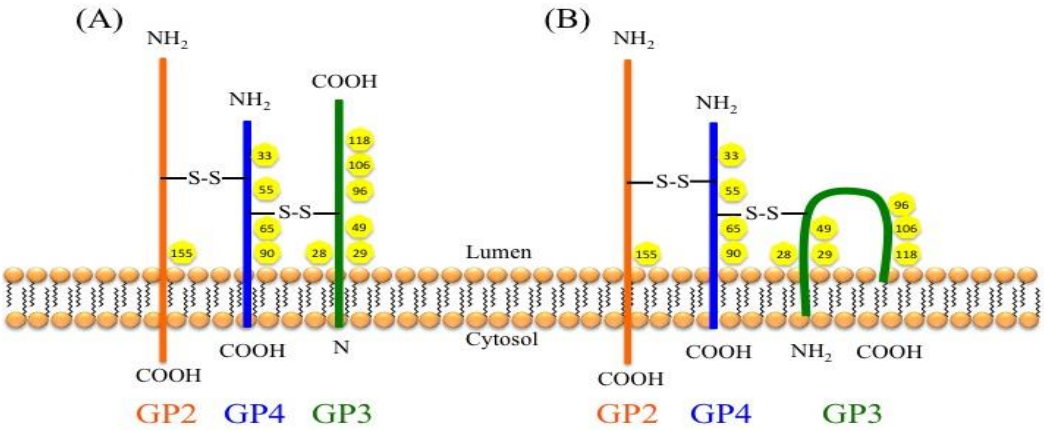


Figure 1.3. Predicted alternative topology models of the GP2/GP3/GP4 heterodimer complex. The GP3 envelope membrane protein is predicted to be either a class II membrane protein (A) or a class IV membrane protein (B). The intermolecular cysteine bridges are predicted with the covalently-linked disulfide bonds (S-S) at the Cys positions. Putative N-glycosylation sites are shown in yellow circles with corresponding amino acid positions. Modified from Wieringa *et al*. (2003) with permission[74].

1.1.3.2. Major envelope proteins

The membrane glycoprotein GP5 and nonglycosylated membrane (M) protein are type IV integral membrane proteins encoded by ORF5b and ORF6, respectively[38,77]. The GP5 and M large envelope proteins form disulfide-linked heterodimers in the virus particles (Figure 1.4.B.). There are four major neutralization determinants of EAV that have been identified in GP5, located at aa 49 (site A), 61 (site B), 67-90 (site C) and 98-106 (site D)

(Figure 1.4.A.)[15,16,20,78–81]. Site D expresses overlapping linear epitopes that may enable it to interact with three other neutralization sites (A-C) to establish conformational epitopes[15,16,81,82]. The nonglycosylated M protein acts as an important scaffold for the correct folding of GP5 in order to generate the neutralizing epitopes[1].

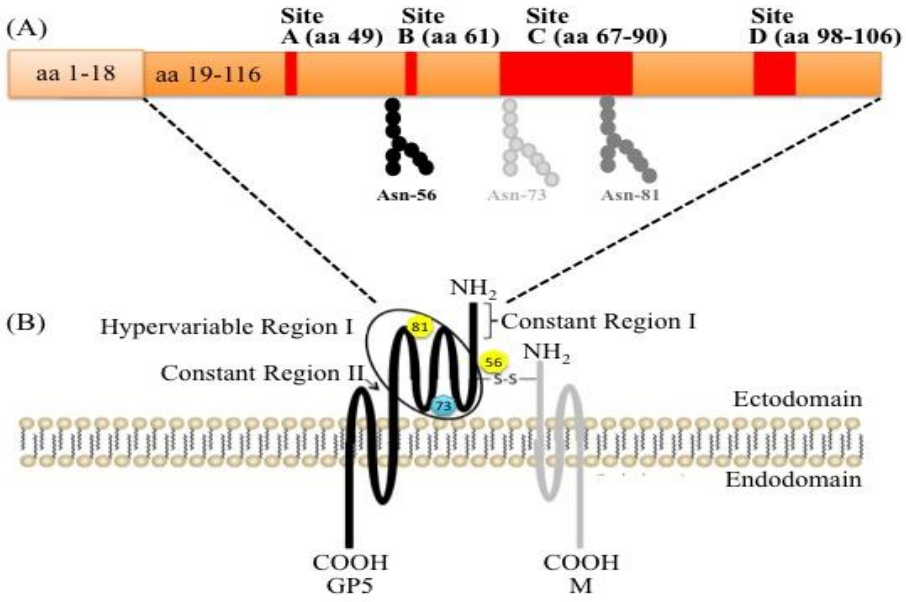


Figure 1.4. Schematic membrane topology of GP5 and M heterodimer structure. (A) Four major neutralization determinants of EAV are shown at aa 49 (site A), 61 (site B), 67-90 (site C) and 98-106 (site D) in the GP5 protein. (B) The covalently-linked heterodimer complex is formed by a disulfide bond (S-S) between Cys-34 in the GP5 protein and Cys-8 in the M protein. Putative N-glycosylation sites are shown in yellow circles with corresponding amino acid positions: Asn-56 and Asn-81. Another N-glycosylation site (Asn-73) that was recently found in isolates from an EVA outbreak in North America[21] is shown in a blue circle. Modified from Balasuriya *et al*. (2013) from Elsevier with permission[1].

In EAV infected cells, the newly synthesized GP5 proteins recruit existing M proteins from the ER to form a disulfide-linked heterodimeric complex[1,75]. The disulfide-linked heterodimeric complex of the two major envelope proteins are formed by the disulfide bridge between two cysteine (Cys) residues at amino acid position Cys-8 of M and Cys-34 of GP5[73]. The heterodimerization of the GP5 and M is indispensable for viral infectivity and production of infectious virus particles[41,69,83]. The GP5/M disulfide-linked heterodimer complex acts as a bridge between the nucleocapsid and the viral envelope during viral assembly[83,84]. Lastly, the GP5/M heterodimerization complex plays an essential role in post-translational modification (i.e. glycosylation) and conformational maturation of the neutralization determinants in the GP5 ectodomain[82,85].

1.2. The replication cycle of EAV

The EAV life cycle begins with the infection of susceptible cells with EAV (Figure 1.5.). EAV can be replicated in various primary cell cultures including equine endothelial cells, equine kidney cells, monocytes, macrophages and a small subpopulation of $CD3^+$ T cells[1,86–88]. EAV has tropism for a broad range of host cell types, and it can be propagated in several continuous cell lines such as baby hamster kidney (BHK-21), African green monkey kidney (Vero, MA-104), rabbit kidney (RK-13) and human kidney cells[1,89–91]. EAV infection is highly cytocidal in both primary and continuous cell lines, and it replicates to produce high titered virus causing cytopathic effect (CPE)[89,90].

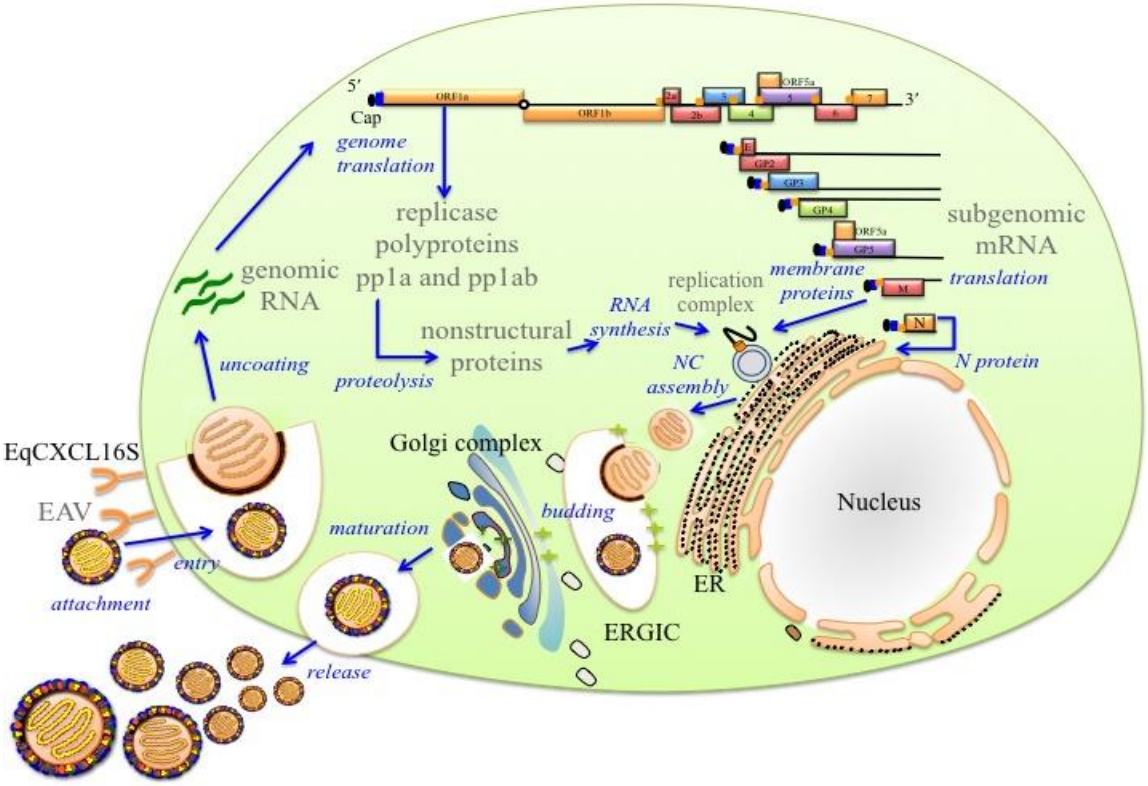


Figure 1.5. Schematic diagram of EAV replication cycle. ER: endoplasmic reticulum; ERGIC: ER-Golgi intermediate compartment; NC: nucleocapsid. Modified from Balasuriya *et al*. (2013) from Elsevier with permission[1].

1.2.1. Virus attachment and entry

The replication cycle of EAV involves virus attachment and entry, biosynthesis (viral genome replication, mRNA transcription and viral protein synthesis [translation]), virus assembly, budding and release[92]. Asagoe *et al*., (1997) have shown that heparin, which is a glycosaminoglycan (GAG), can inhibit EAV infection of RK-13 cells but the mechanism of the inhibitory activity by heparin was not elucidated[93]. The first step in the virus replication cycle is viral attachment and entry that is mediated by a specific receptor on EAV-susceptible cells[92]. Heparin sulphate (HS) is commonly used as an attachment or

binding molecule for numerous viruses and other pathogens as it is ubiquitously present on the surface of animal cells and in the extracellular matrix, where the initial interactions of viruses and target cells occur. Thus, EAV may also be using HS on cell surfaces as the initial attachment molecule[94].

Recently, Sarkar *et al*. (2016) identified EqCXCL16S as an EAV entry receptor on EAV-susceptible cells[95,96]. CXCL16S protein has scavenger receptor properties and is expressed on equine endothelial cells, a subpopulation of $CD3^+$ T cells and a subpopulation of $CD14^+$ monocytes; its distribution in cells of the reproductive tract has not however been determined. The *EqCXCL16* gene is located on equine chromosome 11 (ECA11; positions 49572804 to 49643932), and encodes a type I glycosylated transmembrane protein[95,97]. Two isoforms of EqCXCL16 have been identified (EqCXCL16S and EqCSCL16R). Interestingly, the protein isoform EqCXCL16S but not EqCXCL16R can function as an EAV cellular receptor. Although both molecules have equal chemoattractant potential, EqCXCL16S has significantly higher scavenger receptor and adhesion properties compared to EqCXCL16R. However, EqCXCL16S is certainly not the only receptor because EAV can infect not only equine cells, but also a wide range of commonly used laboratory continuous cell lines from other species of animals[1,86–91]. Similar to other enveloped viruses, EAV tends to bind to a cell surface receptor followed by membrane fusion[95,98]; it is assumed to enter the cell via clathrin-dependent endocytosis[99,100]. However, EAV attachment and entry mechanisms still need to be further investigated. This should include investigation of the possible role of other virus attachment molecules on the cell surfaces, unknown receptors and co-receptors in different cells and the mechanism of fusion between EAV and the target host cell.

Specific viral proteins involved in EAV attachment and entry are not well characterized. However, there have been several studies that investigated the role of major (GP5 and M) and minor envelope (E, GP2, GP3 and GP4) proteins of the virus in EAV attachment and entry. Firstly, GP5 protein was assumed to be a virus attachment protein and acts as the receptor recognition. However, an *in vitro* reverse genetics experiment using a chimeric EAV construct in which the ectodomain of the EAV GP5 protein was replaced by that of other arteriviruses (e.g. PRRSV or LDV) resulted in no alteration in the viral cell tropism[101]. This chimeric virus was still able to propagate in BHK-21 and RK-13 cells. Thus, the ectodomain of the GP5 protein was not the major determinant of EAV tropism in cell culture[101]. A similar study was performed with the ectodomain of the M protein of a PRRS mutant virus, where its exchange for that of EAV or LDV did not alter the tropism for porcine alveolar macrophages. It was concluded that the M protein of arteriviruses is not responsible for receptor binding[84]. Secondly, the E protein is noncovalently associated with the minor structural proteins that integrate the GP2/GP3/GP4 heterotrimeric complex[41]. Although the association of this complex is essential for viral infectivity, it is not required for the formation of EAV particles; studies have suggested that the GP2/GP3/GP4/E complex might be involved in the EAV attachment and entry process and that the virus uses more than one attachment molecule for binding to the host cell[41,69]. Thirdly, EAV was presumed to attach and enter the host cells via a mechanism of receptor-mediated clathrin-dependent endocytosis similar to PRRSV[99,100,102]. There are no definitive studies however that have identified the EAV viral proteins involved in virus attachment and entry or, for that matter, the detailed

mechanisms of viral entry and release of the EAV genome into the cytoplasm to initiate viral replication.

1.2.2. Post-translation in processing of replicase polyproteins

Translation of the EAV replicase polyproteins (pp1a and pp1ab) generates at least 13 nsps after cleavage mediated by three ORF1a-encoded proteases (nsp1, nsp2 and nsp4) (Figure 1.6.)[36,37,103]. The pp1a is translated directly from ORF1a and processed at seven cleavage sites into eight end products, nsp1 to nsp8, with several processing intermediates[44–46]. Expression of ppa1ab requires a -1 ribosomal frameshift mechanism and is processed into four major end products including viral RdRp and RNA helicase by nsp4 proteases[3,63].

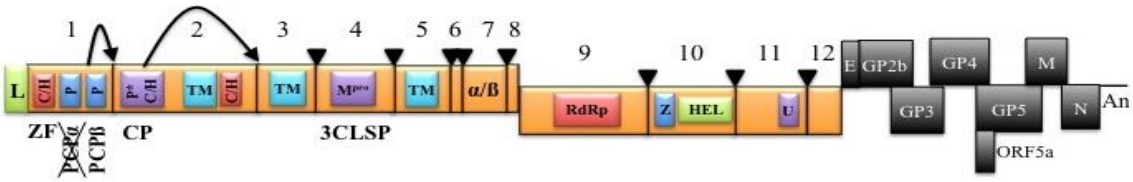


Figure 1.6. Schematic representation of post-translational processing of EAV replicase polyproteins. The order of nsps of the EAV replicase polyproteins is numbered; their distinct characteristics are abbreviated as follows: leader sequence Mpro: 3C-like serine proteases (3CLSP); ZF and Z, zinc-finger and zinc-binding domain, respectively; C/H: cysteine/histidine-rich clusters; TM: predicted transmembrane domains; N and HEL: N-terminus and the nucleoside triphosphate-binding/helicase (HEL), respectively; and U: nidoviral-endonuclease specific for U (NendoU) domain. The PCP1β and CP cleavage sites are shown by black arrows. Multiple 3CLSP cleavage sites are located under the black triangle points. The structural proteins are highlighted in black. Modified from Van Hemert and Snijder. (2008) with permission[104].

The nsp1 is characterized by a zinc finger (ZF) domain and papain-like cysteine protease (PCP) with inactive PCP1α and active PCP1β[37,105]. Although EAV PCP1α becomes inactive by the loss of an active Cys-site during virus evolution[53], the PCP1β acts as a platform for the ZF domain during transcription[57]. The nsp1 is released from the replicase polyprotein by the autocatalytic activity of PCP1β resulting in the cleavage of the nsp1/2 site at its C-terminal domain for the purpose of viral sg mRNA synthesis[44,47,105]. Secondly, the N-terminal domain of nsp2 acts to activate the cysteine protease (CP) that releases the nsp2 from the replicase polyprotein after the cleavage of the nsp2/3 site by the internal cysteine autoprotease[44,49]. Furthermore, the mature nsp2 is responsible for the processing of the nsp4/5 site through either the major or minor processing pathway (Figure 1.7.). Thirdly, the N-terminal domain of nsp4 has chymotrypsin-like serine protease (SP) activity[52]. The nsp4 is the "main protease", also known as a 3C-like serine protease (3CLSP), responsible for the production of the viral RdRp and NTPase/RNA helicase (nsp9 and nsp10, respectively)[37,46,106]. After nsp1 and nsp2 have been autocatalytically released from pp1a, nsp4 initiates further processing, cleavage of the polypeptides: five cleavage sites in nsp3-8 of pp1a and three cleavage sites in nsp3-12 of pp1ab via two alternative pathways (Figure 1.7.)[44–46,49,62,105].

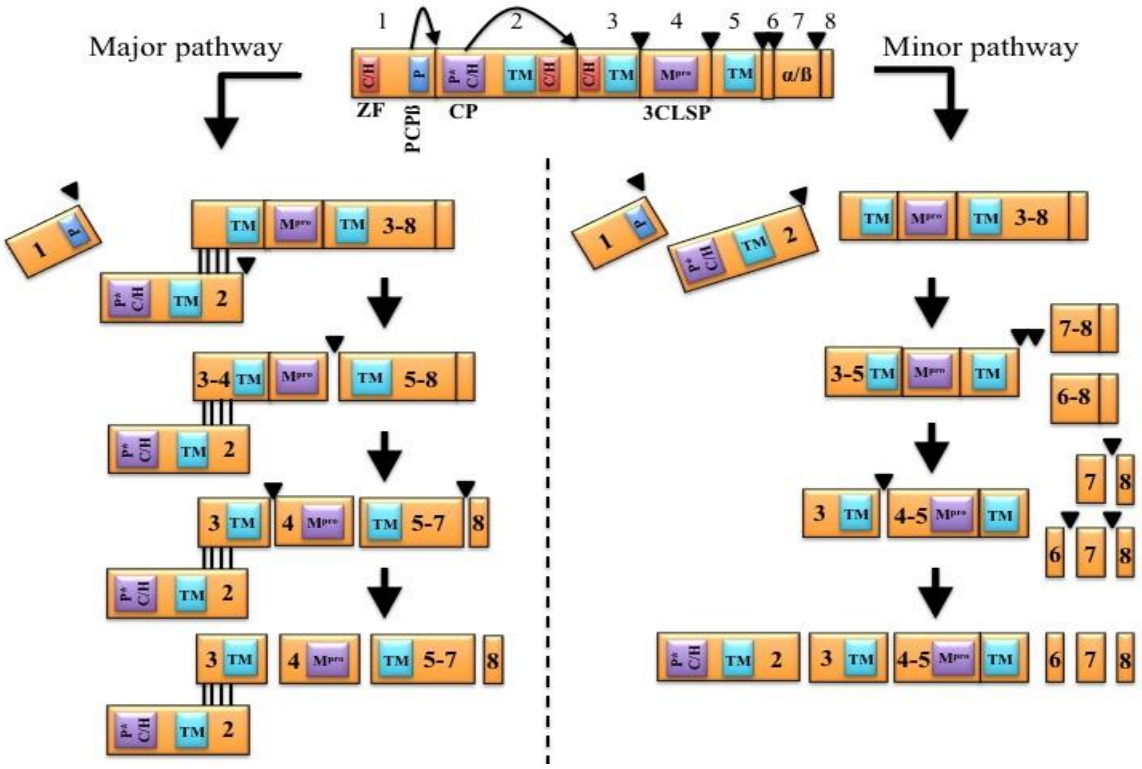


Figure 1.7. Two alternative proteolytic processing pathways of EAV replicase polyprotein pp1a. The important characteristics of EAV replicase polyproteins are abbreviated by P and P*: papain-like cysteine protease (PCPα/β) and cysteine protease (CP), respectively; Mpro: 3C-like serine protease (3CLSP); C/H: cysteine/histidine-rich clusters; and TM: predicted transmembrane domains. The PCP1β and CP cleavage sites are shown by black arrows. Multiple 3CLSP cleavage sites are located under black triangle points. Modified from the *Nidoviruses* book chapter: Van Hemert and Snijder. (2008) with permission[104].

As mentioned above, EAV PCP and EAV CP are responsible for processing the cleavage sites nsp1/2 and nsp2/3, respectively, to release nsp1 and nsp2 from the replicase polyproteins. The 3CLSP processes the rest of the polypeptides in the EAV

pp1a and pp1ab at the EAV 3CLSP-specific known cleavage sites: four at Glu|Gly site (nsp3/4, nsp5/6, nsp7/8/9 and nsp11/12), three at Glu|Ser sites (nsp4/5, nsp6/7 and nsp9/10), and one at Gln|Ser site (nsp10/11) (Figure 1.6.)[37,45,46,62]. There are two alternative (major and minor) pathways that have been identified in the processing of the carboxyl-terminal part of pp1a (nsp3-8) (Figure 1.7.)[45]. In the major processing pathway, the mature nsp2 acts as a cofactor that strongly associates with the processing of nsp3 or nsp3-carrying precursors including nsp3-8 required for further processing the nsp4/5 site by the nsp4 3CLSP[45]. The EAV 3CLSP first cleaves the nsp4/5 site, then processes the nsp3/4 and nsp7/8 sites which results in the rest of the remaining nsp5-7 intermediates remaining uncleaved (Figure 1.7.)[44,46]. However, in the case of the nsp2, this is absent and the 4/5 site cannot be cleaved; in which case the alternative minor processing pathway takes responsibility for further processing[45]. In the minor processing pathway, the nsp4/5 site remains uncleaved; instead, the cleavage processing occurs at the nsp5/6 and nsp6/7 sites[45]. Then, further processing may occur at the nsp3/4 and 7/8 sites (Figure 1.7.)[45].

1.2.3. Genome replication

The typical feature of EAV and other arteriviruses is the formation of paired membranes and double-membrane vesicles (DMVs) after 3 to 6 hour post-infection (hpi)[107–111]. Expression of EAV nsp2-3 is essential for formation of DMVs[56,112]. Pedersen *et al.*, (1999) demonstrated that EAV replicase subunit nsp9, which contains RdRp function, assembles the viral replication complex through membrane-associated DMVs where the RNA synthesis takes place. The outer membranes of EAV-induced DMVs

closely interact with the ER membranes of the target host cells and result in the formation of a reticulovesicular network (RVN)[113]. Electron spectroscopic imaging of DMVs has revealed the presence of EAV RNA genomes and has identified a network connection between EAV-induced RVN and nucleocapsid assembly[114]. Similarly, the presence of newly synthesized viral RNA genomes and many replicase subunits with numerous amounts of DMVs were observed near the perinuclear region in nidovirus-infected cells[112,115,116].

The viral replicase polyproteins are visualized by immunoelectron microscopy to be part of the RTC, and DMVs are closely associated with newly synthesized viral RNA genomes[112,116–118]. Thus, DMVs are predicted to comprise the enzyme complex that plays an important role in viral replication and sg mRNA synthesis[56,112]. The EAV replicase subunits including nsp2-3 and nsp5 that have hydrophobic regions are predicted to impact the structural function of the anchored-RTC by inducing DMV formation[44,45,56,119]. Moreover, a recent study has shown that the isolated RTC-comprised replicase subunits are involved in DMV formation (nsp2-3) or viral replication (nsp9-10) with the accumulation of newly synthesized viral RNA genomes[120], indicating the association with DMV and its essential role in RTC function[1]. As described above, the assembly of the EAV replication complex is based on the expression of replicase polyproteins and the membrane of viral induced- and host cell-derived DMVs that lead to viral RNA synthesis[92].

The EAV genome is a polycistronic RNA molecule that consists of a 5′ cap and 3′ poly (A) tail structures[33]. While the EAV RdRp replicates (+) the viral genomic strand into full-length (-) viral strands (or anti-sense genome) as full-length (-) strand templates

synthesize more (+) viral genomic strands[92], the subgenomic-sized (-) strands are produced to synthesize sg mRNAs during a process of discontinuous RNA synthesis as with other nidoviruses[39,66,102,121]. The recognition of signals at the 3′-end of the full-length (-) strand viral templates is required for the production of new genomic RNA[122,123]. During viral RNA synthesis, more (+) strands are generated than (-) strand RNA genomes[124].

1.2.4. Viral sg mRNA transcription

The EAV genome has a 5′ UTR of approximately 225 nucleotides (nts) that prevents exonuclease-mediated degradation[125]. Furthermore, the 5′ UTR contains a common 5′ leader sequence (211 nt in length)[33]. As in the case of other nidoviruses, EAV sg mRNAs synthesis involves a discontinuous transcription mechanism[121]. This discontinuous extension of (-) stranded RNA is controlled by the transcription regulatory sequence (TRS; Figure 1.8.), a short and conserved sequence element (5′-UCAACU-3′), which is essential for leader-to-body joining via a base-pairing interaction between (+) and nascent (-) stranded RNAs[126–128]. The leader TRS is located in a hairpin loop structure at the 3′-end of the leader sequence[129] and several body TRSs are found upstream of structural protein genes, except for ORFs 2b and 5a[29,39,40,43]. The primary sequence of the body TRS located at an active site in each of the six sg mRNAs will be transcribed; it plays a significant role in discontinuous RNA synthesis[66,126,127,130,131]. The existence of TRSs was also found in the replicase genes[43].

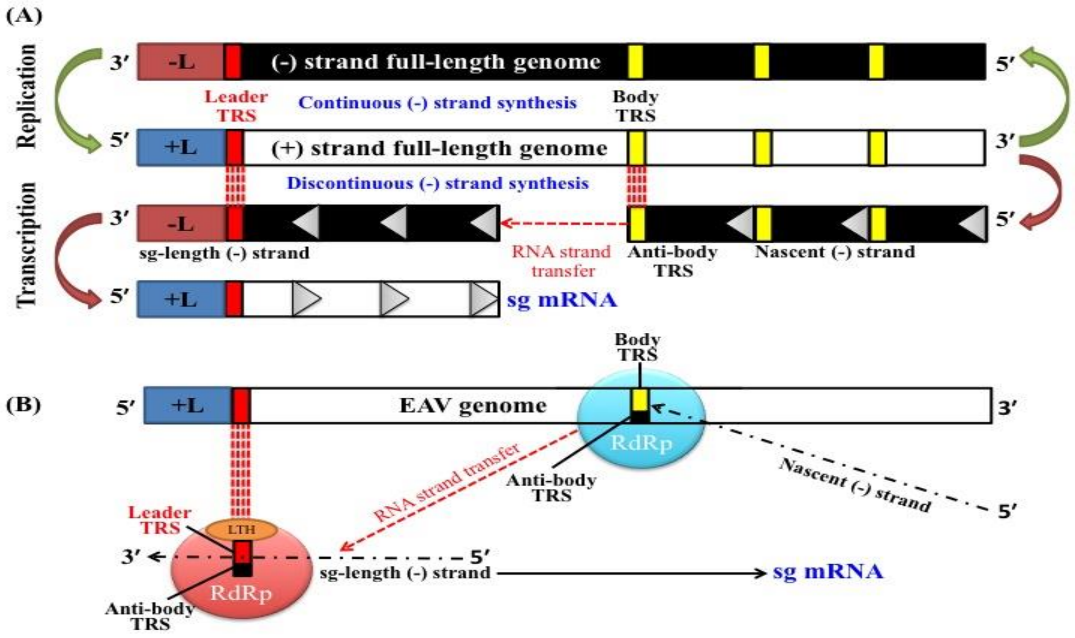


Figure 1.8. Illustration of the replication and transcription model in nidoviruses. (A) After viral replication, the "discontinuous extension of (-) strand RNA model" occurs by using the sg (+) strand as a template for sg mRNA synthesis[121]. The process by which translocation of the nascent (-) strand to the leader TRS region occurs is called RNA strand transfer[127,128]. The (-) strand RNA synthesis reinitiates by base pairing interaction between the antisense body TRS and sense leader TRS added to the anti-leader sequence. The sg-length nascent (-) strand RNA is used as a template strand for sg mRNA synthesis. The (+) and (-) strand RNAs are shown in white and black, respectively. The sense leader and body TRSs are indicated in red and yellow boxes on the (+) strand RNA and antisense leader and body TRSs are shown as black boxes on the (-) strand RNA. (B) Like viral genomic replication, the viral sg mRNA transcription is regulated directly by the RdRp complex. LTH: leader TRS hairpin; The figure (A) modified from Pasternak *et al*. (2006) with permission from the *Journal of General Virology* and figure (B) is reprinted from van den Born *et al*. (2005) with permission[66,129].

The sg (-) strand RNA genomes are used as templates for further sg mRNA synthesis[124]. During discontinuous transcription, the viral RdRp controls termination and reinitiation of transcription, indicating that the nidovirus transcription model assembles by a similar-assisted, copy-choice RNA recombination (Figure 1.8.B.)[128,132,133]. The RdRp complex and 3′-end of nascent (-) strand transcript are released and translocated from the body TRS at the leader sequence region in the 5′-end of the (+) strand genomic template where transcription is reinitiated. During (-) strand synthesis reinitiation, a nested set of sg-length nascent (-) strand templates are extended with the complement of the genomic leader sequence for the subsequent synthesis of the various sg mRNAs[1]. This leader-to-body joining via base-pairing interaction between antisense body TRS at the 3′-end of the nascent (-) strand and the sense leader TRS in the (+) genomic template generates the additional anti-leader sequence and has an impact on RNA strand transfer (Figure 1.8.B.)[127,128]. Additionally, the predicted secondary structure of the leader sequence located at nucleotide position nt 1-211 is also important for the sg RNAs synthesis[129,134]. The leader sequence forms a hairpin structure called leader TRS hairpin (LTH) (Figure 1.8.B.)[128,134]. The synthesis of (-) strand RNA started from the 3′-end of the EAV genome becomes attenuated at one of the body TRS regions[66,67]. If attenuation does not happen, (-) strand RNA synthesis continues to generate a full-length complement of the genome as the intermediate step in EAV replication[121]. Like viral genomic replication, the viral sg mRNA transcription is regulated directly by the RdRp complex and occurs on the membrane of virus-induced and host cell-derived DMVs[92].

1.2.5. Virus assembly, budding and release

The newly synthesized viral RNA genome is encapsulated by nucleocapsid formation[1]. EAV obtains its envelope containing viral envelope proteins through the internal membranes of infected cell, and assembly takes place in the cytoplasmic region near the ER[135]. The budding and maturation processes happen at the ERGIC[1]. The completely formed EAV virions mature in the Golgi complex and are released via an exocytic pathway from the infected host cells (Figure 1.5.)[1].

As previously described, the seven envelope structural proteins (E, GP2, GP3, GP4, ORF5a protein, GP5 and M) are essential for the production of infectious progeny virus[40,73,136]. Moreover, the major envelope proteins (GP5, M and N) form virus-like particles (VLPs)[41]. However, producing VLPs by the cotransfection of infected cells with plasmids encoding EAV GP5, M and N proteins was unsuccessful, indicating that additional factors are involved in virus assembly[40,41]. EAV assembly occurs in association with nucleocapsid formation, and thus EAV assembly will not take place without it[40,41]. However, the interaction between EAV N protein and viral genomic RNA, the formation of icosahedral nucleocapsid, and the copy numbers of EAV N proteins required to form the nucleocapsid remain unknown[92].

The EAV nucleocapsid becomes enveloped though the GP5/M disulfide-linked heterodimer complex which acts as a mediator between the nucleocapsid and the viral envelope during viral assembly[83,84]. This also suggested that the GP5/M heterodimerization complex forms the basic protein matrix of the envelope[77]. From the individual expression of these major enveloped proteins, EAV GP5 and M proteins were observed only in the ER; however, co-expression of the M and GP5 proteins resulted in

the M protein presenting both in ER and ERGIC and the GP5 protein co-localized with the M protein only in the ERGIC[85]. This result suggested that transport of EAV GP5 and M proteins from ER to the ERGIC depends on the formation of a GP5/M heterodimerization complex[85]. Snijder *et al*. (2003) further confirmed this by genetic engineering introducing mutations of Cys residues at amino acid position Cys-8 of M protein and Cys-34 of GP5 protein. As a result, mutations caused the blocking of the GP5/M heterodimerization complex and virus assembly, interrupting the transport of both GP5 and M from Golgi to the ERGIC and production of infectious viral progeny[83]. The authors concluded that the GP5/M heterodimerization complex is a prerequisite for virus assembly and occurs in the ER. Additionally, the interaction between cytoplasmically exposed domains of the GP5/M heterodimerization complex and the synthesized nucleocapsid suggested that EAV nucleocapsids bud in the ERGIC by acquiring lipid membrane-carrying viral envelope proteins[92].

After EAV nucleocapsid formation, the minor envelope proteins (E, GP2, GP3 and GP4) also gather at the lipid membrane of the ERGIC and bud into the lumen[92]. The minor enveloped proteins form a GP2/GP3/GP4 heterotrimeric complex and along with the E protein, they participate in the process of viral attachment/entry and production of infectious viral progeny[1,76]. GP2 and GP4 form a heterodimer, which subsequently covalently associates with GP3[137]. The intermolecular cysteine bridges with an intra-chain disulfide bond at the amino acid position Cys-48, 102 and 137 play a significant role in the conformation and stabilization of the GP2/GP3/GP4 heterotrimeric complex (Figure 1.3.)[74,75,137]. When a protein subunit of the GP2/GP3/GP4 heterotrimeric complex is missing, the formation of the heterotrimeric complex fails and the amount of E protein

in the viral envelope decreases[41]. Furthermore, the absence of E protein will entirely block the formation of the GP2/GP3/GP4 heterotrimeric complex in the EAV particle[41].

With the formation of the EAV nucleocapsid bud into the ERGIC, the virus obtains a lipid membrane consisting of viral major and minor envelope proteins to form the virus particles. The newly produced EAV particles are transported from the intracellular compartments to the plasma membrane of the EAV-infected cell, and then finally, the virus particles exit the cell through exocytosis[92].

1.3. Equine viral arteritis

1.3.1. Epidemiology

EVA is a reproductive and respiratory disease of equids caused by EAV[16,9,13,18,138–141]. EAV was first isolated from the lung of an aborted fetus during an extensive outbreak of the respiratory disease and abortion at a Standardbred breeding farm near Bucyrus, Ohio, in 1953[9,14]. EAV causes a significant economic impact on the equine industry due to the occurrence of abortion, neonatal death and establishment of persistent infection in stallions[9,14,139,142,143]. According to serological surveys, EAV has a worldwide distribution that includes North and South America, Europe, Australia, Africa and various countries in Asia[11]. Iceland, Japan and New Zealand are known as free from the disease[144]. EAV seroprevalence varies according to age and breed. For instance, in the United States a very high percentage of adult Standardbred and Saddlebred horses are seropositive (70-90% and 8-25%, respectively) while a low serprevalence is observed in the Thoroughbred population (<5.4%)[11,145–148]. The 1998 National Animal Health Monitoring System (NAHMS) equine study reported only 0.6% EAV seropositivity in

the Quarter horse population in the United States[149], but an extensive EVA occurrence involving multiple states during 2006 to 2007 probably increased the seroprevalence in the Quarter horse population[150]. In Europe, a high seroprevalence is observed among Warmblood stallions and Spanish purebred horses (55-93% and 17.3%, respectively)[151–153]. Variability in EAV seroprevalence is correlated with age and sex[154]. Recent studies have suggested the linkage of a specific dominant haplotype of the *EqCXCL16* gene located in ECA11 at positions: 49572804 to 49643932 with the susceptibility of $CD3^+$ T lymphocytes to *in vitro* EAV infection and the establishment of persistent infection[86,87,155]. It has been shown that those stallions with $CD3^+$ T lymphocyte susceptiblity phenotype have a higher risk of becoming long-term EAV carriers than those that lack this phenotype[155].

1.3.1.1. Transmission

Transmission of EAV between horses occurs by the horizontal, venereal and vertical routes[11,14,138,139,156,157]. Horizontal transmission happens by the aerosolization of respiratory secretions or fetal fluids at time of abortion. During the acute phase of infection, the virus spreads via the respiratory route by the aerosolization of respiratory secretions[13,138,141,152,158–160]. Nasal shedding continues for 7 to 14 days with viral titers ranging from 10 to >2 x 10^3 plaque-forming units per mL (PFU/mL)[138,161,162]. Venereal transmission occurs through infective semen during natural breeding, artificial insemination (AI) and embryo transfer. The venereal transmission happens via the semen of both acutely and persistently EAV-infected stallions by natural or artificial breeding[157]. Most EVA outbreaks are caused by venereal transmission after a naïve mare is bred to an

EAV carrier stallion[163,164]. Approximately, 10 to 70% of EAV infected stallions can become persistently infected and continuously shed EAV in their semen for several weeks to years or lifelong with the presence of high levels of neutralizing antibodies in serum[11,156,157]. The establishment and maintenance of the EAV carrier state in the stallion is testosterone-dependent and the virus mainly persists in the ampulla of the vas deferens[165]. Carrier stallions constitute the natural reservoir of EAV and are responsible for the maintenance, perpetuation and evolution of EAV in equine populations[11,17,19,21,142,146,164,166]. Previous studies have demonstrated that EAV evolves in the stallion reproductive tract resulting in the emergence of novel viral genetic variants with neutralizing phenotypes that allow for viral immune escape[17,142,164,166]. EAV can also be spread from donor mares inseminated with EAV-contaminated semen to naïve recipient mares through embryo transfer[167]. Vertical (transplacental) transmission occurs from pregnant mares to the fetus during the latter stage of pregnancy[168]. Congenital infection of foals results in development of rapid progressive, fulminating interstitial pneumonia and fibrinonecrotic enteritis[169–173]. Lastly, the lateral transmission of EAV can also take place through contact with EAV-contaminated fomites that include personnel, clothing, vehicles and equipment such as artificial vaginas and phantoms[11,146,160].

1.3.1.2. Immune response to EAV infection

EAV infection induces a strong and protective immune response in the host[154]. Even though the innate immune response has not yet been fully understood, recent studies have demonstrated that EAV regulates the innate immune response by blocking type I interferon (IFN) production[174]. Specifically, the three EAV nonstructural proteins

(nsp 1,2 and 11; mainly nsp1) have the ability to inhibit IFN synthesis. After EAV infection, the horse develops EAV-specific complement-fixing (CF) and viral neutralizing (VN) antibodies[154,175,176]. These CF and VN antibodies are initially detected within 7 to 14 days post-infection (dpi) and reach their peak by 2 to 3 weeks or 2 to 4 months post-infection, respectively[162,176–179]. Although CF antibodies decrease by 8 months post-infection, VN antibodies last for 3 years or longer[180]. Foals receive VN antibodies through the colostrum of immunized mares, which provide them with protection against EAV infection. Although high neutralizing antibody titers remain in serum, EAV still persists in the stallion reproductive tract[97,98]. Recently, it has also been shown that EAV also induces a mucosal antibody response in the reproductive tract with shedding of virus-specific immunoglobulins (IgA, IgM, IgG1, IgG3/5 and IgG4/7) into the seminal plasma as well as the homing of plasma cells in the accessory sex glands[165]. However, despite the systemic neutralizing antibody, local inflammatory and mucosal antibody responses, EAV evades local host immunity mechanisms and persistence is maintained.

1.3.2. Clinical signs

Most primary EAV infections are subclinical or inapparent. However, when clinical signs of EVA develop, they present as an influenza-like illness in adult horses, abortion in pregnant mares and interstitial pneumonia or pneumoenteritis in young foals[11,13,171,181182]. The range and severity of the clinical signs depends on various factors including an individual's genetic background, age, physical and environmental conditions, route of exposure, challenge dose and viral strain[11,140]. Field strains of EAV can vary in

virulence and in the clinical severity of the disease they cause. For instance, some field strains cause moderate to severe disease (e.g. EAV KY84, AZ87, IL93 and PA96 strains), moderate disease (e.g. EAV PA76 and KY77 strains), mild disease (e.g. EAV SWZ64, AUT68, IL94 and CA97 strains) or asymptomatic infection (e.g. EAV KY63 and CA95 strains)[15,16,21,23,96,183–190]. While the highly virulent experimentally-derived horse-adapted Bucyrus strain of EAV (EAV VBS) can cause high morbidity and mortality in healthy mature horses following infection, in natural outbreaks of EVA caused by field strains of EAV, mortality is very uncommon[11,183].

Acutely infected animals may develop a wide range of clinical signs[154]. The average incubation period is 3 to 14 days (6 to 8 days after venereal transmission exposure). Infected horses may develop respiratory disease and reproductive consequences including the development of pyrexia of up to 41°C (39-41°C) which may last for 2 to 9 days. Infected animals may also develop all or any combination of the following clinical signs: depression, anorexia, nasal and/or ocular discharge, conjunctivitis, rhinitis, leukopenia, periobital or supraorbital edema, edema of the limbs, edema involving the scrotum and prepuce of the stallion and mammary glands of the mare, urticarial around the head, sides of neck and/or body and abortion of pregnant mares[9,11,140,191–193]. Other less frequent clinical signs include icterus, photophobia, corneal opacity, coughing, dyspnoea, colic, diarrhea, ataxia, petechiation of the nasal mucosa, conjunctiva and oral mucous membranes, submaxillary and submandibular lymphadenopathy and adventitious edema in the intermandibular space, beneath the sternum or in the shoulder region[9,139,180,191,193–195]. The most frequent clinical signs of EVA include pyrexia and leukopenia. Most naturally infected horses recover

spontaneously. Abortion in pregnant mares may occur from early to late gestation (3 to over 10 months of gestation) and abortion rates have varied from less than 10% to >70%. Congenitally infected foals develop a rapidly progressive pneumoenteric syndrome and severe fulminating interstitial pneumonia[169–171].

1.3.3. Pathogenesis

The pathogenesis of EVA has been studied by following natural outbreaks or by experimentally inoculating horses with different EAV strains[9,14,138,138,139,159,161,177,183,184,193,196–199]. Following initial infection, viral replication occurs within alveolar macrophages and bronchiolar epithelial cells, with subsequent spread to regional (respiratory tract-associated) lymph nodes within 42 hpi[138,146,200]. Within 72 hpi, a cell-associated viremia develops and frequently lasts for 3 to 19 dpi. During the viremic phase, the virus spreads throughout the body replicating in pulmonary macrophages, vascular smooth muscle cells and endothelial cells and causing a systemic panvasculitis[1,138,183,194,198,201–203]. The severe systemic panvasculitis of small vessels with endothelial cell injury increases vascular permeability and leukocyte infiltration. This can cause foci of intimal, subintimal and medial necrosis with hemorrhage, edema and permeation with lymphocytes and neutrophils in the muscular arteries[183,189,197,198,201,202]. The endothelial cells and macrophages are recognized as the principal target cells as well as all the epithelial, mesothelium and smooth muscle cells of the tunica media of smaller arteries, venules and the myometrium[179,202]. However, the relative roles and importance of direct virus-mediated endothelial cell injury and increased vascular permeability in the pathogenesis of EAV infection have not yet been clearly defined[179].

Recent studies of the pathogenesis of EAV infection have demonstrated that the virus localizes in $CD3^+$ T lymphocytes, $CD172a^+$ myeloid cells and a small population of IgM^+ B lymphocytes in the connective tissue[204]. Interestingly, the virus was not detected in epithelial cells in the upper respiratory tract during the acute infection period[204]. The nasopharynx and tubal-nasopharyngeal tonsils are suggested to be important primary sites of EAV replication[189]. Previous *in vitro* and *in vivo* studies with EAV have suggested that increased transcription of genes encoding the proinflammatory interleukin (IL) mediators IL-1β, IL-6, IL-8 and tumor necrosis factor (TNF-α) play a significant role in determining the severity of the disease[186]. Recently, Go *et al*. (2012) demonstrated experimental inoculation of horses with *in vitro* $CD3^+$ T cell susceptible and resistant phenotype with EAV generate different levels of proinflammatory and immunomodulatory cytokine mRNAs. Additionally, horses with *in vitro* $CD3^+$ T lymphocyte resistant phenotype developed more severe clinical signs than the horses with the susceptible $CD3^+$ T lymphocyte phenotype[205].

EAV infection can result in abortion in pregnant mares. Fetuses and placentas may be partially autolyzed at the time of expulsion[9,14,193]. Johnson *et al*. (1991) suggested that abortion in mares is caused by systemic vascular necrosis that weakens uterine and placental blood supply. It has also been suggested that the abortions are caused by the vasculitis of myometrial blood vessels with placental dysfunction[202]. MacLachlan *et al*. (1996) have suggested that abortion happens due to EAV infection directly affecting the fetus causing stress and interfering with the fetal hypothalamic-pituitary axis. There are no pathognomic gross or histopathological lesions associated with abortion. Interlobular pulmonary edema, pleural and pericardial effusion, and petechial and ecchymotic

hemorrhages on the serosal and mucosal surfaces of the small intestine were observed on necropsy of an aborted fetus[202].

With the exception of persistently infected stallions, viral clearance from serum and body tissues usually occurs at ~28 dpi and coincides with an increase in serum neutralizing antibodies[206,207]. However, an *in vivo* study by Vairo *et al*. (2012) recently reported that the virus could still be demonstrated in the tonsil at 28 dpi. In the carrier stallion, EAV primarily persists in the ampullae of the vas deferens[165,189,208,209]. Recently, Carossino *et al*. (2017) has determined the specific tropism of EAV for stromal cells and identified that the virus primarily persists in fibrocytes and T ($CD2^+$, $CD3^+$, $CD5^+$ and $CD8^+$) and B ($CD21^+$) lymphocytes in the ampullae[165]. The study indicated that EAV expresses different host-cell tropism depending on the tissue compartment and duration of infection (i.e. acute vs. chronic)[165]. For instance, EAV infects $CD3^+$ T lymphocytes and CD14+ monocytes, but does not infect $CD21^+$ B lymphocytes in the accessory sex glands of the stallion reproductive tract[87]. The authors concluded that EAV persistence may drive immunosuppression and/or T cell exhaustion to suppress the host immune response[165]. The viral strategy of host immune evasion has to be further determined.

Recently, Carossino *et al*. (2017) investigated the EAV tissue and cellular tropism in the reproductive tract and the inflammatory response induced during persistent infection. They used reproductive tract tissues from normal, short-term and long-term EAV carrier stallions in their studies. EAV tissue and cellular tropism was determined by virus isolation, nucleocapsid-specific single and dual immunostaining (IFA/IHC) and transmission electron microscopy (TEM). The inflammatory response was characterized by histopathology and IHC using a panel of differentiation (CD) markers. The ampulla

was demonstrated to be the tissue with the highest viral titer ($1.0 \times 10^3 – 1.7 \times 10^5$ PFU/g) among the accessory sex glands of carrier stallions. Dual IFA/IHC demonstrated the presence of viral antigen in fibrocytes and lymphocytes (mainly $CD3^+$ [$CD8^+$] and $CD21^+$) within the lamina propria and inflammatory infiltrates in the ampullae. The virus was not detected in the epithelium of the ampullae or any other accessory gland. The inflammatory response was characterized by multifocal lymphoplasmacytic ampullitis which included moderate and high infiltration of $CD4^+$/$CD25^+$ and $CD8^+$ T lymphocytes, respectively; clusters of $CD21^+$ B lymphocytes and moderate Iba-1^+ and $CD83^+$ macrophage and dendritic cell infiltration. Evidence of $CD25^+$ and $FOXP3^+$ T regulatory lymphocytes with lack of expression of IFN-gamma and limited expression of granzyme B in $CD8^+$ T lymphocytes indicate that immunoregulatory mechanisms in conjunction with $CD8^+$ T cell exhaustion are likely involved in maintenance of persistent infection. In summary, EAV was demonstrated to be primarily lymphotropic (including T and B lymphocytes with a restricted homing pattern to the ampulla) and fibrocyte-tropic (vimentin$^+$). Despite the strong local inflammatory and serum neutralizing antibody response, EAV evades host immunity in a tissue under immune surveillance and is capable of establishing long-term persistence in the male reproductive tract. Furthermore, EAV was not detectable in lymphoid tissues including those associated with lymph drainage from the reproductive tract, suggesting that infected T and B lymphocytes exhibit a specific homing pattern to the reproductive tract and restricted migration from reproductive tract tissues to secondary lymphoid organs.

1.3.4. Persistent EAV infection

As previously mentioned, EAV persists exclusively in the reproductive tract of infected stallions. During the acute phase of the disease (average range of first 21 days), the stallions show decreased libido and reduced sperm motility, concentration and percentage of morphologically normal sperm cells[156,190,191]. Subsequently, 10 to 70% of infected stallions and sexually mature colts remain persistently infected and continuously shed EAV in their semen for weeks, months, years or lifelong without alteration in the semen quality[143,156,157]. Viral shedding appears in semen at 5 dpi and is associated with the sperm-rich fraction of the ejaculate[157,210,211]. Viral titers in seminal plasma show a broad range (10^1 to $>10^7$ PFU/mL)[209,211]. Although the duration of the carrier state can vary, it has been arbitrarily divided into three categories: short-term carrier state (several weeks), intermediate carrier state (3 to 7 months) and the long-term persistent or chronic carrier state (years and even lifelong) [11,156].

The mechanism(s) involved in EAV persistence and the factors that lead to viral clearance in stallions is not yet defined. However, the establishment and maintenance of the EAV carrier state in the stallion is testosterone-dependent[11,209,212]. Thus, surgical castration of carrier stallions is the only certain method to cause cessation of viral shedding[212]. Previous studies have suggested that testosterone is strongly associated with an interaction between the host immune response and EAV in the reproductive tract such as immunosuppressive activities or maintenance of a susceptible host cell population that results in maintenance of EAV persistence[208,213–216]. It is proposed that high levels of testosterone in the stallion reproductive tract are important during long-term EAV persistence, but not for viral replication or persistence of intermediate duration[217]. An *in*

vivo study of attempted establishment of persistent EAV infection in prepubertal and peripubertal colts has demonstrated that the carrier state could not be established without the presence of testosterone during intermediate duration[217]. Thus, the carrier state of EAV infection is androgen-dependent, and so persistent infection cannot occur in prepubertal colts, mares, geldings and fetuses[11,140,156,209,212].

As indicated above, carrier stallions provide the means by which EAV evolves[11,17,19,21,142,146,164,166]. EAV continuously evolves via genetic divergence called EAV quasispecies during persistent infection in the stallion reproductive tract[142,163,164]. EAV quasispecies generate novel genetic variants with phenotypic diversity that contributes to the occurrence of EAV[142,163,164,218]. Significant nonsynonymous substitutions at high rates were observed in GP3 and GP5 proteins suggesting that ORF3 and ORF5 are under strong selective pressure in persistently infected stallions[16,17,142]. A limited number of critical amino acid substitutions were especially observed in the V1 variable region of the GP5 protein (aa 61 to 84)[16,218]. The emergence of viral variants from EAV quasispecies present in the long-term carrier may be endowed with novel genetic and phenotypic properties (i.e. neutralization) and could be susceptible for new outbreaks of EVA[17,142,156].

1.4. Diagnosis of EAV infection

1.4.1. Appropriate specimen collection for examination

Equine viral arteritis clinically mimics a number of other infectious and non-infectious diseases and therefore, it is not possible to establish a diagnosis of EVA based solely on the nature of clinical signs[89]. Thus, the clinical diagnosis of EVA should be

confirmed by laboratory diagnosis. During the acute phase of the disease, the most appropriate specimens include nasopharyngeal swabs or washings, conjunctival swabs and blood samples for separation of buffy coat cells (citrated or in ethylenediaminetetraacetic acid [EDTA]) for virus isolation (VI) or PCR[219]. From infected stallions, semen samples containing the sperm-rich fraction are essential for VI or PCR[11]. During the necropsy of aborted fetuses, clinical specimens including placenta, fetal fluids, lung, spleen and lymphoid tissues should be collected for VI, PCR or immunostaining to confirm cases of EAV-induced abortion[219]. All clinical samples listed should be stored and submitted under conditions of refrigeration via overnight delivery for laboratory testing. After a necropsy, tissue samples should be fixed and placed in 10% neutral buffered formalin for virus detection, histopathologic examination and immunohistochemical staining[220]. In suspicious cases of EVA in young foals, various organs and lymph nodes associated with the gastrointestinal and respiratory tracts should be collected[219]. Serum samples are collected for serologic diagnosis, and paired samples obtained at least 14 days apart or preferably a 21-to 28-day interval are recommended[219].

1.4.2. Virus isolation in cell culture

VI is the World Animal Health Organisation (OIE)-approved "gold standard" test for EAV detection, especially in semen samples from stallions and is the prescribed test for international trade[221]. VI should be attempted from specimens collected from clinical or fatal cases using rabbit, equine or monkey kidney cell cultures in accordance with established procedures[222]. The cell system of choice for VI is the rabbit kidney cell line (RK-13), and generally EAV can be isolated on first or second passage. However, some

strains may require additional blind passages in cell cultures[169]. Virus isolates should be further confirmed by cross neutralization assay using EAV-specific equine antiserum or specific standard or real-time RT-PCR[223].

1.4.3. Molecular diagnosis (nucleic acid detection by RT-PCR)

Molecular testing (i.e. detection of viral nucleic acids) has several advantages including high sensitivity, rapid turnaround, convenience and cost[224]. Several molecular diagnostic assays (reverse transcription polymerase chain reaction [RT-PCR]) have been developed including RT-PCR and RT-nested PCR (RT-nPCR), real-time RT-PCR (rRT-PCR) and insulated isothermal RT-PCR (iiRT-PCR) assays for EAV nucleic acid detection from clinical specimens[225–231]. The primers/probes have been designed to target conserved regions of the viral genome, and especially target the 3′-end of ORF1b of the viral polymerase gene and ORF7 (encoding N protein)[225,226,228,229].

The RT-nPCR assay is more sensitive than the standard RT-PCR but it tends to give false positive results due to cross contamination of a sample[219]. In light of this, the rRT-PCR which avoids handling of PCR amplicons has replaced gel-based assays[219]. Previous studies evaluated the rRT-PCR assay and proved an equal or higher sensitivity than VI for EAV detection[232]. Recently, a TaqMan® probe-based iiRT-PCR assay has been developed for qualitative detection of EAV nucleic acid from clinical specimens[233,234]. This system is based on the use of the field-deployable POCKIT™ Nucleic Acid Analyzer (GeneReach USA, Lexington, MA, USA) which allows point-of need use[234,235] including veterinary clinics, racetracks, breeding facilities and diagnostic laboratories[233].

Recently, two RNA *in situ* hybridization (ISH) assays (conventional and RNAscope® ISH) for the detection of EAV RNA in formalin-fixed paraffin-embedded (FFPE) tissues have been described[236]. The conventional RNA ISH assay had a significantly lower sensitivity as compared to the RNAscope®. The use of oligonucleotide probes along with a signal amplification system (RNAscope®) can enhance detection of EAV RNA in FFPE tissues.

1.4.4. Antigen visualization by immunohistochemistry

EAV antigen can also be demonstrated in FFPE tissues from aborted fetuses using nucleocapsid-specific monoclonal antibody (immunohistochemistry [IHC])[236]. Evaluation of 80 FFPE tissues collected from 16 aborted fetuses showed that IHC had a similar sensitivity to RNAscope®. The conventional RNA ISH assay had a significantly lower sensitivity as compared to the RNAscope® and IHC assays.

1.4.5. Serological diagnosis

The virus neutralization test (VNT) is the current OIE-approved "gold standard" test for EAV antibody determination required for international trade[221]. Although the VNT is highly sensitive and accurate, it has a few disadvantages including cost, labor, turnaround time and variability among different laboratories. It requires standardized reagents and protocol[224]. Furthermore, serum cytotoxicity due to anti-cellular antibodies directed against RK-13 cells can interfere with test interpretation at lower serum dilutions[224]. Also, the VNT cannot distinguish antibody responses between vaccinated and naturally infected horses[224]. To get over the disadvantages of VNT, several enzyme-

linked immunosorbent assays (ELISA) have been developed and evaluated to detect EAV-specific antibodies using whole virus, synthetic peptides or recombinant viral proteins (e.g. GP5, M and N) as antigens[215,219,237–243].

The GP5 is the major target as an antigenic protein of virus neutralization that has the potential for development of serological diagnostic assays[224]. Since the M and N proteins consist of conserved sequences among different EAV strains and isolates, these proteins also have the potential to be used as antigens for diagnostic purposes[244]. Firstly, an ELISA using a recombinant fusion protein expressing GP5 (aa 55-98) had a sensitivity and specificity of 99.6% and 91.1%, respectively[239]. Another ELISA based on the ovalbumin-conjugated synthetic peptide containing amino acids 81-106 of GP5 had a sensitivity and specificity of 96.7% and 95.6%, respectively[237]. Secondly, an ELISA using sera from naturally or experimentally infected horses against a cocktail of recombinant GP5, M and N proteins expressed in baculoviruses resulted in a sensitivity and specificity of 100% and 92.3% respectively; it was not able to detect antibodies in the sera from the vaccinated horses[215]. Thirdly, a blocking ELISA using partially purified EAV and MAb against the GP5 protein was developed, and it provided a sensitivity and specificity of 99.4% and 97.7%[238]. Fourthly, an ELISA using the carboxyl terminus of the M protein (aa 88-162) is frequently recognized by EAV-specific equine sera[245]. Lastly, a microsphere immunoassay (MIA) has been developed using a combination of microspheres and three well-developed technologies: bioassays, solution-phase microspheres and flow cytometry[224]. However, all of the above mentioned ELISA assays were neither sensitive nor specific enough to replace the VNT. Recently, a commercial

cELISA assay (VMRD, Pullman, WA) was developed that showed high sensitivity and specificity compared to the VNT, but this needs further evaluation[219].

1.4.6. Differential diagnosis

A range of viral (e.g. equine herpesvirus 1 and 4 infections and equine influenza), and bacterial (e.g. Streptococcus) respiratory diseases that commonly affect horses cannot be distinguished clinically from EVA. Other infectious diseases that can cause systemic signs of illness resembling EVA are equine infectious anemia, dourine, African horse sickness fever and Getah virus infection. EVA also bears many many clinical similarities to the syndrome caused by hoary alyssum (*Berteroa incona*) toxicosis. Abortions or deaths in foals due to EAV need to be differentiated in particular from those caused by equine herpesvirus 1 as well as other infectious causes of abortion in the mare[97,98].

1.5. Treatment of EVA

1.5.1. Antiviral therapy

Even though several antiviral medications have been investigated, there is still no effective antiviral treatment for EAV infection[150,219,246] A potential antiviral compound (phosphorodiamidate morpholino oligomers [PMOs]) has been evaluated[150,246]. PMOs are 20-25 bases in length of single-stranded DNA analogues that tend to be water-soluble and nuclease-resistant[247,248]. PMOs form base pairs with complementary viral RNA target sequences to interfere with viral gene expression and its translation[248,249]. Additionally, peptide-conjugation to the PMO (P-PMO) is involved in transporting more PMO into infected cells[250]. The 5′ UTR of the EAV genome is the most sensitive target site of P-

PMOs[246]. According to *in vitro* studies, 5-10 uM of antisense P-PMO was able to target the EAV 5′ terminus from persistently infected HeLa cells[150]. However, the *in vivo* efficacy of P-PMO therapy needs to be evaluated.

1.5.2. Anti-GnRH vaccination and GnRH antagonists

Previous studies have demonstrated that the application of gonadotropin-releasing hormone [GnRH] antagonists, anti-GnRH vaccination and single-layer centrifugation[251–259] can facilitate viral clearance in some stallions or eliminate semen infectivity in some cases. However, none of them has 100% efficacy in inducing complete viral clearance from the reproductive tract[154]. For instance, Burger *et al*. (2006) have demonstrated in a vaccination study in EAV carrier stallions injecting twice 4 weeks apart with anti-GnRH vaccination resulted in cessation of virus shedding in semen 4 to 6 months after the first immunization and from all subsequent semen collections[154,255]. However, all the vaccinated stallions developed reduced libido, scrotal size, total sperm number and percentage of normal sperm indicating that this protocol would not be a suitable treatment for EAV-infected breeding stallions[154].

1.6. Disease prevention and control

The purpose of EVA control and prevention strategies is to prevent the risk of spread of EAV in horse populations[149,260,261]. In the USA, the standards of EVA control and prevention are described in the Equine Viral Arteritis: Uniform Methods and Rules developed by the USDA-APHIS[260]. Several European countries: United Kingdom (UK), Ireland, Germany, France and Italy established a Code of Practice which provides

guidelines for the control and prevention of several equine diseases including EVA[262]. Globally, the OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals and the Terrestrial Animal Health Code[261,263] provides standards for EVA laboratory testing procedures for international trade of horses and semen. Detailed information on the application of this testing process is also available at the American Association for Equine Practitioners (AAEP) (www.aaep.org).

It is important to correctly identify carrier stallions by determining their serological status and whether they have a prior history against EVA[154]. All stallions should be tested for the presence of serum neutralizing antibodies at least 60 days before the breeding season[154]. If the stallion is seropositive (neutralizing antibody titer $\geq$ 1:4) without a vaccination history or confirmation of seronegative status prior to vaccination, virological assessment of semen is required to determine its status[154]. The collection of two semen samples containing the sperm-rich fraction of the ejaculate is required[154]. The test can be performed on samples obtained on the same collection day, consecutive days, or after an interval of several days or weeks[154]. Another way of identifying EAV carrier stallions is to perform test breeding by which two seronegative mares are bred or inseminated twice, each on two consecutive days, for a total of four covers[154]. These tested mares will be isolated for 28 days until checked for the presence of neutralizing antibodies[154]. If one or both mares show seroconversion after being bred, this is indicative that the stallion is persistently infected or the semen used is infective[154]. Carrier stallions can still be used for breeding purposes under strict conditions: physical isolation and breeding to seropositive mares only as indicated previously[11,157,264–266]. With

exception of the USA, other countries do not accept the introduction of EAV carrier stallions or EAV-infective semen[154].

It is highly recommended to vaccinate horses against EVA if they frequently travel and interact with other horses at competitions[154]. The OIE Terrestrial Animal Health Code provides the general recommendations for the international movement of horses and importation of equine semen[154]. Horses should be subjected to a quarantine period of 28 days without clinical signs of EVA, a stable or declining EAV antibody titer or a seronegative status between two serum samples collected at least 14 days apart[154]. When the state veterinarian receives the notification of EVA outbreak, it is important to suspend the routine activities (training or breeding) and animal movement and isolate all of the clinically affected or in-contact animals by following AAEP guidelines. The immediate quarantine will allow vaccination and laboratory confirmation to prevent wider spread of EVA, associated abortions, neonatal death and persistent infection in stallions from the horizontal mode of transmission of EAV[21,154,267]. Appropriate clinical sample collection should be immediately performed with the cooperation of farm managers and veterinarians for laboratory testing[154]. Disinfection of stalls and equipment is recommended using a disinfectant including phenolic, chlorine, iodine and quaternary ammonium compounds to inactive EAV[154]. The quarantine will be continued for 3 to 4 consecutive weeks after the last confirmed case is reported[154].

1.6.1. Vaccine

There are two commercially available vaccines for the prevention of EVA, a modified live attenuated (MLV; ARVAC®) and an inactivated product (Artervac®). The MLV vaccine was developed by serial passage (P266) of the VBS strain in horse kidney (HK-131), rabbit kidney (RK-111) and equine dermal (ED-24) cells. It has been used since 1985 in the USA and Canada[146,154,159,161,206,268,269] The MLV vaccine has been proven to be safe and effective in providing protective immunity in vaccinated horses[161,175,206,268]. However, MLV vaccination does not necessarily prevent reinfection of first-time vaccinated horses so that revaccination is encouraged[146,161,188]. The MLV vaccine induces neutralizing antibodies within 5 to 8 days post-vaccination, with a peak in antibody titers occurring between 7 to 14 days post-vaccination[146,161,175,188,206,268–270]. Neutralizing antibodies may last for at least 2 years. It has been determined that a titer of at least 1:64 is required for protection[271]. Few vaccinated horses may develop fever with transient lymphopenia, and the attenuated virus can be detected from nasal secretions and buffy coat cells up to 7 days post-vaccination in the vast majority of the horse population[146,146,159,268,270,272,273]. Additional to the MLV product, an inactivated EAV vaccine (Artervac®) has been developed and is licensed in several European countries including the UK, Ireland, France, Hungary and Denmark[154]. As Artervac® has a lower immunogenicity compared to the MLV vaccine, a booster vaccination 3 to 4 weeks after the primary vaccination is recommended with semiannual revaccinations. Although this vaccine stimulates neutralizing antibodies, its efficacy against EAV infection and prevention of the establishment of persistent infection in stallions is substantially less characterized than the MLV vaccine[154].

In the USA, it is highly recommended to immunize seronegative stallions with the MLV vaccine (ARVAC®) at least 28 days prior to the breeding season, followed by an isolation period of 21 days after the primary vaccination to allow maximal development of the VN antibody response and to avoid exposure of other animals during the transient shedding period of the vaccine virus[260]. An *in vivo* study reported that a low level of vaccine virus (<1 PFU/mL) was detected in the semen of one stallion at 4 to 6 days post-vaccination[270]. Annual revaccination prior to the breeding season is highly recommended[154]. Unfortunately, vaccinating persistently infected stallions will not induce clearance of the carrier state[154]. Most importantly, the current VN test is not able to distinguish between vaccinated and naturally infected horses, thus it is critical to confirm the stallions' seronegative status before primary vaccination. The MLV vaccine is effective in the prevention of EVA outbreaks and establishment of carrier state in stallions[154].

Carrier stallions should only be bred to EAV seropositive mares which were naturally infected or vaccinated against EVA[146]. The mares should be vaccinated at least 3 weeks before breeding followed by a 21 day period of isolation after breeding as previously indicated[154]. Neutralizing antibody titers should be determined (≥ 1:64) in EAV seropositive mares at least 30 days before breeding to a persistently infected stallion or using infective semen for artificial insemination or embryo transfer[274,275]. Following breeding, the mares should be kept in isolation for at least 21 days away from EAV seronegative horses[154]. The use of the MLV vaccine is not recommended in pregnant mares especially in the last 2 months of gestation to avoid abortion[276]. Vaccination is also not recommended in foals less than 6 weeks old; they should be vaccinated at ~6 months

of age when maternal acquired antibodies are no longer detectable[154]. So far, the efficacy of the MLV vaccine has been proven in the prevention of EVA outbreaks and persistent infection in stallions[154].

As previously mentioned, the currently available VN test is not able to differentiate vaccinated from infected animals (DIVA). The development of a DIVA vaccine would contribute to the equine industry by facilitating the serological identification of infected from vaccinated animals and supporting surveillance programs for national and international movement of horses[154]. Even though several studies have been undertaken to develop new vaccines, none of them have thus far reached the market. The development of an infectious cDNA clone of the MLV strain (pEAVrMLVB) showed itself to be safe and capable of stimulating protective immune response comparable to the MLV vaccine (ARVAC®)[188]. This infectious clone has the advantage that it could be further manipulated by genetic engineering to generate a second-generation DIVA vaccine[154]. The development of subunit vaccines would also be another suitable alternative as a second-generation DIVA[224]. DNA vaccination has been tested in mice[277]. The vaccine with plasmid DNA encoding ORF5 generated the highest VN antibodies in vaccinated animals[277]. A DNA vaccine expressing a combination of viral genes (ORFs 2, 5 and 7) and equine IL-2 was also tested in horses resulting in high VN antibodies titers, which lasted for at least a year[278]. Lastly, Balasuriya *et al*. (2002) used alphavirus replicon particles (VRP) to co-express the major viral envelope proteins (GP5 and M) as a recombinant vaccine candidate. The *in vivo* study using mice established VRP co-expressing GP5/M heterodimer that was able to induce neutralizing antibodies

against EAV[177]. However, all of the above vaccine candidates require further testing to determine their safety and efficacy.

2. Evolution of RNA viruses

Where did viruses come from and how did they survive? Unfortunately, very few records of viruses are available unlike the situation for other cellular organisms about which we are able to infer from fossils[279]. However, the application of modern technologies such as next-generation sequencing (NGS) and reverse genetic engineering has helped to provide some answers concerning the origins and evolution of viruses[279,280]. By comparing the sequences of viral genomes, it is now possible to estimate the degree of evolutionary relationship between ancestral viruses using bioinformatics, mathematical and phylogenetic techniques[279,280].

Viruses evolve as a result of reassortment or recombination between virus strains in addition to nucleotide substitutions introduced during the replication cycle[279]. RNA viruses in particular rapidly produce large populations in a short period of replication[281]. Most RNA viruses encode for a low fidelity RNA-dependent RNA polymerase (RdRp) enzyme that can result in a mutation rate of 10^3 to 10^5 errors per nucleotide per replication cycle. This rate is 10,000-fold lower than that which occurs during DNA virus replication[282–284]. As a result, DNA viruses evolve more slowly than RNA viruses[279].

A continuous process of genetic variation in RNA viruses is responsible for heterogeneous mutant distributions also termed "mutant swarms or mutant clouds or quasispecies"[280,281]. Viral quasispecies allow RNA viruses to adapt to a given environment with the advantages of dynamics of infection, virus spread, attenuation and

virulence, complexity and self-organization[285,286]. The main features of viral quasispecies, biological implications of quasispecies dynamics and the specific viral system related to viral quasispecies will be further discussed. RNA viruses evolve by four mechanisms in nature and each of these are described in detail below.

