COMPUTATIONAL ANALYSIS OF THE PHYLOGENOMIC RELATIONSHIPS OF HUMAN AND SIMIAN ADENOVIRUSES AND THE CROSS-SPECIES TRANSMISSIONS

by

June Kang

A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Bioinformatics and Computational Biology

Committee:

_____	Dr. Donald Seto, Committee Chair
_____	Dr. Patrick Gillevet, Committee Member
_____	Dr. Haw Chuan Lim, Committee Member
_____	Dr. Iosif Vaisman, Director, School of Systems Biology
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Fernando R. Miralles-Wilhelm, Dean, College of Science
Date: _____	Fall Semester 2020 George Mason University Fairfax, VA

Computational Analysis of the Phylogenomic Relationships of Human and Simian Adenoviruses and the Cross-Species Transmissions

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

June Kang
Bachelor of Science
James Madison University, 2016

Director: Donald Seto, Professor
Bioinformatics and Computational Biology

Fall Semester 2020
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to Hong Xia Zhang and Dongwei Kang whose sacrifices made this document possible.

ACKNOWLEDGEMENTS

I would like to thank the Dr. Donald Seto for his patience, motivation, and valuable guidance. The other members of my committee for their insightful comments and encouragement. Finally, thanks go out to the other members of Dr. Seto's lab (Daniel Negron and Kalpana Dommaraju) for their help and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Adenovirus	AdV
Acute respiratory diseases	ARD
Base pair	bp
Human Adenovirus	HAdV
International Committee on Taxonomy of Viruses	ICTV
Inverted Terminal Repeats	ITR
Multiple Alignment using Fast Fourier Transform	MAFFT
Multiple Sequence Alignment	MSA
Nuclear Factor I	NF-I
Nuclear Factor III	NF-III
Phylogenetic Analysis Using Parsimony	PAUP
Primate Adenovirus	PrAdV
Single nucleotide polymorphism	SNP
Simian Adenovirus	SAdV
Tree Analysis using New Technology	TNT

ABSTRACT

COMPUTATIONAL ANALYSIS OF THE PHYLOGENOMIC RELATIONSHIPS OF HUMAN AND SIMIAN ADENOVIRUSES AND THE CROSS-SPECIES TRANSMISSIONS

June Kang, Ph.D.

George Mason University, 2022

Dissertation Director: Dr. Donald Seto

This thesis describes the

The phylogenetic relationships of HAdV and SAdV are not formally well defined by the International Committee on Taxonomy of Viruses (ICTV). All archived genomes will be analyzed at high resolution using genomics and bioinformatics to provide a comprehensive reference. Given their evolutionary closeness, possible zoonotic and anthroponosis transfers between human and non-human simian hosts will be examined using HAdV-E4 and SAdV-21 as examples, based on their clustering in the phylogenomic trees.

CHAPTER 1: Human and (non-human) Simian Adenoviruses

Adenoviruses (AdV) are medium-size (80 nm – 100 nm), non-enveloped, double-stranded DNA viruses that are apparently found in all vertebrates, including fish, amphibians, reptiles, birds, and mammals (Harrach et al., 2011). They comprise the family *Adenoviridae,* and are further divided into six genera (Harrach et al., 2011). Human adenoviruses (HAdV) and non-human simian adenoviruses (SAdV) are the most studied viruses in the *Adenoviridae*. These belong to the genus *Mastadenovirus* (Harrach et al., 2011). All *Mastadenovirus* share a similar capsid morphology and genome organization (Figure 1.1) (Davison et al., 2003; Norrby, 1969). The AdV capsid is an icosahedron consisting of 252 capsomers with 240 hexon proteins forming the surface and 12 penton bases at the verticals; each penton base bears a fiber protein that is used for cellular recognition and viral entry.

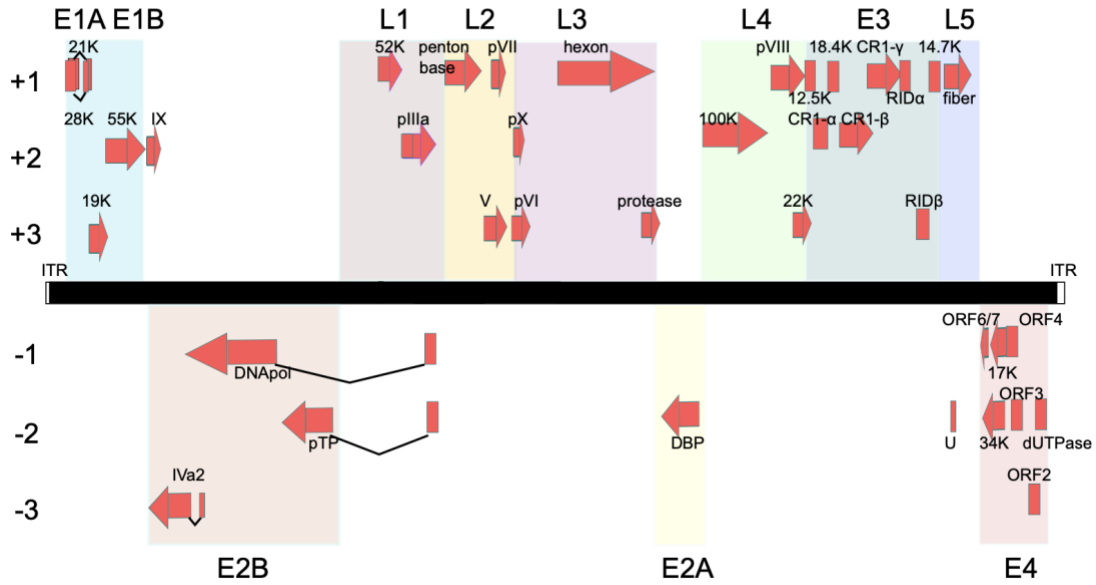


Figure 1.1. Genome organization of a representative adenovirus. All protein-coding genes are expressed within the six reading frames, as shown by colored blocks with direction of transcription noted by arrows. Early (E) and late (L) phases (genes) are highlighted as different color and labeled for better visualization. Genes are arranged in the forward or backward reading frames. The overlapping red arrows within each gene (colored box) represent alternatively spliced gene product.

HAdVs are distributed into species A to G the traditional taxonomic classification scheme based on biological properties including tropism (Davison et al., 2003) and diseases, with an additional seventh species G identified and characterized recently through genome analysis (Jones et al., 2007). There are 51 serotypes as defined by serology-based methods, originally, based on serum neutralization and hemagglutination

assays that targeted the hexon (epsilon epitope) and the fiber (gamma epitope) proteins. Currently, 104 genotypes, including the re-characterized original 51 serotypes, of HAdV are recognized and fully characterized with genome sequence data analysis and vetted by the Human Adenovirus Working Group (http://hadvwg.gmu.edu/). SAdVs are arbitrarily distributed into simian species A to C by different researchers (Chen et al., 2011; Chiu et al., 2013; Malouli et al., 2014; Roy et al., 2012), with some literature noting SAdV species A to G (Panto et al., 2015). However, these simian species are not recognized formally by the International Committee on Taxonomy of Viruses (ICTV) Adenovirus Working Group (Harrach et al., 2011).

HAdVs can cause a range of diseases in the respiratory, ocular, renal and urinary tract infections, and gastrointestinal tracts (Chu & Pavan-Langston, 1979; Gaydos & Gaydos, 1995; Lion, 2014; Wood, 1988). Different species of AdVs can cause diverse infections in various organs (Table 1.1), presumably as a function of cell type recognition and tropism. For example, HAdV species F and G viruses are commonly associated with gastroenteritis (Wood, 1988); species C infects the respiratory tracts and liver, and are also associated with latent infections (Garnett et al., 2009; Shayakhmetov et al., 2005; Yao et al., 2019); and species E viruses cause acute respiratory diseases (ARD), as well as ocular diseases such as conjunctivitis and keratoconjunctivitis (Lion, 2014; Rowe et al., 1953).

Table 1.1 Common diseases that are associated with each HAdV species. Common diseases associated and the tissue infected are listed for each HAdV species[1]. Number of types of HAdV for each species are catalogued.

Species	Type	Tissue	Common disease associated
A	4	GI, Lung, Bladder	Gastroenteritis tract, respiratory tract infection, urogenital tract infection
B	17	Lung, Ocular, GI, Bladder	Respiratory tract infection*, ocular tract infection, gastroenteritis tract, urogenital tract infeciton
C	5	Lung, GI, Bladder	Respiratory tract infection, urogenital tract infection
D	68	Ocular, GI	Ocular tract infection§, gasroenteritis tract,
E	1	Lung, Ocular	Respirtory tract infections*, ocular tract infection
F	2	GI	Gastroenteritis tract
G	1	GI	Gastroenteritis tract

* ARD: acute respiratory disease; commonly associated with B3, E4, B16, and B21

§ EKC: epidemic conjunctivitis; commonly associated with type 8, 19, 37, 53, and 54

Emergent and novel adenoviruses can arise as novel types through genome recombination and may cause pathogenicity. The HAdV genome has been shown to undergo recombination as an evolution mechanism, in which one, several, or all of the serologically important epitopes, comprising major capsid proteins, may be shuffled. In particular, the hexon and fiber proteins are useful markers for characterizing

[1] Citations for common disease associated (Couch et al., 1963; Dehghan, Seto, Jones, et al., 2013; Garnett et al., 2009; Jones et al., 2007; Li et al., 2018; Walsh et al., 2010; Wood, 1988)

recombinants due to their historical characterization using serology typing protocols. The third capsid protein, penton base, is also monitored for determining novel genotypes due to frequent recombination at the hypervariable region (Figure 1.2). Theoretically, and in practice, other genome regions such as the E3 region can also recombine and may be biologically important, but these are not formally used currently for assigning novel genotype designations, mainly due to objections from the adenovirus research community. These are all covered in chapter 2. Understanding recombination in HAdVs is crucial for understanding the role that molecular evolution plays in generating emergent human and zoonotic pathogens.

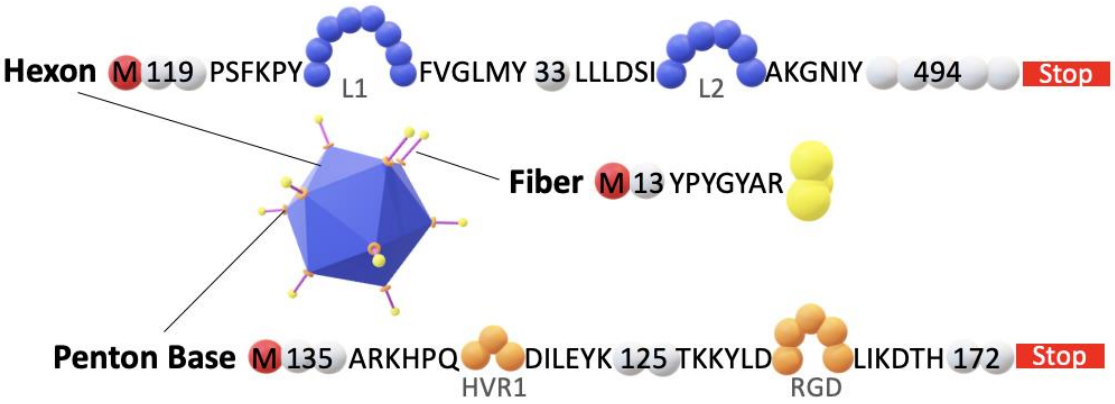


Figure 1.2. The three major adenovirus capsid proteins are presented along with their amino acid sequence and location on the capsid. Hexon, fiber, and penton base proteins are labeled based on amino acid composition, and based on HAdV-D43 as reference.

Epitopes for the capsid proteins were colored based on the 3-D adenovirus model using 3D Builder (Microsoft Corporation, 2013).

Zoonosis is the transmission of a pathogen from non-human animals to humans. A successful zoonotic virus needs to not only demonstrate transfection across various animal species, but also to be able to maintain sustainable infections in the newly-acquired hosts, which may result in novel and possible severe pathologies in the new host. An important example of this is influenza virus. Many emergent and highly contagious viral pathogens apparently have zoonotic origins (Feldmann, 2014; Guan et al., 2003; Towner et al., 2009). For example, the first severe acute respiratory syndrome coronavirus (SARS-CoV) was a recent zoonotic pathogen that crossed species from bats to humans (Lau et al., 2005); it rapidly spread through direct and indirect contacts among humans (Lu et al., 2013), presumably having adapted to the new host. Traces of SARS-CoV have also been isolated in raccoon dogs, ferret badgers, and Asian palm civets (Guan et al., 2003). Ebola virus (EBOV) is another zoonotic pathogen that can infect a range of species such as bats, pigs, great apes, and humans (Feldmann, 2014; Towner et al., 2009). Many other zoonotic viral pathogens were also identified in recent decades (Taylor et al., 2001). Zoonotic viral pathogens are often difficult to study due to its high infection and mortality rate (Gebreyes et al., 2014), that lead to limited presence or may not have appropriate hosts available to study as models. To understand zoonotic viral pathogens better, a model organism is required, preferably one that has been studied

extensively. Therefore, adenoviruses are a good model, as they have been studied and are well-documented in the literature since 1952.

Zoonotic transfers have occurred in the evolution of human adenoviruses, but were unidentified due to the limitation of available tools. An example is human adenovirus type 4 (HAdV-E4). HAdV-E4 was one of the first identified HAdVs, isolated from an outbreak in a military population in 1952 (Rowe et al., 1953). It is a major human respiratory pathogen and is highly contagious, resulting in frequent large outbreaks in certain populations, including the dense and overstressed military recruits in training bases. Inexplicitly, they appeared largely constrained to the U.S. military population (Gaydos & Gaydos, 1995). It was such an important pathogen that a vaccine was developed 2 – 3x to stop the frequent outbreaks (Gaydos & Gaydos, 1995; Graham & Prevec, 1992; Top et al., 1971). This vaccine was very effective in preventing outbreaks of this pathogen (Couch et al., 1963; Gaydos & Gaydos, 1995; Top et al., 1971). Using genomics techniques, it was shown that although HAdV-E4 is a major human pathogen, its genome was nearly identical to chimpanzee AdV genomes. HAdV-E4 is a recombinant of SAdV-E26 (89.2 % of the genome) and HAdV-E4 hexon domain (loop 1) (Dehghan et al., 2013). Chapter 3 will cover the analysis of recent HAdV-E4 isolates and the additional recombination in the ITR region.

This dissertation is designed to reveal the phylogenetic relationships between HAdV and SAdV using high resolution tools providing genomic and phylogenomic analyses. It

characterizes the putative host-adaptation of recent HAdV-E4 strains isolated from the various civilian population by analyzing the replication motifs contained within the ITR as well as the whole genome. Exploring the evolution mechanism and host adaptation of zoonotic pathogens using HAdV-E4 as a model organism will help enhance our understanding of how zoonotic pathogens evolve and adapt, which may lead to novel and, emergent, pathogens. Increased knowledge and information in this area will lead to increased awareness and interest in biosurveillance in the public.

Significance/Implications

Understanding and resolving HAdV and SAdV phylogenetic relationships will lead to the further development of biomedical applications such as gene therapy and vaccine development, both of which often use HAdV as delivery vectors (Graham & Prevec, 1992; Roy et al., 2004, Roy et al., 2006). Additionally, studying the zoonotic genotype HAdV-E4 will aid in understanding the role cross-species transmission plays in the evolution of novel pathogens and in the subsequent host adaptation of viruses in.

Specific Aims

The overall goal is to resolve formally the phylogenetic relationships between HAdVs and SAdVs, currently recognized to be separate lineages by the ICTV. HAdV and SAdV share similar genome organizations, homologous genes, evolutionary gene conservation, diversification, and cross-species host transmissions. The specific aims of this dissertation are as follows:

1) To explore zoonosis by examining the molecular evolution and distribution of recent isolates of HAdV-E4, as they appear to be host-adapted zoonotic pathogens. One approach is through examining the ITR, which contain host-interacting viral replication functions. This may provide understanding of the host adaptative mechanisms that may lead to larger outbreaks outside of the U.S. military population and globally.
2) To analyze recent HAdV-E4 strains collected from a large civilian outbreak in Beijing, as well as from sporadic infections in Hong Kong and Singapore.
3) To conduct in-depth analysis of SAdV-21, a chimpanzee virus known to fall in the same clade as human species B. The evolution and transmission history of SAdV-21 is an interesting and illustrative case study for possible zoonosis and/or anthroponosis in a SAdV.
4) To provide a high-resolution, comprehensive view of the phylogenetic relationships between HAdV and SAdV using whole genome sequences and computational tools.

Definition of Terms

Zoonosis: The transmission of pathogen from non-human animal to human

Anthrozoonosis: The transmission of pathogen from human to non-human animal.

Amphizoonosis: The transmission of pathogen between human and non-human animals.

Outline

- Chapter 1 – Introduction and overview of a model viral organism, human adenovirus, that is historically and currently important with respect to understanding molecular evolution and zoonosis using the genomics and bioinformatics approaches.
- Chapter 2 – Review and update on the current status and understanding of human adenoviruses typing in the era of genomics and bioinformatics tools.
- Chapter 3 –Examine a recognized human adenovirus, HAdV-E4, as a model of zoonosis and its phylogenetic relationship with chimpanzee adenoviruses.
    - The viral replication elements from prototype to current strains will be studied. A recent evolutionary change in the HAdV-E4 inverted terminal repeat (ITR) region containing viral replication motifs indicate host adaptation and may predict outbreaks in the general population.
    - A predicted shift of HAdV-E4 epidemics from the U.S. military to the general civilian populations is explored using genomes from recent outbreaks in Beijing, Hong Kong, and Singapore.

- Chapter 4 – Examine a recognized simian adenovirus, SAdV-21, as a model for zoonosis and its phylogenetic relationship with human adenoviruses.
- Chapter 5 – Resolve of the phylogenetic relationships between human (HAdV) and non-human (SAdV) simian adenoviruses using genomics and bioinformatics approaches.
    - Re-evaluation of the relationships of HAdV and SAdV using taxonomic approaches.
- Chapter 6 – Conclusions

CHAPTER 2: Status of Adenoviruses, Updated

Adenoviruses were first isolated during a poliovirus study in 1952 (Friedman & Bennett, 1957; Hilleman & Werner, 1954; Rowe et al., 1953). Originally, serotyping based on the hexon protein or epsilon epitope was how adenoviruses were identified, characterized, named, and typed. This also included the fiber protein or gamma epitope in later protocols as hexon typing was found to be inconclusive for some types (note to June- this would be the Ads15 and 29 problem- Wigand et al). Since then, adenoviruses typing has transitioned to a DNA sequenced-based, and even high throughput genomic sequencing, protocol (Friedman & Bennett, 1957; Seto et al., 2011, 2013). This chapter will provide a comprehensive overview of the transition from serotyping to genotyping, and highlight the difficulties of these changes. The chapter will present support for recombination as a major molecular evolution pathway for new or re-emerging adenovirus pathogens. Additionally, this chapter will summarize and update the resource provided to the scientific community by the Human Adenovirus Working Group (http://hadvwg.gmu.edu). This is a platform to catalogue and present evaluated novel human adenovirus types using genome sequence-based standardized parameters under a systematic guideline.

Typing human adenovirus using genome sequence data and upkeep of a community resource for human adenoviruses

Typing human adenoviruses using genome sequence data and upkeep of a community resource for human adenoviruses

June Kang[1], Kalpana Dommaraju[1], Elizabeth B. Liu, Shoaleh Dehghan, David W. Dyer, James Chodosh[2*], Donald Seto[1*]

[1]Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

[2]Department of Ophthalmology, Howe Laboratory, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts 02114, USA

*Corresponding author: Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

E-mail address: dseto@gmu.edu (Donald Seto)

Keywords

Molecular evolution; Recombination; Genome sequencing; Serum neutralization; Hemagglutination inhibition; Respiratory pathogens; Hexon; Fiber; Penton base; Zoonosis; Anthroponosis; Human Adenovirus Working Group.

