2018

ECOLOGY, EPIDEMIOLOGY, AND EVOLUTIONARY GENETICS OF CANINE DISTEMPER VIRUS SPILLOVER IN AFRICAN LIONS

Julie Kay Weckworth

ECOLOGY, EPIDEMIOLOGY, AND EVOLUTIONARY GENETICS OF CANINE DISTEMPER VIRUS SPILLOVER IN AFRICAN LIONS

By

Julie Kay Weckworth

Bachelor of Science, University of Cincinnati, Cincinnati, Ohio, 2000

Dissertation presented in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy
in Fish and Wildlife Biology

The University of Montana
Missoula, MT

May 2018

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

L. Scott Mills, Ph.D., Co-Chair
Wildlife Biology, Office of Research and Creative Scholarship

Michael K. Schwartz, Ph.D., Co-Chair
Wildlife Biology, USDA Forest Service

Mark Hebblewhite, Ph.D.
Wildlife Biology

Gordon Luikart, Ph.D.
Wildlife Biology, Flathead Lake Biological Station

Angela Luis, Ph.D.
Wildlife Biology

Weckworth, Julie, Ph.D., Fall 2016 Fish and Wildlife Biology

Ecology, epidemiology, and evolutionary genetics of clinical canine distemper virus emergence in African lions

Co-Chairperson: L. Scott Mills, Ph.D.

Co-Chairperson: Michael K. Schwartz, Ph.D.

The impact of emerging infectious diseases (EIDs) on the health and persistence of wildlife populations is an increasing conservation concern. Large carnivores are particularly vulnerable to EID impacts because they often occur in small, isolated populations with demographic and genetic challenges to long-term persistence. Ecological forces that isolate carnivore populations, e.g. agricultural intensification, simultaneously increase the probability of disease exposure from domestic species and can amplify population susceptibility to infection.

Canine distemper virus emerged as a conservation threat to African lions when an explosive epizootic caused the death or disappearance of a third of the Serengeti lion population in 1994. This same lion population was exposed to CDV on several other occasions without overt clinical infection. For my dissertation, I investigated ecological, epidemiological, and evolutionary factors contributing to the emergence and outcome of CDV infection in this globally important population.

Based on phylodynamic analyses of annotated sequence data I found that the lethal outbreak in 1994 was likely catalyzed by a single spillover event from a canid reservoir, and fueled by repeated transmissions from non-canid hosts, e.g. spotted hyenas. Distinct genotypes were found in canid and non-canid hosts suggesting that there is a host barrier to CDV spillover, which might limit lethal outbreaks in lions. Expanding the spatiotemporal scope of the phylogenetic analysis I found that Serengeti lions were not epidemiologically connected to other carnivore populations at the regional or continental scale. Recurrent CDV infection in Serengeti lions was likely due to local persistence in the domestic and/or wild carnivore community. Finally, based on phylogenetic and selection analyses I identified 25 candidate markers in the CDV genome potentially associated with the pathogenicity of infection in lions during the 1994 outbreak. These were mostly found in functional domains related to transcription and replication, and viral egress, implicating these processes as possible barriers to disease in lions. Mutations at two of the markers were shared with two CDV outbreaks in North America that caused clinical infection in African lions. Surveillance for these two mutations in circulating strains may inform CDV risk assessment in lion populations of conservation concern.

ACKNOWLEDGEMENTS

First, I would like to thank my two brilliant mentors, Scott Mills and Mike Schwartz. Scott gave me full academic license from day one, and coupled that with tireless encouragement. It was a recipe for exponential personal and professional growth. Mike, your engagement and support may have righted a listing ship. Thank you so much for inspiring me and keeping me focused in these final years!

Next I would like to thank all of my committee members past and present. My PhD path has taken many turns over the years, and through all of it I have been lucky to learn from the best in the field. To Gordon Luikart, our very first conversations laid the groundwork for everything in these pages (see Archie et al!). Thank you for the inspiration. To Mark Hebblewhite, thank you for your endurance and thoughtful guidance. Thank you, Angie Luis, for coming on late and assuming a pivotal role in the committee. Thanks also to Marcella Kelly, Vanessa Ezenwa, and Joel Berger who committed their energy and expertise in the earlier iterations of my PhD direction. Above all and to you all, thank you for your investment and time!

I am indebted to my collaborators who entrusted me with their hard-earned data. It is not lost on me the amount of effort it took to amass this resource. Melody Roelke-Parker, you are a genius now, as you were then. This whole dissertation is possible because of your work. Huge thanks go to Craig Packer for being my advocate and for providing a home away from home in Tanzania and Sarah Cleaveland for sharing your resources and energy. I am particularly grateful to the veterinarians and assistants who risked life and limb to catch dogs and sample dogs in the field. I still can't believe that you were able to get ocular and nasal swabs. Thank you Chunde Bigambo, Isra Silaa, Imam Mzimbiri, Paulo Tango, Deo, Ahmed, Jacobo and many, many others. Thank you to the Nelson Mandela African Institution of Science and Technology (NMAIST), the Tanzania Wildlife Research Institute (TAWIRI), Tanzania National Parks (TANAPA), Commission for Science and Technology (COSTECH), the Veterinary Investigation Center (VIC), and the Ngorongoro Conservation Area Authority (NCAA) that made this work possible.

Many people in the U.S. have also made this project possible. I am so grateful for the extraordinary efforts of Vicky Keahy, who lost many loved ones to canine distemper virus. I hope that this work honors their memories and that the results will help predict and prevent future outbreaks in big cats. Thank you to the Texas Department of State Health Services, Dallas Animal Services, and IDEXX.

I am especially grateful for the selfless support provided by Ed Dubovi and colleagues Melissa Laverack, Nancy Cavendar, and others at the Animal Health and Diagnostic Center during my internship at Cornell. Thank you for welcoming me into your lab, sharing your resources, and generally putting up with me.

I would like to thank friends and colleagues past and present from the Mills' (and associated) lab(s) and Wildlife Biology generally. Also, I would like to thank my friends and colleagues from the MEID program for helping me find my academic feet. I explored a lot of academic (and geographic) space, facilitated by this wonderful group, and in so doing, discovered skills and talents that I would not have otherwise. I am forever grateful for this experience and those of you who enriched it.

Thank you to my family, who has always been there, without judgement, in support of whatever ill-conceived plan I have. Especially to my mother, who paid for my undergrad degree and yet never wavered in her support when I was electively unemployed. My father, who was a conservationist of necessity and instilled in me his ethic and respect for the natural world. And to Aunt Mary, who by her example, challenged everyone she met to be their best person - be it in education or in empathy.

Last but not least, to my dear husband and son: Byron and Cassius. Byron quipped in his own PhD acknowledgements that I was in an unenviable position to be with him in the final six months of his PhD journey. He didn't know me well enough to realize that my paybacks are delivered in orders of magnitude. Byron, thank you so much for your tireless support. All the pep talks, the science talks, the shoulder to cry on. For all the nights you made dinner AND cleaned the kitchen (which goes against my core principles) and/or put Cassius in the bath and into bed. Thank you for all of that and so much more. And Cassius, thank you for the perspective that you gave to me. You light up my life.

My scholarship and this project were made possible through the generous support of the National Science Foundation, the Morris Animal Foundation, and the University of Montana.

DEDICATION

This dissertation is dedicated to my dad and Aunt Mary.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

OVERVIEW OF DISSERTATION

BACKGROUND

The frequency of infectious disease emergence is on the rise and presents a growing threat to human health, economy and the conservation of wildlife biodiversity (Cleaveland, Laurenson, & Taylor, 2001; Daszak, Cunningham, & Hyatt, 2000; K. E. Jones et al., 2008; Wiethoelter, Beltrán-Alcrudo, Kock, & Mor, 2015). The trend is largely driven by human-caused ecological changes, such as habitat fragmentation, encroachment, the growing trade in wild animals, and global climate change (Patz, Graczyk, Geller, & Vittor, 2000; Rogalski, Gowler, Shaw, Hufbauer, & Duffy, 2017). As the human population continues to increase, so does the threat of disease emergence at the human, domestic animal and wildlife interface. Understanding the causes and consequences of infectious disease emergence at this interface has surfaced as an urgent research priority to inform effective strategies for their surveillance and management.

Emerging infectious diseases (EIDs) are newly discovered pathogens or endemic pathogens experiencing a rapid expansion in incidence, host range or geographical extent (Morse 1995). The majority of human EIDs are zoonotic, caused by multi-host pathogens and transmitted largely by wildlife (B. A. Jones et al., 2013; K. E. Jones et al., 2008). Conspicuous examples include Ebola virus, human immunodeficiency virus (HIV), high pathogenicity avian influenza (H5N1), and severe acute respiratory sickness (SARS). Arguably the most

important source of EIDs is cross-species spillover, the potential impact of which cannot be overstated.

While much attention has been paid to emerging human zoonoses, multi-host EIDs affecting domestic animals and wildlife targets are cause for alarm (Alexander & Appel, 1994; Cleaveland et al., 2007). Multi-host pathogens cause 77% and 90% of livestock and domestic carnivore disease, respectively (Cleaveland 2001). In wild carnivore populations, Murray (1999) found that of 16 published disease outbreaks causing population declines in wild carnivores, all were caused by multi-host pathogens. Domestic animal populations, with high population densities and constant recruitment, can maintain virulent pathogens that would otherwise burn out in the absence of new susceptible hosts (Holt & Pickering 1982). Such pathogens may spillover to sensitive wildlife populations. Rinderpest virus (RPV), an extreme example, decimated wild ungulates impacting the function of entire ecosystems in sub-Saharan Africa following the importation of infected domestic cattle and continental spread in 1889 (Barrett, Pastoret, Taylor, Scott, & Provost, 2006).

Threatened and endangered species are at special risk from disease emergence. Many wild carnivore populations, including all seven big cat species, are on the decline owing to direct persecution for conflict with humans, poaching for consumption or international trade, and habitat loss. The synergy of these forces creates small, isolated populations with demographic and genetic challenges to long-term persistence. Simultaneously these processes increase the probability of disease exposure from domestic species and amplify population

susceptibility to infection and extinction. For example, sweeping land conversion and prairie dog eradication reduced black-footed ferret numbers to fewer than 18 known individuals in the wild (Thorne & Williams 1988). Concomitant loss of genetic diversity may have contributed to the vulnerability of this population to disease (O'Brien & Evermann 1988) and the last known wild individuals were wiped out by a canine distemper virus outbreak (Thorne & Williams 1988)[1]. Indeed, a novel infectious disease can be the nail in the coffin for a carnivore population already on the decline. Capacity to protect sensitive populations from disease-related declines in multi-host epidemiological systems is contingent upon our understanding of the factors driving pathogen spillover between species.

STUDY SYSTEM

Canine distemper virus (CDV) emergence in African lions in Serengeti National Park resulted in fatal neurological disease and greater than 30% population decline in a single outbreak in 1994 (Roelke-Parker et al., 1996). This alarming event catalyzed a long-term research initiative focused on infectious disease dynamics in the Serengeti carnivore community (Serengeti Carnivore Disease Project). My dissertation research benefits from data collected in this effort and from the long-term population health monitoring of the Serengeti lions (Serengeti Lion Project). Most in-situ wildlife CDV studies are limited to cross-sectional data representing single, independent serological surveys and/or disease outbreak investigations, which limits the temporal and spatial scale of inferences that can be drawn. CDV in Serengeti lions is especially revealing because this

[1] Subsequent captive breeding programs have rescued the black-footed ferret species from extinction.

population is not cryptic like most carnivores and has been studied intensively and continuously, spanning 5 putative CDV 'outbreaks' (Munson et al., 2008; Packer et al., 1999). Despite the challenges inherent in wildlife disease research and a retrospective approach, CDV in Serengeti lions is an unparalleled system in which questions central to the drivers of CDV emergence and spread can be advanced and their application to wildlife disease management can be explored.

RESEARCH OBJECTIVES

A holistic understanding of what drives CDV emergence and spread in complex epidemiological communities is critical for developing preventative and/or rapid response disease control strategies. For my dissertation, I took a multifactor approach, investigating host and viral factors contributing to the emergence, spread, persistence and pathogenicity of CDV infection in African lions. CDV evolves rapidly generating new sequence variation at a pace on par with the processes affecting them. This property enabled me to make ecological, epidemiological, and evolutionary inferences based on the distribution of CDV genetic variation in hosts, time, and space. Specifically, I used high-resolution whole genome sequence data to address the following objectives:

1. *Evaluate the role of host species in emergence and spread of a lethal CDV outbreak in African lions*

Food provisioned domestic species may occur at inflated densities capable of maintaining epidemic pathogens that may otherwise go extinct. Serological and early phylogenetic evidence implicated domestic dogs abutting Serengeti National Park as the origin of a catastrophic CDV outbreak in African lions in 1994

(Carpenter et al., 1998; Cleaveland et al., 2000; Roelke-Parker et al., 1996). However, more recently a phylogenetic analysis with higher resolution presented evidence that distinct CDV strains infected canids and non-canids during the outbreak, and suggested that the wild carnivore community was the source (Nikolin et al., 2017). Understanding the origin of CDV and pathways of transmission is central to designing effective surveillance and control in the population of conservation concern. I used phylodynamic tools with time-, space-, and species-referenced whole CDV genome sequences to reconstruct the origin and cross-species transmission dynamics of CDV in the Serengeti carnivore community during one lethal outbreak.

2. Assess the scale of CDV persistence with respect to Serengeti lions

CDV exhibits epidemic dynamics characterized by peaks in infection separated by population troughs during which levels of infection can be extremely low. As a result, the critical community size (CCS), or minimum host population necessary to maintain a chain of CDV transmission through the inter-epidemic period most of the time is expected to be very large (Bartlett, 1957). The apparent persistence of CDV in small, low density wild carnivore populations has led some to hypothesize that CDV is maintained at very large, regional scales and/or by a 'meta-reservoir' comprised of epidemiologically connected subpopulations of susceptible hosts (Almberg, Cross, & Smith, 2010; Prager et al., 2012; Viana et al., 2015). I used serological and genetic data to assess whether recurrent exposure of CDV in Serengeti lions resulted from local persistence or repeated introductions and characterized the scale of maintenance by assessing

epidemiological connectivity between Serengeti and other carnivores at the regional and continental scale.

3. Identify genetic markers of CDV pathogenicity in African lions

Serological evidence suggests CDV infection occurs in a wide diversity of taxa, transgressing even mammalian Orders, e.g. Primates, Artiodactyla, Carnivora, Rodentia, Cetacea, Proboscidea (Martinez-gutierrez & Ruiz-saenz, 2016). Clinical signs however have been reported in a fraction of apparently susceptible species. What predicts cross-species pathogenicity of CDV infection? I analyzed all available complete CDV genomes, including sequences from African lions clinically infected in three distinct outbreaks, to assess the role of viral genetic factors, e.g. recombination, selection, and neutral genetic forces, in pathogenicity in African lions and other alternative host species.

GENERAL RESULTS

Origin and cross-species transmission of CDV in a fatal outbreak in lions (Chapter 2)

In my phylodynamic analyses, I found evidence that suggests that a catastrophic outbreak of CDV in Serengeti lions was precipitated by a single cross-species spillover event from a canid reservoir to non-canid hosts, followed by multiple cross-species transmission events among non-canids, and possibly onward transmission from lion to lion. That all CDV sequences from non-canid hosts traced back to a common ancestor less than a year before the detection of clinical signs and explosive spread of CDV in lions was a surprising finding. Because of the range of species apparently susceptible to CDV infection, it is

generally thought that there are no or few barriers to cross-species transmission (Parrish et al., 2008). However, I found that distinct genotypes segregated canid and non-canid species in the dataset, which suggests that mutations occurred that may have increased fitness of the virus in the novel host. It is unknown whether the canid-associated strain could infect non-canids, and vice versa.

In contrast to the genetic distinction between sequences from canid and non-canid species, sequences from hosts within species groups, i.e. canids and non-canids, were highly similar. Thus, there did not appear to be barriers to cross-species transmission among more closely related host taxa. In fact, phylogenetic results and a discrete trait analysis supported the hypothesis that cross-species transmission from another non-canid species, possibly the spotted hyena, drove the extremely high CDV prevalence in African lions (85%), as opposed to an exclusive lion-to-lion or pride-to-pride chain of transmission.

Epidemiological history, persistence and scale in Serengeti (Chapter 3)

Serengeti lions have been exposed to CDV at least five times in the last 35 years (Munson et al., 2008; Viana et al., 2015). In this chapter competing, but not mutually exclusive hypotheses were tested to explain CDV re-emergence in Serengeti lions. I found evidence that recurrent CDV exposure in Serengeti lions over time resulted from local CDV persistence and from repeated CDV introductions. Serological data and PCR detections indicated that CDV was present in the Serengeti ecological region in all years of the study, except one, suggesting local persistence. Using genetic data, molecular dating suggested that all CDV strains sequenced from East Africa in this study (sampled in 1993-2011)

trace back to an introduction in 1989. An initial CDV incursion in Serengeti occurred in 1993 causing a brief but extensive outbreak in wild and domestic carnivores (see Chapter 2 for details). This lineage apparently faded out and was replaced by a second invasion of CDV by at least 1997 that was subsequently maintained until at least 2011. Beyond sharing a common origin in 1989, we did not find evidence that coincident CDV epizootics in Serengeti and another population (Laikipia, Kenya) in East Africa were synchronized by migration, i.e. these regional populations were not epidemiologically connected at the time of sampling. CDV lineages from the Horn of Africa and southern Africa were distantly related and not implicated in Serengeti CDV dynamics.

Genetic predictors of CDV clinical outcome in lions (Chapter 4)

My results from Chapter 3 suggest that the same lion population, exposed to slightly different CDV strains from the same genetic background, experienced gravely different clinical outcomes. Serengeti lions in 1993-1994 experienced high morbidity and mortality due to cytopathic effects of CDV infection (hereafter 'Fatal'), while no overt clinical signs of CDV infection or associated declines in individual or population health were observed during other periods of known CDV infection between 1997 and 2011 (hereafter 'Silent'). This provided a unique opportunity to test the hypothesis that viral genetic factors are associated with apparent variation in pathogenicity, and hence observed spillover, in African lion hosts. Using a series of evolutionary analyses, I found that 25 of 49 mutations differentiating the Fatal and Silent East African genotypes were the product of selection, supporting that viral adaptation was involved in the apparent increase in

pathogenicity in lions between these strains. Most mutations were found in functional domains of viral proteins responsible for efficient transcription and viral replication. Additionally, more mutations than expected were found on the Matrix protein, which orchestrates virus assembly and budding which controls onward spread. I investigated whether the adaptive mutations implicated in pathogenicity in lions in East Africa were involved in other CDV outbreaks affecting African lions in captive populations in North America. Was there a common genetic denominator that predicted clinical outcome in lions given infection? Of the 25 mutations, 7 were found in common with a CDV strain that caused clinical disease in African lions and tigers in North America in 2013, and 2 of these were also found in a strain that caused clinical disease in African lions, leopards, tigers, a mountain lion, and a jaguar in North America in 1992. That these mutations occurred almost exclusively in the three documented clinical CDV outbreaks in African lions suggests parallel evolution and is strong support for a functional role of these mutations in CDV pathogenesis in this species. The distribution of the identified adaptive mutations in sequences from other novel species and vaccines developed by passaging isolates in novel species, suggests that some of the adaptive mutations identified may be associated with generalist viral properties.

IMPROVING CDV SURVEILLANCE AND MANAGEMENT IN AFRICAN LIONS

The design of effective disease surveillance and control programs of multi-host pathogens at the human wildlife interface is often limited by an

incomplete understanding of cross-species transmission and reservoir dynamics, the latter of which requires longitudinal data on a population of interest to make valuable inference. CDV in Serengeti lions as a model system reveals important insights relevant to disease surveillance and management in African lions.

I found that spillover of CDV causing clinical disease in African lions in Serengeti was relatively rare, despite continual CDV presence in the carnivore community. During the only known occurrence of clinical CDV disease in Serengeti lions, genetic differences between strains in canids and non-canids suggested that a host barrier to CDV spillover exists and that the need to overcome this barrier may limit lethal outbreaks in lions. My data do not address whether mutations are *necessary and sufficient* to cause mass mortality in lions, though our results suggest that mutations do occur that increase CDV fitness in lions and lead to clinical outcomes. This improved understanding of the consequences of CDV spillover in African lions allows managers to weigh the risks of CDV exposure. Specifically, CDV management in the Serengeti has been focused on preventing lion exposure to CDV (Viana et al., 2015). However, data from Serengeti lions and other wild felid populations suggest that exposure to non-lethal CDV strains can actually increase herd-immunity (Munson et al., 2008; Viana et al., 2015). Thus, natural exposure to CDV may protect Serengeti lions from the rare occurrence of a lethal spillover. While the incidence of CDV in putative reservoir hosts may be a risk factor for CDV spillover in African lions, it does not predict pathogenicity, i.e. the potential clinical or population-level impacts in a new host. Thus surveillance for mutations that we found were

correlated with pathogenicity in lions, in circulating CDV strains would improve our understanding of risk in lion populations of conservation concern.

Reactive disease control strategies in host populations with well-studied social dynamics may selectively vaccinate precise social groups to arrest the spread of a pathogen through the population during an active outbreak. The spread of rabies virus in an Ethiopian wolf population in the Bale Mountains was curbed in this manner to protect a core fraction of the population from infection (Haydon et al., 2006). In my analyses spotted hyenas were implicated as a possible intermediate or liaison host, mediating most African lion CDV incidence during the fatal outbreak in 1994. Thus, if most of CDV exposure in lions occurs via other species, then targeted vaccination of prides with known network contacts would not be effective in protecting the population.

Finally, I detected CDV in Serengeti in almost every year surveyed suggesting that CDV persists in the Serengeti carnivore community, i.e. in the abutting domestic dog population and/or wild carnivores. Although CDV introduction from other populations is possible, during the study period Serengeti carnivores were not epidemiologically connected to other populations in East or southern Africa. This improved understanding of the source of recurrent exposure in lions suggests that surveillance focused on the local carnivore community may be sufficient to assess risk to African lions.

LITERATURE CITED

Alexander, K. A., & Appel, M. J. G. (1994). African wild dogs (Lycaon picturs) endangered by a canine distemper epizootic among domestic dogs near the

Masai Mara National Reserve, Kenya. *Journal of Wildlife Diseases*, *30*(4), 481–485. https://doi.org/10.7589/0090-3558-30.4.481

Almberg, E. S., Cross, P. C., & Smith, D. W. (2010). Persistence of canine distemper virus in the Greater Yellowstone Ecosystem's carnivore community. *Ecological Applications*, *20*(7), 2058–2074. https://doi.org/10.1890/09-1225.1

Barrett, T., Pastoret, P.-P., Taylor, W. P., Scott, G., & Provost, A. (Alain). (2006). *Rinderpest and peste des petits ruminants : virus plagues of large and small ruminants*. Academic.

Bartlett, M. S. (1957). Measles Periodicity and Community Size. *Source Journal of the Royal Statistical Society. Series A (General)*, *120*(1), 48–70. Retrieved from http://www.jstor.org/stable/2342553

Carpenter, M. A., Appel, M. J. G., Roelke-Parker, M. E., Munson, L., Hofer, H., East, M., & O 'brien, S. J. (1998). Genetic characterization of canine distemper virus in Serengeti carnivores. *Veterinary Immunology and Immunopathology*, *65*, 259–266.

Cleaveland, S., Appel, M. G. J., Chalmers, W. S. K., Chillingworth, C., Kaare, M., & Dye, C. (2000). Serological and demographic evidence for domestic dogs as a source of canine distemper virus infection for Serengeti wildlife. *Veterinary Microbiology*, *72*, 217–227.

Cleaveland, S., Mlengeya, T., Kaare, M., Haydon, D., Lembo, T., Laurenson, M. K., & Packer, C. (2007). The conservation relevance of epidemiological research into carnivore viral diseases in the serengeti. *Conservation Biology*,

21(3), 612–622. https://doi.org/10.1111/j.1523-1739.2007.00701.x

Cleaveland, Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1411), 991–9. https://doi.org/10.1098/rstb.2001.0889

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science (New York, N.Y.)*, *287*(5452), 443–9. https://doi.org/10.1126/SCIENCE.287.5452.443

Haydon, D. T., Randall, D. A., Matthews, L., Knobel, D. L., Tallents, L. A., Gravenor, M. B., … Laurenson, M. K. (2006). Low-coverage vaccination strategies for the conservation of endangered species. *Nature*, *443*(7112), 692–695. https://doi.org/10.1038/nature05177

Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., … Pfeiffer, D. U. (2013). Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8399–404. https://doi.org/10.1073/pnas.1208059110

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993. https://doi.org/10.1038/nature06536

Martinez-gutierrez, M., & Ruiz-saenz, J. (2016). Diversity of susceptible hosts in

canine distemper virus infection : a systematic review and data synthesis. *BMC Veterinary Research*, 1–11. https://doi.org/10.1186/s12917-016-0702-z

Munson, L., Terio, K. A., Kock, R., Mlengeya, T., Roelke, M. E., Dubovi, E., … Packer, C. (2008). Climate Extremes Promote Fatal Co-Infections during Canine Distemper Epidemics in African Lions. *PLoS ONE*, *3*(6), e2545. https://doi.org/10.1371/journal.pone.0002545

Nikolin, V. M., Olarte-Castillo, X. A., Osterrieder, N., Hofer, H., Dubovi, E., Mazzoni, C. J., … East, M. L. (2017). Canine distemper virus in the Serengeti ecosystem: molecular adaptation to different carnivore species. *Molecular Ecology*. https://doi.org/10.1111/mec.13902

Packer, C., Altizer, S., Appel, M., Brown, E., Martenson, J., O'Brien, S. J., … Lutz, H. (1999). Viruses of the Serengeti: patterns of infection and mortality in African lions. *Journal of Animal Ecology*, *68*(6), 1161–1178. https://doi.org/10.1046/j.1365-2656.1999.00360.x

Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E., Burke, D. S., Calisher, C. H., … Daszak, P. (2008). Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases. *Microbiology and Molecular Biology Reviews*, *72*(3), 457–470. https://doi.org/10.1128/MMBR.00004-08

Patz, J. A., Graczyk, T. K., Geller, N., & Vittor, A. Y. (2000). Effects of environmental change on emerging parasitic diseases. *International Journal for Parasitology*, *30*(12–13), 1395–1405. https://doi.org/10.1016/S0020-7519(00)00141-7

Prager, K. C., Mazet, J. A. K., Dubovi, E. J., Frank, L. G., Munson, L., Wagner, A. P., & Woodroffe, R. (2012). Rabies Virus and Canine Distemper Virus in Wild and Domestic Carnivores in Northern Kenya: Are Domestic Dogs the Reservoir? *EcoHealth*, *9*(4), 483–498. https://doi.org/10.1007/s10393-013-0815-9

Roelke-Parker, M. E., Munson, L., Packer, C., Kockil, R., CIeaveland, S., Carpenter, M., … G AppeIl, M. J. (1996). A canine distemper virus epidemic in Serengeti lions (Panthera leo). *Nature*, *379*, 441–445.

Rogalski, M. A., Gowler, C. D., Shaw, C. L., Hufbauer, R. A., & Duffy, M. A. (2017). Human drivers of ecological and evolutionary dynamics in emerging and disappearing infectious disease systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1712), 20160043. https://doi.org/10.1098/rstb.2016.0043

Viana, M., Cleaveland, S., Matthiopoulos, J., Halliday, J., Packer, C., Craft, M. E., … Orr, B. (2015). Dynamics of a morbillivirus at the domestic-wildlife interface: Canine Dis- temper Virus in domestic dogs and lions. *PNAS*, *112*(5), 1464–1469.

Wiethoelter, A. K., Beltrán-Alcrudo, D., Kock, R., & Mor, S. M. (2015). Global trends in infectious diseases at the wildlife-livestock interface. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(31), 9662–7. https://doi.org/10.1073/pnas.1422741112

CHAPTER 2

CANINE DISTEMPER VIRUS EMERGENCE: CROSS SPECIES TRANSMISSION AND EVOLUTIONARY DYNAMICS DURING A LETHAL OUTBREAK IN SERENGETI LIONS

ABSTRACT

Canine distemper virus (CDV) is an emerging, multi-host disease that threatens wild felid populations of conservation concern. The process of CDV emergence in novel populations and cross-species transmission during multi-host outbreaks is not well understood. In this study, we found evidence that a large-scale, lethal outbreak of Canine distemper virus in Serengeti lions was precipitated by a single cross-species spillover event from a canid reservoir, followed by multiple cross-species transmissions from the spotted hyena as well as onward transmission from lion to lion. Our data suggests that a CDV lineage associated with these non-canid hosts (found in lions and hyenas) diverged from a canid-associated lineage (found in domestic dogs and bat-eared fox) less than one year before the detection of clinical signs and explosive spread of CDV in lions. Seven amino acid residue differences in biologically relevant regions of the CDV genome separate these lineages, all of which had accumulated in less than 3.6 months of their divergence, and 6.6 months before the explosive onset of mortality in non-canids. The genetic differences between lineages support the hypothesis that CDV evolution is necessary for clinical infection in lions and hyenas and suggest that host barriers to clinical infection can limit lethal outbreaks of CDV in novel host

species. In contrast to the genetic distinction between non-canid and canid CDV lineages, our data show high similarity of CDV sequences regardless of host species within lineages, suggesting frequent cross-species transmission between more closely related hosts. Multiple lines of evidence support the hypothesis that hyenas acted as a source of CDV exposing lions at the pride-level and affirm that onward transmission between lions within and among prides also likely occurred. Understanding the timing and nature of barriers that constrain the outcome of cross-species transmission is critical in designing effective surveillance for and proper response to emerging CDV variants in vulnerable populations.

INTRODUCTION

Canine distemper virus (CDV) is a highly contagious, multi-host pathogen that can cause severe systemic disease and death in individuals (Greene & Appel, 1990), population-level decline (Roelke-Parker et al., 1996), and even local extirpation (Alexander & Appel, 1994). The frequency of spillover to novel host taxa is on the rise in recent years (Black bear (Cottrell, Keel, Brooks, Mead, & Phillips, 2013), rhesus monkey (Qiu, 2011)) including emergence in some critically endangered populations of wild felids (Iberian lynx (Meli et al., 2010), Amur leopard (Sulikhan et al., 2018); Amur tiger (Seimon et al., 2013). The process of CDV emergence in novel host populations is not well understood, owing to the apparently unpredictable nature of spillover and the low density and cryptic nature of many populations in which CDV has emerged.

To effectively survey for and mitigate CDV emergence in a target population of conservation concern it is critical to understand cross-species transmission dynamics: where and in what species CDV originates, how it spreads among susceptible hosts, and what potentially limits this spread. Here we distinguish between two outcomes of cross-species transmission, clinical and subclinical, as they relate to CDV emergence. Numerous species, including domestic cats, domestic pigs, and others, are susceptible to CDV infection of lymphatic cells that results in a transient immune suppression without any overt signs of clinical infection (Appel, Sheffy, Percy, & Gaskin, 1974). Experimentally infected animals with subclinical presentation did not transmit CDV but did acquire immunity (Appel et al., 1974). Clinical infection of an animal is

established following spread in the lymphatic system, when CDV enters and spreads in epithelial cells and/or glial cells of the central nervous system (Noyce, Delpeut, & Richardson, 2013; Sawatsky, Wong, Hinkelmann, Cattaneo, & von Messling, 2012).

The clinical outcome of cross-species transmission has three important consequences for CDV emergence in novel populations. Cross-species transmission leading to a subclinical outcome is likely a dead-end with respect to spillover because onward transmission is precluded if the virus does not infect epithelial cells (Noyce et al., 2013; Sawatsky et al., 2012). Subclinical individuals do not succumb to distemper disease; therefore, CDV exposure should not lead to population decline (in the absence of co-infections). Subclinical individuals become immune, thus reducing the number of susceptible individuals in the population, i.e. increasing herd immunity (Appel et al., 1974). In the face of a lethal outbreak, herd immunity ensures that some individuals survive and decreases the rate of pathogen spread to others in the population. CDV emergence in a novel population is usually recognized when it causes clinical infection because of its impact on individuals and the population in some cases.

CDV is spread primarily by inhalation of aerosol droplets, thus contact is considered requisite for exposure (Greene, 2006). Infectious virions are shed in saliva, nasal-ocular fluid, urine, feces and blood but is probably only viable outside of the host for a short time (e.g. < 3 hours at room temperature (Greene, 2006)). Characteristic symptoms of CDV include ocular and nasal discharge, sneezing, coughing, diarrhea, hyperkeratosis, ataxia, myoclonus, seizures and

death. Recovery from CDV infection confers strong, long lasting immunity. In Africa, it is thought that cross-species transmission occurs at shared kill sites through interference competition and/or via virus deposited and ingested at the carcass. Persistent subclinical infections, i.e. a carrier state, are thought to be nonexistent (Greene, 2006)

In 1994, CDV emerged in the African lion (*Panthera leo*) population in Serengeti National Park (SNP) resulting in the disappearance of an estimated one-third of the population (Roelke-Parker et al., 1996). Clinical CDV infection was also confirmed in domestic dogs (*Canis lupus familiaris*), spotted hyena (*Crocuta crocuta*), bat-eared foxes (*Otocyon megalotis*), and suspected in other wild carnivores (Roelke-Parker et al., 1996). Early detection coupled with a highly visible carnivore community facilitated a careful disease investigation. Hundreds of specimens across multiple species were collected as the outbreak spread among a well-studied carnivore community.

At the time of the outbreak there was enough evidence to conclude that high-density domestic dog populations abutting SNP likely seeded the outbreak. Serological data showed that a sympatric domestic dog population northwest of SNP had been exposed in all of 3 years that were monitored leading up to the outbreak (Cleaveland et al., 2000). African lions on the other hand had not been exposed to CDV in 13 years leading up to the outbreak, suggesting that a wild reservoir was not present in SNP in that period. Moreover, a phylogenetic analysis of sub genomic sequence data (389 bp of P gene, 257 bp H gene) at the time suggested that a single lineage circulated among all affected carnivores, wild and

domestic, with species boundaries presenting no resistance (Carpenter et al., 1998).

Based on this understanding of origin and transmission dynamics, conservationists implemented a disease control program to prevent African lions, the target population, from contacting infectious domestic dogs. Domestic dogs in the immediate vicinity of Serengeti National Park have been vaccinated annually since 1997 (Viana et al., 2015). However, despite these prodigious efforts, the program has not prevented lion exposure as evidenced by periodic seroconversion of monitored lions (Munson et al., 2008; Viana et al., 2015).

In contrast to the earlier analysis, a recent phylogenetic analysis using whole genome sequences (15,690 bp) revealed that in fact two distinct CDV genotypes were in circulation in canid and non-canid species during the outbreak (Nikolin et al., 2017). A genetic distinction between CDV sequences from distantly related hosts during the same outbreak suggests the presence of a genetic barrier to cross-species transmission between these groups. The authors concluded that due to this barrier, it was unlikely that domestic dogs caused the 1993/1994 outbreak in lions (Nikolin et al., 2017). However, interpretation of these data was limited because small sample size and a lack of spatiotemporal data precluded analyses which could date the divergence of lineages and ascertain their origin. Thus, the role of domestic dogs in the 1993/1994 outbreak is still not clear.

What is clear is that domestic dogs did not have high rates of contact with lions in the center of Serengeti National Park, where 17 of 18 studied prides were infected in 1994 (Craft, Hawthorne, Packer, & Dobson, 2008). How then did

CDV reach such high incidence in the lion population? The stochastic susceptible-infected-recovered (S-I-R) model by Craft et al. (2008) supports the hypothesis that spread in the lion population was best explained by cross-species transmission from other wild carnivores however to date no empirical studies have addressed this question. Detailed knowledge of the key host species and pathways of transmission to Serengeti lions is vital for developing a disease control strategy but is lacking.

In this study, we integrate epidemiological data and near whole CDV genome sequences generated directly from clinical specimens to reconstruct the origin and transmission dynamics of the 1993/1994 outbreak at an unprecedented resolution. Here we chronicle the emergence of a lethal CDV lineage and describe its subsequent spread in a complex carnivore community during the 1993/1994 Serengeti outbreak. These results reveal how the 1993/1994 CDV outbreak unfolded in time and space across a complex multi-host system and provide insights into how CDV emerges in vulnerable populations of conservation concern.

MATERIALS AND METHODS

Study specimens

Clinical specimens from which genomes were generated in this study came from 21 individuals representing 4 species, African lion (*Panthera leo*), spotted hyena (*Crocuta crocuta*), bat-eared fox (*Otocyon megalotis*), and domestic dog (*Canis lupus familiaris*). Specimens were collected during a *Canine distemper virus* outbreak between December 1993 and November 1994 in the

Serengeti Ecological Region (SER), an area that includes SNP and adjoining conservation areas, and an unprotected area of dense human settlement northwest of SNP (Fig. 2.1, Table 2.1). The area over which these samples were collected is spatially contiguous, without physical barriers to host movement. However, no canids were sampled inside of SNP boundaries and no non-canid hosts were sampled outside of SNP despite that both occur in both places. Seven of 10 African lion individuals were part of a long-term ecological study on Serengeti lions whose condition and location (and those of their pride mates) were ascertained biweekly. Three lions were unknown to researchers and have presumed origins outside of the study area. Additionally, 16 blood serum samples were collected from spotted hyenas for surveillance during capture for snare removal, research, or disease investigation between December 1992 and June 1995. Specimens were stored in liquid nitrogen or kept at -80°C until used.

RT-qPCR

A reverse transcription quantitative PCR (RT-qPCR) assay was used to screen CDV-suspect specimens for viral RNA (Path-ID™ Multiplex One-Step RT-PCR Kit). The RT-qPCR reaction conditions were as follows: 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C, and 60 sec at 60°C. Each reaction used 2-ul total RNA extract in a 25-ul volume reaction. Oligonucleotide concentrations were used at concentrations according to manufacturer's instructions.

RT-PCR

Twenty of the 21 genomes in this study were generated using amplicon-based deep sequencing. CDV-positive total RNA extracts were reverse

transcribed and PCR amplified in 13-15 overlapping segments using previously published primer sets (Riley & Wilkes, 2015) to obtain whole CDV genomes minus the extreme 3' and 5' non-coding ends (attaining 15,547 bp). Three samples failed to amplify all 15 PCR products and consequently subgenomic sequences were generated for these (i.e. < 15,547 bp). Amplicons were generated using either a one-step (SuperScript™ III One-Step RT-PCR System with Platinum™ *Taq* DNA Polymerase) or two-step (SuperScript™ IV First-Strand Synthesis System with Q5® High-Fidelity DNA Polymerase or QIAGEN Multiplex PCR Kit) protocol. CDV-positive total RNA extracts were DNase-treated prior to RT-PCR (OPTIZYME™ DNase I (RNase-Free), Fisher BioReagents™).

One-step reactions. The one-step reaction was performed according to manufacturer's instructions using 2-ul total RNA in a 25-ul reaction. The reaction conditions were as follows: 55°C for 30 min for reverse transcription, followed by a 2-minute initial denature step and 40 cycles of 15 sec at 94°C for denaturation, 30 sec at 50° C annealing, and 1.2 min at 68°C for elongation. Finally, a 5-minute final extension was performed at 68°C.

Two-step reactions. First strand cDNA synthesis was performed with 8-ul total RNA in a 20-ul reaction according to manufacturer's instructions for gene specific primers (SuperScript™ IV First-Strand Synthesis System). Following RT, either Q5® High Fidelity PCR Kit or QIAGEN® Multiplex PCR Kit was used for PCR. *Q5® High Fidelity PCR Kit* reactions used 2-ul cDNA in a 20-ul reaction. The reaction conditions were as follows: an initial denature at 98°C for

30 sec followed by 40 cycles of 10 sec denaturation at 98°C, annealing for 30 sec at 60°C or 63° C (Table 2.2), and elongation for 1.3 min at 72°C, with a final extension for 2 min at 72°C. For two primer sets these conditions did not produce the desired product and a touchdown PCR was designed with the following conditions: an initial denature at 98°C for 2 min followed by 15 cycles of 10 sec denaturation at 98°C, annealing for 30 sec at 65°C or 68°C (see table) during the first cycle and -1 degree per cycle each of the remaining cycles, an elongation for 2 min at 72°C, followed by 30 cycles of 10 sec denaturation at 98°C, annealing for 30 sec at 50°C, and elongation for 2 min at 72°C, with a final extension for 2 min at 72°C. Some PCR reactions failed consistently with a subset of samples. In this case, primers from contiguous amplicons were combined, e.g. F6 and R7, doubling amplicon length.

PCR reactions using the *QIAGEN® Multiplex PCR Kit* required 2.5-ul cDNA in a 25-ul reaction. The reaction conditions were as follows: an initial denature at 95°C for 15 minutes followed by 40 cycles of 30 sec denaturation at 94°C, annealing for 2 min at 60°C, and elongation for 1.5 min at 72°C, with a final extension for 10 min at 72°C.

Whole genome sequencing

Twenty of 21 libraries were prepped from 13-15 standardized and pooled PCR amplicons. Briefly, pools of 500 ng were sheared on a Covaris E220 Focused-ultrasonicator following the microTUBE protocol for 500 bp in 130-ul. Sonication products were then purified using Agencourt® AMPure® XP PCR purification beads at a 1:1.8 ratio and eluted in 50ul of 10mM Tris-HCl. Libraries

were prepped using the TruSeq® Nano DNA Library Prep kit. Paired-end libraries were sequenced with an Illumina MiSeq producing 2 x 150 bp reads. One library was prepped directly from total RNA (i.e. without RT-PCR) according to manufacturer's protocol (NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina®).

CDV Amplicon Genome Alignment

Paired-end Illumina data generated from amplicons across all 50 samples was trimmed for Illumina adapters and barcodes using Trimmomatic, with bases below Q25 trimmed and bases below Q30 masked aligned to EU716337.1 using BWA. Total coverage and nucleotide counts were calculated on a per-base basis for every position with an alignment corresponding to a minimum of 45 samples (position 77 to 15,623). Consensus sequences were derived from strict, 90%, nucleotide frequency. These sequences have been deposited in GenBank (accession numbers to be included following submission). Intrahost single nucleotide variants (iSNVs), indicating a heterogenous viral population, were derived as any base where a minimum of 10% and less than 90% of the reads indicated an alternate base.