The accumulation of nonsynonymous mutations in the genome will lead to antigenic drift which is one of the main mechanisms involved in the evolution of influenza viruses[287]. Antigenic drift is the result of point mutations in influenza virus gene, encoding HA and NA that cause alternations in the structure of the main viral surface antigens. This is a slow process that is on-going and which causes antigenic differences in strains of influenza virus[288]. The evolutionary process sometimes is under certain limitation; for example, the viral attachment proteins must maintain the eligible configurations and catalytic abilities for the virus to be able to bind to the target host's cell receptor and enzymes (e.g. reverse transcriptase)[279]. The RNA viral genome undergoes virus evolution via the quasispeices effect (the establishment of mutant spectra)[280]. Once the RNA virus enters the susceptible host cell, the virus replicates itself rapidly and generates numerous numbers of virus population. Due to lack of proofreading capability of RdRp, the RNA viruses generate numbers of mutations in the virus population that is called mutant spectra (mutant swarm or mutant clouds). The mutant spectrum is the ensemble of genomes that forms a viral quasispecies, and its complexity and formation are closely related to its biological environment. The complexity of a mutant spectrum can be demonstrated in the nucleotide sequences obtained either by classic molecular cloning and Sanger sequencing or by next-generation sequencing (NGS)[285].

2.1. Mechanisms of virus evolution

2.1.1. Accumulation of mutations

RNA viruses mainly evolve through mutation errors that are generated during viral replication. Although all polymerases cause errors by inserting incorrect nucleotides into the template during viral replication, the DNA-dependent DNA polymerases have proofreading ability in their exonuclease activities that enable them to modify or correct any such mistakes. In contrast, RdRp does not have proofreading ability as previously mentioned[284]. If the errors are located in the protein-coding sequences, then it introduces random substitutions in the encoded amino acid sequence[279]**.** The theory of Darwinian natural selection applies to the mutations that ensure the continuous survival of the viruses[279].

2.1.1.1. Concept of viral quasispecies

Manfred Eigen and Peter Schuster developed the quasispecies theory by looking at the quasispecies (dynamic mutant distributions) which are displayed as a dominant master sequence of the virus population[289]. Then, Wiessmann and his colleagues provided the first evidence of quasispecies distribution in RNA viruses by the bacteriophage Qβ virus using T_1 oligonucleotide fingerprinting through reverse genetics. Serial passages of bacteriophage Qβ virus clones in E. *coli* demonstrated competition between wild-type virus and mutant clones[290,291]. Interestingly, the rate of mutation of bacteriophage Qβ virus was 10^4-fold constantly higher than the values of DNA bacteriophage[292]. This was the first calculation of the rate of mutation for an RNA virus that demonstrates the quasispecies dynamics using bacteriophage Qβ virus. The

quasispecies theory formulates an error-threshold relationship at the maximum error rate compatible with maintaining genetic complexity in a replicative system[289]. The value of the error-threshold is determined by the replication accuracy and the virus fitness of the dominant or master sequence corresponding to the mean fitness value of the error copies[289]. Eigen and other authors proposed that the quasispecies contains the finite replicon populations expanded under nonequilibrium conditions of variable fitness landscapes[293–304]. The quasispecies theory can be mathematically determined based on the dynamics of error copy production and the error threshold relationship and that describes the evolutionary dynamics model[305]. The quasispecies theory introduces a molecular view of evolutionary biology as well as viral quasispecies evolution[285]. The alternative terms for quasispecies include "Mutant swarm, intrahost variation, intrahost diversity, mixture of mutants, nucleotide degeneracy, hyperploidy, heterospecies and heteropopulations"[306].

2.1.1.2. Major characteristics of viral quasispecies

Viral quasispecies is defined as the continuous process of genomic distributions of genetic variation, competition and selection[307,308]. There are two types of selection: positive and negative that influence the generation of viral quasispecies in the virus population. Positive selection is the process of a dominant genotype resulting in the positive evolution of phenotypic traits expressed by the individuals of the evolving population[285]. For instance, the viral quasispecies of a reconstructed foot-and-mouth disease virus (FMDV) generate a dominant mutant distribution by antibody selection[309]. Negative selection is the process of removing a genotype in an evolving population that

causes the negative evolution of phenotypic traits expressed by individuals[285]. However, the outcome of negative selection is sometimes the same as positive selection or vice versa during a competition process[285]. During replication of a viral subpopulation, positive selection increases frequency of subsets of genomes with a net fitness gain[310–312]. This increased fitness of the resulting subpopulations is maintained as memory genomes[313]. The quasispecies memory is present in the mutant spectrum at higher frequencies than mutation rates[313,314]. There are two parameters related to characterization of the viral quasispecies: mutation rate and mutation frequency[285]. Mutation rate is the frequency of occurrence of a mutation during genome replication that can be calculated by mutations per site per replication. It is a value that is independent of the fitness of the parental and mutated genomes[285]. Mutation frequency is the proportion of mutant viruses in a population and is affected by many biochemical and environmental selective factors[285].

Additional to variation and continuity of core information of a biological adaptive system, the viral quasispecies tends to maintain its memory of past events[313]. Evolving viral quasispecies maintains the molecular memory as a form of minority components in their mutant spectra and it is called viral quacispecies memory[313]. The viral quasispecies memory plays an important adaptive role in the host environment such as the immune system and results in the quasispecies dynamics[285]. During viral replication, memory genomes with fitness gain in the most frequent genomes of the same quasispecies is subjected to the Red Queen hypothesis[315,316]. This hypothesis defines that the continuous selection of the most fit genomes in competing viral populations[317–321]. However, the

quasispecies memory level can be gradually lost in the parallel lineages of replicating populations and that is called memory decay[321].

The viral fitness and the effect of population size, known as bottleneck events, are another important characteristic of viral quasispecies[285]. Viral fitness is a major parameter which results in a viral population adapting to a specific environment and producing infectious progeny under the formulation of quasispecies theory and the error threshold relationship[282,285]. Cristina Escarmís and her colleagues have demonstrated the molecular basis of fitness loss associated with successive bottleneck events in biological clones of FMDV. Some numerous, unusual mutations were observed in the large viral population following *in vitro* passage of FMDV; these were not found in any field isolates or clones[322–324] Hundreds of bottleneck passages have driven an evolution of noncytopathic variants of FMDV to be capable of establishing persistent infections in BHK-21 cells[325–327]. The constant steady replication of large viral populations result in a consequence of high fitness called adaptation to new environment; in contrast, the repeated bottleneck events with accumulation of mutations result in low fitness, is called de-adaptation in the environment[280] As the result of a reduction in fitness and the biological changes in viruses, repeated bottleneck event transfers mediate distributions of neutral mutants in mutant spectra[285,328,329]. There are two types of population bottleneck events called plaque-to-plaque transfers and intrahost bottleneck intensity. In the case of plaque-to-plaque transfers, a single infectious genome undergoes replication and diversification[285]. Depending on a different intensity, the number of genomes from a larger ensemble starts replication. Following which, a small number of genomes from the larger ensemble will have limited variants by selective constraint[285]. During the replication cycle of a virus,

viruses infect through host-to-host interaction as well as intrahost bottleneck intensity that influences the intrahost evolutionary events[330–340]. The bottleneck events happen to modify viral population sizes resulting in the heterogeneous and dynamic nature of viral populations that are suitable in the environment by fitness gain[341–343]. Bottleneck events occur in both short-term and long-term periods of virus evolution[285].

Virus evolution does not always occur by modifications in consensus sequence, it also happens through non-equilibrium of a mutant spectrum in variable fitness landscapes, a theoretical construction of the evolutionary process[285,344]. In contrast, population equilibrium plays an important role in maintaining a constant consensus sequence in viral populations[345,346]. The positive selection influences the subpopulation of viral genomes in the mutant spectrum increasing in frequency or replacing the previous distribution and this results in a new consensus adapted to the biological environment[285]. The virus undergoes the evolutionary process in a dynamic mutant spectrum either by evolutionary stasis (rates of evolution of around 10^{-4} substitutions per site per year or lower) or rapid evolution of its consensus sequence (rates of evolution of about 10^{-1} to 10^{-2} substitutions per site per year) depending on its given environment[286]. In the case of avian influenza virus, the evolutionary stasis in the natural target host will not give rise to a disease outbreak, but the rapid evolution of the same virus will cause a disease outbreak in different mammalian hosts when the virus starts to replicate[347]. Another viral quasispecies parameter is the rate of evolution, which is the number of accumulated mutations in viral genomes through a time point[285]. The major difference from the previously mentioned mutation rate and frequency is counting the time factor[285]. The unit of rate of evolution is expressed as the number of substitutions per site per year. This may

affect replication of a viral population during short-term (intrahost) or long-term (interhost) virus evolution[285].

2.1.1.3. Biological outcomes of quasispecies dynamics

As previously mentioned, the major characteristics of viral quasispecies produce several outcomes in the virus-host relationship including host range mutant and viral emergence, escape mutant, mutagen-resistant mutants, virus extinction and long-term virus evolution[285]. Firstly, the viral quasispecies dynamics influence modifications in the cell tropism and host range during virus evolution[285]. Viruses use cellular receptors which are composed of different macromolecules such as proteins, lipoproteins, glycoproteins, glycolipids and glycosoaminoglycans to attach and enter the host cells[348–353]. For instance, canine parvovirus-2 (CPV-2), which causes highly contagious after fatal disease in dogs, is closely related to feline parvovirus (FPV) and therefore assumed to be a host variant of FPV[354–356]. As the result of continuous genetic variation, two or three genetic mutations in FPV caused the emergence of a novel strain of CPV and allowed it to expand its host range to infect dogs[357]. Either amino acid substitutions on external surface proteins or exchange of modules such as encoding relevant protein motifs by recombination affects the modification of the host cell tropism[285]. Presence of a single or few amino acid substitutions in a mutant spectrum may influence the host cell tropism and define the phenotypic characteristics of a virus[285]. After serial passage of FMDV in cell culture, amino acid substitutions were observed in the capsid surface that allows FMDV binding to heparan sulfate (HS) on the cellular surface[348,349,358–360]. These HS-binding variants of FMDV are also present in infected cattle tissues[361]. Modification of antigenic sites also

impact the antibody dependent enhancement (ADE) of viral infection, a mechanism used by several viruses[362]. Depending on the cross-linking of complexes of virus-antibody or virus-activated complement components through infection with cellular molecules such as Fc receptor on certain immune cells, the virus uses ADE mechanism to enhance its infectivity to the susceptible host cells[362]. Analysis of the mutant spectrum composition may enable one to predict the potential evolutionary outcomes related to receptor recognition[285]. Interestingly, host cell tropism and viral pathogenicity are closely associated with substitutions in the level of the structural or nonstructural viral proteins[285]. For instance, amino acid substitutions in the polymerase subunit PB2 of influenza virus promotes adaptation to alternative hosts and enhances the viral replication cycle[363–367].

Secondly, viral quasispecies facilitate the emergence of escape mutants which are the subset of mutants observed in the viral population that are associated with continuous viral replication despite the presence of selective constraint in the host environment[285]. A virus may generate mutations for the purpose of defending itself against the host immune responses or selective constraints. Selective constraints include antiviral drugs, immune components of the innate or the adaptive immune response including those induced by vaccination and interfering RNAs[285]. By considering frequency factors of inhibitor-resistant amino acid mutants, genetic barriers and the fitness cost represented in mutant spectra, the frequency of inhibitor-resistant mutants of viruses is about in the range of 10^{-3} to 10^{-5} [368–371]. Similar to the positive selection events, one or a few point mutations or multiple mutations by recombination events can initiate resistance to an antiviral drug[285]. According to Domingo *et al*. (2012), the drug resistance for human immunodeficiency virus 1 (HIV-1), hepatitis B virus (HBV) and hepatitis C virus (HCV) is systematically

reviewed as follows[285]: (i) Drug resistance is a general phenomenon and the selection of a drug-resistant mutant against a specific antiviral drug is evidence of the phenomenon[368,371–373]. (ii) Multiple different mutations or combinations of mutations called mutational pathways can lead to drug resistance[374,375]. For instance, multiple amino acid replacements at different sites in the protein may decrease affinity for a drug or a target protein complex[285]. (iii) High-level resistance to a drug may require mutations that are different from mutations associated with low-level resistance. Different frequencies of these mutations have been identified to the same drug[376–379]. (iv) Substitutions that are influenced by drug resistance increase viral fitness and absence of the drug decreases fitness[285]. (v) When multiple drug-resistant mutations occur, only some mutations are selected for a fitness-enhancing effect, but not all of the mutations will be directly involved in reducing affinity to the drug[285]. Vaccination is another selective constraint that exerts evolutionary forces on viruses[380]. For instance, Hepatitis A virus (HAV) has a single serotype, and its vaccine escape mutants have been identified in the complex mutant spectra of the viral population[381–384]. HAV requires rare codons in the capsid-coding region to regulate ribosomal traffic for proper capsid folding proteins and to prevent antigenic variants producing a new serotype[317,385]. However, the HAV escape mutants continuously evade adaptive immune responses of vaccinated immunocompromised individuals[348,386,387]. Multiple selective constraints establish the adaptive immune response in the host to block virus replication and remove infected cells through neutralizing antibodies derived from virus specific T cells[285]. Also, investment in immunotherapeutic approaches has the potential of eliminating viruses by targeting highly conserved epitopes or neutralizing antibodies in the conserved regions of the viral

proteins[388,389]. The development of effective antiviral drugs, vaccinations and immunotherapeutic approaches is a challenge when it comes to inhibiting mutant viruses.

Thirdly, evidence of virus extinction observed in mutant spectrum complexity occurs by an increase in the mutation rate above an extinction threshold. This suppresses the replication of high fitness variants through a process called lethal mutagenesis[390]. The hallmarks of lethal mutagenesis are increased frequency of defective genomes and a constant consensus sequence[391–393]. The types of defective mutants are called defectors and are associated with virus extinction caused by lethal mutagenesis[280]. This could be a potential approach for antiviral protocols by targeting mutagenic agents[285]. During lethal mutagenesis of viruses, there is evidence of a precarious mutant spectra suppressing the virus replication of high fitness variants[390]. For example, populations of FMDV that are resistant to polyclonal antibodies tend to have reduced fitness compared with wild types of the virus[394]. The mutant spectrum interactions are affected by the interaction between mutagenic agents and also by potential inhibitors of sequential treatments which can be focused on lethal mutagenesis[280]. Combination therapies with two or more inhibitors are generally used to prevent or reduce the production of inhibitor-resistant mutants[280]. Additionally, combination of therapeutics of mutagenic agents and inhibitors will have the advantage of reducing viral load and further result in virus extinction[280].

Lastly and most importantly, the viral quasispecies dynamics will drive long-term virus evolution *in vivo*[285]. While reviewing published information about the evolution of RNA viruses and quasispecies evolution, two questions arise[285]. Firstly, does the outcome of experimentally designed evolution studies reflect the viral population dynamics in the natural environment? Fortunately, the findings suggest there is no discrepancy in virus

behavior between experimental and natural infections in terms of genetic variation, selection competition within the mutant spectrum or selection of genome subsets[285]. Secondly, why should the consequences of quasispecies dynamics be considered for understanding long-term virus evolution? This is because the mutant spectra determine which type of genomes will be spread into susceptible target hosts[285]. Frequent complex cycles of the quasispecies swarm and transmission influence long-term evolutionary patterns by alternating the constitution of mutant spectra[285]. Viral lineages can undergo slow or rapid evolution; these are related to pathogenicity[395]. For example, the genetic diversity generated by a high polymerase error rate impacts the pandemic occurrence of norovirus[396]. Also, viral fitness level determines the severity of dengue disease[397]. The quasispecies dynamics proposes that RNA viruses evolve as a consequence of disequilibria in mutant distributions instead of linear accumulation of mutations[285]. However, there are contradictions concerning dating of viral ancestors and rates of co-divergence between viruses and hosts[398,399]. A study involving phylogenetic analyses of divergence has been conducted on HIV-1 and HIV-2 based on the common ancestor of retroviruses and using strict or relaxed clock models that sometimes give conflicting results[400]. The steady accumulation of mutations has randomly selected from the quasispecies that was amplified in the susceptible hosts[401]. From the average rates of virus evolution in the range of $1x10^{-1}$ to $3x10^{-5}$ substitution per site per year[402], phylogenetic analyses interestingly showed that the virus isolated in a short time span presented a higher rate of evolution than isolates after a long time[402,403]. The operation of a molecular clock follows the neutral theory of molecular evolution that means the rates of evolution reflect mutation rates[285]. However, this does not support the quasispecies

dynamics concept with reference to dating of viral ancestors, rates of co-divergence of viruses and their hosts[398,399]. These discrepancies might arise because of insufficient data concerning the population disequilibria to support the short-term and long-term evolution under the assumption of a linear accumulation of mutations[285]. Overall, the quasispecies dynamics through disequilibria of mutant swarms plays an important role during short-term intrahost and long-term evolution[285]. In conclusion, the virus evolution and viral population biology require continuous investment with regard to a rapidly replicating, error-prone, viral quasispecies in integrated biological systems[285].

2.1.2. Reassortment

RNA viruses with segmented genomes such as reoviruses, bunyaviruses and influenza viruses evolve through a process called reassortment[279]. When a susceptible host cell is coinfected by two strains of a virus, the progeny virus may have mixture of genome segments from the two parental strains of a virus that are called reassortants[279]. For instance, a susceptible cell is infected by two strains of influenza A virus that have unique envelope glycoproteins: hemagglutinin (H) and neuraminidase (N). After replication, the combination of parental RNA segments encoding H and N proteins now produce new progeny viruses[279]. This continuous reassortment process results in new reassortants of influenza A viruses that were responsible for three pandemics of influenza during the 20th century[279]. The largest scaled pandemic known as “Spanish Flu” occurred in 1918-19 resulting in the number of people killed by this influenza virus “H1N1” was greater than the number of deaths during the First World War[279].

2.1.3. Recombination

Many RNA viruses exchange nucleotide sequences among RNA molecules of different genomes producing new genomes[284]. This process is called RNA recombination; RNA viruses evolve by rearranging their genomes or functional parts among different viruses[284]. Recombination also occurs in DNA viruses[279]. There are two mechanisms of recombination, namely base-pairing dependent and base-pairing independent. Base-pairing dependent occurs between identical nucleotide sequences and base-pairing independent happens between variable nucleotide sequences[284]. Recombination plays a role in viral pathogenesis; for example, the noncytopathic bovine viral diarrhea virus (BVDV) can become cytopathic BVDV causing severe fatal gastrointestinal disease in cattle. The nonstructural protein 3 (NS3) synthesis of BVDV is critically associated with pathogenicity, and the base-pairing independent RNA recombination generates extra protease cleavage sites at the N terminus of the NS3 proteins[284]. Both DNA and RNA viruses use recombination mechanism for their evolution[279]. Many DNA viruses usually do not enter the nucleus during viral replication, except for herpesviruses that enter into the nucleus closely interacting with target host cells to mimicking cellular proteins. For instance, human herpesviruses can similarly modulate the host immune system including major histocompatibility complex proteins and cytokines (e.g. IL-10)[279].

3. Research objectives

Viruses have coexisted with human and animal populations on the earth for millions of years. They are considered a parasite, because the virus cannot replicate itself without the host organism. The virus is a causative agent of infectious disease in allliving

organisms. Most virus infections in mammals are cleared following development of the host immune response. However, some viruses have evolved to establish long-term persistent infection in their natural hosts in the presence of strong host immune response. Such viruses could either establish latent infection (dormant infection) or produce smoldering infection in a specific tissue or organ and continue to produce infectious progeny virus (persistent infection), which can be transmitted to susceptible naïve animals. These viruses have evolved mechanisms to evade the host immune response. In addition, these viruses tend to evolve during persistent infection generating novel genetic and antigenic variants that can precipitate new disease outbreaks. EAV is one of the classic examples of persistent viral infections in the natural host. Following natural and experimental infection, the virus establishes persistent infection in the stallion reproductive tract. Interestingly, 10% to 70% of the stallions that are infected with EAV become persistently infected carriers and continue to shed the virus in semen for several weeks to years to lifelong[11]. The primary objectives of the this research is to study EAV evolution during persistent infection by determining the full-length genome sequences of sequential EAV isolates present in the semen of long-term EAV carrier stallions by next-generation sequencing (NGS). We have included virus strains from both experimentally (49 sequential isolates over a two-year period from seven experimentally infected carrier stallions) and naturally infected (4 sequential isolates over a period of 10 years) stallions. Both experimentally and naturally infected stallions were exposed to the KY84 strain of EAV (EAV KY84). We used next-generation sequencing (NGS) to determine the complete genome sequence of each strain. Subsequently, the annotated full-length genome sequences were obtained, processed and subjected to complex bioinformatics

analyses. In this study, we investigated the intra-host evolutionary dynamics during acute and persistent infection periods. Furthermore, we characterized the virus quasispecies diversity, rate of evolution and selective pressure during the different time periods (e.g. acute and chronic) of EAV infection. Chapter two focuses on the evolution of EAV KY84 strain during acute and persistent infection in the stallion.

EAV infection is highly species-specific and only infects members of the family *Equidae* such as horses, donkeys, mules and zebras[10,11]. Recently, we have isolated a strain of EAV from a feral donkey in Chile. The second objective of this research was to obtain the full-length genome of this EAV donkey isolate using NGS. Thus, the second objective of this study is to analyze genomic, phylogenetic and antigenic characterization of this field strain of EAV from a donkey, which is a different member of the family *Equidae*. Comparative nucleotide and amino acid sequence analysis and phylogenetic analysis demonstrated that there is a unique relationship between EAV strains from horses, as well as other donkey isolates from Italy and South Africa. Subsequently, EAV strain-specific antisera and neutralizing monoclonal antibodies against the prototype EAV Bucyrus strain were used to determine the antigenic variation of the Chilean donkey isolate. Chapter three is focused on the molecular characterization and evolution of the Chilean donkey strain.

As previously mentioned, virus isolation is the OIE-approved "gold standard" test for EAV detection[221]. It is known that the high passage rabbit kidney (RK-13) cell line that is used for EAV isolation is contaminated with noncytopathic bovine viral diarrhea virus 1 (BVDV). During NGS sequencing of the EAV isolates from RK-13 cells, we were able to obtain the full-length genome of the contaminating BVDV. The third

objective of this research was to assemble the complete genome sequence of BVDV-1 strain contaminating the RK-13 cells. Chapter four is dedicated to the comparative nucleotide amino acid sequence analysis of the BVDV-1 strain.

CHAPTER TWO

Genetic Bottleneck and Selection of Equine Arteritis Virus During Acute Infection and Intra-host Quasispecies Diversification During Long-term Persistent Infection in the Reproductive Tract of the Stallion

2.1. Summary

Equine arteritis virus (EAV) is a small enveloped, positive-sense, single-stranded RNA virus in the *Arteriviridae* family. It infects the upper respiratory tract of equids (horses, donkey and zebra) with initial replication in nasopharyngeal epithelium and alveolar macrophages. Systemic spread is through infection of cells of the monocyte lineages and $CD3^+$ T lymphocytes. It establishes persistent infection in the reproductive tract of 10 to 70 % of infected stallions. Such stallions continually shed EAV in the semen ensuring successful transmission of EAV to susceptible naïve horses. In this study, seven EAV seronegative stallions were experimentally inoculated with EAV KY84 strain and followed for 726 dpi. Nasal secretions (swabs), unclotted blood and semen were sequentially collected, and a total of 53 sequential viruses were deep sequenced to elucidate the intra-host micro evolutionary process of EAV after a single transmission event. In addition, the viruses from two long-term EAV carrier stallions (seven to ten years duration) were deep sequenced; these carrier stallions were naturally infected with the same KY84 strain of EAV. The intra-host viral population dynamics of the experimentally infected stallions were compared and contrasted with the two naturally infected stallions. In this study, the analysis of viral sequences in nasal secretions and buffy coat cells showed a lack of extensive positive selection; however, characteristics of the mutant spectra were different in the two sample types. By contrast, semen virus

populations during acute infection had undergone a strong selective bottleneck as reflected by a reduction in population size and multiple sites of the viral genome that was under diversifying selection. During the period of persistent infection, virus solely evolved inside the reproductive tract of the stallions in the presence of strong neutralizing antibody response, resulting in a non-stochastic evolutionary process with a number of atypical high frequency minor variants being identified. This indicates that an active selection pressure continually morphs EAV quasispecies population structures and dynamics during persistent infection. Comparative analysis of the experimental evolution with that of naturally infected stallions showed a higher degree of correlation in the rate of sequence divergence through time with mirroring patterns of selection between experimental and naturally infected stallions.

2.2. Introduction

Equine arteritis virus (EAV) is the causative agent of equine viral arteritis (EVA), a reproductive and respiratory disease of horses, donkeys, mules and zebras[10,11,97,154,404–406,406]. EAV was first isolated from the lung of an aborted fetus during an extensive outbreak of respiratory disease and abortion storm on a Standardbred breeding farm near Bucyrus, OH, USA, in 1953[9–11,14,194]. Depending on the host's age, physical condition, immune status and the strains of virus, EAV infection causes mild to moderate clinical signs including leukopenia, pyrexia, depression, anorexia, dependent edema (scrotum, ventral trunk, and limbs), stiffness of gait, conjunctivitis, lacrimation and swelling around the eyes (periorbital and supraorbital edema and urticaria)[11,97,154,168]. EAV infection can also cause abortion in broodmares

and congenital infection can lead to interstitial pneumonia in young foals and death in newborn foals[171]. Most importantly 10-70% of acutely infected stallions become persistently infected carriers and can continue to shed virus in semen for a long time[11,157,158,208]. The carrier state is testosterone-dependent and can last from several weeks (short-term carrier state), 3 to 7 months (intermediate carrier state) and years or even lifelong (long-term persistent or chronic carrier state) despite the development of a strong serum neutralizing antibody response[1,11,97,98,143,154,156,157,407]. EAV primarily persists in the ampulla of the vas deferens in the reproductive tract of carrier stallions[165]. Persistently infected stallions play a central role in the maintenance, perpetuation and evolution of the virus in the horse population[142]. An increase in the incidence of the disease has been observed in the past twenty years, associated *inter alia* with increased national and international movement of horses and shipment of frozen or chilled semen [19,21,142,160,164,167,184,265–267,408].

EAV is a small, enveloped, positive-sense, single-stranded RNA virus that belongs to the family *Arteriviridae* (genus: *Equartevirus*, order: *Nidovirales*), which includes lactate dehydrogenase-elevating virus (LDV), simian hemorrhagic fever virus (SHFV), porcine reproductive and respiratory syndrome virus (PRRSV), wobbly possum disease virus (WPDV) and African pouched rat arterivirus (APRAV-1)[1,409]. The EAV genome contains approximately 12.7 kb with at least 10 known open reading frames (ORFs). The ORFs 1a and 1b are located in the 5′-three quarter of the genome and encode for two polyproteins (1a and 1ab), which are subsequently processed into 13 nonstructural proteins (nsp1-13; including nsp 7α and 7β) by viral proteases[1]. The ORFs 2a, 2b, 3, 4, 5a, 5b, 6 and 7 are located at the 3′-distal quarter of the genome [38,40] and

encode seven envelope proteins (E, GP2, GP3, GP4, ORF5a protein, GP5 and M) and the nucleocapsid (N) protein, respectively. The GP2, GP3 and GP4 are minor envelope glycoproteins and form a heterotrimer; whereas GP5 and M are the major envelope proteins of the virus and are assembled into a heterodimer in the virus particle. The major neutralization determinants of EAV are located in the N-terminal ectodomain of the GP5 major envelope glycoprotein of EAV. Studies using neutralization-resistant variants to monoclonal antibodies (MAbs) and polyclonal equine antisera in combination with reverse genetics have identified four major neutralization sites in the N-terminal ectodomain of the GP5 major envelope glycoprotein of EAV GP5 (site A [aa 49], site B [aa 61], site c [aa 67 to 90], and site D [aa 98-106])[1,82]. The changes in conformational neutralization epitopes is influenced by the GP5 interaction with the M protein[16,82]. Although there is only one known EAV serotype, field strains differ in their virulence and neutralization phenotypes [21,23,155,162,177,179,183,184,188,189,199]. Phylogenetic analysis based on partial ORF5 sequences segregated EAV isolates from around the world into North American (NA) and the European (EU) lineages, and each lineage is further subdivided into two North American (NA-1, NA-2) and European (EU-1 and EU-2) clades[21,408,410–412].

The lack of proofreading capability of the viral RNA-dependent RNA polymerase (RdRp) enzyme leads to a high mutation rate in the EAV genome resulting in the occurrence of viral quasispecies[282]. Often times, a novel virus variant arises from the quasispecies pool under conditions that regulate the frequency of minor variants. A number of studies have demonstrated intra-host quasispecies variation and the resultant genomic diversity in a number of important human and animal viruses[361,413–417].

Establishment of EAV-persistent infection in the reproductive tract of stallions has been known for decades but very little work has been performed to better understand the molecular basis of natural evolution of this virus during long-term persistent infection in infected stallions. Earlier findings from T_1 oligonucleotide fingerprinting analysis demonstrated the presence of a population level diversity in naturally infected stallion[418]. Furthermore, partial sequence analysis (ORFs 2-7) of sequential viruses from two persistently infected stallions following the 1984 outbreak in Kentucky have shown marked genetic divergence over time with emergence of new phenotypic variants[142]. However, these initial studies on EAV evolution during persistent infection were hampered due to the limitations in sequencing technology and data analysis capabilities [17,19,21,142,164,410]. In recent years, next-generation sequencing (NGS) technologies have improved the processing capacity and allowed for high throughput analysis of viral genomic data[361,419]. This allowed identification of low frequency variant populations in unprecedented detail and opened the door for more precise characterization of viral qausispecies population and its evolution during acute and persistent viral infections.

In this study, the comprehensive analysis of the dynamic events associated with EAV evolution was reported by following intranasal inoculation of seven experimentally inoculated stallions and two naturally infected stallions. The data showed successful transmission of EAV founder populations with distinct evolution during acute and persistent infections and also demonstrated the temporal regulation of variant populations following a single transmission event driven by genetic bottlenecks and episodic selections.

2.3. Materials and methods

Cells and viruses

The high passage (HP) rabbit kidney cell line RK-13 (KY), passage level 399-409, was maintained in Eagle's minimum essential medium (EMEM; Mediatech, Manassas, VA, USA) supplemented with 10% ferritin-supplemented bovine calf serum (Hyclone Laboratories, Logan, UT, USA), 100 U/mL penicillin/streptomycin (Mediatech, Manassas, VA, USA) and 1 μg/mL amphotericin B (Sigma-Aldrich, St. Louis, MO, USA). Equine pulmonary artery endothelial cells (EECs) were maintained in Dulbecco's modified essential medium (Mediatech, Herndon, VA, USA) with sodium pyruvate, 10% fetal bovine serum (Hyclone Laboratories, Inc., Logan, UT, USA), 100 U/μg per mL of penicillin/streptomycin and 200mM L-glutamine[68].

The KY84 strain of EAV was originally isolated from pooled blood collected during the 1984 EVA outbreak (5/1984) from three stallions acutely infected with EAV (stallion E and two other stallions) and was serially passaged three times in horses and subsequently once in EECs to make the working virus stock[419]. This moderately virulent EAV KY84 strain was used as the challenge virus in this study[22,420]. The KY84 strain of EAV has been shown to establish persistent infection in the reproductive tract of stallions and to cause moderate to severe clinical signs of EVA in infected horses[156,208,420]. Tissue culture fluid (TCF) containing the commercial live attenuated vaccine strain of EAV (ARVAC®; Zoetis, Kalamazoo, MI, USA) was used as reference virus to determine the serum neutralizing antibody responses during acute and persistent

infection of the stallions in accordance with the World Organisation for Animal Health (OIE) standardized virus neutralization test (VNT) protocol[421].

Ethics statement

This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The Institutional Animal Care and Use Committee (IACUC) at the University of Kentucky, Lexington, KY, USA approved this protocol (number 2011-0888). Stallions were humanely euthanized by pentobarbital overdose following the American Veterinary Medical Association (AVMA) guidelines for the euthanasia of animals, and all efforts were made to minimize suffering.

Stallions used for experimental infection

Seven sexually mature (4 to 16 years of age) stallions of mixed breed were included in the study (stallion IDs: L136 to L142). Horses were obtained from a local commercial vendor and acclimated to their new environment for approximately two months before the study commenced. During this period, animals were accommodated in individual paddocks and trained to mount a mare or a phantom for collection of semen into an artificial vagina. All stallions exhibited good libido and normal sexual behavior and were seronegative (titer <1:4) for EAV neutralizing antibodies several times prior to intranasal inoculation with the EAV KY84 strain using a previously described protocol[405]. The animals were housed in individual stalls in an isolation

facility for the duration of the study at the University of Kentucky Maine Chance Farm, Lexington, KY, USA.

Experimental infection of stallions with EAV KY84 strain and clinical examination

Stallions were inoculated intranasally with 3.75×10^5 plaque-forming units (PFU) of the EAV KY84 strain in 5.0 mL of EMEM (inoculum) using a fenestrated catheter passed via the posterior nares into the nasopharynx as previously described[199]. All stallions were examined and clinical parameters were recorded by the same veterinarian. Pre-inoculation (7, 5 and 2 days before experimental challenge) clinical examinations were performed once daily to determine baseline values for body temperature and also to certify that the parameters were within normal limits before experimental challenge of the stallions. Specifically, fever and scrotal edema were monitored twice daily (every 12 hours) for the first 15 days after infection, and the highest value of the day was recorded. Clinical signs continued to be monitored once per week during the following four weeks of the experiment (at 21, 28, 35 and 42 dpi). Blood samples were collected at 0, 2, 4, 6, 8, 10, 12, 14, 21, 28, 35 and 42 dpi to determine individual serum neutralizing antibody responses to EAV. The neutralizing antibody titers were determined as per the OIE-prescribed test[263].

Semen collection from experimentally infected stallions

Two ovariectomized mares previously vaccinated with the commercial MLV vaccine against EVA (ARVAC®, Zoetis Animal Health, Inc., Kalamazoo, MI, USA) (serum neutralizing antibody titers ≥1:256) were used to "tease" the stallions. Each

stallion was allowed to mount either one of the mares or a breeding phantom to enable semen collection in an artificial vagina (Botucatu model, Botupharma, Botucatu, SP, Brazil). The artificial vagina was disinfected with a disinfectant (Roccal®-D Plus, Pfizer Inc., New York, NY, USA) between collections. A disposable liner lubricated with sterile lubricant (Priority Care, First Priority Inc., Elgin, IL, USA) was used to collect semen from each stallion to avoid cross-contamination between ejaculate samples. A pre-semen sample was collected and evaluated 2 days before challenge, and then the samples were collected on days 1, 3, 5, 7, 9, 11, 13, 15 and 23 during acute period of infection. Once the stallions became persistently infected, semen samples were collected from each stallion once per month up until 726 days and processed to evaluate the samples. Specifically, the numbers of sequential EAV isolates from each stallion were selected at 5, 9, 107, 170, 345, 380, 548 and 726 dpi (Figure 2.1.).

EVA outbreak in 1984 and semen collection from two naturally infected stallions (D and E)

An extensive outbreak of EVA occurred in Central Kentucky in 1984[158]. This outbreak was the first recorded occurrence of EVA in the North American Thoroughbred population, and subsequent investigation established that stallions could be persistently infected with EAV; furthermore, carrier stallions were critical to the epidemiology of EAV infection[146,156,157]. Two stallions from the index premises (designated D and E) were infected during this outbreak and subsequently became long-term EAV shedders. Semen was collected from the two stallions at regular intervals (stallion D, 6/84 [month/year], 9/85, 12/86, 9/87, 7/88, 1/89, 1/91, 9/92 and 8/94 [seven

sequential samples] and stallion E, 6/84, 9/85, 11/86, 2/88, 1/89 and 1/91 [six sequential samples]) following initial infection of the stallions, which occurred in May 1984. Semen was collected until the stallions ceased shedding EAV, as determined by virus isolation as previously described[157].

Virus isolation

Virus isolation from raw gel-free semen samples was attempted in the RK-13 (KY) cell line according to a standard protocol used by the OIE Reference Laboratory at Maxwell H. Gluck Equine Research Center, University of Kentucky, USA[222]. Briefly, after sonication, serial decimal dilutions (10^{-1} to 10^{-4}) of each sample supernatant were prepared in EMEM (Cellgro®, Mediatech, Inc, Manassas, VA, USA), and 1 mL of each dilution was inoculated into each of two 25-cm^2 flasks containing a confluent monolayer of RK-13 (KY) cells grown in supplemented EMEM and overlaid with the supplemented EMEM containing 0.75% carboxymethylcellulose (CMC; Sigma-Aldrich, St. Louis, MO, USA). Flasks not showing cytopathic effect (CPE) after 4 days were subinoculated onto new RK-13 (KY) monolayers using 1 mL of culture supernatant as the inoculum. Tissue culture fluid (TCF) was harvested and stored at -80°C for viral RNA extraction. First and second passage flasks were stained with a 0.2% crystal violet solution in buffered formalin on day 4 post-inoculation to count plaques and calculate virus titers. The same individual performed all attempted virus isolations. Virus isolates were, confirmed later by an EAV-specific TaqMan® real-time RT-PCR assay as previously described[223].

Phenotypic characterization of EAV isolates against neutralizing monoclonal antibodies and equine sera

The EAV microneutralization assay was performed as previously described[15,16]. Neutralization titers were determined for each virus isolated from the semen of individual carrier stallions using a panel of EAV-specific MAbs and polyclonal equine antisera[422].

Real-time RT-PCR

Briefly, viral RNA was directly isolated from 50 µl of tissue culture fluid (TCF) using a commercial RNA isolation kit (MagMAX™ -96 Viral RNA Isolation Kit, Applied, Foster City, CA, USA) according to the manufacturer's instructions. Viral RNA was eluted in 50 µl of nuclease-free water and stored at -80°C. RNA extracted from TCF derived from RK-13 (KY) cells inoculated with semen samples that were negative for EAV, as well as from nuclease free water were included as negative controls. Viral RNA extracted from TCF containing the KY84 strain of EAV was used as a positive control. A one-tube TaqMan® real-time RT-PCR assay was performed as described by Miszczak *et al.*[223]

Next-generation and Sanger sequencing

The total viral RNA was extracted from the TCF containing EAV KY84 strain (inoculum) or from the cell lysate following freeze-thaw of HP RK-13 (KY) cell monolayer (P1) in 25 cm^2 flasks using a total RNA purification kit (Catalog no. 17200,

Norgen Biotek Corp., Ontario, Canada). The RNA was quantified using a Qubit 2.0 spectrophotometer (Life Technologies, Carlsbad, CA, USA). The cDNA libraries were constructed from 100 ng of total RNA using a TruSeq Stranded total RNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Multiplex libraries were prepared using barcoded primers and a median insert size of 340 base pairs (bp). Libraries were analyzed for size distribution using a Bioanalyzer and quantified by quantitative RT-PCR using a Kapa library quantification kit (Kapa Biosystems, Boston, MA, USA), and relative volumes were pooled accordingly. The pooled libraries were sequenced on an Illumina MiSeq platform with 150 bp end reads following standard Illumina protocols. The Sanger sequencing was performed by RT-PCR amplification of the entire EAV genome in multiple overlapping segments using high-fidelity DNA polymerase enzyme as previously described. Primer sequences used for RT-PCR amplification and sequencing will be available upon request from the corresponding authors[17,21,150]. The nucleotide sequences of the 5' untranslated region (UTR) and 3' UTR termini region were further confirmed by standard Sanger sequencing of the reverse transcription-PCR products. Overlapping sequences were analyzed and assembled into a full-length genome sequence using the Geneious 7.0.6 software (Biomatters Ltd., Auckland, New Zealand).

Read mapping and variant calling

Raw reads in FASTq formats were imported into a CLC Genomics Workbench Version 8.5.1 (CLCBio, Aarhus, Denmark) after quality associated trimming; reads were

paired using the Illumina short read default parameters. Paired reads were then mapped to the reference genome (EAV KY84 in the inoculum) and the frequency of viral quasispecies population in individual samples, were called from reads coverage. The mapping procedures were normalized by subsampling of the reads across data sets using samtools 1.1 to 1000 x coverages to avoid the biases from a high and low read coverage across samples sets. Mapped files in (BAM) were imported back to CLC genomic workbench and numbers of single nucleotide variants (SNVs) were estimated by applying a minimum frequency filter of 1% and ignoring broken reads and nonspecific matches. Reference masking was set up for reads coverage >100,000 and < 100 in individual samples. A technology-associated filter was used to remove any pyro error variant with a minimum length of 3 and a frequency below 0.8.

Nucleotide diversity and evolutionary selection analysis

The statistical measure of nucleotide diversity **(∏)**, which quantifies the average number of pairwise differences within a sample, was calculated using open-source software SNPGenie (https://github.com/chasewnelson/snpgenie). SNPGenie is a Perl script which analyzes SNV caller results to calculate an evolutionary parameter, such as nucleotide diversity (nonsynonymous and synonymous diversity, ∏n and ∏s). Reference sequences in FASTA format, SNV report from CLC Genomics Workbench were used as input files. To produce GTF formats, the script snpgenie-gb2gtf.pl. was used and an estimation of nucleotide diversity ∏ was performed by running snpgenie.pl. from the command line. Values generated for the whole genome were used to show the diversity estimate at different time points. Individual ORFs were used to calculate ∏n

(nonsynonymous diversity) and ∏s (synonymous diversity). A ratio of ∏n/∏s > 1 was considered indicative of positive selection, <1 means purifying selection and 1 indicates neutrality.

For selected ORFs with evidence of positive selection, the codon level was estimated to investigate particular areas under selection pressure. In this analysis for each codon, estimates of the numbers of inferred synonymous (s) and nonsynonymous (n) substitutions are presented along with the numbers of sites that are estimated to be synonymous (S) and nonsynonymous (N). These estimates are produced using the joint maximum likelihood reconstructions of ancestral states under a Muse-Gaut model of codon substitution and model of nucleotide substitution. For estimating maximum likelihood values, a tree topology was automatically computed. The test statistic dN - dS is used for detecting codons that have undergone positive selection, where dS is the number of synonymous substitutions per site (s/S) and dN is the number of nonsynonymous substitutions per site (n/N). A positive value for the test statistic indicates an overabundance of nonsynonymous substitutions. Normalized dN - dS for the test statistic is obtained using the total number of substitutions in the tree (measured in expected substitutions per site) and used for making comparisons across data sets. Maximum likelihood computations of dN and dS were conducted using hypothesis testing and the phylogenies (HyPhy) software package.

In addition, the presence of selection pressure at the codon level is further assessed using an online server at http://www.datamonkey.org. Fast unconstrained Bayesian analysis for inferring selection (FUBAR) algorithms was used to characterize selective pressure that are aggregated in all the branches of the constructed phylogenies,

and the mixed effect model of evolution (MEME) was used to identify individual sites subject to episodic positive selection along a subset of phylogenetic tree branches.

Sequence analysis and SNV detection from consensus sequences

Reads from individual samples were assembled *de novo* as indicated above using ABySS (vb1.3.7). The highest-ranking base at a position and with highest quality scores (highest calculated Q-score) were used to determine the consensus sequence. Once the consensus sequence was determined for each sample, the annotation of individual ORFs, and other genomic features was conducted using an automated gene transfer algorithm, Glimmer 3. The EAV reference genome in the NCBI database (NC_002532) was used to benchmark this procedure. The annotated ORFs and nonstructural proteins encoding regions were manually inspected and curated for correct enzyme cleavage sites, translation frame, start and stop codon coordinates.

Finally, the full genome nucleotide sequences (derived from the consensus) and parts representing a genomic region were aligned using Clustal W plugin in Geneious V6.1.7. For the nucleotide analysis, a default setting was in Clustal W cost matrix, a gap opening cost 15 and gap extension cost of 6.6. For translated amino acid sequence alignment, the GONNET cost matrix was used with a gap opening cost of 10 and a gap extension cost of 0.1. The original EAV KY84 strain used for the inoculum was the reference sequence for multiple sequence alignments. Single nucleotide variants (SNVs) in positions were identified using Geneious V6.1.7 and the effect of each SNP on translation (synonymous and nonsynonymous) was calculated for individual ORFs.

Phylogenetic analysis

A phylogenetic tree was constructed using the same parameters described by Guindon *et al*. (2010)[423]. Evolutionary and genetic divergence analyses were performed by estimating the maximum likelihood of phylogenies using the PhyML program as a plugin for Geneious V6.1.7. These trees were constructed based on the GTR model, with 1000 bootstrap replicates. In addition, trees for molecular evolution were assessed using the maximum likelihood method based on the Tamura-Nei model as implemented in MEGA 7.0.21. Maximum likelihood trees were also generated using TempEst-v1.5 (formerly known as Path-O-Gen) to investigate the temporal signals (clock-likeness) of the sequential genomes generated in this study. These trees were used to estimate the root tip distance by regression analysis. The relationship between root to tip divergence and sample collection date supports the use of molecular clock trees in this study. Molecular clock trees were generated by Bayesian inference analysis, as implemented in MrBayes version 3.2.2[424]. These trees were estimated using the GTR substitution model and gamma distribution rate of variation. The posterior probabilities were calculated using MCMC chain length generating 1,000,000 trees and sampling every 100th tree. The initial 25% of trees were discarded as a burn-in, and the others were used to construct a majority rule tree.

2.4. Results

Establishment of EAV persistent infection in experimentally infected stallions

A major aim of this study was to investigate intra-host evolutionary dynamics of EAV during acute and persistent infection in stallions. To achieve this goal, an experimental model of persistent infection was established by using the moderately virulent EAV KY84 strain. The selection of this strain was based on two important phenotypic traits relevant to this study, one is a moderately virulent phenotype which can be used to assess clinical and behavioral parameters following infection and the second is a demonstrated ability to establish long-term persistent infection in the stallion reproductive tract.

Following infection, all seven stallions developed moderate to severe clinical signs. The detailed description of the clinical signs, virus effect on semen quality, as well as its tissue and cellular tropism have been previously described[211]. Briefly, clinical signs started at 2 dpi, with signs that included fever (38.7°C to 40.8°C [101.7°F to 105.6°F]), leukopenia, ocular, scrotal, preputial and limb edema, nasal and ocular discharge, congestion and petechial or ecchymotic hemorrhages on the oral mucosal membrane and decreased libido. These clinical signs lasted for up to 20 days, and the peak manifestations occurred at a median of 7 dpi. Once stallions recovered from these signs there were no recurrences of signs up to the conclusion of the experimental study (726 dpi). Viral shedding in nasal secretions was detected as early as 2 dpi and lasted up to 19 dpi. Viral titers varied from $\leq 10 - 5.7 \times 10^4$ PFU/mL (mean viral titer of 4.36×10^3 PFU/mL). Similarly, viremia was detected in buffy coat cells at 2 dpi but was extended to between 28 and 42 dpi, and a mean viral infectivity titer varied from $\leq 10 - 5.1 \times 10^3$

PFU/mL (mean viral titers 2.35 x 10^2 PFU/mL). The peak for both nasal shedding and viremia occurred at a median of 6 dpi. The clearance of virus from blood mononuclear cells coincided with the appearance of neutralizing antibodies.

All experimentally infected stallions exhibited decreased libido during the acute phase of infection. However, neither clinical signs of EVA nor reproductive dysfunction were observed in any of the experimentally infected stallions (n=7) following recovery from acute infection. Viral shedding in semen started at 5 dpi, except for one stallion (L141) where it was detected at 3 dpi. Viral infectivity titers in semen varied from 1.0 x 10^1 to 1.88 x 10^7 PFU/mL during the acute infection (up to 21 dpi), with a mean titer of 7.85 x 10^5 PFU/mL. A peak in viral infectivity titer in semen occurred at a median of 9 dpi. Stallions seroconverted from 8 dpi and maintained high serum neutralizing antibody titers until the end of the study (1:64 to >1:512 [median titer of 1:256]).

Two out of seven experimentally infected stallions (L136 and L140) continued to shed EAV in their semen for ≥ 726 dpi (i.e. long-term persistently infected stallions). By the end of the study, viral titers in semen of these two stallions were 6.5 x 10^3 and 7.8 x 10^3 PFU/mL, respectively[165], but infectivity titers in these stallions fluctuated from 1.2 x 10^3 to 8.8 x 10^5 PFU/mL during persistent infection in the reproductive tract (Figure 2.1.).

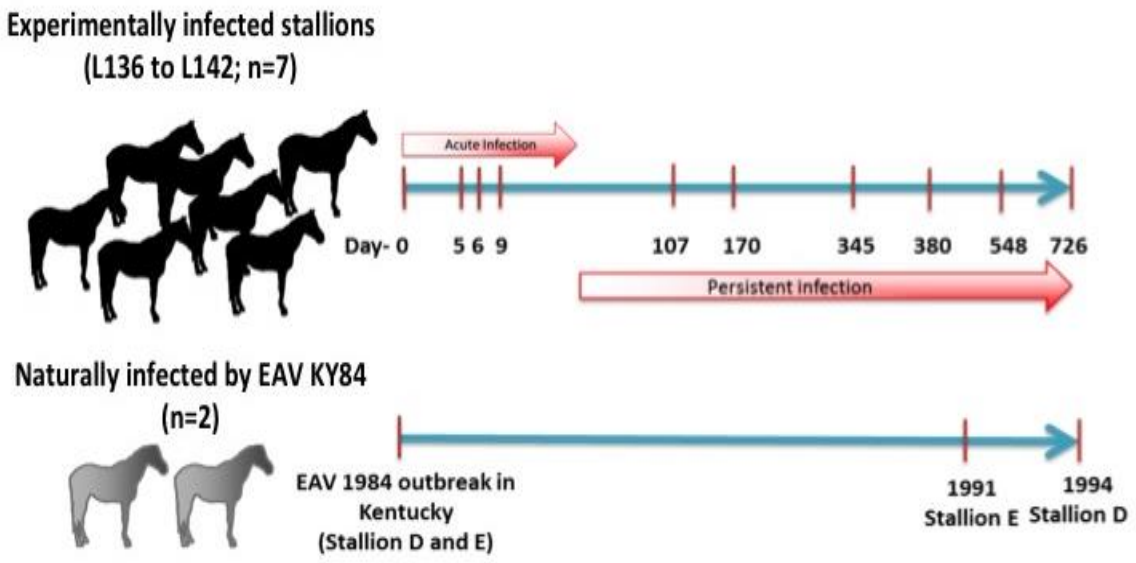


Figure 2.1. Schematic presentation of the experimental design for this study. Darker horses were used to represent experimentally infected stallions (n=7), arrow showing timeline of sampling. Grey shaded horses were used to represent naturally infected stallions and arrow showing timeline of semen collection for isolation of the sequential viruses.

Persistent EAV infection in naturally infected stallions

Stallions D and E were naturally infected and became persistently infected during the extensive outbreak in 1984[142]. These stallions have been monitored up to ten years, and during this period, EAV was shed in semen with titers ranging from 10 x 10^3 to 1.0 x 10^5 PFU/mL (Figure 2.1.). There were neither apparent clinical signs nor any evidence of reproductive dysfunction in these stallions during the follow up to the conclusion of the study[142].

Complete genome sequence of EAV KY84 strain used for experimental infection

The complete genome sequence of EAV KY84 strain used as inoculum was determined using two methods. Consensuses derived independently from Sanger and NGS were compared and showed a 100% pairwise nucleotide identity. A fully annotated and complete sequence of EAV KY84 strain is deposited in GenBank with the accession number (MG137429).

Virus population in the inoculum

In order to understand the transmission dynamics of EAV quasispecies population following experimental infection, the quasispecies population was first characterized in the inoculum (original EAV KY84) by investigating the SNV present at minimum allelic frequency of 1%. The analysis showed that a genetically more homogenous viral population with structure constituted by a major variant (consensus virus; with ≥77% in frequency) and minority SNV population that was circulating from 13 to 25 %. However, this is likely to represent only a subset of the true variation present within the viral inoculum and characterizes the variants present at a frequency of ≥1% (Figure 2.2.).

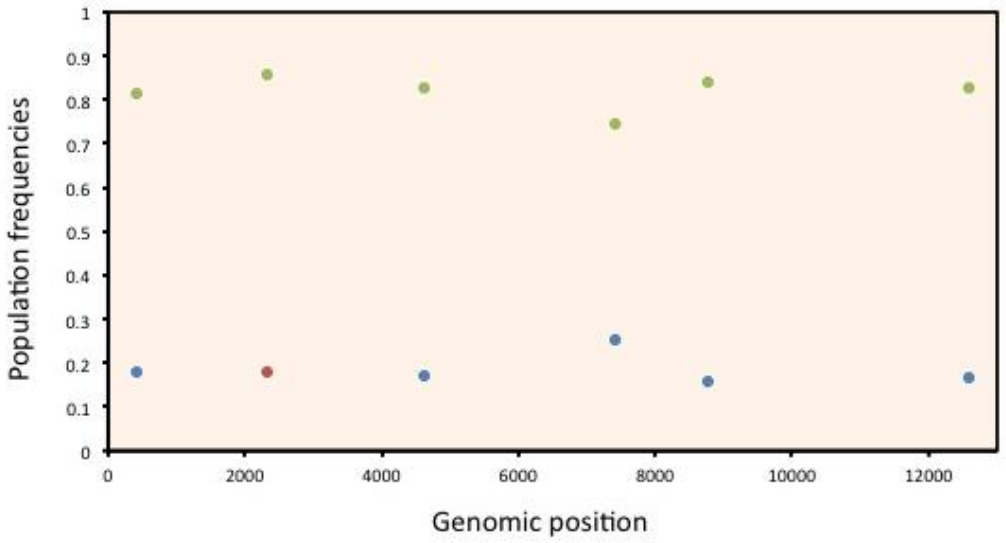


Figure 2.2. SNV population in the inoculum used for experimental infection. Each dot in the figure represents a major and minor SNV population (nucleotide numbers A407G, A2337G, T4616C, T7414C, A8759C and C12582T). Population frequencies on y-axis and genomic position on x-axis are plotted in 2000 bp intervals. Blue dot (a synonymous SNV), red dot (nonsynonymous SNV) and light green (major variant).

The minor variant populations in the inoculum were dominated by synonymous substitutions that are distributed evenly along the EAV genome. One exception to this was a SNV population with nonsynonymous substitution from A->G (T->A) at genomic position 2,337 corresponding to nsp2. The allelic frequency for this specific substitution in the virus inoculum was 13.5 %; whether this substitution is associated with specific adaptive function in cell culture is hard to predict in the absence of experimental data. However, a higher degree of conservation in the dominant nucleotide "A" across cell culture-passaged EAV field isolates (data not shown) suggest that it is unlikely that this substitution is associated with any specific advantage in the cell culture.

Founder virus analysis and nucleotide diversity during acute infection period

Viruses isolated in the first nine days of the experimental infection were uniquely positioned to provide information on the mechanism of adaptive evolutionary process of EAV in a variety of host cells following initial infection. During this period, the virus present in the inoculum had successfully established infection in the respiratory epithelium and blood cells of all seven experimentally inoculated stallions. Virus shedding in the nasal secretions and viremia in blood (virus in buffy coat cells) were detected as early as 2 dpi, and then appeared for the first time in the semen at 5 dpi, suggesting an initial infection of the stallion reproductive tract where EAV established persistent infection.

This adaptation to infect and replicate in various tissues along with an active challenge from host innate and fledgling adaptive immune responses could constitute evolutionary bottlenecks that select specific viral variants to shape an effective population size. In order to test this hypothesis and also elucidate the early event of quasispecies dynamics, representative viruses from time points of highest nasal shedding and viremia (6 dpi) and the initial semen isolates (5 and 9 dpi) were deep sequenced from viral RNA derived in the first TCF passage 1 (P1) in RK-13 (KY) cells. Thus in this section, results from deep sequencing of 28 sequential viruses are systematically presented to highlight the early events in EAV KY84 infection.

First, the variant population present in the inoculum was compared with the 14 sequential viruses from nasal swab and buffy coat samples. For this, the SNV populations (A407G, A2337G, T4616C, T7414C, A8759C and C12582T) were identified in the inoculum and were traced in the nasal swab and buffy coat viruses. Each of these six

minor variant populations was detected at least in five of the seven experimental stallions without causing a major shift in the population frequency (Figure 2.3.).

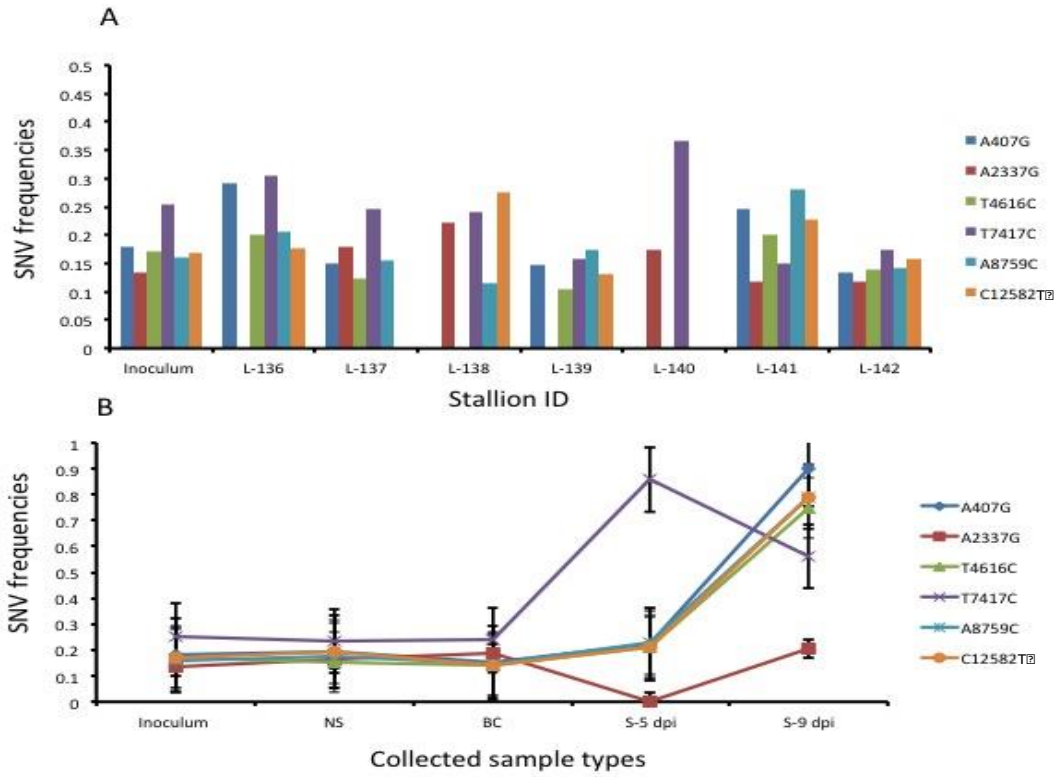


Figure 2.3. Mapping of SNV in nasal secretions (NS), buffy coat (BC) cells and semen (S) in seven experimentally infected stallions during acute infection (0 to 9 dpi). (A) Frequencies of the six SNV (represents variants in the inoculum) in an individual stallion. (B) Average detection frequencies of the six mapped SNV were steady in nasal secretion and buffy coat cells at 6 dpi. However, the SNV populations in the S (semen) at 5 and 9 dpi had undergone a shift in frequencies, which resulted in the fixation of T7417C at 5 dpi and A407G, T4616C, A8759C and C12582T at 9 dpi. The nonsynonymous substitutions from A2337G were below the detection limit at 5 dpi and then attained its average detection frequency at 9 dpi.

These results have confirmed transmission of all virus populations present in the inoculum and successful establishment of infection. Then, the overall similarities were identified between the major variants in the inoculum and initial viruses from the experimentally inoculated stallions by constructing phylogenetic trees and assessing consensus level. The results as demonstrated by the star-like phylogenetic tree structure and a 100% consensus agreement between the virus inoculum and the seven viruses from nasal secretions indicating that a single major founder virus population was responsible for establishment of the experimental infection (Figure 2.4.). However, there were marked differences in quasispecies population composition and structure between the virus inoculum and viruses derived from nasal swabs. In this regard, nasal swab viruses have shown an expansion of the minor variant population by accumulating mutations that were not present in the inoculum. As a result, the quasispecies population frequencies were increased by more than 20-fold in the nasal swab viruses (Figure 2.5.A and B.).

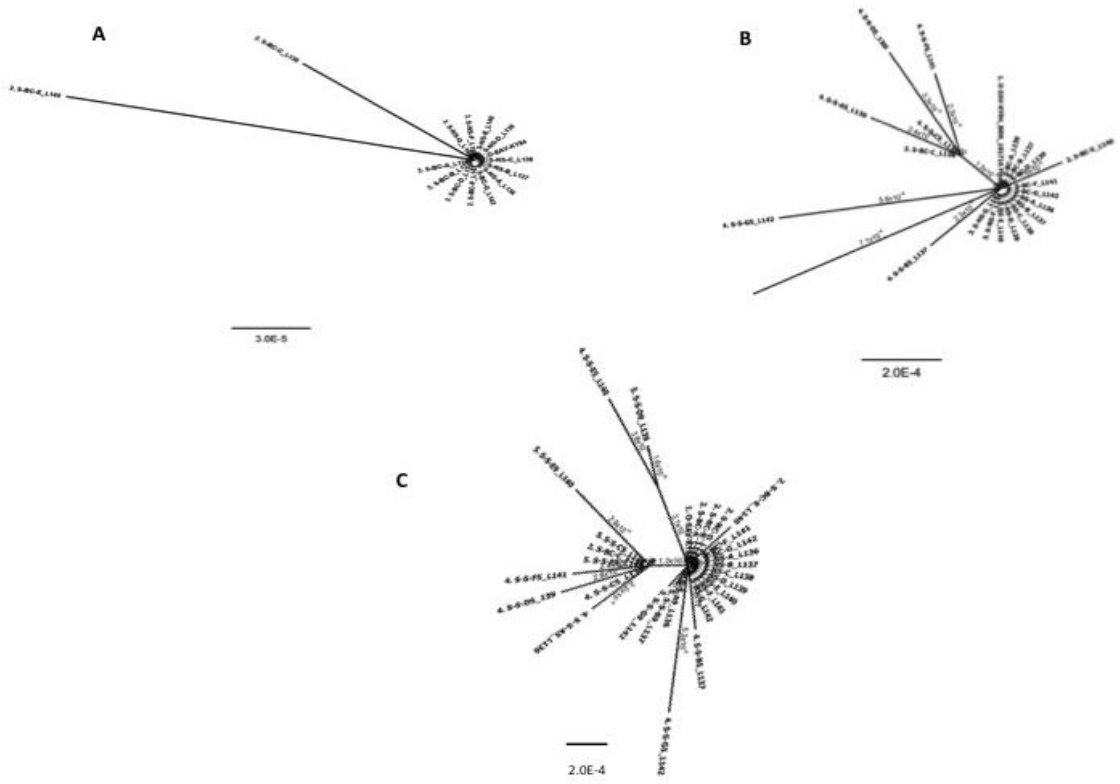


Figure 2.4. Maximum likelihood phylogenetic trees showing a star-like network, which demonstrates the presence a major founder population in the experimental stallions similar to the stock virus. (A) Viruses from nasal secretions and buffy coat at 6 dpi; (B) tree showing divergence of initial semen viruses from the major founder population at 5 dpi; and (C) at 9 dpi, the virus from semen was shown to evolve further in the stallion reproductive tract. The rates of evolution (nucleotide substitutions / site) in each tree are indicated under the bar.

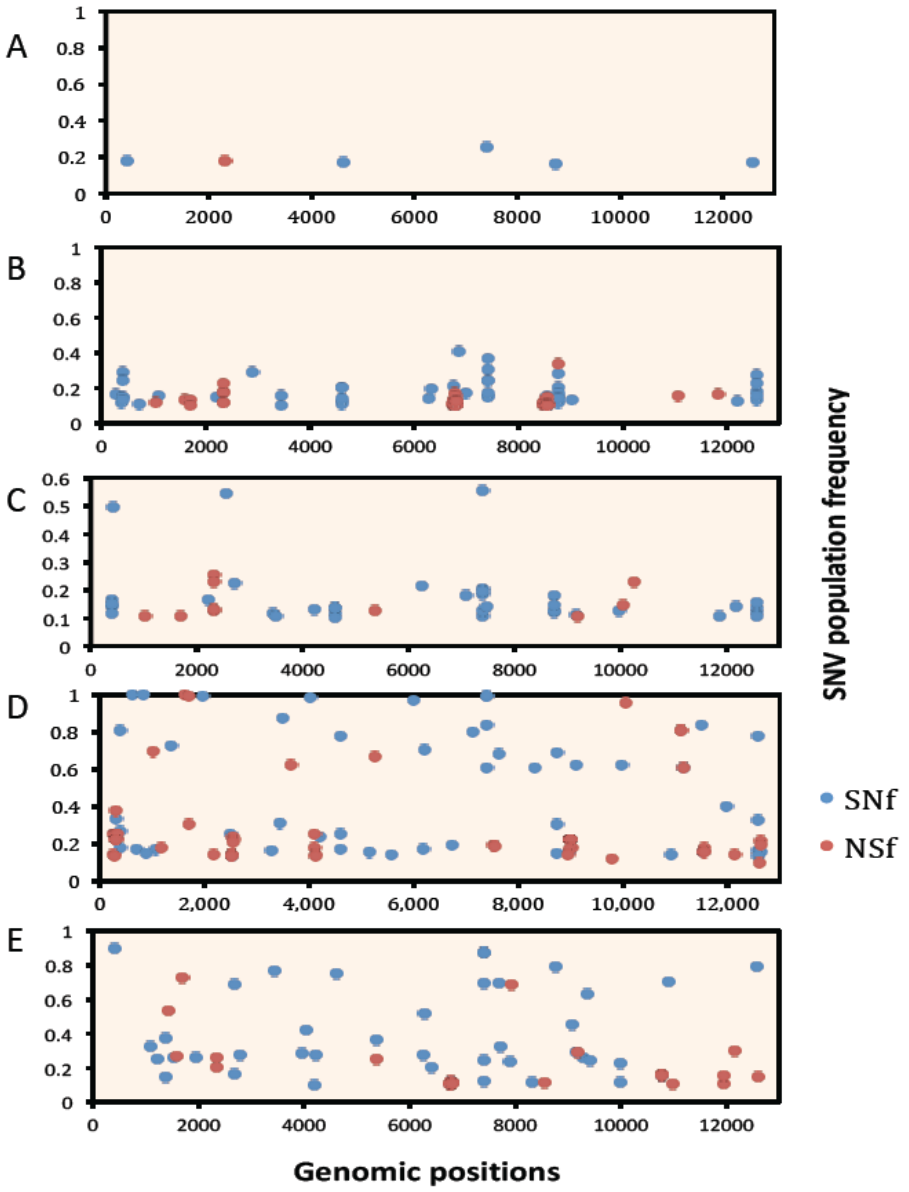


Figure 2.5. SNV population and structure during the acute phase of infection (0 to 9 dpi). Each dot in the figures represents a major and minor variant population. Population frequencies on y-axis and genomic position on x-axis are plotted in 2000 bp interval. Blue dot (synonymous SNV) and red dot (nonsynonymous SNV). (A) Less heterogeneous population in viral inoculum used for experimental infection of stallions, (B) nasal swab viruses at 6 dpi, (C) buffy coat viruses at 6 dpi, (D) semen viruses at 5 dpi and (E) semen viruses at 9 dpi.

To evaluate the level of polymorphism in this expanded quasispecies population, the nucleotide diversity was estimated within the nasal swab viruses by computing nucleotide diversity (∏) statistics. In this method, SNV are weighted at the population level and high frequency polymorphisms will be considered to be more diverse than populations with the same number of low frequency variants. The method also allowed us to calculate within the sample nonsynonymous nucleotide diversity ∏n and synonymous diversity ∏s separately. The ratio of the two estimates was performed to evaluate the presence of a selective pressure. A value of 1 suggested neutrality or absence of selection pressure, whereas a value of >1 suggested the presence of positive selection and a value < 1 suggested a purifying type of selection. Based on these analyses, a higher nucleotide diversity was detected in nasal swab (1.9×10^{-4}) viruses than in the inoculum virus (1.3×10^{-5}). The ∏n/∏s ratio for nasal swab viruses was 3.1×10^{-1}, suggesting a dominant purifying type of selection across the entire viral genome. However, during the SNV enumeration, the most nonsynonymous mutations were aggregated in nsp9 and nsp10 (RNA-dependent RNA polymerase [RdRp] and zinc-binding domain [ZBD]) encoding regions in the viruses present in the nasal secretions.