Abstract

Recognizing and typing novel human adenovirus using only two capsid proteins as identifying parameters have been ambiguous due to frequent recombination occurring in the genome sequence. This report reviews the adenovirus typing transition from serology-based to genome sequence-based and provide support for using high-throughput sequencing in genomic analysis. A new perspective and understanding of HAdV evolution based on genome recombination as a major contributing molecular evolutionary mechanism is highlighted, with several examples of novel human adenovirus types. Additionally, zoonosis and anthroponosis, as an underappreciated pathway for virus pathogen transmission in adenovirology, are also introduced. Lastly, the Human Adenovirus (HAdV) Working Group provides current and updated information on novel types. The HAdV Working Group, working with NIH, was established to provide advice for recognizing GenBank submissions of candidate novel genotypes and this group has maintained a platform for information exchange and a reference resource among adenovirologists.

1. INTRODUCTION

Human adenovirus (HAdV) is a non-enveloped, double-stranded DNA virus that belongs to the genus Mastadenovirus (Harrach et al. 2011; Lion 2014). It is about 65 to 80 nm in diameter. The viral genome is about 26 to 46 kbp long and contains 23 to 46 protein-coding genes (Swenson et al. 2003). The virion is composed of a protein capsid, nucleoprotein core, and minor internal proteins virion includes genetic material. This capsid is an icosahedron that is composed of 252 capsomeres. Each virus particle capsid consists of 228 hexons and 12 fibers attached to the 12 penton bases. Hexon and fiber proteins are the bases for different serotypes, as per the early definition of adenovirus type, which are important for viral identification and classification (Friedman and Bennett 1957).

Human adenoviruses are one of the most studied viral pathogens, as they have the potential to cause a diverse range of illness such as conjunctivitis, keratoconjunctivitis, gastroenteritis, hepatitis, myocarditis, renal and urinary tract infections, and pneumonia (Rajkumar et al. 2015; Cersovsky et al. 1999; Chu and Pavan-Langston 1979; Uhnoo et al. 1984). Adenovirus infections may occur in otherwise healthy individuals regardless of age, and are particularly of concern for people with a weakened immune system (Lion 2014). Most of the infections occur in children younger than six years old and are usually self-limiting diseases (Carballal et al. 2001; Singh-Naz and Rodriguez 1996; Yao et al. 2019). HAdV contributes to 5-7% of the respiratory tract infections globally (Walls, Shankar, and Shingadia 2003; Singh-Naz and Rodriguez 1996; Mahr and Gooding 1999).

Adenoviruses were first isolated by Rowe and his colleagues (Rowe et al. 1953). Today, there are 104 HAdV genotypes reported on the Human Adenovirus Working Group website (http://hadvwg.gmu.edu). Many of the recent adenoviruses reported result from recombination in the hypervariable regions. A recombinant may present as an unusual human pathogen. One example is HAdV-B55, formerly known serologically as HAdV-B11a. HAdV-B55 is a recombinant containing the HAdV-B11 hexon gene and the rest of the HAdV-B14 genome (Walsh et al. 2010a). HAdV-B55 causes highly contagious ARD, but is serologically misidentifies as a renal pathogen, HAdV-B11. Cross-species recombinations between HAdV and SAdV have been described, which have given give rise to emergent novel viruses. For example, SAdV-B35 arose from several genome recombinations, with its genome comprising elements of SAdV-B21, -B27, and HAdV-B16 from non-human simian and human hosts, respectively (Dehghan et al., 2013). Thus, HAdVs apparently utilize genome recombination as a survival mechanism. Understanding recombination in HAdVs is crucial for understanding the role molecular evolution plays in generating emergent human and zoonotic pathogens, as well as in viral pathogenesis and epidemiology.

Human adenoviruses are classified into seven species, from HAdV-A to HAdV-G, based on the DNA homology, biological properties, and serological differences (Davison, Benko, and Harrach 2003; Shenk 2001). The adenoviruses in distinct species generally have similar disease-causing patterns. For example, species A members are usually

associated with gastrointestinal tract infections (Schmitz, Wigand, and Heinrich 1983). Species B and C viruses are commonly linked to respiratory tract infections (Fox, Hall, and Cooney 1977; Wadell 1984). Members of species D are associated with ocular tract infections that result in conjunctivitis and keratoconjunctivitis (ELNIFRO et al. 1999) as well gastroenteritis disease (Uhnoo et al. 1984). Species E viruses have been reported as the causes of respiratory and ocular infections (Wadell 1984). Finally, species F and G viruses have been linked to severe gastroenteritis (Uhnoo et al. 1984; Jones et al. 2007).

2. HADV TYPING: SEROTYPING VS GENOTYPING

Traditionally, the serotypes of human adenovirus were determined by the hemagglutination and serum neutralization assays (Friedman and Bennett 1957; Vincent et al. 2001; Rosen 1960). However, exceptions caused by frequent cross-reaction among viruses have blurred the type boundary (Bakay et al. 2005; Svensson et al. 1983). Proper identification of the adenovirus types is important for epidemiological studies and clinical diagnosis. The purpose of this review is to analyze the historical transition from serology to genotype as the basis of recognizing novel adenoviruses and to introduce the Human Adenovirus Working Group as a new collective to evaluate and process adenoviruses.

Serotype can be defined as a group of distinct microorganisms that react to the common set of antigenic markers. Microorganisms that are identified and typed using serology are distinguishable by unique antigenic markers targeted by lab-derived antibodies. The International Committee on Taxonomy of Viruses (ICTV) Adenovirus Working Group

(https://talk.ictvonline.org/ictv-reports/ictv_9th_report/dsdna-viruses-2011/w/dsdna_viruses/93/adenoviridae) guidelines that suggest any novel adenovirus serotype must have either of the following two criteria: first, a novel strain must display a homologous to heterologous titer ratio greater than 16 in both directions and cannot exhibit cro ss-reaction with each other. Second, a substantial biophysical, biochemical, or phylogenetic difference must exist for an 8 or 16 titer ratio to be classified as a novel type. The serotype-specific antibody is generally polyclonal, which exhibits a type-specific property.

Adenoviruses serotypes can be differentiated by a serum neutralization (SN) test and hemagglutination-inhibition (HI) test. An SN test is a serological method used to detect and quantify the presence of antibodies against the specific viral pathogen (Fontaine et al. 1963). It is a highly sensitive and specific test that has been applied to many different viruses such as influenza A (Gauger and Vincent 2014) and bovine herpesvirus (Holz et al. 2010). SN provides detailed information on recent infections, viral incidence, and infection status of the host. It is easy to perform, and the interpretation is less subjective compared to other tests. An HI test is a method used to titrate the level of antibody response to viral particles in a sample by allowing the fiber protein to bind to the sialic acid receptors on red blood cells (Rosen 1960). An HI test is relatively inexpensive for instruments and supplies. Results can be generated within a few hours. Fibers are involved in neutralization and hemagglutination-inhibition, while hexons can only be detected through neutralization. Penton base is not considered in serotyping.

Genotyping is a process of classify unknown organisms based on the genetic marker using sequencing technologies. Genomic data provide accurate and precise genetic information which can also be used to analyze gene variations, type identification, SNP analysis, etc. Molecular genomics has often sequenced and analyzed the hexon regions and the whole adenovirus. And since then the power of genomic sequencing has been harnessed and increases exponentially. Genomic-based typing is a valid approach for naming and classifying the novel adenoviruses. Serotyping identified the first 51 adenoviruses types. Genotyping identifies all 104 types.

3. GENOME SEQUENCE DEFINES TYPE

3.1. Viruses' reality: viral standards

The two traditional ways of identifying adenoviral genomes are the serum neutralization test (SN) and the hemagglutination-inhibition (HI) test (Friedman and Bennett 1957; Vincent et al. 2001; Rosen 1960; Bakay et al. 2005). The benefit of a serum neutralization test is that it is highly specific and sensitive for hexon identification (Pichla-Gollon et al. 2007). However, SN is very time-consuming and only identifies the hexon gene. Experts in the field ceased using serology because it's time-consuming and technically complex which result in a high error rate (Vincent et al. 2001). It is very labor-intensive, and the results vary depending on the antibody source (Klasse 2014; Vincent et al. 2001). The problem is only a limited number of specific antigens can be

generated from a rabbit. Each rabbit has its own unique immune response, which may create different subset of antibodies. The exact physicochemical antibody properties cannot be recreated due to the dependence of animal source(Fierz 2004). For this reason, there is a lack of commercially available reagents for SN test. SN test is not a practical technique for adenovirus typing.

The HI test is a method used to type the adenovirus fiber. HI test is simple to use and relatively inexpensive compared to other assays indirect fluorescent antibody (IFA) and radioimmuno-assay (RIA). Although HI test has been used in the laboratories for many decades, there are many disadvantages in using SN which includes several technical difficulties (Paula, Oliveira, and Fonseca 2004). The disadvantage of HI is its lack of specificity in adenovirus fiber detection. HI does not provide accurate identification of the infected adenovirus serotypes (Paula, Oliveira, and Fonseca 2004). The manual interpretation of the results is subjective and inconsistent, and a reasonable level of quantification is difficult to achieve. The reagents required for the test has a short expiration date (Rosen 1960; Hierholzer 1973). An unwanted abnormal agglutination patterns would appear after the expiration date which will affect the specificity. Even though HI is cheaper and faster, it is not a suitable test for classification. Currently, there is no laboratory in the country performing the hemagglutination inhibition and only a few still use the serum neutralization test. There are no new antisera made for the recently discovered novel adenoviruses. Serology characterizing is outdated and has limits to the understanding of adenoviruses.

3.2. SN is variable: cross-reactivity

Serum neutralization is an inefficient way to classify human adenovirus due to technical complexity and time-consumption (Fontaine et al. 1963; Klasse 2014). There are three regions, hexon L1, hexon L2, and fiber knob, in the adenovirus that could elicit an antigen response (Figure 2.1) (Pichla-Gollon et al. 2007; Olive et al. 2002; Mautner et al. 1975; Rosen 1960). SN and HI only detect the epsilon (Ɛ) region on hexon and gamma (γ) region on fiber. The antibodies produced do not detect all regions at once(Friedman and Bennett 1957; Fontaine et al. 1963). Cross-reactivity can occur among different adenoviruses with the same hexon, especially with types that have similar epitope regions (Vincent et al. 2001). An example of this is adenovirus type 15, 29, 56, 69, 87, 88, and 94. All these adenoviruses shares the same hexon (Singh et al. 2015). A neutralization test will result in a series of cross-reactivity among these adenoviruses. All four adenoviruses were recognized as the same type of adenovirus even though they are different and elicit different immune responses. The ICTV Human Adenovirus Working Group defines novel adenovirus as two adenoviruses that have >16X bidirectional differences or a substantial biophysical, biochemical, and phylogeny difference with an 8 or 16 tier ratio. It is difficult to classify a novel type when antibodies cross-react with each other.

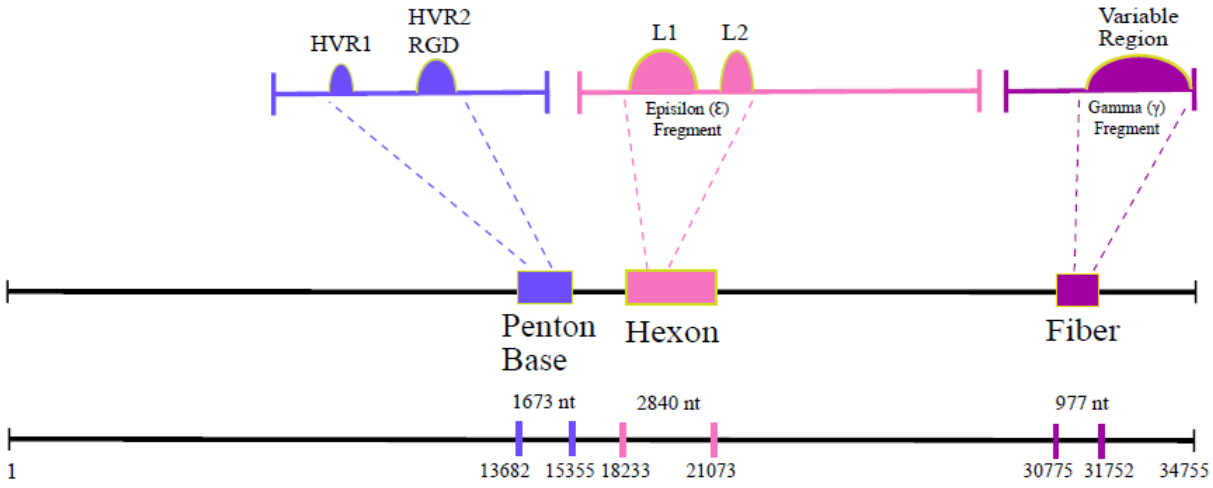


Figure 2.1. Genome organization of adenovirus capsid proteins: hexon, penton base, and fiber. Epitopes for hexon and fiber were labeled based on the location in the HAdV-B55 genome.

4. A NEW VIEW ON EVOLUTION OF HADVS VIA GENOMICS

Adenovirologists have adopted genome sequencing as an alternative way to classify adenovirus (Reddy et al. 1998). Genome sequencing is more cost-effective, time-efficient, and accurate than other methods, such as serum neutralization test and hemagglutination inhibition test. It provides more complete and reliable data for identifying microorganisms (Houldcroft et al. 2018). In fact, many adenoviruses have been identified in recent years by genome sequencing (Walsh et al. 2010a; Zhou et al. 2012).

406 The genome of HAdV-C2 and HAdV-C5 are the first ones to be fully sequenced with
407 approximately 36 kb and encodes about 36 proteins (Flint 1999). Genome sequencing can
408 analyze the strain's genetic makeup. The traditional method using seroepidemiological
409 assays can only detect a few regions and is limited to the hexon and fibers (Friedman and
410 Bennett 1957; Rosen 1960; Vincent et al. 2001). Recent studies have shown that there are
411 altered polypeptides produced by alternative splicing (Zhao, Chen, and Pettersson 2014)
412 and this can be analyzed easily by genomic sequencing. A genomic-based analysis is
413 necessary to analyze these sites. Genomic sequencing analysis should be used in all
414 adenovirus for standard classification.
415
416 **4.1. Recombinants are novel viruses/pathogens: Biology/pathogenicity**
417 Recombination is common in adenoviruses (Robinson et al. 2013). Recombination occurs
418 between the hexon, and fiber genes in adenoviruses (Grodzicker et al. 1974; Mautner et
419 al. 1975; Williams et al. 1975). The specific mechanism of recombination remains
420 unknown due to the fact that it can generate an epidemiologic significant genotype by
421 inserting a pathogenic gene into a nonpathogenic genotype (Huang and Xu 2013). These
422 new recombinants should be considered as a novel virus instead of a variant with unique
423 pathogenicity.
424
425 ***4.1.1. HAdV-B55/11a***
426 HAdV-B55 is a recombination of HAdV-B11 and HAdV-B14 and causes highly
427 contagious acute respiratory pathogen (ARD) (Walsh et al. 2010a). The hexon genome

(2.6 %) in HAdV-B55 is identical to the renal pathogen HAdV-B11 with the rest of the genome (97.4 %) highly similar to HAdV-B14. This recombination creates a "Trojan horse" effect; HAdV-B55 would appear in hosts as a renal pathogen with the specific antibody epitope but causes respiratory diseases and renal pathogens are rarer in populations than respiratory pathogens. Serology recognizes it as HAdV-B11 (Zhu et al. 2009; Yang et al. 2009a) and physicians would prescribe medicines that treat renal pathogens instead of respiratory symptoms. Sanger sequencing revealed the specific strain was a recombinant of HAdV-B14 and HAdV-B11 (Walsh et al. 2010b; Yang et al. 2009b). This is an example of how recombination leads to a novel human adenovirus pathogen, and of how genomics-based methods allow a higher resolution understanding of and a more definitive determination of genotype.

4.1.2. HAdV-D53

HAdV-D53 is a recombination of HAdV-D37 penton base, HAdV-D22 hexon, and HAdV-D8 fiber. HAdV-D53 was isolated from an epidemic keratoconjunctivitis (EKC) outbreak. HAdV-D53 was classified as a HAdV-D22 variant by serotyping when it first appeared. The problem is HAdV-D22 is a nonpathogenic strain and had no record of causing EKC. The common source of EKC at the time was HAdV-D8, -D19, and -D37. The genomic sequencing and bioinformatic analysis reveal that HAdV-D53 has undergone at least five recombination events (Engelmann et al. 2006; Walsh et al. 2009). It is insufficient to name HAdV-D53 as a variant of HAdV-D22 due to the similar capsid.

Correctly identifying and classifying adenoviruses are essential for containing the pathogenicity.

4.1.3. HAdV-D64/19

HAdV-D64 is a highly contagious EKC agent that was misidentified as HAdV-D19 by serology. HAdV-D64 is a recombination of type 22 penton base, 19 hexon, and 37 fiber. HAdV-D19 is a common non-pathogenic virus, but HAdV-D64 causes highly inflammatory infections in the ocular area. HAdV-D64 and other clinical EKC isolates have a ~100% sequence identity in NCBI. HAdV-D19 had a lack of association with keratitis and is not infectious for corneal epithelial cells. HAdV-D64 was later reclassified due to its unique pathogenicity, but substantial confusion was raised in this period of time regarding both the HAdV-D19 origin and pathogen tropism (Zhou et al. 2012).

4.1.4. HAdV-D56/D15/D/29

HAdV-D56 has a similar serotype as HAdV-D15 and HAdV-D29 by molecular typing. HAdV-D15 can cause acute follicular conjunctivitis and respiratory tract infections. HAdV-D56 can cause respiratory-related fatalities and highly contagious ocular pathogens. Genomic-based algorithms have identified it as a new type with specific pathogenicity and recombinants (Robinson et al. 2011). HAdV-D56 is a unique identifier rather than a variant of HAdV-D15 and HAdV-D29. Serology can lead to misleading results. In this case, HAdV-D15 and HAdV-D29 cross-reacted below 8x and 16x which

created more confusion. HAdV-D56 is a novel strain with multiple recombination events and significant pathogenic and biological effects.

5. HUMAN ADENOVIRUS WORKING GROUP

The Human Adenovirus (HAdV) Working Group (http://hadvwg.gmu.edu) is a collaboration within the adenovirus research community with the mission to evaluate candidate novel human adenovirus genotypes using a genome sequence-based algorithm. This is a platform to catalogue and present evaluated novel human adenovirus types using genome sequence-based standardized parameters under a systematic guideline. It was created to provide advice to NCBI for assigning new genotype numbers to proposed human adenoviruses. A running catalog of novel genotypes as well as working group contacts are posted on the HAdV Working Group website (http://hadvwg.gmu.edu). The current parameters and criteria used to identify novel sequences can be located on the "Submit 'candidate' HAdV" page. The submission is forwarded to the chairman, who distributes the information to a committee consists of nine members that meet virtually to evaluate. Once the approval of the candidate strain has been determined, a letter will be sent to the requester, which is submitted with the genome data for inclusion in the GenBank database. The new genotype catalogued and updated on the home page (Table 2.1).

491 Table 2.1. Molecular typing data for human adenoviruses types 53-104. Penton base,
492 hexon, and fiber genotyping identities for each adenovirus genotype are listed, along with
493 metadata.