Genome assembly and alignment

Paired end Illumina data generated for the library prepped from total RNA (i.e. no RT-PCR) was trimmed for Illumina adapters and barcodes using Trimmomatic 0.36 (Bolger, Lohse, & Usadel, 2014) and bases below Phred-scaled quality score (Q) 25 were removed. De novo genome assembly was performed using Velvet for k-mer sizes 25, 29, 31 and 35 (Zerbino & Birney,

2008), and using SPAdes 3.9 with variable k-mers (Bankevich et al., 2012). Resulting contigs were compared to GenBank using BLAST (Morgulis et al., 2008) to filter host (*Crocuta crocuta*) or exogenous DNA from de novo assembly. Contigs were aligned to the reference genome most similar to remaining contigs (Canine Morbillivirus virus isolate 164071, accession EU716337.1) using the Burrows-Wheeler aligner implemented in BWA (Li & Durbin, 2009). A consensus sequence was generated using 90% identity threshold. This consensus sequence has been deposited into GenBank under accession number (upon submission).

Serum neutralization assay

Virus neutralization (VN) assay was performed on hyena serum samples to assess exposure to CDV and followed standard procedures for VN assays in microtiter plates. Two-fold serum dilutions (50-ul) in duplicate are mixed with 100-300 $TCID_{50}$ of CDV (Onderstepoort strain – Baker Institute) in a 50-ul volume. Mixtures are allowed to incubate for at least 1 hr at room temperature. A 100-ul volume of indicator cells (Vero - ATCC) is added to each well and the plates are placed in a CO_2 incubator at app 37°C for 4 days. Wells are scored for the presence or absence of typical CDV cytopathology. Antibody values are given as titers (reciprocal of end-point dilution). Our calculation uses serum dilutions with a 50% end-point determination. In instances where the test samples exhibit toxicity to the indicator cells, the medium in the microtiter plates is changes after 12-18 hrs on test. This change of medium does not affect the Ab titer of the sample, but does reduce the toxic properties of some of the test samples.

Genetic analysis

Sequence alignments were performed using the MUSCLE algorithm in MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Edgar, 2004; Kumar, Stecher, & Tamura, 2016), and included 21 near whole genome sequences (Table 2.1). The Hasegawa-Kishino-Yano model (HKY, (Hasegawa, Kishino, & Yano, 1985)) model of nucleotide substitution with uniform rates among sites was the best fit to the data as determined by the Bayesian Information Criterion (BIC) in jModelTest (Posada, 2008). The phylogenetic relationships of CDV sequences were reconstructed, rooted with the Onderstepoort genome, using the Maximum Likelihood method in MEGA (Kumar et al., 2016).

Haplotype networks were reconstructed from 19 sequences (15,050 – 15,547 bp) using the TCS algorithm in PopARt to visualize the relationships among genotypes in our sample and count the number of mutational steps between them (Clement, M., Snell, Q., Walker, P., Posada, D., & Crandall, 2002; Leigh & Bryant, 2015).

Time-measured MCC

To estimate the evolutionary rate, divergence times, and phylogenetic relationships between CDV RNA sequences in our sample we used a Bayesian Markov Chain Monte Carlo approach implemented in BEAST software package version 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012) with the BEAGLE library (Ayres et al., 2012). The data alignment consisted of 21 CDV RNA sequences (Table 2.1). Prior to this analysis, we explored these sequence data for

evidence of recombination using RDP4 and SimPlot (Lole et al., 1999; Martin & Rybicki, 2000). Additionally, TempEst was used to qualitatively explore whether the sequence data contained sufficient temporal signal for this analysis which relies on the accumulation of mutations between heterochronously sampled sequences (Rambaut, Lam, Carvalho, & Pybus, 2016). Preliminary analyses indicated that the best-fit clock model of sequence evolution was the strict clock (over a relaxed lognormal clock model (BF = 50.99, Table 2.3)). The best-fit demographic tree prior of 4 that were tested was the coalescent constant population (See Table 2.3 for details).

Three independent MCMC chains were run for 100M steps, of which 10M were discarded as burn-in. Tree and parameter files were logged every 10,000 steps. Traces were checked in Tracer for convergence, i.e. ESS>200. The program TreeAnnotator from the BEAST package was used to combine and annotate trees and the resulting maximum clade credibility tree was visualized in FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

Discrete Trait Analysis

We estimated the host state at all branches and internal nodes across the time-measured MCC tree using a discrete phylogenetic diffusion model and 4 host states (African lion, spotted hyena, bat-eared fox, and domestic dog) in BEAST version 1.8.4 (Lemey, Rambaut, Drummond, Suchard, & Ali, 2009). A non-reversible continuous-time Markov chain was specified to allow the estimation of asymmetric host transition rates and the Bayesian Stochastic Search

Variable procedure (BSSVS) was enabled to constrain the number of rates allowed for explaining the history of cross-species spread.

RESULTS

Phylogenetic analysis of 21 CDV near-WGS (11,926-15,547 bp) from multiple hosts sampled during the 1993/1994 outbreak in SER rooted with the Onderstepoort vaccine CDV strain indicates that two unique but closely related CDV lineages were in circulation (similarity 99.92%) (Fig. 2.2). The lineages are separated by taxa such that one monophyletic clade contains only canid species, domestic dog and bat-eared fox sampled in the southeast part of the study area, and the other contains only species that are non-canid, African lion and spotted hyena sampled for the most part in SNP. Hereafter these lineages are referred to as the "canid" and "non-canid" lineages following the distinction made in previous studies on CDV host tropism (McCarthy, Shaw, & Goodman, 2007; Nikolin et al., 2017).

Of 18 lion prides under surveillance, 17 were exposed to CDV, the majority of which suffered associated morbidity and mortality (Craft et al., 2008). CDV sequences generated from lions in 4 prides fall into 3 distinct clusters, hereafter transmission groups as the sequences share common ancestry. Hyena sequences are interspersed among all transmission groups and are basal in 2 of 3 groups that also contain lions. Together these results support that multiple cross-species transmission events contributed to the extent of CDV incidence in the lion population, rather than an exclusive lion-lion or pride-pride transmission

explanation, and suggest that hyenas may have been a source of CDV infection in lions.

The non-canid clade is characterized by a large polytomy to the exclusion of one sequence (PLE-658), suggesting an explosive radiation from a common ancestor of these viral genotypes. PLE-658 was an unknown lioness found dead within the territory of a well-known study pride in May 1994, after most mortality in the lion study population had occurred. That she was unknown to researchers suggests that she was from outside the study area, possibly a nomad or displaced by social upheaval in the wake of widespread mortality. Importantly all other non-canid sequences regardless of time or location of sampling, clustered together, which could suggest that the chain of transmission that led to infection in PLE-658 was a dead end or stuttering chain of transmission.

A time-calibrated maximum clade credibility tree under a strict molecular clock model of evolution dates the most recent common ancestor (MRCA) of the canid and non-canid associated lineages as approximately 1.76 years before the last specimen was collected (i.e., Feb, 1992) (95% HPD: 1.12-2.38 y) (Fig. 2.3). This timing is roughly 10.2 months before the index case in the non-canid lineage, a clinically infected hyena cub necropsied on December 20, 1993. The MRCA between PLE-658 and all other non-canid sequences was approximately 1.46 years before the last specimen (95% HPD: 1.13-1.83 y), or around 6.6 months before the non-canid index case. The approximate timing of the radiation of all non-canid viral genotypes excluding PLE-658 was around 1.23 years before the last specimen (95% HPD: 1.02-1.47 y) or 3.8 months before the non-canid index

case. The median evolutionary rate of CDV during this outbreak was 6.525×10^{-4} substitutions per site per year, i.e. 10.2 substitutions per year, consistent with other CDV reports.

A haplotype network analysis of 18 near-complete CDV genomes (15,050–15,547 bp) supports the genetic distinction between lineages associated with canid and non-canid species and reveals 13 nucleotide substitutions separating the two (Fig. 2.4a). Of these, 7 substitutions are nonsynonymous and occur in the P/V, F, and H regions of the genome (Fig. 2.4b.), genes whose functions facilitate efficient replication, host immune evasion, host cell receptor binding and membrane fusion. The position and character of amino acid changes in the genome are listed in Table 2.4. PLE-658 shares 1 synonymous and 2 nonsynonymous nucleotide substitutions with the canid lineage in the L gene region (asterisks Fig. 2.4b.), a protein that catalyzes the replication of viral genomic RNA. The residues on the L gene that are affected by these substitutions are conserved across all other published CDV sequences. PLE-658 does not show evidence of recombination (data not shown).

Serological assays for CDV of serum from hyenas sampled before, during and after the onset of clinical distemper in non-canids suggest that this population was either not exposed to CDV before the index case, or that CDV presence was below the level of detection given sample size (Table 2.5). All of 12 apparently healthy adult and sub-adult hyenas sampled between December 29, 1992 and September 23, 1993 were CDV seronegative. Three clinically ill hyena cubs (as determined by histopathology, Munson unpublished data) sampled between

December 20, 1993 and July 3, 1994 had CDV titer values that were too low to distinguish from cross-reactivity with other pathogens (titers 8-16). One clinically ill pregnant adult (Linda Munson, unpublished data) had no detectable titers at a dilution of 64 but the assay failed at dilutions lower than this due to toxicity. All of 4 adults sampled between January 6, 1995 and June 20, 1995 were seropositive with titers ranging from 96 to 256.

Ancestral host-state reconstruction predicted that the most probable origin of CDV seeding the 1993/1994 outbreak in non-canids was domestic dogs (PP = 0.67) (Fig. 2.5). Cross species transmission between canids and non-canids following this spillover did not occur, or was below the level of detection by our methods. Conversely, the analysis shows that cross-species transmission was common from hyenas to lions with hyenas being the highest probable predicted host state for all transmission groups containing lions (PP = 0.64-0.72). The ancestral state at two nodes were unresolved (PP = 0.54 and 0.53) possibly owing to sequence gaps in informative regions in sequences at the tips of these nodes. Intra-species transmission was predicted to have occurred in both hyenas and lions. Similarly, intraspecific transmission was predicted in both domestic dogs and bat-eared foxes with high probability (P = 0.99), however resolution at internal nodes is unresolved (P = 0.50 and 0.51) due possibly to the small sample size.

To characterize the spatial diffusion of non-canid lineage of CDV during the outbreak, the locations of individual samples were mapped according to their transmission group determined by phylogenetic clustering. Seronera, a human

settlement at the center of SNP, appears to be the epicenter of spatial diffusion of the non-canid lineage of CDV to peripheral areas of the SER (Fig. 2.6). Within each transmission group in the non-canid lineage, the earliest sample was collected in or near Seronera and those sampled later were collected also in Seronera or at a distance to the north, southeast, southwest or west. One important exception to this pattern is a transmission group that originates in the west corridor on December 23, 1993 (the second earliest non-canid case by 3 days) and terminates in Seronera 1-1.5 months later. This observation is consistent with the hypothesis that CDV originated in the domestic dog population in the Serengeti District, which is less than 17 km away to the north.

That the PLE-658 sequence CDV is basal to all other non-canids, and that she was not from Seronera suggests that a separate chain of transmission of the non-canid lineage circulated outside of Seronera.

DISCUSSION

In this study, we found evidence that a large-scale, lethal outbreak of *Canine distemper virus* in Serengeti lions was precipitated by a single cross-species spillover event from a canid reservoir, followed by multiple cross-species transmissions from a non-canid host (e.g. spotted hyena) as well as onward transmission from lion to lion. Our data suggests that a CDV lineage associated with non-canid hosts diverged from a canid-associated lineage less than 11 months before the detection of clinical signs and explosive spread of CDV in non-canid hosts. Seven amino acid residue differences in biologically relevant regions of the CDV genome separate these lineages, all of which had accumulated in less

than 3.6 months of divergence, and 6.6 months of the explosive onset of CDV-related mortality in non-canids. These results together suggest the presence of a host barrier to CDV infection between canids, and lions and hyenas. In contrast to this genetic distinction between canid and non-canid associated lineages, our data show that sequences between species infected by the same lineage are highly similar suggesting that there is no or a low barrier to clinical CDV infection between more closely related host species in our sample, e.g. between domestic dog and bat-eared fox, and between African lion and spotted hyena. Ancestral host-state reconstruction suggests that hyenas may have been an important source of CDV infecting lions and affirms that onward transmission in the lion population also likely occurred.

Single spillover from canid reservoir

Our data support the hypothesis that a domestic dog population was the origin of the CDV lineage infecting non-canids in 1993/1994. Bayesian analysis of time-stamped genomic data revealed that the two lineages of CDV circulating in 1994 diverged from a common ancestor just months before the detection of clinical signs in Serengeti lions and hyenas. Integration of host species data in ancestral state reconstruction predicted that this common CDV ancestor had a domestic dog origin (PP = 0.67). Multiple incursions of CDV from a canid reservoir were unlikely to have fueled the outbreak in non-canid species given the high genetic similarity of all non-canid species sequences to the exclusion of the canid sequences. This interpretation is supported by the spatial diffusion of the non-canid lineage. We observed a star-like pattern of CDV spread of the non-

canid lineage in which most chains of transmission originate in the Seronera area and terminate peripherally, consistent with the epidemiological observations of previous reports (Cleaveland et al., 2007; Roelke-Parker et al., 1996). That all non-canid sequences from clinically-infected hosts can be traced to a single spillover despite that CDV was circulating in domestic dogs (and African wild dogs in the northern reaches of the ecosystem) for at least 3 years prior to the outbreak (Alexander & Appel, 1994; Cleaveland et al., 2000), suggests that host barriers to clinical infection may limit lethal outbreaks of CDV in novel non-canid species.

Our results shed new light on the origins of the 1993/1994 outbreak. The phylogenetic discovery of Nikolin et al. (2017) that two strains circulated in canid and non-canid species in the Serengeti in 1993/1994 challenged the previous model of CDV emergence in Serengeti lions that had specified a domestic dog origin (Cleaveland et al., 2000). Specifically, their study concluded that genetic differences separating the two strains signify a genetic barrier to cross-species transmission between canids and non-canids, and therefore CDV in non-canids must have come from a non-canid origin. Our analysis of a novel, whole CDV genome dataset from unpassaged specimens upholds the finding that two strains circulated in canids and non-canids in 1993/1994. However, the integration of epidemiological and genomic data in our analysis offered the unique discovery that the CDV lineage causing mortality in non-canid species in 1993/1994 had evolved in a few months from a canid progenitor. This is not surprising given that the lineages share 99.82% similarity. Nevertheless, this result is illuminating

because it reveals the process of emergence of a lethal CDV genotype in spotted hyenas and African lions from a canid origin.

Transmission Dynamics within lineages

Our results indicate that a significant amount of cross-species transmission *within* CDV lineages occurred during the outbreak. Sequences from within lineages were highly similar regardless of host species. For example, in the canid lineage, the bat-eared fox sequences bear only a single nucleotide difference from their most recent common ancestor with domestic dogs. In this case study, genetic similarity of sequences between more closely related host species suggests that host barriers to clinical infection could be mediated by host phylogenetic relationships.

Within the non-canid lineage, CDV phylogeny and haplotype network topologies reveal that hyenas may have been an important source of CDV in Serengeti lions at the population scale during the 1994 outbreak. If lion-to-lion transmission alone were responsible for CDV spread, then we would expect all lions to cluster together in these analyses, irrespective of pride affiliation. However, sequences from 4 sampled prides fall into 3 distinct clusters signifying that cross-species transmission must be responsible for exposure in most prides sampled within the non-canid lineage. Furthermore, hyenas are interspersed in every transmission chain that reached lions and occupy a basal position in two of these, supporting the role of hyenas as a CDV source. Ancestral host state reconstruction supports this interpretation, implicating hyenas as the most

probable source of CDV (of species that were sampled) in all transmission groups of the non-canid CDV lineage.

While spotted hyenas may have played a key role in driving the spread of the non-canid CDV lineage in Serengeti lions at the population-scale, we found evidence for lion-to-lion transmission at the local scale. Specifically, 5 lion sequences, representing two neighboring lion prides and one individual lion of unknown origin, fall in a single cluster suggesting that lion-to-lion transmission occurred. This interpretation is supported by ancestral host state reconstruction that predicts lion-to-lion transmission with high probability (PP = 0.97-0.99) in this cluster. Thus, lion-to-lion transmission may partially explain exposure patterns within and between lion prides.

It is important to note that assigning transmission roles in sparsely sampled communities during a multi-host outbreak is complicated when unsampled susceptible host species are unaccounted for in the data. If a host species that was not sampled, e.g. mongoose, was responsible for CDV spread to both hyenas and lions, then our data might wrongly assign transmission roles to the species that were sampled. Although we cannot rule out the possibility that an unsampled host drove the observed patterns in our data, we believe that our interpretation that hyenas drove the spread of CDV in the lion population is supported by the epidemiological evidence and hyena movement ecology.

The sequence of clinical observations in wild carnivores suggests that CDV circulated in hyenas before causing disease in lions. Histopathology results from lymphoid and meningeal tissues support a CDV diagnosis in a lion-killed

hyena as early as April 1993 (Linda Munson, unpublished data). Haas et al. (1996) report clinical signs in hyena cubs as early as November 1993, whereas the first lion mortalities related to the outbreak were not reported until January 1994 and clinical signs would not be observed in a living lion until February (Roelke-Parker et al., 1996). If only a single spillover can be inferred from the data, and clinical signs in hyenas were observed months ahead of clinical signs in lions, it follows that hyenas could be the source of CDV for lions.

Hyena movement ecology is consistent with the inference that hyenas spread CDV at the population level in lions. CDV emerged in 6 lion prides almost simultaneously despite that these prides were mostly spatially discontinuous and separated by up to 40 km (Craft, Hawthorne, Packer, & Dobson, 2008). The four most isolated of these prides were located on the short grass plains where, at the time of the outbreak, hundreds of thousands of wildebeest were congregating for the wet season. Hyenas commute long distances (40-80km) from den sites in the Seronera area to the short grass plains to take advantage of the abundant prey (Hofer & East, 1993). Commuters travel singly and in any one day, several individuals from a single Seronera clan may be dispersed over hundreds of square kilometers, having traversed multiple lion pride and hyena clan territories to get there (Hofer & East, 1993). No other Serengeti carnivore to our knowledge makes such extensive, long-range movements. Previous studies report that clinical infection was limited to hyena cubs in 1994, which would argue against hyena involvement in long-distance spread (Nikolin et al., 2017). In this study, we report clinical CDV infection in an adult pregnant hyena and her intrauterine offspring

(Munson, unpublished data), suggesting that pregnant or otherwise immunocompromised adult hyenas were competent hosts and capable of spreading CDV.

Our data do not shed light on how CDV spread to canids to the southeast of SNP. Epidemiological evidence alone would suggest that CDV "spilled back" into dogs from wild non-canids as reported earlier (Cleaveland et al., 2007) however our data refute that conclusion.

CDV evolution during the 1994 outbreak

Viral evolution can surmount host barriers determining cross-species transmission and/or clinical severity in hosts (Brault et al., 2007; Hueffer & Parrish, 2003). In this study, we identified 13 nucleotide substitutions coding 7 amino acid residue substitutions in biologically relevant regions of the CDV genome separating the canid and non-canid lineages. Molecular dating places the accumulation of these mutations within 3.5 months after the divergence of the two lineages, indicating a rate of evolution along that branch roughly 1.7 times the average that we estimated (6.525×10^{-4} substitutions per site per year, 95% HPD: 4.27E $\times 10^{-4}$ -9.13 $\times 10^{-4}$). The divergence of CDV genotypes associated with distantly related host species, driven by accumulation of mutations in such a short period, suggests that CDV evolution reduced the host barrier leading to clinical infection in lions and hyenas.

While a distinction between the lineages is clear, we cannot rule out the possibility that the genetic differences were driven by space and genetic drift. The canid and non-canid lineages share a common ancestor ~1y before the outbreak

started. The canids that we sampled were from the southeast portion of the study area, the domestic dogs being from a low-density population that experienced sporadic CDV outbreaks and had not been exposed since 1991 (Cleaveland et al., 2000). Thus, our results suggest that CDV spilled over to this population from the same source as that which spilled over to the non-canids. Due to opportunistic sampling, it is unclear whether the canid and non-canid lineages circulated simultaneously within SNP, or if their distributions were parapatric, with the canid lineage only circulating outside SNP. There is no evidence that the canid community inside SNP was affected by CDV during the outbreak. Could the two lineages have diversified because of drift or adaptation to different places? Because CDV is transmitted directly, with no external life stage, we think that it is more likely that the lineages adapted to their immediate selective environment, their hosts. It is unclear whether the mutations acquired by the non-canid lineage preclude their infectivity in a canid species.

Our data do not address if the observed mutations in the non-canid lineage are necessary and sufficient for clinical infection in lions and hyenas. However, a CDV sequence recently published from a hyena in South Africa (Loots, Du Plessis, Dalton, Mitchell, & Venter, 2017) bears 4 of the mutations that were observed in non-canids in SER in 1993/1994 (V- G134S, H-D178G, H- R519I and Y549H). This is notable because 1) this is the first case of CDV in a hyena ever reported outside of the 1993/1994 outbreak in SNP, and 2) the sequence from South Africa belongs to a distinct and different strain (Africa-1) than the Serengeti sequences (Africa-2) (Loots et al., 2017). That the South African and

East African sequences share rare mutations despite having very distant genetic backgrounds overall, supports that these mutations have functional significance in spotted hyena, and possibly African lion, clinical infection.

Management implications

Though physiological and behavioral differences between host species may limit CDV clinical infection and thus spillover outcome, we show that one or a few evolutionary steps can overcome these barriers in a short time frame. Nevertheless, the occurrence of *lethal* spillover in this system is apparently rare. This improved understanding of CDV spillover in African lions allows managers to weigh the risks of CDV exposure. For example, CDV management in the Serengeti has been focused on preventing lion exposure to CDV from domestic dogs (Viana et al., 2015). However, data from Serengeti lions and other wild felid populations suggest that exposure to non-lethal CDV strains can actually increase herd-immunity (Munson et al., 2008; Viana et al., 2015). Thus, natural exposure to CDV may protect Serengeti lions and hyenas from the rare occurrence of a lethal spillover.

Our results suggest that in the event of a lethal outbreak, reactionary disease control measures must focus on multiple host species, i.e. preventing lion-to-lion or pride-to-pride transmission alone will not be effective. Our results highlight the role of spotted hyenas as a significant source of CDV infection in African lions and driver of long-distance CDV spread of a lethal CDV genotype during this outbreak. The appearance of CDV clinical signs in hyenas months in

advance of clinical signs in lions highlights the potential for hyenas as a sentinel species.

We identified four novel specific mutations separating a lethal non-canid lineage from a canid-associated lineage in this outbreak, in addition to three that have been previously discussed (Nikolin et al., 2017). Although it is unclear if any or some of these mutations predict clinical infection in lions and hyenas, this study provides the genomic positions to motivate future experiments. Until a better understanding of the functional significance can be reached, surveillance for these mutations in circulating CDV strains may inform CDV risk assessment in lion and hyena populations of conservation concern.

LITERATURE CITED

Alexander, K. A., & Appel, M. J. G. (1994). African wild dogs (Lycaon pictus) endangered by a canine distemper epizootic among domestic dogs near the Masai Mara National Reserve, Kenya. *Journal of Wildlife Diseases*, *30*(4), 481–485. https://doi.org/10.7589/0090-3558-30.4.481

Appel, M., Sheffy, B. E., Percy, D. H., & Gaskin, J. M. (1974). Canine distemper virus in domesticated cats and pigs. *American Journal of Veterinary Research*, *35*(No.6), 803–806. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19742225506

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., … Suchard, M. A. (2012). BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Systematic Biology*, *61*(1), 170–173. https://doi.org/10.1093/sysbio/syr100

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology*, *19*(5), 455–77. https://doi.org/10.1089/cmb.2012.0021

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brault, A. C., Huang, C. Y.-H., Langevin, S. A., Kinney, R. M., Bowen, R. A., Ramey, W. N., … Miller, B. R. (2007). A single positively selected West Nile viral mutation confers increased virogenesis in American crows. *Na*, *39*. https://doi.org/10.1038/ng2097

Carpenter, M. A., Appel, M. J. G., Roelke-Parker, M. E., Munson, L., Hofer, H., East, M., & O 'brien, S. J. (1998). Genetic characterization of canine distemper virus in Serengeti carnivores. *Veterinary Immunology and Immunopathology*, *65*, 259–266.

Cleaveland, S., Appel, M. G. J., Chalmers, W. S. K., Chillingworth, C., Kaare, M., & Dye, C. (2000). Serological and demographic evidence for domestic dogs as a source of canine distemper virus infection for Serengeti wildlife. *Veterinary Microbiology*, *72*, 217–227.

Cleaveland, S., Mlengeya, T., Kaare, M., Haydon, D., Lembo, T., Laurenson, M. K., & Packer, C. (2007). The conservation relevance of epidemiological research into carnivore viral diseases in the serengeti. *Conservation Biology*,

21(3), 612–622. https://doi.org/10.1111/j.1523-1739.2007.00701.x

Clement, M., Snell, Q., Walker, P., Posada, D., & Crandall, K. (2002). TCS: Estimating gene genealogies. In *Parallel and Distributed Processing Symposium, International Proceedings, 2* (p. 184).

Cottrell, W. O., Keel, M. K., Brooks, J. W., Mead, D. G., & Phillips, J. E. (2013). First Report of Clinical Disease Associated with Canine Distemper Virus Infection in a Wild Black Bear (*Ursus americana*). *Journal of Wildlife Diseases*, *49*(4), 1024–1027. https://doi.org/10.7589/2013-02-027

Craft, M. E., Hawthorne, P. L., Packer, C., & Dobson, A. P. (2008). Dynamics of a multihost pathogen in a carnivore community. *Journal of Animal Ecology*. https://doi.org/10.1111/j.1365-2656.2008.01410.x

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian hylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–7. https://doi.org/10.1093/nar/gkh340

Greene, C. E. (2006). *Infectious diseases of the dog and cat*. Saunders Elsevier.

Greene, C. E., & Appel, M. J. (1990). Canine Distemper. In C. E. Greene (Ed.), *Infectious Diseases of the Dog and Cat* (pp. 226–241).

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*(2), 160–174. https://doi.org/10.1007/BF02101694

Hofer, H., & East, M. L. (1993). The commuting system of Serengeti spotted hyaenas:how a predator copes with migratory prey. II. Intrusion pressure and commuters' space use. *Anim. Behav.* , *46*, 559–574. Retrieved from http://ac.els-cdn.com/S0003347283712236/1-s2.0-S0003347283712236-main.pdf?_tid=a29b89ec-29e8-11e7-9ebc-00000aab0f26&acdnat=1493146659_d6dde528c10a58a925a036e52abcf4d5

Hueffer, K., & Parrish, C. R. (2003). Parvovirus host range, cell tropism and evolution. *Current Opinion in Microbiology*. https://doi.org/10.1016/S1369-5274(03)00083-3

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*, *33*(7), 1870–4.

Leigh, J. W., & Bryant, D. (2015). popart : full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lemey, P., Rambaut, A., Drummond, A. J., Suchard, M. A., & Ali, Y. (2009). Bayesian Phylogeography Finds Its Roots. *PLoS Computational Biology*, *5*(9), e1000520. https://doi.org/10.1371/journal.pcbi.1000520

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., … Ray, S. C. (1999). Full-length human immunodeficiency

virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology*, *73*(1), 152–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9847317

Loots, A. K., Du Plessis, M., Dalton, D. L., Mitchell, E., & Venter, E. H. (2017). Genome Sequences of Three Vaccine Strains and Two Wild-Type Canine Distemper Virus Strains from a Recent Disease Outbreak in South Africa. *Genome Announcements*, *5*(27). https://doi.org/10.1128/genomeA.00603-17

Martin, D., & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *BIOINFORMATICS APPLICATIONS NOTE*, *16*(6), 562–563. Retrieved from http://

McCarthy, A. J., Shaw, M.-A., & Goodman, S. J. (2007). Pathogen evolution and disease emergence in carnivores. *Proceedings. Biological Sciences / The Royal Society*, *274*(1629), 3165–74. https://doi.org/10.1098/rspb.2007.0884

Meli, M., Simmler, P., Cattori, V. M., Fernando Vargas, Astrid Palomares, Fransisco Lopez-Bao, … Lutz, H. (2010). Importance of canine distemper virus (CDV) infection in free-ranging Iberian lynxes (Lynx pardinus). *Veterinary Microbiology*, *146*(1–2), 132–137. https://doi.org/10.1016/J.VETMIC.2010.04.024

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., Schäffer, A. A., … An, P. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics*, *24*(16), 1757–1764. https://doi.org/10.1093/bioinformatics/btn322

Munson, L., Terio, K. A., Kock, R., Mlengeya, T., Roelke, M. E., Dubovi, E., …

Packer, C. (2008). Climate Extremes Promote Fatal Co-Infections during Canine Distemper Epidemics in African Lions. *PLoS ONE*, *3*(6), e2545. https://doi.org/10.1371/journal.pone.0002545

Nikolin, V. M., Olarte-Castillo, X. A., Osterrieder, N., Hofer, H., Dubovi, E., Mazzoni, C. J., … East, M. L. (2017). Canine distemper virus in the Serengeti ecosystem: molecular adaptation to different carnivore species. *Molecular Ecology*. https://doi.org/10.1111/mec.13902

Noyce, R. S., Delpeut, S., & Richardson, C. D. (2013a). Dog nectin-4 is an epithelial cell receptor for canine distemper virus that facilitates virus entry and syncytia formation. *Virology*, *436*(1), 210–220. https://doi.org/10.1016/J.VIROL.2012.11.011

Noyce, R. S., Delpeut, S., & Richardson, C. D. (2013b). Dog nectin-4 is an epithelial cell receptor for canine distemper virus that facilitates virus entry and syncytia formation. *Virology*, *436*(1), 210–220. https://doi.org/10.1016/J.VIROL.2012.11.011

Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, *25*(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Qiu, W. (2011). Canine Distemper Outbreak in Rhesus Monkeys, China. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid1708.101153

Rambaut, A., Lam, T. T., Carvalho, L. M., & Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution*, *2*(1). https://doi.org/10.1093/ve/vew007

Roelke-Parker, M. E., Munson, L., Packer, C., Kockil, R., CIeaveland, S., Carpenter, M., … G AppeIl, M. J. (1996). A canine distemper virus epidemic in Serengeti lions (Panthera leo). *Nature*, *379*, 441–445.

Sawatsky, B., Wong, X.-X., Hinkelmann, S., Cattaneo, R., & von Messling, V. (2012a). Canine distemper virus epithelial cell infection is required for clinical disease but not for immunosuppression. *Journal of Virology*, *86*(7), 3658–66. https://doi.org/10.1128/JVI.06414-11

Sawatsky, B., Wong, X.-X., Hinkelmann, S., Cattaneo, R., & von Messling, V. (2012b). Canine distemper virus epithelial cell infection is required for clinical disease but not for immunosuppression. *Journal of Virology*, *86*(7), 3658–66. https://doi.org/10.1128/JVI.06414-11

Seimon, T. A., Miquelle, D. G., Chang, T. Y., Newton, A. L., Korotkova, I., Ivanchuk, G., … McAloose, D. (2013). Canine distemper virus: an emerging disease in wild endangered Amur tigers (Panthera tigris altaica). *mBio*, *4*(4), e00410-13. https://doi.org/10.1128/mBio.00410-13

Sulikhan, N. S., Gilbert, M., Yu Blidchenko, E., Naidenko, S. V, Ivanchuk, G. V, Yu Gorpenchenko, T., … Seimon, T. A. (2018). Canine Distemper Virus in a Wild Far Eastern Leopard Panthera pardus orientalis. *Journal of Wildlife Diseases Wildlife Disease Association*, *54*(1), 0–0. https://doi.org/10.7589/2017-03-065

Viana, M., Cleaveland, S., Matthiopoulos, J., Halliday, J., Packer, C., Craft, M. E., … Orr, B. (2015). Dynamics of a morbillivirus at the domestic-wildlife interface: Canine Dis- temper Virus in domestic dogs and lions. *PNAS*,

112(5), 1464–1469.

Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–9. https://doi.org/10.1101/gr.074492.107

Table 2.1. List of specimens used to generate whole genome sequence for this study.

Individual ID	Common name	Pride affiliation	Sampling date	Accession no.
CCR-6	Spotted hyena	NA	12/20/1993	Pending publication
CCR-7	Spotted hyena	NA	12/23/1993	Pending publication
CCR-282	Spotted hyena	NA	01/18/1994	Pending publication
PLE-589	African lion	Campsites	01/22/1994	Pending publication
PLE-653	African lion	Campsite	01/28/1994	Pending publication
PLE-654	African lion	Masai	01/31/1994	Pending publication
PLE-595	African lion	Unknown	02/03/1994	Pending publication
PLE-640	African lion	Campsite	02/07/1994	Pending publication
PLE-635	African lion	Transects	02/18/1994	Pending publication
PLE-656	African lion	Simba numbers	02/20/1994	Pending publication
PLE-658	African lion	Unknown	05/21/1994	Pending publication
CCR-10	Spotted hyena	NA	06/20/1994	Pending publication
CCR-11	Spotted hyena	NA	07/03/1994	Pending publication
CCR-12	Spotted hyena	NA	07/09/1994	Pending publication
OME-8	Bat-eared fox	NA	07/16/1994	Pending publication
OME-9	Bat-eared fox	NA	07/25/1994	Pending publication
PLE-652	African lion	Unknown	07/30/1994	Pending publication
CFA-52	Domestic dog	NA	08/23/1994	Pending publication
CFA-51	Domestic dog	NA	09/09/1994	Pending publication
PLE-641	African lion	NA	11/15/1994	Pending publication
CFA-54	Domestic dog	NA	11/18/1994	Pending publication

Table 2.2. Annealing temperatures (Ta) used in PCR reactions with Q5® High Fidelity PCR kit/ mastermix.

Annealing Temperature*	Primer set
60	F1/R1, F2/R2, F3/R3, F4/R4, F5/R5, F6/R6, F8/R8, F9/R9, F10/R10, F15/R15
63	F7/R7, F11/R11, F12/R12
TD 65/50	F13/R13
TD 68/50	F14/R14

* Q5® High Fidelity PCR conditions only

Table 2.3. Bayes factor comparisons for 5 candidate models using MLE stepping stone sampling (right of diagonal), and path sampling model selection (left of diagonal). The model used is in bold italics.

		Relaxed lognormal	Strict			
		Constant	Constant	Exponential	Skyline	Skyride
Relaxed lognormal	Constant	—	-51.33	-42.53	-34.59	-47.30
Strict	***Constant***	***50.99***	**—**	***8.80***	***16.74***	***4.03***
	Exponential	42.22	-8.78	—	7.94	-4.77
	Skyline	34.17	-16.83	-8.05	—	-12.71
	Skyride	46.91	-4.09	4.69	12.74	—

Table 2.4. Amino acid differences between canid and non-canid lineages.

Gene	**Canid aa**	**Position on gene**	**Non-canid aa**
P/V	Glycine (G)	134	Serine (S)
P/V	Lysine (K)	280	Glutamic acid (E)
F	Arginine (R)	375	Glutamine (Q)
H	Arginine (R)	160	Lysine (K)
H	Aspartic acid (D)	178	Glycine (G)
H	Arginine (R)	519	Isoleucine (I)
H	Tyrosine (Y)	549	Histidine (H)
L	Leucine (L)	93*	Phenylalanine (F)
L	Asparagine (N)	1402*	Histidine (H)

*sequence PLE-658 does not bear this mutation

Table 2.5. CDV serological results from spotted hyenas captured before, during, and after the 1993-1994 CDV outbreak.

Capture period	**Dec 92 – Sept 93**	**Dec 93 – July 94**	**Jan 95 – Jun 95**
Seroprevalence (%)	0	0	100
Clinical CDV signs	None recorded	Lymphoid and CNS lesions consistent with CDV*, and/or RT-PCR positive tissue	None recorded
Sample size	12	3	4

* Munson, unpublished data

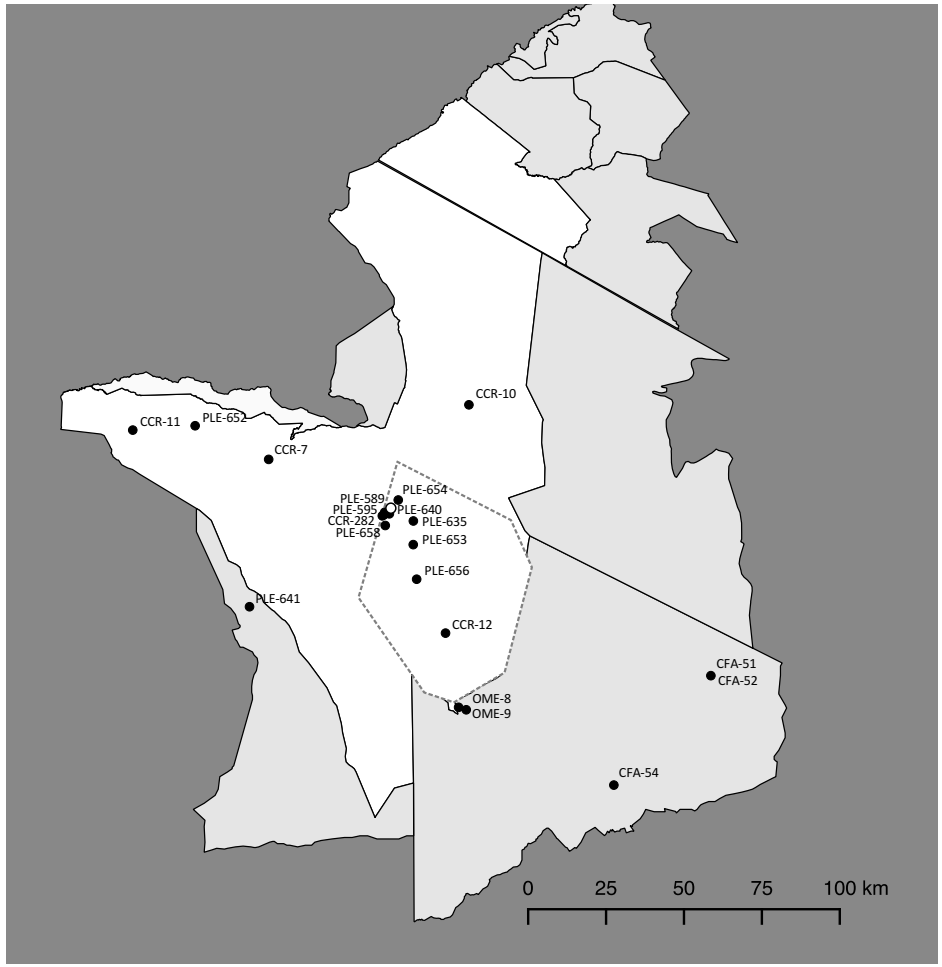

CCR-10
CCR-11
PLE-652
CCR-7
PLE-654
PLE-589
PLE-595
PLE-640
CCR-282
PLE-635
PLE-658
PLE-653
PLE-656
PLE-641
CCR-12
CFA-51
CFA-52
OME-8
OME-9
CFA-54
0
25
50
75
100 km

Figure 2.1. Map of the locations of samples from which CDV sequence was generated for this study. Sample names are prefixed with species code; CFA=domestic dog, CCR=spotted hyena, OME=bat-eared fox, PLE=African lion. White polygons represent national parks, Serengeti National Park to the south and Masai Mara Game Reserve to the north. Light grey polygons represent conservation areas. The dashed line demarcates the Serengeti Lion Project study area. The open circle marks Seronera, a small settlement of SNP staff, wildlife researchers, and tourist lodges.

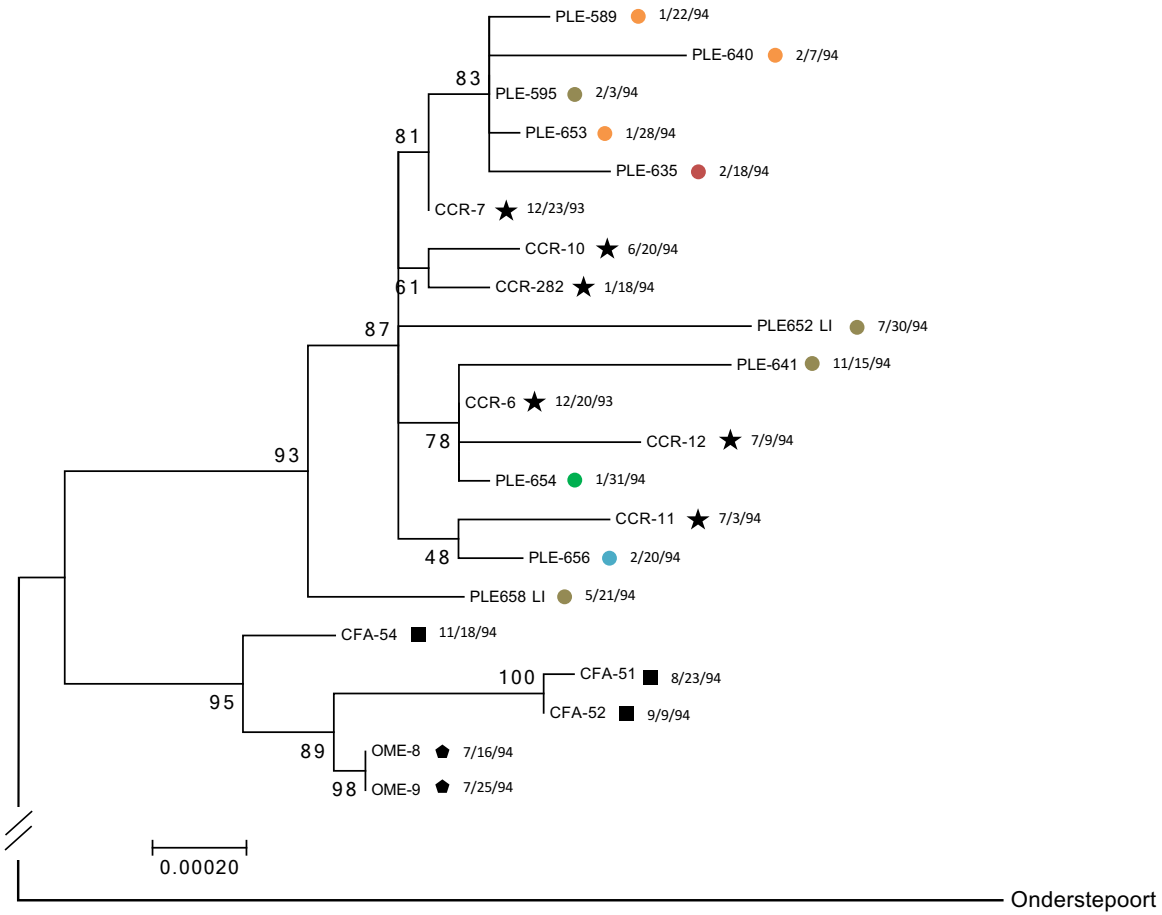


Figure 2.2. Phylogenetic relationship of near whole CDV genome sequences sampled during the 1993/1994 Serengeti CDV outbreak reconstructed using the Maximum Likelihood method in MEGA. Tree tips are labeled with sample name. Symbol

shape indicates host species: square=domestic dog, pentagon=bat-eared fox, star=spotted hyena, circle=African lion. Circle symbol color indicates pride affiliation: orange=Campsites pride, red=Transects, brown=unknown, green=Masai, blue=Simba Numbers. Bootstrap support for nodes shown next to branches (BS=1000).