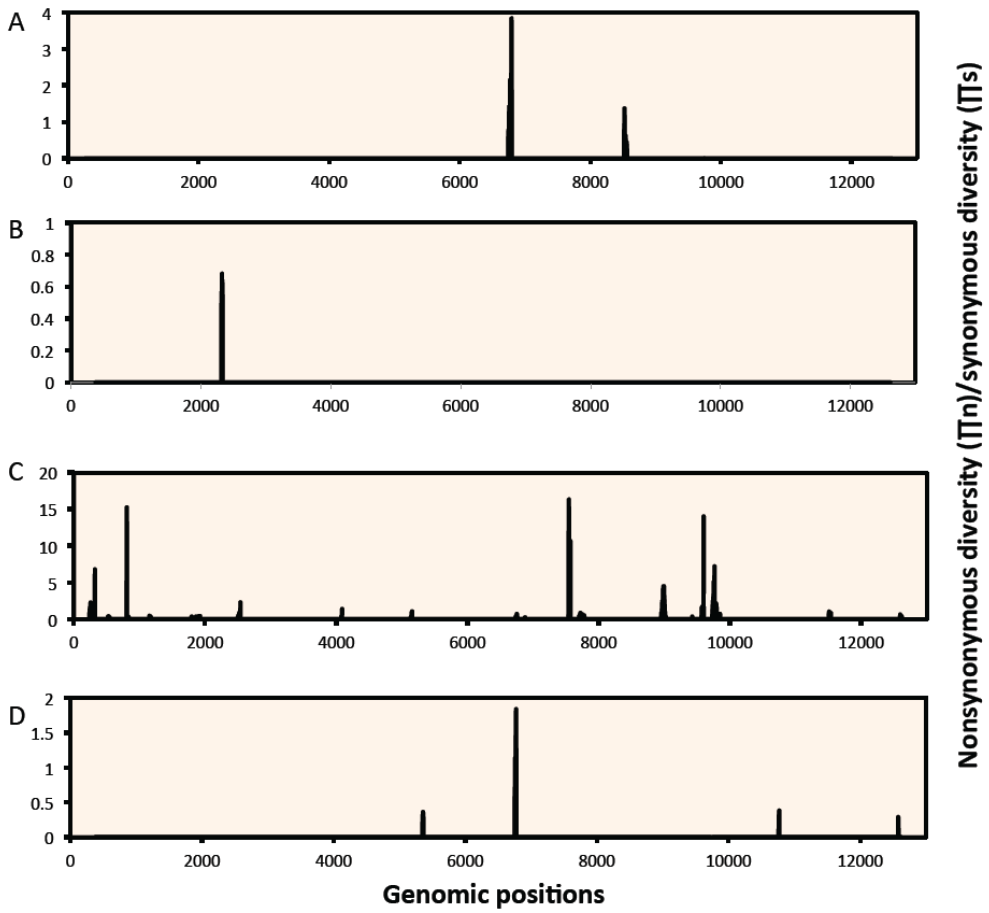


Figure 2.6. Sliding window analyses showing evidence for positive evolutionary selections in the sequential viruses during acute infection. Selection was estimated by calculating the ratio of nonsynonymous and synonymous nucleotide diversity (∏n/∏s). A value of >1 was indicative of positive selection, value of 1 = neutrality and value < 1 was indicative of purifying selection; results were plotted against genomic position (x-axis) in 2000 bp intervals. (A) Viruses from nasal swab, a positive selection was identified in genomic position corresponding to nsp9 and nsp10; (B) viruses from buffy coat, no region was found under positive selection pressure; (C) and (D) semen viruses at 5 and 9 dpi respectively, semen virus at 5 dpi showed strong positive selection at multiple locations, whereas at 9 dpi, the selection pressure decreased except in the nsp9 region.

To investigate whether this is due to positive selection, the sliding window analysis (∏n/∏s) was performed via spanning the entire genome of nasal swab viruses. The results showed that the presence of strong positive selection (∏n/∏s >1) at the genomic regions spanning nucleotides (nts) 6729 to 6816 (within nsp9 peptide) and nts 8469 to 8565 (within nsp10 peptide) (see previous page) (Figure 2.6.A.). Similar analysis was conducted with viruses present in the BC samples at 6 dpi. Interestingly in the case of these samples, a decline in quasispecies diversity was observed as measured by the decreased nucleotide diversity (1.6×10^{-4}) and total population size (Figure 2.7.A.). In addition, a minor shift in the population structure was identified in two stallions (L140 and L138) that resulted in fixation of synonymous nucleotide substitutions at genomic positions 422 [G->A], 2567 [C->T] and 7417 [T->C] (Figure 2.5.C.). However, contrary to what was observed in the case of nasal swab viruses, the sliding window analysis of ∏n/∏s in the buffy coat viruses has shown no supporting evidence for localized positive selection (Figure 2.6.B.). In summary, despite a high level of consensus agreement and same date of isolation (6 dpi), the analyses have identified differences in viral quasispecies diversity, population size and evolutionary selection pattern between viruses from nasal swab and buffy coat samples.

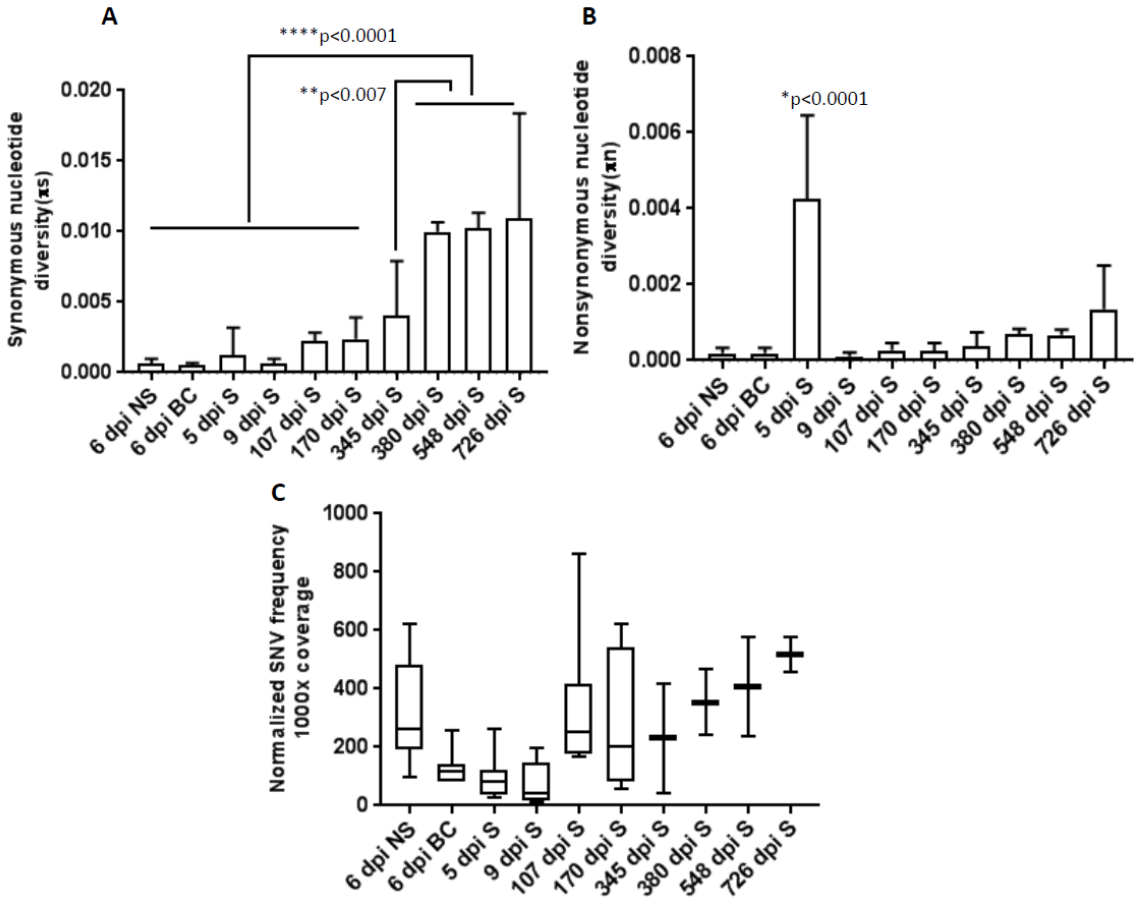


Figure 2.7. Quasispecies diversity and population size of sequential viruses in nasal secretions, buffy coat and semen during acute and persistent infections. (A) Comparison of synonymous nucleotide diversity in the sequential viruses. The diversity in the two long-term carrier stallions (L-136 and L-140) at later dpi (380 and 726) were significantly higher ($p < 0.0001$; one-way ANOVA) than diversity detected in all other samples in this study. (B) Nonsynonymous diversity at 5 dpi in the semen viruses had a significantly higher substitution rate ($p < 0.0001$; one-way ANOVA) than any other sequential virus isolate in this study. (C) The quasispecies population size in sequential viruses has shown a trend of decline in the initial semen samples (at 5 and 9 dpi), where at 9 dpi it attains its lowest point in the experimental period. NS (nasal swab), BC (buffy coat) and S (semen).

Quasispecies diversity and evolutionary selection in the initial semen viruses

Virus first appeared in semen at 5 dpi; the analysis of the viral quasispecies population in these samples revealed marked differences in overall evolutionary dynamics compared to viruses in the initial inoculum, nasal secretions and buffy coat samples. One major observation is semen viruses at 5 dpi have a high rate of nonsynonymous nucleotide diversity occurring at a rate ($\prod n = 4.2 \times 10^{-3}$) (Figure 2.7.B.). This was the highest nonsynonymous substitution rate recorded for the entire experimental period ($p < 0.0001$). Interestingly, the viral quasispecies population size at 5 dpi demonstrated a slight decline compared to what it was in buffy coat viruses and a 3-fold reduction when compared to viruses in the nasal swab (Figure 2.7.C.). This showed that virus populations were bottlenecked during the initial infection of the stallion reproductive tract. In addition, the presence of a high level of nonsynonymous nucleotide diversity within the population ($\prod n=4.2 \times 10^{-3}$) and the shifts of the minor and majority variant population characterizes the type of bottleneck as a selective one (Figure 2.5.C.). The sliding window analyses for $\prod n/ \prod s$ also support the finding with additional evidence of strong selective sweeps in multiple genomic positions of the semen viruses (Figure 2. 6.C.). More interestingly, there were 29 new SNV fixed in the consensus genome that contributed 8 nonsynonymous substitutions in the structural protein encoding genes and in ORFs 1a and 1b. Given the nonstructural region is three-fourths of the entire EAV genome, the rate of nonsynonymous substitutions was higher in the structural protein encoding ORFs, which occupy one-fourth of the EAV genome. The regions of nsp2 peptide and ORF5 each acquired three amino acid substitutions. The changes in the nsp2 peptide region were at nucleotide positions 1,018 [P->L], 1633 [K-

>R] and 1713 [A->T], whereas they were at genomic position 11,133[R->H], 11,162 [V->A] and 11,663[L->P] in ORF5 (Figure 2.5.C.).

At 9 dpi, in contrast to what was observed in 5 dpi semen viruses, the ∏n values dropped significantly ($p<0.0001$), suggesting the selective sweeps observed in the initial semen viruses. Very interestingly, examining the population structure and sliding window analysis of ∏n/∏s showed only a very limited positive selection localized to the nsp9 encoding region of ORF1b (Figure 2.5.E. and Figure 2.6.D.).

Overall, the analysis of the viral population dynamics during the acute infection period (0 to 9 dpi) has identified a selective genetic bottleneck event during establishment of infection in the stallion reproductive tract (around 5 to 9 dpi). This is evidenced by the initial number of limited variant population in the parental virus that gave rise to a higher number of the variant population in nasal and buffy coat viruses with a dominant purifying type of selection. That was followed by a decline in the viral quasispecies population in the initial semen viruses. The presence of a strong positive selection across multiple genomic regions and, at the same time, a shift of the quasispecies population structure was clear indication of a selective bottleneck scenario that promotes variants with an adaptive fitness to establishing infection in the stallion reproductive tract.

Additionally, the phylogenetic tree constructed from the nasal swab and buffy coat viruses showed star-like features that identify a responsible major founder virus population for establishment of the infection process. In contrast, the semen viruses (5 and 9 dpi) were shown to diverge from the founder virus population indicating the emergence of a new dominant variant in the stallion reproductive tract (Figure 2.4.).

Intra-host evolution of EAV KY84 strain in the stallion reproductive tract during persistent infection

Of the 34 sequential EAV semen isolates from seven experimentally infected stallions, 20 were isolated during the persistent infection period. Sequences from these viruses were systematically analyzed to gain a detailed insight into the characteristic feature of the EAV KY84 evolution in the stallion reproductive tract during long-term persistent infection.

During the course of this experiment, five stallions had stopped shedding virus by different dpi (stallions L137 and L138 at 170 dpi and L139, L141 and L142 at 345 dpi). In contrast, stallions L136 and L140 became long-term carriers and continued shedding virus throughout the experimental period (726 dpi). The data showed a general increase in the viral quasispecies population size and diversity in the reproductive tract of these two stallions (Figure 2.8.A.). However, it is interesting to note that quasispecies diversity at later time points (380-726 dpi, where only L136 and L140 were shedding) was significantly higher compared to any other time point in this study ($p<0.007$ to 0.0001) (Figure 2.7.A.). This indicates the presence of a higher degree of heterogeneity in the virus population of the two long-term carrier stallions at later dpi (380-726), compared to those stallions that had stopped shedding. Overall, sequential viruses in this study were more under strong purifying types of evolutionary pressure ($p< 0.0057$) than positive (diversifying) selection (Figure 2.8.A.).

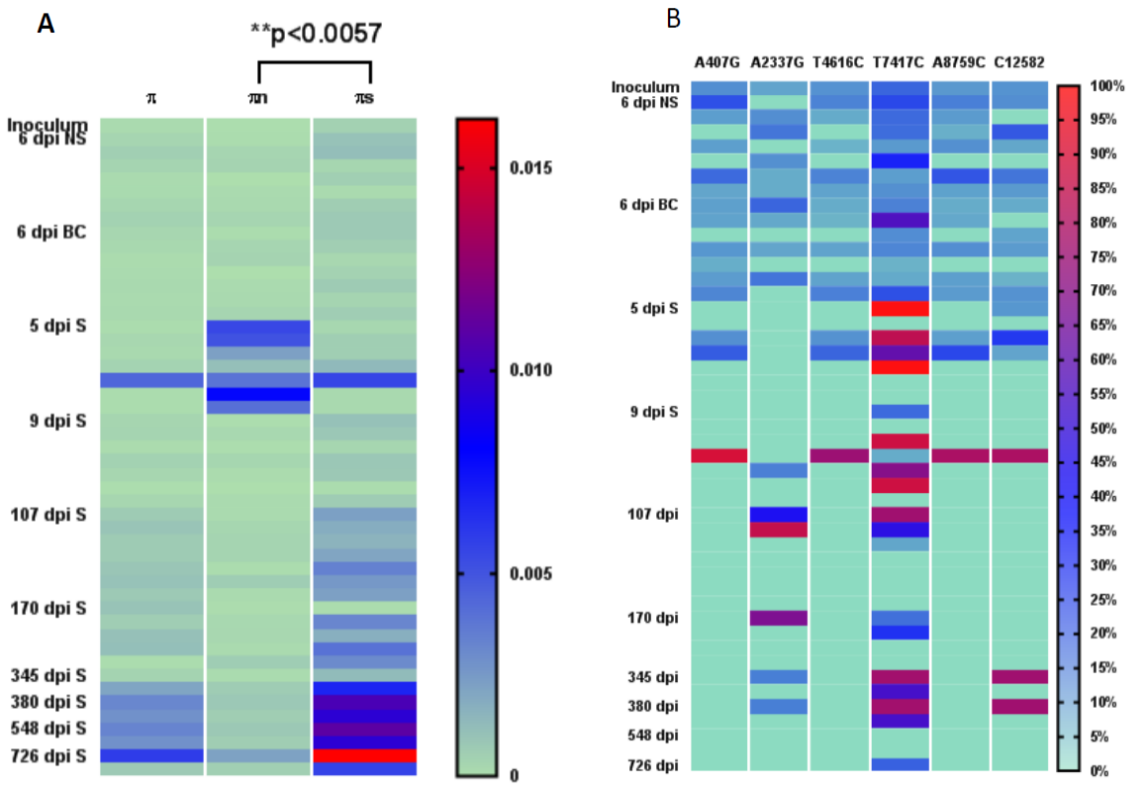


Figure 2.8. Heat map showing the evolution of 49 sequential viruses from seven experimentally infected stallions. At the full genome level, a strong purifying pressure played a significant role in long-term evolution of the sequential viruses (**p< 0.0057). (A) Nucleotide diversity (∏), nonsynonymous nucleotide diversity (∏n) and synonymous nucleotide diversity (∏s) in the 49 sequential viruses. (B) Illustrates the temporal changes in the SNV population detected in the inoculum over the two years of the experimental period. This founder virus population was detected consistently in the nasal swab (NS) and buffy coat (BC) samples, and detection and frequency shown to be variable in the subsequent semen (S) viruses.

According to the Figure 2.9., the initial population structures (in the inoculum, nasal swab and buffy coat viruses) were becoming genetically divergent with the appearance of high frequencies of minor variants. This indicates that the presence of a persistent selective pressure that continually morphs the population structure. The second important observation is the aggregation and fixation of a nonsynonymous variant population in the region of genome that encodes for the minor and major structural proteins (between nucleotide numbers 10,000 to 12,000) (Figure 2.9.). This indicates that mutations that occurred in this region have a greater role in maintaining persistent infections in the stallion reproductive tract. In order to investigate this observation further, the pooled semen viruses were computed in the analyses to find out whether the evolutionary selection was at a particular gene level. This revealed that ORF3 was under positive selection with a $\prod n/\prod s$ value of 1.19. Additionally, the sliding window analyses on pooled semen viruses identified multiple regions within nonstructural and structural genes that are under strong positive selection (Figure 2.10.).

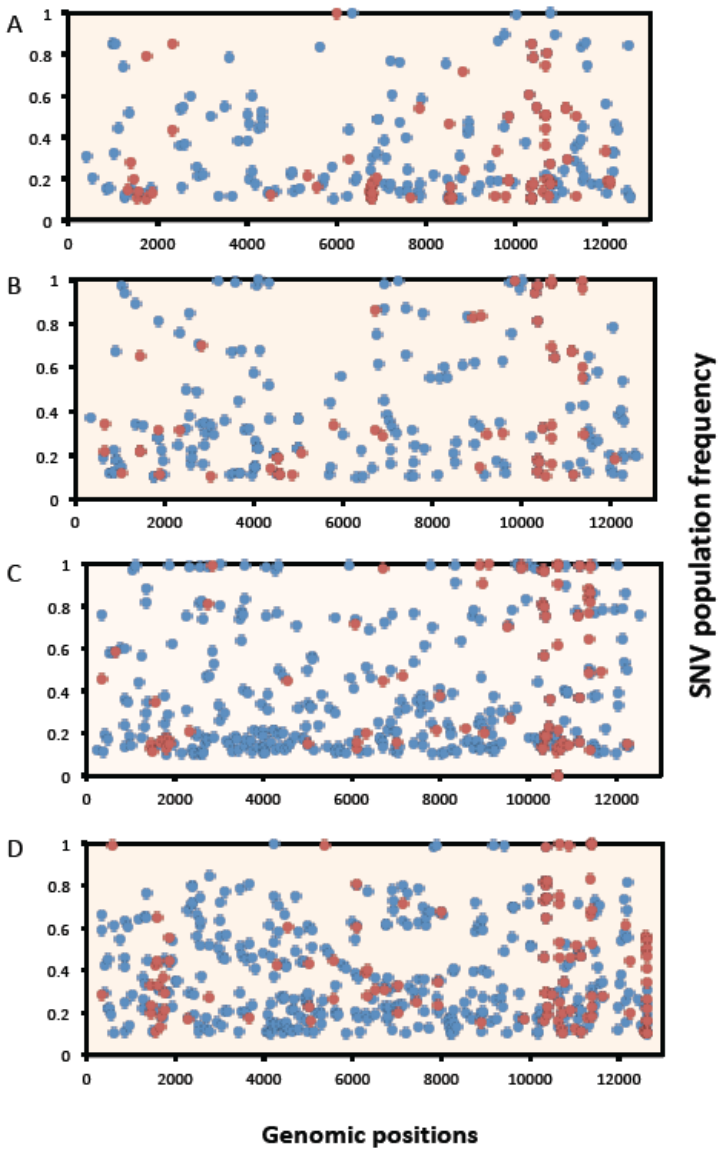


Figure 2.9. Quasispecies population structure during persistent infection (107 to 726 dpi). Each dot in the figures represents a major and minor SNV population identified from the semen viruses of experimentally infected stallions (L136-142) collected during persistently infected period; on y-axis population frequencies and x-axis genomic position are drawn in 2000 bp intervals. Blue dot (a synonymous SNV), red dot (nonsynonymous SNV). SNV population in the semen viruses at (A) 107 dpi, (B) 170 dpi, (C) 380 dpi and (D) 726 dpi.

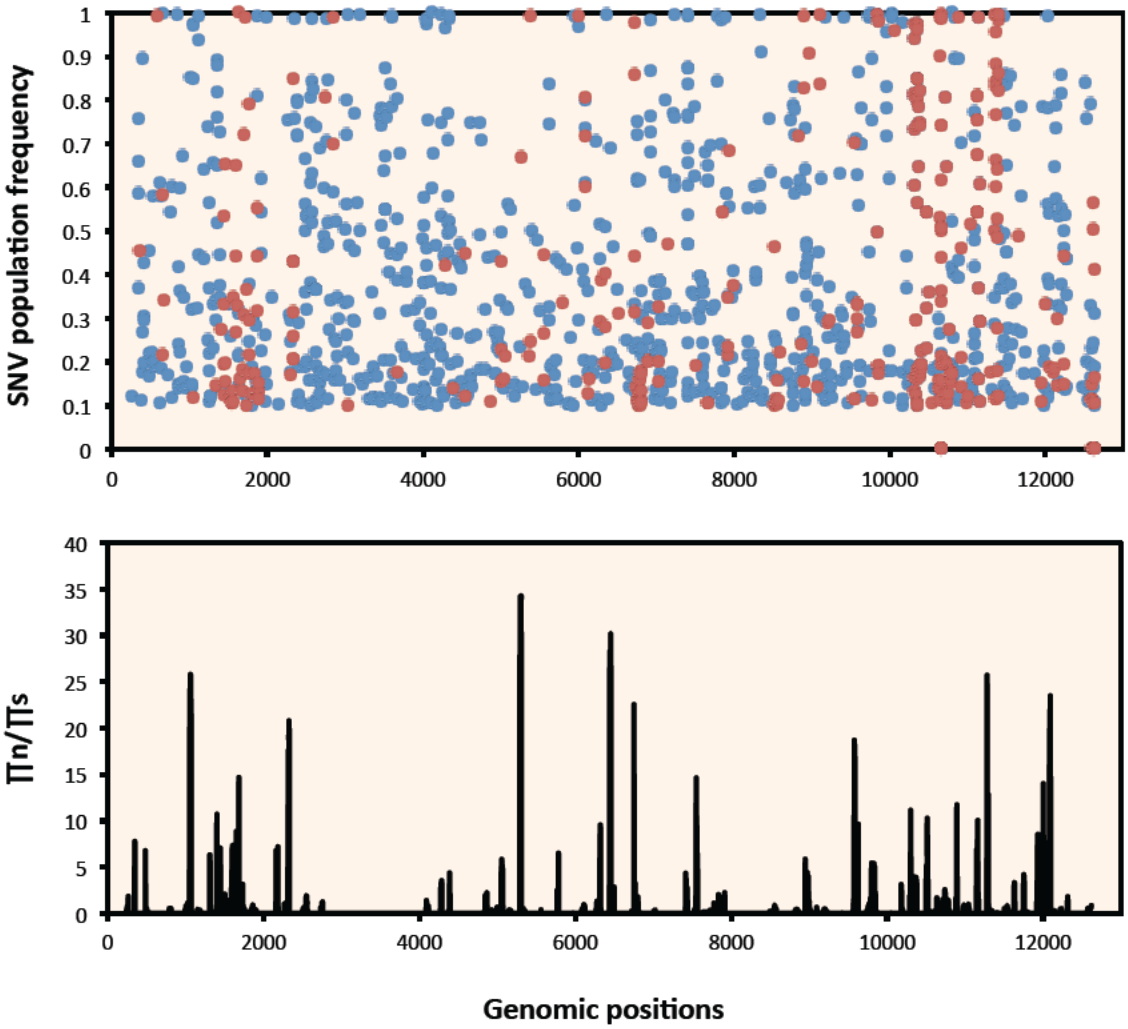


Figure 2.10. Quasispecies populations and positive selection in semen viruses. (A) SNV identified in semen viruses in the experimental period, each dot in the figures represents a major and minor SNV population; on y-axis population frequencies and x-axis genomic position are drawn in 2000 bp intervals. Blue dot (a synonymous SNV), red dot (nonsynonymous SNV). (B) Sliding window analyses for evidence of positive evolutionary selections of the sequential viruses from semen, the results were plotted against genomic position (x-axis) in 2000 bp intervals.

Non-stochastic intra-host evolution of consensus viruses during persistent infection in the stallion reproductive tract

In addition to characterizing the quasispecies population in experimental infection, the role of the major variant population was investigated by constructing a consensus sequence for each of the 48 viruses examined in this study. Comparative analyses of these sequential virus genomes have shown a total of 298 SNV fixed in the consensus genomes, and of these changes 68 were nonsynonymous substitutions. Using this information, the maximum likelihood and Bayesian phylogenetic trees were generated that show the extent of genetic divergence in the sequential viruses (Figure 2.11.A.). These viruses were evolving at a rate of 2.5×10^{-4} nucleotide substitutions /site/year.

A major observation from Bayesian inference analysis was the clustering of viruses in the phylogenetic tree in a non-stochastic manner. At the full genome level, the analyses have shown that EAV KY84 sequential viruses follow a linearly progressive type of evolution; as a result viruses from later dpi had larger branch lengths than their predecessors (Figure 2.11.). This was particularly emphasized in those viruses isolated from >107 dpi (shaded light brown in the phylogenetic tree), where viruses clustered by their date of isolation and the respective stallion. As a result, none of the viruses from two different stallions at different dpi clustered together. This observation was further confirmed with root to tip divergence analysis that demonstrated the clock-likeness of the sequential viruses by fitting the nucleotide substitution rate and date of isolation into a regression model. This showed a 0.90 correlation coefficient and $R^2 = 0.81$ (Figure 2.11.B.).

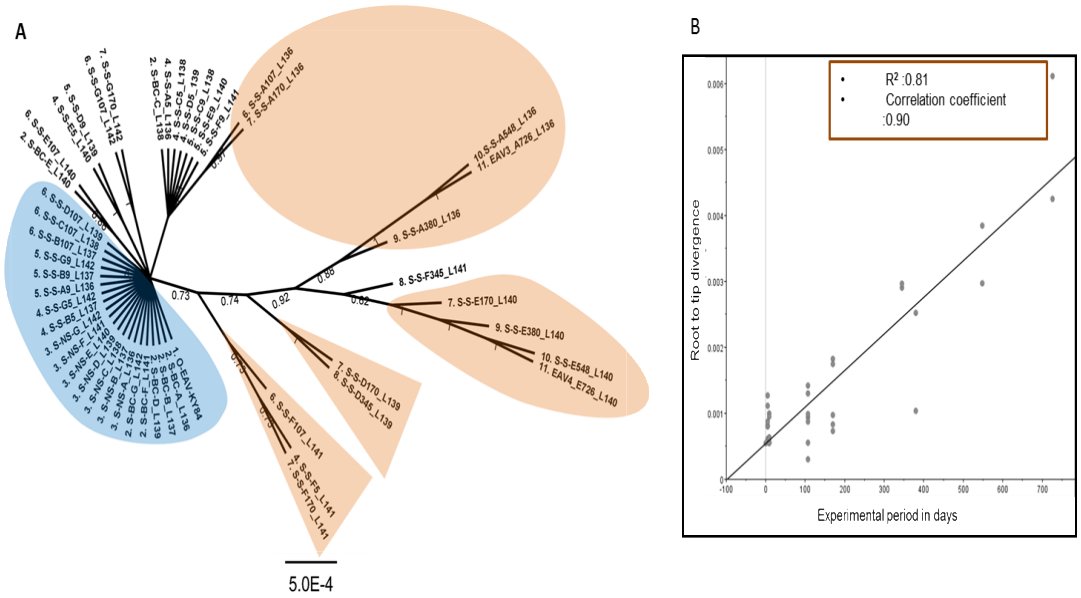


Figure 2.11. Genetic divergence and evolution of EAV KY84 sequential viruses in experimental stallions. (A) Full-genome Bayesian phylogenetic tree constructed from all sequential viruses in the experimental period (726 dpi) and the numbers in each branch showing posterior probabilities calculated from 1000 replicates. Bar under the graph indicates the average pairwise distance of the sequential isolates from the original EAV KY84 strain described as the nucleotide substitution rate /site. The light brown shaded areas showing strict non-stochastic clustering of viral sequence; light blue shaded area shows clustering of sequences mostly from the period of acute infection. (B) Root to tip divergence analysis justified the use of molecular clock analysis on the sequential isolates in this study.

Intra-host evolution and genetic divergence in the structural region of consensus viruses during persistent infection

In order to get a deeper insight on the important genomic region that signifies the evolutionary process, the data were systematically analyzed and these changes weighted by ORFs. The diversity estimation used to measure evolution was pairwise distance matrix to see how these sequential viruses diverged from the inoculum virus (original EAV KY84). This analysis at the level of the full genome assumes equal rate of substitution in various genomic segments over a period of time and hence was not satisfactory to identify the role of a particular region in the evolutionary process. Therefore, the analyses were further divided into by different ORFs and calculated a pairwise distance; this showed ORFs 5 and 3 are the two most evolving regions at rate of 3×10^{-3} nucleotide substitutions/site resulting in sequence divergence of 1 to 1.4% respectively. ORF1ab and ORF7 were shown to be the two most conserved regions with only 0.3 and 0.5 % of divergence from the parental EAV KY84 over the two years of the experimental period.

These changes accumulating in the genomes of sequential isolates have either a purifying or a selective role in the evolution of viral populations in the reproductive tract. To characterize the changes in such a way, the ORF1ab region was broken down into nonstructural protein coding regions (nsp1 to 12) and included the tally of synonymous and nonsynonymous substitutions (Figure 2.12.). The analysis showed that despite a higher number of synonymous changes occurring in the ORF1ab region, most nonsynonymous changes were concentrated in the nsp2 encoding region of ORF1a. The 1.7 kb-long nsp2 had undergone 29 synonymous and 14 nonsynonymous substitutions over a period of 726

days; this constitutes 42 % of the total nonsynonymous substitutions for the whole of nonstructural region.

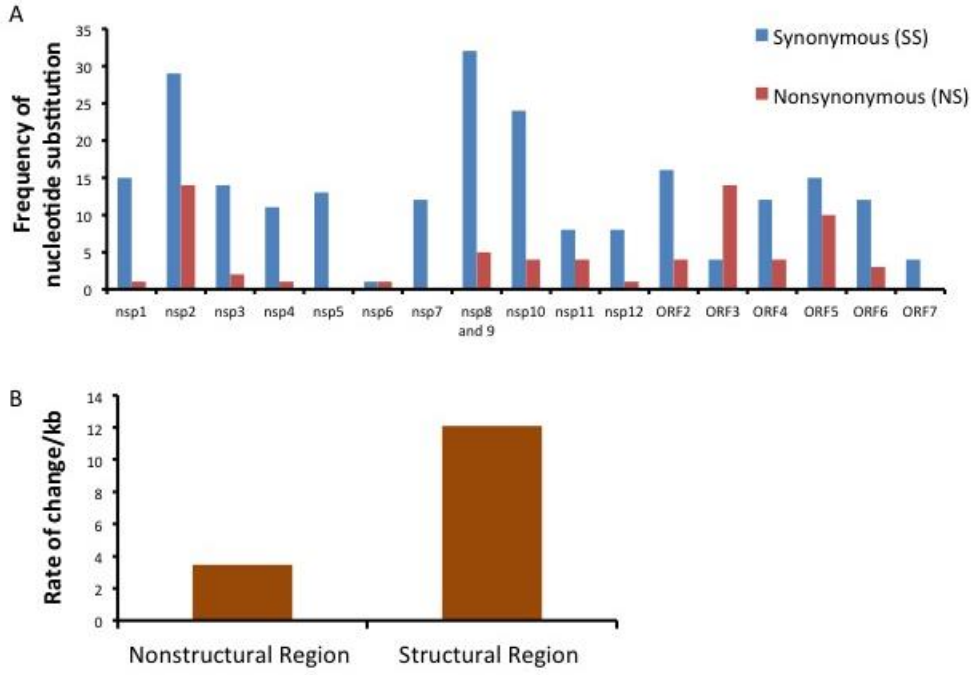


Figure 2.12. Rate of synonymous and nonsynonymous substitutions across genomic regions and ORFs of the sequential viruses. (A) Tally of all nonsynonymous and synonymous substitutions in the ORFs and peptide regions of the sequential viruses. (B) Graph showing the rate of variation by which nucleotide substitution occurred in the region of EAV KY84 sequential genomes, 13 nucleotide substitution/kb in the structural region and only 3 substitutions/kb in nonstructural region.

In the structural protein encoding region (ORFs 2-7) of EAV, the sequential isolates had a higher rate of nonsynonymous substitution compared to nonstructural region (Figure 2.12.B.). The data shows that for every kb in the nonstructural region, there were 3.3 nonsynonymous substitutions over the two year period of the study; whereas there were 12

nonsynonymous substitutions /kb in the structural regions, clearly suggestive of a higher rate of fitness-associated substitution in the structural region. This observation is also strongly supported by ∏s and ∏n analysis, which indicated a strong positive selection in ORF3 and also in ORF5 (Figure 2.13.). The result also demonstrated that the nsp2 peptide had undergone a positive selection (∏n / ∏s >1) indicating an important role for this region in the adaptive evolutionary process.

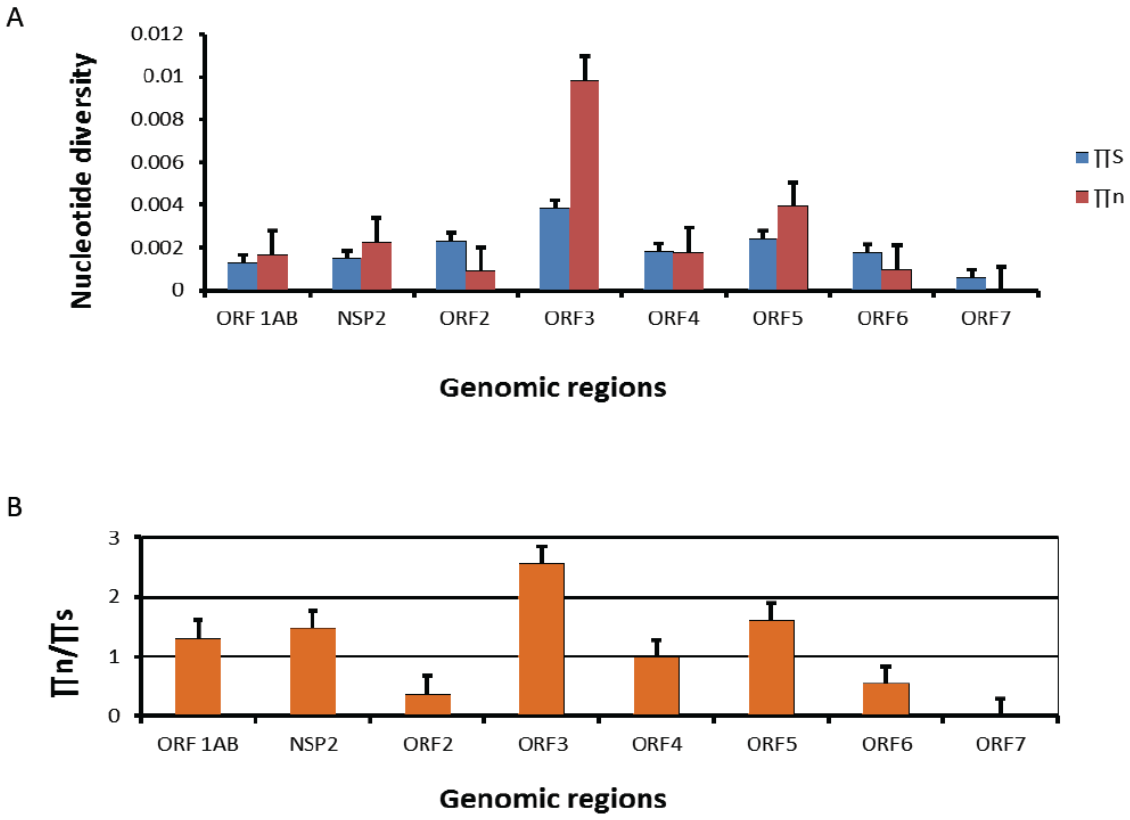


Figure 2.13. Comparison of nonsynonymous to synonymous nucleotide diversity in the experimental horses. (A) Nucleotide diversity estimate by regions of the genome ∏n (nonsynonymous diversity) ∏s (synonymous diversity). (B) Showing the presence of positive selection in ORFs that have ∏n/∏s value > 1.

To further investigate which specific codons were under selection pressure, the sequences from ORF3 and ORF5 were used for a codon level estimate of positive selection. A maximum likelihood reconstruction of ancestral states using a Muse-Gaut model employed under MEGA7 was used to test the statistic dN - dS for detecting codons that have undergone positive selection, where dS is the number of synonymous substitutions per site (s/S) and dN is the number of nonsynonymous substitutions per site (n/N). A positive value for the test statistic indicates an overabundance of nonsynonymous substitutions. Analysis of ORF3 using this model demonstrated most codon positions in this gene were found to be overabundant with nonsynonymous substitutions (Table 2.1.) indicating a positive selection pressure in the codons highlighted yellow. However unlike ORF3, ORF5 contained a higher number of negative dN – dS values per codon site, indicating an overabundant synonymous codon site. The values for ORF5 are given in Table 2.2. and the overabundant nonsynonymous sites are highlighted in yellow.

Table 2.1. Nonsynonymous and synonymous substitution site estimate for ORF3, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions

Codon position	Triplet	Syn (s)	Nonsyn (n)	Syn sites (S)	Nonsyn sites (N)	dS	dN	dN-dS	Normalized dN-dS
3	CAT	0	6	0.472316935	2.32462463	0	2.58106	2.581061875	48.4119334
19	CGC	0	2	0.604089237	2.21492879	0	0.90296	0.902963568	16.93652235
22	TTT	0	1	0.513023997	2.486976	0	0.40209	0.402094752	7.541928601
27	AGC	0	1	0.472316935	2.5230496	0	0.39635	0.396345755	7.43409699
28	GAT	0	3	0.472316935	2.50739043	0	1.19646	1.196463049	22.44157338
30	ATT	0	1	0.752124388	2.24787561	0	0.44486	0.444864473	8.344142954
56	TTC	1	0	0.472316935	2.52768307	2.11722	0	-2.117222413	-39.71188428
84	TTT	0	1	0.502792145	2.49720786	0	0.40045	0.400447243	7.511026929
116	CAT	0	1	0.456816071	2.50674798	0	0.39892	0.39892323	7.482441643
118	CAC	0	1	0.604271254	2.39572875	0	0.41741	0.417409526	7.829181613
119	TTA	0	3	0.901384196	1.69266481	0	1.77235	1.772353267	33.24331323
120	ACT	0	1	0.994498099	2.0055019	0	0.49863	0.498628298	9.352569267
121	ACG	0	1	0.961526306	2.03847369	0	0.49056	0.490563112	9.201293841
134	GAC	1	0	0.472316935	2.52768307	2.11722	0	-2.117222413	-39.71188428
137	CTG	1	0	1.426311798	1.55975712	0.70111	0	-0.701108973	-13.15041737
145	TGC	1	0	0.472316935	2.28646425	2.11722	0	-2.117222413	-39.71188428

Table 2.2. Nonsynonymous and synonymous substitution site estimate for ORF5, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions

Codon position	Triplet	Syn (s)	Nonsyn (n)	Syn sites (S)	Nonsyn sites (N)	dS	dN	dN-dS	Normalized dN-dS
2	CTA	2	0	1.317034821	1.5911086	1.51856	0	-1.5186	-41.39089036
6	GTA	0	1	1	2	0	0.5	0.5	13.62831097
61	CAA	1	0	0.294983231	2.2754949	3.39002	0	-3.39	-92.40058102
67	ACA	1	0	1	2	1	0	-1	-27.25662195
70	GAC	0	1	0.42952183	2.5704782	0	0.38903	0.38903	10.60371656
71	GAA	0	1	0.312863777	2.3364825	0	0.42799	0.42799	11.66566497
73	ATT	0	1	0.870981837	2.1290182	0	0.4697	0.4697	12.80243749
80	TGT	1	0	0.42952183	2.3142412	2.32817	0	-2.3282	-63.45805981
82	GAC	0	2	0.42952183	2.5704782	0	0.77807	0.77807	21.20743312
84	TAT	0	3	0.42952183	2.2752715	0	1.31852	1.31852	35.93850918
91	GTT	1	0	1	2	1	0	-1	-27.25662195
104	GAT	0	1	0.480582678	2.5194173	0	0.39692	0.39692	10.81862132
108	CCT	1	0	1	2	1	0	-1	-27.25662195
109	TTC	1	0	0.42952183	2.5704782	2.32817	0	-2.3282	-63.45805981
110	ATC	1	0	0.685758836	2.3142412	1.45824	0	-1.4582	-39.74665805
111	TAC	1	0	0.42952183	2	2.32817	0	-2.3282	-63.45805981
117	GGC	1	0	1	2	1	0	-1	-27.25662195
121	TTG	2	0	0.668222846	2.036285	2.99301	0	-2.993	-81.57943744
138	CTA	2	0	1.24209339	1.6048531	1.61018	0	-1.6102	-43.88820064
154	ATA	1	0	0.698375107	2.3016249	1.4319	0	-1.4319	-39.02862755
173	CTC	0	1	1	2	0	0.5	0.5	13.62831097
187	CTG	1	0	1.347037414	1.6255227	0.74237	0	-0.7424	-20.2344951

Intra-host evolution and genetic diversity of EAV KY84 strain in naturally infected stallions over a span of 10 years

In addition to the experimentally infected horses, there was the opportunity to analyze virus sequences from two naturally infected stallions. The two Thoroughbred stallions (E and D) were on the farm where the first major EVA outbreak in Thoroughbred horses in Kentucky occurred. These two stallions were initially infected with EAV KY84 (inoculum virus) that was used for experimental inoculation of the seven stallions. However, the two horses subsequently developed the natural carrier state and continued to shed the virus in the semen for a period 7 to 10 years (stallion E, 1984 to 1991 and stallion D, 1984 to 1994, respectively). In this study, the complete genome of four viruses that were isolated from the semen of stallion E in 1984, 1991 and stallion D1984, 1994 (these viruses will be referred as E or D followed by year of isolation) were sequenced and analyzed.

Comparative nucleotide and amino acid analyses showed that E84 and D84 viruses from naturally infected stallions are very similar to the inoculum virus that was used for experimental infection of the other stallions. The only differences identified were 9 substitutions in the D84 and E84 viruses compared to the EAV KY84 strain used to inoculate the stallions. These changes resulted in three nonsynonymous substitutions in D84 virus at position (583 T>I (nsp1), 5376 D>A (nsp9) and 10,886 F>L (GP4) and only one nonsynonymous substitution in E84 virus at 5376 D>A (nsp9) (Figure 2.14.A.).

The quasispecies population structures in E84 and D84 viruses have some collinearity with what was observed with viruses from the semen of the experimental stallions, by having a number of minor variants (Figure 2.15.A and B.). However, the

E91 and D94 viruses showed a clearing of the high frequencies of minor variants with a new distinct population structure of new minor and new major variants (Figure 2.15.C and D.).

Similar to the observation made in the experimental samples, a higher level of nonsynonymous fixation was shown to aggregate in the structural regions of the genome. For a detailed characterization of these changes, the consensus sequence was calculated as genome-wide pairwise distance for stallion D and E viruses. The results showed sequence divergence by 1.6 and 2% over the span of seven year for E91 and ten years for D94 (Figure 2.14.B.).

Previous analyses of viruses from the experimentally infected stallions showed that evolution of ORFs occurred at a variable rate. Hence, the extent of evolution was investigated between the various ORFs in these four sequential viruses. In agreement with earlier observation, different ORFs evolved independently with some accumulating more numerous changes than others. In the 10 year period, ORF3 was shown to have a 4.5% sequence divergence from the original isolates, diverging at a rate of 2 .6 x 10^{-2} nucleotide substitutions / site (Figure 2.14.B.). By contrast, ORF7 was demonstrated to have the lowest evolutionary rate with 2 x 10^{-3} nucleotide substitutions / site, which resulted in a 1.2% pairwise sequence divergence from the original EAV KY84 genome (Figure 2.14.B.).

These changes in various ORFs have brought 461 substitutions, which included 361 synonymous and 96 non-synonymous substitutions (Figure 2.15.C and D.). A higher proportion of the nonsynonymous substitutions was shown in the structural protein

coding region where 19.3 nonsynonymous substitutions / kb, whereas 4.1 amino acid substitutions / kb were seen in the nonstructural region of the EAV genome.

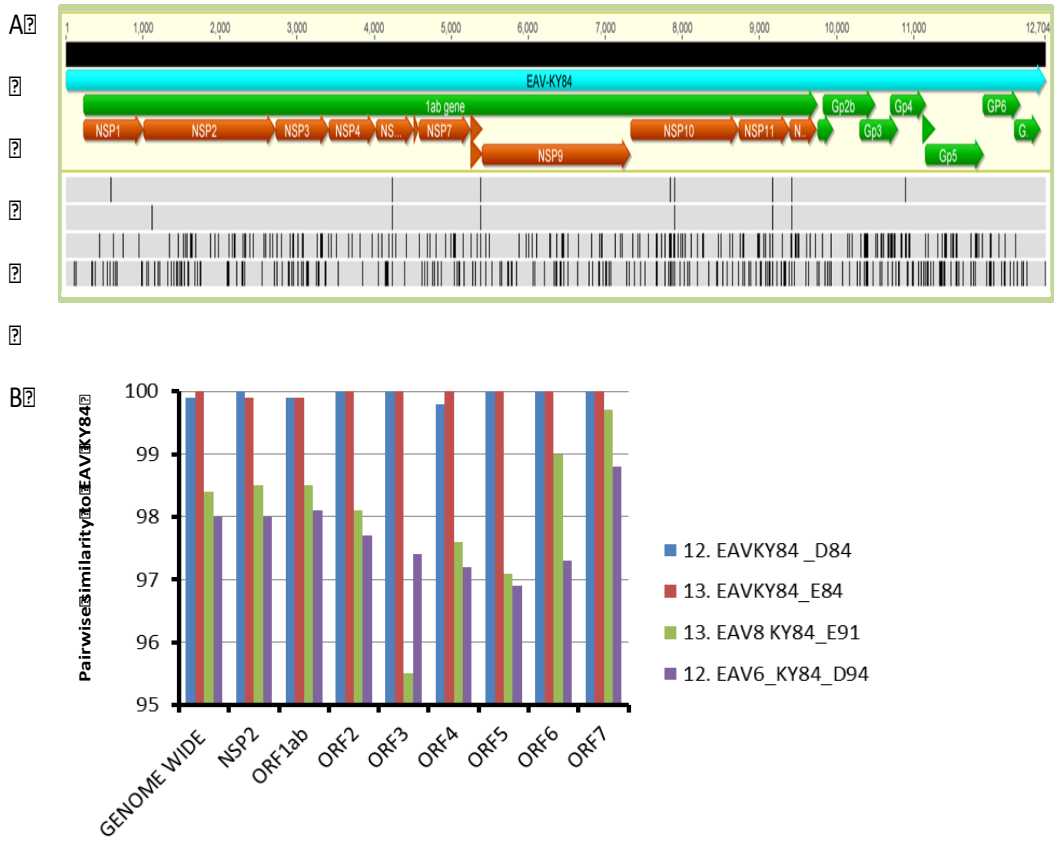


Figure 2.14. Sequence alignments and comparison of sequential viruses in naturally infected stallions D and E. (A) Multiple genome alignment of four sequential viruses with the parental EAV KY84. The genome arrangement with arrowed boxes indicating the location of peptides and important genes. A vertical line indicates disagreement between the four viruses from the two naturally infected stallions with the parental EAV KY84 strain. (B) Bar graph showing the pairwise similarity of these sequential genomes from naturally infected stallions. Genome wide means comparing pairwise distance of consensus sequences from stallion D and E viruses.

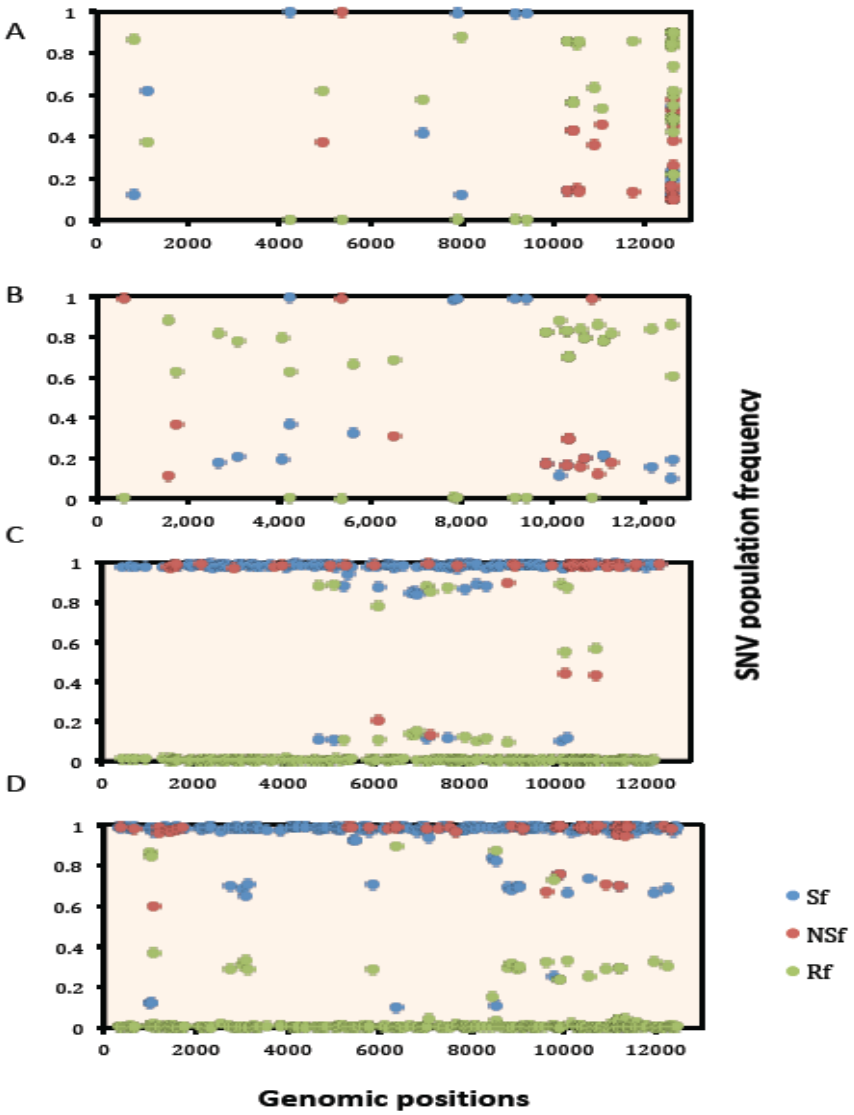


Figure 2.15. Quasispecies population structure in naturally infected stallions D and E. The stallions D and E were infected with EAV KY84 strain are very similar to the inoculum virus used for experimental infection of the seven stallions. Each dot in the figures represents a major and minor SNV population; on y-axis population frequencies and x-axis genomic position are drawn in 2000 bp intervals. Blue dot (a synonymous SNV), red dot (nonsynonymous SNV) and light green (reference variant). (A) SNV population frequencies of E84, (B) SNV population frequencies of D84, (C) SNV population frequencies of E91 and (D) SNV population frequencies of D94.

To examine these changes and their role in the adaptive evolutionary process, the computed ∏n/∏s analysis showed the presence of selection pressure in ORF3 and ORF5 (Figure 2.16.). However, contrary to what is observed in the experimentally infected the stallions, those changes in the nsp2 region were found to be purifying rather than associated with evolutionary fitness. Based on these results, further investigation was focused on the codon level selection estimate on ORF3 where multiple nonsynonymous overabundant codon sites were identified isolates (n=53) of both experimentally and naturally infected horses that are highlighted yellow in Table 2.3.

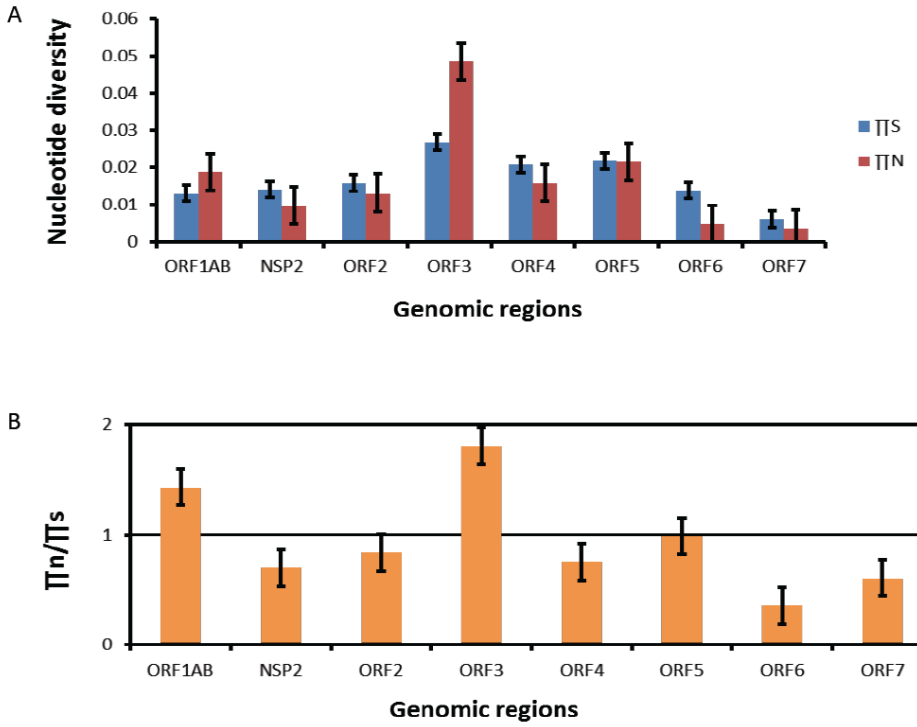


Figure 2.16. Comparison of nonsynonymous to synonymous nucleotide diversity of viruses from naturally infected stallions. (A) Nucleotide diversity estimate by regions and ORFs in the EAV KY84 genome, ∏n (nonsynonymous nucleotide diversity) ∏s (synonymous nucleotide diversity). (B) Showing the presence of positive selection in ORFs that have ∏n/∏s value >1.

Table 2.3. Nonsynonymous and synonymous substitution site estimate for ORF3 was identified in virus isolates from both experimentally and naturally infected horses, values shaded in yellow are codon positions that have more nonsynonymous substitutions than synonymous substitutions

Codon positions	Triplet	Syn (s)	Nonsyn (n)	Syn sites (S)	Nonsyn sites (N)	dS	dN	dN-dS	Normalized dN-dS
3	CAT	0	1	0.474379292	2.3717036	0	0.4216378	0.4216378	6.4809292
16	CTC	0	1	0.846082888	2.1539171	0	0.4642704	0.4642704	7.1362278
17	TAT	0	1	0.474379292	2.0762514	0	0.4816373	0.4816373	7.4031707
18	TTT	0	1	0.545846333	2.4541537	0	0.4074724	0.4074724	6.2631951
19	CGC	0	1	0.846082888	2.0824501	0	0.4802036	0.4802036	7.3811341
24	AGT	0	1	0.474379292	2.5256207	0	0.3959423	0.3959423	6.0859666
25	GTT	0	1	0.959153795	2.0408462	0	0.4899928	0.4899928	7.5316029
27	AGC	0	1	0.55063065	2.4493694	0	0.4082684	0.4082684	6.2754289
28	GAT	0	1	0.474379292	2.5256207	0	0.3959423	0.3959423	6.0859666
30	ATT	0	1	0.982514416	2.0174856	0	0.4956665	0.4956665	7.6188119
54	TCG	1	0	1	1.7150967	1	0	-1	-15.370843
65	GAT	1	0	0.474379292	2.5256207	2.1080178	0	-2.1080178	-32.402011
72	TAT	1	0	0.474379292	2	2.1080178	0	-2.1080178	-32.402011
83	CTG	1	0	1.358992126	1.6056027	0.7358394	0	-0.7358394	-11.310473
85	GCC	1	0	1	2	1	0	-1	-15.370843
88	GTT	0	1	1	2	0	0.5	0.5	7.6854216
93	AGC	1	0	0.474379292	2.5256207	2.1080178	0	-2.1080178	-32.402011
99	CGC	0	1	0.846082888	2.1539171	0	0.4642704	0.4642704	7.1362278
117	ACC	1	0	1	2	1	0	-1	-15.370843
118	CAC	0	1	0.628296404	2.3717036	0	0.4216378	0.4216378	6.4809292
119	TTA	0	1	0.818373064	1.6560062	0	0.6038625	0.6038625	9.2818752
120	ACT	0	2	0.917549929	2.0824501	0	0.9604072	0.9604072	14.762268
121	ACG	0.5	1.5	0.849918247	2.1500818	0.5882919	0.6976479	0.109356	1.6808941
131	GTA	0	1	0.92129199	2.078708	0	0.481068	0.481068	7.3944215
134	GAC	1	0	0.474379292	2.5256207	2.1080178	0	-2.1080178	-32.402011
137	CTG	1	0	0.728380473	2.0427188	1.3729089	0	-1.3729089	-21.102767
145	TGC	1	0	0.474379292	2.2815637	2.1080178	0	-2.1080178	-32.402011
148	GCC	0	1	1	2	0	0.5	0.5	7.6854216
153	CAC	1	0	0.474379292	2.5256207	2.1080178	0	-2.1080178	-32.402011

Mirroring the trend in the comparative evolution between experimentally and naturally infected stallions

As described in the earlier section, one advantage of the sequences from naturally infected isolates extended analyses to a ten-year window and enabled us to understand the EAV evolutionary process over a longer time period. Hence, the evolutionary rate of EAV KY84 was calculated from the combined sequences from the experimental infection with isolates from stallion D and E. A full genome multiple alignment of 53 sequential viruses was used to perform a Bayesian phylogenetic tree (Figure 2.17.). The result showed that EAV KY84 isolates were evolving at rate of 4×10^{-4} nucleotide substitutions / site / year. This is slightly higher nucleotide substitution rate than that individually estimated for the experimental stallions (Figure 2.17.).

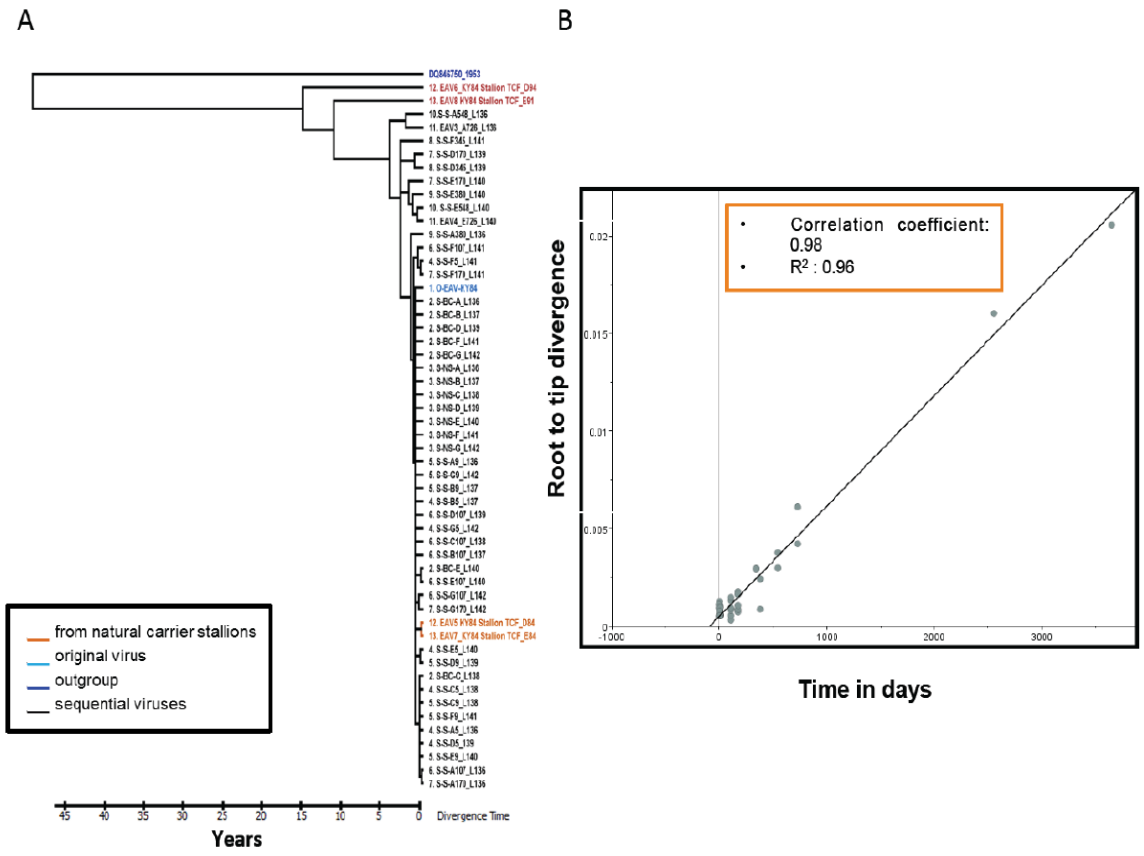


Figure 2.17. Estimation of evolutionary rate of the sequential viruses. (A) Ancestral reconstruction using a Bayesian phylogenetic. (B) Root to tip divergence analysis justifying the use of a molecular clock analysis on the sequential isolates in this study.

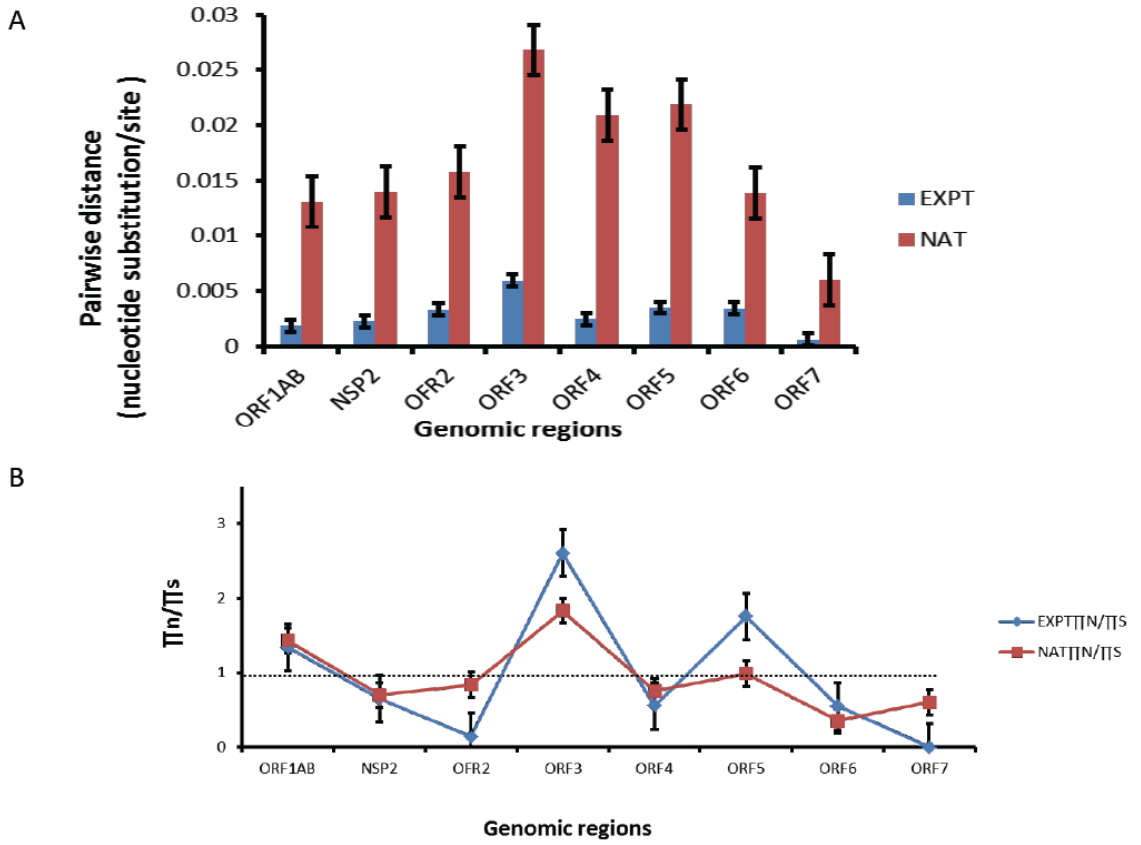


Figure 2.18. Comparative trend in nucleotide diversity and selection between viruses from the natural and experimental infections. (A) Collinearity in trends of nucleotide diversity in EXPT (sequential viruses from experimental infection), NAT (sequential viruses from natural infection). (B) Collinearity in trends of positive and negative selection in EXPT (sequential viruses from experimental infection), NAT (sequential viruses from natural infection) various genomic region sequential viruses. Horizontal line showing where the ratio ∏n/∏s is 1, which indicates equal rate of nonsynonymous and synonymous diversity.

In addition, these combined sequences were used to compare and contrast virus evolution and selection pressure between experimentally and naturally infected stallions. This analysis has showed that the rate of nucleotide substitution / site described as an average pairwise distance in Figure 2.18.A. and the diversity estimate ratio of ∏n/∏s obtained across all ORFs followed a mirroring trend across genomic regions of viruses from experimentally and naturally infected stallions (Figure 2.17.). These comparisons also confirmed the previous observation that ORF3 and 5 have a prominent role in the evolutionary process of viruses in experimental as well as natural infections.

In order to visually identify codons or regions of such importance in ORF3 and 5, the FUBAR (fast-unconstrained Bayesian approximation for inference of selection) was implemented in datamonkey server and fitted the data with a codon substitution model. This method uses an empirical Bayes procedure in the phylogenetic likelihood framework to identify codon regions with statistically significant evidence for diversifying and purifying selection. The analysis has identified seven sites with diversifying (positive selection) in ORF3 with a posterior probability $\geq$ 0.9 at codon positions 3, 19, 27, 28, 30, 119 and 120 (Figure 2.19.). It also identified four sites with evidence of purifying (negative selection) at posterior probability $\geq$ 0.9, codon positions 93, 134, 137 and 145 (Figure 2.19.).

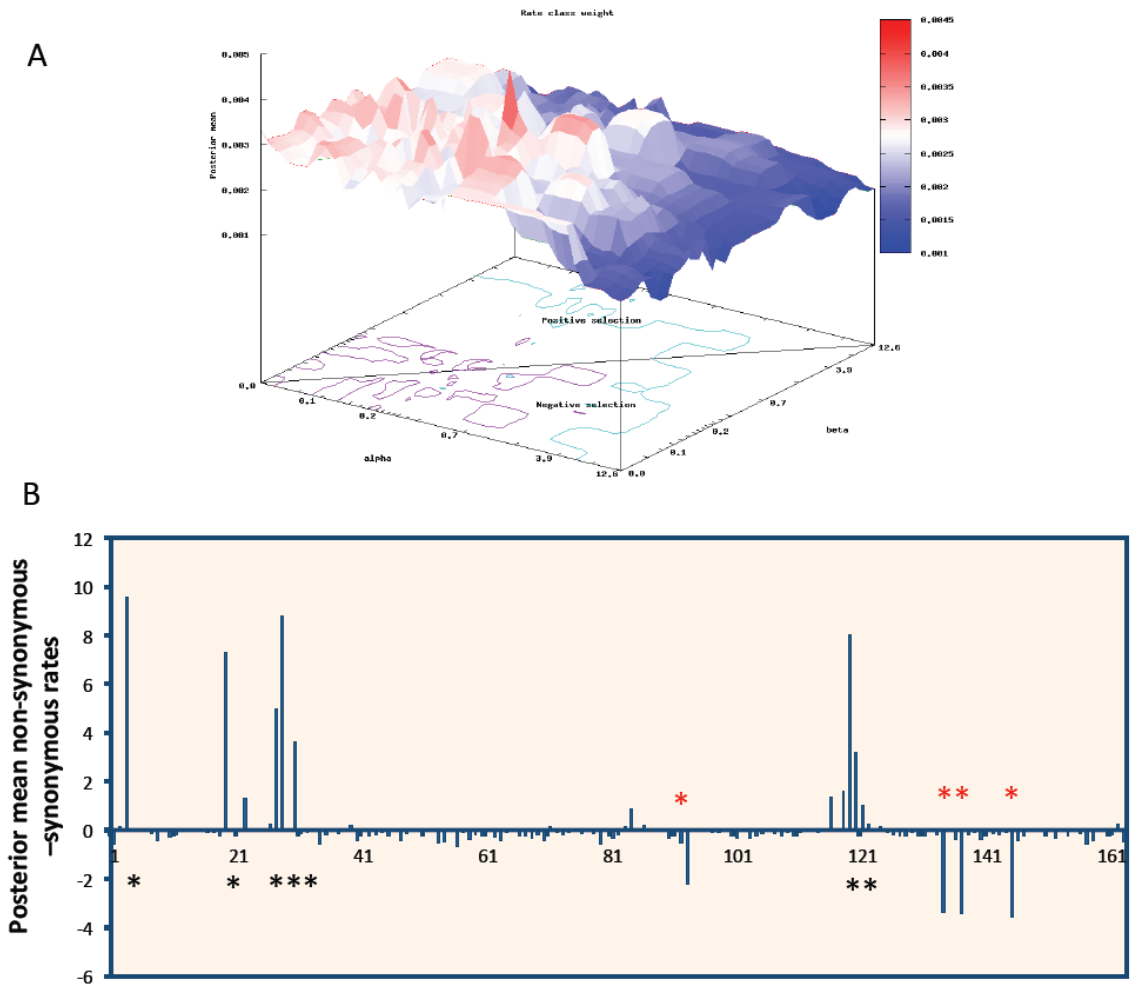


Figure 2.19. Fast-unconstrained Bayesian approximations for inferring selection (FUBAR) analysis of ORF3. (A) A rate class weight analysis of positive and negatively selected codon positions with heat map calculated from posterior mean analysis of nonsynonymous (beta) over synonymous sites (alpha) by the FUBAR algorithm. (B) Codon positions with * (black) are the seven sites with statistically significant evidence for diversifying selection at posterior probability of $\geq$ 0.9 and * (red) the four sites identified with statistically significantly purifying selection at posterior probability of $\geq$ 0.9.