Adenovirus Genotype	Name	Accession #	Year (Publication)	Penton base	Hexon	Fiber
HAdV-D104	P1H1F2/2017/CHN	MH558113	2019	1	1	2
HAdV-D103	P103H33F30/1986/DEU	KF268322	2019	103	33	30
HAdV-D102	P102H38F30/1984/DEU	KF268312	2019	102	38	30
HAdV-D101	P101H37F45/1988/DEU	KF268324	2019	101	37	45
HAdV-D100	P100H17F30/1986/DEU	KF268330	2019	100	17	30
HAdV-D99	P9H46F39/1987/DEU	KF268211	2019	9	46	39
HAdV-D98	P98H46F9/1986/DEU	KF268332	2019	98	46	9
HAdV-D97	P67H28F60/1987/DEU	KF268320	2019	67	28	60
HAdV-D96	P23H32F62/1988/DEU	KF268327	2019	23	32	62
HAdV-D95	P95H9F15/1984/DEU	KF268206	2019	95	9	15
HAdV-D94	P33H15F9/1982/DEU	KF268201	2019	33	15*	9
HAdV-D93	P28H37F38/1988/DEU	KF268334	2019	28	37	38
HAdV-D92	P92H32F27/1986/DEU	KF268325	2019	92	32	27
HAdV-D91	P37H37F17/DEU	KF268208	2019	37	37	17
HAdV-D90	P33H27F67/2017/BGD	98% Partial	TBA	33	27	67
HAdV-C89	P89H2F2/2015/DEU	MH121097	2019	89	2	2
HAdV-D88	P88H15F9/1963/USA	MF476842	TBA	88	15*	9
HAdV-D87	P9H15F25/1967/USA	MF476841	TBA	9	15*	25
HAdV-D86	P9H25F25/1978/SWE	KX868297	2018	9	25*	25
HAdV-D85	P37H19F8/2015/JPN	LC314153	2018	37	19	8
HAdV-D84	P43H17F84/2011/PAN	MF416150	2017	43	17	84
HAdV-D83	P83H9F15/2010/PAR	KX827426.1	2016	83	9	15
HAdV-D82	P56H56F37/2011/JPN	LC066535.1	2015	56	56*	37
HAdV-D81	P65H48F60/2012/JPN	AB765926.1	2014	65	48	60
HAdV-D80	P19,23H28F22/2014/DEU	TBA	TBA	22	28	22
HAdV-B79	P11H34F11/2015/JPN	LC177352	2016	11	34	11
HAdV-B78	P11H11F7/2000/ARG	KT970440	2016	11	11	7
HAdV-B77	P35H34F7/1985/DEU	KF268328	2013	35	34	7
HAdV-B76	P21H21F16/DEU	KF633445	2013	21	21	16
HAdV-D75	P75H26F29/2015/DEU	KY618678	2017	75	26	29
HAdV-D74	P70H74F51/2015/DEU	KY618677	2017	70	74	51
HAdV-D73	P67H45F27/2015/DEU	KY618676	2017	67	45	27
HAdV-D72	P72H30F72/1985/DEU	KF268335	2013	72	30	72
HAdV-D71	P9H20F71/1987/DEU	KF268207	2013	9	20	71
HAdV-D70	P70H70F29/2014/DEU	KP641339	2015	70	70	29
HAdV-D69	P53H15F69/1955/SAU	JN226748	2013	53	15*	69
HAdV-B68	P16H3F16/2004/ARG	JN860678	2011	16	3	16
HAdV-D67	P67H9F67/2005/BGD	AP012302	2013	67	9	67
HAdV-B66	P66H7F3/1987/ARG	JN860676	2012	66	7	3
HAdV-D65	P58H10F9/2004/BGD	AP012285	2012	58	10	9
HAdV-D64	P22H19F37/1993/USA	EF121005	2006	22	19	37
HAdV-D63	P30H30F29/1959/USA	JN935766	2012	30	30	29
HAdV-D62	P62H62F62/1993/GBR	JN162671	2014	62	62	62
HAdV-A61	P31H31F31/2004/JPN	JF964962	2011	31	61	31
HAdV-D60	P60H20F60/2009/CAN	HQ007053	2013	60	20	60
HAdV-D59	P64H25F56/2007/USA	JF799911	2012	64	25*	56
HAdV-D58	P58H58F29/1996/ARG	HQ883276	2011	58	58	29
HAdV-C57	P1H57F6/2001/RUS	HQ003817	2011	1	57	6
HAdV-D56	P56H15F9/2008/FRA	HM770721	2010	56	15*	9
HAdV-B55	P14H11F14/2006/CHN	FJ643676	2009	14	11	14
HAdV-D54	P54H54F8/2000/JPN	AB333801	2009	54	54	8
HAdV-D53	P37H22F8/2005/DEU	FJ169625	2009	37	22	8

* Equivalent to 15, 25, 29, 56, 69, 87, 88, and 94

494

495

6. CONCLUSION

Human adenovirus recombination and cross-reactions amongst types make it difficult to rely solely on serology-based data. Genome sequencing is a better method for identifying, characterizing, and typing human adenoviruses, and also in understanding adenovirus evolution. The Human Adenovirus Working Group is a resource for the adenovirus research community, providing typing evaluation for novel adenovirus candidate types based on genome sequence data.

Conflicts of interest

There are no conflicts of interest.

References

Bakay, Marta, Katalin Szalay, Ilona Beladi, Elisabeth Balint, Anna Lengyel, Eva Adam, and Tibor Olah. 2005. “Cross-Reactivity between Human Adenoviruses in Delayed-Type Hypersensitivity.” *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica* 113 (3): 197–202. https://doi.org/10.1111/j.1600-0463.2005.apm1130307.x.

Carballal, G., C. M. Videla, M. A. Espinosa, V. Savy, O. Uez, M. D. Sequeira, V. Knez, P. V. Requeijo, C. R. Posse, and I. Miceli. 2001. “Multicentered Study of Viral Acute Lower Respiratory Infections in Children from Four Cities of Argentina,

1993-1994." *Journal of Medical Virology* 64 (2): 167–74. https://doi.org/10.1002/jmv.1032.

Cersovsky, Steven, Binn LN, Mitchell FD, Kolavic SA, Polyak CS, Jose Sanchez Jr, and Bester WT. 1999. "A Prospective Study of Acute Respiratory Diseases (ARD) in a Cohort of US Army Basic Trainees, October-November 1998, Fort Jackson, South Carolina." *Medical Surveillance Monthly Report* 5: 10, 12.

Chu, Wing, and Eborah Pavan-Langston. 1979. "Ocular Surfce Manifestations of the Major Viruses." *International Ophthalmology Clinics* 19 (2). https://journals.lww.com/internat-ophthalmology/Fulltext/1979/01920/Ocular_Surfce_Manifestations_of_the_Major_Viruses.10.aspx.

Dehghan, Shoaleh, Jason Seto, Morris S. Jones, David W. Dyer, James Chodosh, and Donald Seto. 2013. "Simian Adenovirus Type 35 Has a Recombinant Genome Comprising Human and Simian Adenovirus Sequences, Which Predicts Its Potential Emergence as a Human Respiratory Pathogen." *Virology* 447 (1–2): 265–73. https://doi.org/10.1016/j.virol.2013.09.009.

Elnifro, E., R. Cooper, P. Klapper, A. Bailey, and A. Tullo. 1999. "Diagnosis of Viral and Chlamydial Keratoconjunctivitis: Which Laboratory Test?" *The British Journal of Ophthalmology* 83 (5): 622–27.

Engelmann, Ilka, Ijad Madisch, Heidi Pommer, and Albert Heim. 2006. "An Outbreak of Epidemic Keratoconjunctivitis Caused by a New Intermediate Adenovirus 22/H8

Identified by Molecular Typing." *Clinical Infectious Diseases* 43 (7): e64–66. https://doi.org/10.1086/507533.

Fierz, Walter. 2004. "Basic Problems of Serological Laboratory Diagnosis." *Methods in Molecular Medicine* 94: 393–427.

Flint, J. 1999. "Organization of the Adenoviral Genome." In *In Adenoviruses: Basic Biology to Gene Therapy*, 17–30. Austin, TX, USA: Landes Company.

Fontaine, M. P., D. Chabas, M. Fontaine, and A. J. Brion. 1963. "A Simplified Technique for the Serum-Neutralization Test in Infectious Bronchitis." *Avian Diseases* 7 (2): 203–6. https://doi.org/10.2307/1588052.

Fox, J. P., C. E. Hall, and M. K. Cooney. 1977. "The Seattle Virus Watch. VII. Observations of Adenovirus Infections." *American Journal of Epidemiology* 105 (4): 362–86.

Friedman, Murray, and C. R. Bennett. 1957. "A Hemagglutination Test for Detection of Adenovirus Antibodies." *Proceedings of the Society for Experimental Biology and Medicine* 94 (4): 712–17. https://doi.org/10.3181/00379727-94-23062.

Gauger, Phillip C., and Amy L. Vincent. 2014. "Serum Virus Neutralization Assay for Detection and Quantitation of Serum-Neutralizing Antibodies to Influenza A Virus in Swine." *Methods in Molecular Biology (Clifton, N.J.)* 1161: 313–24. https://doi.org/10.1007/978-1-4939-0758-8_26.

Grodzicker, Terri, Carl Anderson, Phillip A. Sharp, and Joe Sambrook. 1974. "Conditional Lethal Mutants of Adenovirus 2-Simian Virus 40 Hybrids I. Host Range Mutants of Ad2+ND1." *Journal of Virology* 13 (6): 1237–44.

560 Harrach, B, M Benkö, G.W. Both, M Brown, A.J. Davison, M Echavarria, M Hess, et al.
561 2011. “Virus Taxonomy: Classification and Nomenclature of Viruses.” In , 125–
562 41. San Diego: Elsevier Press.
563 Heemskerk, Bianca, Louise A. Veltrop-Duits, Tamara van Vreeswijk, Monique M. ten
564 Dam, Sebastiaan Heidt, Rene E. M. Toes, Maarten J. D. van Tol, and Marco W.
565 Schilham. 2003. “Extensive Cross-Reactivity of CD4+ Adenovirus-Specific T
566 Cells: Implications for Immunotherapy and Gene Therapy.” *Journal of Virology*
567 77 (11): 6562–66.
568 Hierholzer, J. C. 1973. “Further Subgrouping of the Human Adenoviruses by Differential
569 Hemagglutination.” *The Journal of Infectious Diseases* 128 (4): 541–50.
570 https://doi.org/10.1093/infdis/128.4.541.
571 Hilleman, M. R., and J. H. Werner. 1954. “Recovery of New Agent from Patients with
572 Acute Respiratory Illness.” *Proceedings of the Society for Experimental Biology*
573 *and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*
574 85 (1): 183–88. https://doi.org/10.3181/00379727-85-20825.
575 Holz, Carine Lidiane, Samuel Paulo Cibulski, Thais Fumaco Teixeira, Helena Beatriz C.
576 R. Batista, Diógenes Dezen, Fabrício Souza Campos, Ana Paula M. Varela, and
577 Paulo Michel Roehe. 2010. “Serum Neutralization with Different Types and
578 Subtypes of Bovine Herpesvirus 1 and 5.” *Pesquisa Veterinária Brasileira* 30 (7):
579 515–22. https://doi.org/10.1590/S0100-736X2010000700001.
580 Houldcroft, Charlotte J, Sunando Roy, Sofia Morfopoulou, Ben K Margetts, Daniel P
581 Depledge, Juliana Cudini, Divya Shah, et al. 2018. “Use of Whole-Genome

582 Sequencing of Adenovirus in Immunocompromised Pediatric Patients to Identify
583 Nosocomial Transmission and Mixed-Genotype Infection." *The Journal of*
584 *Infectious Diseases* 218 (8): 1261–71. https://doi.org/10.1093/infdis/jiy323.
585 Huang, Guo-Hong, and Wen-Bo Xu. 2013. "[Recent advance in new types of human
586 adenovirus]." *Bing Du Xue Bao = Chinese Journal of Virology* 29 (3): 342–48.
587 Jones, Morris Saffold, Balázs Harrach, Robert D. Ganac, Mary M. A. Gozum, Wilfred P.
588 dela Cruz, Brian Riedel, Chao Pan, Eric L. Delwart, and David P. Schnurr. 2007.
589 "New Adenovirus Species Found in a Patient Presenting with Gastroenteritis."
590 *Journal of Virology* 81 (11): 5978–84. https://doi.org/10.1128/JVI.02650-06.
591 Klasse, P. J. 2014. "Neutralization of Virus Infectivity by Antibodies: Old Problems in
592 New Perspectives." Edited by Ma Luo. *Advances in Biology* 2014 (September):
593 157895. https://doi.org/10.1155/2014/157895.
594 Lion, Thomas. 2014. "Adenovirus Infections in Immunocompetent and
595 Immunocompromised Patients." *Clinical Microbiology Reviews* 27 (3): 441–62.
596 https://doi.org/10.1128/CMR.00116-13.
597 Mahr, J. A., and L. R. Gooding. 1999. "Immune Evasion by Adenoviruses."
598 *Immunological Reviews* 168 (April): 121–30.
599 Mautner, V., J. Williams, J. Sambrook, P. A. Sharp, and T. Grodzicker. 1975. "The
600 Location of the Genes Coding for Hexon and Fiber Proteins in Adenovirus
601 DNA." *Cell* 5 (1): 93–99.
602 Olive, Melanie, Laurence Eisenlohr, Neal Flomenberg, Susan Hsu, and Phyllis
603 Flomenberg. 2002. "The Adenovirus Capsid Protein Hexon Contains a Highly

Conserved Human CD4+ T-Cell Epitope." *Human Gene Therapy* 13 (10): 1167–78. https://doi.org/10.1089/104303402320138952.

Paula, De, Sérgio Oliveira, and Benedito Antônio Lopes da Fonseca. 2004. "Dengue: A Review of the Laboratory Tests a Clinician Must Know to Achieve a Correct Diagnosis." *Brazilian Journal of Infectious Diseases* 8 (6): 390–98. https://doi.org/10.1590/S1413-86702004000600002.

Pichla-Gollon, Susan L., Mark Drinker, Xiangyang Zhou, Feng Xue, John J. Rux, Guang-Ping Gao, James M. Wilson, Hildegund C. J. Ertl, Roger M. Burnett, and Jeffrey M. Bergelson. 2007. "Structure-Based Identification of a Major Neutralizing Site in an Adenovirus Hexon." *Journal of Virology* 81 (4): 1680. https://doi.org/10.1128/JVI.02023-06.

Rajkumar, Veena, Cheryl S. M. Chiang, Jia Meng Low, Lin Cui, Raymond T. P. Lin, Nancy W. S. Tee, Matthias Maiwald, Chia Yin Chong, Koh Cheng Thoon, and Natalie W. H. Tan. 2015. "Risk Factors for Severe Adenovirus Infection in Children during an Outbreak in Singapore." *Annals of the Academy of Medicine, Singapore* 44 (2): 50–59.

Reddy, P S, N Idamakanti, A N Zakhartchouk, M K Baxi, J B Lee, C Pyne, L A Babiuk, and S K Tikoo. 1998. "Nucleotide Sequence, Genome Organization, and Transcription Map of Bovine Adenovirus Type 3." *Journal of Virology* 72 (2): 1394–1402. https://doi.org/10.1128/JVI.72.2.1394-1402.1998.

Robinson, Christopher M., Gurdeep Singh, Cécile Henquell, Michael P. Walsh, Hélène Peigue-Lafeuille, Donald Seto, Morris S. Jones, David W. Dyer, and James

Chodosh. 2011. “Computational Analysis and Identification of an Emergent Human Adenovirus Pathogen Implicated in a Respiratory Fatality.” *Virology* 409 (2): 141–47. https://doi.org/10.1016/j.virol.2010.10.020.

Robinson, Christopher M., Gurdeep Singh, Jeong Yoon Lee, Shoaleh Dehghan, Jaya Rajaiya, Elizabeth B. Liu, Mohammad A. Yousuf, et al. 2013. “Molecular Evolution of Human Adenoviruses.” *Scientific Reports* 3 (May): 1812. https://doi.org/10.1038/srep01812.

Rosen, L. 1960. “A Hemagglutination-Inhibition Technique for Typing Adenoviruses.” *American Journal of Hygiene* 71 (January): 120–28. https://doi.org/10.1093/oxfordjournals.aje.a120085.

Rowe, W. P., R. J. Huebner, L. K. Gilmore, R. H. Parrott, and T. G. Ward. 1953. “Isolation of a Cytopathogenic Agent from Human Adenoids Undergoing Spontaneous Degeneration in Tissue Culture.” *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* 84 (3): 570–73.

Schmitz, H., R. Wigand, and W. Heinrich. 1983. “Worldwide Epidemiology of Human Adenovirus Infections.” *American Journal of Epidemiology* 117 (4): 455–66.

Singh, Gurdeep, Xiaohong Zhou, Jeong Yoon Lee, Mohammad A Yousuf, Mirja Ramke, Ashrafali M Ismail, Ji Sun Lee, et al. 2015. “Recombination of the Epsilon Determinant and Corneal Tropism: Human Adenovirus Species D Types 15, 29, 56, and 69.” *Virology* 485 (November): 452–59. https://doi.org/10.1016/j.virol.2015.08.018.

Singh-Naz, N., and W. Rodriguez. 1996. “Adenoviral Infections in Children.” *Advances in Pediatric Infectious Diseases* 11: 365–88.

Smith, C. A., L. S. Woodruff, C. Rooney, and G. R. Kitchingman. 1998. “Extensive Cross-Reactivity of Adenovirus-Specific Cytotoxic T Cells.” *Human Gene Therapy* 9 (10): 1419–27. https://doi.org/10.1089/hum.1998.9.10-1419.

Svensson, Lennart, Göran Wadell, Ingrid Uhnoo, Mats Johansson, and Carl-Henrik Von Bonsdorff. 1983. “Cross-Reactivity between Enteric Adenoviruses and Adenovirus Type 4: Analysis of Epitopes by Solid-Phase Immune Electron Microscopy.” *Journal of General Virology,*. Microbiology Society,.

Swenson, PD, G Wadell, A Allard, and JC Hierholzer. 2003. *Manual of Clinical Microbiology*. Vol. 2.8. Washington, DC: ASM Press.

Uhnoo, I., G. Wadell, L. Svensson, and M. E. Johansson. 1984. “Importance of Enteric Adenoviruses 40 and 41 in Acute Gastroenteritis in Infants and Young Children.” *Journal of Clinical Microbiology* 20 (3): 365–72.

Vincent, T, B G Harvey, S M Hogan, C J Bailey, R G Crystal, and P L Leopold. 2001. “Rapid Assessment of Adenovirus Serum Neutralizing Antibody Titer Based on Quantitative, Morphometric Evaluation of Capsid Binding and Intracellular Trafficking: Population Analysis of Adenovirus Capsid Association with Cells Is Predictive of Adenovirus Infectivity.” *Journal of Virology* 75 (3): 1516–21. https://doi.org/10.1128/JVI.75.3.1516-1521.2001.

Wadell, G. 1984. “Molecular Epidemiology of Human Adenoviruses.” *Current Topics in Microbiology and Immunology* 110: 191–220.

Walls, Tony, A. G. Shankar, and Delane Shingadia. 2003. “Adenovirus: An Increasingly Important Pathogen in Paediatric Bone Marrow Transplant Patients.” *The Lancet. Infectious Diseases* 3 (2): 79–86.

Walsh, Michael P., Ashish Chintakuntlawar, Christopher M. Robinson, Ijad Madisch, Balázs Harrach, Nolan R. Hudson, David Schnurr, et al. 2009. “Evidence of Molecular Evolution Driven by Recombination Events Influencing Tropism in a Novel Human Adenovirus That Causes Epidemic Keratoconjunctivitis.” *PLOS ONE* 4 (6): e5635. https://doi.org/10.1371/journal.pone.0005635.