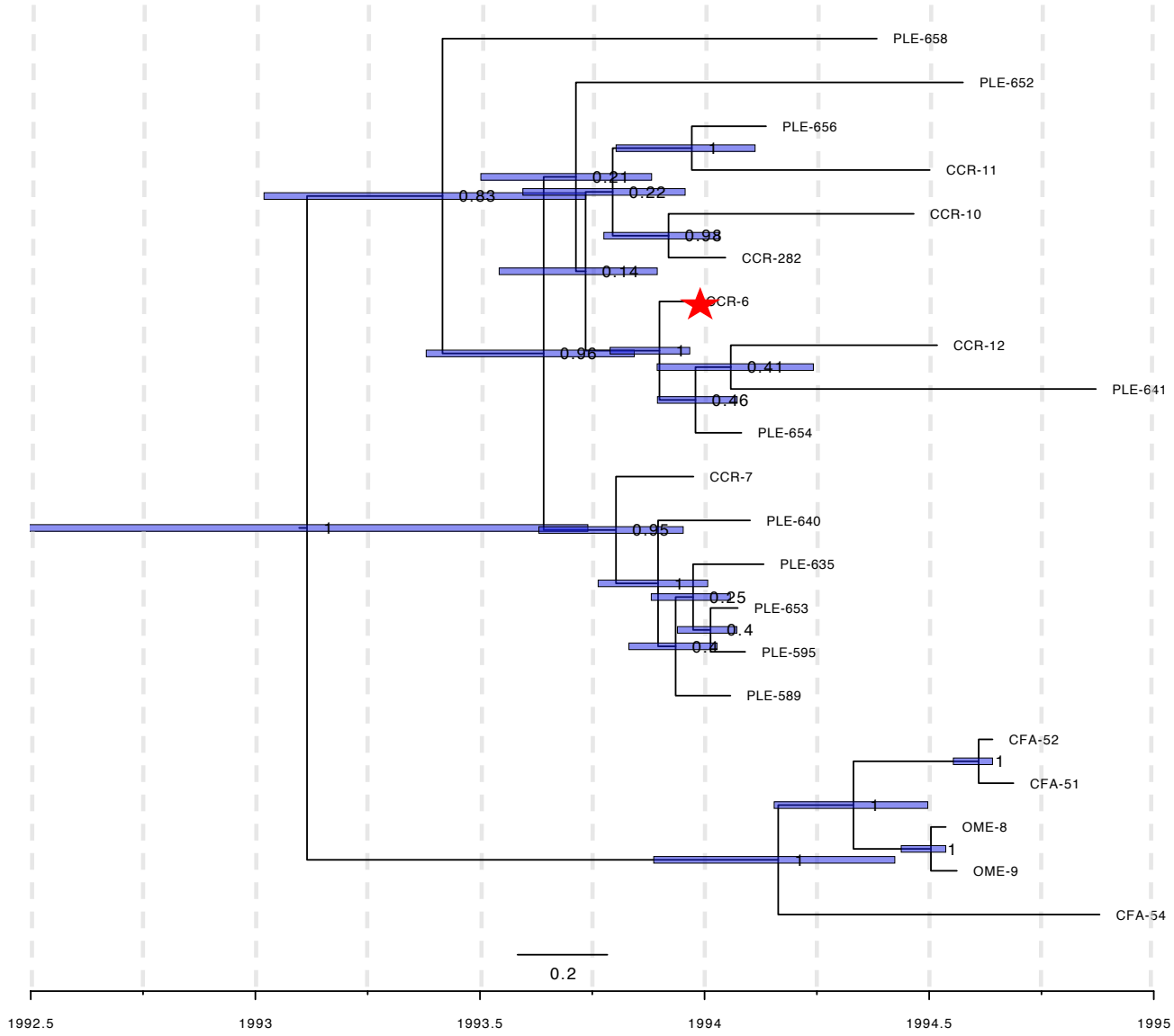

PLE-658
PLE-652
PLE-656
1
CCR-11
0.21
0.83
0.22
CCR-10
0.98
CCR-282
0.14
CCR-6
CCR-12
0.96
1
0.41
PLE-641
0.46
PLE-654
CCR-7
1
PLE-640
0.95
PLE-635
1
0.25
PLE-653
0.4
0.4
PLE-595
PLE-589
CFA-52
1
CFA-51
1
OME-8
1
1
OME-9
CFA-54
0.2
1992.5
1993
1993.5
1994
1994.5
1995

Figure 2.3. Time calibrated maximum clade credibility tree reconstructed from near whole CDV genomes sampled from the Serengeti ecological region in 1993/1994. Purple bars indicate the 95% HPD interval for node age estimates. Nodes are annotated with their posterior probability. The timing of the non-canid index case is indicated with a red star.

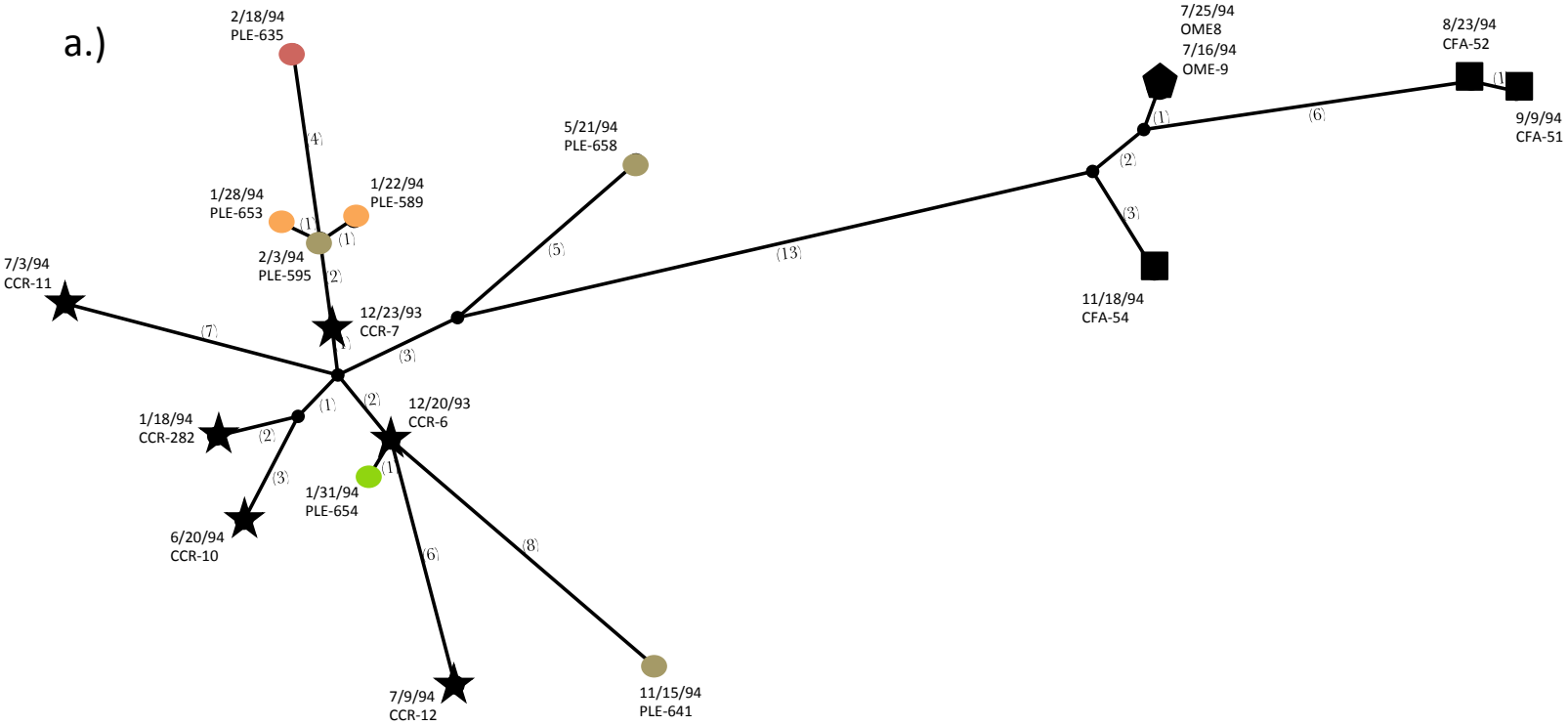

a.)
2/18/94
PLE-635
(4)
1/28/94
PLE-653
(1)
1/22/94
PLE-589
(1)
2/3/94
PLE-595
(2)
5/21/94
PLE-658
(5)
7/25/94
OME8
7/16/94
OME-9
(1)
(2)
(3)
11/18/94
CFA-54
(13)
8/23/94
CFA-52
(1)
9/9/94
CFA-51
(6)
7/3/94
CCR-11
(7)
12/23/93
CCR-7
(3)
(1)
(2)
12/20/93
CCR-6
1/18/94
CCR-282
(2)
(3)
6/20/94
CCR-10
1/31/94
PLE-654
(1)
(6)
7/9/94
CCR-12
(8)
11/15/94
PLE-641
b.)
N
P/V/C
M
F
H
L
*
*
*

Figure 2.4. a). Haplotype network by TCS algorithm of near whole genome CDV sequences collected during the 1993/1994 outbreak. Symbol shape indicates host species: square = domestic dog, pentagon = bat-eared fox, star = spotted hyena, circle = African lion. Circle symbol color indicates pride affiliation: orange = Campsites pride, red = Transects, brown = unknown, green = Masai, blue = Simba Numbers. Two bat-eared foxes have identical sequence, all other symbols represent a single sequence. b). Illustration of position of nucleotide substitutions on CDV genome. Grey boxes represent open reading frames of the CDV N, P/V/C, M, F, H, and L genes. The connecting lines indicate intergenic or untranslated regions. Vertical lines indicate the position of nucleotide substitutions between canid and non-canid strains; red=nonsynonymous, black=synonymous. Asterisks indicate the nucleotide substitutions that occur between canid species and all non-canid species except for PLE-658.

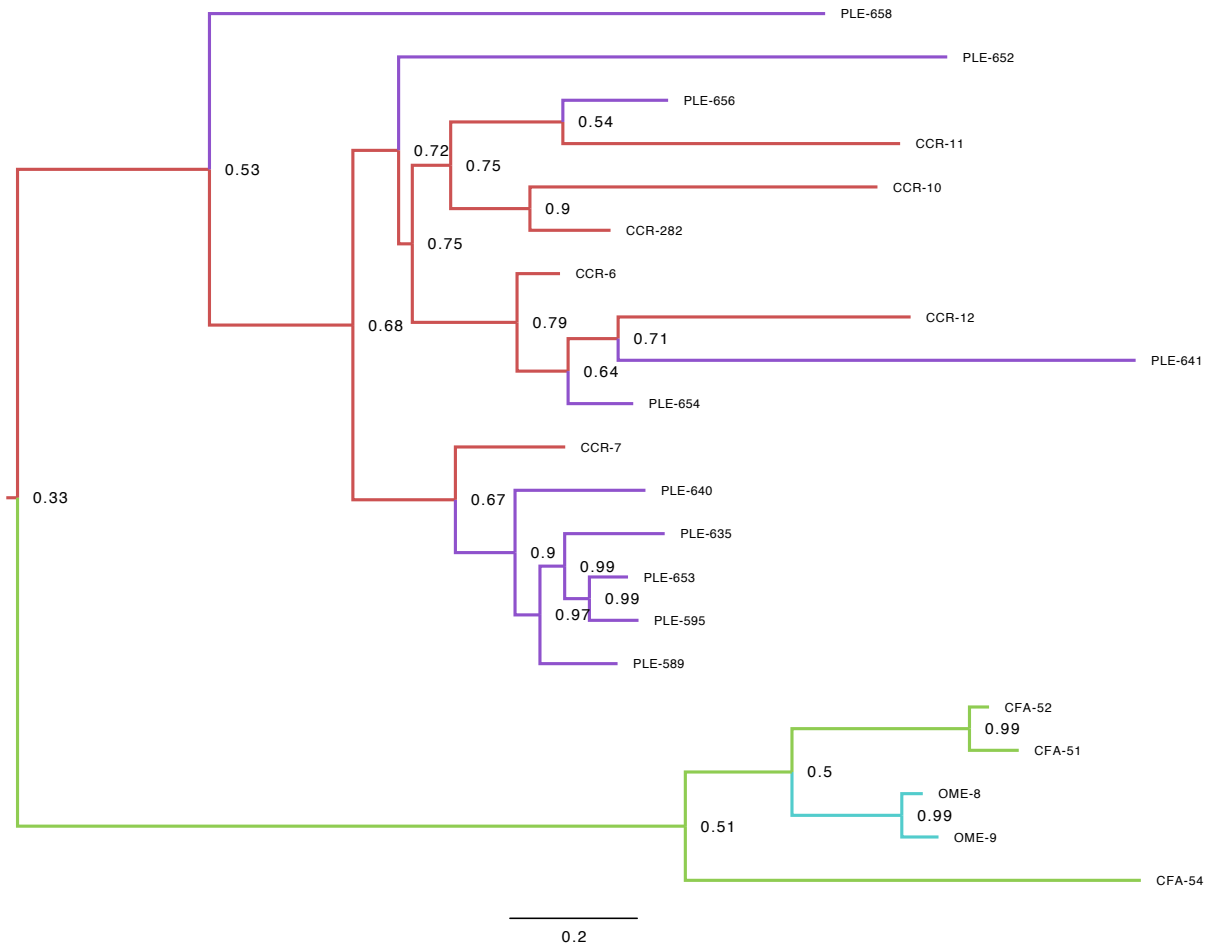

PLE-658
PLE-652
PLE-656
0.54
CCR-11
0.72
0.75
0.53
CCR-10
0.9
CCR-282
0.75
CCR-6
0.68
0.79
CCR-12
0.71
PLE-641
0.64
PLE-654
CCR-7
0.33
0.67
PLE-640
PLE-635
0.9
0.99
PLE-653
0.99
0.97
PLE-595
PLE-589
CFA-52
0.99
CFA-51
0.5
OME-8
0.99
0.51
OME-9
CFA-54
0.2

Figure 2.5. Discrete trait analysis in maximum clade credibility tree framework under asymmetric model of host species transitions. Host state is indicated by branch color; red = spotted hyena, blue = bat-eared fox, purple= African lion, green = domestic dog. Nodes are annotated with posterior probability of the host state of the common ancestors of branches indicated by color at that node.

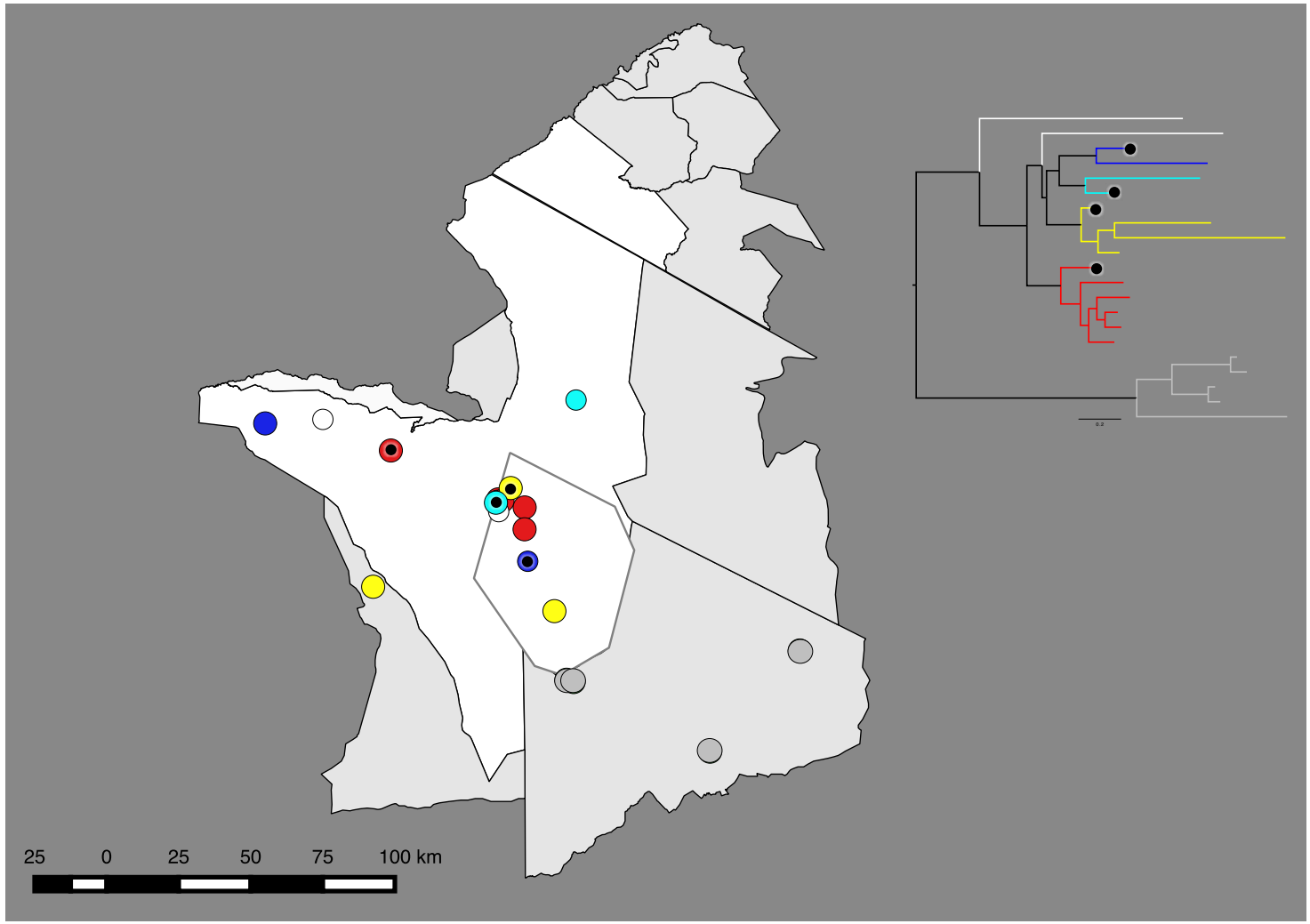

25 0 25 50 75 100 km

Figure 2.6. Star-like spatial diffusion of the non-canid CDV lineage during the 1993/1994 Serengeti outbreak (canid samples are shown in grey for reference). Sample color coding distinguishes phylogenetic clusters, or transmission groups, indicated in the inset maximum clade credibility tree. A black dot indicates the earliest sample collected in each transmission group. In 3 of 4 non-canid CDV transmission groups the earliest sample was located in or near Seronera, while sequences sampled later in the outbreak are found toward the periphery of SNP. The red clade is an exception to the pattern, representing a cluster which appears to originate in the western corridor and spread to the Seronera area early on in the outbreak.

CHAPTER 3

MULTI-SCALE PHYLOGENETIC ANALYSIS OF CANINE DISTEMPER VIRUS REVEALS RECURRENT INFECTION IN SERENGETI CARNIVORES FUELED BY LOCAL PERSISTENCE

ABSTRACT

Designing effective disease control strategies is contingent upon understanding where and at what scale pathogens persist on the landscape relative to vulnerable host populations. Clinical *Canine distemper virus* (CDV) infection was first documented in African lions and other wild carnivores in the Serengeti Ecological Region (SER) in 1994. In the years following, Serengeti carnivore populations (e.g. domestic dog, African lion, African wild dog, spotted hyena, and jackal species) have experienced recurrent CDV exposure and/or clinical infection. Here we use CDV serological and genetic data collected in the SER over a 20-year period to examine whether recurrent CDV exposure in Serengeti carnivores resulted from local CDV persistence or repeated introduction from an external reservoir. To explore where and at what scale CDV is maintained relative to Serengeti, we assessed epidemiological connectivity between SER and populations at the regional and continental scales. We find that CDV was continuously present in SER over the course of the study, apart from 2002 when CDV was not detected by serology or antigen screening in any carnivore population. Our phylodynamic analysis suggests that CDV was introduced in the Serengeti carnivore community twice from a common origin. The first CDV

incursion sparked the only known outbreak of clinical distemper disease in Serengeti lions in 1994 and subsequently went extinct. A second lineage was introduced as early as 1997 and persisted until at least 2011. Together these results suggest that apparent CDV re-emergence in Serengeti resulted both from repeated introduction, during its establishment, and local persistence thereafter. We did not find evidence of epidemiological connectivity between SER and populations at the regional or continental scale during the study period, indicating that the scale of persistence in SER did not include the populations sampled at this time. These results yield important insights regarding the invasion and establishment of CDV in a naïve multi-host community, and improve our understanding of the spatial scale of CDV persistence in SER.

INTRODUCTION

Canine distemper virus (CDV) is a highly infectious, multi-host disease and considered one of the most important disease threats to wild carnivores today (Deem, Spelman, Yates, & Montali, 2009). The frequency of CDV emergence in species of conservation concern is on the rise in recent years (e.g. Amur leopard (Sulikhan et al., 2018), brown hyena (Loots, Mitchell, Dalton, Kotzé, & Venter, 2016)). Understanding where and at what scale CDV is maintained on the landscape is critical for protecting vulnerable populations.

CDV exhibits epidemic dynamics characterized by peaks in infection separated by population troughs during which levels of infection can be extremely low. As a result, the critical community size (CCS), or minimum host population necessary to maintain a chain of CDV transmission through the inter-epidemic period most of the time is expected to be very large. The apparent persistence of CDV in small, low density wild carnivore populations has led researchers to hypothesize that CDV is maintained at very large, regional scales and/or by a "meta-reservoir" comprised of epidemiologically connected subpopulations of susceptible hosts (Almberg, Cross, & Smith, 2010; Prager et al., 2012; Viana et al., 2015).

Canine distemper virus emerged as a conservation threat to wild carnivores in the Serengeti Ecological Region (SER) in 1994, when an explosive, lethal outbreak coincided with clinical signs and mortality reported in African lions, spotted hyenas, bat-eared foxes, and domestic dogs (Roelke-Parker et al., 1996). In the years following, CDV clinical infection and/or exposure has been

reported multiple times in domestic and wild carnivores in the SER, including African lions and domestic dogs (Munson et al., 2008; Viana et al., 2014), African wild dogs (Goller et al., 2010), spotted hyenas and a golden jackal pup (Harrison et al., 2004; Nikolin et al., 2017). The recurrent incidence of CDV in Serengeti carnivores prompted us to ask whether CDV persisted in SER over time, and if so, at what spatial scale was it maintained.

CDV is a single stranded, negative-sense RNA virus. CDV populations are expected to generate new sequence variation in real-time, as they are being affected by environmental changes, owing to rapid mutation rates, large effective population sizes, and fast generation times (Pybus & Rambaut, 2009). This evolutionary characteristic of CDV allowed us to test predictions regarding CDV persistence within SER, and between SER and other populations in Africa, with phylogenetic analysis of whole and partial viral genome sequences.

Figure 3.1 illustrates how viral phylogenetics might be used to address whether a virus persists over time in a population of interest. In the example, virus from a population is sampled at multiple time points corresponding to peaks in infection, i.e. disease outbreaks. If the virus persisted in this population through inter-epidemic periods, we expect that all sequences would trace to a single introduction, i.e. cluster together regardless of time of sampling (see Fig. 3.1a.). Alternatively, if multiple, independent introductions explain recurrent infection in the population, we expected genetic independence of samples from different outbreaks, with shared ancestry in the distant past (Fig. 3.1b.).

Figure 3.2 illustrates how viral phylogenetics might also be used to understand where and at what scale a pathogen is maintained on the landscape with respect to a population of interest. Conceivable meta-population scenarios that could support pathogen maintenance at large spatial scales and the phylogenetic predictions that accompany them are presented (Fig. 3.2). In the first example (Fig. 3.2a.), a single subpopulation meets the CCS and repeatedly spreads infection to smaller populations that do not. This scenario is comparable to the "cities and villages" paradigm that describes prevaccination-era Measles virus dynamics (Grenfell & Bolker, 1998). In Figure 3.2b., the pathogen is maintained in each subpopulation at a local scale following introduction from a common source. Finally, in Fig. 3.2c., a collection of populations of susceptible hosts maintains the pathogen even if all subpopulations individually do not meet the CCS.

In this study, we used serological and genetic data collected over a 20-year period to investigate the pattern and spatial scale of CDV persistence with respect to the Serengeti carnivore community. Our objectives were to 1) to explain recurrent CDV infection in Serengeti carnivores by investigating CDV persistence at a local scale, and 2) to characterize the spatial extent of CDV maintenance by assessing epidemiological connectivity between SER and populations in Africa at the regional and continental scales.

MATERIALS AND METHODS

Study area

The study was carried out in the Serengeti Ecological Region (SER) in northwestern Tanzania, East Africa. The SER comprises Serengeti National Park (SNP, Tanzania) and Maasai Mara National Reserve (MMNR, Kenya), contiguous protected areas (Ikorongo Game Reserve, Grumeti Game Reserve, Maswa Game Reserve, Ngorongoro Conservation Area, Loliondo Game Control Area), and human settlements in districts west of SNP (Maswa, Tarime, Musoma, Bunda, Bariadi, Magu, Meatu). Domestic dogs are prevalent in villages in both the western and eastern districts bordering SNP, but prohibited and found only rarely inside national park boundaries. The density, distribution and movement patterns of domestic dogs around the SNP are linked to humans. Dense agro-pastoralist communities dominate districts bordering the west side of SNP where domestic dog density can be found in excess of $11/km^2$ (Lembo et al., 2008). The large eastern district is more sparsely inhabited by pastoralist Maasai communities where domestic dog density is less than half that of populations on the west at $<5/km^2$ (Lembo et al., 2008). Craft et al. (2017) report higher species richness and abundance of wild carnivores in sympatry with domestic dog populations in the east relative to the west (Craft et al., 2017).

Study specimens

In the SER, domestic dog pups (aged 4-12 months) were sampled annually for serology as part of a long-term disease surveillance program from 1992-2012 (with a one-year gap in sampling in 1995). Sampling design and protocol are described elsewhere (Kaare et al., 2009; Viana et al., 2015). Sampled dogs were

themselves unvaccinated, however they mostly came from dog populations included in a mass vaccination campaign against CDV, rabies and Canine parvovirus implemented from 1996 throughout the study period.

Wild carnivore tissue samples were collected opportunistically in SER from road kill and other mortalities.

Tissues from jackals and domestic dogs were collected during a putative CDV outbreak in 2000 from a ranch 413 km north of Serengeti in northwestern Kenya. Histopathological changes in brains of two jackals sampled supported CDV clinical infection (Prager et al., 2012).

CDV presence/absence

To evaluate CDV persistence over time in the SER carnivore community we looked annually for evidence of recent exposure to CDV in domestic dogs and wild carnivores. In domestic dog populations, evidence of recent exposure, or CDV presence, in each year was determined if ≥1 dog pup (aged 4-12 months) was seropositive (with a titer cutoff of ≥32) and/or if CDV antigen was detected by RT-PCR in a dog of any age. Pups less than 4 months old were excluded because antibodies detected in this age class may be maternally derived. Owners reported dog ages. Evidence of recent exposure to CDV in wild Serengeti carnivores was determined if CDV was detected in ≥1 wild carnivore by RT-PCR in any species in each year.

Serology

Serological assay methods carried out for this study are described elsewhere (Viana et al., 2015). Briefly, virus neutralization assays were carried

out in three labs, Intervet UK, Animal Health Diagnostic Center at Cornell (New York, USA) and University of Glasgow (UK) using similar protocols and viral strains. Serial dilutions of sera were made and co-incubated with a target dose of CDV virus (Onderstepoort strain). Indicator cells were added and further incubated. Infectivity was determined by the presence or absence of typical CDV cytopathology and the antibody titers were recorded as the reciprocal of the end-point dilution. For this study, we used a cutoff value of a 1:32 dilution to define seropositivity for CDV.

RT-qPCR

A reverse transcription quantitative PCR (RT-qPCR) assay was used to screen CDV-suspect specimens for viral RNA (Path-ID^{TM} Multiplex One-Step RT-PCR Kit). A domestic dog was considered CDV-suspect if: 1) it had an especially high CDV titer (e.g. ≥500), or 2) it was sampled in the same village at the same time as a pup that had a positive titer. Most carnivore tissues from necropsy were screened whether disease was implicated as a cause of death or not because CDV can cause aberrant behavior leading to mortality by trauma, e.g. road kill.

The RT-qPCR reaction conditions were as follows: 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C, and 60 sec at 60°C. Each reaction used 2-ul total RNA extract in a 25-ul volume reaction. Oligonucleotide concentrations were used at concentrations as recommended for the Path-ID kit; 400 nM forward primer, 400 nM reverse primer and 250 nM probe.

Whole genome sequencing and sequence datasets

Near whole genome sequences (14,619-15,547 bp) from 7 wild carnivores sampled in SER and Laikipia District, Kenya were generated using methods described in Chapter 2.

Partial and complete CDV genome sequences were retrieved from NCBI (http://www.ncbi.nlm.nih.gov) to evaluate epidemiological connectivity between SER and populations at the regional and continental scales in Africa. The final whole CDV genome sequence dataset (15,547 bp) included samples from wild carnivores in East Africa (this study, n=7), wild and domestic carnivores from Serengeti 1993-1994 (chapter 2, n=20), and wild carnivores from Serengeti (n=1), Ethiopia (n=1), and South Africa (n=2). A partial H genome sequence dataset (404 bp) included samples from wild carnivores in East Africa (this study, n=7), wild and domestic carnivores from Serengeti 1993-1994 (chapter 2 of this dissertation, n=20), spotted hyenas sampled in Serengeti from 1993-2004 (n=6), domestic and wild carnivores from South Africa (n=4), and Ethiopia (n=1). Sequence locations are shown in Figure 3.3.

Phylogenetic reconstruction of CDV whole genome sequences over time

To estimate the evolutionary rate, divergence times, and phylogenetic relationships between time-stamped CDV WGS sequences we used a Bayesian Markov Chain Monte Carlo approach implemented in BEAST software package version 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012) with the BEAGLE library (Ayres et al., 2012). Sequence alignments were performed using the MUSCLE algorithm in MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Edgar, 2004; Kumar, Stecher, & Tamura, 2016).

The best fit nucleotide substitution model was estimated by the Akaike Information Criterion (AIC), Akaike Information Criterion corrected for small sample sizes (AICc) (Hurvich and Tsai 1989), Bayesian Information Criterion (BIC), and Decision Theory Criterion (DT) in jModelTest (Darriba D, Taboada GL, Doallo R, & Posada D, 2012; Guindon S & Gascuel O, 2003) and selected by consensus (GTR + Γ). Prior to analysis, we explored these sequence data for evidence of recombination using RDP4 (Lole et al., 1999; Martin & Rybicki, 2000). Additionally, the program TempEst was used to qualitatively explore whether the sequence data contained sufficient temporal signal for this analysis which relies on the accumulation of mutations between heterochronously sampled sequences (Rambaut, Lam, Carvalho, & Pybus, 2016). Preliminary analyses indicated that the best-fit clock model of sequence evolution was the strict clock (over a relaxed lognormal clock model (BF = 239.51, Table 3.1)). The best-fit demographic tree prior of three that were tested was the coalescent: exponential growth (See Table 1 for details).

Three independent MCMC chains were run for 100M steps, of which 10M were discarded as burn-in. Tree and parameter files were logged every 10,000 steps. Traces were checked in Tracer for convergence, i.e. ESS>200. The program TreeAnnotator from the BEAST package was used to combine and annotate trees and the resulting maximum clade credibility tree was visualized in FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

Phylogenetic reconstruction of partial CDV H gene sequences

Sequence alignments were performed as above for the partial CDV H gene (n=36, 405 bp). Sequence alignments were tested for evidence of positive selection at individual sites in DataMonkey (Delport, Poon, Frost, & Kosakovsky Pond, 2010; Pond, Frost, & Muse, 2005) using SLAC, FEL, REL, and MEME methods (Kosakovsky Pond & Frost, 2005; Murrell et al., 2012). Amino acid site 549 on the partial H gene was found to be under positive selection consistent with previous studies (McCarthy, Shaw, & Goodman, 2007; Nikolin et al., 2017)) and was removed from the alignment. The best fit nucleotide substitution models estimated using the Akaike Information Criterion corrected for small sample sizes (AICc) (Hurvich and Tsai 1989) in jModelTest (Darriba D et al., 2012; Guindon S & Gascuel O, 2003) were HKY + Γ, for partial H gene datasets. Phylogenetic relationships were reconstructed using the Maximum Likelihood method in MEGA (Kumar et al., 2016).

RESULTS

Canine distemper virus presence in domestic dogs and wild carnivores over time

Annual CDV seroprevalence in domestic dogs aged 4-12 months from two populations in SER (west and east) is shown in Figure 3.4. for the period between 1992 and 2012, with the exception of 1995 during which no surveillance was conducted. CDV was detected in at least one SER dog population in every year of sampling except 2002, as determined by seroconversion of pups aged 4-12 months. Peaks in seroprevalence in the west and east in 1993 and 1994,

respectively correspond with an extensive CDV outbreak that affected domestic and wild carnivores in the ecosystem.

Annual CDV detection in domestic and wild carnivores of the SER and other populations in East Africa was determined by RT-PCR in this study and others (Goller et al., 2010; Nikolin et al., 2017; van de Bildt et al., 2002) between 1992 and 2012 is shown in Table 3.2. CDV was detected only sporadically in domestic dogs by this method. In 1992-1994, 2000, and 2010, these detections in dogs were associated with clinical outbreaks. Detections in domestic dogs in 2005-2006 were made in individuals /populations without overt or reported clinical signs at the time. CDV was detected frequently but not continuously in wild carnivores over the study period. Clinical CDV in lions and hyenas was only observed in 1993-1994.

Canine distemper virus history in Serengeti and epidemiological connectivity

Bayesian coalescent analysis under a strict molecular clock model of evolution (n=36 sequences) predicted the time to most recent common ancestor (tMRCA) of all East African samples was in 1989, or 26.77 years before 2017 (95% HPD: 26.12, 29.16) (Fig. 3.5). A minimum of two introductions of CDV into Serengeti carnivores followed, as indicated by distinct clades in the phylogeny. Sequences comprising the first clade were sampled from domestic and wild carnivores in 1993 and 1994 during an explosive, highly lethal outbreak affecting domestic dogs, African lions, spotted hyenas, and bat-eared foxes (Roelke-Parker et al., 1996). Sequences comprising the second Serengeti clade were sampled over a 5-year period in 2006, 2007, and 2011 from wild canid

species. The most recent common ancestor of these sequences was in 2006, or 10.87 years before 2017 (95% HPD: 10.47, 11.32). Genetic similarity between sequences in this clade indicated that CDV persisted from at least 2006-2011.

In 2000, multiple carnivore populations in East Africa experienced clinical CDV outbreaks or peaks in CDV exposure including wild and domestic carnivores in SER and Laikipia, Kenya, and African wild dogs in Mkomazi Game Reserve, Tanzania (Prager et al., 2012; van de Bildt et al., 2002; Viana et al., 2015). Sequences from SER and Laikipia have an estimated MRCA in 1990 (26.9 years prior to 2017 (95% HPD: 25.95, 29.87), 11 years prior to the coincident peaks in infection (Fig. 3.5), indicating that these populations were not epidemiologically connected during the study period. Informative CDV sequence data from Mkomazi Game Reserve was not available for this analysis. At the continental scale, CDV sequences from the Horn of Africa and southern Africa were distantly related to SER, with estimated tMRCA with SER CDV in 1943 (70.63 years prior to 2017 (95% HPD: 59.42, 82.19) and 1947 (73.85 years before 2017(95% HPD: 60.85, 88.42), respectively.

To investigate the timescale of CDV persistence in Serengeti beyond the years represented by sequences in the whole genome dataset, we performed a phylogenetic analysis with publically available partial H gene sequences from Serengeti spotted hyenas sampled between 1993 and 2004 and sequences from this study (Fig. 3.6). CDV sequences from hyenas sampled in 1993, 1994, and 1997 fell with the 1993/4 Serengeti cluster, indicating that the lineage circulating in the 1993/4 outbreak continued to circulate in Serengeti for at least 2 years after

the last confirmed case of CDV associated with the outbreak in 1995. Hyena CDV sequences from 1997, 1999, and 2004 fell in the 2006-2011 cluster of wild Serengeti canids, indicating that this lineage persisted for at least 14 years, between 1997-2011.

DISCUSSION

Exposure and/or clinical infection have been documented in Serengeti carnivores repeatedly since CDV was first documented there in 1993. In this study, we found evidence that recurrent CDV infection over the 20-year study period resulted both from repeated CDV introduction and from local CDV persistence. Molecular dating suggested that CDV was introduced to East Africa in 1989. An initial CDV incursion in Serengeti occurred in 1993 causing a brief but extensive outbreak in wild and domestic carnivores. This lineage apparently faded out by at least 1997and was replaced by a second invasion of CDV that was subsequently maintained until at least 2011. Beyond sharing a common origin in 1990, we did not find evidence that coincident CDV epizootics in Serengeti and another population (Laikipia) in East Africa were synchronized by migration, i.e. these regional populations were not epidemiologically connected at the time of sampling. CDV lineages from the Horn of Africa and southern Africa were distantly related and not implicated in Serengeti CDV dynamics.

Epidemiological history of CDV in SER - introductions and persistence

According to our analysis of whole genome sequences collected after 1993, the arrival of CDV in East Africa occurred around 1989. This finding is consistent with a recent phylodynamic analysis of full CDV H genes which

reported that CDV dispersal from the United States to Tanzania occurred in 1988 (Ke et al., 2015). However, this estimate is not consistent with CDV serological data which suggests that lion and hyena populations of the SER were exposed to CDV in 1981 (Harrison et al., 2004; Packer et al., 1999) and anecdotal reports of neurological signs in Serengeti lions (Schaller, 1972) and recurrent CDV outbreaks among domestic dogs in Nairobi and Mombasa, Kenya prior to 1994 (Alexander & Appel, 1994).

One possible explanation for the discrepancy between the genetic and serologic data is that some lineages were not sampled either because of small sample size or because lineages went extinct prior to the sampling period. Genetic bottlenecks that occur during dramatic population size fluctuations between epidemic cycles can confound divergence time estimates for pathogens with epidemic dynamics (Archie, Luikart, & Ezenwa, 2009). Thus, our estimate may have dated the last CDV population expansion in East Africa. Alternatively, it is possible that the CDV neutralizing antibodies reported in lions and hyenas sampled in 1981 were acquired by exposure to Rinderpest virus (RPV), a closely related Morbillivirus (Viana et al., 2015). RPV antibodies have been shown to neutralize CDV in serological tests (Imagawa, Goret, & Adams, 1960; Jones, Tenorio, Gorham, & Yilma, 1997). Rinderpest virus circulated in common prey species of hyenas and lions in Serengeti. The last known outbreak of RPV in SER was recorded in 1982 (Anderson et al., 1990), conspicuously timed with the last CDV outbreak in the SER in 1981 inferred from the CDV serological data (Alexander et al., 1995; Packer et al., 1999). Thus, the result of our molecular

dating is plausible and if accurate, this study characterizes the behavior of CDV during its invasion and establishment in SER.

Our data suggest that the initial incursion of CDV into SER occurred in the early-mid 1990s through two separate introductions. Bayesian analysis of time-stamped whole genome sequence data fit phylogenetic predictions for both the "repeated introduction" and "persistence" hypotheses to explain CDV recurrence (illustrated in Fig. 3.1), as indicated by two distinct Serengeti clusters and the clustering of sequences from 2006-2011, respectively (Figure 3.5). The first introduction of CDV in SER catalyzed an explosive epizootic in domestic and wild carnivores in 1993 and 1994. This lineage subsequently went extinct, and was replaced by a second invasion that has since persisted. Clustering of partial H gene sequences in our phylogenetic analysis (Figure 3.6) suggests that this second wave arrived as early as 1997, indicating that this lineage has persisted in SER from at least 1997-2011. This finding is supported by our detection of CDV, albeit it at low incidence, in domestic dogs and wild carnivores in the SER continuously over the study period (apart from 2002).

One limitation of our phylogenetic analysis of partial H gene sequences is that this sequence is located in a region of the CDV genome that is associated with host cell receptor recognition (SLAM), and as such may be under selective pressure (Nikolin et al., 2017). Our selection analysis identified one site under positive selection, at residue 549. We subsequently removed the nucleotide position coding this substitution from the analysis, though we cannot rule out the possibility that the topology of the partial H gene tree was somehow affected by

selection. However, Nikolin et al. (2017) made the argument that mutations in this sequence direct canid tropism. If this is true then we would expect the hyena sequences sampled between 1996-2004 to cluster with the non-canid sequences sampled in 1993-1994. That instead some hyenas cluster with canids sampled between 2006-2011 and others with non-canids in 1993/4, suggests that neutral evolutionary processes may have a stronger influence on these phylogenetic results.