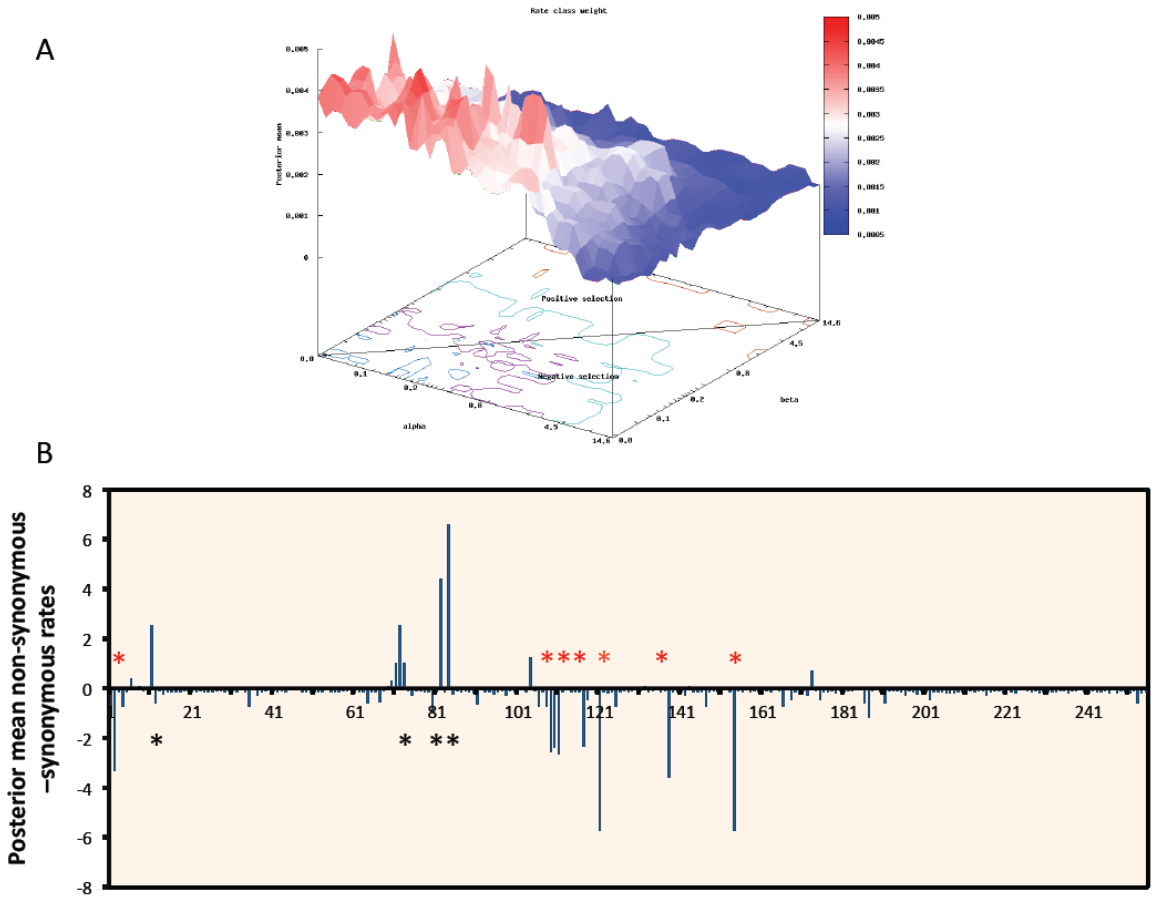


Figure 2.20. Fast-unconstrained Bayesian approximations for inferring selection (FUBAR) analysis of ORF5. (A) A rate class weight analysis of positive and negatively selected codon positions with heat map calculated from posterior mean analysis of nonsynonymous (beta) over synonymous (alpha) sites by FUBAR algorithm. (B) The four codon positions with * (black) are identified with evidence for diversifying selection at posterior probability of $\geq$ 0.9 and * (red) showing the seven sites identified with purifying selection at posterior probability of $\geq$ 0.9.

Similarly in ORF5, FUBAR has identified four statistically significant codon sites with diversifying (positive selection) at codon positions 11, 72, 82 and 84 and eight sites (at codon positions 2, 109, 110, 111, 117, 121, 138 and 154) with purifying (negative selection) at posterior probability of $\geq$ 0.9 (Figure 2.20.). The analysis by FUBAR was based on pooled signals of all the phylogenetic tree branches; however, some studies have suggested using the mixed effect model of evolution (MEME) that is more sensitive than FUBAR for detection of selection on subsets of viral linages[425–427]. For this reason, ORFs 3 and 5 were subjected to MEME analysis as implemented in the adaptive evolution server of datamonkey. By doing so, additional sites of statistically significant value have been identified (Table 2.4.). More interestingly, four out of the seven codon sites in ORF3 were in concordance for evidence of diversifying positive selection in both FUBAR and MEME algorithms (Table 2.4.).

Table 2.4. Analysis of positively selected codon sites in ORF3 and ORF5 of sequential viruses

		Analysis for positive selection	
	Codon positions	[1]FUBAR (posterior probability$\geq$0.9)	[2]MEME ($p \leq 0.05$)
ORF3	3	0.985871	0.0335392
	19	0.976159	0.0143498
	27	0.906460	NS
	28	0.970072	0.0502667
	30	0.973041	NS
	119	0.966733	0.0556173
	120	0.969391	NS
ORF5			
	11	0.900558	NS
	82	0.940842	0.0579575
	84	0.970840	0.0507463

By combining the information in Table 2.4. that identifies the sites under diversifying selection and localization of the codons in Figure 2.19. and Figure 2.20., it is possible to see the pattern of selection in ORF3 and 5. In this regard, the region that contains the first 30 codons of ORF3 (an overlapping region with ORF2) is shown to preferentially undergo a positive selection pressure, and this may indicate a fitness-associated role in this region during persistent infection; whereas in ORF 5, the region between codon positions 100 to 154 as under significant influence of purifying type of selection, which showed that genetic drift is the major driver of evolution in this region.

2.5. Discussion

Experimental evolutionary models constitute the most powerful means of recapitulating natural evolution of viruses, and they have been used to test basic theories/hypotheses and elucidate the mechanisms of viral survival and spread in short-term and long-term viral infections[428–430]. In this study, seven sexually mature stallions were used for experimental inoculation of the moderately virulent strain of EAV KY84 and then these stallions were followed for up to 726 dpi. This appears to be the first study to experimentally characterize the intra-host evolutionary process of the prototype arterivirus in its natural host. Similar studies for other RNA viruses have been shown to elucidate viral evolution across a single transmission event and over the course of the infection period[416,431–433]. However, a major strength of this study was the ability to follow and sample experimentally and naturally EAV infected horses for relatively long periods of time (~up to 2 and 10 years, respectively). The only other study that has looked at the evolution of EAV in longitudinal samples was that of Hedges *et al*.

(1999)[142], where sequential viruses from two naturally infected stallions (D and E) during the 1984 EAV outbreak in Kentucky were characterized. However, that study used traditional RT-PCR amplification, cloning and Sanger sequencing methods to characterize EAV ORFs 2 to 7 in sequential EAV KY84 viruses present in semen. With significant improvement on this study, four sequential viruses were included from horses (D and E) that were naturally infected during the 1984 EAV outbreak and employed NGS technology to evaluate the intra-host evolution and compare it to the analysis from the experimental stallions. In total, this study was based on data from 53 sequential viruses that have been deep sequenced using an Illumina Miseq technology and used to reconstruct the quasispecies population dynamics and overall genetic divergence of EAV in the acute and persistent stages of the infection.

Recent advances in sequencing technology allow viral evolution to be studied with deep insight into viral quasispecies populations[429,434]. As a result, investigating intra-host viral evolutions in longitudinal samples has become a part and parcel of pathogenesis studies[285,397,435]. These types of studies have demonstrated that rapidly evolving RNA viruses exist as genetically related but highly diverse and complex viral populations[436,437]. These observations followed by identification of the viral evolution in dynamic population structure indicate that variants undergo temporo-spatial change and regulate quasispecies composition that can result in the emergence of new phenotypes[438,439].

The first reports on identification of EAV quasispecies populations were those based on T_1 oligonucleotide fingerprinting of viruses isolated from naturally infected stallion semen[142]. Subsequently, others characterized the genetic divergence of this virus

mainly by targeted sequencing of structural proteins encoding ORFs, mainly ORF5. In this study, the EAV KY84 quasispecies populations in the inoculum and in nasal, buffy coat and sequential semen samples were reconstructed. In agreement with observations for other RNA viruses, the major and minor variant populations were defined in the inoculum virus. These variants were traced in nasal swabs and buffy coat samples that comprised successful establishment of experimental infection from the viral inoculum populations. In the initial samples from the experimentally infected stallion, the founder effect had been shown to result in loss of certain populations from the inoculum during viral transmission. This might be associated with the efficiency of the nasopharyngeal inoculation technique used to infect horses in this study. In larger animal models such as horses, the presence of a more extensive surface area in the nasal cavities coupled with a high viral challenge dose in the inoculum could facilitate efficient transmission of viral populations. A similar finding has also been reported in the case of intra-host evolutionary studies of influenza virus where transmission bottleneck events occurring in experimental infections of ferrets and mice were determined by the route of inoculation[434,440].

The presence of a high rate of mutation in RNA viruses is beneficial for rapid adaptation to their microenvironment; at the same time, it has the disadvantage of random generation of non-viable viruses[285,439]. However, contemporary thought on RNA virus evolution suggests that a cloud of potentially beneficial mutations occur at the population level[285,428,439] . An interesting observation from this study was that viruses from nasal swabs and buffy coats have distinct quasispecies population sizes and selection patterns. This was examined as to how this cell population becomes infected. EAV initially

invades the upper respiratory tract and multiplies in the nasopharyngeal epithelium and tonsillar tissue as well as in bronchial epithelium and alveolar macrophages. Then, the virus infects cells of the monocyte linages and CD3 T^+ lymphocytes transport the virus to the regional lymph nodes, where it undergoes a further cycle of replication before it enters into the blood stream[86,87,155,205]. Based on the initial phase of the pathogenesis of EAV, the viral quasispecies in nasal swab samples represent the initial population that established infection in the experimental stallions. Analysis of these viruses showed an expanded population size as a result of a high rate of *de novo* mutations. Very interestingly, despite an overall purifying type of selection, there was a strong diversifying pressure on the nsp9 encoding region of ORF1b (nt 6729 to 6816). This suggests that the low frequency mutation on nsp9 may have a regulatory effect on virus replication and may be associated with the higher population size in nasal swab viruses. Consistent with this observation, shrinkage in population size was detected in the buffy coat samples where there was no selection pressure on the indicated site (nt 6729 to 6816) of the nsp9 peptide.

It is also important to note that the virus population being shed by individual animals under the same sampling conditions, but from different cell types are distinct in their quasispecies composition, selection pattern and population size. The reduction in variant population size observed in buffy coat viruses may reflect a transmission bottleneck scenario that confers fitness to variants with the right type of mutations to infect cells of the monocyte lineages and CD3 T^+ lymphocytes.

Similarly, viruses in semen at 5 dpi show a divergent population structure with the appearance of several new fixations, reduced population size and multiple sites under

positive selection. Given that the neutralizing antibodies first appeared in blood of experimental stallions at 8 dpi, the driving force behind the selective sweep in the initial semen samples is most likely associated with the fitness cost of adapting to the cells in the stallion reproductive tract. EAV has been demonstrated in a variety of cells in the ampullae and other accessory sex glands of stallions[165]. Although these sites of virus persistence are not immunologically privileged, there is strong inflammatory cell infiltration along with plasma cells producing EAV specific antibodies[441]. These observations suggest that infections of various cells in the reproductive tract, as well as adaptation to evade local inflammatory and immune responses may be the driving force for the selection and reduction of the initial viral variant population in the reproductive tract.

It has been shown that persistent infection with EAV is testosterone dependent and castration of stallions leads to clearance of viral persistence[11,209,212]. How this hormone regulates EAV evolution in the reproductive tract, and whether the quasispecies population is any different in the presence of variable concentrations of this hormone, is a subject for future investigation. However, some studies have suggested testosterone is associated with variable effects on reproductive tissue such as maintenance of susceptible host cell population and immune suppression[11,13,209,209]

The study has also demonstrated the presence of continual positive selection in specific regions of the genome that drives evolution of viruses isolated from semen. The non-stochastic nature of EAV evolution in persistent infection and the swarm of quasispecies populations that shows an atypical high frequency of a minor variant population imply ongoing interaction between viral quasispecies populations and the host

immune system. Persistent infection with EAV develops in 10 to 70 % of sexually mature stallions[11,140,156]. This occurs in the presence of a high level of virus neutralizing antibodies in serum that effectively eliminates systemic infection in body tissues with the exception of the stallion reproductive tract. Previous studies have showed that natural transmission of EAV can occur by aerosol or infective semen, however the quasispecies diversity during persistent infection contrasts with the relative genetic stability observed in horizontally transmitted disease outbreaks[21,154,267]. There is some evidence from natural infection of lactate dehydrogenase-elevating virus and simian hemorrhagic fever virus to suggest variants with lower virulence and immunogenicity are selected during persistent infection, whereas more virulent variants seem to be favored during outbreaks[86,87,205,442].

In this study, ORF3 of viruses present in semen have been shown to be under selective pressure during persistent infection. The ORF3 may have some of the T cell epitopes of the virus but unfortunately, the T cell epitopes of this virus have not yet been characterized. The result is also true for viruses from cases of natural infection suggesting an important role for this particular gene in persistent EAV infection. Similar observations have been made in previous works that have characterized natural EAV infections. However, the observation showed a minimal role for ORF3 in the initial stages of infection when viruses first appear in the reproductive tract. The pattern of nonsynonymous substitutions at 5 and 9 dpi revealed that selection localizes in the ORF5 (GP5) and nsp2 peptide regions of ORF1a, suggesting a crucial role for these proteins in the initial adaptation process of variant populations in the stallion reproductive tract. The

selection of ORF3 variants comes later in the persistent infection period when maintenance of persistent infection is needed.

Acknowledgements

This study was supported by the Agriculture and Food Research Initiative competitive grant no.2013-68004-20360 from the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA). The next-generation sequencing using an Illumina Miseq technology was collaborated by Dr. Ganwu Li and Ms. Ying Zheng at the Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University. Dr. Zelalem Mekuria is gratefully acknowledged for the intellectual and creative input in the complex bioinformatics analysis of next-generation sequencing data for this study. Also, Drs. William McCollum and Thomas Swerczek are acknowledged for whom isolated the KY84 strain of equine arteritis virus, which has been used in this research project.

CHAPTER THREE

Genomic, Phylogenetic and Antigenic Characterization of a Novel Field Strain of Equine Arteritis Virus Isolated from a Feral Donkey in Chile

3.1. Summary

Here we describe the isolation and characterization of a new strain of equine arteritis virus (EAV donkey VD7634 strain) from the vas deferens of a wild male donkey from northern Chile in 2013. The full-length genome sequence of donkey VD7634 was obtained using next-generation sequencing (NSG) with Illumina MiSeq technology. Comparative nucleotide sequence analysis indicated that the donkey VD7634 virus had only 77% and 76.8% nucleotide identity with North American and European strains of EAV, respectively. The complete genome of the EAV donkey VD7634 strain was 12,703 bp in length and consists of ten open reading frames (ORFs) encoding 13 nonstructural proteins and eight structural proteins. Comparative nucleotide and amino acid analyses, as well as phylogenetic analysis demonstrated that this donkey strain is distinct from reference EAV strains due to multiple insertions and deletions. However, the virus appears antigenically very similar to the other EAV strains and could be neutralized with polyclonal and monoclonal antibodies against EAV.

3.2. Introduction

Equine arteritis virus (EAV) is the causative agent of equine viral arteritis (EVA), a respiratory and reproductive disease affecting members of the family *Equidae*, which include horses, donkeys, mules and zebras. EAV was first isolated from the lung of an

aborted fetus following extensive outbreak of respiratory disease and abortion on a Standardbred breeding farm near Bucyrus, OH, USA, in 1953. Equine arteritis virus (EAV) is a small, enveloped, positive-sense, single-stranded RNA virus that belongs to the family *Arteriviridae* (genus: *Equartevirus,* order: *Nidovirales*). The EAV genome is approximately 12.7 kb; it contains 10 open reading frames (ORFs) that are flanked by a 5′ leader sequence and a 3′ untranslated regions (UTR)[3,17,19,21,64,199]. ORF1a and 1b encode two polyproteins (pp1a and pp1ab) that are processed post-translationally by three ORF1a-encoded proteinases (nsp1, -2 and -4)[21,55,112]. These polyproteins are cleaved into 13 nonstructural proteins (nsp1-12, including nsp7α and 7β)[21,36,37,55]. The other eight ORFs (2a, 2b and 3, 4, 5a, 5, 6 and 7) encode the envelope proteins E, GP2, GP3, GP4, GP5, ORF5a protein and M and nucleocapsid protein, N, respectively[5,21]. The GP5 protein has been shown to carry the major neutralization determinants of the virus[16,20,21,82,443]. Based on ORF5 phylogenetic analysis, EAV field strains can be broadly divided into two distinct clades: North American (NA) and European (EU) groups[1,40,410,444]. Phylogenetic analysis based on partial ORF5 sequences segregated EAV isolates from around the world to North American (NA) and the European (EU) lineages and each lineage as further such divided into two North American (NA-1 and NA-2) and European (EU-1 and EU-2) clades (Table 3.3.); these sequences segregate into independent branch when subjected to phylogenetic analysis[10].

In 2013, Servicio Agricola y Ganadero (SAG) identified a large herd of wild donkeys naturally infected with EAV in northern Chile. Serum samples from these wild donkeys tested positive for anti-EAV antibodies by ELISA and virus neutralization test (VNT) assays at SAG. Subsequently, 18 necropsy specimens from two male donkeys

(animal ID: Donkeys 7634 and 7635) were collected and these included epididymis, prostate, vas deferens, seminal vesicles and conjunctival swabs. Viral nucleic acid was extracted from the vas deferens of Donkey 7634 and the partial sequence spanning ORFs 6 and 7 were obtained by sequencing the RT-PCR products. BLASTn analysis of ORFS 6 and 7 partial sequences showed 86 % and 90% nucleotide identity with EAV VBS (GenBank accession number: DQ846750), respectively. This confirmed that the donkey was infected with an EAV strain circulating in the wild donkey population. Samples from donkeys 7634 and 7635 were submitted to the EVA OIE reference laboratory at the Maxwell H. Gluck Equine Research Center, University of Kentucky, Lexington, KY, USA. Here we describe the virus isolation, genetic, phylogenetic and antigenic characterization of a novel donkey strain of EAV (EAV donkey VD7634) from a wild donkey in northern Chile.

3.3. Materials and methods

Necropsy specimens

Eighteen frozen necropsy specimens (epididymis [2], prostate [2], vas deferens [2], seminal vesicles [2], conjunctival swab [1] from two donkeys (7634 and 7635) were received at the EVA OIE reference laboratory, Maxwell H. Gluck Equine Research Center, University of Kentucky, Lexington, KY, USA for EAV testing.

Cells and viruses

High passage (HP) rabbit kidney 13 cells RK-13 (KY), passage level 399–409 derived from ATCC® CCL-37™, American Type Culture Collection, Manassas, VA,

USA were maintained in Eagle's minimum essential medium (EMEM, Cellgro®, Mediatech Inc., Herndon, VA, USA) with 10% ferritin-supplemented calf serum (HyClone Laboratories, Inc., Logan, UT, USA), penicillin and streptomycin (100 U/mL and 100 µg/mL) and 0.25 µg/mL of amphotericin B (Gibco®, Carlsbad, CA, USA). The prototype Bucyrus strain of EAV (EAV VBS; ATCC VR-796, Manassas, VA, USA) was used as a control virus in the neutralization assay.

Virus isolation

HP RK-13 (KY) cells were used for isolation of EAV from reproductive tract tissues and the conjuctival swabs, as previously described. Briefly, 10% tissue homogenates of necropsy specimens were prepared, centrifuged at 2,500 x *g* for 15 min, and the supernatant serially ten-fold diluted (10^{-1} to 10^{-3}). The conjunctival swab was immersed in 5 mL of EMEM and the medium filtered through a 0.45 µm filter. One mL of each tissue homogenate and conjunctival filtrate was inoculated in duplicate 25-cm^2 flasks containing confluent monolayers of HP RK-13 (KY) cells. Additionally, 1 mL of each undiluted tissue homogenate was inoculated in duplicate 25-cm^2 flasks. After 1 hour (h) adsorption at 37°C, monolayers were overlaid with EMEM supplemented with 0.75% carboxymethylcellulose, incubated for 4 days at 37°C and checked for the appearance of cytopathic effect (CPE). If there was no detectable CPE, a second blind passage was performed. After decanting the medium, monolayers were stained with a 1% crystal violet solution, and viral titers were expressed as plaque-forming units per gram of tissue (PFU/g). The tissue culture fluids (TCFs) were subjected to a one-way serum neutralization assay using polyclonal anti-EAV equine serum against the prototype VBS

strain of EAV to confirm the identity of the isolates as previously described[223]. The isolates were further confirmed by EAV specific TaqMan® real-time RT-PCR assay as previously described[232].

Microneutralization test

The neutralization phenotypes of the EAV donkey isolate was determined by a microneutralization assay using a panel of neutralizing monoclonal antibodies (mAbs; 5G11, 6D10, 7E5, 10F11, 1H7, 1H9, 5E8, 6A2, 7D4 and 10B4)[15–17,78] and polyclonal equine antisera against EAV VBS, commercial live-attenuated vaccine strain of EAV (MLV, ARVAC®; Zoetis, Kalamazoo, MI, USA), EAV KY84 strain and EAV CW01 strain were performed as previously described[15,78]. The non-neutralizing mAbs 12A4 (specific for EAV non-structural protein-1 [nsp-1]) and 3E2 (specific for N protein) and normal equine serum were used as negative controls. Briefly, duplicate serial two-fold dilutions of each equine antiserum sample from 1:4 to 1:8192 were made in supplemented EMEM containing 10% guinea pig complement (25 µl per well; Rockland Immunochemicals, Gilbertsville, PA, USA) and tested in 96-well plates. Ascitic fluid containing each mAb was diluted from 1:16 to 1:32,384 in unsupplemented EMEM without 10% guinea pig complement as previously described[15–17,78]. An equal volume containing 200 $TCID_{50}$ of the EAV donkey VD7634 strain was added to each well. Plates were incubated at 37°C in 5% CO_2 for 1 h. After incubation, a suspension of HP RK-13 (KY) cells (125 µl per well) was added to each well, and the plates were incubated for 72 hours at 37°C until CPE had fully developed in the virus control wells.

The titer of a sample was recorded as the reciprocal of the highest serum dilution that provided at least 50% protection of the monolayer.

Real-time RT-qPCR

The nucleic acids extracted from the reproductive tract tissue samples (10% tissue suspensions), TCFs and conjunctival swabs were tested by an EAV specific TaqMan® real-time RT-PCR assay as previously described [232]. Briefly, viral RNA was directly isolated from 50 µl of 10% tissue suspension or TCF using a commercial RNA isolation kit (MagMAX™ -96 Viral RNA Isolation Kit, Thermo Fisher, Waltham, MA, USA) according to the manufacturer's instructions. Viral RNA was eluted in 50 µl of nuclease-free water and stored at -80°C. RNA extracted from TCF derived from HP RK-13 (KY) cells inoculated with semen samples that were negative for EAV, as well as from nuclease-free water were included as negative controls. Viral RNA extracted from TCF containing the EAV VBS was used as a positive control. A one-tube TaqMan® real-time RT-PCR assay was performed as described in Miszczak *et al*. (2011)[223].

Sanger sequencing

Viral RNA was extracted from the TCF containing EAV donkey D7634 strain using MagMax™-96 viral RNA Isolation Kit (Thermo Fisher, Waltham, MA, USA). The sequences of the PCR products were determined by the Sanger method. The extracted viral RNA was first reverse-transcribed into cDNA using SuperScript™ III Reverse Transcriptase (Thermo Fisher, Waltham, MA, USA) with gene specific primer (BN12707N) following the manufacturer's instruction, and amplified by Qiagen

HotStarTaq® DNA polymerase using forward and reverse primers described in Supplementary Table 3.1. Fourteen new primers were designed based on published EAV donkey sequences from South Africa[405]. In addition, fifteen EAV specific primers were used for primer walking. ORFs 2-7 sequences including the flanking 5' ORF1b and 3'-UTR (2912 bp) were determined. A similar approach was used to amplify the partial ORF1ab regions (1-1087 [1087 bp], 1592-2093 [502 bp], 4152-4721 [570 bp] and 6436-6923 [488 bp] encoding the nsp1, partial nsp2, partial nsp5, nsp6, partial nsp7 and partial nsp9, respectively. The PCR products were gel-purified using a high pure PCR product purification kit (Roche). Both sense and antisense strands were sequenced. About 5559 base pairs (43.75%) of the entire genome were sequenced by Sanger methods. Then, Illumina MiSeq next-generation sequencing technology was used to obtain the complete genome sequence of EAV donkey D7634 strain.

Next-generation sequencing

Total RNA was isolated from HP RK-13 (KY) cells infected with the EAV donkey VD7634 strain using a total RNA isolation kit (Norgen Biotek Corp., Niagara-on-the-Lake, Ontario, Canada). The RNA was quantified with a Qubit 2.0 spectrophotometer (Life Technologies, Carlsbad, CA, USA) and purified RNA was stored at -80°C. The complete genome sequence of EAV donkey VD7634 strain was determined by using next-generation sequencing (NGS) technology on an Illumina MiSeq platform according to previously established procedures[445]. Briefly, the cDNA libraries were constructed from 100 ng of total RNA using a TruSeq Stranded total RNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions and

previously described protocol . Multiplex libraries were prepared using barcoded primers and a median insert size of 340 base pairs (bp). Libraries were analyzed for size distribution using a Bioanalyzer and quantified by quantitative RT-PCR using a Kapa library quantification kit (Kapa Biosystems, Boston, MA, USA); relative volumes were pooled accordingly. The pooled libraries were sequenced on an Illumina MiSeq platform with 150 bp end-reads following standard Illumina protocols. Sequences were mapped to all known EAV strains and mapped read-sets were used for *de novo* assembly using ABySS, v.1.3.7 (BC Cancer Agency, Vancouver, Canada) and Geneious 7.0.6 software (Biomatters Ltd., Auckland, New Zealand). Sequences obtained by NGS technique were mapped to all available published genome sequences of the family *Arteriviridae* (currently 550 genomes are available at http://www.ncbi.nlm.nih.gov/) using BWA-MWM (v0.7.5a). Sequence ID of mapped reads were extracted using SAMtools (v0.1.19).

Sequence analysis

The new paired-end read sets were extracted from the original paired-end fastq files using seqtk with the list of mapped ID to insure both reads of one pair were selected. The raw reads in FASTA formats were imported into IGV_2.3.79 for mapping and paired using the default parameter for Illumina short read. Comparative nucleotide (nt) and amino acid (aa) sequence alignment, maximum likelihood (ML) phylogenetic reconstruction were performed using Geneious v.7.0.6 (Biomatters Ltd., Auckland, New Zealand) and MEGA v.7.0.26 (http://www.megasoftware.net).

Phylogenetic analysis

Phylogenetic trees were constructed by using the maximum likelihood method based on the Tamura-Nei model[446], applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) distance with Gamma-distributed rates, pairwise gap deletion and boopstrap resampling (1,000 replications)[447].

3.4. Results

Virus isolation and RT-qPCR testing of the reproductive tissues from donkeys

EAV was isolated from the epididymis, prostate, vas deferens and seminal vesicles of donkey number 7634, but not from the tissues of the donkey number 7635 (Supplementary Table 3.2.). The viral isolates were further confirmed by TaqMan® real-time RT-qPCR specific for EAV ORF7. These data unequivocally confirmed that virus isolates from epididymis, prostate, vas deferens and seminal vesicles of VD7634 as EAV. Viral RNA extracted from 10% tissue homogenates were also tested directly by EAV-specific RT-qPCR. Of the nine samples tested from donkey 7634 (Supplementary Table 3.2.), only three samples gave positive results with EAV RT-qPCR (epididymis, vas deferens and seminal vesicles). All the necropsy samples from donkey 7635 were negative for EAV nucleic acid by RT-qPCR. This further confirmed that the donkey 7635 was not persistently infected with EAV.

Analysis of genomic sequence of EAV donkey VD7634 strain

The full-length genome sequence of the EAV donkey VD7634 strain isolated from the vas deferens of donkey 7634 was determined by NGS. The complete genome of EAV donkey VD7634 strain was 12,703 bp in length with ten ORFs which encode 13 nonstructural proteins (nsp) and eight structural proteins (sp) common to all known EAV strains. The annotation of EAV donkey VD7634 strain was predicted based on NCBI-referenced EAV sequence (EAV Utr; accession numbers: NC002532), then compared to both NCBI-referenced EAV sequence and our prototype virulent Bucyrus virus strain (EAV VBS; accession numbers: DQ846750) to determine its general genome features[1]. The sequence of EAV donkey VD7634 strain was deposited in GenBank under the accession number MF598091.

The complete genome sequence of EAV donkey VD7634 strain was 1nt shorter in length and had 75.6 % nucleotide identity to both EAV VBS and EAV Utr (GenBank Accession numbers DQ846750 and NC002532, respectively). The nucleotide and amino acid differences in ORF1a and ORF1b are shown in Table 3.1. The nsp2, nsp3 and nsp12 of donkey VD7634 had 1-3 amino acid deletions as compared to the EAV VBS and EAV Utr. The leader sequence of EAV donkey VD7634 strain is 211 nucleotides, and its 5´ UTR is 6 nt longer than the other two EAV strains. The ORF1a and ORF1ab start codon of EAV donkey VD7634 strain is located at nucleotide position 231.

Table 3.1. Comparative nucleotide and amino acid sequence analysis of donkey VD7634 strain of EAV and published EAV sequences (EAV Utr [NC002532] and EAV VBS [DQ846750]

ORF	Protein	EAV donkey VD7634 strain (MF598091) nucleotide sequence (Start…End)	EAV donkey VD7634 strain (MF598091) nucleotide position (nt)	EAV donkey VD7634 strain protein length (aa)	EAV Utr (NC002532) nucleotide sequence (Start...End)	EAV Utr (NC002532) nucleotide position (nt)	EAV Utr (NC002532) protein length (aa)	EAV VBS (DQ846750) nucleotide sequence (Start...End)	EAV VBS (DQ846750) nucleotide position (nt)	EAV VBS (DQ846750) protein length (aa)
5' TTR	-		1-230	-		1-224	-		-	-
Leader sequence			1-211	-		1-211	-		-	-
ORF1a	1a polyprotein	ATGGCAACCT…GTTAAACTGA	231-5405	1724	ATGGCAACCT…GTTAAACTGA	225-5408	1727	ATGGCAACCT…AGTGAATCAG	225-5399	1725
ORF1b	-	CTGAGAGCGC…ACCCGTGTGA	5402-9748	1448	CTGAGAGCGC…GCCCGTGTGA	5405-9751	1448	CTGAGAGCGC…GCCCGTGTGA	5405-9751	1448
ORF1ab	1ab polyprotein	ATGGCAACCT…ACCCGTGTGA	231-9748	3172	ATGGCAACCT…GCCCGTGTGA	225-9751	3175	ATGGCAACCT…GCCCGTGTGA	225-9751	3175
	nsp1	ATGGCAACCT…GAACTACGGT	231-1010	260	ATGGCAACCT…CAACTACGGC	225-1004	260	ATGGCAACCT…CAACTACGGC	225-1004	260
	nsp2	GGGTACAACC…GTTGATTGGA	1011-2714	568	GGCTACAATC…GCTGATAGGT	1005-2717	571	GGCTACAATC…GCTGATAGGT	1005-2717	571
	nsp3	GGGTGGACCT…TGGTGTTTGA	2715-3412	232	GGATGGATTT…GGTGTTTGAA	2718-3416	233	GGATGGATTT…GGTGTTTGAA	2718-3416	233
	nsp4	GGGCTTGTTC…CAACAGAGAG	3413-4025	204	GGGCTATTCA…CAATAGAGAG	3417-4028	204	GGGCTATTCA…CAATAGAGAG	3417-4028	204
	nsp5	AGTAGCCTTT…CTTCCTTGAG	4026-4511	162	AGCAGCCTTT…CTTCCTGGAG	4029-4514	162	AGCAGCCTTT…CTTCCTGGAG	4029-4514	162
	nsp6	GGAGGTGTGA…CACTCAAGAA	4512-4577	22	GGAGGAGTGA…TACCCAGGAG	4515-4580	22	GGAGGAGTGA…TACCCAGGAG	4515-4580	22
	nsp7	AGTTTGACCG…CAGTTATGAA	4578-5252	225	AGTCTCACTG…GAGCTATGAA	4581-5255	225	AGTCTCACTG…GAGCTATGAA	4581-5255	225
	nsp8	GGGTTAGACA…CCAGTTAAAC	5253-5402	50	GGCCTAGATC…TCAGTTAAAC	5256-5405	50	GGCCTAGATC…TCAGTTAAAC	5256-5405	50
	nsp9	GGGTTAGACA…GCAGTACGAG	5253-7330	693	GGCCTAGATC…GCAGTATGAG	5256-7333	693	GGCCTAGATC…GCAGTATGAG	5256-7333	693
	nsp10	GCGGCTGTAT…GGAAAAACAG	7331-8731	467	AGTGCCGTGT…GGAAAAGCAA	7334-8734	467	AGTGCCGTGT…GGAAAAGCAA	7334-8734	467
	nsp11	TCAAACAAAG…TGTCCAGGAG	8732-9388	219	TCCAACAAAA…TGTCCAAGAG	8735-9391	219	TCCAACAAAA…TGTCCAAGAG	9835-9391	219
	nsp12	GGGATTGATG…AATTGGACCC	9389-9742	118	GGTGTTGATG…CATTGGGCCC	9392-9745	118	GGTGTTGATG…TGGGCCCGTG	9392-9748	119
	Slippary sequence	GTTAAAC	5398-5404	-	GTTAAAC	5399-5405	-	GTTAAAC	5399-5405	-
ORF2a	E	ATGGGCTTAG…ACCTCTTTAA	9748-9951	67	ATGGGCTTAG…ACCTGTTTAA	9751-9954	67	ATGGGCTTAG…ACCTGTTTAA	9751-9954	67
ORF2b	GP2	ATGCTGTGCT…GATTTTGTAG	9821-10504	277	ATGCAGCGCT…GATTTTGTAG	9824-10507	227	ATGCAGCGCT…GATTTTGTAG	9824-10507	227
ORF3	GP3	ATGGATTATG…AAGCTCGTAA	10303-10794	163	ATGGGTCGTG…AGGCTCGTAA	10306-10797	163	ATGGGTCGTG…AGGCTCGTAA	10306-10797	163
ORF4	GP4	ATGAAGAACT…TTATCTATGA	10697-11155	152	ATGAAGATCT…TTATCTATGA	10700-11158	152	ATGAAGATCT…TTATCTATGA	10700-11158	152
ORF5a	ORF5a	ATGTTCTTCT…TGCGCTATGA	11109-11288	59	ATGTTCTTTT…TGCGATATGA	11112-11291	59	ATGTTCTTTT…TGCGATATGA	11112-11291	59
ORF5	GP5	ATGTTATCTA…GGAGCCATAG	11143-11910	255	ATGTTATCTA…GGAGCCATAG	11146-11913	255	ATGTTATCTA…GGAGCCATAG	11146-11913	255
ORF6	M	ATGGGAGCCA…GTTACAATGA	11898-12386	162	ATGGGAGCCA…GCTACAATGA	11901-12389	162	ATGGGAGCCA…GCTACAATGA	11901-12389	162
ORF7	N	ATGGCGTCAA…AGGGCCGTAA	12310-12642	110	ATGGCGTCAA…AGGGCCGTAA	12313-12645	110	ATGGCGTCAA…AGGGCCGTAA	12313-12645	110
3' TTR	-	GACGTGGATA…AAGTATCCAG	12643-12703	-	GACGTGGATA…CCCAGGAACC	12646-12704	-	GACGTGGATA…CCCAGGAACC	12646-12704	-

Characterization of nucleotide hallmarks of EAV donkey VD7634 strain as compared with other EAV strains

Comparative nucleotide sequence analysis of the full-length genome of VD7634 sequence and twenty-nine EAV strains in GenBank determined unique insertions and deletions (Table 3.2.). These included four North American strains and twenty-five European strains (EU-1 [CW and S numbered strains] and EU-2 [French strains]) that were selected by BLASTn analysis. The nucleotide identity with North American and European strains varied from 77% and 76.8%, respectively. Analysis identified a number of unique insertions and deletions that were present in 5´ UTR, 3´ UTR, ORF1a (nsp2 coding region) and ORF3 of EAV donkey VD7634 strain (Table 3.2.).

Table 3.2. Insertion and deletion positions in full-length genome alignments of EAV isolates from North America and Europe

ORF	EAV donkey VD7634 strain nucleotide position	Nucleotide substitution length (bp)	Nucleotide substitution	Description
5' UTR	6	1	C	Insertion
	26	1	G	Insertion
	88-89	2	GT	Consecutive nucleotide insertions
	123	1	T	Insertion, similar to F27, F60, F61, F62, F63 strains
	124	-1	T	Deletion, except F27, F60, F61, F62, F63 strains
	127	-1	G	Deletion
	164	1	C	Insertion, similar to CW01, CW96, F27, F60, F61, F62, F63 strains
	224	1	T	Insertion
nsp2	1399-1401	-3	AAC*	Deletion
	1460-1462	3	AAC	Insertion, similar to CW01, CW96, F60, F61, F62, F63, F27 strains
	1478-1480	3	AGA	Insertion
	1515-1517	3	TTG	Insertion
	1565-1579	- 15	CGTCCACAGTGGTCT*	Insertion except EAV donkey VD7634 strain, CW01, CW96, F60, F61, F62, F63, F27, HK25, HK116, VBS, ARVAC strains
	1621-1635	- 15	GGCGTGTGGCCAAAA*	Deletion
GP3	10,653-10,664	- 12	TGCGAAACCACT	Insertion in F61, F62, F63 strains
	10658-106659	2	AG	Insertion
	10,684-10,695	- 12	AGTGCATTTGGA*	Insertion except EAV donkey VD7634 strain, CW01, CW96, F60, F61, F62, F63, F27, HK25, HK116, VBS, ARVAC strains
	10,669 -10,670	-2	GG	Deletion
3' UTR	12685	1	G	Insertion
	12703	1	C	Insertion

*Most common deletion sequence is shown out of three different deletion sequences in the nucleotide comparison among EAV referenced sequences.

The 5´-leader sequence of the EAV donkey VD7634 strain had three single nucleotide insertions (C-6, G-26 and T-224), one single nucleotide deletion (G-127 based on the EAV VBS sequence) and two consecutive nucleotide insertions (G-88 and T-89). In addition, there was a single nucleotide insertion at T-123 of the EAV donkey VD7634 strain sequence that is similar to some of the French strains (F27, F60-F63). There was one insertion followed by one consecutive single nucleotide deletion compared to the earlier described French strains. Finally, there was a single nucleotide insertion in the 5´ leader sequence at position C-164, which is similar to some of the European strains

(CW96, CW01; F27, F60-F63). The donkey EAV strain had two single nucleotide insertions in the 3´ UTR (G-12,685 and C-12,703).

Interestingly, there was a consecutive 3-nucleotide deletion in the nsp2 of EAV donkey VD7634 strain at nucleotide positions 1399-1401 based on the EAV VBS sequence position numbering (AGC [CW96, CW96, HK25, HK116, Bucyrus and ARVAC] or CGC [F27 and F60-F63] or AAC [S-3685, S-3685, S-3699, S-3711, S-3712, S-3817, S-3854, S-3861, S-3886, S-3943, S-3961, S-4007, S-4216, S-4417, S-4421, S-4227, S-4333 and S-4445]). Then, there were three consecutive nucleotide insertions encoding region of ORF1a at three positions (A-1460, A1461 and C-1462; A-1478, G-1479 and A-1480; and T-1515, T-1516 and G-1517). Seven other EAV strains also had three consecutive nucleotide insertions at nucleotide positions 1460-1462 (CGT [CW96 and CW01] GGG [F27, F60-F63]). The consecutive 15-nucleotide insertion in nsp2 encoding region of ORF1a was presented by S-numbered EAV strains in nucleotide positions at 1565-1579; CGTCCACAGTGGTCT [S-3685, S-3699, S-3711, S-3712, S-3854, S-3861, S-3886, S-3943, S-3961, S-4007, S-4227, S-4333, S-4421 and S-4445] or CGTCCACAGTGGTCC [S-3685, S-3817 and S-4216] or CATCCACAGTGGTCT [S-4417]. Based on the EAV VBS sequence position numbering, EAV donkey VD7634 strain had a consecutive 15-nucleotide deletion at 1621-1635; GGCGCGTGACTAAAA [CW01 and CW96] or GGCGCGTGACAAAGA [F27 and F60-63] or GGCGTGTAGCCAAAA [S-4417] or GGCGTGTGGCCAAAA [HK25, HK116, Bucyrus, ARVAC, S-3685, S-3685, S-3699, S-3711, S-3712, S-3817, S-3854, S-3861, S-3886, S-3943, S-3961, S-4007, S-4216, S-4421, S-4227, S-4333 and S-4445].

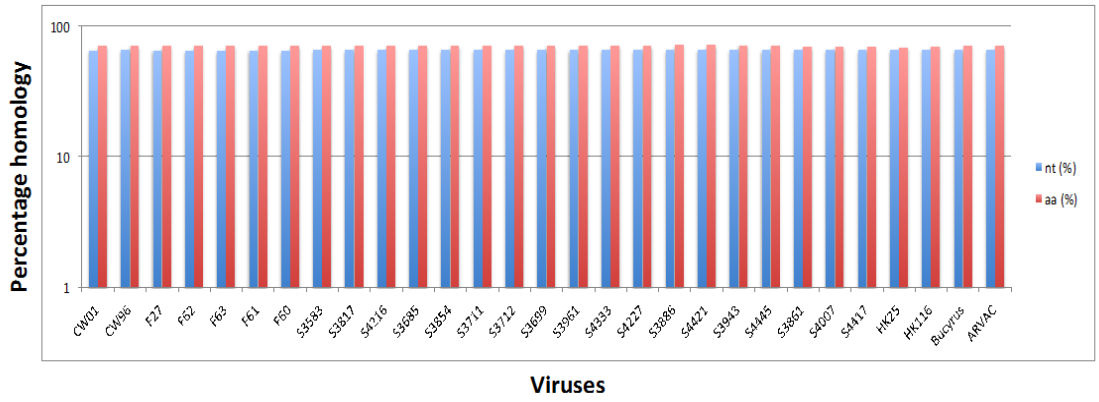
Two consecutive 12-nucletide insertions (10,653-10,664; TGCGAAACCACT [F61] or TGCGAAACCAYT [F62] or TGCRAAACCAYW [F63]; and 10,684-10,695; AGTGCATTTGGA [S-3583, S-3817, S-3685, S-3699, S-3711, S-3712, S-3854, S-3869, S-3943, S-40007, S-4417, S-4445, S-3961 and S-4216] or TAGTGTATTTGGA [S-3886, S-4333 and S-4421] or AGTGCACTTGGA [S-4227]) in some of the European strains were not present in the ORF3 of EAV donkey VD7634 strain and the rest of the EAV sequences available in GenBank. ORF3 of EAV donkey VD7634 strain had two unique nucleotide insertions (A-10,658 and G-10,659) and two unique nucleotide deletions (G-10,669 and G-10,670; according to EAV VBS sequence position numbering).

Additionally, the percentage identity (%) of nucleotide and amino acid sequence alignment among EAV donkey VD7634 strain and other EAV sequences from GenBank (n=29) were calculated and used to create bar graphs (Figure 3.1. and supplementary table 3.5.). The average nucleotide and amino acid identity are 65.55% and 70.72% in ORF1a (nsp2 coding region) and nsp2, respectively (Figure 3.1.A.). Similarly, the average nucleotide identity and amino acid identity were 78.18% and 74.21% in ORF3 and GP3, respectively (Figure 3.1.B.). In nucleotide similarity, ARVAC, HK116, HK25 and Bucyrus NA strains showed 80.8% to 81.4% highest nucleotide identity in ORF3, as well as 77% to 78.2% of amino acid identity in GP3. EAV isolates CW01, CW96, F27 and F60 strains showed second highest nucleotide identity from 78.1% to 79.6% in ORF3 and 72.7% to 73.9% amino acid identity in GP3.

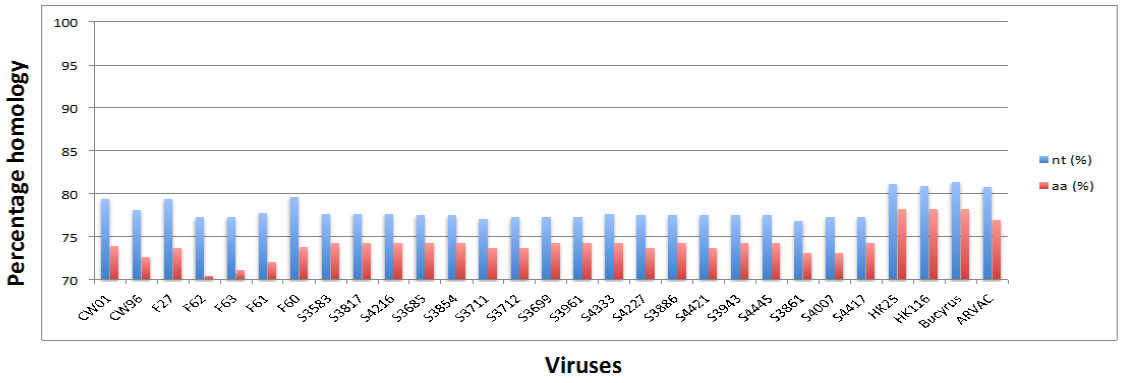
EAV donkey VD7634 strain showed 81.3% to 81.5% and 93% nucleotide identity and amino acid identity to partial ORF5 and GP5 gene from donkey isolates from Republic of South Africa (RSA), respectively (Figure 3.1.C.). Even though two of the

EAV donkey referenced sequences isolated from Italy presented low nucleotide identity from 74.3% to 74.7%, respectively, these two sequences showed 86.6% amino acid identity, which is the second highest similarity in partial GP5. The other EAV sequences had 75.54% nucleotide identity and 86.34% amino acid identity in partial GP5 compared to the EAV donkey VD7634 strain. There were no insertions or deletions in partial ORF5 sequences analyzed from donkeys.

(A) nsp2



(B) GP3



(C) Partial GP5

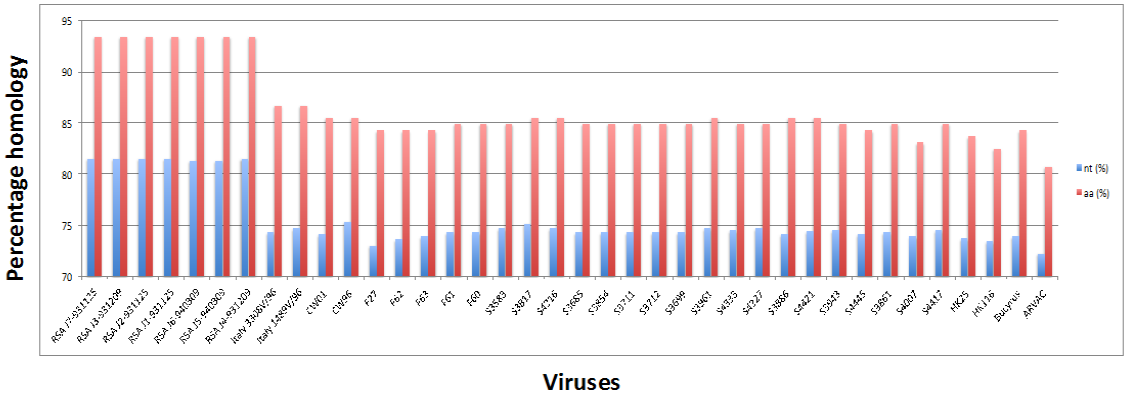


Figure 3. 1. Based on percentage homology among EAV 29 EAV strains; nucleotide sequence identity (%) and amino acid sequence identity (%) analyses were used to create bar graphs. (A) nsp2, (B) GP3 and (C) partial GP5. Blue bar indicates nucleotide identity (%) and red bar indicates amino acid identity (%).

Comparison of partial ORF5 sequence of EAV donkey VD7634 strain and other EAV donkey strains

Nine partial donkey EAV ORF5 sequences (nt 11,295-11,812 [518 bp in length][10] available in GenBank were compared to the EAV donkey VD7634 strain (Supplementary Table 3.3.). The data showed that the EAV donkey VD7634 strain shared variable nucleotide and amino acid identity with other EAV donkey strains: 74.3% to 81.5% nucleotide identity and 86.7% to 93.4% amino acid identity. Phylogenetic analysis was carried out on partial ORF5 sequences (Figure 3.2.). The maximum likelihood phylogenetic tree was clearly divided into two lineage groups: Republic of South Africa (RSA) and Italy.

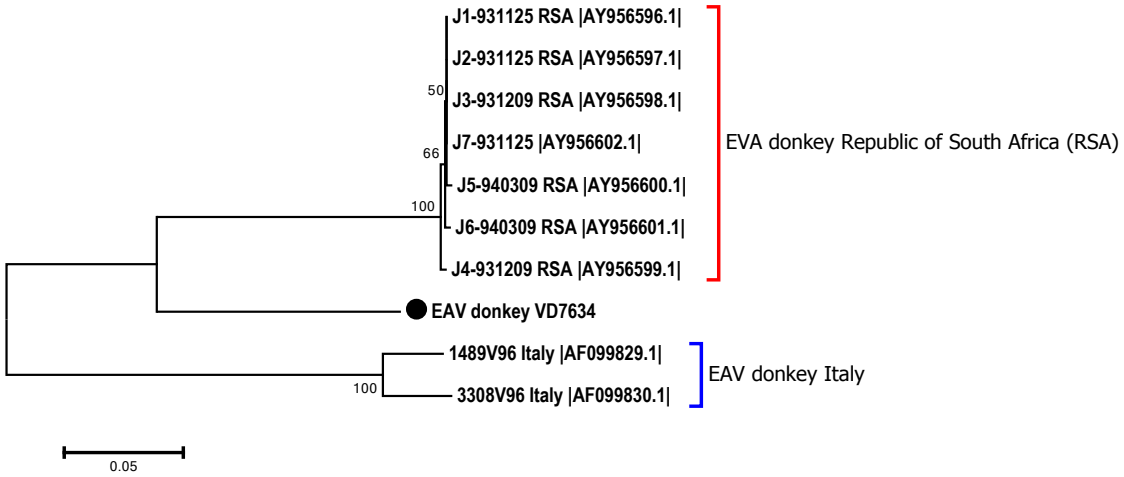


Figure 3.2. Phylogenetic tree based on the partial ORF5 of EAV donkey VD7634 strain and nine EAV donkey reference strains. Phylogenetic trees constructed based on nucleotide sequences of partial ORF5 (nt 11,295-11,812 [518 bp]) of EAV donkey VD7634 strain using the maximum likelihood method implemented in MEGA7.0. Bootstrap values for 1,000 replicates are shown at the nodes. The bar indicates genetic distance.

Both sequence identity and phylogenetic tree indicate that EAV donkey VD7634 strain was more closely related to Republic of South African (RSA) strains than the Italian strains. According to Stadejek *et al*. (2006), the RSA EAV donkey strains sequenced from the semen of naturally infected donkeys were different from reference sequences obtained from viruses in horses in North American and European groups[10]. In contrast, ORF5 sequences of Italian EAV donkey strains were clustered in European groups, indicating that these donkeys had probably become infected from horses[227,412].

Phylogenetic analysis of EAV donkey VD7634 strain

Twenty-nine EAV full-length genome sequences from GenBank were downloaded and used in phylogenetic analysis (Figure 3.3.). These sequences included EAV strains from both North American (NA) and European (EU) lineages. In addition, partial ORF5 sequences from EAV and donkey isolates from South Africa and Italy were included in the analysis. Phylogenetic analysis was conducted using maximum likelihood based on the Tamura-Nej model and neighbor-joining methods[446]. As the result, EAV donkey VD7634 strain was presented as more diverged than other referenced EAV strains.

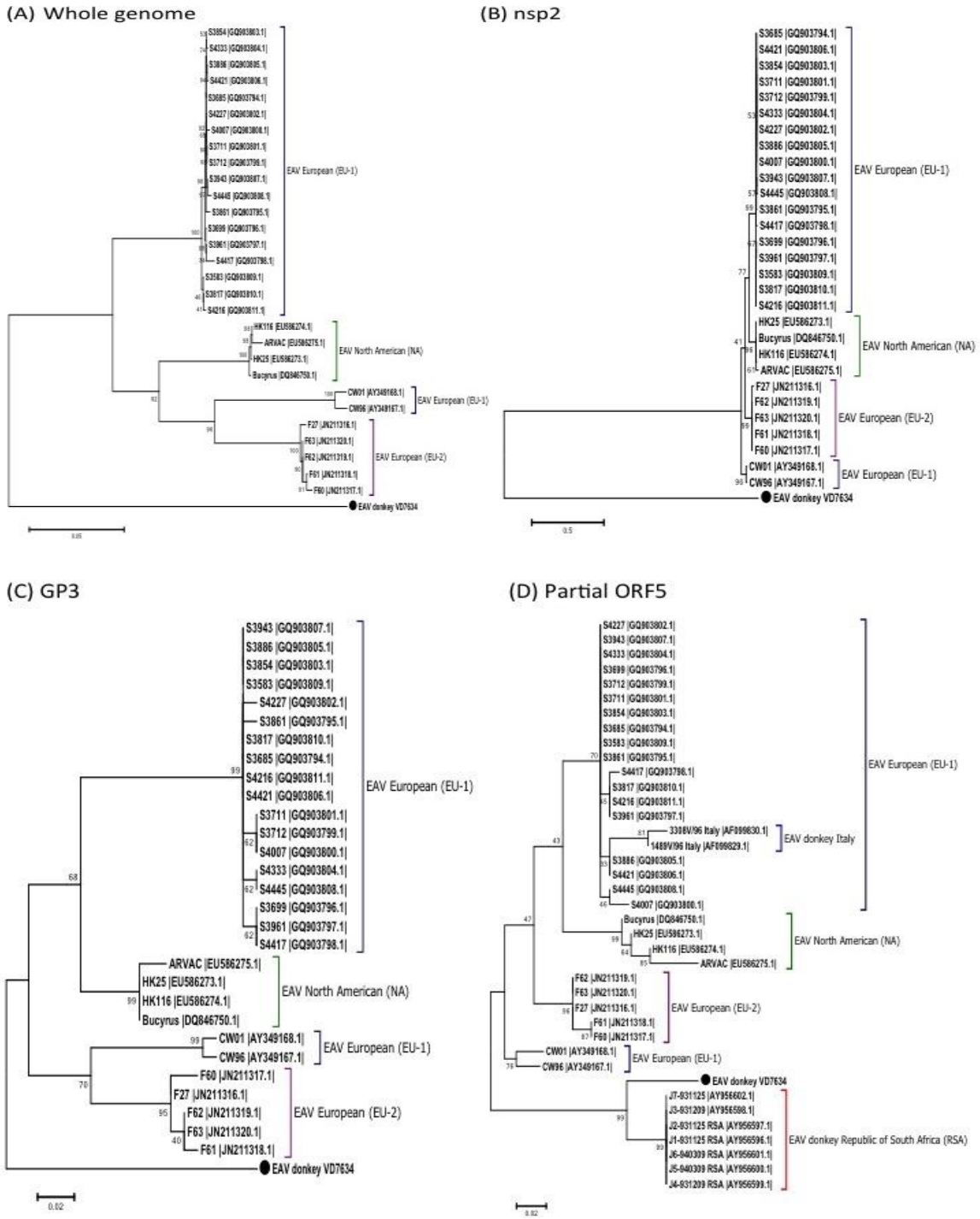


Figure 3.3. Phylogenetic analyses of EAV were inferred based on primary nucleotide alignment. (A) Whole genome of 29 referenced strains and EAV donkey VD7634 strain and amino acid alignment of (B) ORF1a, nsp2 encoding region, (C) ORF3 and (D) Partial GP5 includes nine additional strains of EAV donkey reference strains.

Phylogeny based on full-length genome

Based on BLASTn analysis of GenBank database, twenty-nine EAV whole genome sequences were selected to perform sequence alignment and phylogenetic analysis (Figure 3.3.A.). The EAV donkey VD7634 strain is 12,703 nucleotides in length, slightly shorter than EAV CW strains (12,708 nucleotides; EU-1), EAV French strains (12,710 to 12722 nucleotides; EU-2), EAV North American strains (12704 nucleotides; NA) and EAV S numbered strains (12,731 nucleotides; EU-1). The EAV donkey VD7634 strain shares average homology of 75.37% at the nucleotide level among 29 full genomes of EAV referenced sequences (Supplementary table 3.4.). Analyzing homology at the amino acid level for specific ORFs, least similarity was noted across ORF1a and ORF3 which encode nsp2 and GP3 at 70.72% and 74.21%, respectively. These proteins were previously described with nucleotide hallmarks of unique insertions and deletions. The greatest similarity at the amino acid level was ORF7, which encode nucleocapsid (N) protein at 95.67% among 29 full genomes of EAV referenced sequences (Supplementary table 3.4.). Phylogenetic analysis did not cluster EAV donkey VD7634 with the other EAV full-length genome sequences and segregated it into an independent lineage.

Phylogeny based on ORF1a nsp2 coding region and ORF3

EAV donkey VD7634 strain has unique genetic hallmarks in nsp2 and ORF3, thus these protein sequences at the amino acid level were analyzed phylogenetically to compare among each genome of 29 EAV referenced sequences (n=29) (Figure 3.3. B and C). In these analyses, EAV donkey VD7634 strain segregated into an independent lineage similar to the full-length genome sequences. (Figure 3.3.A.)

Phylogeny based on ORF5 partial sequences including other EAV donkey strains

Phylogenetic analysis of the EAV partial ORF5 sequences including other EAV donkey sequences from Italy (n=2) and Republic of South Africa (n=7) was carried out (Figure 3.3.D.). Interestingly, EAV donkey VD7634 strain was clustered with EAV donkey strains from the Republic of South Africa, and this clustering was supported by a bootstrap value of 99%. It suggests that EAV donkey VD7634 strain was most closely related to EAV donkey strains that originated in the Republic of South Africa. Not surprisingly, partial ORF5 sequences of Italian EAV donkey strains were clustered in the European group (EU-1), but were distantly related to other EU-1 virus isolates in this phylogenetic subgroup.

Neutralization phenotype of EAV donkey VD7634 strain

A panel of 12 monoclonal antibodies (mAbs) and a panel of 6 of polyclonal equine antisera were used to characterize the neutralization phenotypes of EAV donkey VD7634 strain (Table 3.3.). EAV donkey VD7634 strain was neutralized by mAbs 5G11, 6D10, 7E5, 10F11 and 6A2. Especially, mAbs 6D10, 7E5 and 6A2 neutralized the donkey virus to a higher titer. In contrast, mAbs 1H7, 1H9, 5E8, 7D4 and 10B4 did not neutralize EAV donkey VD7634 strain. Only anti-KY84 and anti-CW96 polyclonal equine antisera were able to neutralize the EAV donkey VD7634 strain to a higher titer. Interestingly, anti-ARVAC and anti-VBS polyclonal equine sera did not neutralize donkey VD7634 to a high titer. Furthermore, anti-EAV NVSL sera failed to neutralize the donkey strain.

Table 3.3. Serum neutralization test titers of mAbs and polyclonal equine antisera against EAV donkey VD7634 strain and other EAV isolates

Antibody	Protein specificity/epitope location (amino acid no. and site)	Virus	Virus of SNT result from Balasuriya et al. (2004)[a]				Virus SNT result from Zhang et al. (2010)[b]								
		EAV donkey VD7634 strain	CW96[a]	CW01[a]	EAV 030[a]	ATCC[a]	ARVAC[b]	S3583[b]	S4216[b]	S3685[b]	S3861[b]	S3854[b]	S4333[b]	S3943[b]	S4445[b]
mAbs															
5G11	GP5/102 and 104 [D]	512	512	4096	4096	>4096	<32	1024	512	2048	2048	1024	2048	1024	1024
6D10	GP5/99 [D]	4096	4096	>4096	>4096	>4096	512	4096	2048	2048	4096	4096	>4096	4096	4096
7E5	GP5/102 [D]	**512**	**32**	256	256	256	**32**	128	128	512	128	512	128	256	256
10F11	GP5/ND [D]	4096	4096	>4096	>4096	>4096	2048	2048	1024	4096	4096	4096	4096	4096	4096
1H7	GP5/49 [A]	16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
1H9	GP5/69 [C]	<16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
5E8	GP5/61 [B]	<16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
6A2	GP5/69 [C]	**2048**	**128**	**<32**	**<32**	**<32**	**<32**	**256**	**64**	1024	**256**	2048	512	1024	**32**
7D4	GP5/69 [C]	<16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
10B4	GP5/62-101 deleted [C and D]	32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
Control mAbs															
3E2*	N	16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
12A4*	nsp1	<16	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32	<32
Polyclonal equine sera															
anti-ARVAC	NA	**16**	<8	8	128	8	**256**	16	64	32	32	32	16	8	16
anti-VBS	NA	64	64	16	512	512	256	128	128	128	128	128	64	128	128
anti-KY84	NA	**1024**	512	**32**	>1024	>1024	**64**	256	**64**	**128**	512	**128**	**64**	512	**128**
anti-CW96	NA	1024	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Control polyclonal equine sera															
anti-EAV NVSL positive	NA	16	32	16	128	64	128	64	16	32	32	32	32	32	16
anti-EAV NVSL negative	NA	<8	<8	<8	<8	<8	<8	<8	<8	<8	<8	<8	<8	<8	<8

Neutralization titers are expressed as the inverse of the antibody dilution providing 50% protection of RK-13 cell monolayers against 200 $TCID_{50}$ of virus. Significant differences in neutralization titers (four-fold or greater) are indicated by bold numbers. ND, not determined. NA, not applicable. EAV donkey VD7434 strain SNT result was compared to previously published data.

[*]mAb to the nucleocapsid (N) protein and nonstructural protein 1 (nsp1) of EAV was used as negative control

[a] Viruses of neutralization test result titers were from Balasuriya *et al*. (2004)[17]

[b] Viruses of neutralization test result titers were from Zhang *et al*. (2010)[21]

Previously published data described neutralization phenotypes of the same panel of mAbs and polyclonal equine antisera (exception of anti-CW96 polyclonal equine sera) against different EAV strains: CW96, CW01, EAV030, ATCC, ARVAC, S3583, S4216, S3685, S3861, S3854, S4333 and S3943[17,21]. Interestingly, 6A2 showed significant variability in its ability to neutralize different EAV isolates, but the same mAb neutralized the EAV donkey VD7634 strain to a higher neutralization titers[17,21]. In summary, these neutralization data suggest that EAV donkey VD7634 strain has a distinct neutralization phenotype. Comparative amino acid sequence analysis of the critical neutralization regions of the GP5 protein of this isolate are shown in Figure 3.4.

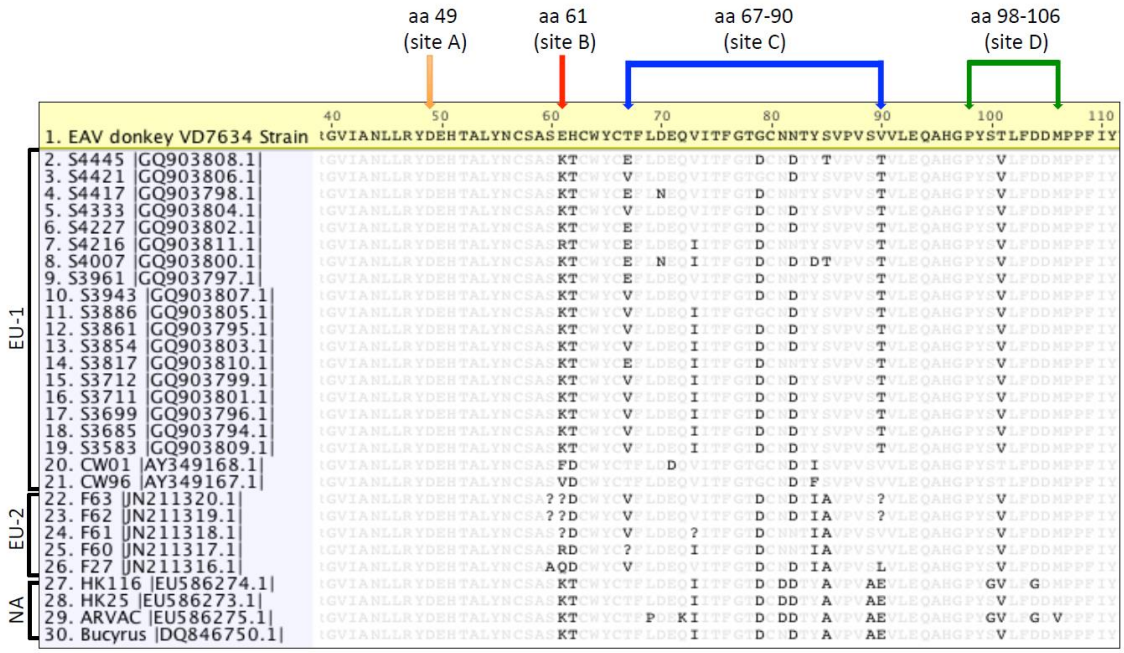


Figure 3.4. Comparative amino acid substitution of ORF5 of EAV donkey VD7634 to 29 full-length EAV strains according to the serum neutralization test (SNT) result by mAb.

Interestingly, the GP5 protein of the EAV donkey VD7634 strain had an amino acid substitution in neutralization site B (aa 61), 13 amino acid substitutions in site C (aa 67, 69-73, 79, 81-82, 84-85, 89 and 90) and 4 amino acid substitutions in site D (aa 100-101, 104 and 106) compared to 29 EAV referenced sequences. These amino acid substitutions in GP5 may contribute to its distinctive neutralization phenotype (Supplementary table 3.6.).

3.5. Discussion

We have successfully determined the complete genome sequence of the novel EAV donkey strain isolated from a naturally infected wild donkey in northern Chile. In this study, we demonstrated the full genomic, phylogenetic and antigenic characterization of the EAV donkey VD7634 strain. The complete genome of EAV donkey VD7634 strain was 12,703 bp in length, which was 1nt shorter in length and had 75.6% nucleotide identity to both NCBI reference EAV sequence and our prototype virus EAV VBS. Comparisons of the EAV donkey VD7634 strain to both NCBI referenced EAV sequence and our prototype EAV VBS defined its genomic annotation. Further sequence alignment among 29 EAV referenced full-genome sequences determined nucleotide and amino acid percentage homology (%) and revealed a unique number of insertions and deletions in both 5′ and 3′ UTRs, nsp2 and GP3.

Six insertions and two deletions were found in the 5′ UTR and two insertions were found in 3′ UTR in EAV donkey VD7634 strain by comparison with 29 EAV referenced sequences. It is known that the 5′ and 3′ UTRs of PRRSV play important roles in viral replication, subgenomic RNA transcription and infectivity[448–450]. Even though 5′

and 3′ UTR sequences are highly conserved, the unique presence of insertions and deletions have been frequently identified in both North American and European genotypes of PRRSV[450,451]. A previous study demonstrated that a deletion of either one or two consecutive nucleotides had occurred in both 5′ and 3′ UTRs in highly pathogenic PRRSV isolates from China[452]. In contrast, a single-nucleotide deletion observed in an *in vitro* serial cell culture passaged PRRSV strain that was overattenuated from a highly pathogenic PRRSV strain[453]. These studies indicate that both 5′ and 3′ UTR sequences of *Arteriviruses* are involved in viral virulence and these may also affect EAV donkey VD7634 strain.

In this study, we determined that the EAV donkey VD7634 strain has a number of insertions and deletions in ORF1a (particularly in the encoding region of nsp2 replicase subunit) and ORF3. Very interestingly, EAV donkey VD7634 strain has similar unique patterns of insertions and deletions to EAV CW strains and French strains[17,19]. For instance, the consecutive 15-nucleotide insertion in ORF1a nsp2 encoding region (nucleotide position at 1565-1579) was present in S-numbered EAV isolates, but EAV donkey VD7634 strain and EAV CW and French strains did not have this particular insertion.

According to previous studies, the ten ORFs of EAV evolve at different rates during persistent infection. Most variations occur among the ORFs encoding viral structural proteins ORF3 and ORF5, which encode GP3 and GP5 enveloped glycoproteins of the virus, respectively[17,19,21]. For instance, EAV F61-63 strains have been shown to have a consecutive 12-nucleotide insertion in ORF3; this was not present in the EAV donkey VD7634 strain at nt location 10,653-10,664[19]. Molecular

characterization of the EAV S-numbered isolates revealed a unique consecutive 12-nucleotide insertion feature at nt position 10,684-10,695 that had not been observed in GP3 of EAV donkey VD7634 strain as well as CW, French and NA strains of EAV. Interestingly, EAV donkey VD7634 strain had two additional unique nucleotide deletions (G-10,669 and G-10,670) following this consecutive 12-nucleotide insertion in S-numbered isolates of EAV.