Walsh, Michael P., Jason Seto, Morris S. Jones, James Chodosh, Wenbo Xu, and Donald Seto. 2010a. “Computational Analysis Identifies Human Adenovirus Type 55 as a Re-Emergent Acute Respiratory Disease Pathogen.” *Journal of Clinical Microbiology* 48 (3): 991–93. https://doi.org/10.1128/JCM.01694-09.

———. 2010b. “Computational Analysis Identifies Human Adenovirus Type 55 as a Re-Emergent Acute Respiratory Disease Pathogen.” *Journal of Clinical Microbiology* 48 (3): 991–93. https://doi.org/10.1128/JCM.01694-09.

Williams, J., T. Grodzicker, P. Sharp, and J. Sambrook. 1975. “Adenovirus Recombination: Physical Mapping of Crossover Events.” *Cell* 4 (2): 113–19.

Yang, Zhaohui, Zhen Zhu, Liuying Tang, Li Wang, Xiaojuan Tan, Pengbo Yu, Yong Zhang, et al. 2009a. “Genomic Analyses of Recombinant Adenovirus Type 11a in China.” *Journal of Clinical Microbiology* 47 (10): 3082. https://doi.org/10.1128/JCM.00282-09.

———. 2009b. “Genomic Analyses of Recombinant Adenovirus Type 11a in China.” *Journal of Clinical Microbiology* 47 (10): 3082–90. https://doi.org/10.1128/JCM.00282-09.

Yao, Li-hong, Chao Wang, Tian-li Wei, Hao Wang, Fen-lian Ma, and Li-shu Zheng. 2019. “Human Adenovirus among Hospitalized Children with Respiratory Tract Infections in Beijing, China, 2017–2018.” *Virology Journal* 16 (1): 78. https://doi.org/10.1186/s12985-019-1185-x.

Zhao, Hongxing, Maoshan Chen, and Ulf Pettersson. 2014. “A New Look at Adenovirus Splicing.” *Virology* 456–457 (May): 329–41. https://doi.org/10.1016/j.virol.2014.04.006.

Zhou, Xiaohong, Christopher M. Robinson, Jaya Rajaiya, Shoaleh Dehghan, Donald Seto, Morris S. Jones, David W. Dyer, and James Chodosh. 2012. “Analysis of Human Adenovirus Type 19 Associated with Epidemic Keratoconjunctivitis and Its Reclassification as Adenovirus Type 64.” *Investigative Ophthalmology & Visual Science* 53 (6): 2804–11. https://doi.org/10.1167/iovs.12-9656.

Zhu, Zhen, Yong Zhang, Songtao Xu, Pengbo Yu, Xiaoping Tian, Li Wang, Zhonghua Liu, et al. 2009. “Outbreak of Acute Respiratory Disease in China Caused by B2 Species of Adenovirus Type 11.” *Journal of Clinical Microbiology* 47 (3): 697. https://doi.org/10.1128/JCM.01769-08.

CHAPTER 3: HAdV-E4 is of Zoonotic Origin with Recent Isolates Displaying a Putative Host Adaptation

Human adenovirus type 4 (HAdV-E4/Ad4) is linked to highly contagious acute respiratory disease infections that were historically limited to population-dense and overstressed military populations (Cersovsky et al., 1999; Hilleman & Werner, 1954). It was the first adenoviruses identified and characterized, sampled in the military personnel in 1953, and was isolated contemporaneously with other human adenoviruses isolated in civilian settings (Hilleman & Werner, 1954). Ad4 contributed to the majority of the respiratory pathogen outbreaks in the military populations (Cersovsky et al., 1999; DuVernoy et al., 2000), and created a significant problem such that vaccines were developed to prevent these outbreaks. This major and historical human viral pathogen is an example of how high-resolution genome data can provide insights into their origins. Ad4 is the first adenovirus identified to have a zoonotic origin (Dehghan et al., 2013). Recent applications of genome sequencing and bioinformatics analysis shows the prototype isolate shares 89.2 % of the genome identity with SAdV-B26, an adenovirus isolated from chimpanzees (Dehghan et al., 2013). Additionally, since Ad4 has a higher genome similarity with other chimpanzee adenoviruses (SAdV-25, SAdV-39, SAdV-30, SAdV-37.1, SAdV-37.2, and SAdV-25.2) and a lower genome similarity with other human adenoviruses (HAdV-B16 and HAdV-B21), it was hypothesized to have a chimpanzee origin (Dehghan et al., 2013), that is, Ad4 was a emergent pathogen resulting

731 from a zoonotic transfer to humans. In this chapter, Ad4 data collected from outbreaks in
732 Singapore, Hong Kong, and Beijing are analyzed and discussed in the context of the
733 prevalence of Ad4 emerging in the general civilian population, presumably due to host
734 adaptation.
735
736 *Inverted Terminal Repeats*
737 One defining difference between HAdVs and adenoviruses hosted by other great apes, in
738 particular chimpanzees, is the sequence of the inverted terminal repeats (ITRs). ITRs are
739 complimentary sequences (100 – 200 bp) at the ends of the linear double-stranded
740 genomes that play an essential and important role in viral DNA replication (Arrand &
741 Roberts, 1979; Smart & Stillman, 1982; Tokunaga et al., 1982; Tolun et al., 1979). ITRs
742 contain the origin of replication which initiates DNA synthesis in adenoviruses (Arrand
743 & Roberts, 1979). The precursor terminal protein (pTP), serves as a primer, forms a pTP-
744 pol complex with the DNA polymerase, and binds to the core origin motif in the ITRs for
745 pre-initiation (Stuiver & van der Vliet, 1990). For optimal HAdV replication, host
746 nuclear factor I (NF-I) enhances the initiation of the adenovirus DNA replication by
747 binding at the NF-1 motif and increasing the stability of the pre-initiation complex
748 binding, thereby stimulating the replication process up to 50-fold (Mul & Van der Vliet,
749 1992). Host nuclear factor III (NF-III) increases the initiation efficiency by binding at the
750 NF-III motif and stimulating the reaction four- to six-fold (van der Vliet et al., 1986). The
751 synergy of NF-I and NF-III can stimulate the initiation up to 200-fold (Hay & Russell,
752 1989; Stillman, 1989; Tsurimoto et al., 1989). In contrast, chimpanzee ITRs have core

753 origin and NF-III that are similar to HAdVs, but a distinctive NF-I region (Dehghan et
754 al., 2013).
755
756 Recently, an increased number of HAdV-E4 outbreaks have appeared in the general
757 public rather than the limited military populations that were historically sites of Ad4
758 outbreaks (Coleman et al., 2019; Li et al., 2018; Zhang et al., 2019). This may be
759 associated with increased infectivity of the virus or host adaptation of the zoonotic-
760 origins HAdV-E4 as recent genomic analyses showed an apparent modification in the
761 ITRs motif region that may have contributed to increased viral infectivity in human hosts
762 (Dehghan et al., 2013). Our study shows that the ITRs region of the newly isolated strains
763 apparently have evolved, by recombination, and producing two separate clusters in the
764 phylogenetic tree. The ITRs regions for HAdV-E4 strains can be separated into
765 prototype-like and host-adapted clades based on the appearance of a NF-I motif which
766 are found in HAdVs but not in the chimpanzee adenoviruses. In contrast, the core origin
767 and NF-III sequences are conserved across both hosts (Figure 3.1). Previous study has
768 shown that the NF-I motif for the prototype-like clade resembled the simian adenovirus
769 species E, a chimpanzee virus (Dehghan et al., 2013). The ITRs for the recent isolates
770 from in the Hong Kong (Zhang et al., 2019) and the Singapore (Coleman et al., 2019)
771 have more resemblance to the human adenoviruses species B (Figure 3.1)
772

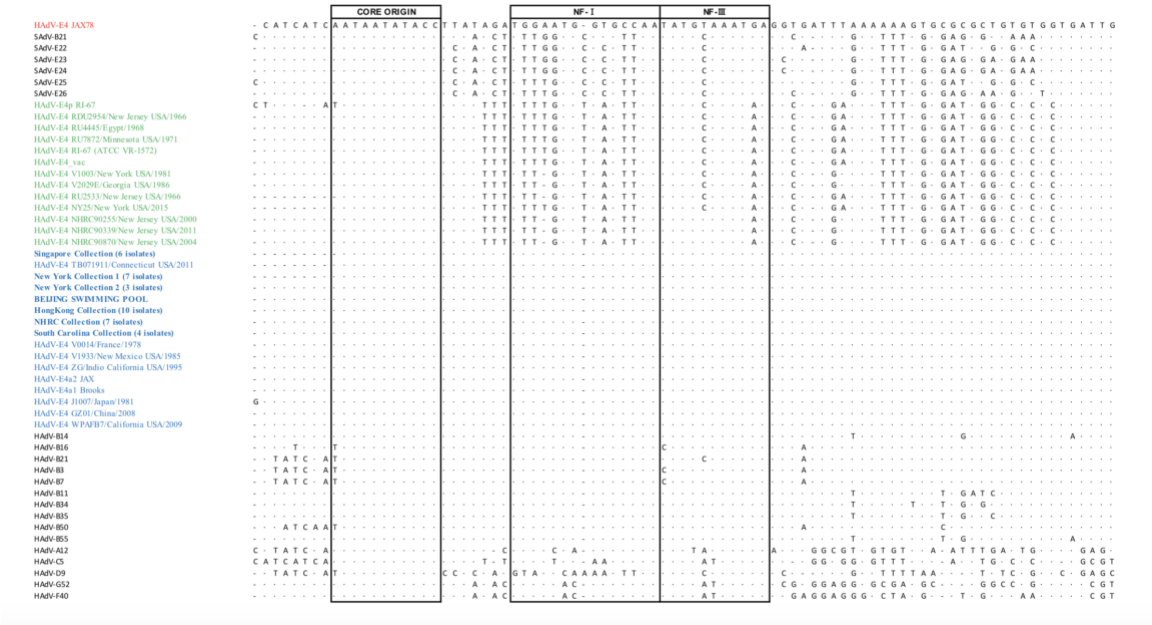
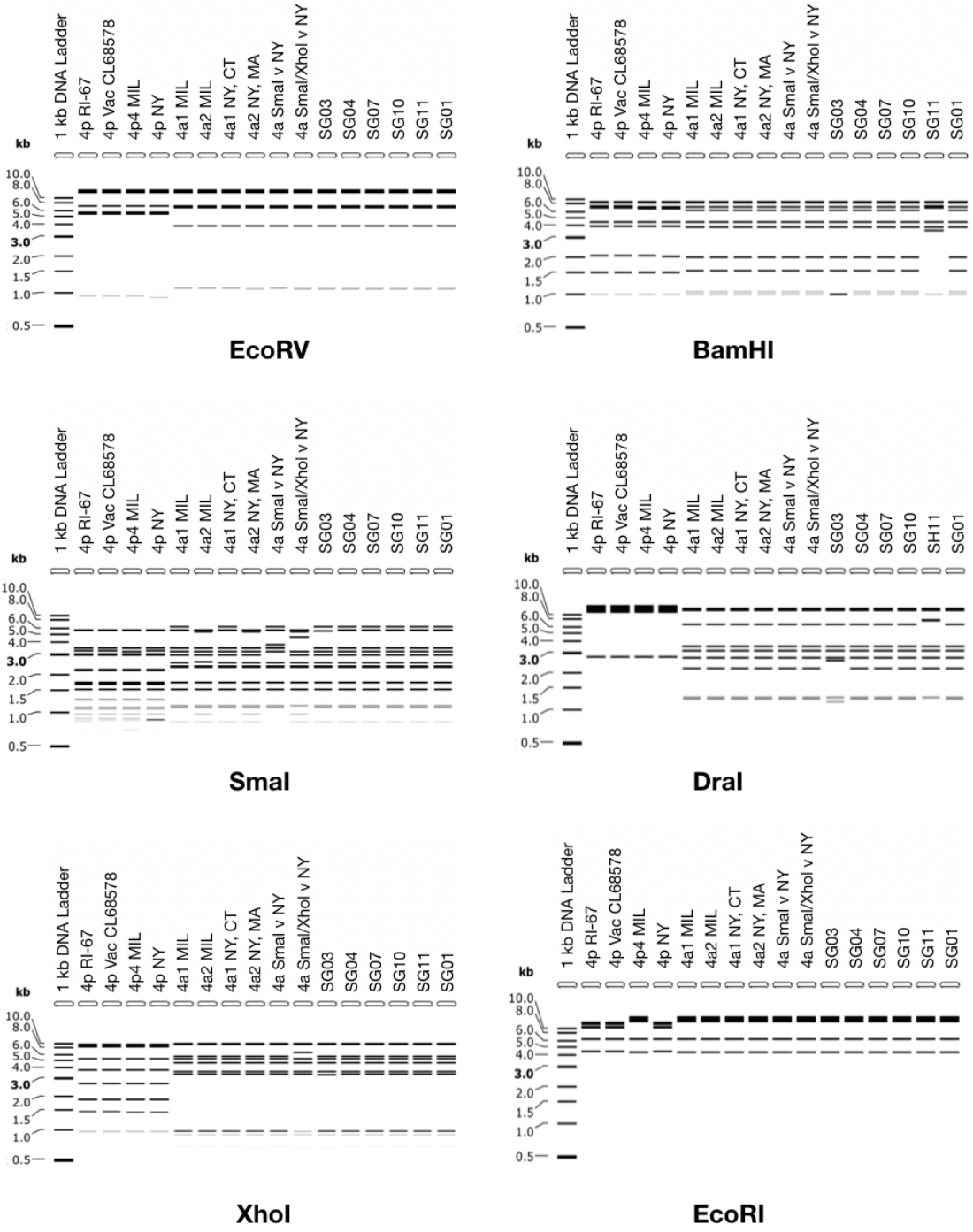


Figure 3.1. Multiple sequence alignment of human adenovirus inverted terminal repeats. The three sequence motifs used for adenovirus DNA replication are boxed and labeled as 1) Core origin; 2) Nuclear factor I (NF-I) motif; and 3) Nuclear factor III (NF-III) motif. HAdV-E4 JAX78, the first recorded HAdV-E4 isolate, is used as a reference at the top and highlighted in red. Note the host-adapted isolates (blue) and the prototype-like strains (green) had noticeable base differences at the NF-I host transcription factor binding motif. Dots and dashes indicate base conservation and missing nucleotides, respectively.

Singapore dataset

An epidemiological analysis of adenoviruses circulating in Singapore was a collaborative work with Duke-NUS Medical School. A total of 533 HAdV-positive clinical samples were collected and genome sequenced during 2012-2018. HAdV-E4 were identified in 15.2% of the cases. As part of the collaboration, our group noted the presence of the NF-

1 ITR recombinant and this study revealed an increase in adenoviruses E4 infections in the civilian population. Further, as shown in Figure 3.2, the circulating genome type was consistent with others reported circulating in the general population (Kajon et al., 2018). This epidemiological study is published (Coleman et al., 2019).



792

793 Figure 3.2. *In silico* restriction enzyme digests for HAdV-E4 genomes collected from

794 Singapore samples collected from 2012 to 2016. Reference sequences are obtained from

NCBI GenBank. Restriction enzyme profiles are generated using SnapGene 5.0 (https://www.snapgene.com). MIL: military isolate; v: variant; Vac: Vaccine strain; SG: Singapore isolate. This shows the Singapore dataset are consistent with the HAdV-E4a variants previously published

Hong Kong Epidemiological Study

A survey of 100 HAdV samples from pediatric patients diagnosed HAdV-positive showed that 10% of isolates are HAdV-E4 (manuscript in prep). Part 1 of this chapter contains publication that examines the ITRs region of the Hong Kong HAdV-E4 dataset. Accordingly, an analysis of the HAdV-E4 transmission adaptations of the NF-I and provides an understanding of the molecular evolution by using phylogenetic analysis. The findings presented in the paper demonstrate the host-adaptations at the NF-I region which enables the novel HAdV-E4 to have an efficient replication in the human host (Dehghan et al., 2013; Mul & Van der Vliet, 1992; Tsurimoto et al., 1989; van der Vliet et al., 1986). This presumably host-adapted strain may have a selective advantage in the immune-naïve civilian populations, as HAdV-E4 was not circulating previously. Part 2 of this chapter analyzed the entire 100 HAdV-positive pediatric samples using various statistical methods and phylogenetic analysis noting that the most prevalent HAdV infections are due to HAdV-B3 and HAdV-E4 (manuscript in preparation). The increase of HAdV-E4 infections in the civilian population should raise concerns, and calls for the biosurveillance of the infectious disease agent.

Beijing swimming pool-associated HAdV-E4 outbreak

Part 3 of this chapter focuses on analyzing a large outbreak of HAdV-E4 at a Beijing university. This outbreak is critical in recognizing that the apparent host-adapted strain may become a major pathogen no longer restricted to the military outbreaks (manuscript in prep). All of these manuscripts are ready to publish.

Part 1. a survey of recent adenoviral respiratory pathogens in Hong Kong reveals emergent and recombinant human adenovirus type 4 (HAdV-E4) circulating in civilian populations

Communication

A Survey of Recent Adenoviral Respiratory Pathogens in Hong Kong Reveals Emergent and Recombinant Human Adenovirus Type 4 (HAdV-E4) Circulating in Civilian Populations

Jing Zhang [1], June Kang [2], Shoaleh Dehghan [2,3], Siddharth Sridhar [4], Susanna K. P. Lau [4], Junxian Ou [1], Patrick C. Y. Woo [4], Qiwei Zhang [1,5,*] and Donald Seto [2,*]

[1] Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, Guangdong 510515, China; 243562394@smu.edu.cn (J.Z.); ojx426@smu.edu.com (J.O.)

[2] Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, VA 20110, USA; qkang@masonlive.gmu.edu (J.K.); shoalehd@gmail.com (S.D.); dseto@gmu.edu (D.S.)

[3] Chemistry Department, American University, Washington, D.C. 20016, USA

[4] Department of Microbiology, The University of Hong Kong, Hong Kong, China; sid8998@gmail.com (S.S.); skplau@hku.hk (S.K.P.L.); pcywoo@hku.hk (P.C.Y.W.)

[5] Guangzhou Key Laboratory of Virology, Institute of Medical Microbiology, Jinan University, Guangzhou 510632, China

* Correspondence: zhangqw@smu.edu.cn (Q.Z.); dseto@gmu.edu (D.S.); Tel: 86-20-6164-8649 (Q.Z.); 703-993-8403 (D.S.)

Received: 18 January 2019; Accepted: 30 January 2019; Published: date

Abstract: Human adenovirus type 4 (HAdV-E4), which is intriguingly limited to military populations, causes acute respiratory disease with demonstrated morbidity and mortality implications. This respiratory pathogen contains genome identity with chimpanzee adenoviruses, indicating zoonotic origins. A signature of these "old" HAdV-E4 is the absence of a critical replication motif, NF-I, which is found in all HAdV respiratory pathogens and most HAdVs. However, our recent survey of flu-like disease in children in Hong Kong reveals that the emergent HAdV-E4 pathogens circulating in civilian populations contain NF-I, indicating recombination and reflecting host-adaptation that enables the "new" HAdV-E4 to replicate more efficiently in human cells and foretells more potential HAdV-E4 outbreaks in immune-naïve civilian populations. Special attention should be paid by clinicians to this emergent and recombinant HAdV-E4 circulating in civilian populations.