Epidemiological connectivity and the spatial scale of persistence

In the metapopulation framework, subpopulations are connected by migration. In the epidemiological context, migration between subpopulations can coordinate peaks in infection. We analyzed sequences from two putative subpopulations in East Africa, SER and a population 413-km to the north in Laikipia, Kenya, that were sampled during a coincident peak in CDV infection. Sequences sampled from these populations were genetically independent and could be traced back to a common source 11 years prior to the coincident peak in infection (Fig. 3.5), i.e. they were not epidemiologically connected. Thus, if a regional CDV meta-reservoir existed in East Africa, one and/or both of these populations were not constituents of it at that time.

At the continental scale, our results suggest that the SER was not epidemiologically connected to CDV populations in Ethiopia or South Africa. The time measured phylogeny indicates that sequences from these three populations are distantly related, sharing a common ancestor in 1943. The phylodynamic analysis of Ke et al. (2015), suggests that CDV spread from the

United States to Tanzania in 1988. If this is accurate, then the common ancestor of all lineages found in Africa today may have existed in the United States around the 1940s.

That SER was not epidemiologically connected to any subpopulations included in this study at the regional or continental scale, would suggest that CDV persistence might occur over a small spatial scale in the SER. A stochastic susceptible-exposed-infected-recovered simulation model of CDV in carnivores in Yellowstone National Park, U.S.A., revealed that CDV could be maintained at small spatial scales if more than one competent host were involved in transmission (Almberg et al., 2010). Indeed, there are 26 wild carnivore species in Serengeti, with clinical CDV documented in at least 9 of them, thus multihost transmission may be a plausible mechanism enabling persistence at a small spatial scale in SER. However, due to our sparse sampling, we cannot rule out the possibility that CDV is maintained at a regional scale and we failed to sample it.

A notable limitation of this study is the lack of CDV sequence data from the domestic dog populations in western SER, and further west and north along the shores of Lake Victoria. The Serengeti carnivore community comprises wildlife populations and two domestic dog populations on either side of Serengeti National Park. Surveillance methods and/or effort in this study differed significantly between these three carnivore subgroups (wildlife, west dogs, and east dogs), precluding our ability to draw conclusions about their relative roles in persistence. However, we can assume that a single detection reveals presence (although ‘no detection’ does not equate to absence). With this condition, our

results indicate that CDV persisted at least in the high-density domestic dog population in western SER during the study period, with the possible exception of 2002. An early serological investigation in SER implicated this dog population as the origin and possible reservoir of CDV seeding the 1993-1994 outbreak (Cleaveland et al., 2000). A recent model incorporating longitudinal serological data however concluded that the domestic dog populations could not solely be responsible for driving CDV dynamics in SER (Viana et al., 2015). The western domestic dog population is contiguous with an even denser domestic dog population along the shores of Lake Victoria (see inset map Fig. 3.3). If CDV persists in domestic dog populations in western Serengeti, it seems likely that the spatial extent of persistence would include this dense, extensive domestic dog population as well.

Conclusion

This study characterizes the behavior of CDV during its invasion and establishment in a naïve multi-host community. Although our results shed light on the spatial scale of persistence of CDV, a thorough exploration of the spatial extent of maintenance in this system would require sampling multiple putative subpopulations, preferably contemporaneously, and ideally at multiple time points. Given the unpredictable and epidemic nature of CDV, and the difficulty in locating and handling clinically infected carnivores, this pursuit is impracticable. Our understanding will be improved as more informative sequence data accumulates in the public domain.

LITERATURE CITED

Alexander, K. A., & Appel, M. J. G. (1994). African wild dogs (Lycaon pictus) endangered by a canine distemper epizootic among domestic dogs near the Masai Mara National Reserve, Kenya. *Journal of Wildlife Diseases*, *30*(4), 481–485. https://doi.org/10.7589/0090-3558-30.4.481

Alexander, K. A., Kat, P. W., Frank, L. G., Holekamp, K. E., Smale, L., House, C., & Appel, M. J. G. (1995). Evidence of canine distemper virus infection among free-ranging spotted hyenas (Crocuta crocuta) in the Masai Mara, Kenya. *Journal of Zoo and Wildlife Medicine*, *26*(2), 201–206. Retrieved from http://www.jstor.org/stable/20095463

Almberg, E. S., Cross, P. C., & Smith, D. W. (2010). Persistence of canine distemper virus in the Greater Yellowstone Ecosystem's carnivore community. *Ecological Applications*, *20*(7), 2058–2074. https://doi.org/10.1890/09-1225.1

Anderson, E. C., Jago, M., Mlengeya, T., Timms, C., Payne and, A., & Hirj, K. (1990). A serological survey of rinderpest antibody in wildlife and sheep and goats in Northern Tanzania. *Epidemiol. Infect*, *105*, 203–214. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2271795/pdf/epidinfect00022-0194.pdf

Archie, E., Luikart, G., & Ezenwa, V. (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology & Evolution*, *24*(1), 21–30. https://doi.org/10.1016/J.TREE.2008.08.008

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., … Suchard, M. A. (2012). BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Systematic Biology*, *61*(1), 170–173. https://doi.org/10.1093/sysbio/syr100

Cleaveland, S., Appel, M. G. J., Chalmers, W. S. K., Chillingworth, C., Kaare, M., & Dye, C. (2000). Serological and demographic evidence for domestic dogs as a source of canine distemper virus infection for Serengeti wildlife. *Veterinary Microbiology*, *72*, 217–227.

Craft, M. E., Vial, F., Miguel, E., Cleaveland, S., Ferdinands, A., & Packer, C. (2017). Interactions between domestic and wild carnivores around the greater Serengeti ecosystem. *Animal Conservation*, *20*(2), 193–204. https://doi.org/10.1111/acv.12305

Darriba D, Taboada GL, Doallo R, & Posada D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772.

Deem, S. L., Spelman, L. H., Yates, R. A., & Montali, R. J. (2009). CANINE DISTEMPER IN TERRESTRIAL CARNIVORES: A REVIEW. *http://dx.doi.org/10.1638/1042-7260(2000)031[0441:CDITCA]2.0.CO;2*. https://doi.org/10.1638/1042-7260(2000)031[0441:CDITCA]2.0.CO;2

Delport, W., Poon, A. F. Y., Frost, S. D. W., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, *26*(19), 2455–2457. https://doi.org/10.1093/bioinformatics/btq429

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian hylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–7. https://doi.org/10.1093/nar/gkh340

Goller, K. V., Fyumagwa, R. D., Nikolin, V., East, M. L., Kilewo, M., Speck, S., … Wibbelt, G. (2010). Fatal canine distemper infection in a pack of African wild dogs in the Serengeti ecosystem, Tanzania. *Veterinary Microbiology*. https://doi.org/10.1016/j.vetmic.2010.05.018

Grenfell, & Bolker. (1998). Cities and villages: infection hierarchies in a measles metapopulation. *Ecology Letters*, *1*(1), 63–70. https://doi.org/10.1046/j.1461-0248.1998.00016.x

Guindon S, & Gascuel O. (2003). A simple, fast and accurate method to estimate large phylo- genies by maximum-likelihood. *Systematic Biology*, *52*, 696–704.

Harrison, T. M., Mazet, J. K., Holekamp, K. E., Dubovi, E., Engh, A. L., Nelson, K., … Munson, L. (2004). ANTIBODIES TO CANINE AND FELINE VIRUSES IN SPOTTED HYENAS (CROCUTA CROCUTA) IN THE MASAI MARA NATIONAL RESERVE. *Journal of Wildlife Diseases Wildlife Disease Association*, *40*(1), 1–10.

Imagawa, D. T., Goret, P., & Adams, J. M. (1960). IMMUNOLOGICAL RELATIONSHIPS OF MEASLES, DISTEMPER, AND RINDERPEST

VIRUSES*. *Public Health Service Virology Virology Ann. N. Y. Acad. Sci. Am. J. Vet. Res. Cancer Inst. J. Nat. Cancer Inst*, *46487*(34), 157–609. Retrieved from http://www.pnas.org/content/46/8/1119.long

Jones, L., Tenorio, E., Gorham, J., & Yilma, T. (1997). Protective vaccination of ferrets against canine distemper with recombinant pox virus vaccines expressing the H or F genes of rinderpest virus. *American Journal of Veterinary Research*, *58*(6), 590–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9185963

Kaare, M., Lembo, T., Hampson, K., Ernest, E., Estes, A., Mentzel, C., & Cleaveland, S. (2009). Rabies control in rural Africa: Evaluating strategies for effective domestic dog vaccination. *Vaccine*. https://doi.org/10.1016/j.vaccine.2008.09.054

Ke, G.-M., Ho, C.-H., Chiang, M.-J., Sanno-Duanda, B., Chung, C.-S., Lin, M.-Y., … Chu, P.-Y. (2015). Phylodynamic analysis of the canine distemper virus hemagglutinin gene. *BMC Veterinary Research*, *11*(1), 164. https://doi.org/10.1186/s12917-015-0491-9

Kosakovsky Pond, S. L., & Frost, S. D. W. (2005). Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Molecular Biology and Evolution*, *22*(5), 1208–1222. https://doi.org/10.1093/molbev/msi105

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*, *33*(7), 1870–4.

Lembo, T., Hampson, K., Haydon, D. T., Craft, M., Dobson, A., Dushoff, J., … Cleaveland, S. (2008). Exploring reservoir dynamics: a case study of rabies in the Serengeti ecosystem. *Journal of Applied Ecology*, *45*(4), 1246–1257. https://doi.org/10.1111/j.1365-2664.2008.01468.x

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., … Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology*, *73*(1), 152–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9847317

Loots, A. K., Mitchell, E., Dalton, D. L., Kotzé, A., & Venter, E. H. (2016). Advances in canine distemper virus (CDV) pathogenesis research: a wildlife perspective. *Ournal of General Virology*. Retrieved from http://www.microbiologyresearch.org/docserver/fulltext/jgv/jgv_pap.000666.zip/jgv.0.000666.pdf?expires=1490811416&id=id&accname=guest&checksum=D3E1FA7FBFAD5C6F366F82329E9089D1

Martin, D., & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *BIOINFORMATICS APPLICATIONS NOTE*, *16*(6), 562–563. Retrieved from http://

McCarthy, A. J., Shaw, M.-A., & Goodman, S. J. (2007). Pathogen evolution and disease emergence in carnivores. *Proceedings. Biological Sciences / The Royal Society*, *274*(1629), 3165–74. https://doi.org/10.1098/rspb.2007.0884

Munson, L., Terio, K. A., Kock, R., Mlengeya, T., Roelke, M. E., Dubovi, E., … Packer, C. (2008). Climate Extremes Promote Fatal Co-Infections during

Canine Distemper Epidemics in African Lions. *PLoS ONE*, *3*(6), e2545. https://doi.org/10.1371/journal.pone.0002545

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genetics*, *8*(7), e1002764. https://doi.org/10.1371/journal.pgen.1002764

Nikolin, V. M., Olarte-Castillo, X. A., Osterrieder, N., Hofer, H., Dubovi, E., Mazzoni, C. J., … East, M. L. (2017). Canine distemper virus in the Serengeti ecosystem: molecular adaptation to different carnivore species. *Molecular Ecology*. https://doi.org/10.1111/mec.13902

Packer, C., Altizer, S., Appel, M., Brown, E., Martenson, J., O'Brien, S. J., … Lutz, H. (1999). Viruses of the Serengeti: patterns of infection and mortality in African lions. *Journal of Animal Ecology*, *68*(6), 1161–1178. https://doi.org/10.1046/j.1365-2656.1999.00360.x

Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, *21*(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Prager, K. C., Mazet, J. A. K., Dubovi, E. J., Frank, L. G., Munson, L., Wagner, A. P., & Woodroffe, R. (2012). Rabies Virus and Canine Distemper Virus in Wild and Domestic Carnivores in Northern Kenya: Are Domestic Dogs the Reservoir? *EcoHealth*, *9*(4), 483–498. https://doi.org/10.1007/s10393-013-0815-9

Pybus, O. G., & Rambaut, A. (2009). Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, *10*(8), 540–550. https://doi.org/10.1038/nrg2583

Rambaut, A., Lam, T. T., Carvalho, L. M., & Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution*, *2*(1). https://doi.org/10.1093/ve/vew007

Roelke-Parker, M. E., Munson, L., Packer, C., Kockil, R., CIeaveland, S., Carpenter, M., … G Appell, M. J. (1996). A canine distemper virus epidemic in Serengeti lions (Panthera leo). *Nature*, *379*, 441–445.

Schaller, G. (1972). *The Serengeti Lion: A Study of Predator-Prey Relations.* Chicago, Illinois: Chicago University Press.

Sulikhan, N. S., Gilbert, M., Yu Blidchenko, E., Naidenko, S. V, Ivanchuk, G. V, Yu Gorpenchenko, T., … Seimon, T. A. (2018). Canine Distemper Virus in a Wild Far Eastern Leopard Panthera pardus orientalis. *Journal of Wildlife Diseases Wildlife Disease Association*, *54*(1), 0–0. https://doi.org/10.7589/2017-03-065

van de Bildt, M. W. G., Kuiken, T., Visee, A. M., Lema, S., Fitzjohn, T. R., & Osterhaus, A. D. M. E. (2002). Distemper outbreak and its effect on African wild dog conservation. *Emerging Infectious Diseases*, *8*(2), 211–3. https://doi.org/10.3201/eid0802.010314

Viana, M., Cleaveland, S., Matthiopoulos, J., Halliday, J., Packer, C., Craft, M. E., … Orr, B. (2015). Dynamics of a morbillivirus at the domestic-wildlife

interface: Canine Dis- temper Virus in domestic dogs and lions. *PNAS*, *112*(5), 1464–1469.

Viana, M., Mancy, R., Biek, R., Cleaveland, S., Cross, P. C., Lloyd-Smith, J. O., & Haydon, D. T. (2014). Assembling evidence for identifying reservoirs of infection. *Trends in Ecology & Evolution*, *29*(5), 270–9. https://doi.org/10.1016/j.tree.2014.03.002

Table 3.1. BEAST model comparisons. Model with highest support in bold italics.

		Relaxed lognormal	Strict		
		Constant	Constant	Exponential	Skyline
Relaxed lognormal	Constant	—	-81.58	-360.5	-49.24
Strict	Constant	239.51	—	-278.92	32.34
	Exponential	*360.01*	*120.5*	—	*311.26*
	Skyline	47.98	-191.53	-312.03	—

Table 3.2. Annual detection by RT-PCR of CDV antigen in individuals sampled in East Africa between 1992 and 2012. PCR detections of CDV from previous studies in blue italics. Years in which CDV was detected in wildlife are shaded.

Year	Species detected[1]	Wild	Domestic	Location[2]
1992	CFA		1	SNP (east)[3]
1993	CCR, *CCR*	2, *1*		SNP
1994	CCR, CFA, OME, PLE, *CCR, CFA, OME, PLE*[a]	29	6, *1*	SNP, SNP (west & east), MMNR, *SNP*
1995	OME, CCR	2		SNP
1996	*CCR*[a]	*1*		*SNP*
1997	*CCR*[a]	*2*		*SNP*
1998				
1999	*CCR*[a]	*1*		*SNP*
2000	CFA, CME, *LPI*[b]			LKP, *MKM*
2001				
2002				
2003				
2004	*CCR*[a]	*1*		*SNP*
2005	CFA		1	SNP (west)
2006	CCR, CFA, CME, PLE	3	5	SNP, SNP (west & east)
2007	LPI, *LPI*[4]	11, *1*		LGCA
2008				
2009				
2010	CFA		2	AMU, ANP
2011	*CAU*[a]	*1*		*SNP*
2012	CFA		1	KTU
TOTALS				
GRAND TOTAL		**47, *8***	**16, *1***	

1 Abbreviations: CCR=Crocuta crocuta, CFA=Canis familiaris, CME=Canis mesomelas, LPI=Lycaon pictus, OME=Otocyon megalotis, PLE=Panthera leo. CAU=Canis aureus

2 Abbreviations: AMU=Arusha Municipality, ANP=Arusha National Park, KTU=Karatu, LGCA=Loliondo Game Control Area, LKP=Laikipia, Kenya, SNP=Serengeti National Park, MMNR=Masai Mara National Reserve, MKM=Mkomazi National Park

3 Cardinal directions refer to domestic dog population in which CDV was detected.

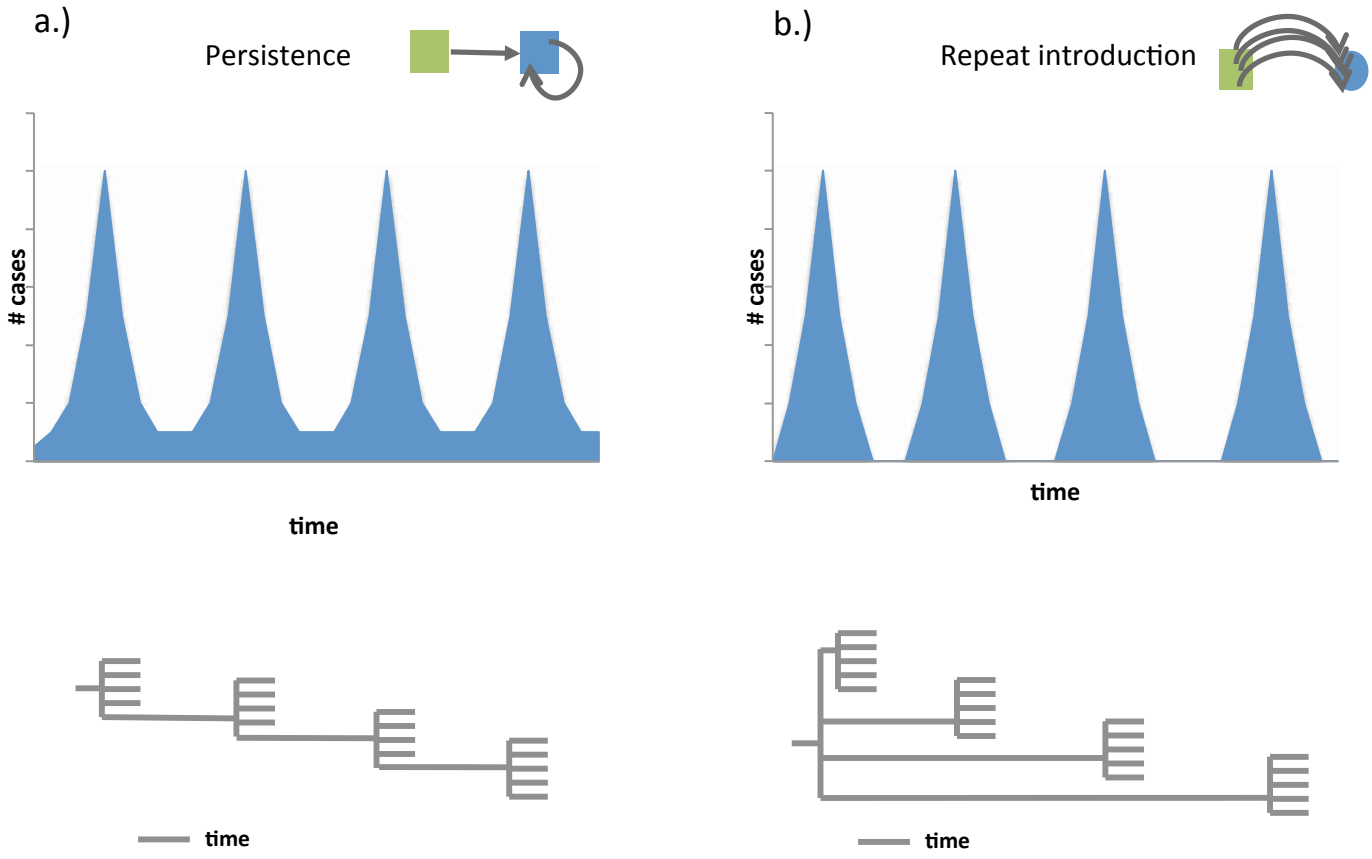
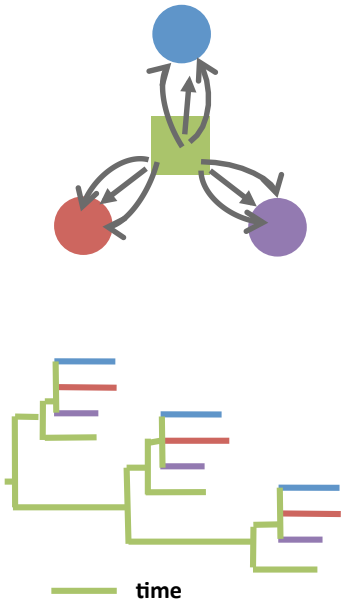

a.)
Persistence
cases
time
time
b.)
Repeat introduction
cases
time
time

Figure 3.1. Phylogenetic expectations of pathogen sequences in a population of interest that experiences periodic exposure due to a) local persistence of virus, or b) multiple, independent introductions from an external source/reservoir. Each hypothesis is graphically illustrated in the inset of the top panel, where green indicates the source and blue the population of interest. Squares indicate maintenance and circle, nonmaintenance populations. The top panels indicate incidence over time of an acute, immunizing pathogen with epidemic dynamics. The bottom panels illustrate phylogenetic expectations of sequences collected from the population during recurrent infection peaks. In population a) the pathogen persists through inter-epidemic periods of low incidence and all sequences sampled through time are expected to trace back to a single introduction. Population b) experiences local extinction during the inter-epidemic periods and subsequent infection results from repeated introductions. Sequences from distinct outbreaks are expected to be genetically distinct from one another, sharing a common ancestor in the distant past

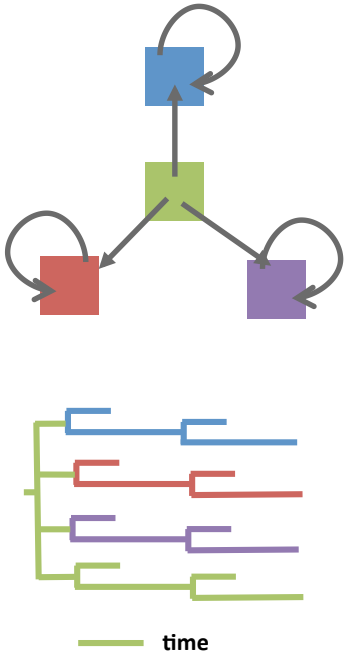
a.)

Regional persistence
single reservoir



b.)

Local persistence



c.)

Regional persistence
meta-reservoir

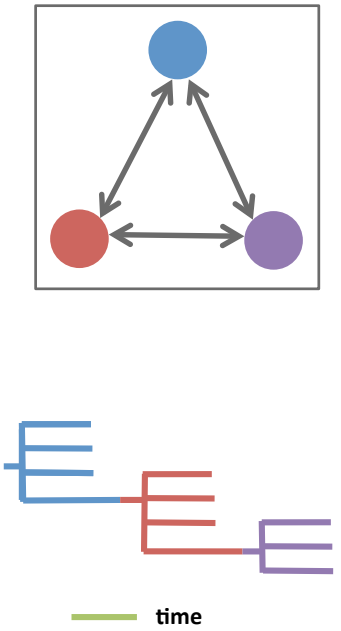


Figure 3.2. Examples of metapopulation systems of disease maintenance and associated phylogenetic predictions. Each hypothesis is graphically illustrated in the top panel, where squares indicate maintenance and circles, nonmaintenance populations. The bottom panels illustrate phylogenetic expectations of sequences collected from the metapopulation during infection peaks. In metapopulation a) a maintenance population transmits the pathogen during peaks in infection to nonmaintenance populations that are connected by migration resulting in genetic similarity of contemporaneously sampled sequences from distinct populations. In b) the pathogen persists independently in each subpopulation following introduction from a common source giving rise to genetically independent populations, i.e. population structure. In c) nonmaintenance subpopulations are epidemiologically connected and persistence in the metapopulation depends on migration between populations generating a continuous branching pattern and transient population structure.

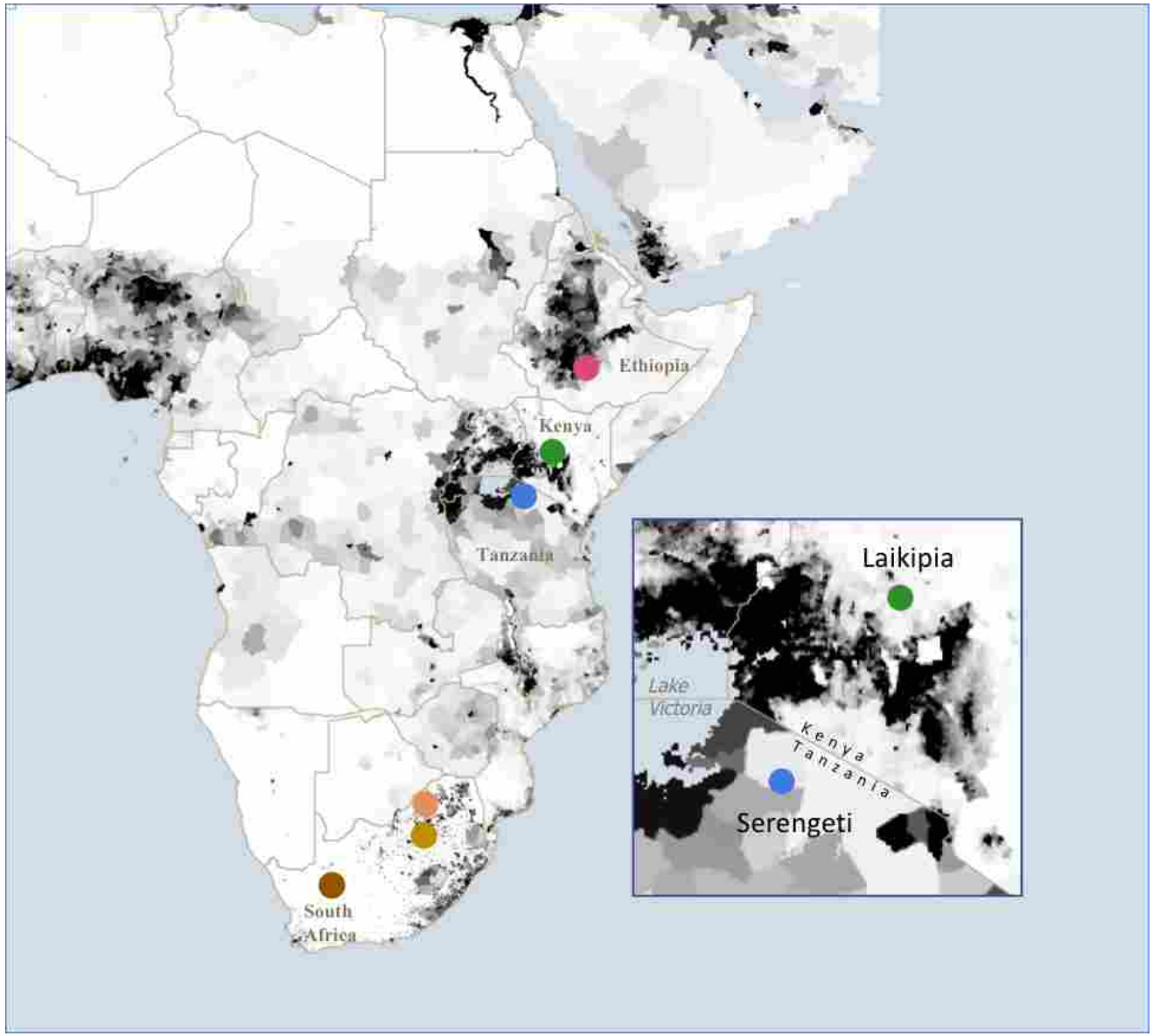

Ethiopia
Kenya
Tanzania
South Africa
Laikipia
Lake Victoria
Kenya
Tanzania
Serengeti

Figure 3.3. Map of Africa showing sample locations of *Canine distemper virus* sequences used in this study. Inset map details East African populations that were sampled during a coincident peak of CDV infection in 2000. Shading indicates human population density, people/square kilometer, from low (white) to high (black).

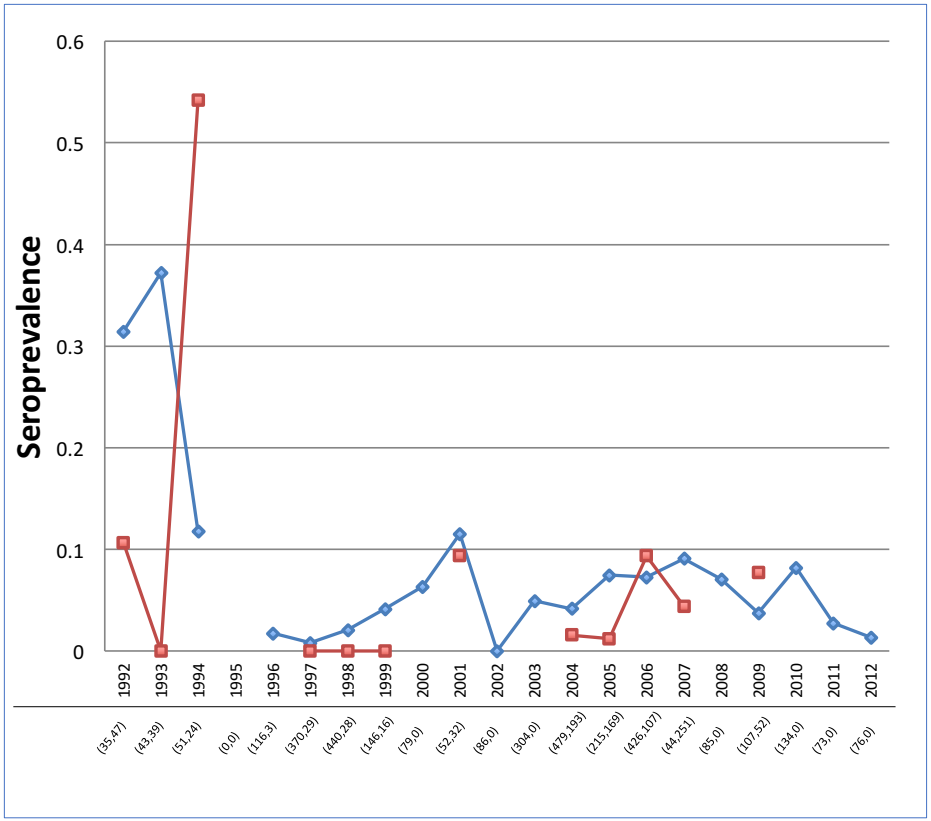


Figure 3.4. Annual CDV seroprevalence in unvaccinated domestic dogs aged 4-12 months from populations on the west side (blue diamonds) and east side (red squares) of Serengeti National Park. Numbers in parentheses indicate sample size in west and east, respectively.

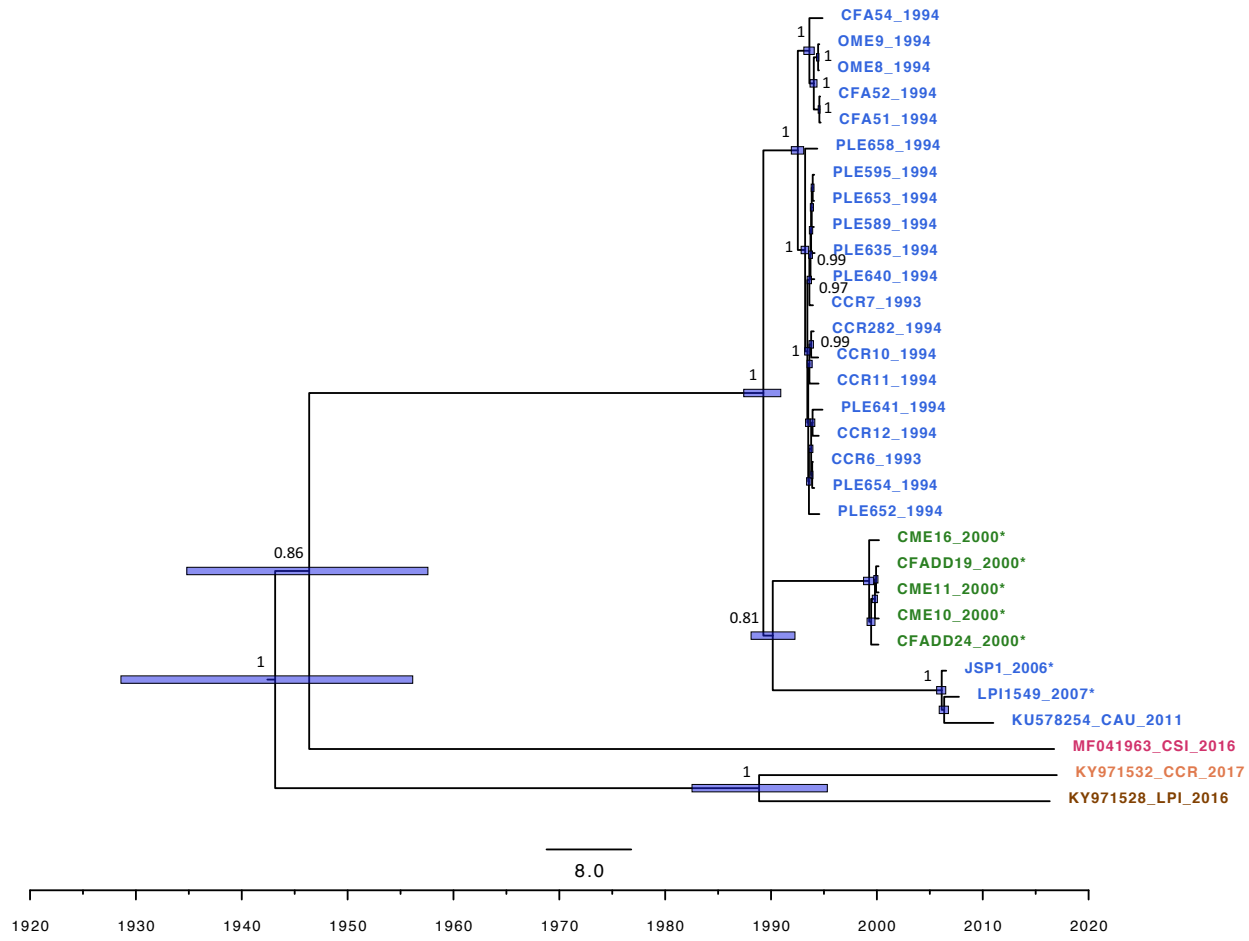

CFA54_1994
OME9_1994
OME8_1994
CFA52_1994
CFA51_1994
PLE658_1994
PLE595_1994
PLE653_1994
PLE589_1994
PLE635_1994
PLE640_1994
CCR7_1993
CCR282_1994
CCR10_1994
CCR11_1994
PLE641_1994
CCR12_1994
CCR6_1993
PLE654_1994
PLE652_1994
CME16_2000*
CFADD19_2000*
CME11_2000*
CME10_2000*
CFADD24_2000*
JSP1_2006*
LPI1549_2007*
KU578254_CAU_2011
MF041963_CSI_2016
KY971532_CCR_2017
KY971528_LPI_2016
0.86
0.81
0.99
0.97
0.99
8.0
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020

Figure 3.5. Time measured maximum clade credibility tree from Bayesian coalescent analysis of whole CDV genome sequences collected in Africa between 1992-2017. Text color refers to sampling location (blue=Serengeti, Tanzania; green=Laikipia, Kenya; red=Ethiopia; brown=Limpopo Province, South Africa, tan=Guateng Province, South Africa). Purple bars indicate the 95% HPD interval for node age estimates. Clade credibility values are shown for nodes >50%.

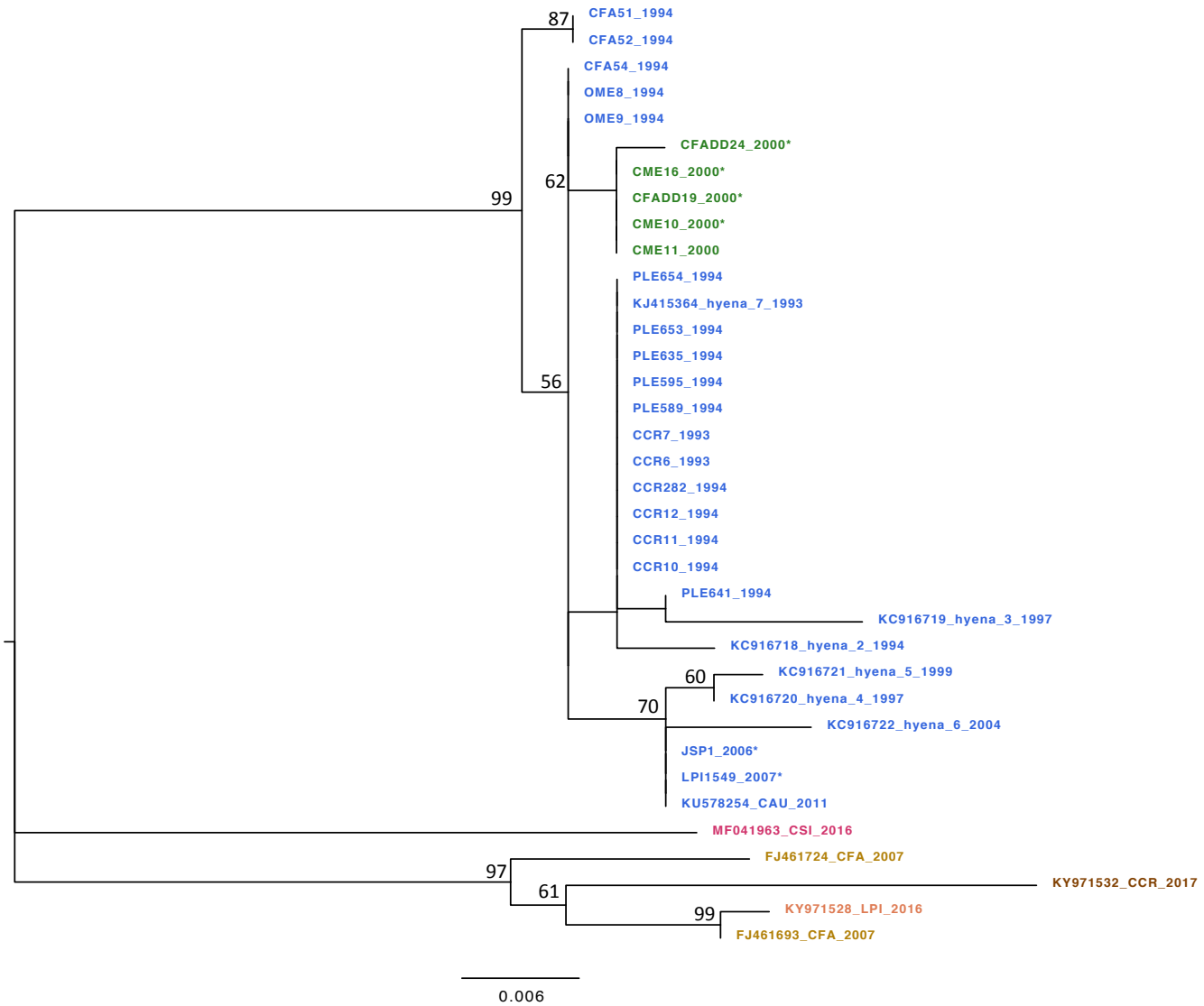

87
CFA51_1994
CFA52_1994
CFA54_1994
OME8_1994
OME9_1994
CFADD24_2000*
62
CME16_2000*
99
CFADD19_2000*
CME10_2000*
CME11_2000
PLE654_1994
KJ415364_hyena_7_1993
PLE653_1994
PLE635_1994
56
PLE595_1994
PLE589_1994
CCR7_1993
CCR6_1993
CCR282_1994
CCR12_1994
CCR11_1994
CCR10_1994
PLE641_1994
KC916719_hyena_3_1997
KC916718_hyena_2_1994
60
KC916721_hyena_5_1999
70
KC916720_hyena_4_1997
KC916722_hyena_6_2004
JSP1_2006*
LPI1549_2007*
KU578254_CAU_2011
MF041963_CSI_2016
FJ461724_CFA_2007
97
KY971532_CCR_2017
61
99
KY971528_LPI_2016
FJ461693_CFA_2007
0.006

Figure 3.6. Phylogenetic relationships of partial CDV H gene sequences (405 bp) from wild and domestic carnivores sampled in Africa from 1993-2017 estimated using the maximum likelihood method. Bootstrap support (1000 iterations) indicated at nodes with greater than 50% support. Text color refers to sampling location (blue=Serengeti, Tanzania; green=Laikipia, Kenya; red=Ethiopia; brown=Limpopo Province, South Africa, salmon=Northern Cape Province, South Africa, tan=Guateng Province, South Africa).

CHAPTER 4

EVOLUTIONARY GENETICS OF CLINICAL CANINE DISTEMPER VIRUS EMERGENCE IN AFRICAN LIONS

ABSTRACT

Emerging pathogens may be highly virulent following spillover in a novel host species, or infection with the same pathogen may cause no overt clinical symptoms in a new host despite successful replication. Serological evidence in wild and captive populations indicates that Canine distemper virus (CDV) infection outcomes in felids are apparently highly variable, from high morbidity and mortality associated with CDV cytopathology, to no observed impact at the individual or population level. Here we use a genome-wide approach to investigate the role of CDV evolution in driving differences in pathogenicity given infection in African lions. In Serengeti National Park, serological evidence indicates that the lion population has been exposed to CDV on at least 5 occasions, while clinical signs and associated mortality were observed only once during an extensive, lethal outbreak. In this study, non-neutral evolution was inferred at 25 sites in the CDV genome that differentiated strains sequenced from clinically infected lions sampled during the fatal Serengeti outbreak, and strains sequenced from sympatric canids during "silent" outbreaks in lions in East Africa. Most of these mutations mapped to functional domains of the RdRp (polymerase) complex and matrix proteins, implicating the processes of transcription and replication, and viral budding as potential barriers to clinical CDV spillover in lions, respectively. We investigated whether these mutations that correlated with pathogenicity in lions in East Africa were involved in other CDV

outbreaks affecting African lions in captive populations in North America. We found that sequences from one outbreak in 2013 shared mutations at seven of the putatively adaptive genetic markers, and mutations at two of these were shared by a third outbreak in 1992, suggesting parallel evolution at these sites correlated with clinical infection in lions. Our results support the hypothesis that viral genetic factors are associated with pathogenicity in African lions given CDV infection, and highlight potential barriers to clinical infection in this novel host species.