The GP5 protein, encoded by ORF5, consists of the known neutralization determinants of EAV[15,16,78]. Previous studies determined that the individual amino acid substitutions within the GP5 ectodomain occur during persistent infection of stallions. This results in the emergence of viruses with different neutralization phenotypes[142,164,166]. Neutralization testing with a large panel of well-characterized mAbs and polyclonal equine sera confirmed different phenotypes and demonstrated the nucleotide variations of ORF5 in EAV donkey VD7634 strain compared to other strains of EAV (CW96, CW01, EAV030, ATCC, ARVAC, S3583, S4216, S3685, S3861, S3854, S4333 and S3943[21,82]. EAV donkey VD7634 strain was found to have a distinct neutralization phenotype. This virus was highly neutralized by any of the mAbs that neutralized other isolates (mAbs: 6D10, 7E5, 10F11 and 6A2) as well as neutralized to a significantly high titers by polyclonal equine sera (anti-KY84) compared with other isolates. Interestingly, the comparative amino acid sequence analysis of GP5 protein determined that EAV donkey VD7634 strain had an amino acid substitution in neutralization site B (aa 61), 13 amino acid substitutions in site C (aa 67, 69-73, 79, 81-82, 84-85, 89 and 90) and 4 amino acid substitutions in site D (aa 100-101, 104 and 106) compared to 29 EAV referenced sequences. These amino acid substitutions in GP5 may correlate to this unique

neutralization phenotype of EAV donkey VD7634 strain associated with EAV persistency, however this needs to be investigated further.

In conclusion, there are approximately 3,000 wild donkeys living in small herds in the Huasco province, Atacama Region, Chile[454]. In the 16th century, Spanish settlers introduced donkeys in Chile for transportation[454]. In order to evaluate the health status of infectious diseases in these animals, SAG collected 312 serologic specimens from randomly selected wild donkeys and subjected them to serum neutralization testing[454]. As result of a serological survey for animal health monitoring, neutralizing EAV antibodies were found in 53% of the obtained samples giving a statistically significant p-value of 0.001. This finding indicates that the virus circulates in the wild donkey population[454]. Yet, EAV has not been detected in domestic horses in Chile[455]. EAV strains have been spread globally through the international movement of equids or shipment of semen and on occasion, have been responsible for epidemics of major economic consequences[266]. Also, EAV outbreaks appear only periodically, hence it is hard to trace the historic origin so we assume that EAV has been in circulation long before it's first isolation[48]. Phylogenies may be used for epidemiologically tracing the virus origin. For instance, EAV Gb_Glos_2012 strain from Great Britain was reported as presumably transferred from an imported Spanish stallion[48].

By reviewing the historical background of wild donkeys in Chile and the relationship of EAV donkey VD7634 strain to other EAV donkey strains and horse strains based on the phylogeny of ORF5, the origin of EAV donkey VD7634 strain may be closely related to EAV donkey strains from Republic of South Africa. The history of donkey movement by Spanish settlers would suggest that EAV donkey VD7634 strain

and EAV donkey strains from Republic of South Africa might have originated from each other or another source centuries ago.

Acknowledgements

This study was supported by the Agriculture and Food Research Initiative competitive grant no.2013-68004-20360 from the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA). Servicio Agricola y Ganadero (SAG) identified EAV positive in wild donkeys and provided clinical samples for this study. The next-generation sequencing using an Illumina Miseq technology was collaborated by Dr. Ganwu Li and Ms. Ying Zheng at the Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University.

CHAPTER FOUR

Complete Genome Sequence of a Noncytopathic Strain of Bovine Viral Diarrhea Virus 1 (BVDV-1), a Contaminant of the High Passage RK-13 Cell Line

***Published in ASM Genome Announc*. 3(5): e01115-15.**

Modified with permission

4.1. Summary

It has been previously reported that the rabbit kidney continuous cell line (RK-13; ATCC CCL-37 cell line, ATCC, Manassas, VA, USA) and its derivative high passage RK-13 cell line (HP-RK-13[KY]) is contaminated with a noncytopathic bovine viral diarrhea virus (ncpBVDV). Our laboratory has been using the HP-RK-13 (KY) cell line for isolation of equine arteritis virus from clinical specimens for many years. In this study, we report the complete genome sequence of the noncytopathic bovine viral diarrhea virus 1 (ncpBVDV-1) contaminating the HP-RK-13 (KY) cell line. The complete genome of this ncpBVDV strain (ncpBVDV HP-RK-13) is 12,271 nucleotides (nts) in length; it contains a single open reading frame (ORF) made up of 11,697 nts [387-12,083]) encoding 3898 amino acid polyproteins. Phylogenetic analysis indicated that this strain belongs to BVDV group 1b.

4.2. Introduction

Bovine viral diarrhea virus (BVDV) is a non-enveloped, positive-sense, single-stranded RNA virus whose genome size is approximately 12.3-12.5 kb. The virus is a member of the genus *Pestivirus* in the family *Flaviviridae* which also includes Border disease virus and classical swine fever virus[456,457]. Two major genotypes of BVDV are recognized (type 1 [BVDV-1] and 2 [BVDV-2]). BVDV-1 strains are genetically subdivided into at least 17 subtypes (a to q) and BVDV-2 strains into four subtypes (a to d)[458–461]. In addition, two distinct biotypes within each genotype have been identified: cytopathic viruses (cpBVDV) that cause cytopathic effects in cultured cells and noncytopathic viruses (ncpBVDV) that do not cause cytopathic effects in cultured cells (7-10). It has previously been reported that a number of cell lines (e.g. of cattle, sheep, goat, deer, bison, rabbit and domestic cat origin) including the RK-13 cell line (CCL-37; American Type Culture Collection [ATCC], Manassas, VA, USA) are persistently infected with ncpBVDV resulting from the use of BVDV contaminated fetal bovine serum in cell culture media[462–466]. Many laboratories use the RK-13 cell line from the ATCC or it derivatives for research and laboratory confirmation of various viral agents. Our laboratory has been using high passage RK-13 cells (P399-409; HP-RK-13 [KY]) for routine laboratory diagnostic investigation for over 50 years[467]. In this study, we determined the complete genome sequence of ncpBVDV virus present in the HP-RK-13 (KY) cells (P404) using next-generation sequencing (NGS) technology on an Illumina MiSeq platform according to previously established procedures[445].

4.3. Materials and methods

Cells and viruses

High passage rabbit kidney 13 (HP399-409; HP-KY-RK-13 [KY]; derived from ATCC® CCL-37™, American Type Culture Collection) cells were maintained in Eagle's minimum essential medium (EMEM, Cellgro®, Mediatech Inc., Herndon, VA, USA) with 10% ferritin-supplemented calf serum (HyClone Laboratories, Inc., Logan, UT, USA), penicillin and streptomycin (100 U/mL and 100 μg/mL) and 0.25 μg/mL of amphotericin B (Gibco®, Carlsbad, CA, USA).

Next-generation sequencing

Total viral RNA was extracted from the tissue culture fluid following freeze thaw of HP-RK-13 (KY) cell monolayers (P404) in 25-cm^2 flasks using a total RNA purification kit (Catalog no. 17200, Norgen Biotek Corp., Ontario, Canada). The RNA was quantified with a Qubit 2.0 spectrophotometer (Life Technologies, Carlsbad, CA, USA) and purified RNA was stored at -80°C. The complete genome sequence of ncpBVDV HP-KY-RK13 strain was obtained by next-generation sequencing (NGS) technology on an Illumina MiSeq platform following a previously established protocol[445]. Briefly, cDNA libraries were constructed from 100 ng of total RNA using a TruSeq Stranded total RNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions and previously described protocol. Multiplex libraries were prepared using barcoded primers and a median insert size of 340 base pairs (bp). Libraries were analyzed for size distribution using a Bioanalyzer and quantified by quantitative RT-PCR using a Kapa library quantification kit (Kapa Biosystems, Boston,

MA, USA); relative volumes were pooled accordingly. The pooled libraries were sequenced on an Illumina MiSeq platform with 150 bp end reads following standard Illumina protocols.

Sequence analysis

Sequences were mapped to all known BVDVs and mapped read sets were used for *de novo* assembly using ABySS, v.1.3.7 (BC Cander Agency, Vancouver, Canada) and Geneious 7.0.6 software (Biomatters Ltd., Auckland, New Zealand). The new paired-end read sets were extracted from the original paired-end fastq files using seqtk with the list of mapped ID to insure both reads of one pair were selected. The raw reads in FASTA formats were imported into IGV_2.3.79 for mapping and paired using the default parameter for illumine short read. Comparative alignment analysis of nucleotide (nt) and amino acid (aa) sequences was performed through Geneious 7.0.6 software (Biomatters Ltd., Auckland, New Zealand).

Phylogenetic analysis

Phylogenetic trees were constructed by using the Maximum Likelihood method based on the Tamura-Nei model[446], applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) distance with Gamma-distributed rates, pairwise gap deletion and boopstrap resampling (1,000 replications)[447].

4.4. Results and discussion

The complete genome of ncpBVDV contaminating the HP-RK-13 (KY) cell line (ncpBVDV HP-KY-RK13 strain) is comprised of 12,271 nucleotides (nt) and contains a 5′ untranslated region (UTR; 386 nts), a single open reading frame (ORF; 11,697 nt [387-12,083]) and a 3′ UTR (188 nt). The single ORF encodes a 3,898 amino acid polyprotein which is predicted to be cleaved into 12 proteins (Table 4.1.[468]). The polyprotein is cleaved into 12 proteins that are respectively enumerated from amino acids to carboxyl terminus as follow: NH2-N^{pro}-C-E^{rns}-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH. Specifically, the polyprotein is cleaved by the viral protease into four structural proteins (C, E^{rns}, E1 and E2) and eight nonstructural proteins (N^{pro}, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B). Based on the putative cleavage sites of structural and nonstructural proteins of other BVDVs (Table 4.1.), the ncpBVDV HP-KY-RK13 strain UTRs and genes were annotated to the following nucleotide positions in the genome: 5′ UTR, 1-386 (386 nt); N^{pro}, 387-887 (501 nt); C, 888-1,196 (309 nt); E^{rns}, 1,197-1,877 (681 nt); E1, 1,878-2,462 (585 nt); E2, 2,463-3,584 (1,122 nt); P7, 3,585-3,794 (210 nt); NS2, 3,795-5,153 (1,359 nt); NS3, 5,154-7,202 (2,049 nt); NS4A, 7,203-7,394 (192 nt); NS4B, 7,395-8,435 (1,041 nt); NS5A, 8,436-9,923 (1,488 nt); NS5B, 9,924-12,083 (2,160 nt); 3′ UTR, 12,084-12,271 (188 nt) (Table 4.2.). The complete genomic sequence of ncpBVDV HP-KY-RK13 strain has been submitted to GenBank under accession number: KT355592.

Comparative nucleotide sequence analysis showed that the ncpBVDV HP-KY-RK-13 strain had 85.2% to 99.7% identity with 11 strains of BVDV-1b and 68.6% to 70.9% identity with eight strains of BVDV-2. This further demonstrated that the

ncpBVDV HP-KY-RK13 strain is very closely related to the previously described ncpBVDV present in the RK-13 cells from Japan (RK13/E$^-$ strain [GenBank accession number: JX419397.1; 12064 nt, 99.7% identity])[469]. The ncpBVDV HP-KY-RK13 strain was compared with eleven other BVDV-1b subtypes available in GenBank; P7 and NS5A proteins are the most variable (78.1% to 99.0% and 79.1% to 99.6%, respectively), whereas the E^{rns}, NS3, NS4A and NS4B proteins are the most conserved. The ncpBVDV HP-KY-RK13 strain had several nucleotide insertions and deletions compared to several of the other BVDV-1b strains (Table 4.3.). However, the specific insertion(s) and/or deletion(s) that are predictable for pathogenesis and establishment of persistent infection in the HP-RK-13 (KY) cell line have not been determined.

A phylogenetic tree was constructed using the complete genome sequence of ncpBVDV HP-KY-RK13 strain along with 34 BVDV strains available in GenBank (Figure 4.1.). The phylogenetic analysis determined that ncpBVDV HP-KY-RK13 strain belongs to BVDV-1b genotype.

Acknowledgements

This study was supported by the Agriculture and Food Research Initiative competitive grant no.2013-68004-20360 from the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA). The next-generation sequencing using an Illumina Miseq technology was collaborated by Dr. Ganwu Li and Ms. Ying Zheng at the Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University.

Table 4.1. Cleavage sites between structural and nonstructural proteins of reference BVDV strains and HP-KY-RK13 strain

ORF	N^{pro}	C	E^{ms}	E1	E2	P7	NS2	NS3	NS4A	NS4B	NS5A	NS5B
SH-28	-LWVTS	CSDEG-QLVTG	ENIT-FGAYA	ASPYC-TGAQG	FPECK-Q**I**AMG	ARVNT-GVVKA	**SKT**NT-	QVTGL	STAEN-ELKEL	**QKIKE**-KIRNL	**SS**NYL-YTMKL	SSWST-
XJ-04	-LWVTS	CSDEG-QLV**E**G	ENIT-FGAYA	ASPYC-TGAQ**S**	FPECK-Q**I**AMG	ARVNT-G**L**VKA	**REI**NT-	-QVTGL	STAEN-ELKEL	AVGDL-KIRNL	**SS**NYL-YTMKL	SSWST-
890	-LWVTS	CSDEG-QLVTG	ENIT**Q**-FGA**H**A	ASPYC-TG**V**QG	FPECK-QMAMG	A**G**VNT-GAVKA	**IPPEE**-	-QVTGL	STAEN-ELKEL	AVGDL-KIRNL	SGNYI-YTMKL	SSWST-
NADL	-LWVT**T**	CSDTK-QVT**M**G	ENIT-FGAYA	ASPYC-TG**V**QG	**KKLFD**-Q**KAL**G	**IQYGS-DV**VKA	**DSGGQ**-	-QVTGL	**SSA**EN-ELKEL	**ASGDV**-KIRNL	SGNYI-Y**A**MKL	SSW**FL**-
ZM-95	MELIS-LWVTS	**C**SDTK-QVT**A**G	ENIT-FGAYA	ASPYC-TGAQG	**YPDCK**-Q**RAS**G	**TQCGA**-GAVKA	**ESGTQ**-HLGWI	LRGPA-Q**VA**GL	**SSA**EN-ELKEL	**AL**GDV-KIRNL	SGNYI-YTMKL	SSW**FM**-VGASS
SD0803	MELIS-**WVASC**	**SDPKN**-QVT**M**G	ENIT-FGAYA	ASPYC-TG**V**QG	**IPECR**-Q**KAS**G	**AQCGA**-GAVKA	DSETQ-GWILR	GPAVC-QVVGL	SSAEN-ELKEL	**AL**GD**T**-KIRNL	SGNYI-YTMKL	SSWFL-V**N**ASS
CC13B	MELIT-LWVSS	CSDTK-QVTMG	ENTI-FGAYA	ASPYC-TGAQG	YPDCK-QMALG	AQYGS-GMVKA	EPGTK-GWILR	GPAVC-QVVGL	SSAEN-ELKEL	AVGDL-KIRNL	SGNYV-YTMKL	SSWFL-VGASS
HP-KY-RK13	MELIT-LWISS	CSDTK-QVTVG	ENIT-FGAYA	ASPYC-TGVQG	YPDCK-QMASG	AQYGA-GMARA	EPGAQ-GWILR	GPAVC-QVVGL	SSAEN-ELKEL	AVGDL-KIRNL	SGNYV-YTMKL	SSWFL-VGASS

Cleavage sites of SH-28, XJ-04, 890, NADL, ZM-95 strains of BVDV from Tao *et al*. (2013)[468].

Table 4.2. Nonstructural and structural proteins of BVDV strains

BVDV strain (GenBank accession number)	BVDV HP-KY-RK13 (KT355592.1)		BVDV SD0803 (JN400273.1)		BVDV CC13B (KF772785.1)		BVDV 1 Japan (JX419397.1)		BVDV ZM-96 (AF526381.3)	
Full-length genome (nt)	122271 nt		12271 nt		12265 nt		12064 nt		12220 nt	
1 ORF amino acid sequence length (aa)	3898 aa		3898 aa		3898 aa		3898 aa		3901 aa	
ORF	Nucleotide length (nt)	Amino acid length (aa)	Nucleotide length (nt)	Amino acid length (aa)	Nucleotide length (nt)	Amino acid length (aa)	Nucleotide length (nt)	Amino acid length (aa)	Nucleotide length (nt)	Amino acid length (aa)
5′ UTR	1-386 (386)	- (128)	1-388 (388)	- (129)	1- 380 (380)	- (126)	1-388 (388)	- (129)	1-329 (329)	- (109)
N^{pro}	387-887 (501)	1-167 (167)	389-892 (504)	1-168 (168)	381-881 (501)	1-167 (167)	389-892 (504)	1-168 (168)	330-833 (504)	1-168 (168)
C	888-1196 (309)	168-270 (103)	893-1198 (306)	169-270 (102)	882-1190 (309)	168-270 (103)	893-1198 (306)	169-270 (102)	834-1145 (312)	169-272 (104)
E^{rns}	1197-1877 (681)	271-497 (227)	1199-1879 (681)	271-497 (227)	1191-1871 (681)	271-497 (227)	1199-1879 (681)	271-497 (227)	1146-1826 (681)	273-499 (227)
E1	1878-2462 (585)	498-692 (195)	1880-2464 (585)	498-692 (195)	1872-2456 (585)	498-692 (195)	1880-2464 (585)	498-692 (195)	1827-2411 (585)	500-694 (195)
E2	2463-3584 (1122)	693-1066 (374)	2465-3586 (1122)	693-1066 (374)	2457-3578 (1122)	693-1066 (374)	2465-3586 (1122)	693-1066 (374)	2412-3536 (1125)	695-1069 (375)
P7	3585-3794 (210)	1067-1136 (70)	3587-3796 (210)	1067-1136 (70)	3579-3788 (210)	1067-1136 (70)	3587-3796 (210)	1067-1136 (70)	3537-3746 (210)	1070-1139 (70)
NS2	3795-5153 (1359)	1137-1589 (453)	3797-5155 (1359)	1137-1589 (453)	3789-5147 (1359)	1137-1589 (453)	3797-5155 (1359)	1137-1589 (453)	3747-5099 (1353)	1140-1590 (451)
NS3	5154-7202 (2049)	1590-2272 (683)	5156-7204 (2049)	1590-2272 (683)	5148-7196 (2049)	1590-2272 (683)	5156-7204 (2049)	1590-2272 (683)	5100-7154 (2055)	1591-2275 (685)
NS4A	7203-7394 (192)	2273-2336 (64)	7205-7396 (192)	2773-2336 (64)	7197-7388 (192)	2273-2336 (64)	7205-7396 (192)	2773-2336 (64)	7155-7346 (192)	2276-2339 (64)
NS4B	7395-8435 (1041)	2337-2683 (347)	7397-8437 (1041)	2337-2683 (347)	7389-8429 (1041)	2337-2683 (347)	7397-8437 (1041)	2337-2683 (347)	7347-8387 (1041)	2340-2686 (347)
NS5A	8436-9923 (1488)	2684-3179 (496)	8438-9925 (1488)	2684-3179 (496)	8430-9917 (1488)	2684-3179 (496)	8438-9925 (1488)	2684-3179 (496)	8388-9875 (1488)	2687-3182 (496)
NS5B	9924-12083 (2160)	3180-3898 (719)	9926-12085 (2160)	3180-3898 (719)	9918-12077 (2160)	3180-3898 (719)	9926-12085 (2160)	3180-3898 (719)	9876-12035 (2160)	3183-3901 (719)
3′ UTR	12084-12271 (188)	- (62)	12086-12271 (186)	- (62)	12078-12265 (188)	- (62)	12086-12271 (186)	- (62)	12036-1220(185)	- (61)

Table 4.3. Insertion and deletion positions in full-length genome alignments of BVDV-1b strains

Compared strain (GenBank accession number)	ORF	Nucleotide position	Nucleotide substitution	HP-KY-RK13 Insertion/Deletion
JL-1	5′ UTR	11	A	deletion
|KF501393.1|		17	T	deletion
		45-48	AAAA	insertion
		75-76	AA	insertion
	E2	2719-2721	GAA	insertion
	NS2	5010-5021	AATATCCTTCAG	deletion
CP7 strain	5′ UTR	45-47	AAA	deletion
|U63479.1|				
	NS2	4358-4384	GGGTGTTCCTTATCCGGACCCTCA	deletion
3156	5′ UTR	45-47	AAA	insertion
|JN704144.1|		75-76	AA	insertion
	NS5A	9406	A	insertion
		9411	A	deletion
		9800-9803	GGGC	insertion
		9812-9815	AAGT	deletion
	NS5B	11960	A	deletion
		11976	A	insertion
	3′ UTR	12090-12092	AAT	insertion
		12116-12120	GAATT	deletion
		12125-12129	TGTAT	insertion
		12168-12176	CCTCAAAAG	deletion
		12185-12206	CTCAACATACACAGCTAAACAG	deletion
		12210-12215	TTGAGA	deletion
Germini	E2	2469-2471	GAC	insertion
|JX297516.1|		2720-2725	AAGAAA	insertion
	3′ UTR	12033	C	insertion

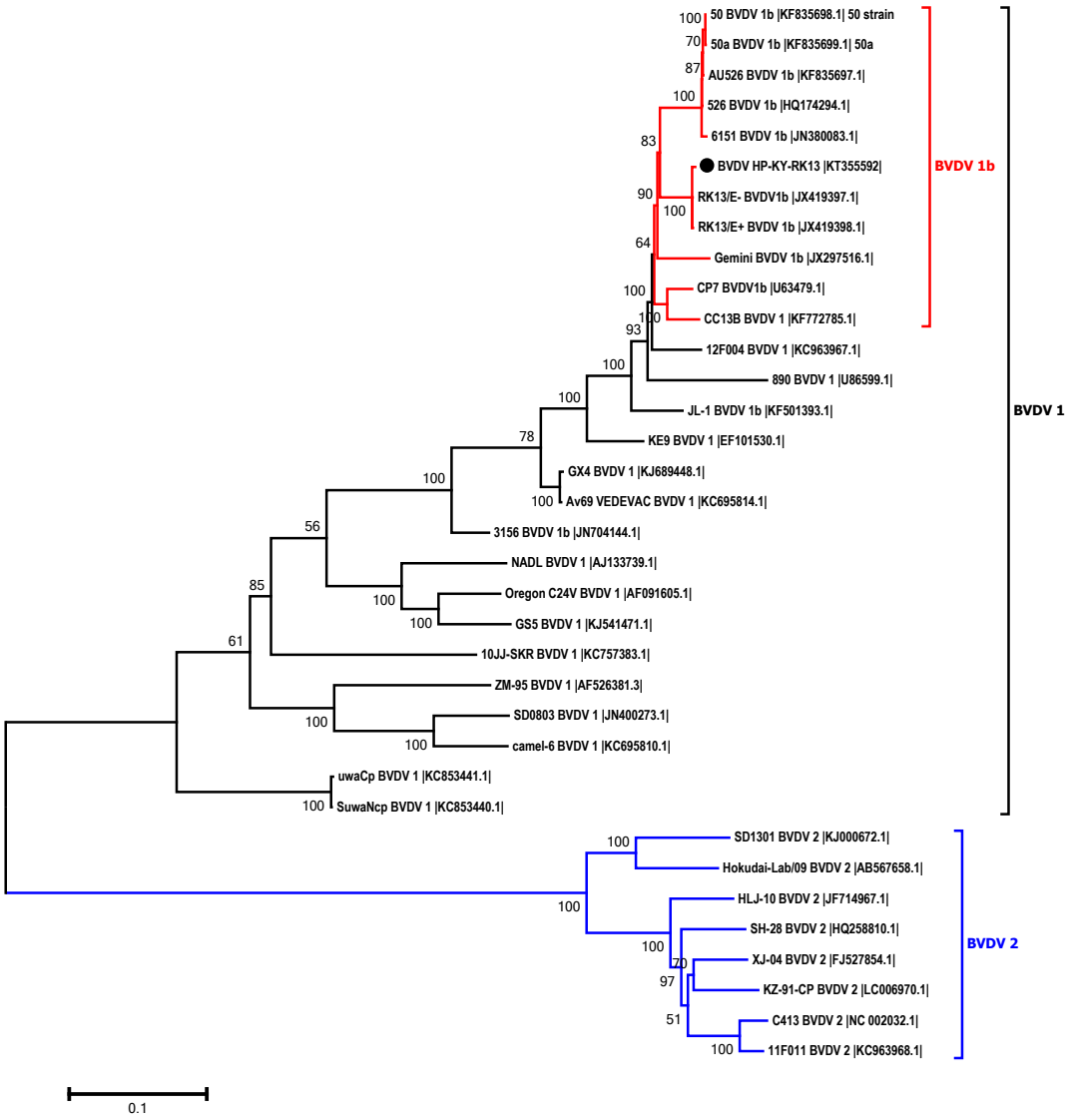


Figure 4. 1. Phylogenetic analyses of whole genome of 34 reference BVDV strains obtained from GenBank and the ncpBVDV HP-KY-RK13 strain identified with a black dot.

CHAPTER FIVE

Summary of Research Findings

Chapter one provides a review of the current literature on equine arteritis virus (EAV) and equine viral arteritis (EVA) which includes viral classification, genome organization, replication cycle and neutralization determinants and the disease caused by the virus including epidemiology, modes of transmission, immune response, clinical signs, pathogenesis, persistent infection in the stallion, diagnosis, treatment, prevention and control. Also, a brief overview of RNA virus quasispecies theory and viral evolutionary mechanisms is provided. Viruses are infectious particles that cannot self-replicate without susceptible and permissive host cells. Like other RNA viruses, the EAV RNA-dependent RNA polymerase enzyme lacks proofreading capability (error-prone), which results in the generation of a large number of viral genomes with mutations. The viral population size, number of mutations and quasispecies effect drive RNA virus evolution. RNA viruses evolve close to their error threshold at a high mutation rate during a short period of viral replication and produce a large viral population and enough mutants to enhance viral survival. Viral populations exist as a dynamic mutant distribution (mutant swarms) that is called quasispecies. The viral quasispecies population constantly evolves in the face of selection pressures to adapt to a given environment. When the virus fitness decreases, the viral population loses the ability to adapt to its environment and it experiences a bottleneck event, which is a sharp reduction in the size of a population due to environmental events such as selective pressure and viral fitness.

In chapter two, EAV evolution was studied during long-term persistent infection established following experimental infection of seven stallions with the KY84 strain of EAV. In addition, two stallions naturally infected with the EAV KY84 strain for a period of 7 to 10 years were also included in the study. This was the first study where the evolution of the full-length EAV genome was studied during long-term persistent infection using next-generation sequencing. The sequencing data were analyzed using contemporary sequence and phylogenetic analysis software. The statistical analyses were performed to measure the intra-host evolutionary dynamics, quasispecies diversity, rate of evolution and selection pressure among viral populations from the ancestral strain (the initial inoculum) through the evolutionary time scale (acute and persistent infection periods). Analysis of viral sequences in nasal secretions and buffy coat cells revealed a lack of extensive positive selection; however, characteristics of the mutant spectra were different in the two sample types. By contrast, virus populations in semen during the acute infection phase were under selective pressure and subjected to bottleneck events, which was reflected in a reduction in population size and the positive selection of viral genomes. During long-term persistent infection, virus evolved in the reproductive tract of the stallions; as a result, a non-stochastic evolutionary process with many atypical high-frequency minor variants was identified. This indicates that an active selection pressure continually morphs EAV quasispecies population structures and dynamics during persistent infection. Comparative analysis of the evolution in experimentally infected stallions with that of naturally infected stallions showed a higher degree of correlation in the rate of sequence divergence through time and mirrored the patterns of selection between experimental and naturally infected stallions. In summary, this study

demonstrated the genetic bottleneck event and selection during the acute infection period and intra-host quasispecies diversification during persistent infection in the stallion reproductive tract.

A total of 53 sequential isolates made during acute and persistent infection were analyzed and compared to the of EAV KY84 strain present in the inoculum. Major and minor viral populations were identified in the inoculum used for experimental infection of the seven stallions. The more homogenous virus populations with major and minor variants were demonstrated in the inoculum virus and viruses isolated during the acute infection period. A distinct quasispecies population size and selection pattern were observed during the initial establishment of the viral population in nasal swab and buffy coat samples. For instance, the reduction in variant population size observed in buffy coat virus isolates may reflect a transmission bottleneck scenario. The results indicate that the selective bottleneck event happens during virus migration from the upper respiratory tract to the regional lymph nodes via the blood stream until EAV reaches its final destination in the male reproductive tract. In the virus population during acute infection period, strong negative evolutionary selection was observed at nucleotide positions of nsp9 located at aa 6729-6816, which play a role in RdRp function and in assembly of the viral replication complex. This suggested that the conserved nsp9 region with low frequency mutation might have a regulatory effect on viral replication.

In contrast, the presence of continual positive selection resulting in quasispecies diversity was observed in specific regions of the genome that drives the evolution of semen virus isolates made during persistent infection. For instance, an atypical high frequency minor variant population implies an ongoing interaction between the viral

qausispecies population and the host immune system. Even though 10 to 70% of stallions become persistently infected carriers, the presence of high levels of serum virus neutralizing antibodies effectively eliminates systemic infection in body tissues except for the reproductive tract of the stallion. The strong diversifying selection was observed especially in ORF3 and ORF5. The results suggest that quasispecies diversity may play a significant role in establishment of persistent infection and the pathogenesis of EAV.

In conclusion, this study identified virus evolution with the dynamic population in sequential samples from both experimentally and naturally infected stallions during acute and persistent infection. Questions to emerge from the research findings are as follows: What leads to persistent infection of EAV in the stallion reproductive tract? Supplementary tables 2.2. and 2.3. may provide a key to understanding the major amino acid changes in structural proteins, particularly GP5, which contains the major neutralization sites of EAV. What will happen to pathogenesis of the virus as it evolves over time? Does the phenomenon of the virus quasispecies population occur similarly in mares? Due to the lack of resources, this *in vivo* study could not be extended to include breeding trials with experimental mares to define the pathogenesis of the infection. However, supplementary tables 2.4. and 2.5. suggest the presence of distinctive neutralizing phenotypes; further investigation of the latter may well help in understanding in pathogenesis of the virus as well as persistent infection of EAV in the stallion reproductive tract.

Chapter three was devoted to a study of the isolation and characterization of a novel field strain of EAV from a persistently infected donkey in northern Chile, described as EAV donkey VD7634 strain. This is the first full-length genome sequence of EAV

obtained from a wild male donkey. The annotation of EAV donkey VD7634 strain was predicted based on the NCBI referenced EAV sequence (accession numbers: NC002532). There are only partial ORF5 sequences (nucleotide location: 11,295-11,812 [518 bp in length]) of donkey EAV isolates available from GenBank. Both sequence identity and phylogenetic analyses indicated that the partial ORF5 sequence of EAV donkey VD7634 strain was closely related to that of a donkey isolate from the Republic of South Africa (RSA). However, further phylogenetic analysis using 29 full-length EAV strains from horses grouped the EAV donkey VD7634 strain with French and CW strains (EU-2 subgroup). Although EAV donkey VD7634 strain diverged from both NA and EU subgroups of 29 referenced EAV sequences of horses, phylogenetic analysis including other arteriviruses indicated that EAV donkey VD7634 strain belongs to the genus *Equartevirus* (Family: *Arteriviridae* and Order: *Nidovirales*). Neutralization testing with a large panel of well-characterized mAbs and polyclonal equine sera confirmed antigenic cross reactivity of this strain. The newly identified EAV donkey VD7634 strain had a unique neutralization phenotype compared to other EAV strains.

Even though EAV has not yet been detected from domestic horses in Chile, this is the first isolate of EAV from a carrier male donkey originating from a wild herd for an unknown period of time. It is very difficult to imagine how these feral donkeys survive, some persistently infected with EAV, in Huasco province, Atacama region in Chile, that is geographically surrounded by arid desert, salt lakes and felsic lava that flows towards the Andes mountains[470]. There are very large deposits of sodium nitrate in the region, which could serve as fertilizer and may enable plants to grow as a food source for the feral donkeys. Although the geographical location limits the possibility of contact of

EAV positive feral donkeys with domestic equids including donkeys and horses, this study might suggest that there could be the potential for an outbreak of EVA. Accordingly, implementation of prevention and control measures against EVA are highly recommended to protect the domestic equine population in Chile from the risk of such an event.

Chapter four, the full-length genome sequence of a noncytopathic bovine diarrhea virus 1 (ncpBVDV-1) contaminating the high passage rabbit kidney continuous cell line (HP-RK-13 [KY]). It has been repeatedly reported that the RK-13 cell line is contaminated with noncytopathic bovine viral diarrhea virus (ncpBVDV) as a result of the use of BVDV contaminated fetal bovine serum in cell culture media. This laboratory has been using the HP-RK-13 (KY) cell line for isolation of EAV from clinical specimens for many years.

In this study, the complete genome of ncpBVDV-1 (ncpBVDV HP-KY-RK13 strain) was determined and the genome organization predicted by comparing cleavage sites to other published BVDV sequences. Several nucleotide insertions and deletions were observed in the ncpBVDV HP-KY-RK13 strain by comparing it to several other BVDV-1b strains; however, its role in the pathogenesis and establishment of persistent virus infection has to be determined. Phylogenetic analysis determined that the ncpBVDV HP-KY-RK13 strain belongs to the BVDV-1b genotype.

APPENDIX 1

List of Abbreviations

$^{\circ}C$	Celsius
3CLSP	3C-like serine proteases
$TCID_{50}$	50% tissue culture infective dose
R^2	Accuracy of the linear regression analysis
COOH	Acidic carboxylic group
Vero or MA-104	African green monkey kidney
APRAV-1	African pouched rat arterivirus
AAEP	American Association for Equine Practitioners
ATCC	American Type Culture Collection
AVMA	American Veterinary Medical Association
aa	Amino acid
BHK-21	Baby hamster kidney-21
bp	Base pair
NH2	Basic amino group
BVDV	Bovine viral diarrhea virus
BC	Buffy coat
HS or HT-7	Canine hepatitis virus-transformed hamster tumor
CO_2	Carbon dioxide
CMC	Carboxymethylcellulose
CD	Cluster of differentiation
CF	Complement-fixing
cDNA	Complementary DNA
Cys	Cysteine
CP	Cysteine protease
cp	Cytopathic
cpBVDV	Cytopathic BVDV
CPE	Cytopathic effect
dpi	Days post-infection
DNA	Deoxyribonucleic acid
DIVA	Differentiate vaccinated from infected animals
DMV	Double-membrane vesicle
EMEM	Eagle's minimum essential medium
ER	Endoplasmic reticulum
E	Enveloped protein
ELISA	Enzyme-linked immunosorbert assay
EAV	Equine arteritis virus
C-X-C motif	Equine chemokine
EqCXCL16	Equine chemokine ligand 16
ECA11	Equine chromosome 11

EEC	Equine endothelial cell
EO	Equine ovary cell
EVA	Equine viral arteritis
ERGIC	ER-Golgi intermediate compartment
EU	European
FUBAR	Fast unconstrained Bayesian analysis for inferring selection
FMDV	Foot-and-mouth disease virus
GP	Glycoprotein
GnRH	Gonadotropin releasing hormone
h	Hour
H	Hemagglutinin
HmLu	Hamster lung
HS	Heparin sulfate
HAV	Hepatitis A virus
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HP-RK-13 (KY)	High passage RK-13 (KY)
hpi	Hours post-infection
huCXCL16	Human chemokine ligand 16
HIV	Human immunodeficiency virus
HyPhy	Hypothesis testing using phylogenies
ARTERVAC®	Inactivated EAV vaccination
pEAVrMLVB	Infectious cDNA clone of the MLV strain of EAV
IACUC	Institutional Animal Care and Use Committee
iiRT-PCR	Insulated isothermal RT-PCR
IFN	Interferon
IL	Interleukin
ICTV	International Committee for Taxonomy of Viruses
KY	Kentucky
KY84	Kentucky 84 strain
kb	Kilobase
kDa	Kilodalton
L	Liter
LDV	Lactate dehydrogenase-elevating virus
LTH	Leader TRS hairpin
MCL	Maximum composite likelihood
ML	Maximum likelihood
M	Membrane protein
mRNA	Messenger RNA
µg/mL	Microgram per milliliter
µL	Microliter
MIA	Microsphere immunoassay

MEME	Mixed effect model of evolution
ARVAC®	Modified live attenuated EAV vaccination
MLV	Modified live attenuated vaccination
mAbs	Monoclonal antibodies
NS	Nasal swab or nasal secretion
NAHMS	National Animal Health Monitoring System
NCBI	National Center for Biotechnology Information
N	Neuraminidase
NGS	Next-generation sequencing
NendoU	Nidoviral-endonuclease specific for U
ncp	Noncytopathic
ncpBVDV	noncytopathic BVDV
nsp	Nonstructural protein
NS3	Nonstructural protein 3
n	Nonsynonymous
∏n	Nonsynonymous nucleotide diversity
n/N	Nonsynonymous substitutions per site
NA	North American
N protein	Nucleocapsid protein
NTPase	Nucleoside-triphosphatase
nt	Nucleotide
∏	Nucleotide diversity
dN	Number of n/N
dS	Number of s/S
ORF	Open reading frame
p	P-value
PCP	Papain-like cysteine protease
P1	Passage 1
P-PMO	Peptide-conjugation to the PMO
PMO	Phosphorodiamidate morpholino oligomer
PFU/mL	Plaque-forming units per milliliter
PCR	Polymerase chain reaction
pp	Polyprotein
PRRSV	Porcine reproductive respiratory syndrome virus
RK-13	Rabbit kidney-13
qRT-PCR	Real-time RT-PCR
rRT-PCR	Real-time RT-PCR
RTC	Replication/transcription complex
RSA	Republic of South Africa
RVN	Reticulovesicular network
RT-PCR	Reverse transcription-polymerase chain reaction
RNA	Ribonucleic acid
RdRp	RNA-dependent RNA polymerase

RT-nPCR	RT-nested PCR
S	Semen
SP	Serine protease
SAG	Servicio Agricola y Ganadro
SHFV	Simian hemorrhagic fever virus
SNV	Single nucleotide variant
ss	Single stranded
SH2	Src homology 2
sp	Structural protein
sg mRNA	Subgenomic viral messanger RNA
s	Synonymous
∏s	Synonymous nucleotide diveresity
s/S	Synonymous substitutions per site
TCF	Tissue culture fluid
TRS	Transcription regulating sequence
TNF	Tumor necrosis factor
USDA-APHIS	United States Department of Agriculture Animal and Plant Health Inspection Service
USA	United States of America
U/mL	Units per milliliter
UTR	Untranslated region
VS	Vas deferens
VN	Viral neutralizing
VBS	Virulent Bucyrus strain
VI	Virus isolation
VNT	Virus neutralization test
VRP	Virus replicon particle
VLP	Virus-like particle
WPDV	Wobbly possum disease virus
OIE	World Organisation for Animal Health
ZF	Zinc finger
ZBD	Zinc-binding domain

APPENDIX 2

Glossary Relevant to the Virus Evolution

Adaptation	The constant steady replication of large viral populations results in a consequence of high fitness
Antigenic drift	One of the main mechanisms that is a result of point mutations in influenza virus genes, encoding HA and NA that cause alternations in the structure of the main viral surface antigens. This is a slow process that is on-going and which causes antigenic differences in strains of influenza virus
Bottleneck event	Modification of viral population sizes resulting in the heterogeneous and dynamic nature of viral populations that are suitable in the environment by fitness gain
Defectors	Types of defective mutants that are associated with virus extinction caused by lethal mutagenesis
Drug resistance	A general phenomenon and the selection of a drug-resistant mutant against a specific antiviral drug is evidence of the phenomenon
Error-threshold	The value that is determined by the replication accuracy and the virus fitness of the dominant or master sequence corresponding to the mean fitness value of the error copies
Exonuclease activities	The DNA-dependent DNA polymerases have proofreading ability that are enable them to modify or correct any such mistakes
Lethal mutagenesis	A process of suppression in the replication of a high fitness variants
Molecular clock	The neutral theory of molecular evolution that means the rates of evolution reflect mutation rates
Mutant spectra	The ensemble of genomes that forms a viral quasispecies and its complexity and formation are closely related to its biological environment.
Mutation frequency	The proportion of mutant viruses in a population, and is affected by many biochemical and environmental selective factors
Mutation rate	The frequency of occurrence of a mutation during genome replication that can be calculated by mutations per site per replication. It is a value that is an independent of the fitness of the parental and mutated genomes

Negative selection	The process of removing a genotype in an evolving population that causes the negative evolution of phenotypic traits expressed by individuals
Next-generation sequencing	The application of modern technologies to study the virus evolution and viral quasispecies population
Non-equilibrium	A mutant spectrum in variable fitness landscapes, a theoretical construction of the evolutionary process
Nonsynonymous mutation	A nucleotide mutation that changes the amino acid sequence of a protein
Population equilibrium	An important role in maintaining a constant consensus sequence in viral populations
Positive selection	The process of a dominant genotype resulting in the positive evolution of phenotypic traits expressed by the individuals of the evolving population
Quasispecies	Also known as mutant swarms or mutant clouds; a continuous process of genetic variations in RNA viruses is responsible for heterogeneous mutant distributions in virus population
Quasispecies dynamics	RNA viruses evolve as a consequence of disequilibria of mutant distributions instead of linear accumulation of mutations
Quasispecies memory	The presence of maintaining virus subpopulations in the mutant spectrum at high frequencies than mutation rates
Rate of evolution	Viral quasispecies parameter which is the number of accumulated mutations in viral genomes through a time point
Reassortant	When a susceptible host cell is coinfected by two strains of a virus, the progeny virus may have a mixture of genome segments from the two parental strains of a virus
Recombination	Viruses evolve by rearranging their genomes or functional parts among different viruses
RNA-dependent RNA polymerase	Most RNA viruses encode for a low fidelity polymerase enzyme that encoded by most RNA viruses that can result in a mutation rate of 10^3 to 10^5 errors per nucleotide per replication cycle.
Selective constraints	Antiviral drugs, immune components of the innate or the adaptive immune response including those induced by vaccination and interfering RNAs

Synonymous mutations	A nucleotide mutation that does not change the amino acid sequence of a protein
Viral fitness	A major parameter which results in a viral population adapting to a specific environment and producing infectious progeny under the formulation of quasispecies theory and the error threshold relationship
Viral quasispecies	It allows RNA viruses to adapt to a given environment with the advantages of dynamics of infection, virus spread, attenuation and virulence, complexity and self-organization
Virus extinction	Mutant spectrum complexity occurs by an increase in the mutation rate above an extinction threshold.

APPENDIX 3

Supplementary Tables

Supplementary table 2.1. GenBank submission information of 53 sequential isolates of EAV KY84 strain from this study

Sample Number	Sample ID	GenBank Submission Name	Collection dpi	Specimen Type	GenBank ID
1	EAV KY84 strain	KY84	0	Tissue culture fluid	MG137429
2	L136 (A) NS D6	KY84-L136A-NS6	6	Nasal swab	MG137434
3	L137 (B) NS D6	KY84-L137B-NS6	6	Nasal swab	MG137443
4	L138 (C) NS D6	KY84-L138C-NS6	6	Nasal swab	MG137448
5	L139 (D) NS D6	KY84-L139D-NS6	6	Nasal swab	MG137453
6	L140 (E) NS D6	KY84-L140E-NS6	6	Nasal swab	MG137460
7	L141 (F) NS D6	KY84-L141F-NS6	6	Nasal swab	MG137469
8	L142 (G) NS D6	KY84-L142G-NS6	6	Nasal swab	MG137476
9	L136 (A) BC D6	KY84-L136A-BC6	6	Buffy coat	MG137435
10	L137 (B) BC D6	KY84-L137B-BC6	6	Buffy coat	MG137444
11	L138 (C) BC D6	KY84-L138C-BC6	6	Buffy coat	MG137449
12	L139 (D) BC D6	KY84-L139D-BC6	6	Buffy coat	MG137454
13	L140 (E) BC D6	KY84-L140E-BC6	6	Buffy coat	MG137461
14	L141 (F) BC D6	KY84-L141F-BC6	6	Buffy coat	MG137470
15	L142 (G) BC D6	KY84-L142G-BC6	6	Buffy coat	MG137477
16	L136 (A) S D5	KY84-L136A-S5	5	Semen	MG137436
17	L137 (B) S D5	KY84-L137B-S5	5	Semen	MG137445
18	L138 (C) S D5	KY84-L138C-S5	5	Semen	MG137450
19	L139 (D) S D5	KY84-L139D-S5	5	Semen	MG137455
20	L140 (E) S D5	KY84-L140E-S5	5	Semen	MG137462
21	L141 (F) S D5	KY84-L141F-S5	5	Semen	MG137471
22	L142 (G) S D5	KY84-L142G-S5	5	Semen	MG137478
23	L136 (A) S D9	KY84-L136A-S9	9	Semen	MG137437
24	L137 (B) S D9	KY84-L137B-S9	9	Semen	MG137446
25	L138 (C) S D9	KY84-L138C-S9	9	Semen	MG137451
26	L139 (D) S D9	KY84-L139D-S9	9	Semen	MG137456
27	L140 (E) S D9	KY84-L140E-S9	9	Semen	MG137463
28	L141 (F) S D9	KY84-L141F-S9	9	Semen	MG137472
29	L142 (G) S D9	KY84-L142G-S9	9	Semen	MG137479
30	L136 (A) S D107	KY84-L136A-S107	107	Semen	MG137438
31	L137 (B) S D107	KY84-L137B-S107	107	Semen	MG137447
32	L138 (C) S D107	KY84-L138C-S107	107	Semen	MG137452
33	L139 (D) S D107	KY84-L107D-S107	107	Semen	MG137457
34	L140 (E) S D107	KY84-L140E-S107	107	Semen	MG137464
35	L141 (F) S D107	KY84-L141F-S107	107	Semen	MG137473
36	L142 (G) S D107	KY84-L142G-S107	107	Semen	MG137480
37	L136 (A) S D170	KY84-L136A-S170	170	Semen	MG137439
38	L139 (D) S D170	KY84-L139D-S170	170	Semen	MG137458
39	L140 (E) S D170	KY84-L140E-S170	170	Semen	MG137465
40	L141 (F) S D170	KY84-L141F-S170	170	Semen	MG137474
41	L142 (G) S D170	KY84-L142G-S170	170	Semen	MG137481
42	L139 (D) S D345	KY84-L139D-S345	345	Semen	MG137459
43	L141 (F) S D345	KY84-L141F-S345	345	Semen	MG137475
44	L136 (A) S D380	KY84-L136A-S380	380	Semen	MG137440
45	L140 (E) S D380	KY84-L140E-S380	380	Semen	MG137466
46	L136 (A) S D548	KY84-L136A-S548	548	Semen	MG137441
47	L140 (E) S D548	KY84-L140E-S548	548	Semen	MG137467
48	L136 (A) S D726	KY84-L136A-S726	726	Semen	MG137442
49	L140 (E) S D726	KY84-L140E-S726	726	Semen	MG137468
50	Stallion D84	KY84 Stallion D84	-	Semen	MG137430
51	Stallion D94	KY84 Stallion D94	-	Semen	MG137431
52	Stallion E84	KY84 Stallion E84	-	Semen	MG137432
53	Stallion E91	KY84 Stallion E91	-	Semen	MG137433

Supplementary table 2.2. Comparative amino acid substitution analysis of both nonstructural and structural proteins between inoculum and semen viruses of EAV KY84 strain from the experimentally infected carrier stallions during acute and persistent infection periods

Open Reading Frames (ORFs)	Protein (aa length)	Amino Acid Substitutions																
		Amino acid position	O-EAV KY84	L136 (A) S D5	L136 (A) S D107	L136 (A) S D170	L136 (A) S D380	L136 (A) S D548	L136 (A) S D726	L140 (E) S D5	L140 (E) S D9	L140 (E) S D107	L140 (E) S D170	L140 (E) S D380	L140 (E) S D548	L140 (E) S D726
Amino Acid Substution:				1	1	2	5	11	21	3	1	1	4	6	7	13
ORF 1ab (255-9751)	**nsp1:** Met1-Gly260 (260)	-	-		-	-	-	-	-	-	-	-	-	-	-	-
	nsp2: Gly261-Gly831 (571)	265	Pro (P)							Leu (L)						
		411	Asp (D)						Gly (G)							
		447	Ala (A)						Thr (T)							
		458	Thr (T)													Ile (I)
		462	Ser (S)						Leu (L)							
		489	Ser (S)													Cys (C)
		509	Thr (T)						Ile (I)							
		549	Ala (A)													Thr (T)
	nsp3: Gly832-Glu1064 (233)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp4: Gly1065-Glu1268 (204)	1148	His (H)							Arg (R)						
	nsp5: Ser1269-Glu1430 (162)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp6: Gly1431-Glu1452 (22)	1439	Ala (A)					Thr (T)	Thr (T)							
	nsp7: Ser1453-Glu1677 (225)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp8: Gly1678-Asn1727 (50)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp9: Gly1678-Arg2370 (693)	226	Asp (D)											Gly (G)	Gly (G)	Gly (G)
		582	Ile (I)					Val (V)	Val (V)							
	nsp10: Ser2371-Gln2837 (467)	842	Glu (E)								Ala (A)					
		863	Ser (S)					Cys (C)	Cys (C)							
	nsp11: Ser2838-Glu3056 (219)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp12: Gly3057-Val3175 (119)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ORF2a (9751-9954)	**E** (67)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ORF2b (9824-10507)	**GP2** (227)	81	Ala (A)	Tyr (Y)					Tyr (Y)	-	-	-	-	-	-	-
ORF3 (10306-10797)†	**GP3** (163)†	3	His (H)						Tyr (Y)			Tyr (Y)	Tyr (Y)	Tyr (Y)	Tyr (Y)	Tyr (Y)
		19	Arg (R)				Cys (C)	Cys (C)	Cys (C)					Cys (C)	Cys (C)	Cys (C)
		22	Phe (F)													Ser (S)
		28	Asp (D)												Asn (N)	Asn (N)
		30	Ile (I)					Thr (T)	Thr (T)							
		84	Phe (F)						Leu (L)							
		116	His (H)													Gln (Q)
		118	His (H)				Arg (R)	Arg (R)	Arg (R)							
		119	Leu (L)		Ser (S)	Ser (S)	Ser (S)	Ser (S)	Ser (S)				Pro (P)	Pro (P)	Pro (P)	
ORF4 (10700-11158)	**GP4** (152)	14	Ala (A)										Val (V)			
		116	Lys (K)													Gln (Q)
ORF5a (11112-11291)	**ORF5a protein** (59)	13	Ala (A)						Val (V)							
ORF5 (11146-11913)	**GP5** (255)	71	Glu (E)						Asp (D)							
		73	Ile (I)			Val (V)	Val (V)	Val (V)	Val (V)				Val (V)	Val (V)	Val (V)	Val (V)
		82	Asp (D)					Asn (N)	Asn (N)							Asn (N)
		84	Tyr (Y)				His (H)	His (H)	His (H)					His (H)	His (H)	His (H)
		104	Asp (D)						Gly (G)							
		173	Leu (L)							Pro (P)						
ORF6 (11901-12389)	**M** (162)	88	Ile (I)					Met (M)	Met (M)							
		135	Lys (K)						Arg (R)							
ORF7 (12313-12645)	**N** (110)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary table 2.3. Comparative amino acid substitution analysis of both nonstructural and structural proteins between inoculum and semen viruses of EAV KY84 strain from the experimentally infected stallions that stopped shedding the virus

Open Reading Frames (ORFs)	Protein (aa length)	Amino Acid Substitutions						Amino Acid Substitutions				Amino Acid Substitutions				Amino Acid Substitutions		
		Amino acid position	O-EAV KY84	L137 (B) S D5	L137 (B) S D9	L137 (B) S D107	L138 (C) S D107	L139 (D) S D5	L139 (D) S D9	L139 (D) S D170	L139 (D) S D345	L141 (F) S D5	L141 (F) S D107	L141 (F) S D170	L141 (F) S D345	L142 (G) S D5	L142 (G) S D107	L142 (G) S D170
Amino Acid Substution:				1	1	3	4	1	1	6	15	2	6	6	10	2	1	3
ORF 1ab (255-9751)	**nsp1**: Met1-Gly260 (260)	145	Val (V)												Met (M)			
	nsp2: Gly261-Gly831 (571)	404	Ser (S)		Arg (R)													
		411	Asp (D)															Gly (G)
		470	Lys (K)													Arg (R)		
		495	Phe (F)						Ser (S)									
		497	Ala (A)													Thr (T)		
		511	Thr (T)				Ile (I)											
		705	Thr (T)				Ala (A)											
	nsp3: Gly832-Glu1064 (233)	844	Val (V)								Ile (I)							
		873	Thr (T)							Ile (I)	Ile (I)							
	nsp4: Gly1065-Glu1268 (204)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp5: Ser1269-Glu1430 (162)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp6: Gly1431-Glu1452 (22)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp7: Ser1453-Glu1677 (225)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	nsp8/9: Gly1678-Asn1727 (50)	1681	Gln (Q)	Arg (R)														
	nsp9: Gly1678-Arg2370 (693)	200	Val (V)			Ala (A)												
		441	Ser (S)							Gly (G)	Gly (G)							
	nsp10: Ser2371-Gln2837 (467)	819	Val (V)										Phe (F)					
		1039	Ser (S)			Gly (G)												
	nsp11: Ser2838-Glu3056 (219)	1142	Leu (L)				Phe (F)											
		1168	Leu (L)							Phe (F)	Phe (F)							
		1190	Thr (T)								Ala (A)							
		1234	His (H)							Tyr (Y)	Tyr (Y)							
	nsp12: Gly3057-Val3175 (119)	1379	Ala (A)												Val (V)			
ORF2a (9751-9954)	**E** (67)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ORF2b (9824-10507)	**GP2** (227)	9	Tyr (Y)							Cys (C)	Cys (C)							
		10	Leu (L)								Ser (S)							
		217	Ser (S)										Phe (F)					
ORF3 (10306-10797)†	**GP3** (163)†										Tyr (Y)							
		3	His (H)															
		19	Arg (R)							Cys (C)	Cys (C)		Cys (C)	Cys (C)	Cys (C)			
		27	Ser (S)			Cys (C)												
		28	Asp (D)												Tyr (Y)			
		118	His (H)								Arg (R)							
		119	Leu (L)								Ser (S)		Ser (S)	Ser (S)	Pro (P)			
		120	Thr (T)															Ile (I)
		121	Thr (T)								Met (M)							
ORF4 (10700-11158)	**GP4** (152)	3	Thr (T)				Ile (I)											
		145	Arg (R)									His (H)	His (H)	His (H)	His (H)			
ORF5a (11112-11291)	**ORF5a protein** (59)	8	Val (V)									Ile (I)	Ile (I)	Ile (I)	Ile (I)			
		13	Ala (A)								Val (V)							
ORF5 (11146-11913)	**GP5** (255)	6	Val (V)					Ala (A)										
		70	Asp (D)														Asn (N)	Asn (N)
		73	Ile (I)											Val (V)	Val (V)			
		82	Asp (D)												Asn (N)			
		84	Tyr (Y)								His (H)				His (H)			
ORF6 (11901-12389)	**M** (162)	66	Val (V)											Phe (F)				
ORF7 (12313-12645)	**N** (110)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary table 2.4. Serum neutralization test titers of mAbs against semen viruses of EAV KY84 strain from both experimentally and naturally infected stallions during acute and persistent infection periods

EAV Strain	Neutralizing mAbs										Control mAbs	
	5G11	6D10	7E5	10F11	1H7	1H9	5E8	6A2	7D4	10B4	3E2	12A4
L136 d5	16	16	16	<16	<16	<16	<16	32	<16	16	<16	<16
L136 d9	16	16	64	16	<16	<16	<16	64	<16	<16	<16	<!6
L136 d107	16	<16	<16	<16	<16	<16	<!6	16	<16	<16	<16	<16
L136 d170	<16	<16	<16	<!6	<16	<16	<16	32	<!6	16	<16	<16
L136 d380	32	256	512	1024	<16	<16	<16	512	<16	<16	<16	<16
L136 d548	32	64	128	512	<16	<16	<16	256	<16	<16	<16	<16
L136 d726	16	16	16	256	16	<16	<16	<16	<16	<16	<16	<16
L139 d5	<16	<16	<16	<16	<16	<16	<16	<16	<16	<16	<16	<16
L139 d9	16	<16	16	16	<16	<16	<16	16	16	16	<16	<16
L139 d107	32	32	32	32	<16	<16	<16	64	<16	<16	<16	<16
L139 d170	32	16	32	16	<16	<16	<16	64	<16	<16	<16	<16
L139 d345	16	16	32	128	<16	<16	<16	128	<16	16	16	16
L140 d5	16	<16	<16	128	64	<16	<16	128	<16	32	32	<16
L140 d9	32	32	64	32	16	<16	<16	32	<16	<16	16	<16
L140 d107	<16	<16	<16	<16	<16	<16	<16	16	<16	<16	<16	<16
L140 d170	<16	<16	<16	16	<16	<16	<16	16	<16	<16	<16	<16
L140 d380	<16	64	64	16	<16	16	<16	16	<16	<16	<16	<16
L140 d548	64	128	128	256	<16	16	<16	128	<16	<16	<16	<16
L140 d726	16	512	128	2048	16	<16	<16	1024	<16	16	<16	<16
L141 d5	32	16	32	<16	16	Ab toxic	<16	64	Ab toxic	Ab toxic	Ab toxic	<16
repeated		16	16	<16	<16	<16	<16	16	<16	<16	<16	<16
L141 d9	16	<16	16	<16	16	16	<16	32	<16	16	32	<16
L141 d107	32	64	2048	16	16	16	<16	64	16	32	32	<16
repeated		<16	16	<16	<!6	<16	<16	32	<16	<16	<!6	<!6
L141 d170	16	16	32	32	<16	16	<16	256	<16	Ab toxic	16	<16
repeated		16	16	<16	<16	<16	<16	16	<16	<16	<16	<16
L141 d345	<16	<16	16	16	<16	<16	<16	<16	<16	<16	<16	<16
L142 d5	32	16	32	Wrong	<16	32	<16	256	16	16	16	<16
repeated		<16	32	<16	<16	<16	<16	32	<16	<16	<16	<16
L142 d9	16	16	16	<16	<16	<16	<16	32	16	16	16	16
L142 d107	32	256	128	32	32	<16	<16	128	<16	<16	<16	<16
L142 d170	16	32	128	128	<16	<16	<16	64	<16	16	16	<16
EAV KY84	2048	4096	2048	2048	<16	<16	<16	<16	<16	<16	<16	<16
EAV VBS	1924	8192	1024	4096	64	<16	<16	256	<16	<16	<16	<16
EAV ARVAC	256	1024	128	1024	<16	<16	<16	<16	<16	<16	<16	<16
EAV KY84 D84	2048	2048	256	4096	<16	<16	<16	128	<16	<16	<16	<16
EAV KY84 D94	64	512	64	64	<16	<16	<16	<16	<16	<16	<16	<16
EAV KY84 E84	32	<16	<16	32	<16	<16	<16	<16	<16	<16	<16	<16
EAV KY84 E91	1024	2048	256	4096	16	<16	<16	64	<16	<16	<16	<16

Supplementary table 2.5. Serum neutralization test titers of polyclonal equine antisera against semen viruses of EAV KY84 strain from both experimentally and naturally infected stallions during acute and persistent infection periods

EAV Strain	Isolate Equine Sera		Polyclonal Equine Antisera				Control Serum
	d14	d726	Anti-KY53	Anti-KY84	Anti-ARVAC	Anti-EAVNVSL	NVSLNEG
L136 d170	1:64	1:32	1:128	1:128	1:8	<1:4	1:8
L136 d380	1:32	1:16	1:32	1:32	1:8	<1:4	1:4
L136 d548	1:8	1:16	1:16	1:16	<1:4	<1:4	<1:4
L136 d726	1:16	1:16	1:32	1:32	1:16	<1:4	<1:4
EAV KY84	1:16	1:32	1:16	1:64	<1:4	<1:4	<1:4
EAV VBS	1:128	1:128	1:64	1:64	<1:4	1:4	<1:4
EAV ARVAC	1:128	1:64	1:128	1:128	<1:512	1:32	<1:4
L139 d5	1:4	1:8	1:16	1:16	1:8	<1:4	<1:4
L139 d345	1:16	1:8	1:32	1:64	1:8	<1:4	<1:4
EAV KY84	1:16	1:32	1:8	1:16	<1:4	<1:4	<1:4
EAV VBS	1:32	1:32	1:64	1:256	<1:4	1:4	<1:4
EAV ARVAC	1:256	1:64	1:128	1:256	1:128	1:16	<1:4
L140 d5	1:4	1:16	1:8	1:8	1:4	<1:4	<1:4
L140 d170	1:4	1:8	1:8	1:8	<1:4	<1:4	<1:4
L140 d380	1:4	1:16	1:8	1:16	<1:4	<1:4	<1:4
L140 d548	1:8	1:8	1:8	1:16	1:4	<1:4	<1:4
L140 d726	1:16	1:16	1:64	1:64	1:4	<1:4	<1:4
EAV KY84	1:8	1:32	1:16	1:64	<1:4	<1:4	<1:4
EAV VBS	1:128	1:16	1:32	1:256	<1:4	1:32	<1:4
EAV ARVAC	1:16	1:32	1:64	1:64	1:256	1:8	<1:4
L141 d170	1:8	1:8	1:8	1:16	<1:4	<1:4	<1:4
L141 d345	1:16	1:16	1:8	1:32	<1:4	<1:4	<1:4
EAV KY84	1:8	1:8	1:4	1:32	<1:4	<1:4	<1:4
EAV VBS	1:256	1:128	1:128	1:64	<1:4	<1:4	<1:4
EAV ARVAC	1:256	1:128	1:128	1:256	1:128	1:8	<1:4
L142 d107	1:32	<1:4	1:16	1:8	<1:4	<1:4	<1:4
L142 d170	1:64	1:32	1:8	1:16	<1:4	<1:4	<1:4
EAV KY84	1:64	1:32	1:8	1:32	<1:4	<1:4	<1:4
EAV VBS	>1:512	1:128	1:64	1:128	<1:4	1:4	<1:4
EAV ARVAC	1:128	1:258	1:128	1:128	1:128	1:16	<1:4

Supplementary table 3.1. Primers used for reverse transcription and PCR amplification

Numbers	Primer	Polarity	Position	Nuleotide sequence (5'-3')
1	BN1P	+	1-21	GCTCGAAGTGTGTATGGTGCC
2	BD501P	+	501-520	CTTGAGCTGTTGCAACACCC
3	JQ548N	-	548--568	CTTGCCTCTTCAATGGCTAAC
4	BD800N	-	800-781	AGTTCGTTGTTCGCGGTAGT
5	JQ1100N	-	1124-1103	AGCCAACTCATCATCACACCAC
6	JQ1524P	+	1524-1542	CAGGCGCCCATCCCAGCACC
7	JQ2106N	-	2140-2121	ACAGTGCAAGGAACATAAGC
8	JQ4002P	+	4002--4021	CTGATTGATGGCTTATCCAA
9	BD4546N	-	4546-4527	GGCGGTGACACTCTCTTTCA
10	JQ4600N	-	4619-4600	GTCATCATCAGTGAGGGCAG
11	JQ6077P	+	6077-6096	GCTGACATCAGCTGTGACGC
12	JQ6577N	-		
13	JQ9745P	+	9745-9767	CGTGTGATGGGCTTAGTGTGGTC
14	BD10330N	-	10330-10311	TTGGCCTAGCGTTGGCATAA
15	BD10449P	+	10449-10468	TAATGTCACCTTCCCGTCGC
16	BD10507N	-	10507-10488	CTACAAAATCTTGCGCCGCG
17	BD10621N	-	10621-10602	AGCATACACAGTGACAGGCC
18	BD10764P	+	10764-10783	TGCACTTTCTACCCATGCCA
19	BD11148N	-	11148-11125	GTTGAGCCCAACGTACCAGT
20	BN11220P	+	11187-11207	GCCATCACATGCTTACTTCTC
21	UB11255P	+	11272-11293	GCCAATTTGCTGCGATATGATG
22	BD11284N	-	11284-11264	GCAGCAAATTGGCAATAACGC
23	BD11361P	+	11361-11380	AGTCATCACCTTCGGCACAG
24	BD11619P	+	11619-11638	TGTACAGCGGGGCTTACTTG
25	BD11800N	-	11800-11781	GCAACTGGCAAACTGGGATC
26	N11877P	+		
27	BN12707N	-	12698-12679	TTGGATCCTGGGTGGCTAATAACTACTT
28	UB12296P	+		
29	BD12322N	-	12305-12286	TTACCAACTGCGGTGTACCC

Supplementary table 3.2. List of total clinical specimens from two male donkeys and results of EAV confirmation test

Animal ID	Donkey 7634		Donkey 7635	
Clinical specimens	Virus isolation	qPCR	Virus isolation	qPCR
Epididymis 1*	-**	+ (very weak)	-	-
Epididymis 2*	+ (1 pfu each in 3 of 4 flasks)	-	-	-
Prostate 1*	+ (1-4 pfu in all 4 flasks)	-	-	-
Prostate 2*	+ (1 pfu each in 3 of 4 flasks)	-	-	-
Vas deferens 1*	+ (1.35x10E4pfu/mL in all 4 flasks)	+	-	-
Vas deferens 2*	+ (2.15x10E4pfu/mL in all 4 flasks)	+ (very weak)	-	-
Seminal vesicles 1*	+ (1 pfu in 1 of 4 flasks)	-	-	-
Seminal vesicles 2*	+ (5.2x10E1pfu/mL in all 4 flasks)	-	-	-
Conjunctival swab 1	-**	-	-	-

*1 mL of 10% tissue suspension into 4 x 25-cm^2 flasks of RK-13 (KY) cells. All virus isolations detected on 1st passage in cell culture

**Negative for virus after 3 blind passages in RK-13 (KY) cells

Supplementary table 3.3. Partial GP5 sequence identities (%) of the novel EAV donkey VD7634 strain to EAV donkey reference strains

Nucleotide position of partial GP5 located at 11,295-11,812 (518 bp in length)

Strain in GenBank with the highest similarity	Nucleotide identity (%)	Amino acid identity (%)	Origin	GenBank accession number
J1-931125	81.5	93.0	RSA	|AY956596.1|
J2-931125	81.5	93.0	RSA	|AY956597.1|
J3-931209	81.5	93.0	RSA	|AY956598.1|
J4-931209	81.5	93.0	RSA	|AY956599.1|
J7-931125	81.5	93.0	RSA	|AY956602.1|
J5-940309	81.3	93.0	RSA	|AY956600.1|
J6-940309	81.3	93.0	RSA	|AY956601.1|
1489V/96	74.7	86.6	Italy	|AF099829.1|
3308V/96	74.3	86	Italy	|AF099830.1|

*Republic of South Africa (RSA)

Supplementary table 3.4. Nucleotide and amino acid sequences of average pairwise identity of EAV donkey VD7634 compared to 29 full-length EAV strains

ORF	Protein	EAV donkey VD7634 strain nucleotide position (bp)	EAV donkey VD7634 strain protein length (aa)	Average pairwise identity of nucleotide alignment (%) n=29	Average pairwise identity of amino acid alignment (%) n=29	Average pairwise identity of nucleotide alignment (%) n=2	Average pairwise identity of amino acid alignment (%) n=2	Average pairwise identity of nucleotide alignment (%) n=5	Average pairwise identity of amino acid alignment (%) n=5	Average pairwise identity of nucleotide alignment (%) n=4	Average pairwise identity of amino acid alignment (%) n=4	Average pairwise identity of nucleotide alignment (%) n=18	Average pairwise identity of amino acid alignment (%) n=18
						EAV donkey and CW strains (EU-1)		EAV donkey and French strains (EU-2)		EAV donkwy and NA strains (NA)		EAV donkey and S numbered strains (EU-1)	
Full length	-	-	-	75.37	-	75.15	-	75.36	-	75.35	-	75.41	-
5' TTR	-	1-230	-	67.56	-	66	-	67.62	-	66.68	-	67.91	-
Leader sequence		1-211	-	-	-	-	-	-	-	-	-	-	-
ORF1a	1a polyprotein	231-5405	1724	72.91	84.12	72.55	84.2	72.72	83.88	72.95	84.05	72.99	84.19
ORF1b	-	5402-9748	1448	75.29	88.4	75.35	88.6	75.2	88.34	74.98	88.68	75.37	88.34
ORF1ab	1ab polyprotein	231-9748	3172	74	-	73.8	-	73.86	-	73.85	-	74.09	-
	nsp1	231-1010	260	78.16	88.41	77.35	89.4	89.12	77.2	77.9	87.8	78.57	88.23
	nsp2	1011-2714	568	65.55	70.72	65.6	71	65.16	69.6	65.68	71.55	65.62	70.81
	nsp3	2715-3412	232	75.74	90.19	75.7	89.7	75.18	89.2	76.65	90.3	75.71	90.5
	nsp4	3413-4025	204	74.37	37.81	73.45	35.8	75.74	40.1	73.86	40.1	74.21	37.33
	nsp5	4026-4511	162	78.37	91.86	77.05	92	76.7	92.48	77.45	90.55	79.18	91.97
	nsp6	4512-4577	22	82.47	92.33	79.55	90.9	89.56	95.5	84.8	95.5	90.9	80.3
	nsp7	4578-5252	225	76.2	93.18	76.15	94.2	76.56	94.7	75.2	92.53	76.32	92.79
	nsp8	5253-5402	50	74.48	94.7	74.7	92	75.46	91.62	78.7	94	73.24	96
	nsp9	5253-7330	693	75.38	44.62	75.25	43.8	75.34	43.34	75.65	44.58	75.34	45.08
	nsp10	7331-8731	467	75.8	90.63	75.8	89.5	75.68	90.6	75.15	90.75	75.98	90.74
	nsp11	8732-9388	219	74.83	84.95	75.55	84.95	74.56	84.68	74.08	85.7	74.99	84.86
	nsp12	9389-9742	118	73.27	85.31	73.3	86.6	73.92	85.96	73.3	85.7	73.08	84.9
	Slippary sequence	5398-5404	-		-		-		-	-	-	-	-
ORF2a	E	9748-9951	67	80.96	78.45	81.15	82.1	81.02	80.06	80.4	80.6	81.04	77.6
ORF2b	GP2	9821-10504	277	83.46	87.39	82.3	87.2	83.36	88.68	83.45	85.03	83.61	87.57
ORF3	GP3	10303-10794	163	78.17	74.21	78.75	73.3	78.28	72.24	81.08	77.9	77.44	74.03
ORF4	GP4	10697-11155	152	77.99	84.06	77.2	82.25	78.72	82.14	78.28	83.75	77.81	84.86
ORF5a	ORF5a	11109-11288	59	91.09	87.05	89.4	84.7	92.72	89.84	92.08	87.68	90.6	86.4
ORF5	GP5	11143-11910	255	78.1	86.1	78.1	86.1	78.34	84.84	77.78	84.1	78.09	85.28
ORF6	M	11898-12386	162	83.83	86.98	83.4	86.4	83.98	87.7	83.75	86.4	83.86	86.97
ORF7	N	12310-12642	110	88.97	95.67	89.05	95.95	88.68	94.04	89.28	94.28	88.97	96.4
3' TTR	-	12643-12703	-	80.92	-	81.7	-	83.3	-	81.7	-	80	-

Supplementary table 3.5. Percent homology (%) of nucleotide and amino acid sequences of EAV donkey VD7634 compare to 29 full-length EAV strains used for bar graph

(A) nsp2

EAV donkey VD7634	nt (%)	aa (%)
CW01	65.4	70.8
CW96	65.8	70.8
F27	65.4	70.8
F62	65.1	70.8
F63	65	70.8
F61	65.3	70.8
F60	65	70.8
S3583	65.7	70.8
S3817	65.7	70.8
S4216	65.7	70.8
S3685	65.6	71
S3854	65.6	71
S3711	65.5	71
S3712	65.5	71
S3699	65.7	70.6
S3961	65.7	70.6
S4333	65.6	70.6
S4227	65.5	70.8
S3886	65.6	71.7
S4421	65.6	71.7
S3943	65.7	71.5
S4445	65.7	71.3
S3861	65.6	70.1
S4007	65.5	69.5
S4417	65.7	69.3
HK25	65.7	68.8
HK116	65.6	70.3
Bucyrus	65.7	71
ARVAC	65.7	71
	1900.9	2050.8
	65.5483	70.7172

(B) GP3

EAV donkey VD7634	nt (%)	aa (%)
CW01	79.4	73.9
CW96	78.1	72.7
F27	79.4	73.7
F62	77.3	70.5
F63	77.3	71.1
F61	77.8	72.1
F60	79.6	73.8
S3583	77.7	74.3
S3817	77.7	74.3
S4216	77.7	74.3
S3685	77.5	74.3
S3854	77.5	74.3
S3711	77.1	73.7
S3712	77.3	73.7
S3699	77.3	74.3
S3961	77.3	74.3
S4333	77.7	74.3
S4227	77.6	73.7
S3886	77.5	74.3
S4421	77.5	73.7
S3943	77.5	74.3
S4445	77.5	74.3
S3861	76.9	73.1
S4007	77.3	73.1
S4417	77.3	74.3
HK25	81.2	78.2
HK116	80.9	78.2
Bucyrus	81.4	78.2
ARVAC	80.8	77.0
	2267.1	2152
	78.17586	74.2069

(C) Partial GP5

EAV donkey VD7634	nt (%)	aa (%)
RSA J7-931125	81.5	93.4
RSA J3-931209	81.5	93.4
RSA J2-931125	81.5	93.4
RSA J1-931125	81.5	93.4
RSA J6-940309	81.3	93.4
RSA J5-940309	81.3	93.4
RSA J4-931209	81.5	93.4
Italy 3308V/96	74.3	86.7
Italy 1489V/96	74.7	86.7
CW01	74.1	85.5
CW96	75.3	85.5
F27	73	84.3
F62	73.6	84.3
F63	73.9	84.3
F61	74.3	84.9
F60	74.3	84.9
S3583	74.7	84.9
S3817	75.1	85.5
S4216	74.7	85.5
S3685	74.3	84.9
S3854	74.3	84.9
S3711	74.3	84.9
S3712	74.3	84.9
S3699	74.3	84.9
S3961	74.7	85.5
S4333	74.5	84.9
S4227	74.7	84.9
S3886	74.1	85.5
S4421	74.4	85.5
S3943	74.5	84.9
S4445	74.1	84.3
S3861	74.3	84.9
S4007	73.9	83.1
S4417	74.5	84.9
HK25	73.7	83.7
HK116	73.4	82.5
Bucyrus	73.9	84.3
ARVAC	72.2	80.7
	2870.5	3280.9
	75.53947	86.33947

Supplementary table 3.6. Comparative amino acid substitutions of ORF5 of EAV donkey VD7634 to 29 full-length EAV strains and SNT result mAb that explains Figure 3. 4.