Keywords: zoonosis; host adaptation; evolution; genome recombination; respiratory pathogens; civilian populations; human adenovirus type 4

1. Introduction

An increase in pediatric patients with influenza-like symptoms in Queen Mary Hospital (July to October 2014) led to a question as to whether the putatively host-adapted human adenovirus type 4 (HAdV-E4) [1] is circulating in Hong Kong. HAdV-E4 was one of the first viral respiratory pathogens to be isolated ("R*(espiratory)*I*(nfection)*-*(number)*67"; 1952) [2]. It is intriguingly extraordinary in several aspects, the first being that it is one of only two adenoviruses for which a vaccine was developed, thrice (1956, 1971, and 2011), at considerable effort and expense, attesting to its potency (https://www.historyofvaccines.org/content/blog/adenovirus-vaccines-reinstated-after-long-absence) [3]. Each vaccine cessation resulted in immediate reemergence as a major pathogen [4]. Second, HAdV-E4 was typically and inexplicably restricted to U.S. military populations [3,5,6], with occasional yet limited infections tallied in pathogen surveys of civilian populations [5–7]. Third, taxonomically, HAdV-E4 differed from other HAdVs with respect to protein chemistry properties [8] and low-resolution genomics data, including partial sequences and restriction enzyme maps [9]. Its classification, with genome sequence, as the sole HAdV comprising clade or "species" E, with 14 simian adenoviruses (SAdVs), adds to its mystique [1,10]. Origins included speculations of being the archetypical HAdV [9] or a HAdV species B and C recombinant. Genome analysis showed RI-67 was a chimpanzee adenovirus (ChAdV) [1,10], with a genome identity of 89.8% and 89.2% to SAdV-E26 and SAdV-E25, respectively. These are much higher than two respiratory pathogens HAdV-B7 at 73.7% and HAdV-C1 at 65.7%, with ocular pathogen HAdV-D9 at 70.4%.

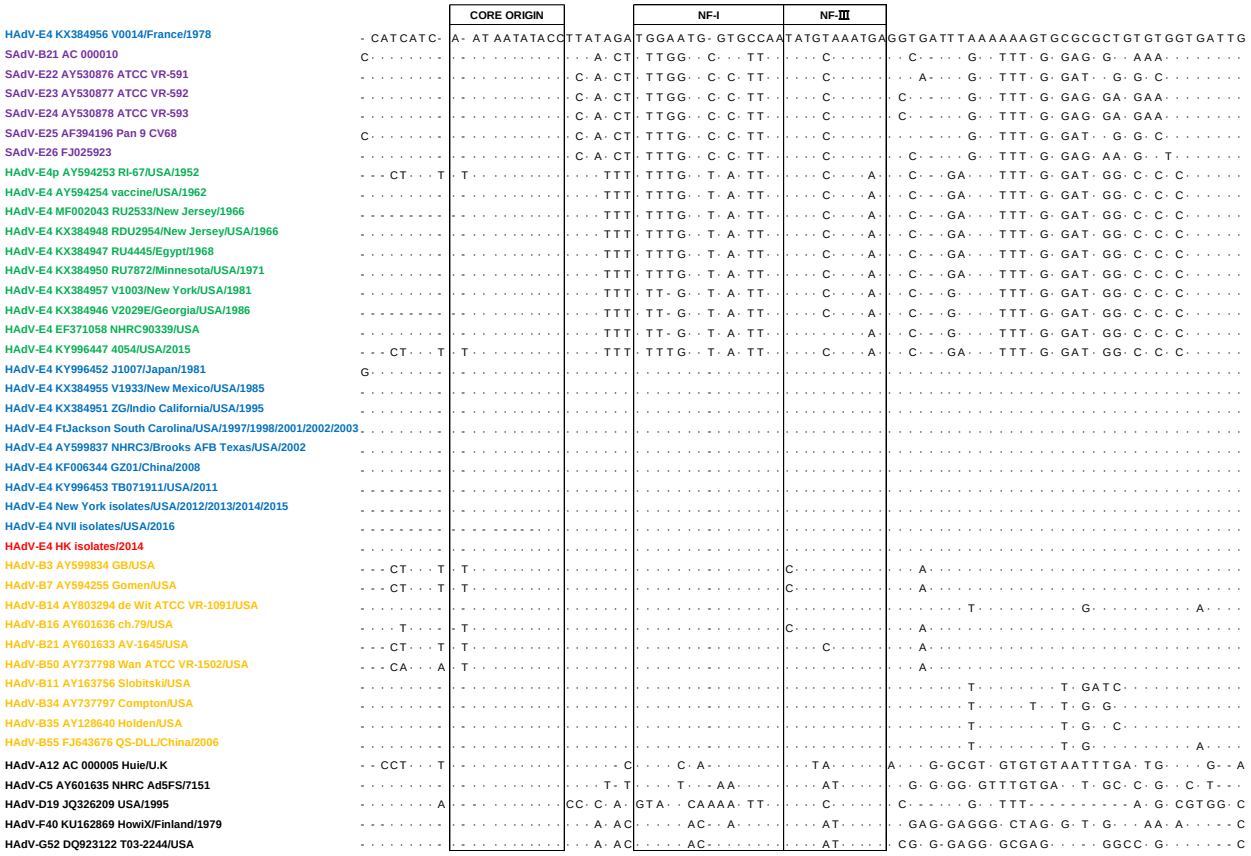
886

887 The zoonotic threat potential of adenoviruses is recognized [11], and HAdV-E4
888 represents the first example. Notably, both the HAdV-E4 prototype and contemporaneous
889 "vaccine" isolates contain a genome signature that is unique to ChAdVs and other
890 SAdVs, i.e., the absence of the Nuclear Factor I (NF-I) binding site that is conserved
891 among the three replication motifs embedded in nearly all HAdV inverted terminal
892 repeats (ITRs) [1]. NF-I is a host transcription factor that binds to nt 23–36 of the HAdV-
893 2 origin of replication [12–20] and is recruited by the human adenoviral replication
894 complex [12,17]. It is firmly established as essential through in vitro and in vivo studies
895 (see [18,19] and *qtd. in*), including replication reconstitution assays, as necessary for
896 efficient adenoviral replication in human cells (see [20] and *qtd. in*). Recent isolates are
897 recombinants containing this HAdV replication motif [1], presumably permitting an
898 expansion of the virus range into the immune-naïve populations [5,6,21–24], and should
899 be noted as a molecular evolution example of a post-zoonotic, host-adaptation of a
900 "novel" and "emergent" human pathogen.
901

902 As noted in the literature, the ITR contains critical conserved DNA replication motifs that
903 include the human transcription factor binding site, NF-I, which is shown to enhance and
904 optimize HAdV replication and growth [2,12–20]. All of the 10 Hong Kong isolates (red)
905 possess this host-adapted ITR, i.e., they have the NF-I binding site that is found in all
906 human respiratory adenoviruses (yellow) (Figure 3.1.1). However, this motif is missing
907 in SAdV ITRs (purple) [18], as well as the "old" HAdV-E4 strains (green), e.g.,

prototype (1952) and “vaccine” (1962) strains, i.e., they have only the core origin and the NF-III motif [1]. The absence of the NF-I motif likely plays a role in limiting the circulation of both SAdVs and prototype HAdV-E4 in human populations, as its acquisition presumably provides efficient and enhanced replication in human cells, i.e., allowed HAdV-E4 to adapt to the new host. This may be the “tipping point” that allows HAdV-E4 entry into the general population that is immunologically-naïve to its epsilon antigen, as HAdV-E4 does not normally circulate outside the U.S. military populations [3,5,6,7]. This is consistent with reports documenting recombination as an evolution mechanism in the genesis of emergent HAdV pathogens [25], illustrated by HAdV-D53 [26] and HAdV-B55 [25]. HAdV-B55 is a recent major respiratory pathogen in China, after years of quiescence, by virtue of its introduction into immunologically-naïve populations, i.e., without antibodies to its HAdV-B11-like epsilon antigen. Another example is the reemergence of a long-dormant respiratory pathogen HAdV-B14 into an immunologically-naïve population [27].

Consistent with this report, recombination has been reported in SAdVs [28,29] and cross-species transmissions between humans and non-human simians have been reported [29–35], including TMAdV across “two passages” of human hosts, with clinical manifestations [35], as well as seroprevalence of adenoviruses across species [30,33–35]. These support the hypothesis and data suggesting HAdV-E4 was of zoonotic origins [1,10].

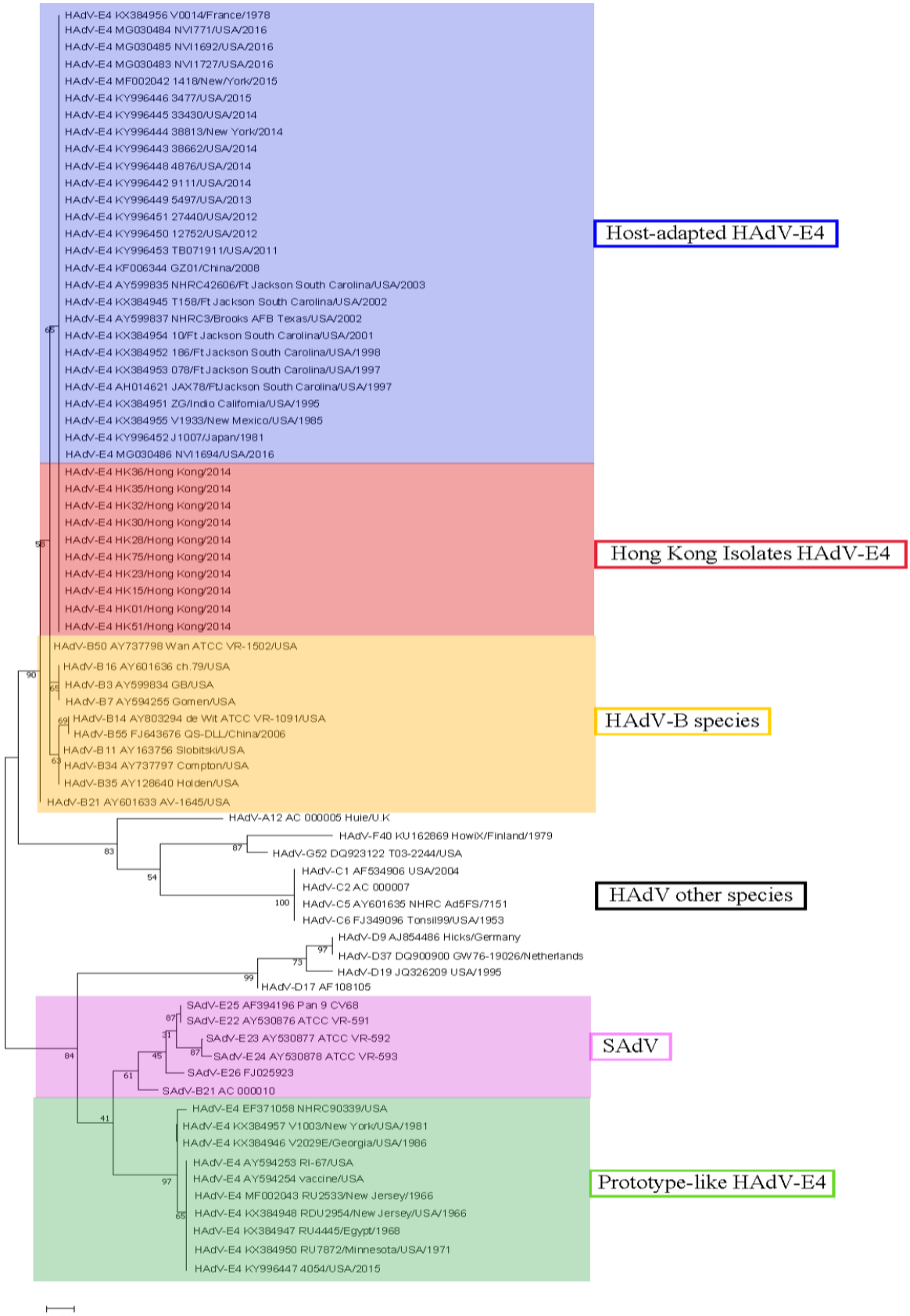


930

931 Figure 4.1.1. Human adenovirus Inverted Terminal Repeats (116–209 bases) embed
932 replication protein-binding motifs within 88 bases: “Core origin”, NF-I (Nuclear Factor
933 I), and NF-III. Host transcription factors NF-I and NF-III enable adenovirus replication in
934 human cells. The first recorded host-adapted HAdV-E4 (Human adenovirus type 4) ITR
935 (inverted terminal repeats) is arrayed along the top (V0014/France/1978), contrasted
936 against differences in motifs from SAdVs (simian adenoviruses) (lavender) and
937 prototype-like HAdV-E4 (green). Dots indicate conserved bases. Recent host-adapted
938 HAdV-E4 isolates (blue), including five isolates from Ft. Jackson, spanning six years,
939 and 10 recent Hong Kong isolates (red) contain these motifs and are nearly

indistinguishable from motifs from HAdV respiratory pathogens (yellow). The majority of HAdVs possess these motifs (black), with the most divergent being HAdV-D19.

Since sequences diverge considerably downstream of the NF-III motif, a phylogenetic analysis of the full ITR (116–209 bases) was completed (Figure 3.1.2). Recent host-adapted "new" HAdV-E4 isolates, including Hong Kong isolates, subclade with human respiratory adenoviruses of species B as well as with other HAdVs. Prototype-like HAdV-E4 isolates subclade with SAdVs.


HAdV-E4 KX384956 V0014/France/1978
HAdV-E4 MG030484 NVI771/USA/2016
HAdV-E4 MG030485 NVI1692/USA/2016
HAdV-E4 MG030483 NVI1727/USA/2016
HAdV-E4 MF002042 1418/New/York/2015
HAdV-E4 KY996446 3477/USA/2015
HAdV-E4 KY996445 33430/USA/2014
HAdV-E4 KY996444 38813/New York/2014
HAdV-E4 KY996443 38662/USA/2014
HAdV-E4 KY996448 4876/USA/2014
HAdV-E4 KY996442 9111/USA/2014
HAdV-E4 KY996449 5497/USA/2013
HAdV-E4 KY996451 27440/USA/2012
HAdV-E4 KY996450 12752/USA/2012
HAdV-E4 KY996453 TB071911/USA/2011
HAdV-E4 KF006344 GZ01/China/2008
HAdV-E4 AY599835 NHRC42606/Ft Jackson South Carolina/USA/2003
HAdV-E4 KX384945 T158/Ft Jackson South Carolina/USA/2002
HAdV-E4 AY599837 NHRC3/Brooks AFB Texas/USA/2002
HAdV-E4 KX384954 10/Ft Jackson South Carolina/USA/2001
HAdV-E4 KX384952 186/Ft Jackson South Carolina/USA/1998
HAdV-E4 KX384953 078/Ft Jackson South Carolina/USA/1997
HAdV-E4 AH014621 JAX78/FtJackson South Carolina/USA/1997
HAdV-E4 KX384951 ZG/Indio Califomia/USA/1995
HAdV-E4 KX384955 V1933/New Mexico/USA/1985
HAdV-E4 KY996452 J1007/Japan/1981
HAdV-E4 MG030486 NVI1694/USA/2016
Host-adapted HAdV-E4
HAdV-E4 HK36/Hong Kong/2014
HAdV-E4 HK35/Hong Kong/2014
HAdV-E4 HK32/Hong Kong/2014
HAdV-E4 HK30/Hong Kong/2014
HAdV-E4 HK28/Hong Kong/2014
HAdV-E4 HK75/Hong Kong/2014
HAdV-E4 HK23/Hong Kong/2014
HAdV-E4 HK15/Hong Kong/2014
HAdV-E4 HK01/Hong Kong/2014
HAdV-E4 HK51/Hong Kong/2014
Hong Kong Isolates HAdV-E4
HAdV-B50 AY737798 Wan ATCC VR-1502/USA
HAdV-B16 AY601636 ch.79/USA
HAdV-B3 AY599834 GB/USA
HAdV-B7 AY594255 Gomen/USA
HAdV-B14 AY803294 de Wit ATCC VR-1091/USA
HAdV-B55 FJ643676 QS-DLL/China/2006
HAdV-B11 AY163756 Slobitski/USA
HAdV-B34 AY737797 Compton/USA
HAdV-B35 AY128640 Holden/USA
HAdV-B21 AY601633 AV-1645/USA
HAdV-B species
HAdV-A12 AC 000005 Huie/U.K
HAdV-F40 KU162869 HowiX/Finland/1979
HAdV-G52 DQ923122 T03-2244/USA
HAdV-C1 AF534906 USA/2004
HAdV-C2 AC 000007
HAdV-C5 AY601635 NHRC Ad5FS/7151
HAdV-C6 FJ349096 Tonsil99/USA/1953
HAdV other species
HAdV-D9 AJ854486 Hicks/Germany
HAdV-D37 DQ900900 GW76-19026/Netherlands
HAdV-D19 JQ326209 USA/1995
HAdV-D17 AF108105
SAdV-E25 AF394196 Pan 9 CV68
SAdV-E22 AY530876 ATCC VR-591
SAdV-E23 AY530877 ATCC VR-592
SAdV-E24 AY530878 ATCC VR-593
SAdV-E26 FJ025923
SAdV-B21 AC 000010
SAdV
HAdV-E4 EF371058 NHRC90339/USA
HAdV-E4 KX384957 V1003/New York/USA/1981
HAdV-E4 KX384946 V2029E/Georgia/USA/1986
HAdV-E4 AY594253 RI-67/USA
HAdV-E4 AY594254 vaccine/USA
HAdV-E4 MF002043 RU2533/New Jersey/1966
HAdV-E4 KX384948 RDU2954/New Jersey/USA/1966
HAdV-E4 KX384947 RU4445/Egypt/1968
HAdV-E4 KX384950 RU7872/Minnesota/USA/1971
HAdV-E4 KY996447 4054/USA/2015
Prototype-like HAdV-E4
0.050

949

Figure 4.1.2. Phylogeny of select HAdV and SAdV ITRs. Recent host-adapted HAdV-E4, including Hong Kong isolates, subclade with HAdVs. Prototype-like HAdV-E4 isolates subclade with SAdVs. Inverted terminal repeats (*ca.*, 116 bases) were aligned using Molecular Evolutionary Genetics Analysis (MEGA) 7.0.26 software (https://www.megasoftware.net), and phylogenetic trees were constructed with a maximum-likelihood method incorporating 1000 bootstrap replications.

Comparison of HAdV-E4 genomes spanning 65 years reveals highly conserved sequences. For example, the two earliest genomes, isolated within 10 years, are 99.9% identical to each other, differing by 18 mismatches and two insertions (one one-base and one three-base) [36]. One striking difference is a recombination resulting in recent HAdV-E4 genomes acquiring a well-characterized replication-enhancing NF-I motif that distinguishes these isolates from the 1952 prototype and a contemporaneous 1965 co-circulating strain. Although the acquisition of this critical replication motif alone may explain the potentially wider distribution of HAdV-E4, the pair-wise sequence analyses also reveal additional potentially important sequence differences in the E3 gene region (data not shown). Proteins encoded in the E3 region are involved in interactions with the host immune response and evasion and represent the most variable region between HAdV types [37–39]. This report highlights the need to re-explore the understudied E3 genes and their critical roles in the human immune response upon HAdV infection.