INTRODUCTION

The impact of emerging infectious diseases on the health and persistence of wildlife populations is an increasing conservation concern. Ecological forces such as agricultural intensification, human-assisted movements, and shifting host and vector distributions primarily drive disease emergence (Hassell, Begon, Ward, & Fèvre, 2017; Kock, 2014; Patz, Graczyk, Geller, & Vittor, 2000). However cross-species contact alone does not often predict pathogenicity (i.e. the ability of a pathogen to cause disease) given infection in a novel host species, or its population-level consequences. Understanding what limits cross-species pathogenicity therefore is critical in predicting the potential impact that an emerging infectious disease may have on a population of interest.

The susceptibility of a species to a given virus is largely determined by the relationship between viral surface proteins and host cell receptors which permit cell entry (Baranowski, Ruiz-Jarabo, & Domingo, 2001). However, essential viral processes post-entry may limit productive infection and onward transmission in a new host. Within-host limits to pathogenicity given infection by a virus can include intracellular trafficking (Hansen, Qing, Kwon, Mah, & Srivastava, 2000), the ability to evade host immune defenses (Wasik, Muñoz-Rojas, Okamoto, Miller-Jensen, & Turner, 2016), the ability to make viral proteins and replicate in the cell (Neumann & Kawaoka, 2006), and the ability to exit the cell and ultimately the host (Fan et al., 2009). Because viruses are obligate intracellular parasites, they depend on host cell co-factors to execute all of these functions. Thus, intrinsic biochemical differences between host species may need to be overcome to induce a disease state given infection.

Canine distemper virus (CDV) is a notorious example of a multi-host pathogen that can cause high morbidity and mortality and is among the most infectious diseases of mammals. Recent emergence in Amur tigers and leopards, brown hyena, and Iberian lynx underscores the conservation threat in populations of high global significance (Loots, Mitchell, Dalton, Kotzé, & Venter, 2016; Meli et al., 2010; Seimon et al., 2013; Sulikhan et al., 2018). CDV virulence can be extremely high in a novel host species following spillover, while clinical signs may not be observed at all in others, despite replication in the new host. For example, CDV infection in ferrets can cause up to 100% mortality, while experimental infection in domestic cats and pigs causes only a transient immune suppression without overt clinical signs (Appel, Sheffy, Percy, & Gaskin, 1974; Harder et al., 1996). What determines pathogenicity following CDV spillover in a novel host species is not well understood.

CDV infection in mammals has been widely documented with serology in both wild and captive settings, yet clinical progression to disease is seldom reported. Of 272 studies including twenty-two susceptible mammalian Families, only 141 (51%) of these reported clinical signs (Martinez-gutierrez & Ruiz-saenz, 2016). Reports of lethal CDV infection in African lions are rare. Two well-documented clinical CDV outbreaks in African lions occurred in the early 1990s and a third occurred in 2013. The first affected exotic big cats, including African lions, in North American zoos and caused 20 percent mortality (Appel et al., 1994). The second occurred in 1994 in the iconic, wild Serengeti lion population when a catastrophic outbreak of clinical CDV reduced the lion population in the Serengeti Ecosystem by 33 percent (Roelke-Parker et al., 1996). Finally, a CDV outbreak occurred in a captive exotic cat population affecting African lions in Texas in

2013 during which 43 percent of CDV-infected felids with clinical presentation died or were euthanized (Vicky Keahy, personal communication).

Even within a single population, exposure to CDV may have drastically different outcomes in hosts during distinct periods of exposure. Longitudinal serological evidence (1984-2013) indicates that in addition to the unmistakable 1994 outbreak, the Serengeti lion population was infected by CDV on at least 5 occasions without any overt signs of clinical CDV infection. Munson et al. (2008) investigated the hypothesis that CDV related mortality in lions in the Serengeti Ecosystem was driven by ecological factors. In 2001, in the nearby Ngorongoro Crater, high mortality was observed in the Ngorongoro Crater lion population following exposure to CDV. The study concluded that high mortality in African lions in the Serengeti and Ngorongoro in 1994 and 2001 resulted from an interaction between recent CDV exposure, climate extremes, and co-infection with Babesia, a ubiquitous hemolytic protozoan (Munson et al., 2008). However, in 2001, no CDV pathology was found in necropsied lions (n = 2) and clinical CDV infection was ruled out (Munson et al., 2008). Though this work advanced our understanding of the possible population-level impacts of CDV exposure, it did not address apparent differences in the clinical pathology leading to those impacts. Thus, why CDV infection only progressed to disease on one occasion in Serengeti lions despite recurring exposure still is not clear.

Here we investigated the hypothesis that pathogenicity in lions given CDV infection is associated with viral genotype. We applied phylogenetic and evolutionary analyses to identify viral genetic markers of adaptation associated with the emergence of clinical CDV disease in African lion hosts. Adaptation to infect new hosts has the

potential to occur on all viral genes. Here we take a whole genome approach, generating the most comprehensive genomic data set possible from CDV sampled during three distinct clinical outbreaks affecting African lions. We first analyzed CDV strains sampled in East Africa that had caused gravely different clinical outcomes in lions – 'Fatal' in 1993-1994 and 'Silent' thereafter. All contemporary East African CDV sequences share a common ancestor in the 1980s (Chapter 3) thus providing a unique opportunity to evaluate evolutionary differences between CDV strains from the same genetic background, in the same lion host population. Subsequently we investigated whether mutations at the sites associated with pathogenicity in East African strains were involved in the emergence of clinical CDV in African lions during outbreaks in North America, and/or with emergence in other novel hosts.

MATERIALS AND METHODS

Study specimens

CDV-positive specimens were collected during lethal outbreaks of CDV in captive African lions and other big cats in North America between 1992-1993 and in 2013. In both outbreaks, lions presented with gastrointestinal, respiratory, and neurological (CNS) signs including myoclonus, seizures, and paraparesis, although some individuals had only CNS signs (Appel et al., 1994; Vicky Keahy, personal communication). CDV was confirmed during the 1992-1993 outbreak by histopathological lesions and virus isolation. Viral isolates were generated from clinical specimens of one leopard (*Panthera pardus*), an African lion (*Panthera leo*), and a raccoon (*Procyon lotor*). Methods for viral isolation were described previously (Appel et al., 1994). Briefly, tissue homogenates or blood lymphocytes from infected animals were

cocultivated with domestic dog blood lymphocytes and incubated, then passaged in dog kidney cells, and finally passaged in dog lung macrophage cultures (Appel et al., 1994). In May 2013, CDV clinical signs and mortality were observed in captive African lions and other big cats at a facility in Texas (USA). Tissues from actively infected lions in 2013 were not available, though CDV involvement was confirmed by PCR from recovered individuals (Rebecca Wilkes, personal communication). A CDV-positive specimen was collected from the urine of a recovered tiger after the 2013 outbreak had subsided. Wild and domestic carnivores were sampled in Texas contemporaneously. Rabies-suspect, rabies-negative raccoons (n = 4) and a grey fox (*Urocyon cinereoargenteus*, n = 1) were sampled in April - May 2013. Domestic dogs (*Canis lupus familiaris*, n = 5) were sampled between November 2012 and December 2013.

In East Africa, CDV-positive specimens were collected from clinically infected African lions, spotted hyenas, domestic dogs, and bat-eared fox during a fatal outbreak in 1993-1994 in the Serengeti Ecological Region (Chapter 2). Lions infected in the 1994 outbreak in Serengeti predominantly presented with CNS signs without obvious involvement of respiratory or gastrointestinal systems, although interstitial pneumonia was observed post-mortem (Roelke-Parker et al., 1996) and CDV antigen was found with PCR in pulmonary and digestive tissues (Chapter 2). Histopathological lesions and virus isolation confirmed CDV involvement in affected tissues from lions, the most prominent findings indicating encephalitis, interstitial pneumonia, and lymphoid depletion in the lymph nodes and spleen associated with cytopathic effect of CDV infection (Roelke-Parker et al., 1996). During silent outbreaks, periods when some lions seroconverted without overt clinical infection, CDV-positive samples were collected from sympatric,

clinically infected canids in Laikipia (Kenya) and Serengeti (Tanzania) including African wild dogs, jackals, and domestic dogs (Chapter 3).

Near complete genome sequences, hereafter WGS, were generated from viral isolates from big cats, including African lion and a raccoon (1992) and clinical samples from mesocarnivores and domestic dogs (2013) according to methods described in Chapter 2. Sequences will be submitted to Genbank under accession numbers in Table S4.1.

The viral loads in recovered big cats in 2013 were too low to generate WGS. A partial H gene sequence (718 bp) was generated from a tiger that initially suffered loss of appetite, loose stools, and seizures, apparently recovered over 10 months, then suddenly and rapidly declined with undiagnosed neurologic symptoms and died. Viral RNA was extracted from a urine specimen using the Qiagen Viral RNA Mini Kit according to manufacturer's instructions. A one-step RT-PCR reaction using previously published primers was performed (SSIII One-Step RT-PCR with Platinum Taq). The RT-PCR reaction conditions were as follows: 10 min at 55°C, followed by 2 min at 94°C , then 40 cycles of 30 sec at 94°C, and 30 sec at 48°C, and 2 minutes at 68°C. Each reaction used 2-ul total RNA extract in a 50-ul volume reaction. Oligonucleotide concentrations were used at concentrations according to manufacturer's instructions. Sequences were generated using the Applied Biosystems Automated 3730xl DNA Analyzer with Big Dye Terminator chemistry. Partial H gene sequence will be submitted to Genbank (Accession number pending).

WGS generated from carnivores sampled in East Africa between 1992-2007 are reported in Chapters 2 and 3. All complete CDV genome sequences with associated host

data available at the time of this study were retrieved from the National Center for Biotechnology (NCBI) GenBank database (http://www.ncbi.nlm.nih.gov/) (n = 78). Sequences and associated data including accession numbers for publically available sequences are listed in Table S4.2. Accession numbers from Chapters 2 and 3 are pending submission.

Genetic markers

To identify genetic markers associated with pathogenicity of CDV in African lions we assessed differences between circulating CDV sequences in East Africa (1992-2011) resulting in apparently different clinical outcomes in lions. Results from our previous analysis suggest that all CDV sequences from East Africa in our dataset can be traced back to a common ancestor in the late 1980s (Chapter 3). We divided the sequence data into two groups: Fatal and Silent. The Fatal group comprised all sequences from the 1993-1994 outbreak, including both non-canids and canids. Despite that 13 mutations separated non-canids and canids in this outbreak, it is unknown whether the genotype found in canids at the time was able to cause clinical disease in lions, therefore we include all sequences from this period in the Fatal group. The Silent group included all other CDV sequences sampled in East Africa between 1997-2011. All sequences in the Silent group came from canid species. For this analysis we assume that African lions that seroconverted during the course of the study were exposed to the same CDV genotype that was circulating in canids at the same location. Consensus genotypes of the groups were compared in a multiple sequence alignment performed using the MUSCLE algorithm in MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Edgar, 2004; Kumar, Stecher, & Tamura, 2016). If a mutation occurred at a

given position in the genome that differentiated the two groups, and/or differentiated non-canids from canids sampled in 1993-1994, it was considered a candidate marker of pathogenicity in lions.

Selection Analysis

To determine whether the mutations differentiating Fatal and Silent groups in East Africa were adaptive rather than the result of founder effects and neutral evolutionary processes, we performed selection analyses on an alignment of all available CDV WGS. First, sequences were screened for recombination using a suite of algorithms available in Recombination Detection Program (RDP4) software (Lole et al., 1999; Martin & Rybicki, 2000), including Recombination Detection Program (RDP) (Martin & Rybicki 2000), GENECONV (Sawyer 1989), BOOTSCAN (Martin et al. 2005b), Max-Chi (Maynard Smith 1992), CHIMAERA (Posada & Crandall 2001) and SISCAN (Gibbs et al. 2000). All tests employed a cutoff of $p \leq 0.05$ and a Bonferroni correction. WGS with significant evidence of recombination events in at least 5 of the 7 methods were removed from further analyses.

Codon based selection detection methods assume that nonsynonymous mutations, those that cause an amino acid change in the protein sequence, have a larger effect on fitness than do synonymous mutations that do not alter the protein. Thus a common measure of selection in protein-coding sequence is the ratio of the rate of nonsynonymous mutations to the rate of synonymous mutations, or $\omega = dN/dS$. When ω is greater than 1, selection is said to be positive. When ω is less than 1, a site is said to be under negative or purifying selection. All tests were performed in the HyPhy software (Pond, Frost, & Muse, 2005) implemented on the DataMonkey web server (Delport, Poon, Frost, &

Kosakovsky Pond, 2010). A p-value cutoff of $p \leq 0.10$ was used to determine significance in all analyses. This cutoff was used because the null model used for comparison in these methods assumes neutral evolution, which is likely to be violated by RNA virus evolution, which is dominated by negative selection.

Positive and negative selection were assessed on a site-by-site basis, i.e. at each codon, using a counting method, SLAC, a fixed effects likelihood method, FEL, (Kosakovsky Pond & Frost, 2005), and a mixed effects model, MEME (Murrell et al., 2012). SLAC and FEL, assume that selection pressure is equal at each codon in the alignment, i.e. at each branch or lineage. This assumption is violated in cases where selection is episodic and/or affects a subset of lineages, as is predicted when testing the hypothesis that viruses evolve differently in different hosts. Therefore, to identify sites experiencing episodic diversifying selection, we used a mixed effects method that allowed ω to vary at each branch.

Duplicate sequences and those with evidence of recombination were removed and the best fitting nucleotide model was selected prior to analyses. All three methods, SLAC, FEL, and MEME were used to identify selection on a site-by-site basis in the coding alignment and corresponding phylogeny. Selection detection methods classify sites with an excess of synonymous mutations, or $dN<dS$, as experiencing negative selection, assuming that nonsynonymous substitutions are being removed from this site because they reduce fitness. However, this assumption does not consider the possibility that positive selection acts at synonymous sites, which has been documented in viruses and other taxa (Cuevas, Domingo-Calap, & Sanjuán, 2012; Novella, Zárate, Metzgar, & Ebendick-Corpus, 2004; Pepin, Domsic, & Mckenna, 2008). Here, we consider that sites

with an excess of synonymous mutations, or dN<dS may be the product of adaptive evolution. In other words, an excess of synonymous mutations at a site which presents as $\omega<1$ may be indicative of positive selection for a substitution that occurs at the third codon position. Thus, to identify sites in the CDV genome that may be adaptive, we consider sites with $\omega<1$ as possible sites of positive selection for synonymous mutations.

Selection acting on mutations at intergenic sites was not investigated because the selection detection methods used only evaluate protein-coding sequence.

Phylogenetic Analysis

Complete genomes sequences were not available from infected lions during the 2013 outbreak, however a partial H gene sequence from a co-located, infected tiger was available. Phylogenetic analysis of partial H gene sequences was performed to determine the origin of CDV at the captive facility where lions were infected in 2013. The results determined the most closely related sequences for which we have complete genomes, which were then used as a proxy for African lion sequences in the WGS phylogenetic analysis. The Maximum Likelihood method was used to reconstruct the phylogenetic relationships of partial CDV H gene sequences (718 bp) in MEGA (Kumar et al., 2016). The General Time Reversible model of nucleotide substitution allowing invariant rates among sites (GTR+I) was the best fit to the data as determined by consensus of the Akaike Information Criterion (AIC), and Akaike Information Criterion corrected for small sample sizes (AICc) (Hurvich & Tsai, 1989) implemented in jModelTest (Posada, 2008).

Phylogenetic analysis using all WGS was performed to determine if sequences causing clinical infection in lions during the three discrete outbreaks were genetically

independent. For visualization, the alignment was pruned by including only one sequence per species per country of origin per year. The Maximum Likelihood method was used to reconstruct the phylogenetic relationships of CDV WGS in MEGA (Kumar et al., 2016). The General Time Reversible model of nucleotide substitution allowing invariant rates among sites and a gamma rate distribution (GTR+I+G) was the best fit to the data as determined by consensus of the Akaike Information Criterion (AIC), Akaike Information Criterion corrected for small sample sizes (AICc) (Hurvich and Tsai 1989), Bayesian Information Criterion (BIC), and Decision Theory Criterion (DT) implemented in jModelTest (Posada, 2008).

Mutations at sites evolving non-neutrally (i.e. $\omega \neq 1$) that differentiated Fatal and Silent strains were mapped to the tips of the phylogeny to visualize the frequency and distribution of these mutations in all available CDV genomes.

RESULTS

Genetic markers

Based on the multiple sequence alignment we identified 49 mutations at nucleotides in the CDV genome sequence that differentiated the Fatal and Silent strains of CDV in circulation in East Africa between 1993-2011 (Table 4.1). Three mutations occurred in the overlapping open reading frames (ORF) of three proteins, two in the P/V/C ORFs and one in P/V. Thus, taking translation into account, 54 total mutations were observed. These mutations occurred in each of the 6 structural genes, the 2 nonstructural genes, and in intergenic untranslated regions (UTRs). Nonsynonymous substitutions accounted for 19 (39%) of the mutations, 27 mutations (54%) were synonymous, and 3 occurred in UTRs (6%).

Recombination and selection analysis

Of 126 whole CDV genome sequences screened, 21 putative recombination events were detected in 11 sequences (Table 4.2). Nine of these putatively recombinant sequences have been described elsewhere (da Fontoura Budaszewski et al., 2016; Han, Liu, & Li, 2008; Ke et al., 2015; Yuan et al., 2017). Evidence of one recombination event in a domestic dog sampled in the USA in 2004 (Accession EU716337) was detected with significant support by five of seven algorithms employed by the RDRP software. Four recombination events were detected in a single raccoon dog genome sampled in 2013 in China (Accession KJ994343). Each event had high support by all seven algorithms, except for one event, which had high support from six of the seven algorithms. All putative recombinant sequences were removed from all further analyses.

Site-by-site selection analyses (with SLAC, FEL and MEME) identified 133 sites with significant pervasive positive or episodic diversifying selection ($p \leq 0.1$) (Table S4.3), while 1,311 sites showed evidence of purifying (negative) selection (with SLAC and FEL) operating across the phylogeny ($p \leq 0.1$) (Table S4.4).

Of the 49 genetic markers differentiating Fatal and Silent strains, 26 were experiencing significant positive and/or negative selection according to the site-by-site selection analyses (Table 4.3). Pervasive positive or episodic diversifying selection accounted for 6 of these sites (23.1%) occurring on the N gene (residue site 451), the P/V gene (site 280), the M gene (site 9), the H gene (site 549), and the L gene (sites 133 and 1402). Twenty sites were classified by SLAC and FEL algorithms as experiencing significant negative selection, i.e. an excess of synonymous substitutions were detected. These occurred on all CDV gene products, except for C. At residue 93 on the L gene, dN-

dS was significantly less than zero, indicating an excess of synonymous mutations which generally results in the consensus amino acid being conserved. However, at this site a nonsynonymous mutation occurred on the branch to the Fatal group. At position 13,386, the Fatal and Silent strains differed because the Silent strain had a unique mutation, while all other sequences had a conserved nucleotide at that site. This site will not be considered further as a candidate marker of pathogenicity in lions.

Phylogenetic analyses and distribution of mutations at non-neutrally evolving sites

Phylogenetic analysis of partial H gene sequences indicated that a tiger infected at the same facility in 2013 in Texas where African lions displayed clinical CDV symptoms was identical to partial H gene sequences from wild mesocarnivores, and distant from domestic dogs sampled contemporaneously in Texas (Figure 4.1). Two distinct global lineages were found circulating in Texas in 2013. America-2 was found in 5 wild carnivores, a tiger, and in one domestic dog, while America-3 was found in 4 domestic dogs.

Phylogenetic analysis of 65 CDV WGS indicated that 10 previously recognized global clades, which correlate mostly by geography, were represented in our data (Figure 4.2A). Sequences from the three outbreaks causing clinical infection in African lions fall into 2 global lineages, Africa-2 and America-2. All sequences from East Africa occur in Africa-2. Both North American outbreaks in captive African lions, in 1992-1993 and 2013, are of the America-2 global lineage and are 95% similar on the H gene, though they occurred 21 years apart.

Amino acid and nucleotides of putatively adaptive genetic markers were mapped to the tips of the phylogeny to visualize the distribution of mutations at these sites over

available CDV WGS, including those infecting African lions in separate outbreaks and in other novel host species (Figure 4.2B). In East Africa, by definition, the consensus sequence of Fatal strains differs from Silent strains at all 25 sites. However, 8 of the 25 selected genetic markers differentiating Fatal and Silent East African strains also segregated the canid and non-canid sequences within the 1993-1994 outbreak, with three exceptions involving a single lion. Lion PLE658 bears the 1994 *canid* genotype at 3 sites on the L gene (Figure 4.2B) and the non-canid genotype at the remaining putatively adaptive genetic markers.

Mutations at 6 putatively adaptive genetic markers (excluding site 13,386) are shared between the Fatal CDV strain (Africa-2) and a strain causing clinical infection in lions and other big cats in 2013. A third outbreak in African lions and other big cats in 1992-1993 shares two of these. CDV strains from both outbreaks outside of Africa belong to the America-2 global lineage.

Of the 10 recognized global lineages in our dataset, sequences in Asia-1 (with the greatest host taxa diversity) shared the most mutations at putatively selected genetic markers with the Fatal strain; 10 total common sites with 5 sites in any one individual sequence. Host species from Asia-1 bearing mutations at 5 sites in common with the Fatal strain include giant pandas (China), and a domestic dog (China). America-2 (comprised mostly of novel hosts, felid and procyonid) had 8 total common sites for the genotype and 7 total sites in any one individual sequence. America-1 (comprised mostly of vaccine strains) had 7 total sites in common, with 6 being the most common sites in any one sequence. Africa-1 (comprised of a spotted hyena and an African wild dog from South Africa) had 6 total along the genotype, with 5 and 4 in each sequence. Two total

mutations across the American 4 genotype were shared with the Fatal strain, while two were found in any one sequence. The South America 1/ Europe strain and the Arctic-like strain each shared 1 mutation with the Fatal strain, and a free-ranging tiger from the Russian Far East in this group shared none. Two genotypes comprised of only domestic dogs, America 3 and Asia 2, both had the consensus nucleotide or amino acid at all candidate markers.

DISCUSSION

In this study, we sought a viral genetic basis associated with pathogenicity of CDV infection in African lions to gain a better understanding of barriers to clinical spillover and predict threats to populations of conservation concern. Putative adaptive evolution was inferred at 25 sites in the CDV genome that differentiated strains sequenced from clinically infected lions sampled during a fatal outbreak, and strains sequenced from sympatric canids during silent outbreaks in East Africa. Most of these mutations mapped to functional domains of the RdRp (polymerase) complex and matrix proteins, implicating the processes of transcription and replication, and viral budding as potential barriers to clinical CDV spillover in lions, respectively. We investigated whether the mutations implicated in clinical spillover in lions in East Africa were involved in other CDV outbreaks affecting African lions in captive populations in North America. We found that sequences from one outbreak in 2013 shared mutations at 7 of the putative adaptive genetic markers, and mutations at 2 of these were shared by a third outbreak in 1992, suggesting parallel evolution at these sites to clinical infection in lions.

Our results support the hypothesis that viral genetic factors are associated with CDV pathogenicity given infection in African lions. We found that 25 putatively adaptive

mutations differentiated Fatal and Silent strains of CDV in African lion populations in East Africa. During the catastrophic 1994, clinical CDV outbreak in lions (samples from which comprise the Fatal group), 13 mutations separated sequences from canid and non-canid species (Chapter 2). Mutations at 8 of these 13 sites overlap the 24 sites differentiating Fatal and Silent strains in East Africa. This finding is consistent with previous work suggesting that mutations within 1994 are explained by adaptive evolution to infection of non-canids at these sites (Nikolin et al., 2017; Chapter 2) and supports the role of viral genetic factors in explaining clinical spillover in lions in East Africa.

Although the putatively adaptive mutations at sites are correlated with clinical outcome of CDV it is not clear whether these mutations are necessary and sufficient to cause clinical infection in lions. One lion sampled during the outbreak provides anecdotal evidence to this end. The sequence of PLE658 is distinctive because this lineage is slightly older than that of all other lions and hyenas sampled in the outbreak, despite being sampled later in the outbreak after most mortality had occurred in the study population (Chapter 2). Thus, even though this lineage may have been present in the Serengeti carnivore community earlier than the other lineages sampled in lions in 1993-1994, it did not ignite the outbreak. Histopathology of tissues from PLE658 was within normal limits in a liver section and lymph node, though lymphoid depletion and rare intranuclear inclusions in the spleen were suggestive of CDV infection (Linda Munson, unpublished data). Immunohistochemistry was inconclusive due to severe autolysis of the tissues. However, the evidence available may suggest subclinical infection and thus a dead-end chain of transmission because epithelial cell infection is required for onward transmission. We found that this individual had an intermediate genotype between canid

and non-canid variants, sharing only 9 of the 13 mutations with other lions. This finding could implicate these specific mutations (L-93, L-1402, and L-2058) as particularly important for determining clinical infection in lions in this outbreak, or suggest that a series of mutations together was necessary for clinical infection. Notably, the sites where these three mutations occur are almost entirely conserved across all other known sequences.

Our data identified putatively adaptive mutations associated with CDV pathogenicity in African lions within a single CDV lineage (Africa-2) and provided a basis for testing the hypothesis that a common mutational signature (i.e. genotype) explains clinical spillover of CDV in other African lion populations (i.e. phenotype), and/or other novel host species. Assessing the distribution of mutations at the 25 identified sites over all available WGS revealed mutations at two sites associated with all outbreaks of clinical CDV in African lions in our dataset. These are a tyrosine to a histidine at residue 549 (Y549H) on the hemagglutinin protein and two synonymous mutations at residue 710 and 1619 on the large protein.

Our finding of Y549H in African lions is consistent with that of McCarthy et al. (2007) which found positive selection at this site and an association of histidine in non-dog hosts (McCarthy, Shaw, & Goodman, 2007). Multiple subsequent reports have supported this pattern, including Nikolin et al. (2012b) which revisited the question of the H-549 association using a larger sample size and narrowed the association to non-canid host, i.e. wild canid hosts were more likely to bear a tyrosine like their domestic counterparts. Furthermore, mutagenesis of a canid strain to bear a histidine at 549 supported the role of this mutation as a determinant of CDV host tropism (Nikolin et al.,

2012a). Nevertheless, there was no clear association of histidine at 549 with sequences when considering all available sequences from Felid species (Terio & Craft, 2013). Our result suggests that from the Family Felidae, Y549H is at least associated with lethal CDV infection in African lions. This discrepancy might be explained by the dominant cross-species pathways of transmission during an outbreak. Specifically, most felid species occur at low densities and have limited social contacts and may be more likely to become infected by contact with another species. The residue at site 549 in a given sequence is likely to be driven by the local reservoir and/or whichever species is fueling a given outbreak. Thus, if the reservoir is a raccoon dog, i.e. a canid, then you might expect Y, if a raccoon, then H. Interestingly, tigers infected in Japan, China and Russia, where the reservoir or source might be expected to be raccoon dog are 549Y (Metzker et al., 2010; Nagao et al., 2012; Seimon et al., 2013; Zhang et al., 2017), while tigers infected in both North American outbreaks putatively fueled by raccoons bear 549H (Appel et al., 1994).

The mutation at residue 710 occurs in the 223-aa conserved region of the L protein, CRIII. This functional domain is host to the catalytic center of phosphodiester bond formation, i.e. where polymerization occurs during transcription and replication (Chattopadhyay, Raha, & Shaila, 2004). That this mutation is shared by all pathogenic strains in lions, and it is conserved in all other available CDV sequences from all hosts and global lineages, suggests parallel evolution and supports the hypothesis that this mutation confers a fitness advantage to CDV in African lions.

Our phylogenetic results indicate that the 1992 and 2013 North American outbreaks were caused by the same global lineage, America-2. Thus, mutations at

common sites between these two outbreaks may be identical by descent, rather than by parallel evolution. Nevertheless, that the same global lineage was responsible for two of three well-documented outbreaks in lions supports the hypothesis that viral genetic factors are associated with lethal CDV spillover in lions. Further, it suggests that this lineage, America-2, could have an intrinsic ability to cause lethal infection in big cats.

The observation of two mutations in common across all lion outbreaks may be conservative. The 1994 East Africa and 2013 North America outbreaks shared mutations at 7 sites, while the 1992 North America strain shared 2 of these. Sequences from 1994 and 2013 were sequenced from clinical samples, whereas the sequences from 1992 were isolated and passaged in various types of canine cells and canine cell lines. Thus any mutations that conferred an advantage in lion cells may have reverted or been lost in adaptation to cell culture on canine cells.

In addition to America-2, putatively adaptive mutations found in the Fatal group were shared with CDV sequences in other lineages, the majority of mutations at shared sites (78.8%) were found in non-dog host species and/or vaccines. CDV vaccines historically were made by repeatedly passaging a clinical isolate in alternative host species in-vivo and in alternative host cells in culture until the isolate acquired so many mutations that it was attenuated in domestic dogs. Thus, the distribution of mutations in common with the Fatal strain in a diversity of species (including vaccines) suggests that some of the putatively adaptive mutations identified in our analysis may not be specific to lion infections, rather related to adaptation to alternative host species more generally. This is consistent with the thesis of Nikolin et al. (2012a, 2012b, 2017) that CDV in non-dog hosts acquire “generalist” mutations. Global lineages comprised only of domestic

dogs do not share any mutations with the Fatal group at sites under selective pressure. Finally, it is interesting to note that the only other felid in our dataset, a wild Amur tiger from the Arctic-like lineage, does not share any putatively adaptive mutations with the Fatal strain responsible for African lion clinical infection.

Putatively adaptive mutations occur primarily in the RdRp complex and the M protein

Functional domains of CDV and other related Paramyxoviruses are well characterized through experimental mutational studies that manipulate viral sequence and observe the resulting change in phenotypes. Furthermore, within the Morbillivirus genus, lessons learned in the study of one species are often transferable to others because the Morbilliviruses share genetic, structural, functional and pathological similarities. The majority of the 25 putatively adaptive mutations correlated with pathogenicity in African lions in East Africa were located in functional domains of the RdRp complex essential for efficient viral transcription and replication (n=17), and on the matrix protein critical for viral spread (n=4). Mutations in these functional domains suggest that these processes may present barriers to clinical spillover in African lions. All mutations identified in this study are mapped in relation to functional domains on a model of the CDV genome in Figure 4.3A.

The RdRp complex is comprised of three viral proteins, the N, P and L (illustrated in Figure 4.3B which was reproduced directly from (Sourimant & Plemper, 2016)) and host cell cofactors (not shown). The RdRp complex is important because it synthesizes viral mRNA and makes copies of the viral RNA genome. The nucleocapsid (N) encapsidates the viral RNA and protects it from degradation and detection by the host innate immune system, but needs to be dealt with for the polymerase to access the viral

genome for polymerization. The large protein (L) contains the catalytic functions of the RdRp necessary for mRNA transcription and viral replication, most of which have been mapped to six regions that are conserved among all Morbilliviruses (Sidhu, Menonna, Cook, Dowling, & Udem, 1993) plus the C-terminus. The phosphoprotein is an essential cofactor in polymerase activity, responsible for chaperoning the polymerase (L) to the nucleocapsid, positioning it at the 3' promoter of the encapsidated viral RNA, and preventing it from falling off during polymerization (Sourimant & Plemper, 2016).

On N, three potentially adaptive mutations in our data occurred in the 125-aa N-tail domain (pink in Figure 4.3A and B). Research suggests that this short amino acid chain protruding from the nucleocapsid is involved in 1) binding the P which in turn recruits the L for transcription and replication (Wang 2013), 2) binding the matrix protein for translocation of the RdRp to the cell surface during egress (Ray 2016), 3) binding a host cell receptor that leads to suppression of immune cell proliferation in the host (Laine 2005), 4) facilitating access to the RNA genome for RdRp, and 5) regulating RdRp activity and as such determining pathogenicity (Cox, Krumm, Thakkar, Sohn, & Plemper, 2017; Thakkar et al., 2018). Mutations in the N-tail thus may affect three vital processes: efficient transcription/replication, the massive suppression of the host immune system, and onward spread. Notably, experimentally mutated isolates with truncated N-tails can replicate in vitro but are attenuated in vivo, exposure to which causes the build up of immunity in the host.

Three mutations occur on the P protein. The first of these mutations, P-280 also occurs in the V protein open reading frame, which is not a part of the RdRp. On the V gene, mutation V-280 occurs immediately following a highly conserved, essential chain

of 47 amino acids that interrupts host innate immunity. A V-deficient CDV caused limited viremia, transient leukopenia, and mild symptoms in experimentally infected ferrets (von Messling, Svitek, & Cattaneo, 2006). Thus, mutation P/V-280 may be determined by its effect on V gene function.

The remaining two mutations on the P gene fall in the PMD, a short domain (of 66 residues) that research suggests both binds L and is pivotal in its correct positioning for polymerization initiation (Chen, Cortay, & Gerlier, 2003; Raha, Kaushik, & Shaila, 2004). Thus, these mutations may optimize recruitment of the polymerase for transcription/replication in a novel host.

On the L protein, six conserved domains (common to all Morbilliviruses) are associated with enzymatic activity of mRNA transcription and virus replication, e.g. catalyzing polymerization and mRNA capping, methylation, and polyadenylation (essential for translation and evading host innate immunity) (Sidhu et al., 1993). In addition there are intrinsically disordered sites having no secondary structure and a connector domain (Liang et al., 2015). Six of eleven mutations associated with the Fatal strain occur in discrete and specific enzymatic domains on the L protein (Table 4.3). These occur in CRII, CRIII, CRIV, CRV, and the C-terminus. The others occur in connector and linker domains and could be involved in secondary structure and proper RNA folding facilitating interaction between the functional domains (Dochow et al., 2012). Mutations on the L gene may optimize the processes of mRNA transcription and viral replication, and indirectly contribute to host immune evasion which requires that capped, polyadenylated mRNAs be produced.

Interestingly, the gene with the second most mutations associated with the Fatal strain was the matrix protein (M). The M protein is essential for viral assembly and egress from the infected host cell, i.e. onward transmission. M interacts with the cytoplasmic tails of the virus glycoproteins, F and H, the N-tail of N, and host cellular actin filaments to assemble virus particles at the cell surface (El Najjar, Schmitt, & Dutch, 2014). The role of matrix protein is especially important in polarized epithelial cells, where it directs viral assembly and egress from the apical side, ensuring onward spread of the virus by releasing it into the respiratory, urinary, and gastrointestinal tract (Dietzel, Anderson, Castan, von Messling, & Maisner, 2011). M can regulate cell fusion, which is an important process in the delicate balancing act between productive cell infection and host immune evasion. Four mutations were found on M that differentiated the Fatal and Silent strains in East Africa, which is more than expected given the small size of the protein. However, no M mutations occurred between canid and non-canid strains within the Fatal group.

Previous work showed that in vaccine strains of CDV, six mutations occur on the M gene as compared to the virulent wildtype, interrupting its function (Dietzel et al., 2011). Recombinant viruses replacing the vaccine M into a wildtype strain resulted in complete attenuation in vivo. These findings demonstrate that a virus with a defective M (as found in vaccines) can replicate in host cells, initiating the host immune response leading to immunity. Thus, if a wildtype CDV M protein cannot interact with the cell cytoskeleton or other host cofactors of a novel host effectively, then a host exposed to wildtype virus may seroconvert without clinical infection.

Mutations on the two glycoproteins of CDV (F and H) associated with clinical spillover of CDV in African lions in East Africa were fewer than expected, with two potentially adaptive mutations on each gene. CDV infection in any organism is initiated when the H protein of CDV binds to the host immune cell receptor, signal lymphocyte activation molecule or SLAM. The H protein is responsible for recognizing and binding the host cell, and works in concert with proximal F proteins to fuse the cell and viral membranes. Because this process is required to initiate infection, and the SLAM binding region has variable affinity to SLAM receptors from different host species, genetic diversity in the SLAM binding region has been considered to be the primary determinant of CDV host tropism. Indeed, SLAM binding defines susceptibility and host tropism, however it does not determine host pathogenicity. For example, CDV recognizes and binds domestic pig and cat SLAM, and replicates in SLAM-positive cells, but does not cause clinical disease (Appel et al., 1974; Harder et al., 1996). Considering the widespread occurrence of CDV seroconversion in taxa ranging from elephants to deer and all Families of carnivores, the ability to bind SLAM does not seem to be the limiting factor in cross-species pathogenicity.

Potential role for intermediate host species

Our phylogenetic results suggest raccoons may have infected African lions in both North American outbreaks (Appel et al., 1994) and previous analyses implicate spotted hyenas as the source of CDV infecting African lions in 1994 (Chapter 2). Persistent infections with CDV are not known to occur in domestic dogs, however recent reports suggest that raccoons in North America and spotted hyenas in East Africa may support subclinical CDV infection (Marescot et al., 2018; Pope, Miller, Riley, Anis, &

Wilkes, 2016). These observations raise the possibility of these species acting as intermediate or liaison host species that form a bridge between canid-adapted sequences and those capable of infecting African lions. In this study, over 8% of the WGS we analyzed had evidence of recombination events, most of which occurred in raccoons in North America. Furthermore, we detected two distinct CDV lineages co-circulating in the Dallas area in 2013 (Texas, USA) consistent with other reports of multiple co-circulating strains (Riley & Wilkes, 2015). A persistent infection would enable the accumulation of mutations over the course of a long infection, possibly adapting the virus to the individual host. Further it would provide opportunity for co-infections and recombination, as has been implicated in clinical infection of giant panda (Han et al., 2008). Persistent infections could increase CDV genetic diversity possibly providing a vessel for adaptation to novel hosts and spread to sensitive African lion populations.

Conclusions

Identifying genetic markers of clinical outcome in African lions is a useful predictive tool and can improve surveillance when 1) the mutations are necessary and sufficient to cause disease, and 2) are present in the susceptible host community before spillover occurs, i.e. “off-the-shelf” (Pepin, Lass, C Pulliam, Read, & Lloyd-Smith, 2010). Specifically, the risks of CDV infection in a population of conservation concern might be determined by screening sympatric, putative reservoir populations for CDV variants bearing the genetic signature of pathogenicity in lions. Our data do not address whether these conditions are met by the putatively adaptive mutations we found to be associated with clinical spillover of CDV in lions. For this, experimental validation of the specific role of the mutations that we identified is critical. However, despite that the

necessity of the identified mutations has not been demonstrated, convergent evolution at at least two sites strongly suggests a functional role in clinical infection in lions. Surveillance for these two mutations at a minimum is warranted. To this end, our analysis revealed that two well-documented outbreaks affecting African lions were caused by the America-2 global lineage. This lineage should be screened for and considered a risk to sensitive felid populations.

The presence of a common genetic signature requisite for clinical spillover in African lions might be precluded if there exists multiple ways, i.e. mutations, to achieve the same phenotypic effect. This may be that different mutations or combinations of mutations in a certain functional domain will have the same affect on the process, or that mutations affecting different processes are sufficient to produce the same phenotype in vivo, i.e. clinical outcome in lions. Identifying sites in the CDV genome under selective pressure in clinical outbreaks of CDV has an intrinsic value in highlighting viral processes that may pose a barrier to CDV spillover in lions and other species.

Limitations of the study

In this study, we identified sites in the CDV genome that potentially are important determinants of pathogenicity in African lion hosts. Our data are strictly correlative and experimental validation is essential to determine if and how mutations at these sites contribute to the expression of disease phenotype in lions.