Specific amino acid substitution site in GP5	Amino acid change numbers	Amino acid substitution position	Amino acid substitution		Variant sequences of EAV strains (Description)
			EAV other strains/isolates	Donkey	
aa 49 (site A)	None		-	-	
aa 61 (site B)	1	**61**	K	E	All of EAV S numbered isolates and NA strains are very different to EAV donkey VD7634 strain
			?	E	EAV F61-63 straisn (? Is undetermined amino acid sequence from GenBank)
			Q	E	EAV F27 strain
			R	E	EAV F60 strain
			F	E	EAV CW01 isolate
			V	E	EAV CW96 isolate
aa 67-90 (site C)	13	**67**	V	T	All of EAV S numbered isolates and French strains are very different to EAV donkey VD7634 strain
			E	T	EAV S3817, S4227, S4216, S3961, S4417, S4445 and S4007 isolates
			?	T	EAV F60 strain (? Is undetermined amino acid sequence from GenBank)
			T	**T**	NA and CW strains 100% identity to EAV donkey VD7634 strain
		69	P	L	EAV ARVAC strain
		70	N	D	EAV S4417 and S4007 isolates
		71	D	E	EAV CW01 isolate
		72	K	Q	EAV ARVAC strain
		73	I	V	S3699, S3861, S3712, S3711, S3854, S3685, S3583, S3886, S3817, S4216, S4007, F60 and NA strains are very different to EAV donkey VD7634 strain
			?	V	EAV F61 strain (? Is undetermined amino acid sequence from GenBank)
			V	**V**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		79	D	G	Rest strains are very different to EAV donkey VD7634 strain
			G	**G**	EAV S3886, S4421 and CW strains are 100% identity to EAV donkey VD7634 strain
		81	D	N	EAV HK116, HK25, and ARVAC (laboratory NA strains)
		82	D	N	Rest of all strains are different to EAV donkey VD7634 strain
			N	**N**	EAV S3817, S4216, S3961 and S4417 isolates and F60-61 strains are 100% identity to EAV donkey VD7634 strain
		84	D	Y	EAV S4007 isolate
			I	Y	EAV F27, F60-63 strains and CW01 isolate
			F	Y	EAV CW96 isolate
			Y	**Y**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		85	T	S	EAV S4007 and S4445 isolates
			A	S	EAV F27, 60-63 and NA strains
			S	**S**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		89	A	S	All EAV NA strians (HK25, HK116, Bucyrus and ARVAC) are differnet to EAV donkey VD7634 strain
			S	**S**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		90	T	V	all S numbered EAV isolates are different to EAV donkey VD7634 strain
			E	V	All EAV NA strians (HK25, HK116, Bucyrus and ARVAC) are differnet to EAV donkey VD7634 strain
			?	V	EAV F62 and F63 strains (? Is undetermined amino acid sequence from GenBank)
			L	V	EAV F27 strain
			V	**V**	EAV F60, F61, CW01, CW96 are 100% identity to EAV donkey VD7634 strain
aa 98-106 (site D)	4	**100**	G	S	EAV HK116 and ARVAC strains are differnet to EAV donkey VD7634 strain
			S	**S**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		101	V	T	Rest of all strains are very different to EAV donkey VD7634 strain
			T	**T**	CW96 and CW01 strains are 100% identity to EAV donkey VD7634 strain
		104	G	D	EAV HK116 and ARVAC strains
			D	**D**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain
		106	V	M	EAV ARVAC strain
			M	**M**	Rest of all EAV strains are 100% identity to EAV donkey VD7634 strain

APPENDIX 4

Experimental Methods

Our standard laboratory protocol for commonly used tissue culture media recipes

Rabbit kidney-13 high passage and low passage from Dr. William McCollum, UKY
Laboratory designated name: RK-13 HP #399-409 and RK-13 LP #194-204

Eagle's minimum essential medium (EMEM;	500 mL
with Earle's salts and L-glutamine; GIBCO# 11095-080)	
Cosmic calf serum 10% (HYCLONE# SH30087.03)	50 mL
Penicillin and streptomycin (GIBCO# 15140-122)	5 mL

Rabbit kidney 13 (ATCC# CCL-37)
Laboratory designated name: CCL-37-RK13

Eagle's minimum essential medium (EMEM;	500 mL
with Earle's salts and L-glutamine; GIBCO# 11095-080)	
Fetal bovine serum 10% (FBS; HYCLONE# SH30396.03)	50 mL
Penicillin and streptomycin (GIBCO# 15140-122)	5 mL
1.0 mM Sodium pyruvate (100 mM [100X]; GIBCO# 11360-070)	5 mL
0.1 mM Non-essential amino acids (10 mM [100X]; GIBCO# 11140-050)	5 mL

Equine pulmonary artery endothelial cells (EPAEC)
Laboratory designated name: EPAEC, currently EEC

Dulbecco's modified Eagle medium (DMEM;	1 L
With high glucose [4.5 g/L], without L-glutamine; CELLGRO# 15-013-CM)	
Fetal bovine serum 10% (FBS; HYCLONE# SH30396.03)	110 mL
Penicillin and streptomycin (10,000 U/mL and μg/mL; GIBCO# 11140-050)	11 mL
0.1 mM Non-essential amino acids (10 mM [100x]; GIBCO# 11140-050)	11 mL
2 mM L-glutamine (200 mM [100X]; GIBCO# 25030)	10 mL

Equine derm cells (ATCC# CCL-57)
Laboratory designated name: ED (Equine derm)

Eagle's minimum essential medium (EMEM;	500 mL
with Earle's salts and L-glutamine; GIBCO# 11095-080)	
Fetal bovine serum 10% (FBS; HYCLONE# SH30396.03)	50 mL
Penicillin and streptomycin (GIBCO# 15140-122)	5 mL
1.0 mM Sodium pyruvate (100 mM [100X]; GIBCO# 11360-070)	5 mL
0.1 mM Non-essential amino acids (10 mM [100X]; GIBCO# 11140-050)	5 mL

Baby hamster kidney-21 (ATCC# CCL-10)
Laboratory designated name: BHK-21

Eagle's minimum essential medium (EMEM;	500 mL
with Earle's salts and L-glutamine; GIBCO# 11095-080)	
Fetal bovine serum 10% (FBS; HYCLONE# SH30396.03)	50 mL
Tryptose phosphate broth (TPB)	50 mL
Penicillin and streptomycin (GIBCO# 15140-122)	5 mL

Cell lines grown with EMEM 10% FSCS in incubator without CO_2

Rabbit kidney-13 high passage and low passage (RK-13 HP #399-409 and RK-13 LP #194-204) from Dr. William McCollum, UKY, Equine derm cells (ATCC Cat# CCL-57) and Baby hamster kidney-21 (ATCC Cat# CCL-10)

10 L stock cell culture media, EMEM 10% FSCS

Materials

Eagle's minimum essential medium (EMEM; NEAA, L-glutamine; MEDIATECH #50-011-PB; 10 L/bottle)	10 L/bottle
Ferritin supplemented calf serum (FSCS; HYCLONE# SH30072.03)	1,000 mL
Penicillin and streptomycin (10,000 U/mL and μg/mL; GIBCO# 11140-050)	100 mL
Fungizone (SIGMA A-9528; one vial)	10 mL
Fungizone (Amphotericin B; 1,000 μg/mL)	50 mg
Sterile distilled water	50 mL
Dissolve fungizone in dH_2O and dispense in 10 mL aliquots. Store at -20°C	
Unreconstituted vial should be stored at 4°C	
Sodium bicarbonate ($NaHCO_3$; SIGMA# S5761-500 g)	6.0 g
Glass fiber prefilter (MILLIPORE# AP2012450)	
Durapore® membrane filters (0.45 μm; MILLIPORE# HVLP14250)	

Materials in order

EMEM	96.1 g
$NaHCO_3$	6.0 g
Penicillin and streptomycin	100 mL
Fungizone (Amphotericin 1,000 μg/mL)	10 mL
Ferritin supplemented calf serum (FSCS)	1,000 mL
dH_2O (tissue culture grade water)	q.s. to 10,000 mL

Method

1. Autoclave and sterilize the containers, filter with membrane filters, hose, 12 bottles (1 L).
2. Add 8 L dH_2O to 12 L bottle wit stir bar.
3. Place on stirrer/hotplate and turn stirrer to setting #3. Make sure stir bar is standing middle in the bottom of 12 L bottle.
4. Add EMEM bottle quickly without losing powder. Rinse twice with dH_2O. Let it mix well until dissolved in about 15 min.
5. Add sodium bicarbonate (color should change from yellow to red), penicillin and streptomycin, fungizone and serum (rinse each container twice with dH_2O).
6. Final concentration of penicillin and streptomycin 100 U/mL and 100 μg/mL and fungizone 1 μg/mL.
7. Instead of #4-5, anti-anti (1 bottle [100X GIBCO# 15240]) can be used.
8. Add dH_2O to 10 L and mix 5 min.
9. In cell culture room, filter through sterile Millipore filter unit using a prefilter. Dispense 1 L/sterile 1 L bottle.
10. Perform sterility check. Store at 4°C.

Carboxymethylcellulose (CMC) overlay media (1.3X EMEM 10% FSCS)

Materials

Carboxymethylcellulose, medium viscosity (SIGMA# C4888)	6.0 g
Distilled water (dH_2O)	140 mL

Method 1

1. Add 140 mL of dH_2O to a 1 L bottle (12), swirl while adding 0.75% carboxymethylcellulose (CMC) and then shake if necessary.
2. Let stand (12 bottles) overnight at room temperature (shake occasionally).
3. Autoclave for 20 min at 121°C (liquid setting).
4. When cooled to approx. 40°C, add filtered 1.3X EMEM 10% FSCS to the 800 mL mark.

10L stock CMC overlay media (1.3X EMEM 10% FSCS)

Materials

EMEM	96.1 g
$NaHCO_3$	6.0 g
Penicillin and streptomycin	100 mL
Fungizone (Amphotericin 1,000 μg/mL)	10 mL
Ferritin supplemented calf serum (FSCS)	1,000 mL
dH_2O (tissue culture grade water)	q.s. to 7,800 mL

Method 2

Prepare **method 1** a day before making the CMC stock

Prepare **EMEM 10% FSCS** by following the protocol except start with 6 L dH_2O and q.s. to 7.8 L

5. Shake the bottle to dissolve cooled carboxymethylcellulose dilution.
6. Perform sterility check. Store at 4°C.
7. Before using, add 0.8 mL gentamicin/bottle (working concentrationg – 50 μg/mL).

Virus transport medium

Hanks's balanced salt solution with phenol red 1X (INVITROGEN# 14170112)	500 mL
HEPES buffer (final concentration 25 mM; INVITROGEN# 15630080)	20 mL
Bovine serum albumin V (final concentration 0.5%; SIGMA# A9647)	2.5 g
Antibiotic-Antimycotic (100X, Penicillin, Steptomycin & Amphotericin B; INVITROGEN# 15240112)	5 mL
Gentamicin sulfate (final concentration 250 mg/L; CELLGRO# 30-005-CR)	2.5 mL

Freezing media

Dimethylsulfoxide (DMSO; SIGMA# D-2650)	10%
EMEM 10% FSCS	90%

Method

5 mL of DMSO + 95 mL of complete medium (depending on which cell lines)

Generally, 1 x 10^6 cell/mL DMSO/1vial

Or
Recovery™ cell culture freezing medium (GIBCO# 12648-010)
Cells from T-75 = averages 8.4 x 10^6 cells/mL DMSO/1vial, makes 5 vials

Dublecco's phosphate-buffered saline (PBS)

Materials

PBS (GIBCO# 21600-069)	95.5 g (1 bottle)
Sterile distilled water (dH_2O)	q.s. to 10,000 mL

Method

1. Add 9 L dH_2O to 12 L bottle with stir bar. Place on stirrer/hotplate and turn stirrer to setting #3. Make sure stir bar is standing middle in the bottom of 12 L bottle.
2. Add PBS bottle quickly (rinse twice with water) and mix for 10 min.
3. Distribute evenly into each 1 L bottle 1 L for about 800ml. Autoclave the PBS contained bottles for 20 min at 121°C (liquid setting).
4. Cool bottles down to room temperature before tightening caps and labeling. Perform sterility check. Store at room temperature.

Trypsin EDTA solution (0.5%)

Materials

PBS (sterile 1X)	180 mL (900 mL)
Trypsin EDTA (no phenol red [10X]; GIBCO# 15400-054)	20 mL (100 mL)

Aliquot out 45 mL in 50 mL conical tube and store at -20°C (for long term) and store at 4°C (for current use).

Guinea pig complement

Material

100 mL/bottle (ROCKLAND# C300-0100).

Method

Thaw and dispense (on ice) into 1 or 2 mL volumes and store at -70°C.

Antibiotics for cell culture

Gentamicin MEDIATECH# 30-005-CR

Gentamicin sulfate, 50 mg/mL, 10 mL/vial and store at 4°C

Fungizon SIGMA A-9528

Amphotericin B (1,000 μg/mL)
Aliquot out 11 mL in 15 mL conical tubes and store at -20°C

Penicillin/Streptomycin GIBCO# 11140-050

10,000 μg/mL Pen and 10,000 μg/mL Strep
Aliquot out 11 mL in 15 mL conical tubes and store at -20°C

Antibiotic-Antimycotic (Anti-Anti 100X) GIBCO# 15240

L-Glutamine 100X GIBCO 25030
Aliquot out 11 mL in 15 mL conical tubes and store at -20°C

MEM NEAA 100X GIBCO 11140 MEDIATECH #50-011-PB
Aliquot out 11 mL in 15 mL conical tubes and store at -20°C

Miscellaneous
70% Ethanol

95% Ethanol	dH_2O	Total
147 mL	53 mL	200 mL
368 mL	132 mL	500 mL
2,944 mL	1,056 mL	4,000 mL

Roccal
Maintain at 1:1200 dilution (20 mL of Roccal : 4 L of dH_2O)

Our standard laboratory protocol for commonly used bacterial media recipes
Liquid media
LB Medium (Luria-Bertani Medium; 500 mL)
Materials

Deionized water	500 mL
Bacto-trypto	5 g
Bacto-yeast extract	2.5 g
NaCl	5 g

Method
1. Stir until the solutes have dissolved. Adjust the pH to 7.0 with 5N NaOH.
2. Sterilize by autoclaving for 20 min on liquid cycle.

Psi Broth (500 mL)
Materials

Bacto-trypto	5 g
Bacto-yeast extract	2.5 g
NaCl	5 g
5 mM MgSO4	
10 nM KCl	
dH_2O	to 500 mL

Autoclave liquid cycle 30 min.

Media containing agar
LB Agar plates
Materials

Deionized water	500 mL
Bacto-trypto	5 g
Bacto-yeast extract	2.5 g
NaCl	5 g

Method

1. Prepare LB medium according to the above recipe. Just before autoclaving, add 7.5 g of Bacto agar. Sterilized by autoclaving for 20 min on liquid cycle.
2. When the medium is removed from the autoclave, swirl it gently to distribute the melted agar throughout the solution (need to be very careful, the media may bio over when swirled). Allow the medium to cool to 45-50°C before adding antibiotic*.
3. To avoid air bubbles, mix the medium by swirling. Pour 20-25 mL of medium in a petri dish (use the tissue culture hood and florescent light should be off).
4. When medium has hardened completely, invert the plates and wrap in aluminum foil and store them at 4°C until needed. The plates should be removed from storage 1-2 hours before they are used and allow them to dry.

*Antibiotics solutions

Antibiotic	Working concentration
Ampicillin	50 μg/mL
Streptomycin	50 μg/mL
Tetracycline	50 μg/mL
Kanamycin	50 μg/mL
Carbenicillin	50 μg/mL
Chloramphenicol	170 μg/mL

Stock solutions should be stored at -20°C. Avoid repeated freezing thawing and exposure to light. Magnesium ions are antagonists of tetracycline. Use media without magnesium salts for selection of bacteria resistant to tetracycline.

Our standard laboratory protocol for commonly used solutions and buffers for molecular biology

0.5 M EDTA (pH 8.0)

Deionized water	115 mL
EDTA (di-sodium)	27.918 g
5N NaOH	15,825 μL

Bring the volume to 150 mL with deionized water

5.0 M Sodium chloride

NaCl	29.22 g
Distilled water	100 mL

1M Tris HCl(pH 8.0)

Tris HCl	Tris base	Molarity
4.44 g/1 mL	2.65 g/1 mL	0.05 M
88.8 g/1 mL	53.0 g/1 mL	1.0 M
8.88 g/100 mL	5.30 g/100 mL	1.0 M
17.76 g/200 mL	10.6 g/200 mL	1.0 M

1M Tris HCl (per liter)

Tris base	121.1 g
Add deionized water to 800 mL	
Desired pH	Volume of HCl
pH 7.4	70 mL
pH 7.6	60 mL
pH 8.0	42 mL
Add deionized water to 1 L	

TE buffer

pH 8.0	**pH 7.6**	**pH 7.4**
10 mM Tris HCl (pH 8.0)	10 mM Tris HCl (pH 7.6)	10 mM Tris HCl (pH 7.4)
1 mM EDTA (pH 8.0)	1 mM EDTA (pH 8.0)	1 mM EDTA (pH 8.0)

TAE (Tris-acetate; 50X concentrated stock solution [per L])

Tris base	242 g
Glacial acetic acid	57.1 mL
0.5 M EDTA (pH 8.0)	100 mL
Deionized water	q.s.
Working concentration 1X (dilute 10 mL in 490 mL of deionized water)	

TAE (50X to 1X)

50mL of 50X TAE buffer + 1960 mL dH_2O = 2 L

TAE (10X to 1X)

200 mL of 50X TAE buffer + 1800 mL of dH_2O = 2 L
100 mL of 50X TAE buffer + 900 mL of dH_2O = 1 L
50 mL of 50X TAE buffer + 450 mL of dH_2O = 500 mL

Ethidium bromide (EtBr; 10 mg/mL)

Method

1. EtBr in agarose: 5.0 µL of EtBr (10 µg/µL) solution per 100 mL of agarose.
2. EtBr staining solution (1XTAE with 0.5 µg/mL EtBr): 25 µL of EtBr stock solution in TAE buffer.

25 bp ladder (INVITROGEN# 10597-011)

25 bp DNA ladder (1 µg/µL)	50 µL
6X loading buffer	100 µL
Nuclease free water	450 µL

100 bp ladder (INVITROGEN# 15628-019)

100 bp DNA ladder (1 µg/µL)	50 µL
6X loading buffer	100 µL
Nuclease-free water	450 µL

1 Kb ladder (INVITROGEN# 15615-024)

1 Kb plus DNA ladder (1 μg/μL)	50 μL
6X loading buffer	100 μL
Nuclease-free water	450 μL

PCR products

PCR products	10 μL
6X loading buffer	2 μL

PCR products to estimate concentration

PCR products	2 μL
6X loading buffer	2 μL
Nuclease-free water	8 μL

Our standard laboratory protocol for commonly used stains

Crystal violet stain (10X stock solution)

Crystal violet (MALLINCKRODT# 8839)	12 g
Methanol (FISHER# A412-20)	600 mL

Method

Large volume of crystal violet (10X stock)

1. Measure 1200 mL of methanol, add 24.0g of crystal violet
2. Stir well with a pipet. Aliquot equally 400 mL to 1 L bottles (3)
3. Store at room temperature

Crystal violet stain (working solution)

1 part of crystal violet (stock solution)	100 mL
9 parts of 10% buffered formalin (FISHER 23-245-685)	900 mL

Our standard laboratory protocol for commonly used antibodies

EAV specific antibodies

Monoclonal antibodies	**Protein specificity/epitope location**
5G11	GP5/102 and 104 [D]
6D10	GP5/99 [D]
7E5	GP5/102 [D]
10F11	GP5/ND [D]
1H7	GP5/49 [A]
1H9	GP5/69 [C]
5E8	GP5/61 [B]
6A2	GP5/69 [C]
7D4	GP5/69 [C]
10B4	GP5/62-101 deleted [C and D]
Control monoclonal antibodies	
3E2	N
12A4	nsp1

Polyclonal equine sera

Anti-ARVAC	NA
Anti-VBS	NA
Anti-KY84	NA
Anti-CW96	NA
Anti-EAV NVSL positive	NA
Anti-EAV NVSL negative	NA

Our standard laboratory protocol for commonly used for tissue culture techniques

Maintenance of continuous cell line

Materials

Trypsin EDTA (keep in the water bath for few min at 37°C)
Confluent T-75 flask of cells
Appropriate cell culture medium

Method

1. Check the cell under the microscope to verify cells are confluent and ready to be passaged (3-7 days old).
2. Pour off media from the flask.
3. Rinse cells 2-3 times with PBS (pH 7.4).
4. Add 1 mL (25-cm^2 flask; T-25), 2 (3) mL (75-cm^2 flask; T-75), 4 (5) mL (150-cm^2 flask; T-150) trypsin-EDTA, and rock gently for 30 sec, then let the flask(s) sit at room temperature for 1 min.
5. Pour off trypsin-EDTA (optional) and incubate the flask(s) another 7-8 min at 37°C.
6. When cells are visualized on the bottom of flask, bang the flask hard to remove cells.
7. (Option) Add media (T-25: 5 mL; T-75: 10 mL; T150: 20 mL) into flask and aliquot out into conical tube. Centrifuge it at 800 rpm for 5 min to get the cell pellet. Throw the media out in the conical tube and break the pellet cell by collide with pipet tube. Add fresh EMEM media (T-25: 8 mL; T-75: 10 mL; T-150: 20 mL) and mix it well to evenly distribute 1:5 dilution of the cell into new flask. Fill up the fresh EMEM media to appropriate amount of media up to 10 mL (T-25), 30 mL (T-75), 60 mL (T-150).
8. Add appropriate amount of media (EMEM 10% FSCS for 1:5 split) up to 10 mL (T-25), 30 mL (T-75), 60 mL (T-150).

Once a week split cells: 150-cm^2 flask with confluent RK-13 cells is made up to approx. 300 - 335 mL EMEM 10% FSCS.
10 mL (T-25), 30 mL (T-75), 60 mL (T-150)
Note: to get a confluent monolayer within 24-36 hours transfer 2 mL of cell suspension into a T-75 flask and add 20 mL of complete medium to the flask (1:5 split).
Generally one T-75 flask is sufficient to seed five 96-well plates for neutralization or microtitration assays.
For specific cell lines: One T-75 flask of RK-13 cells (confluent) is sufficient within 3-4 days for 96-well plates.

Freezing cells

Materials

Freezing medium
Cells in log phase of growth
Trypsin EDTA (keep in the water bath for few min at 37°C)
Nalgene™ Cryo 1°C freezing container (NALGENE# 5100-0001)

Method

1. Pour off the media.
2. Add 3-4 mL of trypsin EDTA solution and rinse the cells (this is to remove the fetal bovine serum).
3. Add 3-4 mL of fresh trypsin EDTA solution and leave it for 3-4 min (preferably at 37°C) and bang on the flask to detach the cells from the surface.
4. Add 10-12 mL of complete medium (the medium contains 10% fetal bovine serum and this will inhibit trypsin) and pour into a 50 mL conical tube.
5. Centrifuge at 1,100 or 1,300 rpm (200 x g) in a tabletop centrifuge for 5 min at 25°C.
6. Aspirate the supernatant and resuspend the cell pellet in freezing medium at approx. 10^7 cells/mL.
7. Freeze cells first in Nalgene™ Cryo 1°C freezing container with isopropanol at -80°C for 1 day.
8. Transfer frozen cell vials to liquid nitrogen container.

Recovery of frozen cells

Materials

Frozen cells (-70°C) in aliquots
Growth medium

Method

1. Rapid thawing of the cells is necessary. Thaw the cells in the water bath at 37°C, by swirling the tube (do not submerge the lid to avoid contamination of the cells).
2. Resuspend in 15 mL of growth medium and centrifuge at 1,200 or 1,300 rpm (200 x g) in a tabletop centrifuge for 5min at 25°C.
3. Aspirate the supernatant and resuspend the cells in 10 mL of growth medium. Transfer the cell suspension into a T-75 flask and add 10 mL of medium. Incubate at 37°C for 2-3 days.

Our standard laboratory protocol for commonly used classical virology techniques

Processing of nasal swabs for virus isolation

Materials

10 mL syringes
Autoclaved plastic forceps
0.45 µm syringe filter CORNING#431220 or MILLIPORE Millex#SLHA033SB
1.7 mL screw cap tubes SARDSTEDT# 72.694.996 (color) #72.694.006 (clear)
15 mL conical tubes

Method

1. Thoroughly swab the nasopharyngeal area with a sterile rayon swabs (1/2" x 1") with plastic shafts (16").
2. After collection, the rayon tip of each swab is cut off into 7 mL of virus transport medium (Hank's balanced salt solution with penicillin, streptomysin, gentamicin sulfate and amphotericin B; LIFE TECH).
3. Squeeze the swab to recover maximum amount of media and centrifuge at 300 x g for 10 min to remove debris.
4. Transfer individual nasal swabs in transport medium to the barrel of a 12 mL disposable syringe and express through 0.45 μm syringe filter (Millipore).
5. Collect the filtrate into a 15 mL conical tube. Bring up the sample volume to 7 mL with plain MEM to normalize samples (e.g. after filtration usually approx. 6 mL is recovered). Leave approx. 2 mL for virus isolation and aliquot the remaining filtrate (approx. 1 mL x 2) into tubes. Freeze at -70°C until tested.
6. Aspirate off media from confluent RK-13 cells in 6-well plates and inoculate 500 μL specimen/well in duplicate.
7. Incubate plates at 37°C in presence of 5% CO_2 for 1-2 hours.
8. Add 10 mL CMC overlay to each flask and incubate at 37°C in the presence of 5% CO_2 for 5 days.
9. Do second passage on day 4 using the same cells. Stain cells on day 5 with crystal violet. Archive the TCF at -80°C.

Processing of buffy coat for virus isolation (without gradient purification)

Method

1. Draw blood by venipuncture into vacutainer tubes containing sodium citrate or EDTA and mix thoroughly.
2. Allow blood to stand for 1-2 hours or sediment at 600-700 rpm (100 x g) for 10 min at 4°C. Remove layer and sediment at 2,000 rpm for 10 min at 4°C. The few red blood cells in pellet cause no apparent concern.
3. Aspirate off supernatant; resuspend pellet in 5 mL EMEM 10% FSCS and vortex thoroughly. If desired, pellet can be washed with PBS before resuspending in EMEM media. Freeze or inoculate.
4. Process as in steps 5-8 above.
   For buffy coat from the EDTA tube
   1. Centrifuge 700 rpm for 10 min at 4°C. You will see plasma, white blood cells and red blood cells from top to bottom in the tube.
   2. Centrifuge 2,000 rpm for 10 min at 4°C.
   3. Remove yellow top plasma and save white blood cells at the bottom (pellet cell).
   4. Add 5 mL of media, mix it well.
   5. Aliquot out 0.5 mL for DNA extraction and save rest.
   6. 1 mL goes for RK-13 LP, RK-13 HP and EEC cells for virus isolation.

Collecting PBMC from buffy coat

Peripheral blood mononuclear cell (PBMC)

1. Spin the sample at 700 rpm for 10 min (1st).
2. Remove plasma only.
3. Spin plasma for 1,900 rpm for 10 min (2nd).
4. Pour out the plasma. Break and mix PBMC (pellet cell) with MEM.
5. Aliquot out in 1.5 mL tube.
6. Freeze at -80°C.

Serum processing

Materials

Vacutainer needles
Vacutainer adapters (TERMUNO, VENOJECT II# P-1316R)
15 mL BD vacutainer tubes (KENDALL MONOJECT [no additive]# 8881 301819)
1.7 mL screw cap tubes (SARDSTEDT# 72.694.006)
Transfer disposable pipets for serology (SAMCO SCIENTIFIC# 202-1S)

Methods

1. Collect blood into vacutainer tubes via jugular venipuncture (about 15 mL blood).
2. Transport blood tubes to the lab. Spray tubes with diluted Roccal-D (1:200). Rinse tubes and dry with paper towel. Let stand at room temperature for 4-5 hours or for 1 hour at 37°C (water bath works faster).
3. Sediment blood cells at 2,000 rpm (300 x g) for 10 min.
4. Collect serum and aliquot into tubes (1.4 mL x 4). Freeze at -20°C until tested for virus or antibody.
5. Process as in steps 5-8 above.

Virus isolation from semen

Protocol

1. Thaw 15 mL tube containing semen sample in warm tap water, then place on ice.
2. Label 8 flasks (T-25) [4 RK-13 (KY) "high passage" and 4 RK-13 (ATCC-CCL37)"low passage"]/semen, $2x10^{-1}$, $1x10^{-2}$, $1x10^{-3}$ for high passage and $2x10^{-1}$, $1x10^{-2}$, $1x10^{-3}$ for low passage.
3. Sonicate (output 20 #4.5 on dial) samples on ice 3 x 15 sec; let vials set on ice for 15 sec. 1 min between sonications (dilute small volumes to 5 mL). Clean sonicator probe with 95% EtOH 2X between each sample. Spin 300 x g (1,900 rpm) for 10 min to remove sperm and debris, save and use supernatant (SSP).
4. Set up 3 tubes (10^{-1} to 10^{-3}) for each sample, 4.5 mL EMEM 10% FSCS per tube.
5. Aspirate off media from T-25 (3-6 day old RK-13 monolayers). Do not rinse. Make 10^{-1} (0.5 mL sample in 4.5 mL diluent), vortex; with new pipet add 0.5 mL of 10^{-1} to tube #2 (10^{-2}) and 1 mL to each of 4 flasks marked 10^{-1}. Continue the same way with dilutions 10^{-2} and 10^{-3} (2 flasks each).

10^{-1}	10^{-2}	10^{-3}
(2.7) 4.5 mL EMEM	(2.7) 4.5 mL EMEM	(2.7) 4.5 mL EMEM
(0.3) 0.5 mL SSP	(0.3) 0.5 mL SSP	(0.3) 0.5 mL SSP

6. Incubate all flasks at 37°C in air incubator (incubator without CO_2) with caps closed for 1 hour (2 hours maximum).
7. Add 10 mL of CMC overlay to each flask and incubate at 37°C in air incubator (incubator without CO_2) with caps closed.
8. Check for CPE on day 4. If bacterial contamination occurs, redo after bacteriological examination and sensitivity testing to find out which antibiotics to used in the tissue culture media. Repeat test as above and also note contamination. Specimen may be filtered through a 0.45 μm filter using 5 or 10 cc syringe, if necessary.
9. If negative, a second passage is performed on day 4 (2^{nd} passage).
10. If sample is toxic, this material is passaged as is (2^{nd} passage).
11. If positive, supernatant is usually harvested from a 10^{-2} flask showing viral activity (store at -20°C) and a reverse SNT is performed against known positive and negative sera (see reverse SNT).
12. After 5 days harvest supernatant from flasks and stain flasks with crystal violet using 2 mL crystal violet (working stock). Stain 1-24 hours, remove stain, rinse flasks with tap water and read.
13. All semen samples are placed in freezer zip lock storage bags and stored at -70°C.
14. Permanently store negative SSP at -20°C and positive SSP at -70°C.

Second passage

1. Aspirate media from 2 "high passage" and 2 "low passage" confluent 3-5 day old RK-13 cells in T-25 flasks for each negative semen sample.
2. From 2 flasks of 10^{-1} passage #1 take up 0.5 mL from each with same pipet (1 mL) and add 1 mL to the new flask for each high and low passage. Do the same for 10^{-2} passage #1, taking up 1 mL from each. Do 10^{-1} to 10^{-2} only. Close caps and incubate at 37°C for 1-2 hours (minimum 1 hour) in air incubator.
3. Add 10 mL of CMC overlay to all flasks, close caps, and incubate at 37°C in air incubator (without CO_2).
4. Read after 3-4 days.
5. After 4 days, harvest media and stain flasks (see passage #1).

<u>Preparation of working virus stock</u>

Materials

EAV
Five confluent monolayers of RK-13 cells in T-150 flasks
EMEM
RK-13 medium
50 mL conical centrifuge tubes
Freezing vials

Method

1. Inoculate monolayers with EAV at an m.o.i. of 1. Resuspend the virus in approximately 50 mL of serum-free MEM and add 10 mL per T-150 flask.
2. At the same time mock infect one T-75 flask of RK-13 cells.
3. Adsorb for 1 hour at 37°C.

4. Add 40 mL of complete RK-13 medium / T-150 flask. Incubate at 37°C for 48 hours or until 95-100% CPE.
5. Centrifuge at 5,000 g (or 3,000 rpm in the tabletop centrifuge) at 4°C for 15 min.
6. Make 1.8 mL and 0.5 mL aliquots of supernatant and store at -70°C.
7. Use one 0.5 mL frozen vial to titrate the virus.

Preparation of high titer virus stock

Method

1. Seed cells into 3, triple-deck flasks and grow until confluent.
2. Remove the medium and leave approx. 50 mL of complete medium contaminating 10% fetal bovine serum. Add 500 μL/mL of virus per flask. If the titer of the virus is known use an m.o.i. of 1 (higher m.o.i. will may produce deletion mutants).
3. Adsorb for 1 hour at room temperature on the orbital shaker.
4. Add 20-40 mL of complete medium containing 10% fetal bovine serum and incubate at 37°C for 3-4 days. Check for signs of infection 2 days after the start of infection.
5. Pellet down the cellular debris (3,000 rpm/4°C/15 min) and aliquot into freezing vials and store at 4°C (stable for up to 6 months) or -70°C (for long term storage). Virus should be stored away from the fluorescent lights since the virus titer appears to decrease when exposed to fluorescent light for a prolonged period of time.
6. Determine the virus titer using the plaque assay procedure.

Microtitration assay for estimation of virus titers

Materials

96-well microtitration plates (FALCON#3072)
Sterile snap cap tubes (FISHER or FALCON#2054)
Multi channel pipette (8 or 12 channels)
Antiserum or ascetic fluid (monoclonal antibodies)
Negative control serum or control ascetic fluid
EMEM (BIOWHITTAKER)
Complete RK-13 medium
Confluent RK-13 cells (one T-75 flask is sufficient for five 96-well plates)
Virus stock to be titrated
Crystal violet staining (working solution)

Method

1. Each test virus is done in triplicate. Label the 96 well plates with cell controls.
2. Dilute the test virus from 10^{-1} to 10^{-8} in EMEM (10-fold serial dilution). Use 200 μL of virus and 1800 μL of EMEM.

10^{-1}	10^{-2} ...Continue...	10^{-8}
1800 μL EMEM	1800 μL EMEM	1800 μL EMEM
200 μL Virus	200 μL Virus	200 μL Virus

3. Add 50 µL of EMEM to all the wells except to control wells (column 10, 11 and 12) where 100 µL of EMEM should be added.
4. Trypsinize a confluent T-75 flask or RK-13 cells and resuspend the pellet in 10 mL of RK-13 cell medium. Take 2 mL of the cell suspension and resuspend in 10 mL (1:5) of RK-13 medium (or 4 mL in 20 mL would be better. This gives enough volume to work).
5. Add 50 µL of diluted virus to test wells in triplicate. Start with the highest dilution (10^{-8}) and go to the lowest dilution (10^{-1}; therefore, can use the same pipet tip).
6. Add 100 µL of resuspended RK-13 cells to each well. Incubate at 37°C for four days.
7. Feed the plates with 50 µL of growth media on day 2 and day 3 (if necessary) to maintain monolayers.
8. When CPE occurs (usually 4 days), dump out the media into infectious waste container and stain with crystal violet (1X working solution) for 20-30 min and then wash gently with tap water to remove excess crystal violet (three times). Air dry plates and determine titer by Reed and Munch method.
9. 50% end point values, by Reed and Munch Method. Underlined numbers indicate logarithmic characteristic of end point dilution. In triplicates.

3 – <u>2</u> – 0 – 0 = 0.25	3 – <u>1</u> – 2 – 0 = 0.5
3 – <u>2</u> – 1 – 0 = 0.5	3 – 1 – <u>3</u> – 0 = 0.2
3 – 2 – <u>2</u> – 0 = 0.0	<u>3</u> – 0 – 0 – 0 = 0.5
3 – 2 – <u>3</u> – 0 = 0.3	<u>3</u> – 0 – 1 – 0 = 0.7
<u>3</u> – 1 – 0 – 0 = 0.75	<u>3</u> – 0 – 2 – 0 = 0.8
3 – <u>1</u> – 1 – 0 = 0.0	3 – <u>0</u> – 3 – 0 = 0.0

10. Report virus titers as $TCID_{50}$ per 50 µL.

<u>Plaque assay</u>

Materials

6-well plates
RK-13 cells
Carboxymethylcellulose (CMC) overlay media (1.3X EMEM 10% FSCS)
Virus for titration

Method

1. Prepare RK-13 cell in 6-well plates
2. Make virus dilution (10^{-1} to 10^{-8}) in complete 10% EMEM growth media.
3. Aspirate the medium and add 200 µL of diluted virus to each well in duplicate (start from 10^{-8} and go to 10^{-1}). Mix gently by rocking the plate.
4. Incubate the plate at 37°C for 1 hour to allow virus particles to adsorb into cells.
5. Following 1 hour incubation period, overlay cells with 4 mL (per well) of CMC overlay media adding to the side of the plate.
6. Allow plates to sit undisturbed on a leveled surface for 4 days until visible plaques develop.
7. After 4 days, discard supernatant and stain with crystal violet staining solution.
8. Count >10 to <100 plaques per well.
9. PFU/mL – (number of plaques in each well [duplicate]/2) x (highest dilution that gives number of plaques) x (dilution factor).

10. E.g. [(25 + 30)/2] x 10^{-5} x 5 (200 µL) = 137.5 x 10^{-5} = 1.37 x 10^{-7} PFU/mL.

Our standard laboratory protocol for commonly used serology technique

Serum neutralization test (SNT)

Method

1. Inactivate sera at 56°C for 30 min.
2. Label plate, tube after calculate working dilution depends on result of 200 $TCID_{50}$ titration. Ex) 50 µL x 96 well plate = 4800 µL = 5 mL / plate needed
   Ex) Virus: $10^{5.5}$ 200 $TCID_{50}$ / 50 µL
   200/1E5.5 (log10) -3.19 (round up to -3.20) = 1E3.20 = 1584 = 1/1600
3. Dilute the virus to 1/1600 for total volume: 30 mL (4 plates)

1/100	(1/10)	1/100	(1/4)	1/400	(1/2)	1/800	(1/2)	1/1600
200 µL	Virus	1 mL		5 mL		10 mL		15 mL
1800 µL	MEM	9 mL		15 mL		10 mL		15 mL

4. Prepare back titration (virus working dilution [w])

10^{-1}	10^{-2}	10^{-3}
100 µL Virus	900 µL	900 µL
900 µL MEM	100 µL	100 µL

5. Label 4 plates, depends on number of antibodies. Antibody dilution starts at 1:4. Duplicate the antibody and virus. Triplicate may used including toxicity control.

	A	B	C	D	E	F	G	H
1	1:16							
2	1:32							
3	1:64							
4	1:128							
5	1:256							
6	1:512							
7	1:1024							
8	1:2048							
9	1:4096							
10	1:8192							
11	1:16384							
12	1:32768							

	A	B	C	D	E	F	G	H
1	Back titration				Cell control			
2	w	-1	-2	-3				
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

6. Add 50 µL of 1XMEM to plate, except add 100 µL on cell control by using multichannel pipet (6 or 12 channels).
7. Add 50 µL acitis (Ab) by single channel pipet individually. Then, mix 10 times by multichannel pipet.
8. Add 50 µL virus (final working dilution).
   a. Pour working dilution stock into reservoir.
   b. Load with multichannel pipet.
9. Incubate 30 min at 37°C in CO_2 incubator.
10. During incubation, prepare cells (use 10% EMEM FSCS).
    a. 30 mL for 1 T-75 flask will make 4 plates.
    b. 40 mL for 4 T-25 flask will make 4-5 plates.
11. Pour cells in reservoir. Use multichannel pipet.
    a. Add 50 µL to plates.
    b. Except, add 100 µL to cell control.
12. After 24 hour incubation, add 50 µL of EMEM FSCS to all plates.
13. Stop assay when working dilution (w) is showing 100% CPE.

Virus neutralization test (VNT)

Method

1. Inactivate sera at 56°C for 30 min.
2. Put 25 µL of serum diluent (10% GP complement in EMEM 10% FSCS [keep on ice until needs]) in all wells.
3. Put 25 µL of test serum to the first well of three rows (including row for checking toxicity*). Dilute with 25 µL multichannel pipet.
4. Add 25 µL of virus (200 $TCID_{50}$) diluted in EMEM 10% FSCS to all wells except control ("see cell controls").
5. Incubate plates at 37°C (5% CO_2 incubator). Read after 48 hours, and do final read at 72 hours.

"Cell controls": positive or negative sera and virus titration controls

Cell controls: add 50 µL serum diluent to 2 wells.

Serum controls: add 25 µL serum diluent to 1 well/each test serum, add 25 µL test serum. Known positive or negative serum: set up in same way as test sera.

Virus titration

Add 25 µL serum diluent into each of 4 wells for cell controls (Row#1).

Add 25 µL of working virus dilution to each of 4 wells (Row#2).

Add 25 µL of 1:10 dilution of working virus in EMEM 10% FSCS to each of 4 wells (Row#3).

Add 25 µL of 1:100 dilution of working virus in EMEM 10% FSCS to each of 4 wells (Row#4).

Add 25 µL of 1:100 dilution of working virus in EMEM 10% FSCS to each of 4 wells (Row#5).

Add 125 µL of trypsinized cells to all 20 wells.

*If serum is toxic in serum control well (1:4 dilution), then repeat test with a row to check for toxicity.

Reverse SNT for EAV verification – confirmation test

Materials

Known positive serum, inactivated and diluted 1:4 in PBS (should come from field sample(s), not from vaccinated horses)

Known negative serum inactivated and diluted 1:4 in PBS

Stock virus (Virulent Bucyrus strain)

Unknown virus isolate

Method

1. Add 0.3 mL of known negative serum to 4 tubes for each virus sample to be tested (including stock virus).
2. Add 0.3 mL of known positive serum to 3 tubes for each virus sample to be tested (including stock virus).
3. Make up 1:10 dilution of guinea pig complement in EMEM 10% FSCS with antibiotics (10 mL for each sample).

4. Make up 10^{-1} to 10^{-5} dilutions of all virus samples to be tested using 1.8 mL of media in #3 and 0.2 mL of virus (It may be necessary to dilute stock virus to 10^{-3} before starting these dilutions).
5. To positive serum (3 tubes), add 0.3 mL of 10^{-1}, 10^{-2} and 10^{-3} virus dilutions.
6. To negative serum (4 tubes), add 0.3 L of 10^{-3}, 10^{-4} and 10^{-5} virus dilutions.
7. Incubate serum-virus mixture for 1 hour at 37°C.
8. 14 of T-25 flasks [7 RK-13 (KY) "high passage" and 7 RK-13 (ATCC-CCL37) "low passage"] are used for each virus to be tested. Aspirate off media from 3-5 day RK-13 flasks and add 0.25 mL of above serum-virus mixture to 2 flasks/dilution.
9. Incubate for 2 hours at 37°C (closed caps, no CO_2), rock flasks after 1 hour.
10. Add 10 mL of CMC overlay/flask. Incubate for 4-5 days according to development of CPE, remove media, and stain with crystal violet plaque stain.

Treatment of serum with toxicity due to (Duvaxyn) antibodies to RK-13 cells

This procedure has been used successfully to eliminate (or reduce to readable at 1:4) toxicity from serum samples demonstrating the vaccine-induced "European" toxicity at 1:4 and greater in the EAV SNT. These horses have been vaccinated with Duvaxyn EHV 1 and 4 (the EHV-1 component is made in RK-13 cells).

Always treat a known negative and a known positive serum whenever this treatment is performed on toxic serum specimens. Preferably, use serum that has not been previously inactivated. Use sterile technique for the following method below.

Method

1. Decant growth media from as flask or 7-10 day old RK-13 (KY) cells (P399-409). Wash the flask three times by 10 mL of PBS. Add 1 mL trypsin EDTA/flask, rotate to spread over cell monolayer, then incubate approx. 6-8 min at 37°C.
2. When cells have detached, add 2 mL EMEM with 10% FSCS/flask. Mix well with 5 mL pipet, and then divide between three 1.5 mL microcentrifuge tubes.
3. Spin tubes at 20,000 rpm for 15 sec in microcentrifuge. Use a Pasteur pipet to remove and discard all supernatant.
4. Add 0.5 mL of toxic serum or control serum per tube; mix with pipet until cells are resuspended; label tube appropriately. Incubate for 1 hour at 37°C. Flick tubes after 30 min to resuspend.
5. Spin 2,500 rpm for 15 sec. Use a Pasteur pipet to transfer each serum to a new, labeled 1.5 mL microcentrifuge tube.
6. Inactivate at 56°C for 30 min if serum has not been inactivated previously. Proceed with EAV SNT test, running sera (toxic, negative, positive) as usual and in pre-seeded (planted 18-24 hours prior) plates.

Our standard laboratory protocol for commonly used molecular biology/recombinant DNA techniques

Primer design

The approach for designing of specific and efficient primers remains somewhat empirical; there are no hard and fast rules. With experience and by using a good primer designing program (e.g. MacDNASIS v3.5; Hitachi) the majority of the primers can be made to work. Following are some guidelines to design specific and efficient primers.

1. Where possible, select primers with an average G+C content of around 50%.
2. Where possible, select primers with a random base distribution. Try to avoid primers with stretches of polypurines, polypyrimidines, or other unusual sequences.
3. Melting temperature (T_m) should be between 55-57°C.
   Increasing the T_m therefore, the annealing temperature enhances discrimination against incorrectly annealed primers and reduces misextension of incorrect nucleotides at the 3′-end of the primers. Therefore, increasing the annealing temperature will help to increase the specificity. An applicable annealing temperature is 5°C below the true T_m of the amplification primer. The best annealing temperatures is in the range of 55-72°C. Quick way to calculate annealing temperature is:
   Annealing temp (°C) = (# of A and T [A+T] x 2 + # of G and C [G+C] x 4)-5°C.
4. Try to design both positive and negative primers with same T_m (1-2°C difference will not matter).
5. Make sure the primer ends (3′) with a G or C.
6. Primer length should be between 19 and 30 bases (19 to 22 bases are preferred).
7. If possible select a unique area of the genome to design the primer.
8. Avoid sequences with significant secondary structure (computer programs are very useful for this).
9. Check the primers against each other for complementarily. Avoiding primers with 3′ overlaps will reduce the incidence of "primer-dimer" artifacts.
10. If possible do a simple homology matching (compare the primers with the whole genome using a computer program) and see whether primer(s) have <50% homology in the area closer to be amplified (or to be desired area). This will reduce the number of nonspecific products of same size.

Agarose gel electrophoresis

Materials

Agarose (SEAKEM® LE agarose; FMC BIO PRODUCTS)
50X TAE buffer (Tris Acetate EDTA, pH 8.0)
Gel casting plate
Gel sealing tape
Buffer tank

Method

1. Make 1,000 mL of 1X TAE, by measuring 20 mL of 50X stock into a 1,000 mL graduated cylinder and q.s. to 1,000 mL mark with deionized water.
2. Weight 1 g of agarose (for 1% gel) and dissolve it in 100 mL of 1X TAE. Melt agarose solution in microwave until completely dissolved. Let stand at room temperature (or at 45°C in a water bath) to remove air bubbles but not enough to solidify.
3. Prepare gel-casting plate by taping the ends with gel sealing tape. Place the sample comb in the proper position on the plate.
4. Pour 100 mL of 1% agarsoe on to the gel casting plate slowly and allow it to solidify (volume of agarose needed depends on the size of the gel casting plate). The gel should be between 0.3 and 0.5 cm thick. After gel has set, final gelling carried out at 4°C for 30 min. There should not be any air bubbles in the gel.
5. Remove the tapes and keep the gel plate in the buffer tank. Add 1X TAE buffer into the reservoirs until it covers the surface of gel at a depth of 3-4 mm.
6. Slowly remove the sample comb and load the samples into the wells (Mix 1-20 μL of RNA or DNA with 6X gel loading buffer; loading volume depends on the sample and make sure gel loading buffer get diluted to 1X with sample).
7. Voltage 70V for 2-3 hours (95V for 45-60 min) or until the first band has migrated at least 2.5 inches away from the wells. The voltage must never exceed 100V. For low melt Agarose use low voltage (70V). Always run the gel in constant voltage.
8. Stain gel in ethidium bromide, letting it sit for 5-10 min. Destain gel by placing in deionized distilled water for 5 min.
9. Transfer the gel onto a UV light source. Usually, a transilluminator is used to facilitate this step. Place a piece of plastic wrap between the gel and surface of the illuminator. Use a face shield, gloves and body shield to minimize the UV exposure. Take a picture for records.
10. Locate the DNA or RNA bands and excise by using a sharp scalpel blade (e.g. to obtain the desired band).
11. Range of separation in gels containing different concentrations of agarose.

% agarose (% [w/v])	Efficient range of separation of linear DNA molecules (Kb)
0.3	5-60
0.6	1-20
0.7	0.8-10
0.9	0.5-7
1.2	0.4-6
1.5	0.2-3
2.0	0.1-2

Viral RNA extraction
Followed manufacturer's recommended protocol

Total viral RNA extraction
Material
MagMax-96 viral RNA isolation kit (THERMO FISHER# AM1836)
Purpose: Purification of viral RNA (and DNA) from clinical specimens (e.g.) to do reverse transcription (RT) PCR amplification.
Materials
Kit storage condition

Amount	Component	Storage Temperature (°C)
1 unit	Processing Plate & Lid	Room temperature
16.0 mL	Lysis/binding concentration	Room temperature
36.0 mL	Wash solution 1 (Add 12mL of Isopropanol)	4°C
40.0 mL	Wash solution 2 (Add 32mL of 100% ethanol)	4°C
9.0 mL	Elution buffer	4°C
1.1 mL	RNA binding beads	4°C
125 µL	Carrier RNA	-20°C
1.1 mL	Lysis/Binding Enhancer	-20°C

Required materials that not included with kit
RNase Zap (RNA remover)
100% Ethanol, Isopropanol
96-well plates magnetic stand
96-well U-bottom plate & lids
96-well plates shaker
General use needs (e.g. pipette, tips)

Protocol
1. Prepare your Lysis/Binding Solution preparation:

Specimens	Lysis/Binding Soln. Conc.	Carrier RNA	100% Isopropanol	Total Volume (µL)
13	845	13	845	1703

   a. Calculate the total volume of lysis/binding solution per your sample number.
   b. Add 13 µL l of lysis/binding solution in each cell by single channel pipet.
2. Prepare your Bead Mix preparation:

Component	Per Reaction (µL)	Specimens	Total volume (µL)
RNA Binding Beads	10	13	130
Lysis/Binding Enhancer	10	13	130

   a. Make 2 extra per your number of samples.

b. Add 20 μL of bead mix in each cell by single channel pipet on well A

3. Add 150 μL of Wash Buffer 1 into row B and C using by 300 multichannel pipette.
4. Add 150 μL of Wash Buffer 2 into row D and E using by 300 multichannel pipette.
5. Add 50 μL of Elution Buffer into row F using by 300 multichannel pipette.
6. Add 50 μL of your RNA sample into row A in the hood room.
7. Turn the machine and add Combs above the plate in the machine.
8. RNA extraction machine takes about 30 min to extract the RNA.
9. Turn off your machine, and aliquot out your RNA in row F.
10. Aliquot out your negative control in each cell at the end of plate. If you are running negative sample, add your positive sample at the end of plate.
11. Turn on your RNA extraction machine in Sheila's lab, click "start" button twice. It takes about 30 min to extract your RNA.

MagMax-96 viral RNA isolation kit (THERMO FISHER# AM1836)
Materials
Add 12 mL of 100% isopropanol to the wash solution 1
Add 32 mL of 100% ethanol to wash solution 2
Carrier RNA and Lysis/binding solution are stored at -20°C.

Reagent preparation
Mix 1 μL of carrier RNA + 65 μL of lysis/binding solution + 65 μL of isopropanol / 1 sample
Mix 10 μL of RNA binding beads + 10 μL of lysis/binding enhancer / 1 sample

Method

1. Add 50 μL sample to 130 μL lysis/binding solution.
    a. Pipet 130 μL of lysis/binding solution (carrier RNA and isopropanol are added).
    b. Transfer up to 50 μL of sample. Avoid aerosol cross contamination.
    c. Shake plate for 1 min on orbital shaker.
2. Add 20 μL beads and mix for 5 min.
    a. Vortex bead mix.
    b. Add 20 μL of beads. Mix to each sample.
    c. Shake the plate for 5 min to lyse viruses and bind RNAs to RNA binding beads.
3. Capture RNA binding beads and carefully discard the supernatant.
    a. Move processing plate to magnetic stand to capture RNA binding beads. Leave the plate for 3 min. RNA binding beads will form pallets.
    b. Carefully, discard the supernatant without disturbing beads.
4. Wash twice with 150 μL of wash solution 1.
    a. Add 150 μL of wash solution 1. Shake the place for 1 min.
    b. Processing plate on magnetic stand for 1 min.
    c. Discard supernatant without disturbing beads carefully.

d. Remove processing plate off the magnetic stand. Repeat 4a. to 4c. second time with 150 µL of wash solution 1.

5. Wash twice with 150 µL of wash solution 2.
   a. Repeat 4a. to 4d. with wash solution 2.
6. Dry the beads by shaking plate for 2 min to allow remaining from wash solution 2. If there is residual solution, shake the plate for another 2 min.
7. Elute the RNA/DNA in 20-50 µL of elution buffer.
   a. Add 20-50 µL elution buffer, shake vigorously for 3 min.
   b. Capture the RNA binding beads. The purified RNA will be in the supernatant.
   c. Transfer the supernatant (contains RNAs) to nuclease-free container. Store the purified RNA at -80°C.

RNA extraction protocol for Next-generation sequencing prep
Materials
Norgen Biotek Total RNA purification Kit (NORGEN BIOTEK# 17200 - 50 prep)

Buffer RL	40 mL
Wash Solution A	38 mL
Elution Solution A	6 mL
Mini Spin Columns	50
Collection Tubes	50
Elution tubes (1.7 mL)	50
Product Insert	1

Customer-Supplied Reagents and Equipment
Benchtop microcentrifuge
96-100% Ethanol
B-mercaptoethanol (optional)

Protocol

1. Cell Lysate Preparation from Plasma/Serum.
   a. 1.0 mL of sample vial centrifuged at 9,000 g force at 4°C (in walk-in room) for 20 min, discard the pellets.
   b. Transfer the supernatant, and then use 200 µL of the sample for RNA extraction.
   c. Add 700 µL of Lysis Solution. Lyse viral-infected cells by vortexing for 15 sec. Ensure that mixture becomes transparent before proceeding to the next step.
   d. Add 400 µL of 100% ethanol and mix by vortex for 10 sec.
2. Binding RNA to Column.
   a. Assemble a column with one of the provided collection tubes.
   b. Apply up to 600 µL of the lysate with the ethanol (from step 1) onto the column and centrifuge for 1 min (3,500 x g).
      Note: If the entire lysate volume has not passed, spin for an additional minute at 14,000 x g
   c. Discard the flow through. Reassemble the spin column with its collection tube.

d. Depending on your lysate volume, repeat Step 2b and 2c as necessary.
   Optional step: For maximum removal of residual DNA that may affect sensitive downstream application, use One-Column DNA Removal Protocol (Norgen's RNase-Free DNase I Kit#25710)

3. Column Wash.
   a. Apply 400 µL of 100% ethanol to the column and centrifuge for 1 min.
   b. Discard the flow through and reassemble the spin column with its collection tube
   c. Repeat steps 3a and 3b to wash column a second time.
   d. Wash column a third time by adding another 400 µL of 100% ethanol.
   e. Discard the flow through and reassemble the spin column with its collection tube.
   f. Spin the column for 2 min in order to thoroughly dry the resin. Discard the collection tube.
4. RNA Elution.
   a. Place the column into a fresh 1.7 mL Elution tube provided with the kit.
   b. Add 50 µL of Elution Solution A to the column.
   c. Centrifuge for 2 min at 200 x g, followed by 1 min at 14,000 x g.
   Note: If the entire 50 µL has not been eluted, spin the column at 14,000 x g for 1 additional min.
5. Storage of RNA. The purified RNA sample may be stored at -80°C for long-term storage.

cDNA synthesis
Followed manufacturer's recommended protocol

Synthesis of first cDNA strand with oligo dT

Material

Kit: SuperScript III First-Strand Synthesis System for RT-PCR (THERMO FISHER# 18080-051 LIFE TECHNOLOGY)

Purpose: Use SuperScript III First-Strand synthesis with Oligo (dT) and Gene Specific Primers (GSP)

Protocol

Template RNA (1 pg-500 ng recommended) - Use about 1,000 ng (1 µg)

Template mix

Components	Volume per reaction (µL)	Number of reactions	Master mix
Olgo dT primer (50 um)	1.0	20	20
10 mM dNTP	1.0	20	20
Template RNA	3.0	20	60
DEPC-treated water	5.0	20	100
	10.0		200

Aliquot out 10 μL per tube and keep them in ice.
Run the first thermo cycler program: SuperScript III at 65°C for 5 min.

RT Master Mix

Components	Voume per reaction (μL)	Number of reactions	Master Mix
10X RT buffer	2.0	20	40
25 mM $MgCl_2$	4.0	20	80
0.1 M DTT	2.0	20	40
RNaseOUT (40 U/μL)	1.0	20	20
Superscript III RT (200 U/μL)	1.0	20	20
	10.0		200

Add 10 μL of cDNA Synthesis Mix to each RNA/primer mixture, mix gently, and collect by brief centrifugation. Incubate as follows.

Oligo dT: 60 min at 50 °C

Terminate the reactions at 85 °C for 5 min. Chill on ice.
Collect the reactions by brief centrifugation. Add 1 μL of Rnase H to each tube and incubate the tubes for 20 min at 37°C.
cDNA synthesis reaction can be stored at -30 to -10°C or used for PCR immediately.

Roche Pure PCR template prep
High Pure PCR Template Preparation Kit Version 20 for 100 isolations (ROCHE# 11 796 828 001)

Preparation of working solution
Proteinase K: Dissolve Proteinase K in 4.5 ml double distilled water, aliquot solution. Store at -15°C to -25°C. Stable for 12 months. Sample Lysis and DNA Binding Protocol step 1.
Inhibitor removal buffer: Add 20 mL absolute ethanol to inhibitor Removal Buffer Lable and date bottle accordingly after adding ethanol. Store at +15°C to +25°C. Stable through the expiration date printed on label. Washing and Elution protocol step 1.
Wash Buffer: Add 80 mL my absolute ethanol to Wash Buffer Label and date bottle accordingly after adding ethanol. Store at +15°C to 25°C Stable through expiration date printed on kit label. Washing and Elution protocol step 2 and 3.

Protocol
1. Thaw 200 μL of EAV KY84 SS cDNA Template x 2vials.
2. Add 200 μL or Binding Buffer. Add, 40 μL Proteinase K, mix immediately, incubate for 10 min at 70°C in each vial.
3. Add 100 μL isopropanol. Mix well and apply each vial mixture into one High Pure filter tube.

4. Centrifuge for 1 min at 8,000 x g.
5. Discard flow through and collection tube.
6. Add 500 µL of inhibitor removal buffer.
7. Centrifuge for 1 min at 8,000 x g.
8. Discard flow through and collection tube.
9. Add 500 µL of Wash Buffer.
10. Centrifuge for 1 min at 8,000 x g.
11. Discard flow through and collection tube.
12. Repeat #9-10.
13. Discard flow through.
14. Centrifuge for 10 seconds at 13,000 x g.
15. Discard collection tube.
16. Add new tube and 200 µL of elution buffer at 70°C E2: 100 µL.
17. Centrifuge for 1 minute at 8,000 x g.
18. Purified Template DNA is ready to go.

Reverse transcription
Followed manufacturer's recommended protocol

Expand Long Template PCR System (ROCHE# 11 681 834 001)

Preparation

All buffers at + 37°C to +56°C before use. Vortex all other reagents thoroughly.

If the initial reaction produces too many primer dimers, try using two separate master mixes:

Master Mix 1 contains dNTPs and primers

Master Mix 2 contains buffer, template cDNA and enzyme

Add these two to the tube on ice, then just before starting the reaction.

Protocol

Amplification cDNA. DNA System 3: >12kb

Master Mix 1 (for 10 viruses)

Components	Volume (µL) per reaction	Number of reactions	Master mix (µL)	Final concentration
10mM dNTP Mix	10.0	5	50	500 uM
Downstream	1.0	5	5	800 nM (400 nm)
Upstream primer	1.0	5	5	800 nM (400 nm)
10x PCR Buffer with MgCl2 (Buffer 3)	5.0	5	25	
Expand Long Tem- 0.75 µL plate Enzyme mix	0.75	5	3.75	
Template DNA after High Pure PCR Tem	5.0	5	25	

Prep Kit

Sterile DW	27.3	5	136.25
	50.0		250

DNA quantification (ng/μL)

Qiagen One Step RT-PCR kit (QIGEN# 210210)
Protocol

Component	Volume	Reaction	Master Mix
Master mix			
Rnase-Free water(provided)	14.0 μL	13	182
5X Qiagen One Step RT-PCR Buffer (contains $MgCl_2$)	5.0 μL	13	65
dNTP Mix (containing 10 mM of each dNTP)	1.0 μL	13	13
Qiagen OneStep RT-PCR Enzyme Mix	1.0 μL	13	13
Rnase Inhibitor (optional, not provided)	0.5 μL	13	6.5
	21.0 μL		279.5

Aliquot out the master mix into tubes (6)

Primer 1	1.0 μL
Primer 2	1.0 μL
Template RNA	
Template RNA, added at step 4	1.5 μL
Total volume	25 μL

Add each forward and reverse primer to master mix tubes (6).

Thermocyclor		Duration	Temperature
Reverse transcription		30 min	50°C
Initial PCR activation step		15 min	95°C
3 step cycling:	**Denaturation**	30 sec	94°C
	Annealing	30 sec	55°C
	Extension	1 min	72°C
Final extension		10 min	72°C

QIAGEN HotStar Taq DNA Polymerase PCR

Components	Volume (μL)	# Reactions	Volume (μL)
10X PR Buffer (with 15 mM $MgCl_2$)	5.0	12	60

25 mM $MgCl_2$ (to get final conc. 2.5 mM)	2.0	12	24
dNTP Mix (10 mM of each dNTP)	1.0	12	12
QIAGEN HotStart DNA Polymerase	0.3	12	3.6
Nuclease-free water	38.7	12	464.4
Master mix	47.0		564

Aliquot out 47 μL into tubes (10).

Forward primer (20 μm)	**1.0 μL**
Reverse primer (20 μm)	**1.0 μL**
Template DNA	1.0 μL

Add each forward and reverse primer to master mix tubes (10).

Thermal Cycler Conditions		
Initial activation step	95°C	15 min
3-step cycling (40 cycles)		
Denaturation	94°C	30 sec
Annealing	55°C	30 sec
Extension	72°C	1 min
Final extension	72°C	10 min

Roche pure PCR purification

Purification of DNA Fragments from Agarose Gel

Protocol

1. Cut desired DNA band from gel using an ethanol-cleaned scalpel or razor blade.
   a. Minimize gel volume by visualizing DNA and cutting the smallest possible gel slice on a UV light box.
2. Place excised agarose gel slice in a sterile 2.0 mL microcentrifuge tube.
   a. Determine gel mass by first pre-weighting the tube, and then reweighting with the excised gel slice
3. Add 300 μL Binding Buffer for every 100 mg agarose gel slice to the microcentrifuge tube.
   a. 1 g = 1,000 mg/ 100 mg = 0.1 g.
4. Dissolve agarose gel slice in order to release the DNA.
   a. Vortex the microcentrifuge tube 15-30 sec to resuspend the gel slice in the binding buffer.
   b. Incubate the suspension for 10 min at 56°C in water bath.
   c. Vortex the tube briefly every 2-3 min during incubation.
5. After the agarose gel slice is completely dissolved.
   a. Add 150 μL isopropanol for every 100 mg agarose gel slice to the tube.
      i. 1g = 1,000 mg/ 100mg = 0.1 g
   b. Vortex thoroughly.
6. Insert one high pure filter tube into one collection tube.
   a. Pipette the entire contents of the microcentrifuge tube into the upper reservoir of the filter tube.

b. Do not exceed 700 µL total volume. If mixture is >700 µL, split the volume and use two separate Filter Tubes for each portion.