971 However, whether in concert with E3 gene changes or alone, the NF-I acquisition, given
972 "Occam's Razor", may be the critical evolutionary modification and "tipping point" that
973 is necessary for a wider distribution of HAdV-E4 in immune-naïve civilian populations in
974 light of the extensive literature reported for HAdV replication requirements [12–20]. This
975 host adaptation may be significant as HAdV-E4 is a highly contagious pathogen with
976 demonstrated morbidity and mortality implications and has now been reported in global
977 civilian populations [6,7,24]. During the preparation of this report, to emphasize further
978 the potential of HAdV-E4 emergence in the global general population,
979 seroepidemiological data identifying HAdV-E4 infections in China and Sierra Leone
980 (Africa) were published [40]. Explicitly, the authors note "An apparent increase in the
981 frequency of human adenovirus type 4 (HAdV-4) infections among general populations
982 has been observed over the past 10 years" [40].
983

984 **2. Materials and Methods**

985 *2.1. Clinical Specimens and Virus Culture*

986 Nasopharyngeal swab specimens were collected from both outpatients and inpatients who
987 presented with flu-like symptoms, and are archived at Queen Mary Hospital (Hong
988 Kong). Adenoviruses were detected by PCR and were identified further by molecular
989 typing using partial sequence data from the hexon and fiber genes, as previously reported
990 [41]. Then, 10 virus strains that were identified as type 4 were cultured in human lung

991 cell lines (A549), and genomic DNA was extracted using the Viral DNA Extraction Kit
992 from Omega Bio-Tek, Inc. (Norcross, GA, USA), as previously noted [41].
993

994 *2.2. DNA Sequencing and Bioinformatic Analyses*

995 ITRs were sequenced using genomic DNA as a template, with either a 5'-primer "Ad4-
996 ITR-L" (5'-AACTCTTCTCGCTGGCACTCAA-3') or a 3'-primer "Ad4-ITR-R" (5'-
997 CCGCCCCTAACAGTCGCC-3'). These Sanger chemistry-derived ITR sequence data
998 generated with an ABI3730 DNA sequencer were aligned and parsed for sequence
999 motifs. Phylogenetic analysis was undertaken with Molecular Evolutionary Genetics
1000 Analysis software (MEGA) 7.0.26 (https://www.megasoftware.net), for which FASTA
1001 files were inputted for sequence alignments utilizing a Maximum Composite Likelihood
1002 method to generate neighbor-joining, bootstrapped phylogenetic trees with 1000
1003 iterations (Tamura et al., 2007). All other MEGA program parameters were set by
1004 default. Bootstrap values above 80 are considered robust. For pair-wise genome
1005 comparisons, zPicture (http://zpicture.dcode.org/) was used to provide detailed global
1006 insights.
1007

1008 *2.3. Sequence Data*

1009 All sequences are available in GenBank from which whole genome data were retrieved
1010 and aligned, and ITR sequences extracted. This resulted in a data set of ITR sequences
1011 ranging from 116 to 209 bases, as lengths were variable among the types. A subset of

1012 sequences spanning the first 1–66 bases that included the conserved HAdV viral
1013 replication sequence motifs were selected for detailed analysis.
1014
1015 Genome sequence data were accessed from GenBank. HAdV-E4 genomes include RI-67
1016 or “prototype” (Ft Leonard Wood, Missouri; 1952; AY594253); CL68578 or “vaccine”
1017 (Camp Lejeune, North Carolina; 1965; AY487947); V0014 (France; 1978; KX384956);
1018 JAX78 (Ft Jackson, South Carolina; 1997; KX384953); and NHRC3 (Brooks AFB,
1019 Texas; 2002; AY599837); NVI771 (USA; 2016; MG030484); NVI692 (USA; 2016;
1020 MG030485); NVI727 (USA; 2016; MG030483); 1418 (New York; 2015; MF002042);
1021 3477 (USA; 2015; KY996446); 33430 (USA; 2014; KY996445); 38813 (New York;
1022 2014; KY996444); 38662 (USA; 2014; KY996443); 4876 (USA; 2014; KY996448);
1023 9111 (USA; 2014; KY996442); 5497 (USA; 2013; KY996449); 27440 (USA; 2012;
1024 KY996451); 12752 (USA; 2012; KY996450); TB071911 (USA; 2011; KY996453);
1025 GZ01 (China; 2008; KF006344); NHRC42606 (Ft Jackson, South Carolina; 2003;
1026 AY599835); T158 (Ft Jackson, South Carolina; 2002; KX384945); 10 (Ft Jackson, South
1027 Carolina; 2001; KX384954); 186 (Ft Jackson, South Carolina; 1998; KX384952); 078 (Ft
1028 Jackson, South Carolina; 1997; KX384953); ZG (Indio, California; 1995; KX384951);
1029 V1933 (New Mexico; 1985; KX384955); J1007; Japan; 1981; KY996452); NVI1694
1030 (USA; 2016; MG030486); NHRC90339 (USCG Training Ctr (Cape May), Delaware;
1031 EF371058); V1003 (New York; 1981; KX384957); V2029E (Georgia; 1986;
1032 KX384946); RU2533 (New Jersey; 1966; MF002043); RDU2954 (New Jersey; 1966;
1033 KX384948); RU4445 (Egypt; 1968; KX384947); RU7872 (Minnesota; 1971;

KX384950); and 4054 (USA; 2015; KY996447). Additional representative HAdV genomes include HAdV-B50 (Wan; AY737798); HAdV-B16 (ch.79; AY601636); HAdV-B3 (GB; AY599834); HAdV-B7 (Gomen; AY594255); HAdV-B14 (de Wit; AY803294); HAdV-B55 (QS-DL; FJ643676); HAdV-B11 (Slobitski; AY163756); HAdV-B34 (Compton; AY737797); HAdV-B35 (Holden; AY128640); HAdV-B21 (AV-1645; AY601633); HAdV-A12 (Huie; AC_000005); HAdV-F40 (HowiX; FH162869); HAdV-G52 (T03-2244; DQ923122); HAdV-C1 (AF534906); HAdV-C2 (AC_000007); HAdV-C5 (7151; AY601635); HAdV-C6 (Tonsil 99; FJ349096); HAdV-D9 (Hicks; AJ854486); HAdV-D37 (GW76-19026; DQ900900); HAdV-D19 (JQ326209); and HAdV-D17 (AF108105). Simian adenoviral genomes include SAdV-B21 (AC_000010); SAdV-E22 (AY530876); SAdV-E23 (AY530877); SAdV-E24 (AY530878); SAdV-E25 (AF394196); and SAdV-E26 (FJ025923). The ITR sequences of Hong Kong HAdV-E4 isolates are archived in GenBank (KY062614–KY062623).

Acknowledgments: D.S. thanks Q.Z. and the Southern Medical University for an appointment as Visiting Professor, and also Professor K.Y. Yuen (Hong Kong University) for the lecture invitation that initiated this collaboration. D.S. is grateful to Mr. Harold Suarez del Toro and the staff, particularly Mr. Margarito Ricardez de la Cruz of the Biwa Hotel in Tulum (Mexico) for providing a warm, tranquil, and hospitable environment that was conducive to a writing sabbatical in Winter 2018-19.

Funding: This work was supported by the National Natural Science Foundation of China (31570155, 31370199) and Guangzhou Healthcare Collaborative Innovation Major Project (201803040004, 201803040007).

Author contributions: Q.Z. and D.S. conceived of this project, as well as designed and analyzed the data, and prepared the manuscript. J.Z., J.O., J.K., and S.D. performed the analyses, analyzed the data, and contributed to the manuscript; S.S., S.K.P.L., and P.C.Y.W provided specimens and clinical metadata.

Competing interests: The authors declare no competing interests.

Reference

1. Dehghan, S.; Seto, J.; Liu, E.B.; Walsh, M.P.; Dyer, D.W.; Chodosh, J.; Seto, D. Computational analysis of four human adenovirus type 4 genomes reveals molecular evolution through two interspecies recombination events. *Virology* **2013**, *443*, 197–207.
2. Hilleman, M.R.; Werner, J.H. Recovery of new agent from patients with acute respiratory illness. Proceedings of the Society for Experimental Biology and Medicine. *Exp. Biol. Med.* **1954**, *85*, 183–188.

3. Potter, R.N.; Cantrell, J.A.; Mallak, C.T.; Gaydos, J.C. Adenovirus-associated deaths in us military during postvaccination period, 1999–2010. *Emerg. Infect. Dis.* **2012**, *18*, 507–509.
4. Barraza, E.M.; Ludwig, S.L.; Gaydos, J.C.; Brundage, J.F. Reemergence of adenovirus type 4 acute respiratory disease in military trainees: Report of an outbreak during a lapse in vaccination. *J. Infect. Dis.* **1999**, *179*, 1531–1533.
5. Kajon, A.E.; Moseley, J.M.; Metzgar, D.; Huong, H.S.; Wadleigh, A.; Ryan, M.A.; Russell, K.L. Molecular epidemiology of adenovirus type 4 infections in us military recruits in the postvaccination era (1997–2003). *J. Infect. Dis.* **2007**, *196*, 67–75.
6. Kajon, A.E.; Lamson, D.M.; Bair, C.R.; Lu, X.Y.; Landry, M.L.; Menegus, M.; Erdman, D.D.; St George, K. Adenovirus type 4 respiratory infections among civilian adults, northeastern united states, 2011–2015. *Emerg. Infect. Dis.* **2018**, *24*, 201–209.
7. Hang, J.; Vento, T.J.; Norby, E.A.; Jarman, R.G.; Keiser, P.B.; Kuschner, R.A.; Binn, L.N. Adenovirus type 4 respiratory infections with a concurrent outbreak of coxsackievirus a21 among united states army basic trainees, a retrospective viral etiology study using next-generation sequencing. *J. Med. Virol.* **2017**, *89*, 1387–1394.
8. Wadell, G. Classification of human adenoviruses by sds-polyacrylamide gel electrophoresis of structural polypeptides. *Intervirology* **1979**, *11*, 47–57.
9. Li, Q.G.; Wadell, G. The degree of genetic variability among adenovirus type 4 strains isolated from man and chimpanzee. *Arch. Virol.* **1988**, *101*, 65–77.

10. Purkayastha, A.; Ditty, S.E.; Su, J.; McGraw, J.; Hadfield, T.L.; Tibbetts, C.; Seto, D. Genomic and bioinformatics analysis of hadv-4, a human adenovirus causing acute respiratory disease: Implications for gene therapy and vaccine vector development. *J. Virol.* **2005**, *79*, 2559–2572.
11. Bailey, E.S.; Fieldhouse, J.K.; Choi, J.Y.; Gray, G.C. A mini review of the zoonotic threat potential of influenza viruses, coronaviruses, adenoviruses, and enteroviruses. *Front. Public Health* **2018**, *6*, 104.
12. Nagata, K.; Guggenheimer, R.A.; Hurwitz, J. Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. USA.* **1983**, *80*, 6177–6181.
13. Rawlins, D.R.; Rosenfeld, P.J.; Wides, R.J.; Challberg, M.D.; Kelly, T.J., Jr. Structure and function of the adenovirus origin of replication. *Cell* **1984**, *37*, 309–319.
14. Diffley, J.F.; Stillman, B. Purification of a cellular, double-stranded DNA-binding protein required for initiation of adenovirus DNA replication by using a rapid filter-binding assay. *Mol. Cell. Biol.* **1986**, *6*, 1363–1373.
15. Rosenfeld, P.J.; Kelly, T.J. Purification of nuclear factor i by DNA recognition site affinity chromatography. *J. Biol. Chem.* **1986**, *261*, 1398–1408.
16. De Vries, E.; van Driel, W.; van den Heuvel, S.J.; van der Vliet, P.C. Contactpoint analysis of the hela nuclear factor i recognition site reveals symmetrical binding at one side of the DNA helix. *EMBO J.* **1987**, *6*, 161–168.

17. Kenny, M.K.; Balogh, L.A.; Hurwitz, J. Initiation of adenovirus DNA replication. I. Mechanism of action of a host protein required for replication of adenovirus DNA templates devoid of the terminal protein. *J. Biol. Chem.* **1988**, *263*, 9801–9808.
18. Hatfield, L.; Hearing, P. Redundant elements in the adenovirus type 5 inverted terminal repeat promote bidirectional transcription in vitro and are important for virus growth in vivo. *Virology* **1991**, *184*, 265–276.
19. Mul, Y.M.; Verrijzer, C.P.; van der Vliet, P.C. Transcription factors nfi and nfiii/oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. *J. Virol.* **1990**, *64*, 5510–5518.
20. Stillman, B.W.; Topp, W.C.; Engler, J.A. Conserved sequences at the origin of adenovirus DNA replication. *J. Virol.* **1982**, *44*, 530–537.
21. Kandel, R.; Srinivasan, A.; D'Agata, E.M.; Lu, X.; Erdman, D.; Jhung, M. Outbreak of adenovirus type 4 infection in a long-term care facility for the elderly. *Infect. Cont. Hosp. Epidemiol.* **2010**, *31*, 755–757.
22. Chen, H.L.; Chiou, S.S.; Hsiao, H.P.; Ke, G.M.; Lin, Y.C.; Lin, K.H.; Jong, Y.J. Respiratory adenoviral infections in children: A study of hospitalized cases in southern taiwan in 2001–2002. *J. Trop. Pediatr.* **2004**, *50*, 279–284.
23. Hang, J.; Vento, T.J.; Norby, E.A.; Jarman, R.G.; Keiser, P.B.; Kuschner, R.A.; Binn, L.N. Adenovirus type 4 respiratory infections with a concurrent outbreak of coxsackievirus a21 among united states army basic trainees, a retrospective viral etiology study using next-generation sequencing. *J. Med. Virol.* **2017**, *89*, 1387–1394.

24. Li, J.; Lu, X.; Sun, Y.; Lin, C.; Li, F.; Yang, Y.; Liang, Z.; Jia, L.; Chen, L.; Jiang, B.; et al. A swimming pool-associated outbreak of pharyngoconjunctival fever caused by human adenovirus type 4 in beijing, China. *Int. J. Infect. Dis.* **2018**, *75*, 89–91.
25. Walsh, M.P.; Seto, J.; Jones, M.S.; Chodosh, J.; Xu, W.; Seto, D. Computational analysis identifies human adenovirus type 55 as a re-emergent acute respiratory disease pathogen. *J. Clin. Microbiol.* **2010**, *48*, 991–993.
26. Walsh, M.P.; Chintakuntlawar, A.; Robinson, C.M.; Madisch, I.; Harrach, B.; Hudson, N.R.; Schnurr, D.; Heim, A.; Chodosh, J.; Seto, D.; et al. Evidence of molecular evolution driven by recombination events influencing tropism in a novel human adenovirus that causes epidemic keratoconjunctivitis. *PloS ONE* **2009**, *4*, e5635.
27. Carr, M.J.; Kajon, A.E.; Lu, X.; Dunford, L.; O'Reilly, P.; Holder, P.; De Gascun, C.F.; Coughlan, S.; Connell, J.; Erdman, D.D.; et al. Deaths associated with human adenovirus-14p1 infections, europe, 2009–2010. *Emerg. Infect. Dis.* **2011**, *17*, 1402–1408.
28. Hoppe, E.; Pauly, M.; Robbins, M.; Gray, M.; Kujirakwinja, D.; Nishuli, R.; Boji Mungu-Akonkwa, D.D.; Leendertz, F.H.; Ehlers, B. Phylogenomic evidence for recombination of adenoviruses in wild gorillas. *J. Gen. Virol.* **2015**, *96*, 3090–3098.
29. Dehghan, S.; Seto, J.; Jones, M.S.; Dyer, D.W.; Chodosh, J.; Seto, D. Simian adenovirus type 35 has a recombinant genome comprising human and simian

adenovirus sequences, which predicts its potential emergence as a human respiratory pathogen. *Virology* **2013**, *447*, 265–273.

30. Pauly, M.; Akoua-Koffi, C.; Buchwald, N.; Schubert, G.; Weiss, S.; Couacy-Hymann, E.; Anoh, A.E.; Mossoun, A.; Calvignac-Spencer, S.; Leendertz, S.A.; et al. Adenovirus in rural cote d'ivoire: High diversity and cross-species detection. *Ecohealth* **2015**, *12*, 441–452.
31. Hoppe, E.; Pauly, M.; Gillespie, T.R.; Akoua-Koffi, C.; Hohmann, G.; Fruth, B.; Karhemere, S.; Madinda, N.F.; Mugisha, L.; Muyembe, J.J.; et al. Multiple cross-species transmission events of human adenoviruses (HADV) during hominine evolution. *Mol. Biol. Evol.* **2015**, *32*, 2072–2084.
32. Wevers, D.; Metzger, S.; Babweteera, F.; Bieberbach, M.; Boesch, C.; Cameron, K.; Couacy-Hymann, E.; Cranfield, M.; Gray, M.; Harris, L.A.; et al. Novel adenoviruses in wild primates: A high level of genetic diversity and evidence of zoonotic transmissions. *J. Virol.* **2011**, *85*, 10774–10784.
33. Chiu, C.Y.; Yagi, S.; Lu, X.; Yu, G.; Chen, E.C.; Liu, M.; Dick, E.J., Jr.; Carey, K.D.; Erdman, D.D.; Leland, M.M.; et al. A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. *mBio* **2013**, *4*, e00084.
34. Ersching, J.; Hernandez, M.I.; Cezarotto, F.S.; Ferreira, J.D.; Martins, A.B.; Switzer, W.M.; Xiang, Z.; Ertl, H.C.; Zanetti, C.R.; Pinto, A.R. Neutralizing antibodies to human and simian adenoviruses in humans and new-world monkeys. *Virology* **2010**, *407*, 1–6.

35. Chen, E.C.; Yagi, S.; Kelly, K.R.; Mendoza, S.P.; Tarara, R.P.; Canfield, D.R.; Maninger, N.; Rosenthal, A.; Spinner, A.; Bales, K.L.; et al. Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. *PLoS Pathog.* **2011**, *7*, e1002155.
36. Purkayastha, A.; Su, J.; McGraw, J.; Ditty, S.E.; Hadfield, T.L.; Seto, J.; Russell, K.L.; Tibbetts, C.; Seto, D. Genomic and bioinformatics analyses of hadv-4vac and hadv-7vac, two human adenovirus (hadv) strains that constituted original prophylaxis against hadv-related acute respiratory disease, a reemerging epidemic disease. *J. Clin. Microbiol.* **2005**, *43*, 3083–3094.
37. Singh, G.; Robinson, C.M.; Dehghan, S.; Schmidt, T.; Seto, D.; Jones, M.S.; Dyer, D.W.; Chodosh, J. Overreliance on the hexon gene, leading to misclassification of human adenoviruses. *J. Virol.* **2012**, *86*, 4693–4695.
38. Benedict, C.A.; Norris, P.S.; Prigozy, T.I.; Bodmer, J.L.; Mahr, J.A.; Garnett, C.T.; Martinon, F.; Tschopp, J.; Gooding, L.R.; Ware, C.F. Three adenovirus e3 proteins cooperate to evade apoptosis by tumor necrosis factor-related apoptosis-inducing ligand receptor-1 and -2. *J. Biol. Chem.* **2001**, *276*, 3270–3278.
39. Lichtenstein, D.L.; Toth, K.; Doronin, K.; Tollefson, A.E.; Wold, W.S.M. Functions and mechanisms of action of the adenovirus e3 proteins. *Int. Rev. Immunol.* **2004**, *23*, 75–111.
40. Wang, B.; Li, J.; Wu, S.; Chen, Y.; Zhang, Z.; Zhai, Y.; Guo, Q.; Zhang, J.; Song, X.; Zhao, Z.; et al. Seroepidemiological investigation of hadv-4 infection among

healthy adults in china and in sierra leone, west africa. *Emerg. Microbes Infect.* **2018**, *7*, 200.

41. Han, G.; Niu, H.; Zhao, S.; Zhu, B.; Wang, C.; Liu, Y.; Zhang, M.; Yang, S.; Liu, F.; Wan, C.; et al. Identification and typing of respiratory adenoviruses in guangzhou, southern china using a rapid and simple method. *Virol. Sin.* **2013**, *28*, 103–108.