LITERATURE CITED

Appel, M. J. G., Yates, R. A., Foley, G. L., Bernstein, J. J., Santinelli, S., Spelman, L. H., … Summers, B. A. (1994). Canine Distemper Epizootic in Lions, Tigers, and Leopards in North America. *Journal of Veterinary Diagnostic Investigation*, *6*(3),

277–288. https://doi.org/10.1177/104063879400600301

Appel, M., Sheffy, B. E., Percy, D. H., & Gaskin, J. M. (1974). Canine distemper virus in domesticated cats and pigs. *American Journal of Veterinary Research*, *35*(No.6), 803–806. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19742225506

Baranowski, E., Ruiz-Jarabo, C. M., & Domingo, E. (2001). Evolution of cell recognition by viruses. *Science (New York, N.Y.)*, *292*(5519), 1102–5. https://doi.org/10.1126/SCIENCE.1058613

Chattopadhyay, A., Raha, T., & Shaila, M. . (2004). Effect of single amino acid mutations in the conserved GDNQ motif of L protein of Rinderpest virus on RNA synthesis in vitro and in vivo. *Virus Research*, *99*(2), 139–145. https://doi.org/10.1016/J.VIRUSRES.2003.11.003

Chen, M., Cortay, J. C., & Gerlier, D. (2003). Measles virus protein interactions in yeast: new findings and caveats. *Virus Research*, *98*(2), 123–129. https://doi.org/10.1016/J.VIRUSRES.2003.09.003

Cox, R. M., Krumm, S. A., Thakkar, V. D., Sohn, M., & Plemper, R. K. (2017). The structurally disordered paramyxovirus nucleocapsid protein tail domain is a regulator of the mRNA transcription gradient. *Science Advances*, *3*(2), e1602350. https://doi.org/10.1126/sciadv.1602350

Cuevas, J. M., Domingo-Calap, P., & Sanjuán, R. (2012). The Fitness Effects of Synonymous Mutations in DNA and RNA Viruses. *Molecular Biology and Evolution*, *29*(1), 17–20. https://doi.org/10.1093/molbev/msr179

da Fontoura Budaszewski, R., Streck, A. F., Nunes Weber, M., Maboni Siqueira, F., Muniz Guedes, R. L., & Wageck Canal, C. (2016). Influence of vaccine strains on

the evolution of canine distemper virus. *Infection, Genetics and Evolution*, *41*, 262–269. https://doi.org/10.1016/J.MEEGID.2016.04.014

Delport, W., Poon, A. F. Y., Frost, S. D. W., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, *26*(19), 2455–2457. https://doi.org/10.1093/bioinformatics/btq429

Dietzel, E., Anderson, D. E., Castan, A., von Messling, V., & Maisner, A. (2011). Canine distemper virus matrix protein influences particle infectivity, particle composition, and envelope distribution in polarized epithelial cells and modulates virulence. *Journal of Virology*, *85*(14), 7162–8. https://doi.org/10.1128/JVI.00051-11

Dochow, M., Krumm, S. A., Crowe, J. E., Moore, M. L., Plemper, R. K., & Plemper, R. K. (2012). Independent structural domains in paramyxovirus polymerase protein. *The Journal of Biological Chemistry*, *287*(9), 6878–91. https://doi.org/10.1074/jbc.M111.325258

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–7. https://doi.org/10.1093/nar/gkh340

El Najjar, F., Schmitt, A., & Dutch, R. (2014). Paramyxovirus Glycoprotein Incorporation, Assembly and Budding: A Three Way Dance for Infectious Particle Production. *Viruses*, *6*(12), 3019–3054. https://doi.org/10.3390/v6083019

Fan, S., Deng, G., Song, J., Tian, G., Suo, Y., Jiang, Y., … Chen, H. (2009). Two amino acid residues in the matrix protein M1 contribute to the virulence difference of H5N1 avian influenza viruses in mice. *Virology*, *384*(1), 28–32. https://doi.org/10.1016/J.VIROL.2008.11.044

Han, G.-Z., Liu, X.-P., & Li, S.-S. (2008). Cross-species recombination in the haemagglutinin gene of canine distemper virus. *Virus Research*, *136*(1–2), 198–201. https://doi.org/10.1016/J.VIRUSRES.2008.04.022

Hansen, J., Qing, K., Kwon, H.-J., Mah, C., & Srivastava, A. A. (2000). Impaired Intracellular Trafficking of Adeno-Associated Virus Type 2 Vectors Limits Efficient Transduction of Murine Fibroblasts. *JOURNAL OF VIROLOGY*, *74*(2), 992–996. Retrieved from http://jvi.asm.org/content/74/2/992.full.pdf

Harder, T. C., Kenter, M., Vos, H., Siebelink, K., Huisman, W., Van Amerongen, G., … Osterhaus, A. D. M. E. (1996). Canine distemper virus from large diseased felids: biological properties and phylogenetic relationships. *Journal of General Virology*, *77*(1996), 397–405. https://doi.org/Doi 10.1099/0022-1317-77-3-397

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and Disease Emergence: Dynamics at the Wildlife-Livestock-Human Interface. *Trends in Ecology & Evolution*, *32*(1), 55–67. https://doi.org/10.1016/j.tree.2016.09.012

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

Ke, G.-M., Ho, C.-H., Chiang, M.-J., Sanno-Duanda, B., Chung, C.-S., Lin, M.-Y., … Chu, P.-Y. (2015). Phylodynamic analysis of the canine distemper virus hemagglutinin gene. *BMC Veterinary Research*, *11*(1), 164. https://doi.org/10.1186/s12917-015-0491-9

Kock, R. (2014). *The Onderstepoort journal of veterinary research. Onderstepoort Journal of Veterinary Research* (Vol. 81). OpenJournals Publishing. Retrieved from https://ojvr.org/index.php/ojvr/article/view/739/1062

Kosakovsky Pond, S. L., & Frost, S. D. W. (2005). Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Molecular Biology and Evolution*, *22*(5), 1208–1222. https://doi.org/10.1093/molbev/msi105

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*, *33*(7), 1870–4.

Liang, B., Li, Z., Jenni, S., Rahmeh, A. A., Morin, B. M., Grant, T., … Whelan, S. P. J. (2015). Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell*, *162*(2), 314–327. https://doi.org/10.1016/J.CELL.2015.06.018

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., … Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology*, *73*(1), 152–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9847317

Loots, A. K., Mitchell, E., Dalton, D. L., Kotzé, A., & Venter, E. H. (2016). Advances in canine distemper virus (CDV) pathogenesis research: a wildlife perspective. *Ournal of General Virology*. Retrieved from http://www.microbiologyresearch.org/docserver/fulltext/jgv/jgv_pap.000666.zip/jgv.0.000666.pdf?expires=1490811416&id=id&accname=guest&checksum=D3E1FA7FBFAD5C6F366F82329E9089D1

Marescot, L., Benhaiem, S., Gimenez, O., Hofer, H., Lebreton, J.-D., Olarte-Castillo, X. A., … East, M. L. (2018). Social status mediates the fitness costs of infection with

canine distemper virus in Serengeti spotted hyenas. *Functional Ecology*. https://doi.org/10.1111/1365-2435.13059

Martin, D., & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *BIOINFORMATICS APPLICATIONS NOTE*, *16*(6), 562–563. Retrieved from http://

Martinez-gutierrez, M., & Ruiz-saenz, J. (2016). Diversity of susceptible hosts in canine distemper virus infection : a systematic review and data synthesis. *BMC Veterinary Research*, 1–11. https://doi.org/10.1186/s12917-016-0702-z

McCarthy, A. J., Shaw, M.-A., & Goodman, S. J. (2007). Pathogen evolution and disease emergence in carnivores. *Proceedings. Biological Sciences / The Royal Society*, *274*(1629), 3165–74. https://doi.org/10.1098/rspb.2007.0884

Meli, M., Simmler, P., Cattori, V. M., Fernando Vargas, Astrid Palomares, Fransisco Lopez-Bao, … Lutz, H. (2010). Importance of canine distemper virus (CDV) infection in free-ranging Iberian lynxes (Lynx pardinus). *Veterinary Microbiology*, *146*(1–2), 132–137. https://doi.org/10.1016/J.VETMIC.2010.04.024

Metzker, M. L., Thomas, T., Luciani, F., Metzker, M., Esteller, M., Skalsky, R., … Turnbaugh, P. (2010). Sequencing technologies — the next generation. *Nature Reviews Genetics*, *11*(1), 31–46. https://doi.org/10.1038/nrg2626

Munson, L., Terio, K. A., Kock, R., Mlengeya, T., Roelke, M. E., Dubovi, E., … Packer, C. (2008). Climate Extremes Promote Fatal Co-Infections during Canine Distemper Epidemics in African Lions. *PLoS ONE*, *3*(6), e2545. https://doi.org/10.1371/journal.pone.0002545

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond,

S. L. (2012). Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genetics*, *8*(7), e1002764. https://doi.org/10.1371/journal.pgen.1002764

Nagao, Y., Nishio, Y., Shiomoda, H., Tamaru, S., Shimojima, M., Goto, M., … Maeda, K. (2012). An Outbreak of Canine Distemper Virus in Tigers (Panthera tigris): Possible Transmission from Wild Animals to Zoo Animals. *J. Vet. Med. Sci*, *74*(6), 699–705. https://doi.org/10.1292/jvms.11-0509

Neumann, G., & Kawaoka, Y. (2006). Host Range Restriction and Pathogenicity in the Context of Influenza Pandemic. *Emerging Infectious Diseases*, *12*(6), 881–886. https://doi.org/10.3201/eid1206.051336

Nikolin, V. M., Olarte-Castillo, X. A., Osterrieder, N., Hofer, H., Dubovi, E., Mazzoni, C. J., … East, M. L. (2017). Canine distemper virus in the Serengeti ecosystem: molecular adaptation to different carnivore species. *Molecular Ecology*. https://doi.org/10.1111/mec.13902

Nikolin, V. M., Osterrieder, K., von Messling, V., Hofer, H., Anderson, D., Dubovi, E., … East, M. L. (2012a). Antagonistic Pleiotropy and Fitness Trade-Offs Reveal Specialist and Generalist Traits in Strains of Canine Distemper Virus. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0050955

Nikolin, V. M., Wibbelt, G., Michler, F.-U. F., Wolf, P., & East, M. L. (2012b). Susceptibility of carnivore hosts to strains of canine distemper virus from distinct genetic lineages. *Veterinary Microbiology*, *156*(1–2), 45–53. https://doi.org/10.1016/j.vetmic.2011.10.009

Novella, I. S., Zárate, S., Metzgar, D., & Ebendick-Corpus, B. E. (2004). Positive Selection of Synonymous Mutations in Vesicular Stomatitis Virus. *Journal of*

Molecular Biology, *342*(5), 1415–1421. https://doi.org/10.1016/J.JMB.2004.08.003

Patz, J. A., Graczyk, T. K., Geller, N., & Vittor, A. Y. (2000). Effects of environmental change on emerging parasitic diseases. *International Journal for Parasitology*, *30*(12–13), 1395–1405. https://doi.org/10.1016/S0020-7519(00)00141-7

Pepin, K. M., Domsic, J., & Mckenna, R. (2008). Genomic evolution in a virus under specific selection for host recognition. *Infection, Genetics and Evolution*, *8*, 825–834. https://doi.org/10.1016/j.meegid.2008.08.008

Pepin, K. M., Lass, S., C Pulliam, J. R., Read, A. F., & Lloyd-Smith, J. O. (2010). Identifying genetic markers of adaptation for surveillance of viral host jumps. *Nature Publishing Group*, *8*. https://doi.org/10.1038/nrmicro2440

Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, *21*(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Pope, J. P., Miller, D. L., Riley, M. C., Anis, E., & Wilkes, R. P. (2016). Characterization of a novel Canine distemper virus causing disease in wildlife. *Journal of Veterinary Diagnostic Investigation*. https://doi.org/10.1177/1040638716656025

Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, *25*(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Raha, T., Kaushik, R., & Shaila, M. S. (2004). Phosphoprotein P of Rinderpest virus binds to plus sense leader RNA: regulation by phosphorylation. *Virus Research*, *104*(2), 191–200. https://doi.org/10.1016/J.VIRUSRES.2004.04.004

Riley, M. C., & Wilkes, R. P. (2015). Sequencing of emerging canine distemper virus strain reveals new distinct genetic lineage in the United States associated with

disease in wildlife and domestic canine populations. *Virology Journal*, *12*, 219. https://doi.org/10.1186/s12985-015-0445-7

Roelke-Parker, M. E., Munson, L., Packer, C., Kockil, R., CIeaveland, S., Carpenter, M., … G AppeII, M. J. (1996). A canine distemper virus epidemic in Serengeti lions (Panthera leo). *Nature*, *379*, 441–445.

Seimon, T. A., Miquelle, D. G., Chang, T. Y., Newton, A. L., Korotkova, I., Ivanchuk, G., … McAloose, D. (2013). Canine distemper virus: an emerging disease in wild endangered Amur tigers (Panthera tigris altaica). *mBio*, *4*(4), e00410-13. https://doi.org/10.1128/mBio.00410-13

Sidhu, M. S., Menonna, J. P., Cook, S. D., Dowling, P. C., & Udem, S. A. (1993). Canine Distemper Virus L Gene: Sequence and Comparison with Related Viruses. *Virology*, *193*(1), 50–65. https://doi.org/10.1006/VIRO.1993.1102

Sourimant, J., & Plemper, R. K. (2016). Organization, Function, and Therapeutic Targeting of the Morbillivirus RNA-Dependent RNA Polymerase Complex. *Viruses*, *8*(9). https://doi.org/10.3390/v8090251

Sulikhan, N. S., Gilbert, M., Yu Blidchenko, E., Naidenko, S. V, Ivanchuk, G. V, Yu Gorpenchenko, T., … Seimon, T. A. (2018). Canine Distemper Virus in a Wild Far Eastern Leopard Panthera pardus orientalis. *Journal of Wildlife Diseases Wildlife Disease Association*, *54*(1), 0–0. https://doi.org/10.7589/2017-03-065

Terio, K. A., & Craft, M. E. (2013). Canine Distemper Virus (CDV) in Another Big Cat: Should CDV Be Renamed Carnivore Distemper Virus? *mBio*, *4*(5), e00702-13-e00702-13. https://doi.org/10.1128/mBio.00702-13

Thakkar, V. D., Cox, R. M., Sawatsky, B., da Fontoura Budaszewski, R., Sourimant, J.,

Wabbel, K., … Plemper, R. K. (2018). The Unstructured Paramyxovirus Nucleocapsid Protein Tail Domain Modulates Viral Pathogenesis through Regulation of Transcriptase Activity. *Journal of Virology*, *92*(8), e02064-17. https://doi.org/10.1128/JVI.02064-17

von Messling, V., Svitek, N., & Cattaneo, R. (2006). Receptor (SLAM [CD150]) recognition and the V protein sustain swift lymphocyte-based invasion of mucosal tissue and lymphatic organs by a morbillivirus. *Journal of Virology*, *80*(12), 6084–92. https://doi.org/10.1128/JVI.00357-06

Wasik, B. R., Muñoz-Rojas, A. R., Okamoto, K. W., Miller-Jensen, K., & Turner, P. E. (2016). Generalized selection to overcome innate immunity selects for host breadth in an RNA virus. *Evolution*, *70*(2), 270–281. https://doi.org/10.1111/evo.12845

Yuan, C., Liu, W., Wang, Y., Hou, J., Zhang, L., & Wang, G. (2017). Homologous recombination is a force in the evolution of canine distemper virus. *PLOS ONE*, *12*(4), e0175416. https://doi.org/10.1371/journal.pone.0175416

Zhang, H., Shan, F., Zhou, X., Li, B., Zhai, J.-Q., Zou, S.-Z., … Luo, M.-L. (2017). Outbreak and genotyping of canine distemper virus in captive Siberian tigers and red pandas. *Scientific Reports*, *7*(1), 8132. https://doi.org/10.1038/s41598-017-08462-4

Table 4.1. Mutations at nucleotides in the CDV genome sequence that differentiate the Fatal and Silent strains of CDV in circulation in East Africa between 1993-2011.

Position (nt)	Gene	Amino acid	Silent	Fatal	Synonymous or Nonsynonymous	AA-Silent	AA-Fatal
132	N	9	A	G	non	T	A
470	N	121	A	G	syn	-	-
1361	***N***	***418***	***T***	***C***	***syn***	***-***	***-***
1382	**N**	**425**	**T**	**C**	**syn**	**-**	**-**
1458	N	451	T	C	non	F	L
1505	N	466	C	A	syn	-	-
1914	P/V/C	38/38/31	A	C	non/non/non	Q/Q/K	H/H/T
2200	***P/V/C***	***134/134/126***	***G***	***A***	***non/non/syn***	***G/G/-***	***S/S/-***
2638	***P/V***	***280/280***	***A***	***G***	***non***	***K/K***	***E/R***
2832	P	344	A/G	C	syn	-	-
2934	P	378	G	A	syn	-	-
3292	P	498	T	C	non	Y	H
3410	P-M UTR	n.a.	A	T	n.a.	-	-
3457	M	9	A	G	non	Q	R
3461	M	10	T	C	syn	-	-
3821	M	130	G	A	syn	-	-
4436	M	335	C	T	syn	-	-
4900	M-F UTR	n.a.	A	G	n.a.	-	-
5001	FSP	23	C	T	non	H	Y
5226	F	98	G	A	non	A	T

5978	F	348	T	C	syn	-	-
6058	F	375	G	A	non	R	Q
6694	F	587	T	C	non	V	A
6710	F	592	T	C	syn		
6905	F	657	C	A	syn	-	-
6974	F-H UTR	n.a.	T	C	n.a.	-	-
7088	H	4	T	C	non	Y	H
7167	H	30	A	G	non	Q	R
7870	H	264	A	T	syn	-	-
8320	H	414	T	C	syn	-	-
8634	H	519	G	T	non	R	I
8723	***H***	***549***	***T***	***C***	***non***	***Y***	***H***
9210	L	61	A	C	non	M	L
9290	L	87	C	T	syn	-	-
9308	***L***	***93***	***C***	***T***	***non***	***L***	***F***
9428	L	133	T	C	syn	-	-
10844	L	605	T/C	G	syn	-	-
11159	***L***	***710***	***C***	***T***	***syn***	***-***	***-***
12170	***L***	***1047***	***T***	***C***	***syn***	***-***	***-***
12986	***L***	***1319***	***T***	***C***	***syn***	***-***	***-***
13142	L	1371	A	G	syn	-	-
13233	***L***	***1402***	***A***	***C***	***non***	***-***	***-***
13524	L	1499	A	G	non	I	V
13886	L	1619	C	T	syn	-	-
14132	L	1701	A	G	syn	-	-
15056	L	2009	T	C	syn	-	-

15185	L	2052	C	T	syn	-	-
15203	L	2058	A	C	syn	-	-
15326	L	2099	T	A	syn	-	-

* Mutation occurs in overlapping open reading frames.

Table 4.2. Recombination events detected in alignment of 126 whole CDV genome sequences using seven algorithms in the Recombination Detection Program (RDP).

Acc. no.	#	Species	Country	Year	Clade	RDP pval	GC pval	BS pval	MC pval	Ch pval	SS pval	3Seq pval
AB474397	1	Dog	Japan	nd	Asia-2	2.07E-27	1.02E-25	1.01E-22	4.88E-08	8.66E-08	1.84E-04	1.38E-11
AB474397	2	Dog	Japan	nd	Asia-2	2.18E-27	2.85E-26	3.27E-20	7.19E-07	7.04E-07	3.18E-06	1.38E-11
AB462810	1	Dog	Japan	nd	Asia-2	2.07E-27	1.02E-25	1.01E-22	4.88E-08	8.66E-08	1.84E-04	1.38E-11
AB462810	2	Dog	Japan	nd	Asia-2	2.18E-27	2.85E-26	3.27E-20	7.19E-07	7.04E-07	3.18E-06	1.38E-11
AY443350	1	Raccoon	USA	2000	Amer-2	2.22E-12	3.32E-18	2.88E-06	1.05E-11	7.38E-13	3.96E-16	1.09E-16
AY443350	2	Raccoon	USA	2000	Amer-2	1.43E-12	2.65E-13	5.67E-07	2.18E-03	1.80E-03	none	1.65E-08
AY445077	1	Raccoon	USA	1998	Amer-1*	9.19E-07	9.92E-07	1.03E-03	4.11E-04	2.96E-03	none	2.99E-05
AY445077	2	Raccoon	USA	1998	Amer-1*	1.45E-18	3.69E-12	1.32E-06	6.67E-06	1.62E-06	1.16E-10	6.91E-12
AY466011	1	Raccoon	USA	1998	Amer-1*	9.19E-07	9.92E-07	1.03E-03	4.11E-04	2.96E-03	none	2.99E-05
AY466011	2	Raccoon	USA	1998	Amer-1*	1.45E-18	3.69E-12	1.32E-06	6.67E-06	1.62E-06	1.16E-10	6.91E-12
AY542312	1	Raccoon	USA	1998	Amer-1*	9.19E-07	9.92E-07	1.03E-03	4.11E-04	2.96E-03	none	2.99E-05
AY542312	2	Raccoon	USA	1998	Amer-1*	1.45E-18	3.69E-12	1.32E-06	6.67E-06	1.62E-06	1.16E-10	6.91E-12
AY649446	1	Raccoon	USA	2001	Amer-2	1.43E-12	2.65E-13	5.67E-07	2.18E-03	1.80E-03	none	1.65E-08
AY649446	2	Raccoon	USA	2001	Amer-2	2.44E-12	6.73E-03	2.87E-07	7.13E-16	1.21E-04	5.04E-07	6.91E-12
EU716337	1	Dog	USA	2004	Amer-2	9.04E-03	none	none	7.57E-05	3.65E-05	9.49E-04	1.29E-05
JX681125	1	Fox	China	2006	Asia-1	4.05E-71	4.11E-69	1.93E-63	5.05E-18	2.21E-17	2.84E-17	2.07E-11
KJ123771	1	Dog	USA	2004	Amer-2	none	1.90E-04	4.11E-05	2.33E-07	2.23E-08	2.49E-15	2.69E-10
KJ994343	1	Raccoon dog	China	2013	Asia-1	6.89E-33	1.78E-31	3.93E-27	1.49E-05	1.18E-05	3.77E-06	2.07E-11
KJ994343	2	Raccoon dog	China	2013	Asia-1	1.05E-09	4.02E-02	7.44E-09	2.55E-07	1.07E-07	1.44E-05	9.67E-06
KJ994343	3	Raccoon dog	China	2013	Asia-1	1.22E-06	6.09E-05	4.47E-05	4.39E-02	3.51E-02	none	9.82E-06
KJ994343	4	Raccoon dog	China	2013	Asia-1	5.31E-05	5.72E-03	1.79E-06	3.96E-02	9.71E-03	1.33E-03	1.45E-04

* Lineage is conventionally determined by hemagglutinin (H) gene diversity. In this case the H gene sequence was the recombinant portion of the genome and thus defined the genotype, however, the majority of genome is America-2.Abbreviations: Acc. no. = Accession number, # = recombination event per sequence, GC = GENECONV, BS = BootScan, MC = MaxChi, Ch = Chimaera, SS = SiScan

Table 4.3. Genomic positions of adaptive genetic markers and associated functional domains

Position (nt)	Gene	AA	Silent aa	Fatal aa	Fatal non-canid only	Functional domains	Domain function	Virus process
1361	N	418	-	-	Yes	N-tail, box 1	Binds unidentified host receptor suppressing immune cell proliferation	Transcription & replication
1458	N	451	F	L	No	N-tail	Binds P recruiting RdRp complex to prevent slipping during transcription & replication	Transcription & replication
1505	N	466	-	-	No	N-tail	Binds P recruiting RdRp complex to prevent slipping during transcription & replication	Transcription & replication
2638	P/V	280	K	E/R	Yes	PNT / V-Zbd	P - binds L - essential co-factor of transcription and replication / V - interrupts host innate immune response, virulence factor	Transcription & replication
2832	P	344	-	-	No	PMD	L binding site - essential co-factor of transcription and replication	Transcription & replication
2934	P	378	-	-	No	PMD	L binding site - essential co-factor of transcription and replication	Transcription & replication
3457	M	9	Q	R	No	M - NTD	Matrix general - modulates fusion and budding, binds F, H, RNP, and cellular actin, assembles components at surface	Assembly & fusion
3461	M	10	-	-	No	M - NTD	Matrix general - modulates fusion and budding, binds F, H, RNP, and cellular actin, assembles components at surface	Assembly & fusion

3821	M	130	-	-	No	M - NTD	Matrix general - modulates fusion and budding, binds F, H, RNP, and cellular actin, assembles components at surface	Assembly & fusion
4436	M	335	-	-	No	M - CTD	Matrix general - modulates fusion and budding, binds F, H, RNP, and cellular actin, assembles components at surface	Assembly & fusion
5978	F	348	-	-	No	F1	Extracellular domain of glycoprotein, mediates fusion with host cell, interacts with H	Fusion
6905	F	657	-	-	No	F cytoplasmic tail	Cytoplasmic tail interacts with M at cell surface during fusion and budding	Assembly & fusion
7870	H	264	-	-	No	no data	No data	Host receptor binding
8723	H	549	Y	H	Yes	SLAM binding	Host immune cell recognition and entry	Host receptor binding
9290*	L	87	-	-	No	LRI	No data	Transcription & replication
9308	L	93	L	F	Yes**	LRI	No data	Transcription & replication
9428	L	133	-	-	No	LRI	No data	Transcription & replication
10844	L	605	-	-	No	CRII	Polymerase activity	Transcription & replication
11159	L	**710**	-	-	No	CRIII	Catalytic center for polymerization, i.e. transcription and replication of vRNA	Transcription & replication
12170	L	**1047**	-	-	Yes	CRIV	mRNA capping, essential for translation and immune escape	Transcription & replication
12986	L	**1319**	-	-	Yes	CRV	Methylation of viral mRNA	Transcription & replication
13233	L	1402	N	H	Yes**	Connector	spacing the catalytic domains, may interact with P to stabilize conformation of RdRp complex	Transcription & replication

13886	L	1619	-	-	No***	Connector	Spacing the catalytic domains, may interact with P to stabilize conformation of RdRp complex	Transcription & replication
14132	L	1701	-	-	No	Linker	Separates regions of the L	Transcription & replication
15203	L	2058	A	C	Yes**	Mtase	Methylation of viral mRNA	Transcription & replication
15326	L	2099	-	-	No	Mtase	Methylation of viral mRNA	Transcription & replication

* Purifying selection across tree, but amino acid change in 1994 non-canid strain
** Differentiates canids and non-canids with exception of PLE-658
***Mutation occurred in the Silent strain

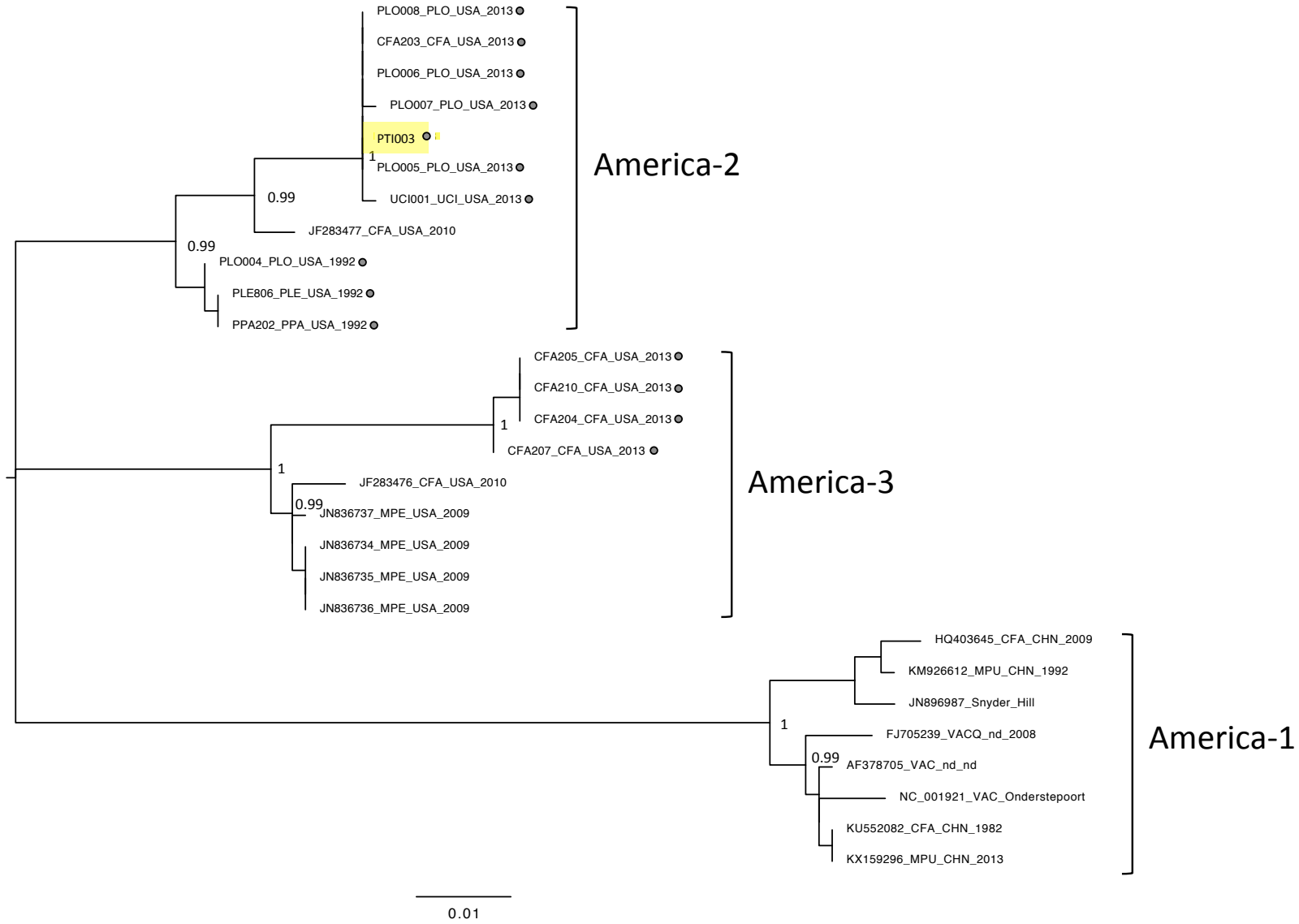

PLO008_PLO_USA_2013
CFA203_CFA_USA_2013
PLO006_PLO_USA_2013
PLO007_PLO_USA_2013
PTI003
PLO005_PLO_USA_2013
UCI001_UCI_USA_2013
JF283477_CFA_USA_2010
PLO004_PLO_USA_1992
PLE806_PLE_USA_1992
PPA202_PPA_USA_1992
America-2
CFA205_CFA_USA_2013
CFA210_CFA_USA_2013
CFA204_CFA_USA_2013
CFA207_CFA_USA_2013
JF283476_CFA_USA_2010
JN836737_MPE_USA_2009
JN836734_MPE_USA_2009
JN836735_MPE_USA_2009
JN836736_MPE_USA_2009
America-3
HQ403645_CFA_CHN_2009
KM926612_MPU_CHN_1992
JN896987_Snyder_Hill
FJ705239_VACQ_nd_2008
AF378705_VAC_nd_nd
NC_001921_VAC_Onderstepoort
KU552082_CFA_CHN_1982
KX159296_MPU_CHN_2013
America-1
1
0.99
1
0.99
1
0.99
0.01

Figure 4.1. Maximum likelihood (ML) phylogenetic tree of partial H gene sequences (718 bp) shows relationship of CDV sequence from recovered tiger (in yellow) sampled at captive facility where African lions were clinically infected. Circles at tree tips indicate sequences generated in this study. Major recognized clads indicated with brackets. Bootstrap support indicated at nodes.

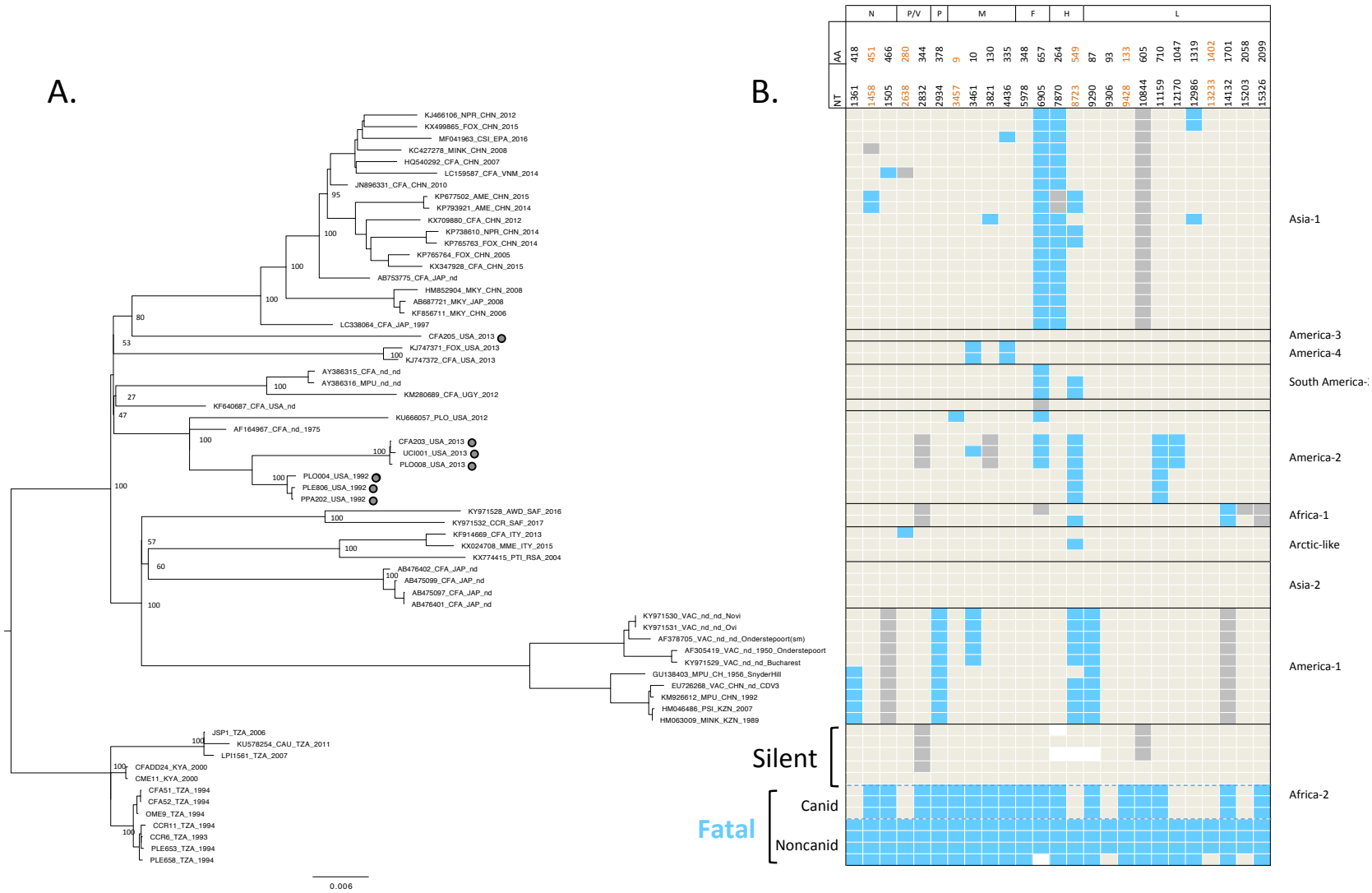
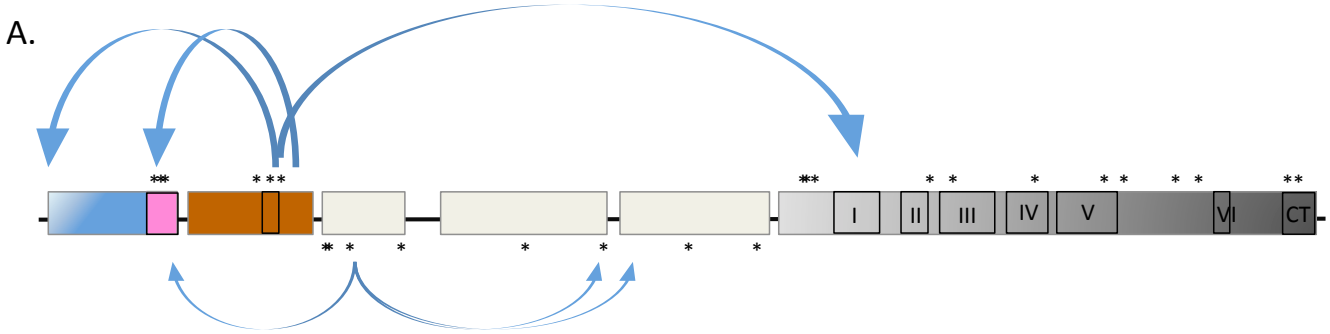

A.
KJ466106_NPR_CHN_2012
KX499865_FOX_CHN_2015
MF041963_CSI_EPA_2016
KC427278_MINK_CHN_2008
HQ540292_CFA_CHN_2007
LC159587_CFA_VNM_2014
JN896331_CFA_CHN_2010
KP677502_AME_CHN_2015
KP793921_AME_CHN_2014
KX709880_CFA_CHN_2012
KP738610_NPR_CHN_2014
KP765763_FOX_CHN_2014
KP765764_FOX_CHN_2005
KX347928_CFA_CHN_2015
AB753775_CFA_JAP_nd
HM852904_MKY_CHN_2008
AB687721_MKY_JAP_2008
KF856711_MKY_CHN_2006
LC338064_CFA_JAP_1997
CFA205_USA_2013
KJ747371_FOX_USA_2013
KJ747372_CFA_USA_2013
AY386315_CFA_nd_nd
AY386316_MPU_nd_nd
KM280689_CFA_UGY_2012
KF640687_CFA_USA_nd
KU666057_PLO_USA_2012
AF164967_CFA_nd_1975
CFA203_USA_2013
UCI001_USA_2013
PLO008_USA_2013
PLO004_USA_1992
PLE806_USA_1992
PPA202_USA_1992
KY971528_AWD_SAF_2016
KY971532_CCR_SAF_2017
KF914669_CFA_ITY_2013
KX024708_MME_ITY_2015
KX774415_PTI_RSA_2004
AB476402_CFA_JAP_nd
AB475099_CFA_JAP_nd
AB475097_CFA_JAP_nd
AB476401_CFA_JAP_nd
KY971530_VAC_nd_nd_Novi
KY971531_VAC_nd_nd_Ovi
AF378705_VAC_nd_nd_Onderstepoort(sm)
AF305419_VAC_nd_1950_Onderstepoort
KY971529_VAC_nd_nd_Bucharest
GU138403_MPU_CH_1956_SnyderHill
EU726268_VAC_CHN_nd_CDV3
KM926612_MPU_CHN_1992
HM046486_PSI_KZN_2007
HM063009_MINK_KZN_1989
JSP1_TZA_2006
KU578254_CAU_TZA_2011
LPI1561_TZA_2007
CFADD24_KYA_2000
CME11_KYA_2000
CFA51_TZA_1994
CFA52_TZA_1994
OME9_TZA_1994
CCR11_TZA_1994
CCR6_TZA_1993
PLE653_TZA_1994
PLE658_TZA_1994
0.006
B.
N
P/V
P
M
F
H
L
AA
NT
418 451 466 280 344 378 9 10 130 335 348 657 264 549 87 93 133 605 710 1047 1319 1402 1701 2058 2099
1361 1458 1505 2638 2832 2934 3457 3461 3821 4436 5978 6905 7870 8723 9290 9306 9428 10844 11159 12170 12986 13233 14132 15203 15326
Asia-1
America-3
America-4
South America-1
America-2
Africa-1
Arctic-like
Asia-2
America-1
Silent
Africa-2
Fatal
Canid
Noncanid

Figure 4.2. Phylogenetic relationships among globally distributed CDV and distribution of possible mutations correlated with pathogenicity in African lions. A) Maximum likelihood (ML) phylogenetic tree of whole CDV genomes generated in this study with publically available sequences collected globally (15,584 bp). ML tree is rooted on the Africa-2 clade for clarity. Major recognized clades indicated in right margin. Black circles at tree tips indicate sequences generated in this study. B) Mutations at sites experiencing non-neutral evolution ($\omega \neq 1$) differentiating strains in circulation during a lethal outbreak from those in circulation during silent outbreaks in East Africa are mapped to tips. Nucleotide and amino acid sites per gene are indicated in the table header, orange font = positive selection, black = purifying selection. Nucleotide or amino acid at a position per sequence is indicated by a colored box: blue = consensus genotype in non-canids infected by Fatal strain (reference allele) , taupe = alternative allele 1, grey = alternative allele 2, white = no data. (Thus a blue box indicates that the sequence shares the same allele at this position as the Fatal strain, and a gray box indicates that the sequence shares a mutation at the same site as the Fatal strain though it has a different allele.)



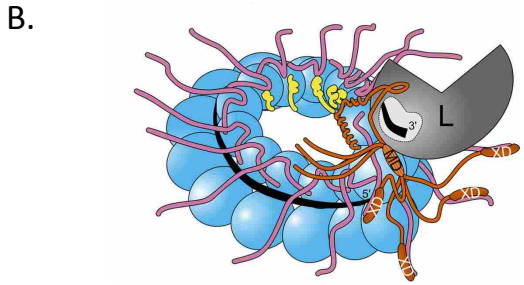


Figure 4.3. Adaptive mutations associated with pathogenicity in African lions and their relationship to functional domains of the Canine distemper virus genome. A) Model of CDV gene organization – N, P/V/C, M, F, H, and L, from left to right. Colored blocks

highlight the genes of the RdRp complex – N (core = blue, tail = pink), P (orange), and L (grey-black). Mutations related to the RdRp complex are shown above, and those related to assembly, binding and fusion shown below their respective genes. Black boxes indicate conserved regions of interest, on L these are implicated in enzymatic activities. Blue arrows indicate interacting gene regions. B) Model of nonspecific binding of the RdRp complex to the nucleocapsid which encapsidates the viral RNA (black ribbon) initiating polymerization for transcription or replication. Figure reproduced from (Sourimant & Plemper, 2016).

Table S4.1. List of sequences generated in this study and associated data.

Accession no.	Label	Origin country	Date sampled	Year sampled	Specimen origin	Tissue type
Pending submission	PLE806_USA_1992	USA	9/1/92	1992	Isolate	Cell culture supernatant
Pending submission	PLO004_USA_1992	USA	01/01/1992	1992	Isolate	Cell culture supernatant
Pending submission	PLO005_USA_2013	USA	03/02/2013	2013	Clinical	Brain
Pending submission	PLO006_USA_2013	USA	03/13/2013	2013	Clinical	Brain
Pending submission	PLO007_USA_2013	USA	03/26/2013	2013	Clinical	Brain
Pending submission	PLO008_USA_2013	USA	04/09/2013	2013	Clinical	Brain
Pending submission	PPA202_USA_1992	USA	10/15/1992	1992	Isolate	Cell culture supernatant
Pending submission	UCI001_USA_2013	USA	03/29/2013	2013	Clinical	Brain
Pending submission	CFA203_USA_2013	USA	01/22/2013	2013	Clinical	No data
Pending submission	CFA204_USA_2013	USA	06/24/2013	2013	Clinical	No data
Pending submission	CFA205_USA_2013	USA	06/07/2013	2013	Clinical	No data
Pending submission	CFA207_USA_2013	USA	11/12/2013	2013	Clinical	No data
Pending submission	CFA210_USA_2013	USA	11/12/12	2013	Clinical	cDNA

Table S4.2. Accession numbers and associated data used in this study.