7. Centrifuge 30-60 sec at maximum speed in a standard tabletop centrifuge at +15 to +26°C.
8. Discard the flow through solution.
   a. Reconnect filter tube with the same collection tube
9. Add 500 µL wash buffer to the upper reservoir.
   a. Centrifuge 1 min at maximum speed (as above)
10. Discard the flow through solution.
    a. Recombine filter tube with the same collection tube.
    b. Add 200 µL wash buffer.
    c. Centrifuge 1 min at maximum speed.
       i. This second 200 µL wash step ensures optimal purity and complete removal of Wash Buffer from the glass fibers.
11. Discard the flow through solution and collection tube
12. Add 50-100 µL of nucleic acid free water to the paper reservoir of the filter tube (70 µL E1/ 30 µL E2 at 70°C).
13. The microcentrifuge tube now contains the purified DNA.
    a. Either use the eluted DNA directly or store the eluted DNA at +2°C to +8°C or -15°C to -25°C for later analysis.

5′ RACE System for Rapid Amplification of cDNA ends (THERMO FISHER# 18374-058 Invitrogen (10rxn) LIFE TECHNOLOGY)

Materials

Reagent	Volume	Storage
10X PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl]	500 µL	-20°C
25 mM $MgCl_2$	500 µL	-20°C
*Note: 10X PCR buffer does not contain $MgCl_2$. Therefore, $MgCl_2$ must be added to the first strand reaction mix		
10 mM dNTP mix [10 mM each dATP, dCTP, dGTP, dTTP]	100 µL	-20°C
0.1 M DTT	100 µL	-20°C
SuperScript II Reverse Transcriptage (200 units/ µL)	10 µL	-20°C
Rnase mix	10 µL	-20°C
5X tailing buffer [50mM Tris-HCl (pH 8.4), 125 mM KCl, 7.5 mM $MgCl_2$]	500 µL	-20°C
2 mM dCTP	50 µL	-20°C
Terminal deoxynucleotidyl transferase (TDT)	15 µL	-20°C
5' RACE abridged anchor primer (AAP, 10 µM)	80 µL	-20°C
Universal amplification primer (UAP, 10 µM)	40 µL	-20°C
Abridged universal amplification primer (AUAP, 10 µM)	40 µL	-20°C
DEPC-treated water	1.25 mL	-20°C
Control gene-specific primer 1 (GSP1, 1 µM)	25 µL	-20°C
Control nested gene-specific primer 2 (GSP2, 10 µM)	80 µL	-20°C
Control PCR primer, gene-specific primer 3 (GSP3, 10 µM)	20 µL	-20°C

Control DNA (2x104 copies/ μL; ~0.1 pg/ μL)	100 μL	-20°C
Control RNA (50 ng/ μL)	10 μL	-20°C

DNA Purification System	**Volume**	**Storage**
S.N.A.P. Columns	10 columns	4°C
Collection tubes	10 tubes	4°C
Binding solution (6M sodium iodide)	30 mL	4°C
Wash buffer concentrate	1 mL	4°C

Additional materials required

Sterilized, RNase-free thin-walled 0.2 or 0.5 mL PCR tubes
Automatic pipettes capable of dispensing 1 to 20 μL and 20 to 200 μL
Sterilized, RNase-free disposable tips for automatic pipettes
Sterilized, distilled water
Absolute ethanol
GSP 1 (cDNA primer, user-defined)
GSP 2 (nested primer for PCR amplification of dC-tailed cDNA, user-defined, appropriately engineered)
Taq DNA polymerase *Note: Purchased: AccuPrime Pfx DNA Polymerase Cat# 12344-024 Invitrogen
37°C, 42°C, 65°C and 70°C water baths or heat blocks (or use thermal cycler)
Microcentrifuge capable of generating a relative centrifugal force of 13,000 x g
Programmable thermal cycler
Mineral oil (if necessary for your thermal cycler)

Method 1

1X Wash Buffer for S.N.A.P. Procedure

*1X wash buffer must be prepared from the wash buffer concentrate.

1. Pipette 1mL (individually wrap) of the wash buffer, concentrate into a 50 ml conical tube.
2. Add 18 mL of molecular biology grade distilled water and 21 mL of absolute ethanol. Mix thoroughly.
3. Transfer half into a second 50 mL conical and store and store at 4°C.

70% Ethanol Wash for S.N.A.P. Procedure

1. Add 35 mL of absolute ethanol and 15 mL of molecular biology grade distilled water to a 50 mL conical tube.
2. Transfer half into a second 50 mL conical and store and store at 4°C.

*Note: Molecular Biology grade distilled water = Nuclease-Free Water (Cat#AM9937 AMBION)

Method 2

First Strand cDNA Synthesis

*Add following to a 0.5 mL microcentrifuge tube (or thin-walled PCR tube if using a thermal cycler)

Components	Voume per reaction (µL)	Number of reactions	Mol
GSP1 (20 um)	1	2	2.5pmoles
Sample RNA (5 µg)	5	2	
DEPC-treated water	9.5	2	

*** Note: RNA of EAV measurement

Aliquot out each 0.5 mL microcentrifuge tube #1 and #2

*Incubate the mixture 10 min at 70°C to denature RNA. Chill 1min on ice.

Collect the contents of the tube by brief centrifugation and add the following in the order given:

Components	Voume per reaction (µL)	Number of reactions
10X PCR buffer	2.5	2
25 mM $MgCl_2$	2.5	2
10 mM dNTP mix	1	2
0.1 M DTT	2.5	2
Final volume (8.5)**	8.5	

***Total final volume of step 1 and 2: 24 µL

Mix gently, and collect the reaction by brief centrifugation. Incubate for 1 min at 42°C.

Add 1 µL of SuperScript II RT immediately. Mix gently and incubate for 50 min at 42°C.

30 min incubation is usually sufficient for shot (<4 kb) mRNAs.

Longer transcripts require at least 50 min to synthesize enough cDNA for a consistent signal in long PCR.

Incubate at 70°C for 15 min to terminate the reaction.

Centrifuge 10 to 20 sec and place the reaction at 37°C (at room temp right away).

Add 1 µL of RNase mix gently but thoroughly, and incubate for 30 min at 37°C.

Collect the reaction by brief centrifugation and place on ice or keep the reaction storeed at -20°C.

* Final composition of the reaction:

20 mM Tris-HCl (PH 8.4)

50 mM KCl

2.5 mM $MgCl_2$

10 mM DTT

100 nM cDNA primer (GSP1)
400 μM each dATP, dCTP, dGTP, dTTP
1-5 ug (~40ng/ μL) RNA
200 units SuperScript II RT

Thermocyclor cycle	Temp	Time
Denature RNA	70°C	10 min
Incubation 1	42°C	1 min
Incubation 2	42°C	50 min
	70°C	15 min
Incubation 3	37°C	30 min

Method 3
Materials
Collection tube
S.N.A.P. Column (spin cartridge)
40X wash buffer concentration
Binding buffer

S.N.A.P. Column Purification of cDNA
1. Equilibrate the binding solution to room temperature.
2. For each sample to be purified, equilibrate ~100 μL of sterilized, distilled water at 65°C for use in step 9.
3. Add 120 μL of binding solution (6 M Nal) to the first strand reaction.
4. Transfer the cDNA/Nal solution to a S.N.A.P. column. Centrifuge at 13,000 x g for 20 sec.
5. Remove the cartridge insert from the tube and transfer the flow through to a microcentrifuge tube. Save the solution until recovery of the cDNA is confirmed. Place the cartridge insert back into the tube.
6. Add 400 μL of cold (4°C) 1X wash buffer to the spin cartridge. Centrifuge at 13,000 x g for 20 sec. Discard the flowthrough. Repeat this wash step three additional times.
7. Wash the cartridge two times with 400 μL of COLD (4°C) 70% ethanol as described in step 6.
8. After removing the final 70% ethanol wash from the tube centrifuge at 13,000 x g for 1 min.
9. Transfer the spin catridge insert into a fresh sample recovery tube.
Add 50 μL of sterilized, distilled, water (preheated to 65°C) to the spin cartridge. Centrifuge at 13,000 x g for 20 sec to elute the cDNA.

Important Note
*Failure to remove all the ethanol can result in poor recovery of the DNA.
** It is very important that the distilled water be at 65°C in order to maximize recovery of DNA.

Method 4

TdT Tailing of cDNA

1) Add the following components to each tube and mix gently:

Components	Voume per reaction (μL)	Number of reactions
DEPC-treated water	6.5	2
5X tailing buffer	5	2
2 mM dCTP	2.5	2
S.N.A.P.-purified cDNA sample	10	2
Final volume	24	

*** Aliquot out to PCR tubes #3 and #4**

2) Incubate for 2 to 3 min at 94°C. Chill 1 min on ice.

Collect the contents of the tube by brief centrifugation and place on ice.

3) Add 1 μL TdT to the each tubes, mix gently, and incubate for 10 min at 37°C.

4) Heat inactivate the TdT for 10 min at 65°C.

Collect the contents of the reaction by brief centrifugation and place on ice.

*Final composition of the reaction:

10 mM Tris-HCl (PH 8.4)

25 mM KCl

1.5 mM $MgCl_2$

200 μM dCTP

cDNA

TdT

Thermocyclor cycle	Temp	Time
Incubation 4	94°C	3 min
Incubation 5	37°C	10 min
TdT Inactivation	65°C	10 min

Method 5

PCR of dC-tailed cDNA

1. Equilibrate the thermal cycler block to 94°C. In most cases, the "good start" procedure gives specific amplification products. For some target and primer sets, "hot start" has been reported to improve the specificity of the reaction.
2. Add the following to a 0.2 or 0.5 mL thin-wall PCR tube sitting on ice.

Components	Voume per reaction (µL)	Number of reactions
Sterilized, distilled water	31.5	2
10X PCR buffer [200mM Tris-HCl (pH8.4), 500mM KCl	5	2
25 mM $MgCl_2$	3	2
10 mM dNTP mix	1	2
Nested GSP2 (prepared as 10 µM solution)*	2	2
Abridged Anchor Primer (10 µM)	2	2
dC-tailed (**PCR tubes #3 and #4**)	5	2
Final volume	49.5	

* Aliquot out to PCR tubes #5 and #6

3. Add 0.5 µL of Taq DNA polymerase (5 units/µL) immediately before mixing. **New enzyme: DNA Taq Polymerase, Native (LIFE TECH#18038-018)
4. Mix the contents of the tube (Taq DNA polymerase is added immediately before going into the thermal cycler).

Overlay with 50 to 100 µL of mineral oil (if necessary).

*Final composition of the reaction:
20 mM Tris-HCl (PH8.4)
500 mM KCl
1.5 mM $MgCl_2$
400 nM GSP2
400 nM Abridged Anchor Primer
200 µM each dATP, dCTP, dGTP, dTTP
Tailed cDNA
2.5 units Taq DNA polymerase

5. Transfer tubes directly from ice to the thermal cycler preequilibrated to the initial denaturation temperature (94°C)
6. Perform 30 to 35 cycles of PCR. A typical cycling protocol for cDNA with < 1 kb amplified region is:

Thermocycler cycle	Temp	Cycle	Time
	94°C	1	2 min
Denaturation	94°C		1 min
Annealing of primers	55°C	35	1 min
Primer extension	72°C		2 min
Final extension	72°C	1	5 min
Indefinite hold	5°C		Until samples are removed

7. Analyze 5-20 µL of 5' RACE products by agarose gel electrophoresis according to standard protocols, using appropriate size standards. Either TAE or TBE electrophoresis buffer may be used for the procedure. The volume of the sample used for analysis will depend on the volume and thickness of the sample well. If products will be extracted for reamplification, ultraviolet (UV) visualization of ethidium bromide-stained products should be performed using either a long wavelength (356 nm) UV or 302 nm wavelength source to minimize DNA nicking.

Nested GSP2 (10 uM solution) Preparation

N1V1 = N2V2 (20 uM to 10 µM Dilution)

20 uM : 50 µL = 10 µM:V2

V2= 25µL

Working stock (1:1)

25 µL Nucleic Acid-Free water + 25 µL of 20uM Primer

= Total 50 µL of Nested GSP2 (10 µM solution) of BN 291N

Performing 1% agarose gel electrophoresis

Materials for small box (10 well combs)	Amt. of materials
TAE (1X) Buffer) [diluted from 10X Tris-Acetate-EDTA]	50 mL
Agarose (1%) [Seakem GTG Agarose Cat#50074 Cambrex)	0.5 g
Ethidium Bromide [GIBCO]	2.5 µL

*When you pour the gel, use pipet and get use to the height of the gel.

You can make 100 mL gel stock and can use about 30 mL for small box (10-well combs)

Method 6

Nested Amplification

1. Dilute a 5 µL aliquot of the primary PCR (tube# 5 and #6) into 495 µL Nucleic Acid Free Water
   * Aliquote out to PCR tubes #5b and #6b (diluted primary PCR)
2. Equilibrate the thermal cycler block to 94°C**.**
   Add the following to a 0.2 or 0.5 mL thin-wall PCR tube sitting on ice.

Components	**Voume per reaction (µL)**	**Number of reactions**
Sterilized, distilled water	33.5	2
10X PCR buffer [200 mM Tris-HCl (pH8.4), 500 mM KCl]	5	2
25 mM $MgCl_2$	3	2
10 mM dNTP mix	1	2
Nested GSP2 (prepared as 10 uM solution)	1	2
UAP (10 µM)	1	2
Dilution of primary PCR product (**PCR tubes #5a and #6a**)	5	2
Final volume	49.5	

*** Aliquot out to PCR tubes #9 and #10**

3. Add 0.5 µL of Taq DNA polymerase (5 units/µL) immediately before mixing.
4. Mix the contents of the tube (Taq DNA polymerase is added immediately before going into the thermal cycler)

Overlay with 50 to 100 µL of mineral oil (if necessary)

*Final composition of the reaction:

20 mM Tris-HCl (pH 8.4)
50 mM KCl
1.5 mM $MgCl_2$
200 nM nested GSP
200 nM UAP or AUAP
200 µM each dATP, dCTP, dGTP, dTTP
Diluted primary PCR product
2.5 units Taq DNA polymerase

5. Transfer tubes directly from ice to the thermal cycler preequilibrated to the initial denaturation temperature.
6. Perform 30 to 35 cycles of PCR.

Thermocycler cycle	Temp	Cycle	Time
	94°C	1	2 min
Denaturation	94°C		1 min
Annealing of primers	55°C	35	1 min
Primer extension	72°C		2 min
Final extension	72°C	1	5 min
Indefinite hold	5°C		Until samples are removed

7. Analyze 5 to 20 μL of the amplified sample, using agarose gel electrophoresis, ethidium bromide staining, and the appropriate molecular size standards.

3′ RACE System for Rapid Amplification of cDNA ends (THERMO FISHER#18373-019)

Materials

Contents	Amount (μL)
10X PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl]	500
25 mM $MgCl_2$	500
10 mM dNTP mix (10mM each dATP, dCTP, dGTP, dTTP)	100
0.1M DTT	100
SuperScript II Reverse Transcriptase (RT, 200units/ μL)	20
Adapter primer (AP, 10 μM)	20
Universal amplification primer (AUAP, 10 μM)	20
Abridged universal amplification primer (AUAP, 10μM)	20
E. *coli* RNASE H (2units/ μL)	20
DEPC-treated water	1.2 mL
Control RNA (50ng/ μL)	10
Control gene-specific primer (GSP, 10 μM)	20

Additional preparations

GSP (user-defined, appropriately engineered)

Method 1

First Strand cDNA Synthesis

This procedure converts 1 to 5 μg of total RNA into first strand cDNA. Poly (A) RNA may be used in this protocol.

1. Mix and quickly centrifuge each component before use.
2. Combine 1-5 µg of total RNA or 50 ng of poly (A) RNA and DEPC-treated water to a final volume of 11 µL in a 0.5 mL microcentrifuge tube.
3. Add 1 µL of the 10 µM AP solution, (adapter primer) mix gently, and collect reaction by brief centrifugation.
4. Heat the mixture to 70°C for 10 min and chill on ice for at least 1min.
Collect the contents of the tube by brief centrifugation and add the following:

Components	Volume (µL)	Reaction	Mastermix
10X PCR buffer	2	5	8
25 mM $MgCl_2$	2	5	8
10 mM dNTP mix	1	5	4
0.1 M DTT	2	5	8
Total	7		28

Aliquot out 7 µL into 4tubes (#1-4)
Final composition of the reaction:
20 mM Tris-HCl (pH 8.4 at 22°C)
50 mM KCl
2.5 mM $MgCl_2$
10 mM DTT
500 nM AP
500 µM each dATP, dCTP, dGTP, dTTP
1-5 µg (<50 ng/µL) of RNA

5. Mix gently and collect the reaction by brief centrifugation. Equilibrate the mixture to 42°C for 2 to 5 min.
6. Add 1 µL of SuperScript II RT. Incubate the tube in a 42°C water bath or heat block for 50 min.
7. Terminate the reaction by incubating at 70°C for 15 min.
8. Chill on ice. Collect the reaction by brief centriugation. Add 1 µL of RNase H to the tube. Mix, and incubate for 20 min at 37°C before proceeding to Protocol 2.
9. The reaction mixture may be stored at -20°C and STOP or continue the Method 2.

Method 2.
Amplification of the Target cDNA

1. To a fresh 0.5 mL microcentrifuge tube, add the following:

Component	Volume (µL)	Reaction	Master Mix
10X PCR buffer	5	5	25
25 mM $MgCl_2$	3	5	15
Autoclaved, distilled water	36.5	5	182.5
10 mM dNTP mix	1	5	5
GSP (prepared as 10 uM solution)	1	5	5
UAP (10 µM)	1	5	5

Taq DNA polymerase (2 to 5 units/μL)		0.5	5	2.5
	Total	48		240

Final composition of the reaction:
20 mM Tris-HCl (pH 8.4)
50 mM KCl
1.5 mM $MgCl_2$
200 nM GSP
200 nM AUAP or UAP
200 μM each dATP, dCTP, dGTP, dTTP
0.04 to 0.1 unit/ μL Taq DNA polymerase

2. Add 2 μL from the cDNA synthesis reaction to the tube. Collect the reaction briefly by centrifugation.
3. Perform 20 to 35 cycles of PCR, using the protocol accompanying Taq DNA polymerase.

Temp	Cycle	Time
94°C	1	3 min
94°C		45 sec
55°C	35	30 sec
72°C		1 min 30 sec
72°C	1	10 min
4°C		Until samples are removed

4. Analyze 10 to 20 μL of the amplified sample, using agarose gel electrophoresis and ethidium bromide staining, and the appropriate molecular size standards. If the positive control RNA was used, a single 720 bp band will be visible.

Method 3.
Gel purification.
Follow previous protocol.

Miniprep
Followed manufacturer's recommended protocol

QIAprep Spin Miniprep Kit (QIAGEN# 27104)
***Note: Optional: Add LyseBluse reagent to Buffer P1 at a ratio of 1 to 1,000.
Add the provided Rnase A solution to Buffer P1, mix, and store at 2-8°C.
(*Total 20 mL of Buffer P1 + 20 μL of Lyseblue + 200 μL of RNase solution)
Add ethanol (100%) to Buffer PE before use.

All centrifugation steps are carried out at 13,000 rpm (~17,900) in a conventional tabletop microcentrifuge.

Method

1. Take the bacterial culture tubes out from shaker incubator overnight (37°C, 200 rpm).
2. Using 2 mL pipet, aliquot out 1.5 mL of bacterial growth in 2 mL microcentrifuge tubes.
3. Centrifugation the tubes at >800 rpm (6,800 x g) for 3 min at room temperature.
4. Discard supernatant without disturbing pelleted bacterial cells.
5. Add 250 µL of Buffer P1 and scratch the tube on the rack to mix thoroughly. Solution will turn cloudy.
6. Add 250 µL of Buffer P2 and mix thoroughly by inverting the tube 4-6 times (gently flipping the tube) until the solution becomes clear. Do not allow the lysis reaction to proceed for more than 5 min. If using LyseBlue reagent, the solution will turn blue. Solution turns blue, then clear
7. Add 350 µL of N3 Buffer and mix thoroughly by inverting the tube 4-6 times (gently flipping the tube). If using LyseBlue reagent, the solution will turn colorless. Solution will be cloudy.
8. Centrifuge for 10 min at 13,000 rpm (~17,900 x g) in a tabletop microcentrifuge.
9. Aliquot out all the supernatant without disturbing pellet (bacterial DNA, no needed) and try to avoid to pipetting white floating stuff. You only want the supernatant into Qiaprep spin column by decanting or pipetting.
10. Add 500 µL of PB buffer to wash the QIAprep spin column. Centrifuge for 30-60 sec and discard the flow through, or apply vacuum to the manifold to draw. Centrifuge for 30-60 sec and discard the flow through, or apply vacuum to the manifold to draw the solution through the QIAprep spin column and switch off the vacuum source.
11. Wash the QIAprep spin column by adding 750 µL Buffer PE. Centrifuge for 30-60 sec and discard the flow through, or apply vacuum to the manifold to draw the solution through the QIAprep spin column and switch off the vacuum source. Transfer the QIAprep spin column to the collection tube.
12. Centrifuge hard for 1 min to remove residual wash buffer.
13. Place the QIAprep column in a clean 1.5 mL microcentrifuge tube. To elute DNA, add 50 µL Buffer EB (10mM Tris-HCl, pH 8.5) or nucleic acid-free water to the center of the QIAprep spin column, let it stand for 1min, and centrifuge for 1 min.
14. Do DNA quantification.

Preparation of Restrict Enzyme Reaction

Reagent	Amount
Nucleic Acid Free water	7 µL
10X FD (fast digest) Buffer	2 µL
Miniprep DNA (Elute 2)	10 µL
Eco R1 (Restriction Enzyme)	1 µL
	20 µL

Incubate the tubes in 37°C water bath for 15 min. Then, ready to load the gel.

Cloning
Followed manufacturer's recommended protocol

Generating Blunt Ends

Materials

T4 DNA Polymerase THERMO FISHER# 18005-016

1. To a Sterile 1.5 mL microcentrifuge tube on ice, add:

Component	Amount
5X T4 DNA polymerase buffer	20 µL
0.5 mM dNTP mix	20 µL
DNA [5' Race Final PCR product pooled (Tube#9,10,11,12)]	6 µL
T4 DNA polymerase (5U/µL)	2 µL
Autoclaved, distilled water	52 µL
Total volume:	**100 µL**

2. Mix gently, and incubate at 11°C for 15 min.
3. Place reaction on ice. Store at -20°C.
4. Terminate reaction by phenol extraction, chloroform extraction and ethanol precipitation.
   Heat inactivation is not recommended, as heat inactivation may cause ends to "breathe" and be more susceptible to 3' to 5' exonuclease activity.

Phenol extraction, chloroform extraction and ethanol precipitation

1. Extract the sample twice with phenol/chloroform/isoamyl alcohol (25:24:1).
   Increase the volume of the sample to about 300 µL by adding sterile deionized water into the sample. For each extraction, mix the DNA sample with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) until an emulsion is formed.
2. The centrifuge the tube(s) at 12,000 x g for approximately 15 sec at room temperature. The upper aqueous phase contains the DNA and should be transferred to a fresh sterile microcentrifuge tube after each extraction.
3. Add sodium chloride (to a final concentration of 0.1M; use 20 µL of 5M NaCl) or sodium acetate (final concentration, 0.25 M) to the sample. Then, precipitate the DNA by adding a 2.5 volume of cold absolute alcohol (750-800 µL). Incubate for 30-45 min at -70°C, then centrifuge (12,000 x g) for 10 min at 4°C.
4. Discard the supernatant, dry DNA pallet under vacuum (or in a 37°C oven) and resuspend the pellet in a convenient volume (20-200 µL) of sterile deionized water and measure the optical density of the solution. Calculate the concentration of the plasmid and the total amount of DNA obtained. DNA is ready for restriction enzyme digestion and PCR analysis.

Zero Blunt TOPO PCR Cloning

Materials

Zero Blunt TOPO PCR Cloning Kit THERMO FISHER# K2800-020SC

Produce Blunt PCR Product

1. Set up the following 6 µL TOPO Cloning reaction

Reagent	Amount*
Fresh PCR Product	0.5-4 µL
Salt Solution	1 µL
Sterile Water	Add to a total volume of 5 µL
PCR Blunt II TOPO	1 µL
Total	6 µL

*For transformation of chemically competent E. *coli* only

2. Mix gently and incubate for 5 min at room temperature

3. Place tubes on ice. Proceed to Transformation and Analysis.

Transformation

1. Thaw 1 vial of One Shot E. *coli* cells on ice for each transformation.
2. Add 2 µL of the TOPO Cloning reaction to each vial of One Shot cells to be transformed, and mix gently.
3. Incubate the vials on ice for 5-30 min.
4. Heat-shock the cells for 30 sec at 42°C without shaking.
5. Add 250 µL of room temperature S.O.C. medium to the cells.
6. Cap the tubes and shake them at 37°C for 1 hour.
7. Spread 10-50 µL from each transformation on pre-warmed LB plates containing 50 µg/mL kanamycin or pre-warmed Low Salt LB plates containing 25 µg/mL Zeocin selective antibiotic.
   Refer to the Zero Blunt TOPO PCR Cloning Kit manual for a Low Salt LB medium recipe.
8. Incubate plates overnight at 37°C.

An efficient TOPO Cloning reaction should produce several hundred colonies. Pick ~10 colonies for analysis. Proceed to Analyze Positive Clones.

Analyze Positive Clones

1. Culture the 10 colonies overnight in LB medium containing 50 µg/mL kanamycin or low Salt LB medium containing 25 µg/mL Zeocin selective antibiotic.
2. Isolate plasmid DNA using your method of choice. For ultrapure plasmid DNA, we recommend the PureLink HQ Mini Plasmid Purification Kit (Cat# K2100-01).
3. Analyze the plasmid by restriction analysis or by sequencing to confirm the presence and orientation of the insert.

NGS data analysis command line coding method

Protocol from Dr. Ganwu Li's laboratory at Iowa State University

1. In DataAnalysis file, separate each sample file.
2. Choose individual file to work on.
3. Turn on Terminal
4. Type ls : view files
5. Type cd Desktop: enter to desktop
6. Type ls : view files on desktop
7. Type cd DataAnalysis to enter into the folder
8. Type ls : to view files in DataAnalysis
9. Choose EAV folder and Type cd (EAV-folder name) : to enter into the folder
10. Type ls : to view files in EAV file
11. Choose a sample folder and Type cd (sample folder name)
12. Create sample_list file : bash namelist.sh
13. View sample_list file : less sample_list
14. Trim the raw NGS data materials : bash batch_trim_Truseq.sh
    a. First batch use the program: Trueseq.sh
    b. Second batch use the program: Nextera.sh
    c. You can check the condition of trimmed data looking at surviving %
15. Enter to cd trimmed/ : new folder
16. Type ls to view the trimmed folder
17. Type mkdir other : create other folder to keep unnecessary sample files
18. Type mv sample name R1 and R2 /others/ : to move two unnecessary sample files to other folder
    a. D107_S20_FUP_L001_R1_001.fastq.gz ./others/
    b. D107_S20_RUP_L001_R2_001.fastq.gz ./other/
19. Create sample_list file in trimmed folder : bash namelist.sh
20. View sample_list file to check the right sample name : less sample_list
21. Copy EAVKY84.fasta file and paste into Trimmed folder in each sample folder
22. Do partial De Novo assembly to collect reads : bash general102314.sh EAVKY84.fasta
    a. To complete de Novo assembly, you need four different programs including Seqman pro (commercial). But in my case, I don't do complete *de novo* assembly in my data analysis..
23. Enter into a folder cd (trimmed_sample name)
24. Type ls to view the folder (trimmed_sample name)
25. Type cd single_id txt .. : Copy single_id txt file in trimmed sample name folder
26. Type cd .. : to come out of trimmed sample name folder
27. Type ls : to view trimmed sample name folder and see if there is folder (single_id txt)
28. Use seqtk subseq to get the reads Type and create two files (sample_R1 and R2.fastq) :
    a. seqtk subseq trimmed_S-S-D107_S20_L001_R1_001.fastq.gz single_id.txt > sample_R1.fastq
    b. seqtk subseq trimmed_S-S-D107_S20_L001_R2_001.fastq.gz single_id.txt > sample_R2.fastq
29. Finally, you will do mapping of sample_R1 & R2.fastq and EAVKY84 with bwa program :
    a. EAVKY84 will be your reference EAV sequence

b. Otherwise, you can use 33 EAV full length genome sequences for the reference sequence
c. To do mapping, you can change the parameter for reads (how long or short you want). But right now, I stick with one parameter to be simple
d. Create Index : bwa index EAVKY84.fasta
e. Then Type and create sam file : bwa mem -M EAVKY84.fasta sample_R1.fastq sample_R2.fastq > sample.sam
f. Now it will do mapping

30. You need to convert sam file to bam file to visualize in IGV :
    a. Type and convert sam to bam file : samtools view -hbS sample.sam > sample.bam
    b. Type and create sorted bam file : samtools sort sample.bam >sample.sorted.bam
    c. Type : samtools index sample.sorted.bam
        i. Samtools index for bam file is different from bwa index
        ii. After sorting you will have (sample.sorted.bam.bai)
        iii. Sample.sorted.bam is your final bam file that visualize via IGV program
31. Now you can view in IGV
    a. First, you need to convert your EAVKY84.fasta into genome file
    b. Press genome and click create genome file
        i. Load EAVKY84.fasta
        ii. Name and describe the file
        iii. Save EAVKY84.genome in sample folder
    c. Second, click load from file, then you load sample.sorted.bam file
    d. You can see % at the ambiguous sites (ambiguity)
32. Now, type below to create consensus sequence :
    a. samtools mpileup -uf EAVKY84.fasta sample.sorted.bam |bcftools call -c | vcfutils.pl vcf2fq > cns.fq
    b. To view your consensus sequence type : less cns.fq
    c. To convert your consensus sequence from fastq to fasta, type : seqtk seq -a cns.fq > cns.fasta

33. To find the numbers and location of your ambiguous sites, type :
    a. perl ~/Desktop/boracode/ambcount.pl cns.fasta

At ISU, Ying helped me to download these below software programs and set up PATH through command line (terminal) to do NGS data analysis: Also I need Ying's self written script to run these programs (boracode folder)
Trimmomatic.0.36 : for trim, clean the data to run the batch_trim_TruSeq or Nextera
IGV: visualizing for map
BWA : mapping
SAM tools : manipulate the mapping files SAM file will convert to BAM file
Seqtk: manipulate the fastq files
BWA (not Abyss and SeqPro): to run the *de novo* assembly
Lofreq: Variants calling software find the variant percentages

REFERENCES

1. Balasuriya, U. B. R., Go, Y. Y. & MacLachlan, N. J. Equine arteritis virus. *Vet. Microbiol.* **167,** 93–122 (2013).
2. de Vries, A. A. F., Horzinek, M. C., Rottier, P. J. M. & de Groot, R. J. The Genome Organization of the Nidovirales: Similarities and Differences between Arteri-, Toro-, and Coronaviruses. *Semin. Virol.* **8,** 33–47 (1997).
3. Boon, J. A. den *et al.* Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J. Virol.* **65,** 2910–2920 (1991).
4. Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. Nidovirales: Evolving the largest RNA virus genome. *Virus Res.* **117,** 17–37 (2006).
5. Snijder, E. J. & Spaan, W. J. Arteriviruses. in *Fields Virology.* (eds. Knipe, D. M. & Howley, P. M.) 1337–1355 (Lippincott Williams & Wilkins, 2007).
6. Cavanagh, D. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* **142,** 629–633 (1997).
7. Dunowska, M., Biggs, P. J., Zheng, T. & Perrott, M. R. Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). *Vet. Microbiol.* **156,** 418–424 (2012).
8. International Committee on Taxonomy of Viruses (ICTV). Available at: https://talk.ictvonline.org/taxonomy/. (Accessed: 4th November 2017)
9. Doll, E. R., Bryans, J. T., Wilson, J. C. & Mccollum, W. H. Isolation of a filterable agent causing arteritis of horses and abortion by mares; its differentiation from the equine abortion (influenza) virus. *Cornell Vet.* **47,** 3–41 (1957).
10. Stadejek, T., Mittelholzer, C., Oleksiewicz, M. B., Paweska, J. & Belák, S. Highly diverse type of equine arteritis virus (EAV) from the semen of a South African donkey: short communication. *Acta Vet. Hung.* **54,** 263–270 (2006).
11. Timoney, P. J. & McCollum, W. H. Equine viral arteritis. *Vet. Clin. North Am. Equine Pract.* **9,** 295–309 (1993).
12. Bell, S. A., Balasuriya, U. B. R. & MacLachlan, N. J. Equine Viral Arteritis. *Clin. Tech. Equine Pract.* **5,** 233–238 (2006).
13. Glaser, A. L., Vries, A. A. F. de, Rottier, P. J. M., Horzinek, M. C. & Colenbrander, B. Equine arteritis virus: A review of clinical features and management aspects. *Vet. Q.* **18,** 95–99 (1996).
14. Doll, E. R., Knappenberger, R. E. & Bryans, J. T. An outbreak of abortion caused by the equine arteritis virus. *Cornell Vet.* **47,** 69–75 (1957).
15. Balasuriya, U. B. R., MacLachlan, N. J., De Vries, A. A. F., Rossitto, P. V. & Rottier, P. J. M. Identification of a Neutralization Site in the Major Envelope Glycoprotein (GL) of Equine Arteritis Virus. *Virology* **207,** 518–527 (1995).
16. Balasuriya, U. B. R. *et al.* Neutralization Determinants of Laboratory Strains and Field Isolates of Equine Arteritis Virus: Identification of Four Neutralization Sites in the Amino-Terminal Ectodomain of the GLEnvelope Glycoprotein. *Virology* **232,** 114–128 (1997).
17. Balasuriya, U. B. R. *et al.* Genetic characterization of equine arteritis virus during persistent infection of stallions. *J. Gen. Virol.* **85,** 379–390 (2004).
18. MacLachlan, N. J. & Balasuriya, U. B. Equine viral arteritis. *Adv. Exp. Med. Biol.* **581,** 429–433 (2006).

19. Miszczak, F. *et al.* Emergence of novel equine arteritis virus (EAV) variants during persistent infection in the stallion: origin of the 2007 French EAV outbreak was linked to an EAV strain present in the semen of a persistently infected carrier stallion. *Virology* **423,** 165–174 (2012).
20. Zhang, J., Timoney, P. J., Maclachlan, N. J. & Balasuriya, U. B. R. Identification of an additional neutralization determinant of equine arteritis virus. *Virus Res.* **138,** 150–153 (2008).
21. Zhang, J. *et al.* Molecular epidemiology and genetic characterization of equine arteritis virus isolates associated with the 2006-2007 multi-state disease occurrence in the USA. *J. Gen. Virol.* **91,** 2286–2301 (2010).
22. W. H. McCOLLUM & Timoney, P. J. Experimental observations on the virulence of isolates of equine arteritis virus. in 558–559 (1998).
23. Pronost, S. *et al.* Description of the first recorded major occurrence of equine viral arteritis in France. *Equine Vet. J.* **42,** 713–720 (2010).
24. Porterfield, J. S. *et al.* Togaviridae. *Intervirology* **9,** 129–148 (1978).
25. Zeegers, J. J., Van der Zeijst, B. A. & Horzinek, M. C. The structural proteins of equine arteritis virus. *Virology* **73,** 200–205 (1976).
26. Giles, J. C., Perrott, M. R. & Dunowska, M. Primary possum macrophage cultures support the growth of a nidovirus associated with wobbly possum disease. *J. Virol. Methods* **222,** 66–71 (2015).
27. Kuhn, J. H. *et al.* Reorganization and expansion of the nidoviral family Arteriviridae. *Arch. Virol.* **161,** 755–768 (2016).
28. Westaway, E. G. *et al.* Togaviridae. *Intervirology* **24,** 125–139 (1985).
29. de Vries, A. A. *et al.* All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* **18,** 3241–3247 (1990).
30. Godeny, E. K., de Vries, A. A., Wang, X. C., Smith, S. L. & de Groot, R. J. Identification of the leader-body junctions for the viral subgenomic mRNAs and organization of the simian hemorrhagic fever virus genome: evidence for gene duplication during arterivirus evolution. *J. Virol.* **72,** 862–867 (1998).
31. Spaan, W. J. M. *et al.* Comparative and evolutionary aspects of Coronaviral, Arteriviral, and Toroviral genome structure and expression. In: Brinton MA, Heinz FX (eds) New aspects of positive-strand RNA viruses. in 12–19 (1990).
32. Lai, M. M. & Cavanagh, D. The molecular biology of coronaviruses. *Adv. Virus Res.* **48,** 1–100 (1997).
33. Snijder, E. J. & Meulenberg, J. J. The molecular biology of arteriviruses. *J. Gen. Virol.* **79 (Pt 5),** 961–979 (1998).
34. Brinton, M. A. *et al.* Family Arteriviridae. In: van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (eds) Virus taxonomy. in 851–857 (Academic Press, 2000).
35. Enjuanes, L., Spaan, W. J. M., Snijder, E. J. & Cavanagh, D. Order Nidovirales. In: van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (eds) Virus taxonomy. in 827–834 (Academic Press, 2000).
36. van Aken, D., Zevenhoven-Dobbe, J., Gorbalenya, A. E. & Snijder, E. J. Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is

essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J. Gen. Virol.* **87,** 3473–3482 (2006).
37. Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81,** 853–879 (2000).
38. de Vries, A. A., Chirnside, E. D., Horzinek, M. C. & Rottier, P. J. Structural proteins of equine arteritis virus. *J. Virol.* **66,** 6294–6303 (1992).
39. Firth, A. E. *et al.* Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production. *J. Gen. Virol.* **92,** 1097–1106 (2011).
40. Snijder, E. J., van Tol, H., Pedersen, K. W., Raamsman, M. J. & de Vries, A. A. Identification of a novel structural protein of arteriviruses. *J. Virol.* **73,** 6335–6345 (1999).
41. Wieringa, R. *et al.* Structural protein requirements in equine arteritis virus assembly. *J. Virol.* **78,** 13019–13027 (2004).
42. van Berlo, M. F., Horzinek, M. C. & van der Zeijst, B. A. Equine arteritis virus-infected cells contain six polyadenylated virus-specific RNAs. *Virology* **118,** 345–352 (1982).
43. den Boon, J. A., Kleijnen, M. F., Spaan, W. J. & Snijder, E. J. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J. Virol.* **70,** 4291–4298 (1996).
44. Snijder, E. J., Wassenaar, A. L. & Spaan, W. J. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J. Virol.* **68,** 5755–5764 (1994).
45. Wassenaar, A. L., Spaan, W. J., Gorbalenya, A. E. & Snijder, E. J. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J. Virol.* **71,** 9313–9322 (1997).
46. Snijder, E. J., Wassenaar, A. L. M., Dinten, L. C. van, Spaan, W. J. M. & Gorbalenya, A. E. The Arterivirus Nsp4 Protease Is the Prototype of a Novel Group of Chymotrypsin-like Enzymes, the 3C-like Serine Proteases. *J. Biol. Chem.* **271,** 4864–4871 (1996).
47. Tijms, M. A., Nedialkova, D. D., Zevenhoven-Dobbe, J. C., Gorbalenya, A. E. & Snijder, E. J. Arterivirus Subgenomic mRNA Synthesis and Virion Biogenesis Depend on the Multifunctional nsp1 Autoprotease. *J. Virol.* **81,** 10496–10505 (2007).
48. Steinbach, F. *et al.* Re-emergence of a genetic outlier strain of equine arteritis virus: Impact on phylogeny. *Virus Res.* **202,** 144–150 (2015).
49. Snijder, E. J., Wassenaar, A. L. M., Spaan, W. J. M. & Gorbalenya, A. E. The Arterivirus Nsp2 Protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J. Biol. Chem.* **270,** 16671–16676 (1995).
50. Posthuma, C. C. *et al.* Formation of the Arterivirus Replication/Transcription Complex: a Key Role for Nonstructural Protein 3 in the Remodeling of Intracellular Membranes. *J. Virol.* **82,** 4480–4491 (2008).
51. Manolaridis, I. *et al.* Structure and Genetic Analysis of the Arterivirus Nonstructural Protein 7α. *J. Virol.* **85,** 7449–7453 (2011).
52. Barrette-Ng, I. H. *et al.* Structure of Arterivirus nsp4 the smallest chymotrypsin-like proteinase with an α/β C-terminal extension and alternate conformations of the oxyanion hole. *J. Biol. Chem.* **277,** 39960–39966 (2002).

53. Boon, J. A. den *et al.* Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. *J. Virol.* **69,** 4500–4505 (1995).
54. Snijder, E. J., Wassenaar, A. L. M., Boon, J. A. D. & Spaan, W. J. M. Proteolytic Processing of the Arterivirus Replicase. in *Corona- and Related Viruses* 443–451 (Springer, Boston, MA, 1995). doi:10.1007/978-1-4615-1899-0_71
55. Snijder, E. J. & Meulenberg, J. J. The arterivirus replicase. The road from RNA to protein(s), and back again. *Adv. Exp. Med. Biol.* **440,** 97–108 (1998).
56. Snijder, E. J., van Tol, H., Roos, N. & Pedersen, K. W. Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. *J. Gen. Virol.* **82,** 985–994 (2001).
57. Tijms, M. A., Dinten, L. C. van, Gorbalenya, A. E. & Snijder, E. J. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc. Natl. Acad. Sci.* **98,** 1889–1894 (2001).
58. van Aken, D. *et al.* Expression, purification, and in vitro activity of an arterivirus main proteinase. *Virus Res.* **120,** 97–106 (2006).
59. Aken, D. van, Snijder, E. J. & Gorbalenya, A. E. Mutagenesis Analysis of the nsp4 Main Proteinase Reveals Determinants of Arterivirus Replicase Polyprotein Autoprocessing. *J. Virol.* **80,** 3428–3437 (2006).
60. van den Born, E., Posthuma, C. C., Knoops, K. & Snijder, E. J. An infectious recombinant equine arteritis virus expressing green fluorescent protein from its replicase gene. *J. Gen. Virol.* **88,** 1196–1205 (2007).
61. Dinten, L. C. van, Wassenaar, A. L., Gorbalenya, A. E., Spaan, W. J. & Snijder, E. J. Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. *J. Virol.* **70,** 6625–6633 (1996).
62. Dinten, L. C. van, Rensen, S., Gorbalenya, A. E. & Snijder, E. J. Proteolytic Processing of the Open Reading Frame 1b-Encoded Part of Arterivirus Replicase Is Mediated by nsp4 Serine Protease and Is Essential for Virus Replication. *J. Virol.* **73,** 2027–2037 (1999).
63. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* **17,** 4847–4861 (1989).
64. van Dinten, L. C., den Boon, J. A., Wassenaar, A. L., Spaan, W. J. & Snijder, E. J. An infectious arterivirus cDNA clone: identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 991–996 (1997).
65. Posthuma, C. C. *et al.* Site-Directed Mutagenesis of the Nidovirus Replicative Endoribonuclease NendoU Exerts Pleiotropic Effects on the Arterivirus Life Cycle. *J. Virol.* **80,** 1653–1661 (2006).
66. Pasternak, A. O., Spaan, W. J. M. & Snijder, E. J. Nidovirus transcription: how to make sense…? *J. Gen. Virol.* **87,** 1403–1421 (2006).
67. Sawicki, S. G., Sawicki, D. L. & Siddell, S. G. A Contemporary View of Coronavirus Transcription. *J. Virol.* **81,** 20–29 (2007).

68. Hedges, J. F., Balasuriya, U. B. R., Timoney, P. J., McCollum, W. H. & MacLachlan, N. J. Genetic variation in open reading frame 2 of field isolates and laboratory strains of equine arteritis virus. *Virus Res.* **42,** 41–52 (1996).
69. Molenkamp, R. *et al.* The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J. Gen. Virol.* **81,** 2491–2496 (2000).
70. Wieringa, R., Vries, A. A. F. de, Raamsman, M. J. B. & Rottier, P. J. M. Characterization of Two New Structural Glycoproteins, GP3 and GP4, of Equine Arteritis Virus. *J. Virol.* **76,** 10829–10840 (2002).
71. Hedges, J. F., Balasuriya, U. B. R. & MacLachlan, N. J. The Open Reading Frame 3 of Equine Arteritis Virus Encodes an Immunogenic Glycosylated, Integral Membrane Protein. *Virology* **264,** 92–98 (1999).
72. Johnson, C. R., Griggs, T. F., Gnanandarajah, J. & Murtaugh, M. P. Novel structural protein in porcine reproductive and respiratory syndrome virus encoded by an alternative ORF5 present in all arteriviruses. *J. Gen. Virol.* **92,** 1107–1116 (2011).
73. Snijder, E. J. *et al.* Unique and Conserved Features of Genome and Proteome of SARS-coronavirus, an Early Split-off From the Coronavirus Group 2 Lineage. *J. Mol. Biol.* **331,** 991–1004 (2003).
74. Wieringa, R., Vries, A. A. F. de & Rottier, P. J. M. Formation of Disulfide-Linked Complexes between the Three Minor Envelope Glycoproteins (GP2b, GP3, and GP4) of Equine Arteritis Virus. *J. Virol.* **77,** 6216–6226 (2003).
75. Vries, A. A. de, Raamsman, M. J., Dijk, H. A. van, Horzinek, M. C. & Rottier, P. J. The small envelope glycoprotein (GS) of equine arteritis virus folds into three distinct monomers and a disulfide-linked dimer. *J. Virol.* **69,** 3441–3448 (1995).
76. Das, P. B. *et al.* Glycosylation of minor envelope glycoproteins of porcine reproductive and respiratory syndrome virus in infectious virus recovery, receptor interaction, and immune response. *Virology* **410,** 385–394 (2011).
77. De Vries, A. A., Post, S. M., Raamsman, M. J., Horzinek, M. C. & Rottier, P. J. The two major envelope proteins of equine arteritis virus associate into disulfide-linked heterodimers. *J. Virol.* **69,** 4668–4674 (1995).
78. Balasuriya, U. B. R., Rossitto, P. V., DeMaula, C. D. & MacLachlan, N. J. A 29K envelope glycoprotein of equine arteritis virus expresses neutralization determinants recognized by murine monoclonal antibodies. *J. Gen. Virol.* **74,** 2525–2529 (1993).
79. Chirnside, E. D., de Vries, A. A. F., Mumford, J. A. & Rottier, P. J. M. Equine arteritis virus-neutralizing antibody in the horse is induced by a determinant on the large envelope glycoprotein GL. *J. Gen. Virol.* **76,** 1989–1998 (1995).
80. Deregt, D., de Vries, A. A. F., Raamsman, M. J. B., Elmgren, L. D. & Rottier, P. J. M. Monoclonal antibodies to equine arteritis virus proteins identify the GL protein as a target for virus neutralization. *J. Gen. Virol.* **75,** 2439–2444 (1994).
81. Glaser, A. L., de Vries, A. A. F. & Dubovi, E. J. Comparison of equine arteritis virus isolates using neutralizing monoclonal antibodies and identification of sequence changes in GL associated with neutralization resistance. *J. Gen. Virol.* **76,** 2223–2233 (1995).
82. Balasuriya, U. B. R. *et al.* Characterization of the neutralization determinants of equine arteritis virus using recombinant chimeric viruses and site-specific mutagenesis of an infectious cDNA clone. *Virology* **321,** 235–246 (2004).

83. Snijder, E. J., Dobbe, J. C. & Spaan, W. J. M. Heterodimerization of the Two Major Envelope Proteins Is Essential for Arterivirus Infectivity. *J. Virol.* **77,** 97–104 (2003).
84. Verheije, M. H., Welting, T. J. M., Jansen, H. T., Rottier, P. J. M. & Meulenberg, J. J. M. Chimeric Arteriviruses Generated by Swapping of the M Protein Ectodomain Rule Out a Role of This Domain in Viral Targeting. *Virology* **303,** 364–373 (2002).
85. Balasuriya, U. B. R. *et al.* Expression of the Two Major Envelope Proteins of Equine Arteritis Virus as a Heterodimer Is Necessary for Induction of Neutralizing Antibodies in Mice Immunized with Recombinant Venezuelan Equine Encephalitis Virus Replicon Particles. *J. Virol.* **74,** 10623–10630 (2000).
86. Go, Y. Y. *et al.* Genome-Wide Association Study among Four Horse Breeds Identifies a Common Haplotype Associated with In Vitro CD3+ T Cell Susceptibility/Resistance to Equine Arteritis Virus Infection. *J. Virol.* **85,** 13174–13184 (2011).
87. Go, Y. Y. *et al.* Complex Interactions between the Major and Minor Envelope Proteins of Equine Arteritis Virus Determine Its Tropism for Equine CD3+ T Lymphocytes and CD14+ Monocytes. *J. Virol.* **84,** 4898–4911 (2010).
88. McCollum, W. H., Doll, E. R., Wilson, J. C. & Johnson, C. B. Propagation of equine arteritis virus in monolayer cultures of equine kidney. *Am J Vet Res* **22,** 731–735 (1961).
89. Hyllseth, B. A plaque assay of equine arteritis virus in BHK-21 cells. *Arch. Für Gesamte Virusforsch.* **28,** 26–33 (1969).
90. McCollum, W. H., Doll, E. R., Wilson, J. C. & Cheatham, J. Isolation and propagation of equine arteritis virus in monolayer cell cultures of rabbit kidney. *Cornell Vet* **52,** 452–458 (1962).
91. Shin, K., Hiroomi, A., Manabu, O. & Hiroshi, S. Studies on equine viral arteritis, 1: Characterization of the virus and trial survey on antibody with vero cell cultures. *Jpn. J. Vet. Sci.* (1975).
92. Zhang, J. Chapter one: general information. in *Permissiveness of selected cell lines to equine arteritis virus: establishment, characterization, and significance of persistent infection in hela cells* 1–40 (University of Kentucky, 2005).
93. Asagoe, T. *et al.* Effect of heparin on infection of cells by equine arteritis virus. *J. Vet. Med. Sci.* **59,** 727–728 (1997).
94. Lu, Z., Sarkar, S., Zhang, J. & Balasuriya, U. B. R. Conserved arginine residues in the carboxyl terminus of the equine arteritis virus E protein may play a role in heparin binding but may not affect viral infectivity in equine endothelial cells. *Arch. Virol.* **161,** 873–886 (2016).
95. Sarkar, S. *et al.* Equine Arteritis Virus Uses Equine CXCL16 as an Entry Receptor. *J. Virol.* **90,** 3366–3384 (2016).
96. Moore, B. D., Balasuriya, U. B. R., Hedges, J. F. & MacLachlan, N. J. Growth Characteristics of a Highly Virulent, a Moderately Virulent, and an Avirulent Strain of Equine Arteritis Virus in Primary Equine Endothelial Cells Are Predictive of Their Virulence to Horses. *Virology* **298,** 39–44 (2002).
97. Balasuriya, U. B. R. Equine Viral Arteritis. *Vet. Clin. Equine Pract.* **30,** 543–560 (2014).

98. Balasuriya, U. B. R. & MacLachlan, N. J. Equine viral arteritis. in *Equine Infectious Diseases* (eds. Sellon, D. C., Long, M. . & Saunders, W. B.) 169–181 (2013).
99. Kreutz, L. C. & Ackermann, M. R. Porcine reproductive and respiratory syndrome virus enters cells through a low pH-dependent endocytic pathway. *Virus Res.* **42,** 137–147 (1996).
100. Nauwynck, H. J., Duan, X., Favoreel, H. W., Van Oostveldt, P. & Pensaert, M. B. Entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages via receptor-mediated endocytosis. *J. Gen. Virol.* **80,** 297–305 (1999).
101. Dobbe, J. C., van der Meer, Y., Spaan, W. J. M. & Snijder, E. J. Construction of Chimeric Arteriviruses Reveals That the Ectodomain of the Major Glycoprotein Is Not the Main Determinant of Equine Arteritis Virus Tropism in Cell Culture. *Virology* **288,** 283–294 (2001).
102. Nitschke, M. *et al.* Equine arteritis virus is delivered to an acidic compartment of host cells via clathrin-dependent endocytosis. *Virology* **377,** 248–254 (2008).
103. Fang, Y. & Snijder, E. J. The PRRSV replicase: Exploring the multifunctionality of an intriguing set of nonstructural proteins. *Virus Res.* **154,** 61–76 (2010).
104. Van Hemert, M. . & Snijder, E. J. Nidoviruses. in (eds. Perlman, S., Gallagher, T. & Snijder, E. J.) (2008).
105. Snijder, E. J., Wassenaar, A. L. & Spaan, W. J. The 5' end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. *J. Virol.* **66,** 7040–7048 (1992).
106. Gorbalenya, A. E., Koonin, E. V. & Lai, M. M.-C. Putative papain-related thiol proteases of positive-strand RNA viruses Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, α- and coronaviruses. *FEBS Lett.* **288,** 201–205 (1991).
107. Stueckemann, J. A. *et al.* Replication of Lactate Dehydrogenase-elevating Virus in Macrophages. *J. Gen. Virol.* **59,** 245–262 (1982).
108. Wood, O., Tauraso, N. & Liebhaber, H. Electron Microscopic Study of Tissue Cultures Infected with Simian Haemorrhagic Fever Virus. *J. Gen. Virol.* **7,** 129–136 (1970).
109. Breese, S. S. & McCollum, W. H. *Electron microscopic characterization of equine arteritis virus*. (Karger, 1970).
110. Pol, J. M. & Wagenaar, F. Morphogenesis of Lelystad virus in porcine lung alveolar macrophages. American Association of Swine Practitioners Newsletter. 29 (1992).
111. Pol, J. M. A., Wagenaar, F. & Reus, J. E. G. Comparative morphogenesis of three PRRS virus strains. *Vet. Microbiol.* **55,** 203–208 (1997).
112. Pedersen, K. W., Meer, Y. van der, Roos, N. & Snijder, E. J. Open Reading Frame 1a-Encoded Subunits of the Arterivirus Replicase Induce Endoplasmic Reticulum-Derived Double-Membrane Vesicles Which Carry the Viral Replication Complex. *J. Virol.* **73,** 2016–2026 (1999).
113. Knoops, K. *et al.* Ultrastructural Characterization of Arterivirus Replication Structures: Reshaping the Endoplasmic Reticulum To Accommodate Viral RNA Synthesis. *J. Virol.* **86,** 2474–2487 (2012).

114. Monastyrska, I. *et al.* An autophagy-independent role for LC3 in equine arteritis virus replication. *Autophagy* **9,** 164–174 (2013).
115. Snijder, E. J. *et al.* Ultrastructure and Origin of Membrane Vesicles Associated with the Severe Acute Respiratory Syndrome Coronavirus Replication Complex. *J. Virol.* **80,** 5927–5940 (2006).
116. Meer, Y. van der *et al.* Localization of Mouse Hepatitis Virus Nonstructural Proteins and RNA Synthesis Indicates a Role for Late Endosomes in Viral Replication. *J. Virol.* **73,** 7641–7657 (1999).
117. Stertz, S. *et al.* The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* **361,** 304–315 (2007).
118. Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K. & Baker, S. C. RNA Replication of Mouse Hepatitis Virus Takes Place at Double-Membrane Vesicles. *J. Virol.* **76,** 3697–3708 (2002).
119. Meer, Y. van der, Tol, H. van, Locker, J. K. & Snijder, E. J. ORF1a-Encoded Replicase Subunits Are Involved in the Membrane Association of the Arterivirus Replication Complex. *J. Virol.* **72,** 6689–6698 (1998).
120. Hemert, M. J. van, Wilde, A. H. de, Gorbalenya, A. E. & Snijder, E. J. The in Vitro RNA Synthesizing Activity of the Isolated Arterivirus Replication/Transcription Complex Is Dependent on a Host Factor. *J. Biol. Chem.* **283,** 16525–16536 (2008).
121. Sawicki, S. G. & Sawicki, D. L. Coronaviruses use Discontinuous Extension for Synthesis of Subgenome-Length Negative Strands. in *Corona- and Related Viruses* 499–506 (Springer, Boston, MA, 1995). doi:10.1007/978-1-4615-1899-0_79
122. Lai, M. M. C. & Holmes, K. V. Coronaviruses. in *Fields Virology* (eds. Knipe, D. M. & Howley, P. M.) 1163–1185 (Lippincott Williams and Wilkins, 2001).
123. Siddell, S. G., Ziebuhr, J. & Snijder, E. J. Coronaviruses, Toroviruses, and Arteriviruses. in *Topley & Wilson's Microbiology and Microbial Infections* (John Wiley & Sons, Ltd, 2010). doi:10.1002/9780470688618.taw0245
124. Sawicki, D. L., Wang, T. & Sawicki, S. G. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J. Gen. Virol.* **82,** 385–396 (2001).
125. Gallie, D. R. A tale of two termini:: A functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* **216,** 1–11 (1998).
126. Pasternak, A. O., Born, E. van den, Spaan, W. J. M. & Snijder, E. J. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J.* **20,** 7220–7228 (2001).
127. Pasternak, A. O., Born, E. van den, Spaan, W. J. M. & Snijder, E. J. The Stability of the Duplex between Sense and Antisense Transcription-Regulating Sequences Is a Crucial Factor in Arterivirus Subgenomic mRNA Synthesis. *J. Virol.* **77,** 1175–1183 (2003).
128. Van Marle, G. *et al.* Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci.* **96,** 12056–12061 (1999).
129. Born, E. van den, Posthuma, C. C., Gultyaev, A. P. & Snijder, E. J. Discontinuous Subgenomic RNA Synthesis in Arteriviruses Is Guided by an RNA Hairpin Structure Located in the Genomic Leader Region. *J. Virol.* **79,** 6312–6324 (2005).

130. Pasternak, A. O., Gultyaev, A. P., Spaan, W. J. M. & Snijder, E. J. Genetic Manipulation of Arterivirus Alternative mRNA Leader-Body Junction Sites Reveals Tight Regulation of Structural Protein Expression. *J. Virol.* **74,** 11642–11653 (2000).
131. Pasternak, A. O., Spaan, W. J. M. & Snijder, E. J. Regulation of Relative Abundance of Arterivirus Subgenomic mRNAs. *J. Virol.* **78,** 8102–8113 (2004).
132. Brian, D. A. & Spaan, W. J. M. Recombination and Coronavirus Defective Interfering RNAs. *Semin. Virol.* **8,** 101–111 (1997).
133. Nagy, P. D. & Simon, A. E. New Insights into the Mechanisms of RNA Recombination. *Virology* **235,** 1–9 (1997).
134. Born, E. V. D., Gultyaev, A. P. & Snijder, E. J. Secondary structure and function of the 5′-proximal region of the equine arteritis virus RNA genome. *RNA* **10,** 424–437 (2004).
135. Magnusson, P., Hyllseth, B. & Marusyk, H. Morphological studies on equine arteritis virus. *Arch. Gesamte Virusforsch.* **30,** 105–112 (1970).
136. Molenkamp, R., Rozier, B. C. D., Greve, S., Spaan, W. J. M. & Snijder, E. J. Isolation and Characterization of an Arterivirus Defective Interfering RNA Genome. *J. Virol.* **74,** 3156–3165 (2000).
137. Wieringa, R., deVries, A. A. F., Post, S. M. & Rottier, P. J. M. Intra- and Intermolecular Disulfide Bonds of the GP2b Glycoprotein of Equine Arteritis Virus: Relevance for Virus Assembly and Infectivity. *J. Virol.* **77,** 12996–13004 (2003).
138. McCollum, W. H., Prickett, M. E. & Bryans, J. T. Temporal distribution of equine arteritis virus in respiratory mucosa, tissues and body fluids of horses infected by inhalation. *Res. Vet. Sci.* **12,** 459–464 (1971).
139. Cole, J. R. *et al.* Transmissibility and abortogenic effect of equine viral arteritis in mares. *J. Am. Vet. Med. Assoc.* **189,** 769–771 (1986).
140. Timoney, P. J. & McCollum, W. H. Equine Viral Arteritis. *Can. Vet. J.* **28,** 693–695 (1987).
141. Glaser, A. L., De, A. V., Rottier, P. J., Horzinek, M. C. & Colenbrander, B. [Equine arteritis virus: clinical symptoms and prevention]. *Tijdschr. Diergeneeskd.* **122,** 2–7 (1997).
142. Hedges, J. F., Balasuriya, U. B. R., Timoney, P. J., McCollum, W. H. & MacLachlan, N. J. Genetic Divergence with Emergence of Novel Phenotypic Variants of Equine Arteritis Virus during Persistent Infection of Stallions. *J. Virol.* **73,** 3672–3681 (1999).
143. Timoney, P. J. & McCollum, W. H. Equine viral arteritis: further characterization of the carrier state in stallions. *J. Reprod. Fertil. Suppl.* 3–11 (2000).
144. McFadden, A. M. *et al.* Evidence for absence of equine arteritis virus in the horse population of New Zealand. *N. Z. Vet. J.* **61,** 300–304 (2013).
145. McCollum, W. H. & Bryans, J. T. Serological Identification of Infection by Equine Arteritis Virus in Horses of Several Countries1. 256–263 (1974). doi:10.1159/000393539
146. Timoney, P. J. & McCollum, W. H. Equine viral arteritis — epidemiology and control. *J. Equine Vet. Sci.* **8,** 54–59 (1988).
147. McCue, P. M. *et al.* Prevalence of equine viral arteritis in California horses. *California Veterinarian* (1991). Available at:

https://geoscience.net/research/002/193/002193221.php. (Accessed: 8th November 2017)
148. McKenzie, J. Survey of stallions for equine viral arteritis. *Surveillance* (1996).
149. National Animal Health Monitoring System (NAHMS). *Equine Viral Arteritis (EVA) and the US Horse Industry*. (US Department of Agriculture, Animal and Plant Health Inspection Service (USDA, APHIS, VS, CEAH), 2000).
150. Zhang, J., Stein, D. A., Timoney, P. J. & Balasuriya, U. B. R. Curing of HeLa cells persistently infected with equine arteritis virus by a peptide-conjugated morpholino oligomer. *Virus Res.* **150,** 138–142 (2010).
151. Moraillon, A. & Moraillon, R. Results of an epidemiological investigation on viral arteritis in France and some other European and African countries. *Ann. Rech. Veterinaires Ann. Vet. Res.* **9,** 43–54 (1978).
152. Burki, F., Hofer, A. & Nowotny, N. Objective Data Plead to Suspend Import-Bans for Seroreactors Against Equine Arteritis Virus Except for Breeder Stallions. *J. Appl. Anim. Res.* **1,** 31–42 (1992).
153. Cruz, F. *et al.* Seroprevalence and factors associated with seropositivity to equine arteritis virus in Spanish Purebred horses in Spain. *Equine Vet. J.* **48,** 573–577 (2016).
154. Balasuriya, U. B. R., Carossino, M. & Timoney, P. J. Equine viral arteritis: A respiratory and reproductive disease of significant economic importance to the equine industry. *Equine Vet. Educ.* n/a-n/a (2016). doi:10.1111/eve.12672
155. Go, Y. Y., Bailey, E., Timoney, P. J., Shuck, K. M. & Balasuriya, U. B. R. Evidence that In Vitro Susceptibility of CD3+ T Lymphocytes to Equine Arteritis Virus Infection Reflects Genetic Predisposition of Naturally Infected Stallions To Become Carriers of the Virus. *J. Virol.* **86,** 12407–12410 (2012).
156. Timoney, P. J., McCollum, W. H., Roberts, A. W. & Murphy, T. W. Demonstration of the carrier state in naturally acquired equine arteritis virus infection in the stallion. *Res. Vet. Sci.* **41,** 279–280 (1986).
157. Timoney, P. J. *et al.* The carrier state in equine arteritis virus infection in the stallion with specific emphasis on the venereal mode of virus transmission. *J. Reprod. Fertil. Suppl.* **35,** 95–102 (1987).
158. McCollum, W. H., Timoney, P. J. & Tengelsen, L. A. Clinical, virological and serological responses of donkeys to intranasal inoculation with the KY-84 strain of equine arteritis virus. *J. Comp. Pathol.* **112,** 207–211 (1995).
159. McCollum, W. H. Pathologic features of horses given avirulent equine arteritis virus intramuscularly. *Am. J. Vet. Res.* **42,** 1218–1220 (1981).
160. Guthrie, A. J. *et al.* Lateral transmission of equine arteritis virus among Lipizzaner stallions in South Africa. *Equine Vet. J.* **35,** 596–600 (2003).
161. McCollum, W. H. Responses of horses vaccinated with avirulent modified-live equine arteritis virus propagated in the E. Derm (NBL-6) cell line to nasal inoculation with virulent virus. *Am. J. Vet. Res.* **47,** 1931–1934 (1986).
162. Balasuriya, U. B. R. *et al.* Equine Arteritis Virus Derived from an Infectious cDNA Clone Is Attenuated and Genetically Stable in Infected Stallions. *Virology* **260,** 201–208 (1999).
163. Balasuriya, U. B. *et al.* Serologic and molecular characterization of an abortigenic strain of equine arteritis virus isolated from infective frozen semen and an aborted equine fetus. *J. Am. Vet. Med. Assoc.* **213,** 1586–9, 1570 (1998).

164. Balasuriya, U. B. R. *et al.* Genetic stability of equine arteritis virus during horizontal and vertical transmission in an outbreak of equine viral arteritis. *J. Gen. Virol.* **80,** 1949–1958 (1999).
165. Carossino, M. *et al.* Equine Arteritis Virus Has Specific Tropism for Stromal Cells and CD8+ T and CD21+ B Lymphocytes but Not for Glandular Epithelium at the Primary Site of Persistent Infection in the Stallion Reproductive Tract. *J. Virol.* **91,** e00418-17 (2017).
166. Balasuriya, U. R., Hedges, J. & Maclachlan, N. J. Molecular Epidemiology and Evolution of Equine Arteritis Virus. in *The Nidoviruses* 19–24 (Springer, Boston, MA, 2001). doi:10.1007/978-1-4615-1325-4_2
167. Broaddus, C. C. *et al.* Infection of embryos following insemination of donor mares with equine arteritis virus infective semen. *Theriogenology* **76,** 47–60 (2011).
168. Balasuriya, U. B. & MacLACHLAN, N. J. Equine Viral Arteritis. in *Equine infectious diseases.* (eds. Sellon, D. C. & Long, M. .) 153–164 (Saunders/Elsevier, 2007).
169. Golnik, W., Michalska, Z. & Michalak, T. Natural equine viral arteritis in foals. *Schweiz. Arch. Tierheilkd.* **123,** 523–533 (1981).
170. Carman, S., Rae, C. & DuBovi, E. Ontario. Equine arteritis virus isolated from a Standardbred foal with pneumonia. *Can. Vet. J.* **29,** 937 (1988).
171. Vaala, W. E., Hamir, A. N., Dubovi, E. J., Timoney, P. J. & Ruiz, B. Fatal, congenitally acquired infection with equine arteritis virus in a neonatal Thoroughbred. *Equine Vet. J.* **24,** 155–158 (1992).
172. Wilkins, P. A., Del Piero, F., Lopez, J. & Cline, M. Recognition of bronchopulmonary dysplasia in a newborn foal. *Equine Vet. J.* **27,** 398 (1995).
173. Del Piero, F. *et al.* Equine viral arteritis in newborn foals: clinical, pathological, serological, microbiological and immunohistochemical observations. *Equine Vet. J.* **29,** 178–185 (1997).
174. Go, Y. Y. *et al.* Equine Arteritis Virus Does Not Induce Interferon Production in Equine Endothelial Cells: Identification of Nonstructural Protein 1 as a Main Interferon Antagonist. *BioMed Research International* (2014). doi:10.1155/2014/420658
175. McCollum, W. H. Vaccination for equine viral arteritis. in 143–151 (1969).
176. Fukunaga, Y. & McCollum, W. H. Complement-fixation reactions in equine viral arteritis. *Am. J. Vet. Res.* **38,** 2043–2046 (1977).
177. Balasuriya, U. B. R. *et al.* Alphavirus replicon particles expressing the two major envelope proteins of equine arteritis virus induce high level protection against challenge with virulent virus in vaccinated horses. *Vaccine* **20,** 1609–1617 (2002).
178. Balasuriya, U. B., Dobbe, J. C. & Heidner, H. W. Characterization of the virulence determinants of equine arteritis virus using an infectious cDNA clone derived from the virulent Bucyrus strain. in (2005).
179. Balasuriya, U. B. R. & MacLachlan, N. J. The immune response to equine arteritis virus: potential lessons for other arteriviruses. *Vet. Immunol. Immunopathol.* **102,** 107–129 (2004).
180. Gerber, H. S., Steck, F., Hofer, A., Walther, L. & Friedly, U. Clinical and serological investigations on equine viral arteritis. in (1978).