Part 2. molecular epidemiology of adenovirus type 4 acute respiratory cases amongst Hong Kong pediatric patients in 2014

Molecular Epidemiology of Adenovirus Type 4 Acute Respiratory Cases Amongst Hong Kong Pediatric Patients in 2014

June Kang[1a], Jing Zhang[2], Siddhartha Sridhar[3], Susanna K. P. Lau[3], Junxian Ou[2], Yuqian Yan[2], Shan Zhao[2], Wendong Lan[2], Wenyi Guan[2], Patrick C. Y. Woo[3*], Qiwei Zhang[2, 4*], and Donald Seto[1*]

[1]Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

[2]Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou Guangdong 510515, China

[3]Department of Microbiology, The University of Hong Kong, Hong Kong, China

[4]Guangzhou Key Laboratory of Virology, Institute of Medical Microbiology, Jinan University, Guangzhou 510632, China

*Corresponding authors: Donald Seto (dseto@gmu.edu); Patrick C.Y Woo (pcywoo@hkucc.hku.hk); Qiwei Zhang (zhangqw@smu.edu.cn).

Keywords: Adenovirus; Pediatric disease; Infectious disease; Human pathogen; Genotyping; Molecular epidemiology; Respiratory disease.

Short running title: epidemiology study of pediatric patients infected with adenovirus isolated in 2014.

Abstract

To better understand the molecular evolution and epidemiology distribution of adenoviruses in Hong Kong, we investigated 100 HAdV-positive hospital pediatric samples collected during the summer of 2014. Clinical samples were cultured and pathogens identified using PCR and subsequent genomic sequencing. Statistical methods and phylogenetic analyses were performed to characterize the epidemiological data collected. This study found that HAdV-B3 accounted for 74% of the adenoviral-related respiratory infections in these pediatric patients. HAdV-E4 (10%) was also found in this population, not usually observed in civilian populations. Patients infected with HAdV-E4 had a high risk for upper respiratory tract infections (OR, 2.33; 95% CI, 0.51-10.67) and coinfection with other pathogens (OR, 2.33; 95% CI, 0.51-10.67). In managing epidemics, it is prudent to be attentive and circumspect to emerging or re-emerging HAdV pathogens.

BACKGROUND

Human adenoviruses (HAdV) are double-stranded DNA viruses that can cause a wide spectrum of human illnesses including respiratory, ocular, gastrointestinal, and urinary infections (Harrach et al., 2011; Lion, 2014; Purkayastha et al., 2005; Rowe et al., 1953; Wood, 1988). There are seven HAdV species (A-G) that group together more than 103 genotypes(Harrach et al., 2011). The diverse illnesses presented by HAdVs vary with species and genotypes (Harrach et al., 2011). Among the seven species, species C infects the respiratory tracts and are also associated with latent infections (Garnett et al., 2009; Shayakhmetov et al., 2005; Yao et al., 2019). HAdV species F and G are commonly associated with gastroenteritis (Wood, 1988). Species B and E can cause acute respiratory disease (ARD), as well as ocular diseases such as conjunctivitis (Lion, 2014; Rowe et al., 1953). Approximately 5-7% of the respiratory illness diagnosed in pediatric patients have been attributed to HAdV species B and E infections (Carballal et al., 2001; Hong et al., 2001). Species C and B are commonly observed in children hospitalized with ARD (Carballal et al., 2001; Larrañaga et al., 2000; Videla et al., 1999).

HAdV-B3, -E4, and -B7 are commonly associated with ARD outbreaks (Chang et al., 2008; Gerber et al., 2001; Li et al., 2018). HAdV-B3 is associated with milder symptoms in both adults and children while HAdV-E4 and HAdV-B7 can cause frequent outbreaks with more severe symptoms (Chang et al., 2008; Gray et al., 2007; Schmitz et al., 1983). Vaccines for HAdV-E4 and -B7 are used in the US military trainee population (Gaydos & Gaydos, 1995). Unlike HAdV-B3 and HAdV-B7, which have been commonly reported

globally in the civilian population, the HAdV-E4 infections are infrequently observed in the general public.

HAdV-E4 was one of the first identified HAdVs, isolated from a military outbreak in 1952 (Rowe et al., 1953). It is a major human respiratory pathogen and is highly contagious, resulting in frequent large outbreaks in certain populations, including the dense and overstressed military recruits on training bases (Barraza et al., 1999; Cersovsky et al., 1999). Inexplicitly, they appeared largely constrained to the U.S. military population (Barraza et al., 1999; Dudding et al., 1973; Gaydos & Gaydos, 1995; Kajon et al., 2007). More recently, HAdV-E4 outbreaks have been observed in the larger civilian community (Coleman et al., 2019; Li et al., 2018; Zhang et al., 2019). Due to the lack of reports of HAdV infections in the Hong Kong region, it is difficult to understand whether HAdV-E4 is circulating or re-emerging in the region. Therefore, it is useful of adenovirus epidemiological and clinical studies with more accurate and rapid HAdV typing globally. To describe the clinical HAdV genotypes epidemiologically to survey and characterize the HAdV-associated respiratory infections in Hong Kong; the present study designed using hexon, fiber, and penton base sequencing to identify the retrospective HAdV infections among 100 in-house patients from the Queen Mary Hospital in Hong Kong.

MATERIALS AND METHODS

This study was approved by the Institutional Ethics Committee of the Queen Mary Hospital to study the archived HAdV-positive specimens collected in 2014. All the

protocols were performed in accordance with the approved guidelines. Legal guardians of all minor patients have given signed informed consent to participate in this study. Sample collection and data records were de-identified and anonymized to protect the patients' privacy.

Retrospective Patient Sample Collection

From July 2014 to October 2014, there was a sudden increase in pediatric outpatients and inpatients with influenza-like symptoms in Queen Mary Hospital, Hong Kong. xxx nasopharyngeal swab specimens were collected from xxx patients, and respiratory vial pathogens were detected. Taq PCR Master Mix kits (TaKaRa Bio, Inc., Japan), QIAamp DNA Blood Mini Kit (QIAGEN, Inc., China), and AxyPrep PCR Clean-Up Kit (AxYgen Biosciences, Union City, CA) were used according to manufacturer instructions. DNA ladder, used as a reference, was based on DL10000 and DL 2000 marker (TAKARA Corp.). Total nucleic acid was extracted from respiratory specimens using a QIAamp DNA Blood Mini Kit (QIAGEN, Inc., China). Respiratory pathogens, including influenza A virus (infA), influenza B virus (infB), parainfluenza virus (PIV), respiratory syncytial virus (RSV), enterovirus (EV), human metapneumovirus (hMPV), bocavirus (BOV), rhinovirus (RHV), mycoplasma pneumonia (MP), and Chlamydia pneumonia (CP), were detected using Taq real-time PCR kit (TaKaRa Bio, Inc., Japan). Respiratory viruses other than the HAdV-positive cases were excluded from our study because these samples were required for study at Hong Kong's Queen Mary Hospital.

Human Adenovirus Culture and Isolation

The HAdV-positive throat swab specimens collected from 100 patients were inoculated onto A549 cells. Samples were cultured in a maintenance medium (Minimal Essential Medium containing 2% fetal bovine serum, 100 U/mL penicillin G, and 100 ug/mL streptomycin) at 37ºC with the atmosphere containing 5% (v/v) carbon dioxide and cytopathic effect (CPE). After monitoring for 5-7 days, cell cultures with no CPE observed were frozen and thawed for three rounds. A549 cells were passaged in again to check the CPE.

PCR Primers Design, Amplification, and Molecular Typing of Adenovirus Specimens

HAdV and other common respiratory pathogens were detected by the real-time PCR kits (TaKaRa Bio, Inc, Japan). The selected reference sequences, including hexon, penton base, and fiber, were extracted from the NCBI (https://www.ncbi.nlm.nih.gov) and aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to locate the conserved regions. Specific primers (Invitrogen, Inc., Guangzhou, China) targeting the three capsid genes were based on the conserved regions and were used in the PCR reaction (Table S1).

Whole Genome Sequencing and Phylogenetic Analysis

All of the HAdV-E4 isolates were sequenced in the Queen Mary Hospital using Illumina sequencing, for further characterization. The adaptors for forward and reverse fastQ files

were trimmed, and the reads were assembled using Geneious Prime version 2020.2.4 (https://www.geneious.com). Phylogenetic relationships among HAdV-E4 sequences were determined using the Molecular Evolutionary Genetics Analysis (MEGA) version X software (http://megasoftware.net). Neighbor-joining statistic methods with 1000 bootstraps replicates are applied using the maximum composite likelihood model and are in the units of the number of base substitutions per site.

Statistical Methods

HAdV genotype occurrence frequencies are measured using logistic regression to examine HAdV-positive case incidences by time period. Potential risk factors associated with upper respiratory tract illness (URTI) and coinfections were calculated for different categories, including gender, age group, and HAdV genotype. Odds ratio (OR) with the standard error (SE) of the log odds ratio and 95% confidence interval (CI) are being calculated using MEDCALC easy-to-use statistical software (https://www.medcalc.org/calc/odds_ratio.php). The test of significance for the standard normal deviate (z-value) is calculated as ln(OR)/SE{ln(OR)} with P-value being the area of the normal distribution outside of the z-value.

RESULTS

Clinical samples collected from 100 HAdV-positive hospitalized patients (93 pediatric, 1 adult, and 6 unknown) at Queen Mary Hospital from July to October 2014 were characterized with respect to genotype using DNA sequencing of the three capsid

proteins (hexon, penton base, and fiber). PCR sequencing were attempted on all 100 samples, in which 94% of the isolates were successfully genotyped. Of these, 4% of the genomes are recombinants, and 2% of the samples were unsequenceable. Figure 3.2.1 presents the distribution of HAdV genotype frequency by month.

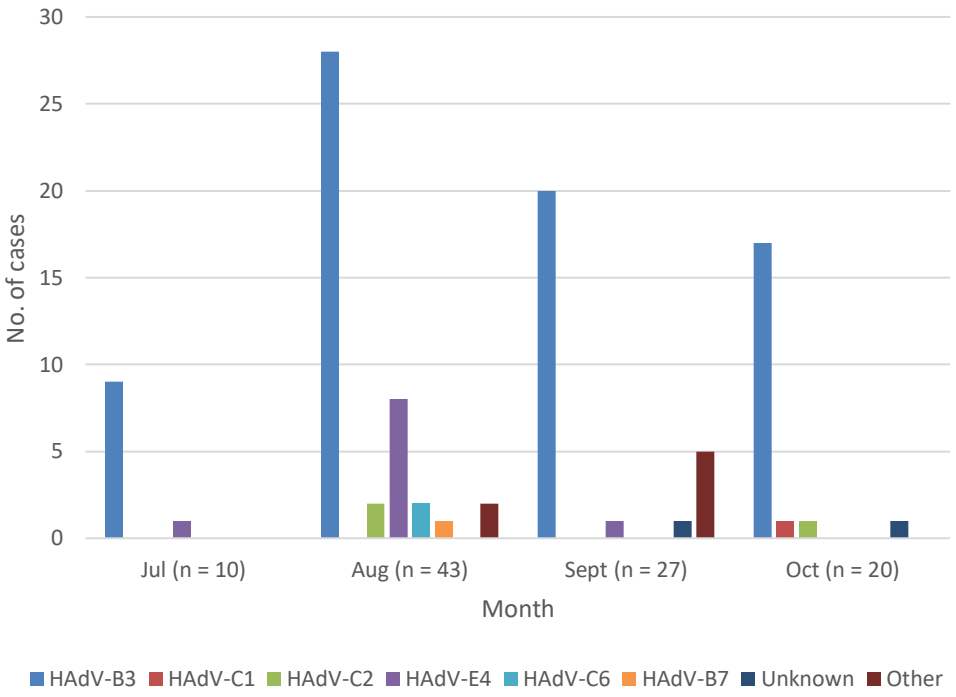


Figure 3.2.1. HAdV genotypes detected for HAdV-positive patients hospitalized in Hong Kong from July to October 2014. The “Unknown”group represents samples with respect to sequencing. “Other” represents samples with genome recombination in the capsid gene hexon, penton base, and fiber.

Among the 100 retrospective cases detected between July and October 2014, HAdV-B3 was most prevalent (74%), followed by HAdV-E4 (10%) and HAdV-C2 (6%) (Figure 3.2.1). Only 1% were found to be HAdV-B7 in this outbreak ,which differs with data from a Singapore clinical study of 15.7% (Coleman et al., 2019; Rajkumar et al., 2015). Four HAdV-positive cases infected by recombinants were subsequently identified to be novel genotypes (Table 3.2.1). The novel recombination includes isolates HK39 (penton base type 7, hexon type 55, fiber type 7 [P7H55F7]), HK42 (P1H1F3), HK61 (P1H5F2), and HK76 (P1H2F2).

The clinical presentation demonstrated that 74% of patients diagnosed with upper respiratory tract infection (URTI) and 17% of HAdV-positive cases were coinfected with enterovirus/rhinovirus (EV/RV, 14%), respiratory syncytial virus (RSV, 2%), and influenza C virus (IVC, 1%). Patients had a 93% hospitalized rate and a mean of 3.5 hospitalized days were indicated. Among the three of the patients with prolonged hospitalization (more than 7 seven consecutive days), one patient had pneumonia with secondary bacterial infection and passed away on day 36 and two of the immunocompromised patients had an average of 15 hospitalized days. Among the 93 pediatric patients (≤ 16 years), 79.6% were diagnosed with URTI which is corresponding with 73.1 % of HAdV-B3 and 9.7 % of HAdV-E4 associated infection. No gender differences are found for the total and pediatric data (Table 3.2.1).

Table 3.2.1. Clinical and HAdV genotype information for HAdV-positive patients hospitalized in Hong Kong from July to October 20.

Variables	Total n = 100 (%)	Pediatric[a] n = 93 (%)
Gender		
Female	49 (49.0)	45 (48.4)
Male	50 (50.0)	48 (51.6)
Clinical Presentation		
URTI	74 (74.0)	74 (79.6)
Gastrointestinal	5 (5.0)	5 (5.4)
Ocular	-	-
Coinfection	17 (17.0)	10 (10.7)
Immunocompromised[b]	2 (2.0)	2 (2.2)
Disease severity		
Hospitalized number	93 (93.0)	92 (98.9)
Average hospitalized days (%)	3.5	3.5 (3.8)
Prolonged hospitalization[c]	3 (3.0)	3 (3.2)
HAdV Genotype		
B3	74 (74.0)	68 (73.1)
C1	1 (1.0)	1 (1.1)
C2	6 (6.0)	6 (6.5)
E4	10 (10.0)	9 (9.7)
C6	2 (2.0)	2 (2.2)
B7	1 (1.0)	1 (1.1)
Recombination[d]	4 (4.0)	4 (4.3)
Undetectable[e]	2 (2.0)	2 (2.2)

Abbreviations: HAdV, human adenovirus; URTI, upper respiratory tract infection
[a] Pediatric (≤ 16 years)
[b] Immunocompromised adenoviremia transplant recipient
[c] Hospitalized for more than 7 consecutive days
[d] recombination in the capsid gene hexon, fiber, and penton base
[e] Genotyping failed

The risk factors for pediatric patients with URTI and coinfections with other pathogens were assessed between July to October 2014. Pediatric patients infected with HAdV-C2 (odds ratio [OR], 3.46; 95% confidence interval [CI], 0.18-65.05), HAdV-E4 (OR, 2.07; 95% CI, 0.24-17.99), or unknown recombinants (OR, 2.39; 95% CI, 0.12-44.08) were found to have an increased risk for URTI (Table 3.2.2). Pediatric patients have an

1407 increased risk of coinfection with other pathogens when diagnosed with HAdV-C2 (OR,
1408 2.33; 95% CI, 0.04-20.0) and HAdV-E4 (OR, 2.33; 95% CI, 0.51-10.67).
1409

Table 3.2.2. Risk factor association for pediatric patients with upper respiratory tract infection resulted from adenovirus infections in Hong Kong, July to October 2014.

Risk Factors	Total No. (%)	URTI* n = 73 (%)	Unadjusted OR (95% CI)	z-value‡ (P-value)
Gender				
Female	45 (48.4)	39 (53.4)	Reference	--
Male	48 (51.6)	34 (46.6)	0.37 (0.13-1.08)	1.82 (0.07)
Age Group				
<1 year	7 (7.5)	4 (5.5)	Reference	--
1-5 years	53 (57.0)	42 (57.5)	3.5 (0.66-18.43)	1.48 (0.14)
6-10 years	29 (31.2)	23 (31.5)	2.88 (0.50–16.48)	1.19 (0.24)
11-16 years	4 (4.3)	4 (5.5)	7 (0.27-178.48)	1.18 (0.24)
HAdV Genotype				
C1	1 (1.1)	0 (0)	--	--
C2	6 (6.5)	6 (8.2)	3.46 (0.18-65.05)	0.83 (0.41)
C6	2 (2.2)	1 (1.4)	0.26 (0.02-4.41)	0.93 (0.35)
B3	68 (73.1)	54 (74.0)	Reference	--
B7	1 (1.1)	0 (0)	--	--
E4	9 (9.7)	8 (11.0)	2.07 (0.24-17.99)	0.66 (0.51)
Recombination §	4 (4.3)	4 (5.5)	2.39 (0.12-47.08)	0.58 (0.57)
Undetectable‖	2 (2.2)	1 (1.4)	0.26 (0.02-4.41)	0.93 (0.35)

Abbreviations: HAdV, human adenovirus

* Upper respiratory tract infections

† A standard normal deviate (z-value) is calculated as ln(OR)/SE{ln(OR)}

§ Recombination in the capsid gene hexon, fiber, and penton base

‖ Genotyping failed

Table 3.2.3. Risk factor association for pediatric patients with coinfection resulted from adenovirus infections in Hong Kong, July to October 2014.