Accession no	LABEL	Species	Origin country	Date sampled	Year sampled
AB462810	AB462810_CFA_JAP_nd	CFA	JAP	No data	No data
AB474397	AB474397_CFA_JAP_nd	CFA	JAP	No data	No data
AB475097	AB475097_CFA_JAP_nd	CFA	JAP	No data	No data
AB475099	AB475099_CFA_JAP_nd	CFA	JAP	No data	No data
AB476401	AB476401_CFA_JAP_nd	CFA	JAP	No data	No data
AB476402	AB476402_CFA_JAP_nd	CFA	JAP	No data	No data
AB476402	AB476402_CFA_JAP_nd	CFA	JAP	No data	No data
AB490670	AB490670_CFA_JAP_nd	CFA	JAP	No data	No data
AB490672	AB490672_CFA_JAP_nd	CFA	JAP	No data	No data
AB490674	AB490674_CFA_JAP_nd	CFA	JAP	No data	No data
AB490676	AB490676_CFA_JAP_nd	CFA	JAP	No data	No data
AB490678	AB490678_CFA_JAP_nd	CFA	JAP	No data	No data
AB490679	AB490679_CFA_JAP_nd	CFA	JAP	No data	No data
AB490680	AB490680_CFA_JAP_nd	CFA	JAP	No data	No data
AB490681	AB490681_CFA_JAP_nd	CFA	JAP	No data	No data
AB687720	AB687720 _MKY_JAP_2008	MKY	JAP	7/1/08	2008
AB687721	AB687721_MKY_JAP_2008	MKY	JAP	7/1/08	2008
AB753775	AB753775_CFA_JAP_nd	CFA	No data	No data	No data
AB753776	AB753776_CFA_nd_nd	CFA	No data	No data	No data
AB823706	AB823706_CFA_nd_nd	CFA	No data	No data	No data

AB823707	AB823707_CFA_nd_nd	CFA	No data	No data	No data
AF014953	AF014953_VAC_nd_nd	VAC	No data	No data	No data
AF164967	AF164967_CFA_nd_1975	CFA	No data	No data	No data
AF305419	AF305419_VAC_nd_1950_Onderstepoort	VAC	No data	No data	No data
AF378705	AF378705_VAC_nd_nd_Onderstepoort(sm)	VAC	No data	No data	No data
AY386315	AY386315_CFA_nd_nd	CFA	No data	No data	No data
AY386316	AY386316_MPU_nd_nd	MPU	No data	No data	No data
AY443350	AY443350_PLO_USA_2000	PLO	USA	1/1/00	2000
AY445077	AY445077_PLO_USA_1998	PLO	USA	1/1/98	1998
AY466011	AY466011_PLO_USA_1998	PLO	USA	1/1/98	1998
AY542312	AY542312_PLO_USA_1998	PLO	USA	1/1/98	1998
AY649446	AY649446_PLO_USA_2001	PLO	USA	1/1/01	2001
EU716337	EU716337_CFA_USA_2004	CFA	USA	9/4/04	2004
EU726268	EU726268_VAC_CHN_nd_CDV3	MINK	CHN	No data	No data
GU138403	GU138403_MPU_CH_1956_SnyderHill	MPU	No data	No data	No data
HM046486	HM046486_PSI_KZN_2007	PSI	KZN	1/1/07	2007
HM063009	HM063009_MINK_KZN_1989	MINK	KZN	1/1/89	1989
HM852904	HM852904_MKY_CHN_2008	MKY	CHN	1/1/08	2008
HQ540292	HQ540292_CFA_CHN_2007	CFA	CHN	No data	2007
HQ540293	HQ540293_FOX_CHN_2006	FOX	CHN	No data	2006
JN896331	JN896331_CFA_CHN_2010	CFA	CHN	1/1/10	2010
JN896987	JN896987_Snyder_nd_nd	Snyder	No data	No data	No data
JX681125	JX681125_FOX_CHN_2006	FOX	CHN	7/1/06	2006

KC427278	KC427278_MINK_CHN_2008	MINK	CHN	7/6/08	2008
KF640687	KF640687_CFA_USA_nd	CFA	USA	No data	No data
KF856711	KF856711_MKY_CHN_2006	MKY	CHN	1/1/06	2006
KF914669	KF914669_CFA_ITY_2013	CFA	ITY	5/15/13	2013
KJ123771	KJ123771_CFA_USA_2004	CFA	USA	1/1/04	2004
KJ466106	KJ466106_NPR_CHN_2012	NPR	CHN	1/1/12	2012
KJ747371	KJ747371_FOX_USA_2013	FOX	USA	1/1/13	2013
KJ747372	KJ747372_CFA_USA_2013	CFA	USA	1/1/13	2013
KJ848781	KJ848781_NPR_CAN_2014	NPR	CAN	4/1/14	2014
KJ994343	KJ994343_NPR_CAN_2013	NPR	CAN	10/1/13	2013
KM280689	KM280689_CFA_UGY_2012	CFA	UGY	1/1/12	2012
KM926612	KM926612_MPU_CHN_1992	MPU	CHN	1/1/92	1992
KP677502	KP677502_AME_CHN_2015	AME	CHN	1/4/15	2015
KP738610	KP738610_NPR_CHN_2014	NPR	CAN	10/1/14	2014
KP765763	KP765763_FOX_CHN_2014	FOX	CAN	7/3/14	2014
KP765764	KP765764_FOX_CHN_2005	FOX	CHN	7/2/05	2005
KP793921	KP793921_AME_CHN_2014	AME	CHN	7/6/05	2014
KU578253	KU578253_AWD_TZA_2007	AWD	TZA	1/1/07	2007
KU578254	KU578254_CAU_TZA_2011	CAU	TZA	1/1/11	2011
KU578255	KU578255_CCR_TZA_1994	CCR	TZA	1/1/94	1994
KU578256	KU578256_PLE_TZA_1994	PLE	TZA	1/1/94	1994
KU578257	KU578257_CFA_TZA_1994	CFA	TZA	1/1/94	1994
KU666057	KU666057_PLO_USA_2012	PLO	USA	11/24/12	2012
KX024708	KX024708_MME_ITY_2015	MME	ITY	9/28/15	2015
KX024709	KX024709_MME_ITY_2015	MME	ITY	10/5/15	2015
KX347928	KX347928_CFA_CHN_2015	CFA	CHN	10/10/15	2015

KX499865	KX499865_FOX_CHN_2015	FOX	CHN	7/7/05	2015
KX709880	KX709880_CFA_CHN_2012	CFA	CHN	7/18/12	2012
KX774415*	KX774415_PTI_RSA_2004	PTI	RSA	1/1/04	2004
KY971528	KY971528_AWD_SAF_2016	AWD	SAF	5/1/16	2016
KY971529	KY971529_VAC_nd_nd_Bucharest	VAC	No data	No data	No data
KY971530	KY971530_VAC_nd_nd_'Novi'	VAC	No data	No data	No data
KY971531	KY971531_VAC_nd_nd_Ovi	VAC	No data	No data	No data
KY971532	KY971532_CCR_SAF_2017	CCR	SAF	1/1/17	2017
LC159587	LC159587_CFA_VNM_2014	CFA	VNM	8/16/14	2014
MF041963	MF041963_CSI_EPA_2016	CSI	EPA	9/30/16	2016
NC_001921	NC_001921_VAC_nd_nd	VAC	No data	No data	No data
LC338064	LC338064_CFA_JAP_1997	CFA	JAP	No data	1997
CH2-Pending submission	CCR10_TZA_1994	CCR	TZA	06/20/1994	1994
CH2-Pending submission	CCR11BH_TZA_1994	CCR	TZA	07/03/1994	1994
CH2-Pending submission	CCR12_TZA_1994	CCR	TZA	07/09/1994	1994
CH2-Pending submission	CCR282_TZA_1994	CCR	TZA	01/18/1994	1994
CH2-Pending submission	CCR6_TZA_1993	CCR	TZA	12/20/1993	1994
CH2-Pending submission	CCR7BR_TZA_1994	CCR	TZA	12/23/1993	1994
CH2-Pending submission	CFA51_TZA_1994	CFA	TZA	09/09/1994	1994
CH2-Pending submission	CFA52_TZA_1994	CFA	TZA	08/23/1994	1994
CH3-Pending submission	CFA54_TZA_1994	CFA	TZA	11/18/1994	1994
CH3-Pending submission	CFADD19_KYA_2000	CFA	KYA	02/28/2000	2000
CH3-Pending submission	CFADD24_KYA_2000	CFA	KYA	02/28/2000	2000
CH3-Pending submission	CME10_KYA_2000	JSP	KYA	03/06/2000	2000
CH3-Pending submission	CME11_KYA_2000	JSP	KYA	03/03/2000	2000
CH3-Pending submission	CME16_KYA_2000	JSP	KYA	03/17/2000	2000

CH3-Pending submission	JSP1_TZA_2006	JSP	TZA	07/20/2006	2006
CH3-Pending submission	LPI1561LU_TZA_2007	LPI	TZA	10/03/2007	2007
CH3-Pending submission	LPI1561_TZA_2007	LPI	TZA	10/03/2007	2007
CH2-Pending submission	OME8SP_TZA_1994	OME	TZA	07/16/1994	1994
CH2-Pending submission	OME9_TZA_1994	OME	TZA	07/25/1994	1994
CH2-Pending submission	PLE589_TZA_1994	PLE	TZA	01/22/1994	1994
CH2-Pending submission	PLE595_TZA_1994	PLE	TZA	02/03/1994	1994
CH2-Pending submission	PLE635LN_TZA_1994	PLE	TZA	02/18/1994	1994
CH2-Pending submission	PLE641_TZA_1994	PLE	TZA	11/15/1994	1994
CH2-Pending submission	PLE652_TZA_1994	PLE	TZA	07/30/1994	1994
CH2-Pending submission	PLE653_TZA_1994	PLE	TZA	01/28/1994	1994
CH2-Pending submission	PLE654_TZA_1994	PLE	TZA	01/31/1994	1994
CH2-Pending submission	PLE656_TZA_1994	PLE	TZA	02/20/1994	1994
CH2-Pending submission	PLE658_TZA_1994	PLE	TZA	05/21/1994	1994

Table S4.3. Positively selected sites along the CDV genome.

Codon	SLAC dN-dS	SLAC p-value	FEL dN-dS	FEL p-value	MEME ω^+	MEME p-value
34	-0.002	0.741	0.112	0.95	>100	0.1
72	-2.954	0.905	-1.092	0.542	>100	0.06
111	7.404	0.351	2.176	0.443	83.425	0.046
134	12.261	0.136	3.96	0.061	>100	0.01
353	1.231	0.667	0.407	0.892	>100	0.001
354	-5.91	0.905	-2.898	0.3	>100	0
372	-2.468	0.889	-0.961	0.578	>100	0.029
383	2.374	0.712	0.755	0.387	>100	0.094
451	9.14	0.269	3.401	0.068	>100	0.055
456	16.062	0.097	6.358	0.026	>100	0.032
467	7.034	0.374	2.32	0.192	>100	0.039
517	4.539	0.526	1.598	0.266	>100	0.094
560	-7.434	0.934	-2.15	0.389	78.944	0.093
613	12.337	0.132	4.701	0.039	>100	0.025
617	2.374	0.712	0.755	0.387	>100	0.008
632	2.38	0.716	0.742	0.439	>100	0.063
643	2.468	0.667	0.751	0.395	>100	0.094
646	2.38	0.691	0.772	0.358	>100	0.084
655	4.934	0.445	1.545	0.225	>100	0.018
669	2.569	0.658	0.675	0.429	>100	0.054
671	9.834	0.201	3.843	0.063	>100	0.084
672	2.268	0.725	0.732	0.442	>100	0.052

677	2.269	0.725	0.733	0.442	>100	0.072
718	11.953	0.168	3.739	0.051	>100	0.07
744	9.723	0.218	2.795	0.078	>100	0.102
753	2.402	0.71	0.802	0.375	>100	0.082
758	9.523	0.23	2.759	0.088	>100	0.113
779	9.873	0.198	3.34	0.053	>100	0.072
792	9.873	0.198	3.448	0.049	>100	0.069
803	7.26	0.355	2.533	0.118	>100	0.022
819	14.374	0.108	4.277	0.034	>100	0.05
839	3.622	0.626	1.356	0.339	>100	0.097
851	-3.722	0.925	-1.594	0.431	>100	0.026
903	-3.723	0.925	-0.931	0.57	>100	0.01
1014	2.268	0.725	0.732	0.442	>100	0.015
1022	-13.925	0.978	-4.377	0.182	>100	0.098
1034	-3.735	0.853	-1.673	0.446	>100	0.008
1039	10.487	0.238	4.07	0.059	>100	0.08
1043	-1.449	0.815	-0.194	0.922	>100	0.081
1054	-2.47	0.889	-0.726	0.62	>100	0.027
1105	-7.205	0.958	-2.038	0.297	>100	0.05
1184	-14.058	0.991	-6.074	0.031	>100	0.048
1213	1.796	0.637	0.472	0.819	>100	0.086
1231	4.754	0.513	1.541	0.268	>100	0.002
1255	0.84	0.695	0.297	0.886	83.125	0.087
1257	4.808	0.503	1.629	0.208	>100	0.043
1315	3.295	0.998	1.021	0.96	>100	0.024
1342	2.706	0.646	0.853	0.382	>100	0.059

1368	16.117	0.101	6.087	0.016	>100	0.008
1369	14.258	0.135	5.427	0.049	>100	0.068
1371	12.32	0.133	4.097	0.032	>100	0.032
1372	16.434	0.105	5.95	0.043	>100	0.061
1376	9.834	0.201	2.99	0.065	>100	0.082
1378	6.709	0.38	2.984	0.276	>100	0.077
1380	4.936	0.445	1.424	0.197	>100	0.064
1386	26.271	0.027	11.733	0.006	>100	0.005
1391	11.996	0.181	4.299	0.077	>100	0.1
1393	17.225	0.06	5.393	0.015	>100	0.024
1408	10.934	0.23	4.184	0.057	>100	0.075
1416	17.837	0.049	5.487	0.023	>100	0.02
1418	15.929	0.101	6.934	0.012	>100	0.009
1419	-5.177	0.869	-1.29	0.632	>100	0.031
1423	9.249	0.268	3.935	0.052	>100	0.071
1424	7.38	0.356	2.969	0.093	>100	0.118
1429	12.292	0.134	3.894	0.036	>100	0.052
1437	14.109	0.137	6.085	0.017	>100	0.027
1452	7.196	0.323	3.012	0.074	>100	0.091
1471	13.661	0.223	7.247	0.068	>100	0.025
1481	11.977	0.185	4.005	0.047	>100	0.043
1502	1.755	0.937	0.594	0.714	>100	0.07
1529	7.394	0.376	2.733	0.123	>100	0.007
1560	4.932	0.445	1.563	0.181	>100	0.043
1567	-7.404	0.963	-2.903	0.158	>100	0.077
1579	-2.168	0.851	-0.349	0.872	>100	0.045

1588	9.533	0.233	2.756	0.089	>100	0.113
1676	2.468	0.667	0.654	0.38	>100	0.025
1719	-2.953	0.905	-1.163	0.526	>100	0.021
1731	6.825	0.378	2.332	0.171	>100	0.062
1911	2.461	0.593	0.61	0.751	>100	0.007
1942	2.269	0.725	0.777	0.377	>100	0.036
1972	-2.488	0.89	-0.726	0.611	>100	0.025
1974	-3.872	0.857	-1.323	0.594	>100	0.008
1976	-27.686	1	-9.023	0.003	>100	0.05
2001	4.936	0.445	1.409	0.256	>100	0.032
2002	9.518	0.229	3.7	0.048	>100	0.066
2004	9.02	0.288	3.828	0.072	>100	0.094
2008	2.468	0.667	0.753	0.35	>100	0.082
2019	2.472	0.667	0.698	0.41	>100	0.015
2051	2.471	0.667	0.8	0.339	>100	0.09
2071	4.767	0.478	1.561	0.193	>100	0.006
2077	2.1	0.593	0.915	0.638	>100	0.034
2164	2.269	0.725	0.86	0.357	>100	0.09
2236	-1.934	0.81	-0.697	0.73	>100	0.074
2303	-2.429	0.772	-0.936	0.773	>100	0.002
2339	3.65	0.627	1.355	0.34	>100	0.027
2404	4.816	0.514	1.503	0.275	>100	0.005
2426	-7.369	0.96	-3.185	0.145	>100	0.085
2442	9.988	0.197	3.086	0.096	>100	0.089
2444	9.655	0.223	3.253	0.09	>100	0.114
2487	9.891	0.198	3.882	0.061	>100	0

2498	1.538	0.667	1.048	0.691	>100	0.001
2514	6.826	0.381	2.015	0.201	>100	0.089
2537	8.031	0.191	3.006	0.076	>100	0.099
2553	-9.775	0.979	-4.219	0.112	>100	0.077
2568	-7.452	0.963	-3.087	0.176	>100	0.039
2576	26.766	0.024	12.217	0.001	>100	0.002
2587	2.283	0.725	0.734	0.442	>100	0.028
2647	2.39	0.712	0.757	0.387	>100	0.044
2767	0	0.741	0.052	0.981	>100	0.002
2780	-2.484	0.889	-0.755	0.595	>100	0.096
2849	4.934	0.445	1.803	0.157	>100	0.076
2850	-12.341	0.988	-3.756	0.082	>100	0.066
2861	1.646	1	0.506	0.992	>100	0.016
2928	2.37	0.719	0.813	0.44	>100	0.061
3014	7.327	0.306	2.962	0.07	>100	0.088
3223	2.273	0.725	0.861	0.357	>100	0.09
3245	12.295	0.161	3.926	0.065	>100	0.087
3258	-9.733	0.979	-2.649	0.238	>100	0.061
3270	-7.419	0.963	-3.035	0.179	>100	0.09
3285	2.379	0.712	0.759	0.386	>100	0.093
3433	-12.366	0.988	-3.816	0.09	>100	0.067
3480	-2.474	0.889	-0.617	0.684	>100	0.083
3520	-15.508	0.993	-4.644	0.102	>100	0.064
3702	2.38	0.691	0.77	0.359	>100	0.044
3991	-19.792	0.993	-7.643	0.027	>100	0.008
4023	2.475	0.667	0.766	0.404	>100	0.072

4036	2.274	0.725	0.778	0.377	>100	0.017
4283	-1.735	0.803	-0.781	0.71	>100	0.045
4373	4.622	0.428	2.959	0.181	>100	0.015
4378	1.552	0.603	0.486	0.87	57.254	0.057
4511	1.418	0.608	0.814	0.754	67.165	0.05
4691	-1.326	0.808	-0.834	0.747	39.953	0.06
4738	2.38	0.716	0.742	0.439	>100	0.058

* cutoff $p \leq 0.10$

Highlighted rows indicate sites that differentiate Fatal and Silent strains

Table S4.4. Negatively selected sites along the CDV genome.

Codon	SLAC dN-dS	SLAC p-value	FEL dN-dS	FEL p-value
1	-38.886	0	-1.798	1
2	-9.873	0.111	-3.539	0.023
6	-11.736	0.085	-3.563	0.031
10	-14.658	0.066	-5.244	0.017
14	-9.873	0.111	-3.67	0.024
21	-14.809	0.037	-4.608	0.008
22	-9.873	0.119	-3.141	0.039
25	-9.873	0.117	-3.69	0.022
27	-4.936	0.333	-1.942	0.098
28	-25.517	0.064	-8.681	0.019
30	-9.873	0.111	-2.951	0.03
39	-21.33	0.009	-6.395	0.003
40	-9.873	0.111	-3.047	0.037
42	-5.991	0.275	-2.131	0.099
43	-9.873	0.124	-3.747	0.032
48	-5.446	0.309	-2.486	0.077
51	-8.125	0.225	-3.196	0.068
53	-5.991	0.275	-2.161	0.098
57	-5.503	0.309	-2.092	0.093
59	-9.873	0.111	-2.908	0.038
66	-12.264	0.117	-5.382	0.039
68	-12.079	0.08	-5.149	0.017
75	-5.991	0.275	-2.193	0.085

80	-17.974	0.021	-5.625	0.008
82	-6.058	0.279	-2.381	0.092
85	-9.873	0.117	-3.788	0.021
86	-11.749	0.149	-3.49	0.069
87	-11.687	0.114	-4.167	0.033
89	-11.763	0.149	-3.654	0.065
104	-9.873	0.111	-3.878	0.025
105	-11.505	0.082	-3.573	0.028
114	-7.257	0.265	-5.652	0.022
122	-5.991	0.275	-2.218	0.084
124	-5.991	0.275	-2.3	0.093
127	-4.936	0.333	-1.908	0.1
132	-6.07	0.281	-2.382	0.097
135	-11.983	0.075	-3.491	0.028
141	-12.12	0.14	-4.899	0.043
142	-9.873	0.117	-3.665	0.022
145	-9.873	0.111	-2.84	0.032
147	-11.812	0.11	-4.318	0.03
148	-11.726	0.083	-3.515	0.029
149	-11.983	0.075	-4.566	0.02
150	-17.605	0.025	-5.334	0.009
153	-14.809	0.037	-5.829	0.006
155	-19.84	0.083	-7.466	0.023
157	-9.873	0.111	-3.659	0.027
165	-37.666	0.008	-15.923	0.001
168	-19.745	0.012	-7.659	0.001

170	-10.665	0.095	-3.053	0.037
171	-11.814	0.11	-4.263	0.031
172	-14.809	0.037	-5.152	0.006
173	-30.247	0.007	-14.65	0.001
178	-14.108	0.043	-4.495	0.017
181	-9.873	0.111	-3.725	0.021
182	-9.873	0.111	-3.029	0.036
184	-9.873	0.111	-3.57	0.023
186	-17.974	0.021	-7.513	0.003
191	-9.873	0.119	-3.04	0.041
194	-5.449	0.309	-2.023	0.096
195	-16.219	0.029	-5.466	0.005
197	-10.665	0.095	-3.906	0.025
202	-11.792	0.148	-3.708	0.064
203	-5.507	0.309	-2.148	0.091
207	-14.809	0.037	-4.523	0.008
210	-5.51	0.309	-2.181	0.09
220	-11.006	0.096	-4.641	0.015
222	-10.73	0.095	-3.384	0.025
226	-11.983	0.075	-3.512	0.028
227	-11.975	0.114	-5.185	0.023
228	-14.809	0.037	-5.839	0.007
230	-16.068	0.029	-5.237	0.006
231	-13.441	0.071	-4.292	0.018
234	-9.873	0.111	-2.961	0.037
236	-14.809	0.037	-4.54	0.015

237	-10.665	0.095	-3.357	0.032
239	-17.974	0.021	-6.448	0.005
242	-9.096	0.224	-3.46	0.056
246	-11.983	0.075	-3.434	0.031
249	-5.408	0.322	-2.132	0.095
252	-6.07	0.278	-2.382	0.092
256	-11.983	0.075	-4.93	0.016
257	-22.189	0.074	-8.731	0.019
262	-9.873	0.111	-2.986	0.037
263	-6.068	0.278	-2.404	0.091
265	-9.873	0.116	-3.787	0.021
267	-9.873	0.111	-3.881	0.02
268	-5.991	0.275	-2.176	0.086
271	-8.999	0.196	-3.826	0.05
282	-9.873	0.111	-2.93	0.039
283	-9.873	0.111	-3.556	0.023
284	-13.857	0.045	-5.383	0.011
288	-11.738	0.083	-3.469	0.029
289	-23.965	0.006	-9.768	0.001
290	-14.809	0.037	-4.426	0.012
291	-19.745	0.014	-7.999	0.001
297	-6.055	0.279	-2.346	0.093
302	-10.863	0.136	-4.474	0.032
308	-6.072	0.278	-2.434	0.09
311	-9.873	0.111	-3.049	0.037
312	-11.983	0.085	-3.526	0.043

315	-10.665	0.095	-3.966	0.024
323	-6.058	0.281	-2.375	0.098
327	-14.809	0.039	-5.621	0.005
328	-14.809	0.037	-4.46	0.012
336	-9.873	0.111	-3.548	0.023
340	-9.873	0.111	-3.744	0.021
348	-14.809	0.04	-6.027	0.004
349	-9.873	0.111	-3.011	0.029
352	-11.983	0.075	-3.424	0.034
355	-9.873	0.111	-2.932	0.04
356	-11.983	0.085	-3.49	0.043
358	-17.974	0.021	-6.668	0.004
359	-24.681	0.004	-11.47	0
360	-19.745	0.012	-7.661	0.001
362	-11.983	0.075	-3.477	0.034
363	-10.687	0.095	-3.29	0.026
369	-14.809	0.037	-5.391	0.007
370	-5.507	0.309	-2.364	0.083
371	-10.971	0.097	-4.499	0.017
373	-29.618	0.001	-9.847	0
374	-9.873	0.111	-2.844	0.032
375	-12.127	0.079	-5.206	0.017
376	-9.873	0.111	-3.796	0.025
378	-9.873	0.111	-3.82	0.027
381	-9.873	0.111	-2.873	0.032
382	-9.873	0.111	-2.86	0.032

384	-14.686	0.038	-5.553	0.01
387	-9.873	0.111	-2.905	0.034
389	-18.103	0.022	-7.581	0.003
390	-23.505	0.007	-7.547	0.002
393	-8.394	0.203	-3.459	0.056
394	-9.873	0.111	-3.058	0.035
396	-12.132	0.078	-5.029	0.016
403	-11.779	0.082	-3.608	0.028
406	-19.745	0.012	-6.366	0.002
412	-9.873	0.111	-3.011	0.037
414	-12.136	0.14	-5.116	0.041
415	-9.873	0.111	-3.465	0.031
417	-15.997	0.029	-5.045	0.009
418	-9.873	0.111	-3.814	0.022
419	-5.991	0.275	-2.372	0.1
422	-14.809	0.044	-5.774	0.008
428	-39.49	0.001	-18.636	0
429	-5.991	0.275	-2.344	0.096
430	-12.098	0.141	-5.105	0.041
438	-10.701	0.095	-3.32	0.025
439	-9.873	0.111	-2.872	0.041
440	-12.079	0.078	-4.808	0.017
449	-10.665	0.095	-3.202	0.034
453	-5.991	0.275	-2.398	0.089
454	-6.046	0.28	-2.395	0.092
455	-19.541	0.032	-6.751	0.019

459	-15.509	0.077	-5.859	0.071
462	-5.991	0.275	-2.243	0.095
466	-12.325	0.112	-3.728	0.099
470	-12.122	0.078	-4.774	0.018
474	-17.974	0.021	-6.725	0.004
477	-15.215	0.094	-6.468	0.055
484	-15.709	0.066	-5.157	0.051
486	-11.983	0.075	-3.508	0.028
487	-13.456	0.052	-4.102	0.016
488	-6.063	0.281	-2.392	0.097
492	-9.873	0.111	-3.815	0.02
497	-11.004	0.096	-4.533	0.016
499	-9.873	0.111	-3.694	0.034
500	-24.681	0.004	-9.755	0
509	-17.974	0.021	-7.316	0.002
514	-9.873	0.111	-3.573	0.029
515	-9.873	0.111	-2.99	0.037
518	-11.983	0.075	-4.549	0.018
526	-11.859	0.081	-3.737	0.026
527	-17.683	0.024	-5.711	0.006
530	-15.506	0.078	-7.321	0.041
533	-43.178	0.001	-21.186	0
566	-12.74	0.104	-3.826	0.07
567	-12.04	0.142	-6.692	0.023
575	-23.658	0.022	-7.578	0.01
576	-6.07	0.278	-2.378	0.092

577	-17.833	0.023	-6.014	0.006
673	-5.991	0.275	-2.197	0.097
684	-11.836	0.082	-4.13	0.023
696	-9.873	0.111	-2.957	0.03
698	-6.046	0.282	-2.395	0.097
699	-14.809	0.037	-6.611	0.005
701	-16.396	0.031	-7.832	0.002
702	-4.936	0.333	-2.11	0.089
703	-9.873	0.111	-3.529	0.023
712	-6.068	0.278	-2.406	0.091
720	-11.983	0.075	-3.4	0.023
733	-10.882	0.135	-3.703	0.046
737	-21.633	0.023	-6.673	0.028
739	-14.809	0.037	-4.729	0.01
746	-14.809	0.037	-4.454	0.011
784	-9.873	0.116	-3.768	0.021
790	-14.809	0.037	-4.775	0.009
826	-9.873	0.111	-3.912	0.026
828	-5.991	0.275	-2.171	0.086
830	-29.957	0.002	-10.804	0.001
831	-11.983	0.075	-4.274	0.02
837	-11.896	0.081	-3.842	0.025
840	-9.788	0.19	-4.995	0.092
841	-16.295	0.028	-6.325	0.005
842	-8.999	0.247	-3.367	0.062
844	-9.873	0.111	-3.09	0.029

847	-6.006	0.284	-2.548	0.091
854	-9.873	0.111	-3.78	0.021
858	-6.055	0.282	-2.346	0.099
860	-5.991	0.275	-2.171	0.098
867	-29.618	0.002	-12.763	0
878	-11.995	0.081	-4.676	0.02
881	-5.332	0.309	-2.743	0.074
882	-14.809	0.037	-5.823	0.005
884	-14.809	0.037	-4.448	0.012
885	-9.873	0.111	-2.916	0.034
891	-9.873	0.125	-3.766	0.031
893	-10.665	0.095	-3.887	0.025
901	-9.873	0.112	-2.942	0.03
904	-12.353	0.111	-5.097	0.047
906	-23.965	0.006	-7.855	0.001
907	-9.873	0.111	-2.927	0.062
908	-19.745	0.012	-6.22	0.002
916	-23.658	0.007	-7.74	0.002
920	-22.477	0.073	-6.638	0.026
921	-37.157	0.008	-14.025	0.002
923	-10.892	0.095	-4.426	0.016
924	-11.983	0.075	-4.681	0.017
925	-9.873	0.125	-3.759	0.031
926	-9.873	0.117	-3.969	0.02
927	-5.509	0.309	-2.314	0.085
928	-4.936	0.333	-1.944	0.098

930	-12.341	0.111	-4.023	0.067
932	-9.873	0.111	-3.749	0.026
936	-14.809	0.037	-4.437	0.011
937	-9.873	0.111	-3.973	0.019
941	-12.579	0.107	-5.028	0.042
950	-8.815	0.193	-3.06	0.066
951	-9.873	0.111	-2.919	0.031
953	-6.07	0.281	-2.382	0.097
955	-12.116	0.14	-4.929	0.043
956	-8.24	0.22	-3.315	0.064
960	-14.57	0.039	-4.384	0.016
963	-6.07	0.281	-2.382	0.097
964	-9.873	0.111	-3.588	0.029
966	-11.983	0.075	-3.614	0.026
970	-11.983	0.075	-3.557	0.022
971	-19.745	0.014	-6.254	0.004
974	-9.873	0.116	-3.983	0.02
977	-9.873	0.111	-3.103	0.036
979	-5.991	0.275	-2.291	0.093
980	-14.809	0.037	-4.561	0.009
982	-9.873	0.111	-3.677	0.028
983	-9.873	0.111	-3.078	0.037
991	-10.665	0.095	-4.169	0.023
993	-9.873	0.125	-3.88	0.03
999	-5.991	0.275	-2.17	0.086
1001	-6.073	0.281	-2.482	0.094

1002	-11.983	0.075	-3.483	0.036
1005	-16.513	0.046	-5.838	0.013
1006	-12.341	0.111	-4.695	0.05
1008	-29.045	0.002	-10.815	0
1011	-10.665	0.095	-3.188	0.035
1012	-11.755	0.084	-3.534	0.031
1013	-9.873	0.111	-2.911	0.031
1016	-11.983	0.075	-4.285	0.022
1017	-8.15	0.222	-3.185	0.068
1025	-11.761	0.084	-3.767	0.028
1026	-5.991	0.275	-2.193	0.085
1028	-14.809	0.037	-6.356	0.004
1032	-19.745	0.012	-7.804	0.001
1038	-11.983	0.075	-4.741	0.017
1040	-19.745	0.012	-7.949	0.002
1041	-29.618	0.002	-11.635	0
1045	-12.129	0.079	-5.313	0.016
1046	-9.873	0.111	-3.011	0.029
1047	-12.028	0.129	-4.78	0.077
1048	-14.727	0.064	-5.453	0.015
1050	-19.745	0.012	-6.23	0.003
1052	-14.848	0.063	-5.513	0.015
1053	-9.873	0.111	-2.914	0.039
1055	-9.873	0.111	-3.456	0.026
1057	-19.745	0.012	-6.228	0.003
1063	-9.873	0.111	-2.895	0.04

1070	-14.809	0.037	-6.139	0.006
1075	-13.514	0.119	-4.989	0.021
1076	-6.068	0.281	-2.406	0.097
1077	-5.991	0.275	-2.259	0.094
1078	-6.071	0.278	-2.414	0.091
1079	-23.965	0.007	-9.309	0.001
1084	-11.983	0.075	-4.57	0.022
1085	-23.417	0.017	-10.431	0.002
1088	-19.882	0.146	-7.741	0.099
1090	-6.067	0.279	-2.436	0.09
1092	-23.965	0.006	-7.463	0.002
1098	-29.618	0.001	-10.262	0
1100	-4.936	0.342	-1.944	0.1
1101	-5.51	0.309	-2.181	0.09
1107	-9.873	0.111	-3.671	0.028
1109	-14.809	0.037	-6.092	0.004
1114	-14.809	0.1	-6.994	0.038
1116	-11.014	0.096	-4.436	0.016
1118	-12.097	0.078	-4.709	0.018
1119	-6.064	0.279	-2.308	0.095
1122	-17.635	0.024	-5.522	0.008
1125	-19.745	0.012	-6.481	0.002
1126	-11.173	0.124	-3.79	0.043
1128	-17.974	0.021	-6.296	0.005
1132	-10.681	0.095	-3.266	0.026
1133	-8.997	0.251	-3.364	0.063

1142	-9.873	0.111	-3.753	0.026
1149	-11.797	0.111	-4.172	0.033
1156	-17.637	0.024	-5.6	0.008
1160	-14.809	0.037	-4.476	0.009
1161	-13.437	0.131	-6.688	0.054
1162	-19.745	0.013	-6.609	0.002
1163	-17.974	0.021	-7.147	0.003
1164	-9.873	0.111	-2.909	0.038
1165	-11.983	0.075	-3.408	0.035
1166	-11.983	0.075	-4.619	0.014
1167	-9.873	0.111	-3.822	0.02
1168	-11.983	0.075	-5.375	0.015
1169	-18.094	0.053	-7.699	0.013
1178	-9.873	0.111	-3.785	0.026
1179	-11.994	0.112	-5.136	0.023
1180	-11.983	0.075	-3.553	0.027
1184	-14.058	0.091	-6.074	0.031
1185	-17.974	0.021	-7.336	0.004
1186	-16.021	0.029	-5.218	0.006
1188	-9.873	0.111	-3.965	0.024
1189	-11.983	0.085	-4.44	0.03
1191	-17.974	0.021	-5.476	0.005
1196	-9.873	0.124	-3.806	0.031
1197	-11.983	0.075	-3.424	0.029
1199	-9.873	0.117	-3.853	0.021
1200	-5.991	0.275	-2.198	0.085

1204	-14.809	0.037	-5.297	0.008
1205	-9.802	0.117	-2.921	0.04
1209	-6.071	0.278	-2.41	0.091
1212	-24.681	0.004	-8.371	0.001
1224	-10.737	0.094	-3.932	0.025
1226	-14.809	0.037	-5.493	0.007
1229	-5.991	0.275	-2.176	0.097
1235	-14.809	0.037	-5.701	0.005
1238	-25.373	0.02	-8.708	0.011
1249	-17.325	0.045	-5.574	0.031
1254	-11.983	0.075	-3.496	0.033
1256	-9.869	0.111	-3.906	0.024
1260	-12.133	0.14	-5.066	0.042
1264	-9.873	0.111	-3.57	0.023
1284	-19.745	0.013	-6.205	0.002
1287	-15.706	0.066	-6.341	0.026
1290	-23.525	0.07	-8.831	0.019
1292	-5.991	0.284	-2.194	0.098
1299	-14.809	0.037	-5.288	0.006
1300	-11.772	0.111	-4.137	0.033
1302	-19.745	0.012	-6.156	0.002
1305	-9.873	0.113	-3.304	0.025
1306	-11.983	0.075	-3.493	0.033
1314	-14.534	0.039	-4.543	0.015
1318	-11.983	0.075	-4.311	0.022
1320	-5.991	0.275	-2.344	0.091

1321	-11.801	0.11	-4.152	0.033
1325	-8.191	0.222	-3.231	0.067
1327	-5.509	0.309	-2.177	0.09
1334	-11.01	0.096	-4.73	0.015
1337	-9.873	0.111	-3.899	0.02
1338	-9.873	0.111	-2.85	0.039
1339	-8.829	0.192	-3.274	0.059
1340	-12.13	0.14	-5.077	0.041
1341	-19.745	0.012	-8.679	0.001
1343	-9.873	0.111	-2.85	0.04
1345	-12.135	0.14	-5.013	0.042
1348	-5.498	0.309	-2.151	0.091
1354	-10.665	0.095	-3.115	0.036
1356	-17.974	0.021	-5.558	0.005
1357	-5.991	0.275	-2.171	0.098
1361	-9.5	0.12	-3.733	0.035
1362	-11.983	0.075	-3.408	0.035
1363	-6.058	0.281	-2.375	0.098
1365	-9.873	0.111	-2.857	0.05
1405	-9.873	0.111	-4.015	0.02
1426	-17.037	0.082	-56.106	0.024
1431	-24.681	0.004	-9.933	0
1435	-19.745	0.045	-8.109	0.017
1442	-10.665	0.095	-3.164	0.035
1448	-11.983	0.075	-3.414	0.037
1462	-4.936	0.342	-2.031	0.096

1468	-15.556	0.073	-6.234	0.042
1489	-14.858	0.081	-4.275	0.085
1494	-9.873	0.111	-2.856	0.051
1498	-24.681	0.004	-9.115	0.001
1503	-17.974	0.021	-6.843	0.006
1512	-21.33	0.009	-8.656	0.001
1513	-10.865	0.092	-3.37	0.031
1514	-9.873	0.114	-3.106	0.028
1515	-9.873	0.111	-3.532	0.025
1517	-5.991	0.275	-2.198	0.085
1525	-10.719	0.095	-3.301	0.026
1527	-5.991	0.275	-2.176	0.086
1530	-11.983	0.085	-3.422	0.045
1533	-23.297	0.071	-8.86	0.019
1535	-14.353	0.062	-5.806	0.014
1544	-11.983	0.075	-4.49	0.021
1545	-5.991	0.284	-2.272	0.095
1546	-9.873	0.111	-2.896	0.034
1547	-6.07	0.281	-2.382	0.097
1557	-9.035	0.194	-3.651	0.053
1561	-15.352	0.034	-5.563	0.008
1565	-17.974	0.021	-5.393	0.008
1574	-5.991	0.275	-2.428	0.093
1578	-10.957	0.134	-4.43	0.033
1583	-9.873	0.125	-3.739	0.032
1586	-5.506	0.309	-2.235	0.088

1587	-18.122	0.053	-7.982	0.012
1594	-14.809	0.037	-4.54	0.01
1611	-9.873	0.112	-3.009	0.029
1616	-17.97	0.054	-8.647	0.009
1618	-17.974	0.021	-5.3	0.008
1621	-9.873	0.111	-3.678	0.022
1623	-14.809	0.037	-6.232	0.004
1625	-12.086	0.141	-4.98	0.042
1632	-12.132	0.078	-5.001	0.016
1638	-24.667	0.067	-8.438	0.02
1639	-6.07	0.278	-2.378	0.092
1644	-9.873	0.111	-4.973	0.012
1646	-11.746	0.149	-3.584	0.067
1651	-14.809	0.037	-4.826	0.007
1654	-9.873	0.117	-3.856	0.021
1657	-11.983	0.075	-4.956	0.016
1659	-9.873	0.111	-3.027	0.036
1663	-33.98	0.001	-11.592	0
1671	-14.809	0.042	-4.636	0.012
1673	-6.011	0.281	-2.3	0.095
1681	-9.041	0.194	-3.822	0.05
1687	-11.983	0.085	-5.208	0.022
1693	-25.269	0.065	-8.683	0.019
1695	-24.681	0.004	-7.674	0.001
1696	-14.809	0.037	-4.49	0.011
1697	-5.991	0.275	-2.22	0.084

1698	-8.966	0.197	-3.742	0.052
1700	-11.983	0.075	-3.584	0.027
1703	-9.873	0.125	-3.874	0.03
1706	-25.555	0.064	-8.753	0.019
1709	-12.133	0.14	-4.886	0.043
1711	-8.146	0.223	-2.707	0.086
1713	-11.983	0.085	-4.555	0.029
1715	-14.809	0.037	-5.388	0.01
1716	-14.809	0.038	-4.636	0.007
1721	-11.733	0.084	-3.479	0.032
1725	-11.736	0.084	-3.5	0.032
1727	-24.681	0.004	-9.666	0
1728	-11.983	0.085	-4.905	0.026
1732	-5.991	0.275	-2.171	0.098
1735	-4.936	0.333	-1.935	0.099
1736	-10.665	0.095	-3.807	0.026
1741	-14.809	0.037	-4.501	0.008
1743	-6.053	0.282	-2.359	0.098
1747	-19.745	0.012	-7.441	0.001
1750	-5.991	0.275	-2.175	0.097
1751	-24.74	0.019	-8.215	0.027
1754	-6.055	0.282	-2.346	0.099
1759	-5.991	0.275	-2.149	0.087
1763	-14.809	0.037	-5.972	0.006
1764	-9.873	0.111	-3.65	0.024
1767	-12.129	0.078	-5.274	0.015

1768	-14.809	0.037	-4.429	0.011
1769	-17.654	0.024	-5.44	0.008
1773	-9.873	0.111	-2.909	0.038
1775	-21.915	0.009	-11.507	0
1777	-18.185	0.022	-7.728	0.003
1778	-9.873	0.111	-3.769	0.021
1784	-14.809	0.037	-4.49	0.008
1786	-11.871	0.146	-5.188	0.038
1792	-9.873	0.111	-2.915	0.038
1797	-4.936	0.333	-1.916	0.1
1800	-14.809	0.037	-5.635	0.005
1804	-9.873	0.111	-3.648	0.029
1805	-11.983	0.075	-4.678	0.019
1806	-11.983	0.075	-4.524	0.022
1807	-5.991	0.275	-2.193	0.097
1810	-12.341	0.111	-3.917	0.091
1811	-11.983	0.081	-4.501	0.019
1812	-9.873	0.111	-3.806	0.025
1813	-21.543	0.023	-6.627	0.03
1820	-5.991	0.284	-2.213	0.097
1825	-10.919	0.135	-4.277	0.035
1829	-11.983	0.075	-3.571	0.021
1830	-24.681	0.004	-9.281	0.001
1833	-12.134	0.14	-5.002	0.042
1838	-9.873	0.111	-2.85	0.032
1842	-14.809	0.037	-5.624	0.007