181. Glaser, A. L., Chirnside, E. D., Horzinek, M. C. & de Vries, A. A. F. Equine arteritis virus. *Theriogenology* **47,** 1275–1295 (1997).
182. Johnson, B., Baldwin, C., Timoney, P. & Ely, R. Arteritis in equine fetuses aborted due to equine viral arteritis. *Vet. Pathol.* **28,** 248–250 (1991).
183. MacLachlan, N. J. *et al.* Fatal experimental equine arteritis virus infection of a pregnant mare: immunohistochemical staining of viral antigens. *J. Vet. Diagn. Investig. Off. Publ. Am. Assoc. Vet. Lab. Diagn. Inc* **8,** 367–374 (1996).
184. McCollum, W. H. *et al.* Features of an outbreak of equine viral arteritis on a breeding farm associated with abortion and fatal interstitial pneumonia in neonatal foals. in *8th International Conference* 559–560 (1998).
185. Moore, B. D. *et al.* Differentiation of strains of equine arteritis virus of differing virulence to horses by growth in equine endothelial cells. *Am. J. Vet. Res.* **64,** 779–784 (2003).
186. Moore, B. D. *et al.* Virulent and avirulent strains of equine arteritis virus induce different quantities of TNF-α and other proinflammatory cytokines in alveolar and blood-derived equine macrophages. *Virology* **314,** 662–670 (2003).
187. Zhang, J. *et al.* Amino acid substitutions in the structural or nonstructural proteins of a vaccine strain of equine arteritis virus are associated with its attenuation. *Virology* **378,** 355–362 (2008).
188. Zhang, J. *et al.* Development and Characterization of an Infectious cDNA Clone of the Modified Live Virus Vaccine Strain of Equine Arteritis Virus. *Clin. Vaccine Immunol.* **19,** 1312–1321 (2012).
189. Vairo, S. *et al.* Clinical and virological outcome of an infection with the Belgian equine arteritis virus strain 08P178. *Vet. Microbiol.* **157,** 333–344 (2012).
190. Vairo, S., Saey, V., Bombardi, C., Ducatelle, R. & Nauwynck, H. The Recent European Isolate (08P178) of Equine Arteritis Virus causes Inflammation but not Arteritis in Experimentally Infected Ponies. *J. Comp. Pathol.* **151,** 238–243 (2014).
191. Collins, J. K. *et al.* Equine viral arteritis at a veterinary teaching hospital. *Prev. Vet. Med.* **4,** 389–397 (1987).
192. Jaksch, W., Sibalin, M., Taussig, E., Pichler, L. & Bürki, F. [Natural cases and experimental transmissions of equine-virus-arteritis in Austria]. *DTW Dtsch. Tierarztl. Wochenschr.* **80,** 374–80 concl (1973).
193. Jones, T. C. Clinical and pathologic features of equine viral arteritis. *J. Am. Vet. Med. Assoc.* **155,** 315–317 (1969).
194. Bryans, J. T., Doll, E. R. & Jones, T. C. The lesions of equine viral arteritis. *Cornell Vet.* **47,** 52–68 (1957).
195. Clayton, H. 1986 outbreak of EAV in Alberta, Canada. *J Equine Vet Sci* **7,** 101 (1987).
196. Crawford, T. B. & Henson, J. B. Viral Arteritis of Horses. in *Comparative Pathophysiology of Circulatory Disturbances* 175–183 (Springer, Boston, MA, 1972). doi:10.1007/978-1-4684-3213-8_11
197. Estes, P. C. & Cheville, N. F. The ultrastructure of vascular lesions in equine viral arteritis. *Am. J. Pathol.* **58,** 235–253 (1970).
198. Prickett, M. E., McCollum, W. H. & Bryans, J. T. The Gross and Microscopic Pathology Observed in Horses Experimentally Infected with the Equine Arteritis Virus1. 265–272 (1974). doi:10.1159/000393541

199. Balasuriya, U. B. R. *et al.* Development and characterization of an infectious cDNA clone of the virulent Bucyrus strain of Equine arteritis virus. *J. Gen. Virol.* **88,** 918–924 (2007).
200. Crawford, T. B. & Henson, J. B. Immunofluorescent, Light-Microscopic and Immunologic Studies of Equine Viral Arteritis. 282–302 (1974). doi:10.1159/000393543
201. Lopez, J. W., Del Piero, F., Glaser, A. & Finazzi, M. Immunoperoxidase histochemistry as a diagnostic tool for detection of equine arteritis virus antigen in formalin fixed tissues. *Equine Vet. J.* **28,** 77–79 (1996).
202. Del Piero, F. Equine Viral Arteritis. *Vet. Pathol.* **37,** 287–296 (2000).
203. Plagemann, P. G. W. & Moennig, V. Lactate Dehydrogenase-Elevating Virus, Equine Arteritis Virus, and Simian Hemorrhagic Fever Virus: A New Group of Positive-Strand RNA Viruses. in *Advances in Virus Research* (eds. Maramorosch, K., Murphy, F. A. & Shatkin, A. J.) **41,** 99–192 (Academic Press, 1992).
204. Vairo, S., Van den Broeck, W., Favoreel, H., Scagliarini, A. & Nauwynck, H. Development and use of a polarized equine upper respiratory tract mucosal explant system to study the early phase of pathogenesis of a European strain of equine arteritis virus. *Vet. Res.* **44,** 22 (2013).
205. Go, Y. Y. *et al.* Assessment of correlation between in vitro CD3+ T cell susceptibility to EAV infection and clinical outcome following experimental infection. *Vet. Microbiol.* **157,** 220–225 (2012).
206. McCollum, W. H. Development of a modified virus strain and vaccine for equine viral arteritis. *J. Am. Vet. Med. Assoc.* **155,** 318–322 (1969).
207. Fukunaga, Y., Imagawa, H., Tabuchi, E. & Akiyama, Y. (Japan R. A. Clinical and virological findings on experimental equine viral arteritis in horses. *Exp. Rep. Equine Health Lab. Jpn.* (1981).
208. Neu, S. M., Timoney, P. J. & McCOLLUM, W. H. Persistent infection of the reproductive tract in stallions experimentally infected with equine arteritis virus. in (ed. Powell, D. G.) 149–154 (The University of Kentucky, 1987).
209. Little, T. V., Holyoak, G. R., McCollum, W. H. & Timoney, P. J. (Department of V. S. Output of equine arteritis virus from persistently infected stallions is testosterone-dependent. (1992).
210. Neu, S. M., Timoney, P. J. & Lowry, S. R. Changes in semen quality in the stallion following experimental infection with equine arteritis virus. *Theriogenology* **37,** 407–431 (1992).
211. Campos, J. R. *et al.* Semen quality of stallions challenged with the Kentucky 84 strain of equine arteritis virus. *Theriogenology* **82,** 1068–1079 (2014).
212. McCollum, W. H., Little, T. V., Timoney, P. J. & Swerczek, T. W. Resistance of castrated male horses to attempted establishment of the carrier state with equine arteritis virus. *J. Comp. Pathol.* **111,** 383–388 (1994).
213. McDonald, L. E. Veterinary endocrinology and reproduction. *Vet. Endocrinol. Reprod.* (1980).
214. Shivaji, S., Scheit, K. H. & Bhargava, P. M. *Immunosuppressive factors in seminal plasma*. (1990).

215. Hedges, J. F. *et al.* Detection of antibodies to equine arteritis virus by enzyme linked immunosorbant assays utilizing GL, M and N proteins expressed from recombinant baculoviruses. *J. Virol. Methods* **76,** 127–137 (1998).
216. Holyoak, G. R., Giles, R. C., McCollum, W. H., Little, T. V. & Timoney, P. J. Pathological changes associated with equine arteritis virus infection of the reproductive tract in prepubertal and peripubertal colts. *J. Comp. Pathol.* **109,** 281–293 (1993).
217. Holyoak, G. R., Little, T. V., McCollum, W. H. & Timoney, P. J. Relationship between onset of puberty and establishment of persistent infection with equine arteritis virus in the experimentally infected colt. *J. Comp. Pathol.* **109,** 29–46 (1993).
218. Patton, J. F. *et al.* Phylogenetic characterization of a highly attenuated strain of equine arteritis virus from the semen of a persistently infected Standardbred stallion. *Arch. Virol.* **144,** 817–827 (1999).
219. Balasuriya, U. B. Equine arteritis virus. in *Molecular detection of animal viral pathogens* (2016).
220. Del Piero, F. Diagnosis of Equine Arteritis Virus Infection in Two Horses by Using Monoclonal Antibody Immunoperoxidase Histochemistry on Skin Biopsies. *Vet. Pathol.* **37,** 486–487 (2000).
221. OIE. *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. (Office International des Epizooties, 2004).
222. Senne, D. A., Pearson, J. E. & Cabrey, E. A. Equine viral arteritis: a standard procedure for the virus neutralization test and comparison of results of a proficiency test performed at five laboratories. in *89th annual meeting of the United States Animal Health Association* 29–34
223. Miszczak, F. *et al.* Evaluation of Two Magnetic-Bead-Based Viral Nucleic Acid Purification Kits and Three Real-Time Reverse Transcription-PCR Reagent Systems in Two TaqMan Assays for Equine Arteritis Virus Detection in Semen. *J. Clin. Microbiol.* **49,** 3694–3696 (2011).
224. Go, Y. Y. Literature review. in *Molecular and genomic approaches to understanding host-virus interactions in shaping the outcome of equine arteritis virus infection. University of Kentucky Doctoral Disserations.* 1–55 (University of Kentucky, 2011).
225. Chimside, E. D. & Spaan, W. J. M. Reverse transcription and cDNA amplification by the polymerase chain reaction of equine arteritis virus (EAV). *J. Virol. Methods* **30,** 133–140 (1990).
226. Gilbert, S. A., Timoney, P. J., McCollum, W. H. & Deregt, D. Detection of equine arteritis virus in the semen of carrier stallions by using a sensitive nested PCR assay. *J. Clin. Microbiol.* **35,** 2181–2183 (1997).
227. Ramina, A. *et al.* Detection of equine arteritis virus in semen by reverse transcriptase polymerase chain reaction–ELISA. *Comp. Immunol. Microbiol. Infect. Dis.* **22,** 187–197 (1999).
228. Starick, E. Rapid and sensitive detection of equine arteritis virus in semen and tissue samples by reverse transcription-polymerase chain reaction, dot blot hybridisation and nested polymerase chain reaction. *Acta Virol.* **42,** 333–339 (1998).

229. St-Laurent, G., Morin, G. & Archambault, D. Detection of equine arteritis virus following amplification of structural and nonstructural viral genes by reverse transcription-PCR. *J. Clin. Microbiol.* **32,** 658–665 (1994).
230. Balasuriya, U. B. R. *et al.* Detection of equine arteritis virus by real-time TaqMan® reverse transcription-PCR assay. *J. Virol. Methods* **101,** 21–28 (2002).
231. Westcott, D. G. *et al.* Use of an internal standard in a closed one-tube RT-PCR for the detection of equine arteritis virus RNA with fluorescent probes. *Vet. Res.* **34,** 165–176 (2003).
232. Miszczak, F. *et al.* Evaluation of two magnetic-bead-based viral nucleic acid purification kits and three real-time reverse transcription-PCR reagent systems in two TaqMan assays for equine arteritis virus detection in semen. *J. Clin. Microbiol.* **49,** 3694–6 (2011).
233. Balasuriya, U. B. R. Type A Influenza Virus Detection from Horses by Real-Time RT-PCR and Insulated Isothermal RT-PCR. in *Animal Influenza Virus* 393–402 (Humana Press, New York, NY, 2014). doi:10.1007/978-1-4939-0758-8_34
234. Tsai, Y.-L. *et al.* Development of TaqMan Probe-Based Insulated Isothermal PCR (iiPCR) for Sensitive and Specific On-Site Pathogen Detection. *PLOS ONE* **7,** e45278 (2012).
235. Chang, H.-F. G. *et al.* A thermally baffled device for highly stabilized convective PCR. *Biotechnol. J.* **7,** 662–666 (2012).
236. Carossino, M. *et al.* Detection of equine arteritis virus by two chromogenic RNA in situ hybridization assays (conventional and RNAscope(®)) and assessment of their performance in tissues from aborted equine fetuses. *Arch. Virol.* **161,** 3125–3136 (2016).
237. Nugent, J. *et al.* Development and evaluation of ELISA procedures to detect antibodies against the major envelope protein (GL) of equine arteritis virus. *J. Virol. Methods* **90,** 167–183 (2000).
238. Cho, H. J. *et al.* Detection of antibodies to equine arteritis virus by a monoclonal antibody-based blocking ELISA. *Can. J. Vet. Res.* **64,** 38–43 (2000).
239. Chirnside, E. D., Francis, P. M., De Vries, A. A. F., Sinclaira, R. & Mumford, J. A. Development and evaluation of an ELISA using recombinant fusion protein to detect the presence of host antibody to equine arteritis virus. *J. Virol. Methods* **54,** 1–13 (1995).
240. Chirnside, E. D., Francis, P. M. & Mumford, J. A. Expression cloning and antigenic analysis of the nucleocapsid protein of equine arteritis virus. *Virus Res.* **39,** 277–288 (1995).
241. Kondo, T., Fukunaga, Y., Sekiguchi, K., Sugiura, T. & Imagawa, H. Enzyme-Linked Immunosorbent Assay for Serological Survey of Equine Arteritis Virus in Racehorses. *J. Vet. Med. Sci.* **60,** 1043–1045 (1998).
242. Starick, E., Ginter, A. & Coppe, P. ELISA and Direct Immunofluorescence Test to Detect Equine Arteritis Virus (EAV) using a Monoclonal Antibody Directed to the EAV-N Protein. *J. Vet. Med. Ser. B* **48,** 1–9 (2001).
243. Wagner, H. M., Balasuriya, U. B. R. & James MacLachlan, N. The serologic response of horses to equine arteritis virus as determined by competitive enzyme-linked immunosorbent assays (c-ELISAs) to structural and non-structural viral proteins. *Comp. Immunol. Microbiol. Infect. Dis.* **26,** 251–260 (2003).

244. Chirnside, E. D., Wearing, C. M., Binns, M. M. & Mumford, J. A. Comparison of M and N gene sequences distinguishes variation amongst equine arteritis virus isolates. *J. Gen. Virol.* **75,** 1491–1497 (1994).
245. Jeronimo, C. & Archambault, D. Importance of M-Protein C Terminus as Substrate Antigen for Serodetection of Equine Arteritis Virus Infection. *Clin. Diagn. Lab. Immunol.* **9,** 698–703 (2002).
246. van den Born, E., Stein, D. A., Iversen, P. L. & Snijder, E. J. Antiviral activity of morpholino oligomers designed to block various aspects of Equine arteritis virus amplification in cell culture. *J. Gen. Virol.* **86,** 3081–3090 (2005).
247. Summerton, J. & Weller, D. Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense Nucleic Acid Drug Dev.* **7,** 187–195 (1997).
248. Summerton, J. Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* **1489,** 141–158 (1999).
249. Stein, D., Foster, E., Huang, S.-B., Weller, D. & Summerton, J. A Specificity Comparison of Four Antisense Types: Morpholino, 2′-O-Methyl RNA, DNA, and Phosphorothioate DNA. *Antisense Nucleic Acid Drug Dev.* **7,** 151–157 (1997).
250. Abes, S. *et al.* Vectorization of morpholino oligomers by the (R-Ahx-R)4 peptide allows efficient splicing correction in the absence of endosomolytic agents. *J. Controlled Release* **116,** 304–313 (2006).
251. Dowsett, K. F. *et al.* A preliminary study of immunological castration in colts. *J. Reprod. Fertil. Suppl.* **44,** 183–190 (1991).
252. Fortier, G., Vidament, M., DeCraene, F., Ferry, B. & Daels, P. F. The effect of GnRH antagonist on testosterone secretion, spermatogenesis and viral excretion in EVA-virus excreting stallions. in (2002).
253. Clement, F. *et al.* Immunocastration in stallions: effect on spermatogenesis and behaviour. *Anim. Reprod. Sci.* **89,** 230–233 (2005).
254. Turkstra, J. A. *et al.* Effects of GnRH immunization in sexually mature pony stallions. *Anim. Reprod. Sci.* **86,** 247–259 (2005).
255. Burger, D. *et al.* Immunization against GnRH in adult stallions: effects on semen characteristics, behaviour and shedding of equine arteritis virus. *Proc. Ninth Int. Symp. Equine Reprod. 1-4 94 2006 9 Int. Symp. Equine Reprod. Kerkrade NLD 2006-08-06-2006-08-11 107-111* (2006).
256. Morrell, J. M. & Geraghty, R. M. Effective removal of equine arteritis virus from stallion semen. *Equine Vet. J.* **38,** 224–229 (2006).
257. Janett, F., Stump, R., Burger, D. & Thun, R. Suppression of testicular function and sexual behavior by vaccination against GnRH (Equity™) in the adult stallion. *Anim. Reprod. Sci.* **115,** 88–102 (2009).
258. Wenzinger, B., Kähn, W. & Bleul, U. [The use of a GnRH vaccine in mares and stallions to influence undesirable behavior: a retrospective study of 31 cases]. *Schweiz. Arch. Tierheilkd.* **152,** 373–377 (2010).
259. Morrell, J. *et al.* Single-Layer Centrifugation Reduces Equine Arteritis Virus Titre in the Semen of Shedding Stallions. *Reprod. Domest. Anim.* **48,** 604–612 (2013).
260. United States Department of Agriculture - Animal and Plant Health Inspection Service (USDA-APHIS). *Equine viral arteritis: Uniform Methods and Rules.* (United

States Department of Agriculture - Animal and Plant Health Inspection Service (USDA-APHIS), 2004).
261. OIE. *Infection with equine arteritis virus. Terrestrial Animal Health Code, 24th edn*. 1–4 (Office International des Epizooties, 2015).
262. HBLB. *Code of practice for equine viral arteritis. In: Codes of Practice*. 19–30 (HBLB, 2016).
263. OIE. *Equine viral arteritis. In: Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 7th edn*. 1–16 (Office International des Epizooties, 2013).
264. Timoney, P. J., McCollum, W. H., Roberts, A. W. & McDonald, M. J. Status of equine viral arteritis in Kentucky for 1986. *Vet. Rec.* **120,** 282 (1987).
265. Timoney, P. J. The Increasing Significance of International Trade in Equids and Its Influence on the Spread of Infectious Diseases. *Ann. N. Y. Acad. Sci.* **916,** 55–60 (2000).
266. Timoney, P. J. Factors Influencing the International Spread of Equine Diseases. *Vet. Clin. North Am. Equine Pract.* **16,** 537–551 (2000).
267. Olguin Perglione, C. *et al.* Equine viral arteritis outbreak in Argentina. in (2010).
268. Harry, T. O. & McCollum, W. H. Stability of viability and immunizing potency of lyophilized, modified equine arteritis live-virus vaccine. *Am. J. Vet. Res.* **42,** 1501–1505 (1981).
269. Doll, E. R., Bryans, J. T., Wilson, J. C. & McCollum, W. H. Immunization against equine viral arteritis using modified live virus propagated in cell cultures of rabbit kidney. *Cornell Vet.* **48,** 497–524 (1968).
270. Summers-Lawyer, K. A. *et al.* Response of Stallions to Primary Immunization with a Modified Live Equine Viral Arteritis Vaccine. *J. Equine Vet. Sci.* **31,** 129–138 (2011).
271. Fukunaga, Y., Wada, R., Matsumura, T., Sugiura, T. & Imagawa, H. Induction of Immune Response and Protection from Equine Viral Arteritis (EVA) by Formalin Inactivated-virus Vaccine for EVA in Horses. *J. Vet. Med. Ser. B* **37,** 135–141 (1990).
272. Mckinnon, A. O. *et al.* Vaccination of stallions with a modified live equine viralarteritis virus. *J. Equine Vet. Sci.* **6,** 66–69 (1986).
273. Timoney, P. J. *et al.* The outcome of vaccinating five pregnant mares with a commercial equine viral arteritis vaccine. *Equine Vet. Educ.* **19,** 606–611 (2007).
274. Fukunaga, Y. *et al.* Effect of the modified Bucyrus strain of equine arteritis virus experimentally inoculated into horses. *Exp. Rep. Equine Health Lab. Jpn.* (1982).
275. Fukunaga, Y. *et al.* Immune potency of lyophilized, killed vaccine for equine viral arteritis and its protection against abortion in pregnant mares. *J. Equine Vet. Sci.* **16,** 217–221 (1996).
276. Broaddus, C. C. *et al.* Evaluation of the safety of vaccinating mares against equine viral arteritis during mid or late gestation or during the immediate postpartum period. *J. Am. Vet. Med. Assoc.* **238,** 741–750 (2011).
277. Tobiasch, E. *et al.* Large Envelope Glycoprotein and Nucleocapsid Protein of Equine Arteritis Virus (EAV) Induce an Immune Response in Balb/c Mice by DNA Vaccination; Strategy for Developing a DNA-Vaccine Against EAV-Infection. *Virus Genes* **22,** 187–199 (2001).
278. Giese, M. *et al.* Stable and Long-Lasting Immune Response in Horses after DNA Vaccination against Equine Arteritis Virus. *Virus Genes* **25,** 159–167 (2002).

279. John, C. & Venetia, S. Origins and evolution of viruses. in *Virology: Principles and Applications* 258–271 (Wiley, 2007).
280. Andino, R. & Domingo, E. Viral quasispecies. *Virology* **479–480,** 46–51 (2015).
281. Holland, J. J., Torre, J. C. D. L. & Steinhauer, D. A. RNA Virus Populations as Quasispecies. in *Genetic Diversity of RNA Viruses* 1–20 (Springer, Berlin, Heidelberg, 1992). doi:10.1007/978-3-642-77011-1_1
282. Domingo, E. & Holland, J. J. Rna Virus Mutations and Fitness for Survival. *Annu. Rev. Microbiol.* **51,** 151–178 (1997).
283. Drake, J. W. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci.* **90,** 4171–4175 (1993).
284. Flint, S. J., Enquist, L. W., Racaniello, V. R. & Skalka, A. M. Synthesis of RNA from RNA templates. in *Principles of Virology* **1. Molecular Biology,** 195–199 (American Society for Microbiology Press, 2009).
285. Domingo, E., Sheldon, J. & Perales, C. Viral Quasispecies Evolution. *Microbiol. Mol. Biol. Rev.* **76,** 159–216 (2012).
286. Castro, C., Arnold, J. J. & Cameron, C. E. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res.* **107,** 141–149 (2005).
287. Shimizu, K. [Mechanisms of antigenic variation in influenza virus]. *Nihon Rinsho Jpn. J. Clin. Med.* **58,** 2199–2205 (2000).
288. Parham, P. Glossary. in *The immune system* G:1-G:30 (Garland Science, 2015).
289. Eigen, M. & Schuster, P. *The hypercycle. A principle of natural selforganization.* (Springer, 1979).
290. Domingo, E., Flavell, R. A. & Weissmann, C. In vitro site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Qbeta. *Gene* **1,** 3–25 (1976).
291. Batschelet, E., Domingo, E. & Weissmann, C. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* **1,** 27–32 (1976).
292. Drake, J. W. Comparative rates of spontaneous mutation. *Nature* **221,** 1132 (1969).
293. Alves, D. & Fontanari, J. F. Error threshold in finite populations. *Phys. Rev. E* **57,** 7008–7013 (1998).
294. Eigen, M. Natural selection: a phase transition? *Biophys. Chem.* **85,** 101–123 (2000).
295. Nowak, M. A. *Evolutionary dynamics*. (Belknap Pres, 2006).
296. Manrubia, S. C., García-Arriaza, J., Domingo, E. & Escarmís, C. Long-range transport and universality classes in in vitro viral infection spread. *EPL Europhys. Lett.* **74,** 547 (2006).
297. Ochoa, G. Error thresholds in genetic algorithms. *Evol. Comput.* **14,** 157–182 (2006).
298. Park, J.-M., Muñoz, E. & Deem, M. W. Quasispecies theory for finite populations. *Phys. Rev. E* **81,** 011902 (2010).
299. Saakian, D. B., Biebricher, C. K. & Hu, C.-K. Phase diagram for the Eigen quasispecies theory with a truncated fitness landscape. *Phys. Rev. E* **79,** 041905 (2009).

300. Saakian, D. B. & Hu, C.-K. Exact solution of the Eigen model with general fitness functions and degradation rates. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4935–4939 (2006).
301. Saakian, D. B., Muñoz, E., Hu, C.-K. & Deem, M. W. Quasispecies theory for multiple-peak fitness landscapes. *Phys. Rev. E* **73,** 041913 (2006).
302. Schuster, P. Genotypes and Phenotypes in the Evolution of Molecules*. *Eur. Rev.* **17,** 281–319 (2009).
303. Takeuchi, N. & Hogeweg, P. Error-threshold exists in fitness landscapes with lethal mutants. *BMC Evol. Biol.* **7,** 15; author reply 15 (2007).
304. Wilke, C. O., Ronnewinkel, C. & Martinetz, T. Dynamic fitness landscapes in molecular evolution. *Phys. Rep.* **349,** 395–446 (2001).
305. Page, K. M. & Nowak, M. A. Unifying Evolutionary Dynamics. *J. Theor. Biol.* **219,** 93–98 (2002).
306. Domingo, E., Holland, J. J., Biebricher, C. & Eigen, M. Quasispecies: the concept and the word. in *Molecular evolution of the viruses* (eds. Gibbs, A., Calisher, C. & García-Arenal, F.) 171–180 (Cambridge University Press, 1995).
307. Domingo, E. *Quasispecies: Concept and Implications for Virology*. (Springer Science & Business Media, 2006).
308. Perales, C., Lorenzo-Redondo, R., López-Galíndez, C., Martínez, M. A. & Domingo, E. Mutant spectra in virus behavior. *Future Virol.* **5,** 679–698 (2010).
309. Perales, C., Martín, V., Ruiz-Jarabo, C. M. & Domingo, E. Monitoring Sequence Space as a Test for the Target of Selection in Viruses. *J. Mol. Biol.* **345,** 451–459 (2005).
310. Escarmís, C., Dávila, M. & Domingo, E. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* **285,** 495–505 (1999).
311. Lorenzo-Redondo, R., Bordería, A. V. & Lopez-Galindez, C. Dynamics of in vitro fitness recovery of HIV-1. *J. Virol.* **85,** 1861–1870 (2011).
312. Novella, I. S. *et al.* Exponential increases of RNA virus fitness during large population transmissions. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 5841–5844 (1995).
313. Ruiz-Jarabo, C. M., Arias, A., Baranowski, E., Escarmís, C. & Domingo, E. Memory in Viral Quasispecies. *J. Virol.* **74,** 3543–3547 (2000).
314. Briones, C. & Domingo, E. *Minority Report: Hidden Memory Genomes in HIV1 Quasispecies and Possible Clinical Implications*. **10,** (2008).
315. Arias, A., Ruiz-Jarabo, C. M., Escarmís, C. & Domingo, E. Fitness Increase of Memory Genomes in a Viral Quasispecies. *J. Mol. Biol.* **339,** 405–412 (2004).
316. Van Valen, L. A New Evolutionary Law. *Evol Theo* **1,** 1—30 (1973).
317. Aragonès, L., Guix, S., Ribes, E., Bosch, A. & Pintó, R. M. Fine-Tuning Translation Kinetics Selection as the Driving Force of Codon Usage Bias in the Hepatitis A Virus Capsid. *PLOS Pathog.* **6,** e1000797 (2010).
318. Clarke, D. K. *et al.* The red queen reigns in the kingdom of RNA viruses. *Proc. Natl. Acad. Sci.* **91,** 4821–4824 (1994).
319. Kashiwagi, A. & Yomo, T. Ongoing Phenotypic and Genomic Changes in Experimental Coevolution of RNA Bacteriophage Qβ and Escherichia coli. *PLOS Genet.* **7,** e1002188 (2011).
320. Liu, J., Chen, K., Wang, J.-H. & Zhang, C. Molecular Evolution of the Primate Antiviral Restriction Factor Tetherin. *PLOS ONE* **5,** e11904 (2010).

321. Ruiz-Jarabo, C. M., Miller, E., Gómez-Mariano, G. & Domingo, E. Synchronous Loss of Quasispecies Memory in Parallel Viral Lineages: A Deterministic Feature of Viral Quasispecies. *J. Mol. Biol.* **333,** 553–563 (2003).
322. Acharya, R. *et al.* The three-dimensional structure of foot-and-mouth disease virus at 2.9 A resolution. *Nature* **337,** 709–716 (1989).
323. Escarmís, C. *et al.* Genetic Lesions Associated with Muller's Ratchet in an RNA Virus. *J. Mol. Biol.* **264,** 255–267 (1996).
324. Escarmís, C., Perales, C. & Domingo, E. Biological Effect of Muller's Ratchet: Distant Capsid Site Can Affect Picornavirus Protein Processing. *J. Virol.* **83,** 6748–6756 (2009).
325. Escarmís, C., Lázaro, E., Arias, A. & Domingo, E. Repeated Bottleneck Transfers Can Lead to Non-cytocidal Forms of a Cytopathic Virus: Implications for Viral Extinction. *J. Mol. Biol.* **376,** 367–379 (2008).
326. de la Torre, J. C., DáVila, M., Sobrino, F., Ortín, J. & Domingo, E. Establishment of cell lines persistently infected with foot-and-mouth disease virus. *Virology* **145,** 24–35 (1985).
327. Herrera, M., Grande-Pérez, A., Perales, C. & Domingo, E. Persistence of foot-and-mouth disease virus in cell culture revisited: implications for contingency in evolution. *J. Gen. Virol.* **89,** 232–244 (2008).
328. Marcus, P. I., Rodriguez, L. L. & Sekellick, M. J. Interferon Induction as a Quasispecies Marker of Vesicular Stomatitis Virus Populations. *J. Virol.* **72,** 542–549 (1998).
329. Meyerhans, A. *et al.* Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58,** 901–910 (1989).
330. Artenstein, M. S. & Miller, W. S. Air sampling for respiratory disease agents in army recruits. *Bacteriol. Rev.* **30,** 571–572 (1966).
331. Couch, R. B., Cate, T. R., Douglas, R. G., Gerone, P. J. & Knight, V. Effect of route of inoculation on experimental respiratory viral disease in volunteers and evidence for airborne transmission. *Bacteriol. Rev.* **30,** 517–529 (1966).
332. Ali, A. *et al.* Analysis of Genetic Bottlenecks during Horizontal Transmission of Cucumber Mosaic Virus. *J. Virol.* **80,** 8345–8350 (2006).
333. Betancourt, M., Fereres, A., Fraile, A. & García-Arenal, F. Estimation of the Effective Number of Founders That Initiate an Infection after Aphid Transmission of a Multipartite Plant Virus. *J. Virol.* **82,** 12416–12421 (2008).
334. Bull, R. A. *et al.* Sequential Bottlenecks Drive Viral Evolution in Early Acute Hepatitis C Virus Infection. *PLOS Pathog.* **7,** e1002243 (2011).
335. Foy, B. D. *et al.* Development of a new Sindbis virus transducing system and its characterization in three Culicine mosquitoes and two Lepidopteran species. *Insect Mol. Biol.* **13,** 89–100 (2004).
336. Haaland, R. E. *et al.* Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtype A and C HIV-1. *PLOS Pathog.* **5,** e1000274 (2009).
337. Li, H. & Roossinck, M. J. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. *J. Virol.* **78,** 10582–10587 (2004).

338. Quer, J. *et al.* Effect of Bottlenecking on Evolution of the Nonstructural Protein 3 Gene of Hepatitis C Virus during Sexually Transmitted Acute Resolving Infection. *J. Virol.* **79,** 15131–15141 (2005).
339. Scholle, F., Girard, Y. A., Zhao, Q., Higgs, S. & Mason, P. W. trans-Packaged West Nile Virus-Like Particles: Infectious Properties In Vitro and in Infected Mosquito Vectors. *J. Virol.* **78,** 11605–11614 (2004).
340. Smith, D. R., Adams, A. P., Kenney, J. L., Wang, E. & Weaver, S. C. Venezuelan equine encephalitis virus in the mosquito vector Aedes taeniorhynchus: Infection initiated by a small number of susceptible epithelial cells and a population bottleneck. *Virology* **372,** 176–186 (2008).
341. Duarte, E. A. *et al.* Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. *J. Virol.* **67,** 3620–3623 (1993).
342. Novella, I. S., Quer, J., Domingo, E. & Holland, J. J. Exponential Fitness Gains of RNA Virus Populations Are Limited by Bottleneck Effects. *J. Virol.* **73,** 1668–1671 (1999).
343. Pariente, N., Sierra, S., Lowenstein, P. R. & Domingo, E. Efficient Virus Extinction by Combinations of a Mutagen and Antiviral Inhibitors. *J. Virol.* **75,** 9723–9730 (2001).
344. Whittaker, D. Evolution 101: Fitness Landscapes | BEACON.
345. Holland, J. & Domingo, E. Origin and evolution of viruses. *Virus Genes* **16,** 13–21 (1998).
346. Holland, J. *et al.* Rapid evolution of RNA genomes. *Science* **215,** 1577–1585 (1982).
347. Gorman, O. T., Bean, W. J. & Webster, R. G. Evolutionary Processes in Influenza Viruses: Divergence, Rapid Evolution, and Stasis. in *Genetic Diversity of RNA Viruses* 75–97 (Springer, 1992). doi:10.1007/978-3-642-77011-1_6
348. Baranowski, E., Ruiz-Jarabo, C. M., Pariente, N., Verdaguer, N. & Domingo, E. Evolution of Cell Recognition by Viruses: A Source of Biological Novelty with Medical Implications. in *Advances in Virus Research* **62,** 19–111 (Academic Press, 2003).
349. Bergelson, J. M. Receptors. 73–86 (2010). doi:10.1128/9781555816698.ch5
350. Crowell, R. L. *Virus attachment and entry into cells.* (ASM, 1986).
351. Marsh, M. & Helenius, A. Virus Entry: Open Sesame. *Cell* **124,** 729–740 (2006).
352. Schneider-Schaulies, J. Cellular receptors for viruses: links to tropism and pathogenesis. *J. Gen. Virol.* **81,** 1413–1429 (2000).
353. Wimmer, E. *Cellular receptors for animal viruses.* (Cold Spring Harbor Laboratory, 1994).
354. W.r, K. An enteric disease of dogs resembling feline panleucopaenia. Letter to the Editor. *Aust. Vet. J.* (1978).
355. Parrish, C. R. Emergence, natural history, and variation of canine, mink, and feline parvoviruses. *Adv. Virus Res.* **38,** 403–450 (1990).
356. Appel, M. J., Scott, F. W. & Carmichael, L. E. Isolation and immunisation studies of a canine parco-like virus from dogs with haemorrhagic enteritis. *Vet. Rec.* **105,** 156–159 (1979).
357. Cornell Univeristy. Animal health article: canine parvovirus. (2014).

358. Baranowski, E., Ruiz-Jarabo, C. M. & Domingo, E. Evolution of Cell Recognition by Viruses. *Science* **292,** 1102–1105 (2001).
359. Baranowski, E. *et al.* Cell Recognition by Foot-and-Mouth Disease Virus That Lacks the RGD Integrin-Binding Motif: Flexibility in Aphthovirus Receptor Usage. *J. Virol.* **74,** 1641–1647 (2000).
360. Sa-Carvalho, D. *et al.* Tissue culture adaptation of foot-and-mouth disease virus selects viruses that bind to heparin and are attenuated in cattle. *J. Virol.* **71,** 5115–5123 (1997).
361. Wright, C. F. *et al.* Beyond the Consensus: Dissecting Within-Host Viral Population Diversity of Foot-and-Mouth Disease Virus by Using Next-Generation Genome Sequencing. *J. Virol.* **85,** 2266–2275 (2011).
362. Takada, A. & Kawaoka, Y. Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. *Rev. Med. Virol.* **13,** 387–398 (2003).
363. Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 18590–18595 (2005).
364. Gabriel, G. *et al.* Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. *Nat. Commun.* **2,** ncomms1158 (2011).
365. Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular Basis for High Virulence of Hong Kong H5N1 Influenza A Viruses. *Science* **293,** 1840–1842 (2001).
366. Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* **67,** 1761–1764 (1993).
367. Yamada, S. *et al.* Biological and Structural Characterization of a Host-Adapting Amino Acid in Influenza Virus. *PLoS Pathog.* **6,** (2010).
368. Domingo, E. RNA virus evolution and the control of viral disease. in *Progress in Drug Research* 93–133 (Birkhäuser Basel, 1989). doi:10.1007/978-3-0348-9146-2_5
369. Domingo, E. & Holland, J. J. Complications of RNA heterogeneity for the engineering of virus vaccines and antiviral agents. *Genet. Eng. (N. Y.)* **14,** 13–31 (1992).
370. Drosopoulos, W. C. & Prasad, V. R. Increased Misincorporation Fidelity Observed for Nucleoside Analog Resistance Mutations M184V and E89G in Human Immunodeficiency Virus Type 1 Reverse Transcriptase Does Not Correlate with the Overall Error Rate Measured In Vitro. *J. Virol.* **72,** 4224–4230 (1998).
371. *Antiviral drug resistance*. (John Wiley and Sons Inc, 1996).
372. Herrmann, E. C. & Herrmann, J. A. A Working Hypothesis—Virus Resistance Development as an Indicator of Specific Antiviral Activity. *Ann. N. Y. Acad. Sci.* **284,** 632–637 (1977).
373. Pillay, D. & Zambon, M. Antiviral drug resistance. *BMJ* **317,** 660–662 (1998).
374. Spyrakis, F., BidonChanal, A., Barril, X. & Javier Luque, F. Protein Flexibility and Ligand Recognition: Challenges for Molecular Modeling. *Curr. Top. Med. Chem.* **11,** 192–210 (2011).
375. Tzeng, S.-R. & Kalodimos, C. G. Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* **21,** 62–67 (2011).
376. Agudo, R. *et al.* A Multi-Step Process of Viral Adaptation to a Mutagenic Nucleoside Analogue by Modulation of Transition Types Leads to Extinction-Escape. *PLOS Pathog.* **6,** e1001072 (2010).

377. Heinz, B. A. *et al.* Genetic and molecular analyses of spontaneous mutants of human rhinovirus 14 that are resistant to an antiviral compound. *J. Virol.* **63,** 2476–2485 (1989).
378. Heinz, B. A. & Vance, L. M. The antiviral compound enviroxime targets the 3A coding region of rhinovirus and poliovirus. *J. Virol.* **69,** 4189–4197 (1995).
379. Verbinnen, T. *et al.* Tracking the Evolution of Multiple In Vitro Hepatitis C Virus Replicon Variants under Protease Inhibitor Selection Pressure by 454 Deep Sequencing. *J. Virol.* **84,** 11124–11133 (2010).
380. Chong, Y. L., Padhi, A., Hudson, P. J. & Poss, M. The Effect of Vaccination on the Evolution and Population Dynamics of Avian Paramyxovirus-1. *PLOS Pathog.* **6,** e1000872 (2010).
381. Feng, Z. & Lemon, S. M. Hepatitis A virus. in *The picornaviruses* (eds. Ehrenfeld, E., Domingo, E. & Roos, R. P.) 383–396 (ASM, 2010).
382. Knowles, N. J., Hovi, T., King, A. M. Q. & Stanway, G. Overview of taxonomy. in (eds. Ehrenfeld, E., Domingo, E. & Roos, R. P.) 19–32 (ASM, 2010).
383. Cristina, J. & Costa-Mattioli, M. Genetic variability and molecular evolution of Hepatitis A virus. *Virus Res.* **127,** 151–157 (2007).
384. Sánchez, G., Bosch, A., Gómez-Mariano, G., Domingo, E. & Pintó, R. M. Evidence for quasispecies distributions in the human hepatitis A virus genome. *Virology* **315,** 34–42 (2003).
385. Aragonès, L., Bosch, A. & Pintó, R. M. Hepatitis A Virus Mutant Spectra under the Selective Pressure of Monoclonal Antibodies: Codon Usage Constraints Limit Capsid Variability. *J. Virol.* **82,** 1688–1700 (2008).
386. Pérez-Sautu, U. *et al.* Hepatitis A Virus Vaccine Escape Variants and Potential New Serotype Emergence. *Emerg. Infect. Dis.* **17,** 734–737 (2011).
387. Domingo, E. Mechanisms of viral emergence. *Vet. Res.* **41,** 38 (2010).
388. Angus, A. G. & Patel, A. H. Immunotherapeutic potential of neutralizing antibodies targeting conserved regions of the HCV envelope glycoprotein E2. *Future Microbiol.* **6,** 279–294 (2011).
389. Novella, I. S. *et al.* Use of substituted and tandem-repeated peptides to probe the relevance of the highly conserved RGD tripeptide in the immune response against foot-and-mouth disease virus. *FEBS Lett.* **330,** 253–259 (1993).
390. Torre, J. C. de la & Holland, J. J. RNA virus quasispecies populations can suppress vastly superior mutant progeny. *J. Virol.* **64,** 6278–6281 (1990).
391. Grande-Pérez, A., Gómez-Mariano, G., Lowenstein, P. R. & Domingo, E. Mutagenesis-Induced, Large Fitness Variations with an Invariant Arenavirus Consensus Genomic Nucleotide Sequence. *J. Virol.* **79,** 10451–10459 (2005).
392. Iranzo, J. & Manrubia, S. C. Stochastic extinction of viral infectivity through the action of defectors. *EPL Europhys. Lett.* **85,** 18001 (2009).
393. González-López, C., Gómez-Mariano, G., Escarmís, C. & Domingo, E. Invariant aphthovirus consensus nucleotide sequence in the transition to error catastrophe. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **5,** 366–374 (2005).
394. Borrego, B., Novella, I. S., Giralt, E., Andreu, D. & Domingo, E. Distinct repertoire of antigenic variants of foot-and-mouth disease virus in the presence or absence of immune selection. *J. Virol.* **67,** 6071–6079 (1993).

395. Zhang, Y. *et al.* Emergence and transmission pathways of rapidly evolving evolutionary branch C4a strains of human enterovirus 71 in the Central Plain of China. *PloS One* **6,** e27895 (2011).
396. Bull, R. A., Eden, J.-S., Rawlinson, W. D. & White, P. A. Rapid Evolution of Pandemic Noroviruses of the GII.4 Lineage. *PLOS Pathog.* **6,** e1000831 (2010).
397. OhAinle, M. *et al.* Dynamics of Dengue Disease Severity Determined by the Interplay Between Viral Genetics and Serotype-Specific Immunity. *Sci. Transl. Med.* **3,** 114ra128-114ra128 (2011).
398. Sharp, P. M. & Simmonds, P. Evaluating the evidence for virus/host co-evolution. *Curr. Opin. Virol.* **1,** 436–441 (2011).
399. Simmonds, P. & Domingo, E. Virus evolution. *Curr. Opin. Virol.* **1,** 410–412 (2011).
400. Coffin, J. M. Genetic Variation in Retroviruses. in *Virus Variability, Epidemiology and Control* 11–33 (Springer, Boston, MA, 1990). doi:10.1007/978-1-4757-9271-3_2
401. Villaverde, A. *et al.* Fixation of mutations at the VP1 gene of foot-and-mouth disease virus. Can quasispecies define a transient molecular clock? *Gene* **103,** 147–153 (1991).
402. Domingo, E. Virus evolution. in *Fields virology* (eds. Knipe, D. M. & Howley, P. M.) **5,** (Lippincott Williams & Wilkins, 2007).
403. Sobrino, F. *et al.* Fixation of mutations in the viral genome during an outbreak of foot-and-mouth disease: heterogeneity and rate variations. *Gene* **50,** 149–159 (1986).
404. Paweska, J. T. Effect of the South African asinine-94 strain of equine arteritis virus (EAV) in pregnant donkey mares and duration of maternal immunity in foals. *Onderstepoort J. Vet. Res.* **64,** 147–152 (1997).
405. Paweska, J. T. Equine viral arteritis in donkeys in South Africa. *J. S. Afr. Vet. Assoc.* **65,** 40 (1994).
406. Paweska, J. T. & Barnard, B. J. Serological evidence of equine arteritis virus in donkeys in South Africa. *Onderstepoort J. Vet. Res.* **60,** 155–158 (1993).
407. Balasuriya, U. B. R. *et al.* Host Factors that Contribute to Equine Arteritis Virus Persistence in the Stallion: an Update. *J. Equine Vet. Sci.* **43,** S11–S17 (2016).
408. Zhang, J. *et al.* Genetic variation and phylogenetic analysis of 22 French isolates of equine arteritis virus. *Arch. Virol.* **152,** 1977–1994 (2007).
409. Snijder, E. J., Kikkert, M. & Fang, Y. Arterivirus molecular biology and pathogenesis. *J. Gen. Virol.* **94,** 2141–2163 (2013).
410. Balasuriya, U. B. R., Timoney, P. J., McCollum, W. H. & MacLachlan, N. J. Phylogenetic Analysis of Open Reading Frame 5 of Field Isolates of Equine Arteritis Virus and Identification of Conserved and Nonconserved Regions in the GLEnvelope Glycoprotein. *Virology* **214,** 690–697 (1995).
411. Mittelholzer, C. *et al.* Extended Phylogeny of Equine Arteritis Virus: Division into New Subgroups. *J. Vet. Med. Ser. B* **53,** 55–58 (2006).
412. Stadejek, T. *et al.* Genetic diversity of equine arteritis virus. *J. Gen. Virol.* **80,** 691–699 (1999).
413. Sobel Leonard, A. *et al.* Deep Sequencing of Influenza A Virus from a Human Challenge Study Reveals a Selective Bottleneck and Only Limited Intrahost Genetic Diversification. *J. Virol.* **90,** 11247–11258 (2016).

414. Murcia, P. R. *et al.* Intra- and interhost evolutionary dynamics of equine influenza virus. *J. Virol.* **84,** 6943–6954 (2010).
415. Grad, Y. H. *et al.* Within-host whole-genome deep sequencing and diversity analysis of human respiratory syncytial virus infection reveals dynamics of genomic diversity in the absence and presence of immune pressure. *J. Virol.* **88,** 7286–7293 (2014).
416. Palmer, B. A. *et al.* Analysis of the evolution and structure of a complex intrahost viral population in chronic hepatitis C virus mapped by ultradeep pyrosequencing. *J. Virol.* **88,** 13709–13721 (2014).
417. Martinez, M. A., Nevot, M., Jordan-Paiz, A. & Franco, S. Similarities between Human Immunodeficiency Virus Type 1 and Hepatitis C Virus Genetic and Phenotypic Protease Quasispecies Diversity. *J. Virol.* **89,** 9758–9764 (2015).
418. Murphy, T. W. *et al.* Genomic variability among globally distributed isolates of equine arteritis virus. *Vet. Microbiol.* **32,** 101–115 (1992).
419. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet. TIG* **24,** 133–141 (2008).
420. McCollum, W. H. & Timoney, P. J. *The pathogenic qualities of the 1984 strain of equine arteritis virus*. 34–47 (Grayson Foundation, Inc, 1985).
421. OIE. *Equine viral arteritis*. (2016).
422. Balasuriya, U. B., Snijder, E. J. & MacLachlan, N. J. Phenotypic characterization of equine arteritis virus with an infectious cDNA clone. *Unpublished data.* (2000).
423. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59,** 307–321 (2010).
424. Ronquist, F. *et al.* MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **61,** 539–542 (2012).
425. Murrell, B. *et al.* FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30,** 1196–1205 (2013).
426. Brayne, A. B., Dearlove, B. L., Lester, J. S., Kosakovsky Pond, S. L. & Frost, S. D. W. Genotype-Specific Evolution of Hepatitis E Virus. *J. Virol.* **91,** (2017).
427. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8,** e1002764 (2012).
428. Moya, A., Holmes, E. C. & González-Candelas, F. The population genetics and evolutionary epidemiology of RNA viruses. *Nat. Rev. Microbiol.* **2,** 279–288 (2004).
429. Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E. & Desai, M. M. The dynamics of molecular evolution over 60,000 generations. *Nature* **551,** 45–50 (2017).
430. Lauring, A. S. & Andino, R. Quasispecies Theory and the Behavior of RNA Viruses. *PLOS Pathog.* **6,** e1001005 (2010).
431. Vrancken, B., Suchard, M. A. & Lemey, P. Accurate quantification of within- and between-host HBV evolutionary rates requires explicit transmission chain modelling. *Virus Evol.* **3,** vex028 (2017).
432. Parameswaran, P. *et al.* Intrahost Selection Pressures Drive Rapid Dengue Virus Microevolution in Acute Human Infections. *Cell Host Microbe* **22,** 400–410.e5 (2017).
433. Murcia, P. R. *et al.* Evolution of Equine Influenza Virus in Vaccinated Horses. *J. Virol.* **87,** 4768–4771 (2013).
434. Moncla, L. H., Florek, N. W. & Friedrich, T. C. Influenza Evolution: New Insights into an Old Foe. *Trends Microbiol.* **25,** 432–434 (2017).

435. Drummond, A. J., Pybus, O. G., Rambaut, A., Forsberg, R. & Rodrigo, A. G. Measurably evolving populations. *Trends Ecol. Evol.* **18,** 481–488 (2003).
436. Raghwani, J. *et al.* Exceptional Heterogeneity in Viral Evolutionary Dynamics Characterises Chronic Hepatitis C Virus Infection. *PLoS Pathog.* **12,** e1005894 (2016).
437. Cornman, R. S. *et al.* Population-genomic variation within RNA viruses of the Western honey bee, Apis mellifera, inferred from deep sequencing. *BMC Genomics* **14,** 154 (2013).
438. Anderson, J. A. *et al.* HIV-1 Populations in Semen Arise through Multiple Mechanisms. *PLOS Pathog.* **6,** e1001053 (2010).
439. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439,** 344–348 (2006).
440. Varble, A. *et al.* Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell Host Microbe* **16,** 691–700 (2014).
441. Carossino, M. *et al.* Equine Arteritis Virus Elicits a Mucosal Antibody Response in the Reproductive Tract of Persistently Infected Stallions. *Clin. Vaccine Immunol. CVI* **24,** (2017).
442. Plagemann, P. W. Lactate dehydrogenase-elevating virus and related viruses. in *Fields Virology* (eds. Fields, B. N., Knipe, D. M. & Howley, P. M.) 1105–1120 (1996).
443. Yamaguchi, S., Kanno, T., Akashi, H. & Kondo, T. Identification of Two Neutralization Antigenic Sites in G_L Protein of Equine Arteritis Virus by Means of Monoclonal Antibodies. *J. Equine Sci.* **8,** 7–11 (1997).
444. Hornyák, Á., Bakonyi, T., Tekes, G., Szeredi, L. & Rusvai, M. A Novel Subgroup among Genotypes of Equine Arteritis Virus: Genetic Comparison of 40 Strains. *J. Vet. Med. Ser. B* **52,** 112–118 (2005).
445. Chen, Q. *et al.* Isolation and Characterization of Porcine Epidemic Diarrhea Viruses Associated with the 2013 Disease Outbreak among Swine in the United States. *J. Clin. Microbiol.* **52,** 234–243 (2014).
446. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10,** 512–526 (1993).
447. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33,** 1870–1874 (2016).
448. Verheije, M. H., Olsthoorn, R. C. L., Kroese, M. V., Rottier, P. J. M. & Meulenberg, J. J. M. Kissing Interaction between 3′ Noncoding and Coding Sequences Is Essential for Porcine Arterivirus RNA Replication. *J. Virol.* **76,** 1521–1526 (2002).
449. Yin, Y. *et al.* Conserved nucleotides in the terminus of the 3′ UTR region are important for the replication and infectivity of porcine reproductive and respiratory syndrome virus. *Arch. Virol.* **158,** 1719–1732 (2013).
450. Wang, L. & Zhang, Y. Novel porcine reproductive and respiratory syndrome virus strains in the United States with deletions in untranslated regions. *Arch. Virol.* **160,** 3093–3096 (2015).
451. Sun, Z. *et al.* Identification of dispensable nucleotide sequence in 3' untranslated region of porcine reproductive and respiratory syndrome virus. *Virus Res.* **154,** 38–47 (2010).

452. Zhou, Z. *et al.* The epidemic status and genetic diversity of 14 highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV) isolates from China in 2009. *Vet. Microbiol.* **150,** 257–269 (2011).
453. Lu, W. *et al.* Attenuation and Immunogenicity of a Live High Pathogenic PRRSV Vaccine Candidate with a 32-Amino Acid Deletion in the nsp2 Protein. *Journal of Immunology Research* (2014). doi:10.1155/2014/810523
454. Moreira, R., García, A., Valencia, J. & Moreno, V. Equine Viral Arteritis in feral donkeys (Equus asinus) of the Atacama Region, Chile. *J. Equine Vet. Sci.* **39,** S63–S64 (2016).
455. Berrios, P. Actualización sobre enfermedades virales de los equinos. *Mon Electr Patol Vet* **2,** 34–59 (2005).
456. Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162,** 156–159 (1987).
457. Heinz, F. X. *et al.* Family Flaviridae. in (eds. Fauquet, C. M., Mayo, M., Maniloff, J., Desselberger, U. & Ball, L. A.) 981–998 (2004).
458. Peterhans, E., Bachofen, C., Stalder, H. & Schweizer, M. Cytopathic bovine viral diarrhea viruses (BVDV): emerging pestiviruses doomed to extinction. *Vet. Res.* **41,** 44
459. Deng, Y. *et al.* Genomic characterization of a bovine viral diarrhea virus 1 isolate from swine. *Arch. Virol.* **159,** 2513–2517 (2014).
460. Peletto, S. *et al.* Detection and phylogenetic analysis of an atypical pestivirus, strain IZSPLV_To. *Res. Vet. Sci.* **92,** 147–150 (2012).
461. Giangaspero, M., Harasawa, R., Weber, L. & Belloli, A. Genoepidemiological Evaluation of *Bovine viral diarrhea virus 2* Species Based on Secondary Structures in the 5' Untranslated Region. *J. Vet. Med. Sci.* **70,** 571–580 (2008).
462. Nuttall, P. A., Luther, P. D. & Stott, E. J. Viral contamination of bovine foetal serum and cell cultures. *Nature* **266,** 266835a0 (1977).
463. Bolin, S. R. *et al.* Detection of a cell line contaminated with hog cholera virus. *J. Am. Vet. Med. Assoc.* **205,** 742–745 (1994).
464. Bolin, S. R., Matthews, P. J. & Ridpath, J. F. Methods for Detection and Frequency of Contamination of Fetal Calf Serum with Bovine Viral Diarrhea Virus and Antibodies against Bovine Viral Diarrhea Virus. *J. Vet. Diagn. Invest.* **3,** 199–203 (1991).
465. Levings, R. L. & Wessman, S. J. Bovine viral diarrhea virus contamination of nutrient serum, cell cultures and viral vaccines. *Dev. Biol. Stand.* **75,** 177–181 (1991).
466. Dezengrini, R., Weiblen, R. & Flores, E. F. Selection and characterization of canine, swine and rabbit cell lines resistant to bovine viral diarrhea virus. *J. Virol. Methods* **137,** 51–57 (2006).
467. Timoney, P. J., Bruser, C. A., McCollum, W. H., Holyoak, G. R. & Little, T. V. *International Workshop on the Diagnosis of Equine Arteritis Virus Infection,*. 26–27 (M. H. Gluck Equine Research Center, 2004).
468. Tao, J., Wang, Y., Wang, J., Wang, J. & Zhu, G. Identification and genetic characterization of new bovine viral diarrhea virus genotype 2 strains in pigs isolated in China. *Virus Genes* **46,** 81–87 (2013).

469. Muhsen, M., Aoki, H., Ikeda, H. & Fukusho, A. Biological properties of bovine viral diarrhea virus quasispecies detected in the RK13 cell line. *Arch. Virol.* **158,** 753–763 (2013).
470. 10 Facts About The Atacama Desert. *South America Travel Blog by Quasar Expeditions* (2016). Available at: https://www.quasarex.com/blog/10-facts-atacama-desert. (Accessed: 4th December 2017)

VITA

Bora Nam

Academic Degree

B.S. University of Kentucky Lexington, KY
Department of Animal and Food Sciences
College of Agriculture, Food and Environment
Major: Equine science and management
Minor: Agricultural economics
Graduation date: December 20, 2013

Professional Positions Held (Research Experience)

Maxwell H. Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky

Jan 15, 2015 - Current	Graduate Research Assistant (Masters degree)
Jan 15, 2014 - Jan 14, 2015	Senior Laboratory Technician/Optional Practical Training (OPT)
Jan 22, 2010 - Dec 20, 2013	Undergraduate Student Lab Assistant

Peer-Reviewed Publications

1. Carossino, M., Wagner, B., Loynachan, A. T., Cook, R. F., Canisso, I. F., Lakshman, C., Edwards, C. L., **Nam, B.**, Timoney, J. F., Timoney, P. J., Balasuriya, U. B. R. 2017. Equine arteritis virus elicits a mucosal antibody response in the reproductive tract of the persistently infected stallion. Submitted to Clinical and Vaccine Immunology (*CVI*).
2. Carossino, M., Loynachan, A. T., Canisso, I. F., Cook, R. F., Campos, J. R., **Nam, B.**, Go, Y. Y., Squires, E. L., Troedsson, M. H. T., Balasuriya, U. B. R. 2017. Equine Arteritis Virus has Specific Tropism for Stromal Cells and $CD8^+$ T and $CD21^+$ B Lymphocytes but not Glandular Epithelium at the Primary Site of Persistent Infection in the Stallion Reproductive Tract. *J. Virol*. 91(13):e00418-17.
3. Carossino, M., Lee, P. A., **Nam, B.**, Skillman, A., Shuck, K. M., Timoney, P. J., Tsai, Y., Ma, L., Chang, H. G., Wang, H. T., Balasuriya, U. B. R. 2016. Development and Evaluation of a Reverse Transcription-Insulated Isothermal Polymerase Chain Reaction (RT-iiPCR) Assay for Detection of Equine Arteritis Virus in Equine Semen and Tissue Samples using the POCKIT™ System. *J. Virol. Methods*. 234: 7-15.
4. **Nam, B**., Li, G., Zheng, Y., Zhang, J. Shuck, K. M., Timoney, P. J., Balasuriya, U. B. R. 2015. Complete Genome Sequence of Noncytopathic Bovine Viral Diarrhea Virus 1 (BVDV-1) Contaminating High Passage RK-13 Cell Line. *ASM Genome Announc*. 3(5): e01115-15.
5. Balasuriya, U. B. R., Lee, P.A., Tiwari, A., Skillman, A., **Nam, B.,** Chambers, T. M., Tsai, Y., Ma, L., Yang, P., Chung, S., Chang, H. G., Wang, H. T. 2014. Rapid Detection of Equine Influenza Virus H3N8 Subtype by Insulated

Isothermal RT-PCR (iiRT-PCR) Assay using the POCKIT™ Nucleic Acid Analyzer. *J. Virol. Methods*. 207: 66-72.

Publications in Preparation

1. **Nam, B.**, Mekuria, Z., Carossino, M., Li, G., Zheng, Y., Zhang, J., Cook, F., Campos, J. R., Shuck, K. M., Squires, E. L., Troedsson, M. H. T., Bailey, E., Timoney, P. J. and Balasuriya, U. B. R. 2017. Genomic bottleneck and selection in equine arteritis virus acute infection and intra-host quasispecies diversification in long-term persistent infection of the stallion reproductive tract.
2. **Nam, B.**, Mekuria, Z., Shuck, K. M., Moreno, V., Benson, C.M., Garcia, A., Abumada, C., Zuniga, R.M., Li, G., Zheng, Y., Zhang, J., Carossino, M., Timoney, P. J., Balasuriya, U. B. R. 2017. Genomic, phylogenetic, and antigenic characterization of a novel field strain of equine arteritis virus isolated from a wild donkey in Chile.

Scientific Abstracts and Presentations

1. #**Nam, B.** and #Mekuria, Z. Genetic and phenotypic modulation of equine arteritis virus during long-term persistence in the stallion. *5th USDA-NIFA Symposium "Outcomes of the USDA-NIFA Grant on Identification of Genetic Factors Responsible for the Establishment of EAV Carrier State in Stallions"* November 10th, 2017, Marriot Griffin Gate, Lexington, Kentucky, USA.
2. #**Nam, B.**, Mekuria, Z., Carossino, M., Li, G., Zheng, Y., Zhang, J., Cook, R. F., and Balasuriya, U. B. R. Evolution of equine arteritis virus during persistent infection in stallions. *American Society for Virology*, 36nd annual meeting June 23th-28th, 2017, University of Wisconsin-Madison at the Monona Terrace Convention Center, Madison, Wisconsin, USA.
3. #**Nam, B.**, Mekuria, Z., Carossino, M., Li, G., Zheng, Y., Zhang, J., Cook, R. F., and Balasuriya, U. B. R. Evolution of equine arteritis virus during persistent infection in experimentally infected stallions. *XIVth International Nidovirus Symposium*, June 4st-8th, 2017, Kansas City, Missouri, USA.
4. #Carossino, M., Loynachan, A. T., Campos, J. R., **Nam, B.**, Canisso, I. F., Go, Y. Y., Timoney, P. J., Shuck, K. M., Henney, P., Troedsson, M. H., Cook, R. F., Swerczek, T., Squires, E. L., Bailey, E. and Balasuriya, U. B. R. Equine Arteritis Virus Tissue and Cellular Tropism During Persistent Infection in the Stallion Reproductive Tract. *XIVth International Nidovirus Symposium*, June 4st-8th, 2017, Kansas City, Missouri, USA.
5. #Carossino, M., Loynachan, A. T., Campos, J. R., **Nam, B.**, Canisso, I. F., Go, Y. Y., Timoney, P. J., Shuck, K. M., Henney, P., Troedsson, M. H., Cook, R. F., Swerczek, T., Squires, E. L., Bailey, E. and Balasuriya, U. B. R. Sites of equine arteritis virus persistence in the stallion's reproductive tract and characterization of the local inflammatory response to the virus. *Conference of Research Workers in Animal Diseases*, December 6th-8th, 2015, Chicago Marriott, Downtown Magnificent Mile, Illinois, USA.
6. #Carossino, M., Loynachan, A. T., Campos, J. R., **Nam, B.**, Canisso, I. F., Go, Y. Y., Timoney, P. J., Shuck, K. M., Troedsson, M. H., Cook, R. F., Swerczek, T., Squires, E. L., Bailey, E. and Balasuriya, U. B. R. Sites of equine arteritis virus

localization in the reproductive tract during with long-term persistence in the stallion. *58th American Association of Veterinary Laboratory Diagnosticians (AAVLD)* and *the 119th United States Animal Health Association (USAHA)* annual meeting, October 22th-28th, 2015, Rhode Island Convention Center, Providence, Rhode Island, USA.

7. #Carossino, M., Loynachan, A. T., Campos, J. R., **Nam, B.**, Canisso, I. F., Go, Y. Y., Timoney, P. J., Shuck, K. M., Troedsson, M. H., Cook, R. F., Swerczek, T., Squires, E. L., Bailey, E. and Balasuriya, U. B. R. Characterization of the local inflammatory response in the reproductive tract of the equine arteritis virus carrier stallion. *58th American Association of Veterinary Laboratory Diagnosticians (AAVLD)* and *the 119th United States Animal Health Association (USAHA)* annual meeting, October 22th-28th, 2015, Rhode Island Convention Center, Providence, Rhode Island, USA.
8. #Bailey, E., Go, Y. Y., Eberth, J. E., Shuck, K., Cook, F., **Nam, B.**, Kalbfleisch, T. S., Timoney, P., and Balasuriya, U. B. R. Immune response phenotypes for equine viral arteritis (EVA) associated with alleles of CXCL16. *Plant and Animal Genome XXIII meeting*, January 10th-14th, 2015, Town and Country Hotel and Convention Center, San Diego, CA, USA.
9. #Balasuriya, U. B. R., Skillman, A., Shuck, K., **Nam, B**., Timoney, P. J., Tsai, Y., Ma, L., Yang, P. C., Chang, H., Lee, P. Y. A., Chang, H. F. G., and Wang, H. T. T. Evaluation of a Newly Developed Insulated Isothermal RT-PCR Assay and a Real-Time RT-PCR Assay for the Detection of Equine Arteritis Virus Nucleic Acid in Equine Semen. 57**th** *American Association of Veterinary Laboratory Diagnosticians (AAVLD)* and *the 118th United States Animal Health Association (USAHA)* annual meeting, October 16th-22th, 2014, Crown Center Sheraton Hotel, Kansas City, Missouri, USA.
10. #Balasuriya, U. B. R., Skillman, A., Shuck, K. M., **Nam, B.**, Timoney, P. J., Tsai, Y. L., Ma, L. J., Yang, P. C., Chang, H. H., Lee, P. Y. A, Grace Chang, H. F. G., and #Wang, H. T. T. 2014. 2014. Rapid insulated isothermal RT-PCR assay for point of need detection of equine arteritis virus nucleic acid in equine semen. *3rd European Association of Veterinary Laboratory Diagnosticians Congress*, October 12th-15th, 2014, Pisa, Italy.
11. **Nam, B**., Mondal, S. P., Timoney, P. J. and #Balasuriya, U. B. R, Comparison of complete genome sequence of KY84 strain with other virulent and avirulent strains of equine arteritis virus. *XIIIth International Nidovirus Symposium*, June 1st-5th, 2014, Salamanca, Spain.
12. #**Nam, B.** and Balasuriya, U. B. R., Complete Genome Sequence Analysis of the KY84 Strain of Equine Arteritis Virus. *National Conference on Undergraduate Research*, April 3th-5th, 2014, University of Kentucky, Lexington, Kentucky, USA.
13. #Balasuriya, U. B. R., Tiwari, A., Skillman, A., **Nam, B.**, Yun-Long Tsai, Y. L., Ma' L. J., Yang, P. C., Alison Lee, P. Y. A., Chung, S., Chang, H. F. G, and Thomas Wang, H. T. T. Validation of a field-deployable POCKIT™ nucleic acid detection system for specific and sensitive point-of-need detection of equine influenza virus (H3N8). *American Association of Veterinary Laboratory Diagnosticians (AAVLD)* and *the United States Animal Health Association*

(USAHA) annual meeting, October 17th-23th, 2013, Town & Country Resort and Convention Center, San Diego, California, USA.

14. #Li, Y., #**Nam, B.**, Henney, P. J., Cook, R. F., Timoney, P. J., Tobin T. and Balasuriya, U. B. R. 2013. Comparison of *in vitro* antiviral activity of human herpesvirus DNA polymerase inhibitors against equine herpesvirus-1 and characterization of drug resistant mutant viruses. *American Society for Virology*, 32nd annual meeting July 20th-24th, 2013, Penn State University Park campus, State College, Pennsylvania, USA.

#Presenter

Nucleotide Sequence Deposition to Genbank Database

Total Number of Genbank Accessions: 55

Full-length Viral Genomic sequences: 55

1. GenBank accession # KT355592 – BVDV 1b HP-KY-RK13
2. GenBank accession # MF598091 – EAV Donkey VD7634
3. GenBank accession # MG137429 – EAV KY84
4. GenBank accession # MG137430 – EAV KY84 Stallion D84
5. GenBank accession # MG137431 – EAV KY84 Stallion D94
6. GenBank accession # MG137432 – EAV KY84 Stallion E84
7. GenBank accession # MG137433 – EAV KY84 Stallion E91
8. GenBank accession # MG137434 – EAV KY84-L136A-BC6
9. GenBank accession # MG137435 – EAV KY84-L136A-NC6
10. GenBank accession # MG137436 – EAV KY84-L136A-S5
11. GenBank accession # MG137437 – EAV KY84-L136A-S9
12. GenBank accession # MG137438 – EAV KY84-L136A-S107
13. GenBank accession # MG137439 – EAV KY84-L136A-S170
14. GenBank accession # MG137440 – EAV KY84-L136A-S380
15. GenBank accession # MG137441 – EAV KY84-L136A-S548
16. GenBank accession # MG137442 – EAV KY84-L136A-S726
17. GenBank accession # MG137443 – EAV KY84-L137B-BC6
18. GenBank accession # MG137444 – EAV KY84-L137B-NS6
19. GenBank accession # MG137445 – EAV KY84-L137B-S5
20. GenBank accession # MG137446 – EAV KY84-L137B-S9
21. GenBank accession # MG137447 – EAV KY84-L137B-S107
22. GenBank accession # MG137448 – EAV KY84-L138C-BC6
23. GenBank accession # MG137449 – EAV KY84-L138C-NS6
24. GenBank accession # MG137450 – EAV KY84-L138C-S5
25. GenBank accession # MG137451 – EAV KY84-L138C-S9
26. GenBank accession # MG137452 – EAV KY84-L138C-S107
27. GenBank accession # MG137453 – EAV KY84-L139D-BC6
28. GenBank accession # MG137454 – EAV KY84-L139D-NS6
29. GenBank accession # MG137455 – EAV KY84-L139D-S5
30. GenBank accession # MG137456 – EAV KY84-L139D-S9
31. GenBank accession # MG137457 – EAV KY84-L139D-S107
32. GenBank accession # MG137458 – EAV KY84-L139D-S170
33. GenBank accession # MG137459 – EAV KY84-L139D-S345

34. GenBank accession # MG137460 – EAV KY84-L140E-BC6
35. GenBank accession # MG137461 – EAV KY84-L140E-NS6
36. GenBank accession # MG137462 – EAV KY84-L140E-S5
37. GenBank accession # MG137463 – EAV KY84-L140E-S9
38. GenBank accession # MG137464 – EAV KY84-L140E-S107
39. GenBank accession # MG137465 – EAV KY84-L140E-S170
40. GenBank accession # MG137466 – EAV KY84-L140E-S380
41. GenBank accession # MG137467 – EAV KY84-L140E-S548
42. GenBank accession # MG137468 – EAV KY84-L140E-S726
43. GenBank accession # MG137469 – EAV KY84-L141F-BC6
44. GenBank accession # MG137470 – EAV KY84-L141F-NS6
45. GenBank accession # MG137471 – EAV KY84-L141F-S5
46. GenBank accession # MG137472 – EAV KY84-L141F-S9
47. GenBank accession # MG137473 – EAV KY84-L141F-S107
48. GenBank accession # MG137474 – EAV KY84-L141F-S170
49. GenBank accession # MG137475 – EAV KY84-L141F-S345
50. GenBank accession # MG137476 – EAV KY84-L142G-BC6
51. GenBank accession # MG137477 – EAV KY84-L142G-NS6
52. GenBank accession # MG137478 – EAV KY84-L142G-S5
53. GenBank accession # MG137479 – EAV KY84-L142G-S9
54. GenBank accession # MG137480 – EAV KY84-L142G-S107
55. GenBank accession # MG137481 – EAV KY84-L142G-S170

Viral Genes: 55

1. Bovine viral diarrhea virus – 1
2. Equine arteritis virus – 54

Honors and Fellowships

1. Dean's list for the fall semester of 2013, College of Agriculture, Food, and Environment, University of Kentucky, on September 20th, 2013.
2. Nominated for the Intern of the Year Award from 2013 Equine Science and Management Program, College of Agriculture, Food, and Environment, University of Kentucky, on September 3th, 2013.
3. Undergraduate Travel Grant from the Office of Undergraduate Research, University of Kentucky, on July 10th, 2013.