1843	-5.991	0.284	-2.175	0.099
1844	-9.873	0.111	-3.644	0.022
1848	-19.745	0.012	-6.399	0.003
1850	-14.809	0.037	-4.587	0.007
1851	-14.809	0.037	-5.422	0.006
1855	-11.89	0.082	-5.134	0.016
1860	-12.132	0.079	-5.072	0.018
1864	-9.873	0.111	-3.075	0.035
1868	-9.873	0.111	-3.734	0.021
1871	-19.417	0.028	-8.355	0.004
1881	-10.665	0.095	-3.13	0.036
1886	-5.991	0.275	-2.17	0.098
1888	-11.844	0.11	-4.2	0.032
1891	-12.764	0.075	-4.867	0.021
1894	-9.873	0.111	-2.884	0.041
1895	-5.991	0.275	-2.307	0.093
1897	-11.983	0.081	-3.437	0.029
1898	-9.873	0.111	-3.872	0.025
1901	-12.133	0.078	-4.993	0.017
1906	-9.873	0.111	-3.574	0.025
1907	-13.586	0.091	-3.947	0.089
1908	-11.892	0.146	-4.9	0.04
1910	-9.873	0.116	-3.818	0.021
1913	-12.134	0.14	-4.942	0.043
1919	-17.974	0.025	-5.683	0.011
1924	-9.873	0.111	-2.85	0.032

1925	-5.876	0.343	-3.199	0.078
1937	-9.873	0.111	-3.73	0.021
1938	-9.873	0.111	-3.939	0.021
1939	-11.983	0.075	-4.731	0.019
1946	-6.067	0.281	-2.532	0.092
1950	-9.873	0.111	-3.623	0.022
1951	-11.857	0.083	-4.078	0.025
1960	-10.665	0.095	-3.058	0.037
1963	-9.873	0.111	-3.095	0.031
1968	-9.873	0.111	-2.926	0.04
1976	-27.686	0.007	-9.023	0.003
1978	-22.213	0.018	-7.83	0.009
1981	-19.798	0.03	-6.107	0.022
1984	-12.341	0.111	-3.743	0.076
1988	-14.658	0.065	-5.548	0.015
1990	-13.655	0.081	-4.83	0.023
1993	-11.983	0.081	-3.696	0.026
1998	-11.983	0.085	-3.484	0.044
2010	-12.081	0.116	-3.879	0.091
2016	-11.554	0.116	-4.139	0.034
2017	-14.809	0.037	-5.956	0.005
2018	-10.114	0.118	-3.783	0.021
2020	-12.242	0.121	-4.146	0.091
2022	-34.614	0.001	-13.981	0
2024	-9.89	0.111	-3.718	0.026
2026	-9.89	0.125	-4.034	0.028

2030	-17.307	0.045	-5.384	0.036
2041	-21.729	0.022	-9.808	0.007
2043	-4.945	0.333	-1.944	0.098
2069	-9.785	0.114	-3.61	0.036
2079	-11.833	0.11	-4.295	0.031
2086	-13.551	0.12	-5.175	0.022
2102	-18.175	0.022	-7.477	0.004
2107	-6.082	0.281	-2.584	0.09
2108	-14.835	0.044	-5.855	0.008
2110	-14.835	0.037	-4.404	0.008
2118	-9.89	0.111	-2.854	0.039
2119	-8.897	0.19	-3.132	0.063
2121	-9.89	0.111	-2.926	0.039
2127	-4.945	0.342	-1.992	0.098
2131	-9.864	0.112	-2.869	0.032
2138	-8.138	0.224	-3.195	0.068
2142	-6.067	0.281	-2.484	0.094
2145	-10.665	0.095	-3.16	0.035
2147	-18.185	0.052	-8.649	0.01
2152	-17.974	0.021	-5.383	0.009
2153	-5.991	0.275	-2.255	0.1
2156	-21.452	0.009	-7.028	0.001
2157	-12.08	0.078	-4.748	0.018
2161	-14.622	0.043	-4.647	0.011
2166	-17.974	0.023	-5.44	0.007
2168	-17.974	0.021	-5.805	0.007

2169	-14.809	0.037	-5.822	0.007
2171	-5.991	0.275	-2.068	0.09
2174	-23.608	0.007	-8.245	0.002
2178	-14.809	0.037	-5.535	0.006
2181	-11.983	0.081	-4.096	0.023
2188	-6.072	0.281	-2.683	0.087
2196	-14.937	0.037	-4.732	0.01
2203	-14.937	0.037	-4.482	0.012
2204	-32.305	0.003	-14.395	0
2206	-8.626	0.227	-4.415	0.093
2214	-5.556	0.309	-2.18	0.091
2229	-14.937	0.041	-6.162	0.007
2230	-19.916	0.012	-8.161	0.001
2232	-12.035	0.113	-4.844	0.026
2235	-9.958	0.111	-3.739	0.028
2241	-12.234	0.128	-5.249	0.065
2249	-12.448	0.111	-4.891	0.048
2253	-14.937	0.04	-6.112	0.004
2256	-9.958	0.111	-3.505	0.023
2259	-12.086	0.085	-4.359	0.031
2260	-8.97	0.19	-3.125	0.064
2272	-27.338	0.007	-9.637	0.003
2280	-23.827	0.07	-8.932	0.019
2285	-9.211	0.13	-2.676	0.048
2288	-18.13	0.021	-7.083	0.004
2291	-14.937	0.037	-6.012	0.006

2296	-9.958	0.111	-2.897	0.035
2302	-14.917	0.101	-4.847	0.095
2305	-19.916	0.012	-6.026	0.004
2307	-12.448	0.111	-3.726	0.076
2312	-17.723	0.094	-7.799	0.022
2313	-9.958	0.111	-3.989	0.019
2315	-19.916	0.012	-6.137	0.002
2322	-0.008	0.002	0	1
2323	-6.043	0.284	-2.229	0.097
2324	-9.958	0.111	-3.822	0.025
2328	-9.958	0.111	-3.025	0.032
2331	-9.101	0.194	-3.8	0.051
2343	-8.197	0.224	-3.257	0.066
2346	-14.937	0.037	-5.271	0.006
2347	-14.107	0.075	-4.769	0.024
2348	-14.937	0.04	-5.618	0.005
2359	-12.218	0.078	-4.958	0.017
2360	-18.657	0.025	-5.914	0.006
2369	-14.579	0.113	-7.54	0.024
2370	-21.758	0.025	-8.787	0.01
2374	-24.373	0.068	-8.415	0.02
2376	-12.086	0.075	-4.397	0.021
2378	-19.886	0.013	-7.746	0.002
2379	-9.958	0.111	-3.152	0.027
2380	-27.865	0.007	-9.245	0.007
2381	-24.35	0.068	-7.492	0.023

2383	-6.043	0.275	-2.154	0.099
2393	-8.932	0.192	-3.07	0.066
2395	-9.958	0.111	-3.617	0.029
2399	-17.783	0.024	-5.547	0.007
2407	-19.916	0.014	-6.622	0.003
2415	-9.958	0.111	-3.471	0.03
2425	-18.13	0.021	-6.659	0.006
2427	-14.937	0.04	-6.445	0.004
2431	-9.958	0.111	-3.878	0.025
2433	-12.086	0.085	-4.584	0.029
2447	-12.24	0.14	-5.039	0.042
2450	-17.157	0.097	-6.077	0.029
2451	-24.895	0.005	-8.87	0.001
2470	-9.958	0.111	-3.975	0.024
2477	-9.86	0.113	-3.558	0.037
2480	-19.916	0.012	-7.91	0.002
2481	-19.916	0.012	-6.702	0.003
2501	-19.785	0.012	-7.68	0.002
2505	-9.893	0.111	-3.023	0.036
2509	-14.829	0.101	-4.884	0.073
2513	-9.893	0.125	-3.805	0.031
2515	-14.839	0.04	-5.868	0.005
2519	-14.987	0.065	-7.198	0.008
2521	-10.717	0.095	-4.006	0.024
2528	-0.006	0.013	0	1
2532	-9.751	0.115	-3.533	0.037

2535	-9.936	0.111	-2.957	0.04
2555	-14.904	0.037	-4.709	0.011
2556	-13.544	0.052	-4.176	0.016
2560	-14.904	0.037	-4.824	0.007
2563	-12.06	0.085	-4.515	0.03
2579	-11.082	0.096	-4.553	0.016
2581	-11.462	0.11	-3.621	0.025
2583	-17.773	0.024	-5.507	0.008
2585	-11.08	0.096	-4.559	0.016
2589	-13.686	0.082	-5.549	0.017
2590	-16.099	0.03	-4.999	0.006
2592	-6.109	0.278	-2.384	0.092
2602	-12.06	0.081	-4.789	0.017
2603	-12.06	0.075	-3.695	0.033
2605	-18.089	0.021	-6.858	0.005
2608	-15.803	0.066	-4.595	0.08
2609	-11.829	0.083	-3.518	0.029
2615	-9.936	0.111	-3.654	0.029
2619	-9.936	0.111	-3.538	0.028
2630	-18.089	0.021	-5.304	0.008
2634	-11.323	0.111	-3.691	0.029
2639	-14.904	0.043	-6.016	0.007
2640	-9.936	0.111	-2.936	0.038
2641	-24.119	0.006	-7.597	0.002
2651	-6.03	0.275	-2.307	0.092
2652	-18.089	0.021	-5.371	0.005

2655	-14.904	0.037	-5.672	0.006
2660	-9.936	0.111	-3.598	0.028
2664	-29.699	0.002	-10.143	0
2665	-12.06	0.085	-4.42	0.031
2667	-9.119	0.241	-3.375	0.063
2676	-6.03	0.275	-2.616	0.08
2678	-24.84	0.004	-7.962	0.001
2686	-11.059	0.096	-4.434	0.017
2689	-12.162	0.078	-5.101	0.016
2694	-11.818	0.149	-3.571	0.067
2697	-10.733	0.095	-3.866	0.026
2698	-24.119	0.006	-7.448	0.002
2701	-12.189	0.078	-5.167	0.016
2703	-9.936	0.111	-2.874	0.032
2704	-12.06	0.075	-3.417	0.023
2709	-9.936	0.111	-2.863	0.035
2714	-9.936	0.111	-2.905	0.04
2717	-15.808	0.066	-4.605	0.071
2721	-12.06	0.085	-4.395	0.031
2726	-12.16	0.08	-4.99	0.018
2727	-15.995	0.058	-6.524	0.035
2728	-10.967	0.134	-3.824	0.043
2729	-18.089	0.021	-7.042	0.006
2732	-17.728	0.058	-5.461	0.025
2734	-13.527	0.104	-5.103	0.021
2737	-9.936	0.111	-3.612	0.029

2739	-10.747	0.095	-3.322	0.025
2741	-5.546	0.309	-2.178	0.091
2747	-14.904	0.04	-5.784	0.005
2748	-12.06	0.075	-3.737	0.028
2750	-10.877	0.093	-3.204	0.034
2752	-12.06	0.075	-4.3	0.016
2759	-12.209	0.079	-5.018	0.018
2763	-12.06	0.075	-3.444	0.031
2770	-9.936	0.111	-3.542	0.023
2784	-6.03	0.275	-2.184	0.097
2798	-9.873	0.111	-3.606	0.028
2804	-9.873	0.111	-3.792	0.023
2813	-16.164	0.028	-6.395	0.005
2814	-0.002	0.001	0	1
2822	-12.12	0.079	-4.755	0.02
2823	-8.96	0.246	-3.238	0.061
2824	-10.667	0.096	-3.267	0.027
2826	-15.706	0.066	-6.664	0.034
2830	-19.745	0.012	-8.158	0.001
2838	-9.019	0.196	-3.926	0.048
2839	-10.161	0.113	-3.101	0.035
2840	-15.774	0.037	-5.629	0.008
2842	-17.974	0.021	-8.816	0.002
2845	-5.991	0.284	-2.193	0.098
2850	-12.341	0.111	-3.756	0.082
2855	-11.983	0.085	-4.502	0.03

2859	-11.983	0.075	-4.55	0.022
2860	-6.071	0.278	-2.417	0.091
2872	-21.83	0.01	-9.239	0.001
2873	-8.708	0.194	-3.435	0.058
2876	-5.991	0.275	-2.176	0.097
2877	-14.809	0.037	-5.458	0.006
2888	-11.861	0.112	-4.377	0.03
2900	-17.974	0.021	-5.785	0.006
2901	-9.873	0.117	-3.796	0.021
2902	-17.974	0.021	-7.231	0.005
2906	-8.999	0.196	-3.597	0.055
2915	-9.873	0.111	-3.66	0.022
2916	-8.17	0.222	-3.002	0.074
2920	-9.873	0.111	-3.461	0.038
2934	-14.249	0.042	-4.519	0.015
2941	-17.974	0.021	-5.522	0.009
2944	-19.745	0.012	-6.538	0.002
2945	-6.068	0.278	-2.404	0.091
2946	-10.665	0.095	-4.558	0.018
2948	-11.741	0.083	-3.489	0.029
2958	-6.055	0.279	-2.346	0.093
2960	-14.809	0.037	-4.738	0.008
2965	-9.873	0.111	-2.941	0.034
2967	-9.873	0.111	-3.663	0.022
2968	-8.241	0.215	-3.325	0.061
2970	-23.965	0.006	-7.515	0.002

2975	-11.76	0.083	-3.52	0.029
2982	-6.006	0.281	-2.548	0.086
2987	-11.983	0.075	-4.609	0.022
2991	-9.873	0.111	-3.905	0.019
2998	-25.511	0.064	-8.683	0.019
3002	-17.898	0.023	-6.458	0.005
3003	-17.974	0.021	-6.764	0.005
3005	-9.066	0.135	-2.621	0.05
3009	-17.974	0.021	-5.751	0.007
3011	-12.341	0.111	-3.721	0.097
3012	-6.07	0.281	-2.382	0.097
3018	-9.873	0.111	-2.959	0.033
3021	-23.515	0.007	-7.579	0.002
3023	-11.983	0.075	-3.631	0.034
3025	-21.89	0.023	-10.087	0.003
3030	-10.665	0.095	-3.149	0.035
3039	-9.873	0.116	-3.585	0.023
3040	-14.809	0.04	-6.441	0.004
3041	-14.809	0.037	-4.498	0.009
3044	-9.873	0.111	-3.746	0.027
3049	-9.873	0.111	-2.913	0.038
3053	-9.873	0.111	-3.775	0.026
3056	-16.043	0.029	-5.071	0.006
3058	-9.873	0.111	-3.502	0.023
3062	-14.809	0.04	-5.861	0.005
3064	-9.873	0.117	-3.753	0.022

3066	-19.745	0.012	-6.454	0.002
3067	-11.983	0.085	-4.324	0.032
3068	-5.991	0.275	-2.192	0.085
3070	-5.991	0.284	-2.17	0.099
3073	-11.983	0.075	-3.9	0.026
3077	-5.991	0.275	-2.398	0.099
3082	-11.983	0.075	-4.727	0.021
3083	-6.07	0.281	-2.378	0.098
3087	-19.745	0.012	-6.33	0.003
3088	-9.873	0.111	-3.431	0.039
3091	-17.974	0.021	-5.408	0.007
3092	-5.991	0.275	-2.195	0.085
3094	-8.624	0.203	-3.067	0.067
3098	-8.67	0.199	-2.955	0.072
3099	-6.066	0.281	-2.499	0.093
3100	-5.991	0.275	-2.175	0.097
3103	-8.941	0.192	-3.288	0.06
3111	-5.991	0.275	-2.123	0.1
3115	-14.809	0.037	-5.985	0.006
3118	-11.983	0.075	-3.501	0.033
3120	-10.68	0.095	-3.271	0.026
3122	-17.974	0.021	-5.364	0.008
3128	-11.854	0.081	-3.809	0.025
3130	-13.51	0.049	-4.022	0.016
3131	-11.014	0.096	-4.448	0.016
3135	-9.873	0.111	-3.081	0.035

3141	-17.675	0.057	-5.445	0.025
3143	-11.983	0.075	-3.518	0.027
3150	-11.983	0.085	-3.494	0.043
3151	-9.873	0.111	-3.612	0.027
3154	-9.873	0.113	-3.605	0.022
3161	-11.983	0.075	-3.483	0.028
3163	-11.983	0.075	-3.496	0.03
3166	-11.983	0.085	-3.546	0.042
3167	-5.991	0.275	-2.178	0.086
3169	-11.86	0.083	-3.802	0.028
3170	-11.738	0.083	-3.735	0.026
3174	-12.137	0.079	-5.002	0.018
3175	-11.734	0.083	-3.479	0.029
3176	-9.873	0.111	-2.949	0.037
3178	-10.705	0.095	-3.504	0.023
3179	-9.803	0.199	-3.511	0.058
3181	-9.873	0.111	-3.612	0.022
3188	-13.652	0.113	-5.21	0.02
3190	-6.077	0.284	-2.195	0.098
3191	-12.15	0.14	-4.771	0.045
3195	-6.085	0.278	-2.415	0.091
3196	-12.007	0.075	-3.591	0.029
3197	-11.855	0.112	-4.442	0.029
3198	-25.965	0.063	-8.868	0.019
3199	-9.892	0.111	-3.636	0.029
3200	-24.013	0.006	-10.678	0.001

3201	-14.838	0.04	-5.97	0.004
3202	-10.686	0.095	-3.972	0.025
3210	-9.892	0.111	-2.934	0.031
3218	-14.838	0.037	-4.396	0.016
3226	-9.892	0.111	-3.581	0.023
3227	-24.731	0.004	-8.015	0.001
3228	-9.892	0.111	-2.938	0.04
3230	-14.838	0.037	-5.865	0.006
3231	-29.677	0.001	-10.741	0
3233	-12.007	0.075	-3.418	0.029
3235	-12.164	0.079	-4.954	0.018
3239	-14.765	0.038	-5.284	0.008
3244	-11.755	0.15	-3.495	0.069
3246	-18.847	0.061	-6.216	0.065
3249	-13.432	0.14	-6.662	0.081
3254	-8.978	0.239	-3.309	0.062
3259	-14.838	0.037	-4.363	0.012
3261	-14.838	0.037	-5.522	0.007
3262	-19.785	0.012	-7.308	0.002
3264	-12.007	0.075	-4.775	0.013
3267	-13.071	0.152	-9.284	0.045
3281	-12.007	0.075	-3.46	0.028
3284	-23.928	0.124	-8.055	0.093
3288	-12.028	0.079	-4.656	0.018
3289	-19.785	0.012	-6.471	0.002
3291	-6.003	0.275	-2.196	0.085

3296	-14.838	0.037	-6.304	0.004
3299	-6.079	0.281	-2.487	0.094
3303	-8.15	0.224	-2.705	0.086
3308	-11.769	0.083	-3.474	0.029
3309	-9.892	0.111	-2.93	0.034
3311	-30.017	0.002	-11.043	0
3319	-36.371	0	-17.039	0
3327	-36.02	0	-18.234	0
3336	-11.785	0.149	-3.545	0.068
3339	-8.192	0.223	-3.258	0.066
3344	-9.892	0.111	-2.907	0.04
3347	-9.892	0.111	-3.576	0.029
3350	-14.838	0.037	-4.436	0.016
3353	-12.007	0.075	-4.602	0.024
3357	-15.737	0.066	-4.921	0.062
3371	-14.838	0.04	-5.892	0.005
3372	-9.892	0.117	-3.687	0.022
3373	-9.892	0.111	-3.075	0.035
3374	-11.783	0.083	-3.586	0.028
3375	-9.892	0.117	-3.751	0.022
3376	-12.007	0.085	-3.534	0.043
3378	-17.774	0.056	-6.104	0.02
3385	-24.731	0.004	-8.749	0
3386	-23.338	0.071	-8.867	0.019
3390	-12.007	0.085	-3.463	0.044
3391	-11.831	0.111	-4.204	0.032

3392	-14.838	0.037	-5.816	0.004
3396	-18.01	0.021	-5.685	0.004
3399	-5.49	0.309	-2.12	0.092
3404	-19.785	0.012	-7.269	0.002
3405	-12.149	0.14	-5.25	0.039
3407	-24.013	0.006	-9.574	0.001
3408	-18.01	0.021	-5.275	0.008
3412	-14.838	0.037	-5.655	0.005
3415	-18.231	0.022	-8.49	0.003
3417	-9.892	0.111	-3.827	0.025
3418	-14.838	0.037	-6.029	0.006
3427	-11.771	0.084	-3.581	0.031
3429	-12.159	0.078	-5.329	0.015
3431	-12.007	0.075	-4.558	0.014
3432	-8.978	0.249	-3.338	0.063
3433	-12.366	0.111	-3.816	0.09
3441	-9.537	0.193	-3.423	0.059
3444	-16.494	0.03	-6.143	0.005
3448	-9.892	0.111	-2.92	0.038
3456	-18.192	0.022	-8.269	0.003
3459	-22.515	0.073	-6.642	0.026
3461	-9.892	0.111	-2.866	0.042
3462	-24.013	0.006	-7.228	0.003
3468	-6.078	0.281	-2.501	0.093
3469	-4.946	0.342	-1.992	0.098
3472	-12.007	0.085	-4.524	0.029

3473	-18.01	0.021	-8.177	0.003
3474	-19.785	0.013	-7.936	0.001
3476	-9.048	0.195	-3.721	0.053
3477	-16.328	0.028	-6.502	0.005
3478	-14.838	0.043	-5.902	0.008
3481	-8.413	0.211	-2.826	0.079
3485	-9.892	0.111	-3.575	0.023
3486	-16.015	0.03	-5.003	0.007
3487	-24.013	0.007	-9.915	0.001
3496	-6.003	0.275	-2.18	0.097
3500	-14.797	0.037	-5.96	0.005
3501	-19.745	0.012	-7.52	0.001
3503	-23.965	0.006	-7.526	0.001
3506	-14.809	0.044	-5.951	0.008
3509	-14.681	0.065	-5.442	0.016
3510	-9.873	0.111	-3.689	0.022
3511	-12.141	0.079	-5.362	0.016
3515	-12.086	0.08	-4.91	0.019
3518	-17.974	0.021	-5.322	0.007
3519	-18.185	0.019	-5.596	0.005
3520	-15.508	0.077	-4.644	0.102
3522	-9.873	0.111	-2.903	0.032
3523	-11.983	0.085	-4.507	0.03
3524	-12.341	0.111	-3.925	0.072
3526	-17.974	0.021	-6.772	0.005
3529	-11.885	0.082	-3.873	0.027

3539	-9.873	0.117	-4.151	0.018
3542	-15.997	0.029	-4.753	0.01
3545	-22.213	0.018	-10.328	0.003
3547	-9.873	0.111	-3.709	0.023
3549	-11.983	0.075	-4.531	0.018
3553	-9.873	0.111	-2.883	0.04
3554	-9.873	0.111	-3.493	0.038
3556	-18.184	0.052	-7.682	0.013
3557	-5.991	0.275	-2.194	0.097
3562	-14.809	0.037	-4.568	0.009
3565	-9.873	0.111	-4.102	0.018
3566	-10.665	0.095	-4.28	0.022
3567	-17.285	0.045	-6.358	0.034
3568	-9.873	0.111	-3.138	0.035
3571	-43.077	0.006	-14.562	0.001
3573	-9.873	0.111	-3.552	0.023
3577	-9.873	0.111	-2.924	0.049
3581	-21.106	0.01	-6.757	0.002
3582	-19.745	0.012	-7.138	0.003
3585	-21.401	0.009	-6.722	0.002
3589	-17.974	0.021	-7.018	0.004
3590	-19.745	0.012	-6.069	0.003
3596	-19.745	0.012	-7.481	0.001
3607	-14.598	0.039	-4.559	0.015
3608	-9.873	0.117	-3.816	0.021
3617	-9.873	0.111	-3.003	0.033

3620	-24.681	0.004	-10.522	0
3623	-9.873	0.111	-3.608	0.029
3624	-14.809	0.041	-5.027	0.01
3625	-23.965	0.007	-10.577	0.001
3630	-5.991	0.275	-2.218	0.084
3631	-17.974	0.021	-5.507	0.006
3633	-11.983	0.085	-4.496	0.03
3634	-9.873	0.112	-3.684	0.028
3635	-14.809	0.037	-4.728	0.007
3639	-5.991	0.284	-2.194	0.098
3646	-9.873	0.111	-2.872	0.05
3647	-9.873	0.111	-3.675	0.035
3650	-10.665	0.095	-3.208	0.034
3651	-14.809	0.037	-4.8	0.008
3652	-9.873	0.111	-3.003	0.03
3662	-24.681	0.004	-10.513	0
3668	-11.983	0.075	-3.48	0.034
3671	-12.1	0.078	-5.162	0.016
3672	-11.983	0.075	-4.01	0.028
3680	-33.26	0.001	-11.711	0
3681	-9.873	0.111	-3.596	0.023
3686	-10.659	0.095	-3.211	0.027
3691	-14.809	0.037	-6.276	0.005
3692	-11	0.096	-4.293	0.017
3698	-8.386	0.22	-3.302	0.064
3699	-5.991	0.275	-2.195	0.097

3701	-19.745	0.012	-6.315	0.002
3703	-39.49	0	-16.54	0
3705	-14.809	0.037	-4.724	0.007
3714	-11.983	0.075	-3.659	0.026
3718	-9.873	0.111	-2.876	0.031
3720	-25.448	0.065	-8.683	0.019
3721	-10.671	0.11	-4.424	0.018
3725	-6.071	0.281	-2.582	0.09
3728	-14.809	0.037	-5.488	0.005
3732	-5.495	0.309	-2.133	0.092
3737	-8.797	0.194	-3.174	0.062
3741	-11.983	0.075	-4.489	0.019
3743	-17.616	0.024	-5.459	0.007
3750	-11.002	0.096	-4.476	0.016
3755	-17.669	0.024	-5.462	0.008
3756	-9.873	0.113	-3.259	0.026
3758	-5.991	0.275	-2.124	0.1
3762	-17.974	0.021	-7.505	0.003
3766	-23.965	0.007	-9.692	0.001
3767	-14.809	0.037	-4.741	0.01
3772	-16.506	0.03	-6.866	0.003
3776	-23.965	0.006	-7.757	0.002
3780	-9.873	0.111	-3.512	0.023
3783	-9.873	0.111	-3	0.03
3790	-9.873	0.111	-3.673	0.021
3793	-9.873	0.111	-2.972	0.037

3795	-11.983	0.075	-4.775	0.017
3799	-9.873	0.124	-3.74	0.032
3802	-9.873	0.111	-3.516	0.023
3805	-44.953	0.005	-13.873	0.002
3809	-11.748	0.083	-3.571	0.028
3810	-14.809	0.037	-5.297	0.008
3811	-5.991	0.284	-2.326	0.093
3813	-14.809	0.037	-5.724	0.009
3814	-11.983	0.081	-3.477	0.029
3817	-9.873	0.111	-2.937	0.03
3820	-11.983	0.075	-3.456	0.031
3827	-14.809	0.037	-4.521	0.011
3830	-17.974	0.023	-5.241	0.007
3832	-9.019	0.195	-3.649	0.054
3835	-10.936	0.091	-4.017	0.024
3838	-9.873	0.116	-3.73	0.022
3847	-22.103	0.008	-7.508	0.002
3849	-9.873	0.111	-2.966	0.039
3850	-19.745	0.012	-8.624	0.001
3858	-12.124	0.079	-5.184	0.017
3860	-4.936	0.333	-1.971	0.097
3863	-6.072	0.281	-2.683	0.087
3865	-9.873	0.111	-3.76	0.027
3867	-10.626	0.096	-3.172	0.028
3869	-8.188	0.219	-3.264	0.065
3870	-12.1	0.079	-4.859	0.019

3872	-9.873	0.111	-3.912	0.02
3876	-4.936	0.333	-1.94	0.098
3877	-9.873	0.111	-3.628	0.024
3882	-0.018	0.001	0	1
3886	-5.998	0.275	-2.192	0.097
3887	-11.996	0.075	-3.411	0.029
3895	-17.296	0.045	-6.719	0.016
3896	-8.879	0.191	-3.218	0.061
3906	-6.07	0.279	-2.309	0.095
3907	-12.148	0.078	-4.935	0.017
3909	-13.527	0.049	-3.951	0.017
3917	-14.825	0.037	-5.83	0.007
3918	-14.825	0.037	-5.433	0.006
3923	-17.441	0.072	-5.059	0.021
3924	-13.493	0.057	-5.068	0.016
3925	-10.945	0.095	-4.702	0.015
3938	-19.795	0.012	-7.155	0.003
3943	-9.898	0.111	-3.774	0.026
3946	-9.898	0.111	-3.899	0.022
3948	-9.898	0.119	-2.899	0.044
3950	-6.007	0.275	-2.31	0.092
3953	-14.846	0.037	-6.434	0.005
3959	-11.765	0.084	-3.504	0.032
3960	-12.168	0.079	-5.276	0.017
3963	-19.795	0.012	-7.529	0.001
3964	-18.02	0.021	-5.664	0.007

3967	-18.02	0.025	-6.932	0.008
3968	-11.773	0.149	-3.576	0.067
3969	-18.214	0.053	-7.782	0.012
3970	-14.846	0.037	-5.807	0.004
3974	-29.693	0.001	-9.735	0
3980	-12.013	0.075	-3.428	0.037
3981	-14.475	0.086	-4.956	0.025
3984	-11.274	0.118	-3.851	0.04
3985	-9.898	0.111	-3.744	0.028
3988	-19.795	0.012	-6.317	0.002
3991	-19.792	0.045	-7.643	0.027
3993	-14.846	0.037	-4.631	0.009
3995	-10.902	0.136	-3.76	0.045
4002	-6.007	0.275	-2.184	0.097
4003	-6.007	0.284	-2.192	0.098
4004	-12.013	0.081	-3.528	0.028
4010	-18.02	0.021	-5.432	0.007
4015	-9.898	0.111	-3.828	0.026
4016	-14.846	0.037	-4.537	0.008
4018	-5.51	0.309	-2.125	0.092
4021	-14.846	0.037	-5.601	0.007
4027	-6.007	0.284	-2.22	0.097
4030	-9.898	0.111	-3.894	0.025
4033	-12.013	0.085	-4.483	0.03
4034	-12.013	0.075	-3.407	0.029
4038	-10.692	0.095	-3.987	0.024

4042	-6.007	0.275	-2.181	0.097
4045	-10.921	0.095	-4.947	0.013
4052	-10.667	0.096	-3.286	0.026
4053	-11.762	0.084	-3.501	0.032
4059	-14.846	0.037	-4.702	0.01
4063	-29.693	0.001	-9.766	0
4066	-9.898	0.111	-3.603	0.036
4068	-24.026	0.006	-11.362	0.001
4069	-9.898	0.111	-3.725	0.026
4072	-6.085	0.281	-2.615	0.089
4075	-9.898	0.111	-3.628	0.022
4080	-11.77	0.083	-3.594	0.028
4086	-6.086	0.278	-2.384	0.092
4087	-12.122	0.08	-4.643	0.02
4092	-6.073	0.279	-2.384	0.092
4093	-9.898	0.111	-2.974	0.037
4094	-12.372	0.111	-3.738	0.1
4095	-24.744	0.004	-8.614	0
4098	-9.898	0.111	-2.876	0.032
4099	-6.007	0.275	-2.431	0.088
4100	-16.821	0.044	-4.351	0.02
4104	-12.013	0.075	-3.5	0.022
4106	-10.692	0.095	-3.163	0.035
4109	-12.013	0.075	-4.558	0.022
4111	-9.633	0.136	-3.444	0.049
4114	-19.795	0.012	-7.79	0.002

4119	-9.898	0.111	-2.938	0.038
4124	-6.007	0.284	-2.152	0.1
4126	-4.949	0.333	-1.974	0.097
4134	-12.013	0.075	-4.548	0.024
4140	-9.898	0.117	-3.833	0.021
4142	-12.013	0.085	-3.642	0.041
4146	-6.079	0.279	-2.31	0.095
4149	-18.02	0.021	-7.15	0.005
4152	-11.824	0.111	-4.13	0.034
4156	-6.007	0.275	-2.177	0.086
4157	-6.082	0.281	-2.502	0.093
4158	-9.898	0.116	-3.976	0.02
4161	-6.087	0.281	-2.419	0.096
4168	-6.007	0.275	-2.125	0.088
4169	-18.02	0.021	-6.902	0.006
4172	-9.898	0.111	-3.812	0.025
4174	-5.421	0.322	-2.134	0.095
4179	-12.013	0.075	-5.893	0.008
4183	-6.088	0.278	-2.41	0.091
4186	-17.974	0.021	-5.333	0.01
4187	-9.873	0.117	-4.051	0.019
4191	-17.974	0.021	-5.492	0.007
4192	-9.873	0.112	-2.981	0.039
4193	-11.769	0.149	-3.637	0.065
4194	-11.983	0.075	-3.399	0.032
4195	-8.122	0.225	-3.319	0.064

4204	-12.282	0.086	-4.501	0.03
4207	-11.983	0.085	-3.448	0.044
4212	-17.974	0.021	-5.271	0.007
4213	-14.809	0.037	-5.405	0.01
4222	-11.983	0.075	-3.437	0.034
4223	-9.873	0.125	-3.899	0.03
4226	-12.028	0.113	-5.83	0.017
4227	-17.974	0.023	-5.372	0.007
4228	-6.071	0.278	-2.414	0.091
4230	-5.991	0.275	-2.123	0.1
4232	-6.136	0.274	-2.184	0.097
4236	-12.199	0.078	-5.219	0.015
4237	-5.539	0.31	-2.178	0.091
4245	-12.057	0.075	-3.414	0.037
4246	-11.081	0.133	-4.424	0.033
4249	-11.122	0.128	-4.588	0.03
4253	-12.634	0.087	-4.649	0.029
4258	-6.029	0.275	-2.182	0.097
4263	-9.911	0.113	-3.734	0.027
4264	-14.901	0.037	-4.574	0.008
4267	-14.901	0.037	-5.868	0.006
4273	-9.934	0.111	-3.027	0.036
4279	-12.183	0.079	-5.337	0.016
4282	-24.835	0.004	-8.71	0
4284	-11.833	0.083	-3.579	0.028
4287	-9.934	0.111	-3.705	0.026

4288	-9.934	0.111	-3.701	0.021
4299	-6.029	0.275	-2.176	0.098
4300	-11.818	0.149	-3.566	0.067
4305	-11.024	0.131	-3.694	0.046
4308	-8.786	0.214	-3.421	0.055
4309	-11.054	0.096	-4.385	0.017
4310	-5.542	0.309	-2.174	0.091
4311	-14.901	0.037	-4.461	0.008
4317	-5.532	0.309	-2.152	0.091
4321	-18.086	0.021	-5.844	0.006
4323	-24.835	0.004	-8.156	0
4328	-14.901	0.037	-5.722	0.005
4331	-18.086	0.023	-5.346	0.007
4332	-9.099	0.195	-3.751	0.052
4335	-15.909	0.146	-7.585	0.068
4343	-9.934	0.111	-3.686	0.028
4355	-14.901	0.037	-4.539	0.011
4356	-6.109	0.281	-2.391	0.097
4357	-6.029	0.275	-2.278	0.094
4359	-16.388	0.05	-5.865	0.014
4362	-19.868	0.012	-6.273	0.005
4364	-24.115	0.006	-9.671	0.001
4371	-24.115	0.006	-7.613	0.003
4374	-9.934	0.111	-3.801	0.026
4379	-9.934	0.117	-3.792	0.021
4381	-9.016	0.172	-2.796	0.052

4385	-9.934	0.111	-3.655	0.027
4386	-12.057	0.075	-3.423	0.037
4389	-5.54	0.309	-2.202	0.09
4390	-11.081	0.096	-4.483	0.016
4394	-12.057	0.075	-3.475	0.031
4395	-13.748	0.074	-3.927	0.027
4396	-13.915	0.067	-3.519	0.026
4401	-9.934	0.111	-3.904	0.02
4403	-6.11	0.278	-2.692	0.082
4404	-45.594	0.005	-17.396	0.001
4406	-9.934	0.111	-3.606	0.029
4411	-6.095	0.279	-2.381	0.092
4415	-6.108	0.278	-2.385	0.092
4423	-11.817	0.083	-3.577	0.028
4429	-8.235	0.221	-2.733	0.085
4436	-14.769	0.065	-5.333	0.017
4438	-11.935	0.109	-4.223	0.032
4445	-6.029	0.275	-2.2	0.085
4448	-44.703	0	-19.642	0
4452	-16.608	0.03	-7.093	0.003
4453	-9.934	0.111	-3.873	0.022
4462	-9.934	0.111	-3.732	0.027
4466	-6.029	0.275	-2.628	0.07
4472	-9.047	0.196	-3.84	0.05
4476	-12.206	0.079	-5.205	0.017
4480	-9.934	0.111	-2.892	0.039

4481	-9.383	0.125	-3.508	0.039
4484	-14.901	0.037	-4.626	0.007
4486	-9.934	0.111	-3.85	0.026
4487	-17.802	0.024	-5.591	0.007
4499	-12.057	0.085	-3.481	0.044
4502	-12.057	0.075	-3.703	0.02
4504	-22.596	0.073	-6.651	0.026
4506	-9.934	0.111	-3.599	0.023
4510	-9.934	0.116	-3.708	0.022
4515	-12.057	0.075	-4.597	0.018
4516	-11.123	0.089	-3.622	0.027
4520	-9.934	0.111	-3.911	0.025
4522	-18.273	0.022	-7.893	0.004
4524	-19.309	0.151	-7.816	0.098
4526	-6.109	0.278	-2.391	0.092
4527	-12.194	0.079	-4.774	0.02
4529	-12.213	0.078	-5.479	0.014
4530	-11.96	0.081	-3.886	0.025
4533	-14.901	0.037	-4.486	0.008
4534	-25.96	0.064	-8.632	0.019
4535	-9.093	0.195	-3.765	0.052
4536	-14.901	0.041	-4.666	0.012
4541	-8.96	0.198	-3.72	0.052
4550	-27.188	0.003	-9.263	0.001
4551	-11.806	0.084	-3.483	0.032
4554	-9.934	0.111	-3.15	0.035

4556	-6.029	0.275	-2.278	0.082
4565	-14.475	0.066	-5.429	0.016
4567	-4.936	0.333	-1.944	0.098
4575	-27.634	0.008	-8.211	0.012
4578	-17.974	0.025	-5.028	0.018
4581	-17.974	0.025	-5.496	0.012
4589	-9.873	0.111	-2.906	0.032
4591	-14.771	0.037	-4.458	0.016
4593	-11.983	0.075	-4.401	0.023
4594	-14.809	0.037	-5.855	0.005
4595	-9.873	0.111	-3.659	0.022
4597	-11.006	0.096	-4.688	0.015
4600	-5.507	0.309	-2.169	0.091
4604	-34.554	0	-12.452	0
4605	-23.572	0.007	-7.936	0.002
4613	-8.971	0.251	-3.571	0.056
4615	-9.873	0.111	-3.68	0.021
4616	-10.665	0.095	-3.87	0.026
4620	-14.809	0.037	-4.422	0.012
4621	-14.809	0.038	-4.469	0.008
4630	-11.741	0.084	-3.514	0.032
4639	-17.974	0.021	-7.973	0.004
4647	-9.873	0.111	-3.848	0.026
4665	-5.991	0.284	-2.176	0.099
4670	-21.697	0.022	-6.764	0.022
4672	-9.873	0.111	-3.557	0.037

4676	-12.121	0.079	-4.702	0.02
4688	-6.067	0.279	-2.484	0.089
4689	-9.873	0.112	-2.918	0.031
4690	-17.343	0.034	-6.053	0.01
4692	-9.873	0.117	-3.931	0.02
4698	-15.509	0.077	-5.832	0.071
4699	-13.454	0.093	-4.113	0.07
4703	-18.213	0.018	-5.574	0.005
4709	-20.368	0.07	-10.12	0.034
4712	-15.676	0.068	-6.473	0.047
4719	-9.873	0.111	-4.195	0.022
4720	-14.809	0.037	-4.702	0.01
4722	-10.665	0.095	-3.149	0.035
4725	-11.742	0.149	-3.547	0.068
4726	-6.072	0.278	-2.683	0.082
4729	-9.873	0.111	-2.883	0.05
4730	-9.873	0.111	-3.909	0.024
4733	-11.107	0.088	-4.008	0.024
4739	-14.809	0.037	-4.955	0.006
4741	-15.997	0.029	-6.043	0.006
4743	-9.816	0.112	-3.608	0.036
4754	-40.445	0.017	-13.962	0.011
4755	-14.809	0.037	-5.343	0.007
4758	-8.126	0.224	-2.779	0.083
4759	-6.07	0.281	-2.382	0.097
4764	-10.665	0.095	-3.907	0.025

4765	-23.965	0.006	-8.2	0.002
4767	-8.892	0.199	-3.797	0.05
4770	-11.983	0.075	-4.9	0.018
4772	-5.991	0.275	-2.398	0.099
4775	-11.983	0.075	-3.468	0.031
4777	-9.873	0.111	-2.979	0.039
4781	-13.58	0.048	-4.011	0.016
4786	-9.873	0.111	-3.97	0.018
4789	-14.809	0.037	-4.977	0.009
4799	-14.809	0.037	-6.074	0.005
4808	-10.96	0.133	-3.845	0.042
4809	-11.967	0.112	-5	0.024
4810	-9.873	0.111	-3.205	0.026
4817	-15.706	0.066	-4.661	0.07

* cutoff $p \leq 0.10$

Highlighted rows indicate sites that differentiate Fatal and Silent strains