Risk Factors	Total No. (%)	Coinfection* n = 17 (%)	OR (95% CI)	z-value† (P-value)
Gender				
Female	45 (48.4)	9 (52.9)	Reference	--
Male	48 (51.6)	8 (47.1)	0.80 (0.28–2.29)	0.42 (0.68)
Age Group				
<1 year	7 (7.5)	2 (11.8)	Reference	--
1-5 years	53 (57.0)	10 (58.8)	0.58 (0.10-3.44)	0.60 (0.55)
6-10 years	29 (31.2)	4 (23.5)	0.40 (0.06-2.81)	0.92 (0.36)
11-16 years	4 (4.3)	0 (0)	--	--
HAdV Genotype				
C1	1 (1.1)	0 (0)	--	--
C2	6 (6.5)	1 (5.9)	0.93 (0.10-8.73)	0.06 (0.95)
C6	2 (2.2)	0 (0)	--	--
B3	68 (73.1)	12 (70.6)	Reference	--
B7	1 (1.1)	0 (0)	--	--
E4	9 (9.7)	3 (17.6)	2.33 (0.51-10.67)	1.09 (0.27)
Recombination §	4 (4.3)	0 (0)	--	--
Undetectable‖	2 (2.2)	1 (5.9)	4.67 (0.27-79.96)	1.06 (0.29)

Abbreviations: HAdV, human adenovirus

* Coinfection with other pathogens (e.g. RSV)

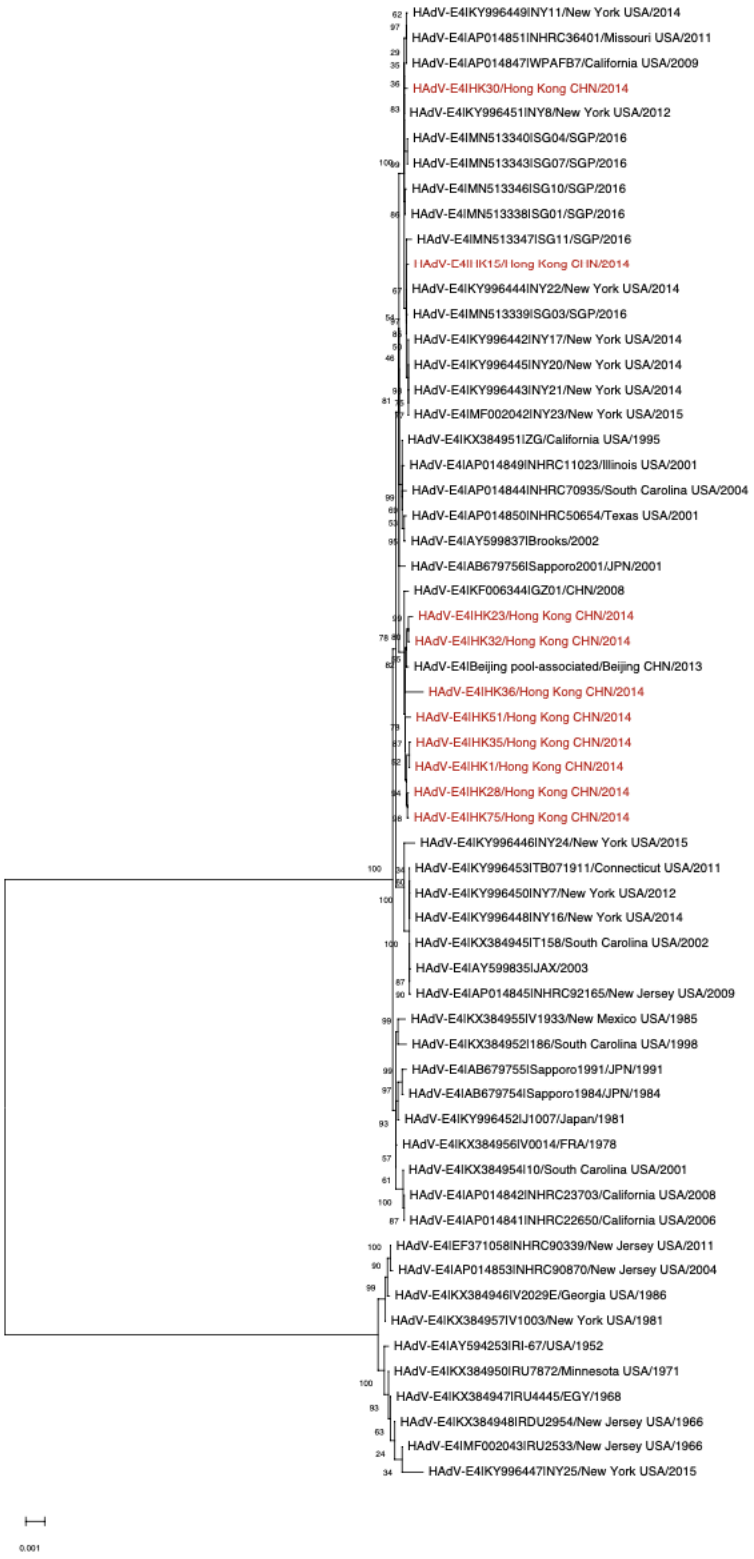
† A standard normal deviate (z-value) is calculated as ln(OR)/SE{ln(OR)}

§ Recombination in the capsid gene hexon, fiber, and penton base

‖ Genotyping failed

In the summer of 2014, cases of HAdV-E4 were observed in the Hong Kong population. All ten HAdV-E4 strains were sequenced and annotated. The whole genome evolutionary

relationships of HAdV-E4 isolates were analyzed using a neighbor-joining phylogenetic method. The optimal tree is 0.053 for the sum of branch length. The phylogenetic tree naturally diverges into two separate clusters of prototype-like isolates and host-adapted clade as described in the previous study (unpublished). These Hong Kong isolates belong to the host-adapted subclade and are closely related to the reported HAdV-E4 strains circulating in the civilian populations. HK15 and HK30 are sisters to a group of strains isolated in New York and Singapore in the recent decade. The other eight Hong Kong isolates share the most common ancestor with strains isolated in China, specifically a Beijing pool-associated strain which caused an outbreak that infected 55 patients at the University of Beijing (Li et al., 2018).


HAdV-E4|KY996449|NY11/New York USA/2014
HAdV-E4|AP014851|NHRC36401/Missouri USA/2011
HAdV-E4|AP014847|WPAFB7/California USA/2009
HAdV-E4|HK30/Hong Kong CHN/2014
HAdV-E4|KY996451|NY8/New York USA/2012
HAdV-E4|MN513340|SG04/SGP/2016
HAdV-E4|MN513343|SG07/SGP/2016
HAdV-E4|MN513346|SG10/SGP/2016
HAdV-E4|MN513338|SG01/SGP/2016
HAdV-E4|MN513347|SG11/SGP/2016
HAdV-E4|HK15/Hong Kong CHN/2014
HAdV-E4|KY996444|NY22/New York USA/2014
HAdV-E4|MN513339|SG03/SGP/2016
HAdV-E4|KY996442|NY17/New York USA/2014
HAdV-E4|KY996445|NY20/New York USA/2014
HAdV-E4|KY996443|NY21/New York USA/2014
HAdV-E4|MF002042|NY23/New York USA/2015
HAdV-E4|KX384951|ZG/California USA/1995
HAdV-E4|AP014849|NHRC11023/Illinois USA/2001
HAdV-E4|AP014844|NHRC70935/South Carolina USA/2004
HAdV-E4|AP014850|NHRC50654/Texas USA/2001
HAdV-E4|AY599837|Brooks/2002
HAdV-E4|AB679756|Sapporo2001/JPN/2001
HAdV-E4|KF006344|GZ01/CHN/2008
HAdV-E4|HK23/Hong Kong CHN/2014
HAdV-E4|HK32/Hong Kong CHN/2014
HAdV-E4|Beijing pool-associated/Beijing CHN/2013
HAdV-E4|HK36/Hong Kong CHN/2014
HAdV-E4|HK51/Hong Kong CHN/2014
HAdV-E4|HK35/Hong Kong CHN/2014
HAdV-E4|HK1/Hong Kong CHN/2014
HAdV-E4|HK28/Hong Kong CHN/2014
HAdV-E4|HK75/Hong Kong CHN/2014
HAdV-E4|KY996446|NY24/New York USA/2015
HAdV-E4|KY996453|TB071911/Connecticut USA/2011
HAdV-E4|KY996450|NY7/New York USA/2012
HAdV-E4|KY996448|NY16/New York USA/2014
HAdV-E4|KX384945|T158/South Carolina USA/2002
HAdV-E4|AY599835|JAX/2003
HAdV-E4|AP014845|NHRC92165/New Jersey USA/2009
HAdV-E4|KX384955|V1933/New Mexico USA/1985
HAdV-E4|KX384952|186/South Carolina USA/1998
HAdV-E4|AB679755|Sapporo1991/JPN/1991
HAdV-E4|AB679754|Sapporo1984/JPN/1984
HAdV-E4|KY996452|J1007/Japan/1981
HAdV-E4|KX384956|V0014/FRA/1978
HAdV-E4|KX384954|I10/South Carolina USA/2001
HAdV-E4|AP014842|NHRC23703/California USA/2008
HAdV-E4|AP014841|NHRC22650/California USA/2006
HAdV-E4|EF371058|NHRC90339/New Jersey USA/2011
HAdV-E4|AP014853|NHRC90870/New Jersey USA/2004
HAdV-E4|KX384946|V2029E/Georgia USA/1986
HAdV-E4|KX384957|V1003/New York USA/1981
HAdV-E4|AY594253|RI-67/USA/1952
HAdV-E4|KX384950|RU7872/Minnesota USA/1971
HAdV-E4|KX384947|RU4445/EGY/1968
HAdV-E4|KX384948|RDU2954/New Jersey USA/1966
HAdV-E4|MF002043|RU2533/New Jersey USA/1966
HAdV-E4|KY996447|NY25/New York USA/2015
0.001

1430

Figure 3.2.2. Evolutionary relationships of the HAdV-E4 species clade, using whole genome sequence analysis. The neighbor-joining method with the maximum composite likelihood model was used through MEGA X (http://megasoftware.net). Color denote the ten HAdV-E4 strains isolated and sequenced. Bootstrap values are displayed at all nodes for 1000 replicates, with 80% noted as robust. The scale bar represents the number of base substitutions per site.

DISCUSSION

A 10% HAdV-E4 infection rate is identified in pediatric patients during this outbreak. It is unusual to identify a high number of HAdV-E4 infections in the civilian population. Again, the number of HAdV-E4-associated ARD cases is unusual, as, in the past, HAdV-E4 infections were found in the adult military basic trainee population. Within civilian populations, infections were sporadic, few, and limited in scope (Kajon et al., 2018). However, our results indicate an increased susceptibility in the presumably immune-naïve civilian populations may be forthcoming. Previous quantitative study in Singapore showed that 15.2% of 533 AdV clinical samples were identified as HAdV-E4 (Coleman et al., 2019) which is consistent with our results. In addition, the first large-scale outbreak of HAdV-E4 in a non-US military population isolated from the Beijing University swimming pool was characterized recently (Li et al., 2018). This again proves that HAdV-E4 is not limited to the military trainees, it has adapted and evolved to circulate in the civilian population.

The HAdV-E4 isolates were sequenced for further analysis. Whole-genome evolutionary tree revealed eight of the HAdV-E4 are closely related to the Beijing swimming pool-related strain isolated in 2013. Two of the isolates share a common ancestor with the strains isolated from Singapore and the United States. It must be mentioned that previous study suggests HAdV-E4 isolates recombine at the inverted terminal repeats region that contains the viral replication motifs (Dehghan et al., 2013). This may have contributed to the host adaptations outside of the U.S. military populations. The recent outbreaks in the larger civilian communities globally (Coleman et al., 2019; Kajon et al., 2018; Li et al., 2018; Zhang et al., 2019) suggests that HAdV-E4 infections are not limited to the confined military population or any geographical regions and a broader biosurvilliance should be considered for HAdV-E4.

Eight out of ten HAdV-E4 pediatric patients (OR, 2.07; 95% CI, 0.24-17.99) developed URTI (pharyngitis) which is consistent with the common symptoms of HAdV-E4 infection (Barraza et al., 1999; Cersovsky et al., 1999; Purkayastha et al., 2005). The diagnostics for one of the HAdV-E4 patient was not available. One of the HAdV-E4 strain was isolated from a transplant recipient that had adenoviremia with high fever and were also coinfected with RSV. Patients infected with HAdV-E4 (OR, 2.33; 95% CI, 0.51-10.67) is highly susceptible to have coinfections with other pathogens. Due to the lack of prior genotype surveillance in Hong Kong, it is uncertain that 33.33% (3/9) of the coinfection rate in HAdV-E4 is a new or pre-existed condition.

This molecular epidemiological study showed that the most prevalent genotyping identified is HAdV-B3 (74%). HAdV-B3 is commonly associated with ARD, which is consistent with high number of URTI cases identified. Studies have shown HAdV-B3 is the most common type implicated with HAdV infections in both the children and adults globally (Chang et al., 2008; Gray et al., 2007; Schmitz et al., 1983; Wang et al., 2013). HAdV-B3 (17.65%, 12/68) is also the second most prevalent genotype with coinfection in this study which is consistent with the previous adenoviral coinfection study in Taiwan (19%, 24/126) (Wang et al., 2013). One of the patients identified with HAdV-B3 induced pneumonia contracted a secondary bacterial infection and passed away on the 36 days.

Surprisingly, only 1% of HAdV-E7 was isolated during the clinical study which is inconsistent with the Singapore (15.4%), Korea (41%), and United States (59%) comprehensive study (Coleman et al., 2019; Hong et al., 2001; Scott et al., 2016). Four novel recombinants were identified and didn't match any of the known genotypes (103 types) published (http://hadvwg.gmu.edu). Patients with novel recombinants had URTI (3/4) and diarrhea (1/4) symptoms. The data contributes to a clear understanding of HAdV outbreak during the summer of 2014. It provides new insights on the emerged HAdV-E4 infections in the general population. However, one disadvantage of this study focuses on hospitalized patients. HAdV patients with less severe symptoms and were not hospitalized could not be account for.

In summary, an increase of HAdV-E4 infections are observed in the general population. This may post threats to the public health. Furthermore, patients with HadV-B3 and -E4 are likely to result in URTI and are more prone to secondary coinfections. It is important to highlight that a high-incidence of HAdV-E4 infections are circulating globally. Consequently, we strongly suggest the health community to have routine genotyping surveillance over the HAdV prevalence in the general population. And vaccine adaptations and antiviral therapies for the specific types in pediatric patients should be considered.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

References

Barraza, E. M., Ludwig, S. L., Gaydos, J. C., & Brundage, J. F. (1999). Reemergence of adenovirus type 4 acute respiratory disease in military trainees: Report of an outbreak during a lapse in vaccination. *The Journal of Infectious Diseases, 179 6,* 1531–1533.

Carballal, G., Videla, C. M., Espinosa, M. A., Savy, V., Uez, O., Sequeira, M. D., Knez, V., Requeijo, P. V., Posse, C. R., & Miceli, I. (2001). Multicentered study of viral

acute lower respiratory infections in children from four cities of Argentina, 1993-1994. *Journal of Medical Virology*, *64*(2), 167–174. https://doi.org/10.1002/jmv.1032

Cersovsky, S., LN, B., FD, M., SA, K., CS, P., Sanchez Jr, J., & WT, B. (1999). A prospective study of acute respiratory diseases (ARD) in a cohort of US Army basic trainees, October-November 1998, Fort Jackson, South Carolina. *Medical Surveillance Monthly Report*, *5*, 10, 12.

Chang, S.-Y., Lee, C.-N., Lin, P.-H., Huang, H.-H., Chang, L.-Y., Ko, W., Chang, S.-F., Lee, P.-I., Huang, L.-M., & Kao, C.-L. (2008). A community-derived outbreak of adenovirus type 3 in children in Taiwan between 2004 and 2005. *Journal of Medical Virology*, *80*(1), 102–112. https://doi.org/10.1002/jmv.21045

Coleman, K. K., Wong, C. C., Jayakumar, J., Nguyen, T. T., Wong, A. W. L., Yadana, S., Thoon, K. C., Chan, K. P., Low, J. G., Kalimuddin, S., Dehghan, S., Kang, J., Shamsaddini, A., Seto, D., Su, Y. C. F., & Gray, G. C. (2019). Adenoviral Infections in Singapore: Should New Antiviral Therapies and Vaccines Be Adopted? *The Journal of Infectious Diseases*, *jiz489*. https://doi.org/10.1093/infdis/jiz489

Dehghan, S., Seto, J., Liu, E. B., Walsh, M. P., Dyer, D. W., Chodosh, J., & Seto, D. (2013). Computational analysis of four human adenovirus type 4 genomes reveals molecular evolution through two interspecies recombination events. *Virology*, *443*(2), 197–207. https://doi.org/10.1016/j.virol.2013.05.014

Dudding, B. A., Top, F. H. J., Winter, P. E., Buescher, E. L., Lamson, T. H., & Leibovitz, A. (1973). Acute respiratory disease in military trainees: The adenovirus surveillance program, 1966-1971. *American Journal of Epidemiology*, *97*(3), 187–198. https://doi.org/10.1093/oxfordjournals.aje.a121499

Garnett, C. T., Talekar, G., Mahr, J. A., Huang, W., Zhang, Y., Ornelles, D. A., & Gooding, L. R. (2009). Latent species C adenoviruses in human tonsil tissues. *Journal of Virology*, *83*(6), 2417–2428. PubMed. https://doi.org/10.1128/JVI.02392-08

Gaydos, C. A., & Gaydos, J. C. (1995). Adenovirus vaccines in the U.S. military. *Military Medicine*, *160*(6), 300–304.

Gerber, S. I., Erdman, D. D., Pur, S. L., Diaz, P. S., Segreti, J., Kajon, A. E., Belkengren, R. P., & Jones, R. C. (2001). Outbreak of adenovirus genome type 7d2 infection in a pediatric chronic-care facility and tertiary-care hospital. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *32*(5), 694–700. https://doi.org/10.1086/319210

Gray, G. C., McCarthy, T., Lebeck, M. G., Schnurr, D. P., Russell, K. L., Kajon, A. E., Landry, M. L., Leland, D. S., Storch, G. A., Ginocchio, C. C., Robinson, C. C., Demmler, G. J., Saubolle, M. A., Kehl, S. C., Selvarangan, R., Miller, M. B., Chappell, J. D., Zerr, D. M., Kiska, D. L., … Erdman, D. D. (2007). Genotype prevalence and risk factors for severe clinical adenovirus infection, United States 2004-2006. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *45*(9), 1120–1131. https://doi.org/10.1086/522188

Harrach, B., Benkö, M., Both, G. W., Brown, M., Davison, A. J., Echavarria, M., Hess, M., Jones, M. S., Kajon, A., Lehmkuhl, H. D., Mautner, V., Mittal, S. K., & Wadell, G. (2011). *Virus Taxonomy: Classification and Nomenclature of Viruses*. 125–141.

Hong, J.-Y., Lee, H.-J., Piedra, P. A., Choi, E.-H., Park, K.-H., Koh, Y.-Y., & Kim, W.-S. (2001). Lower Respiratory Tract Infections due to Adenovirus in Hospitalized Korean Children: Epidemiology, Clinical Features, and Prognosis. *Clinical Infectious Diseases*, *32*(10), 1423–1429. https://doi.org/10.1086/320146

Kajon, A. E., Lamson, D. M., Bair, C. R., Lu, X., Landry, M. L., Menegus, M., Erdman, D. D., & St. George, K. (2018). Adenovirus Type 4 Respiratory Infections among Civilian Adults, Northeastern United States, 2011–20151. *Emerging Infectious Diseases*, *24*(2), 201–209. https://doi.org/10.3201/eid2402.171407

Kajon, A. E., Moseley, J. M., Metzgar, D., Huong, H.-S., Wadleigh, A., Ryan, M. A. K., & Russell, K. L. (2007). Major Articles-Molecular Epidemiology of Adenovirus Type 4 Infections in US Military Recruits in the Postvaccination Era (1997–2003). *The Journal of Infectious Diseases*, *196*(1), 67–75. https://doi.org/10.1086/518442

Larrañaga, C., Kajon, A., Villagra, E., & Avendaño, L. F. (2000). Adenovirus surveillance on children hospitalized for acute lower respiratory infections in Chile (1988-1996). *Journal of Medical Virology*, *60*(3), 342–346.

Li, J., Lu, X., Sun, Y., Lin, C., Li, F., Yang, Y., Liang, Z., Jia, L., Chen, L., Jiang, B., & Wang, Q. (2018). A swimming pool-associated outbreak of pharyngoconjunctival

fever caused by human adenovirus type 4 in Beijing, China. *International Journal of Infectious Diseases*, *75*, 89–91. https://doi.org/10.1016/j.ijid.2018.08.009

Lion, T. (2014). Adenovirus infections in immunocompetent and immunocompromised patients. *Clinical Microbiology Reviews*, *27*(3), 441–462. https://doi.org/10.1128/CMR.00116-13

Purkayastha, A., Ditty, S. E., Su, J., McGraw, J., Hadfield, T. L., Tibbetts, C., & Seto, D. (2005). Genomic and Bioinformatics Analysis of HAdV-4, a Human Adenovirus Causing Acute Respiratory Disease: Implications for Gene Therapy and Vaccine Vector Development. *Journal of Virology*, *79*(4), 2559. https://doi.org/10.1128/JVI.79.4.2559-2572.2005

Rajkumar, V., Chiang, C. S. M., Low, J. M., Cui, L., Lin, R. T. P., Tee, N. W. S., Maiwald, M., Chong, C. Y., Thoon, K. C., & Tan, N. W. H. (2015). Risk Factors for Severe Adenovirus Infection in Children during an Outbreak in Singapore. *Annals of the Academy of Medicine, Singapore*, *44*(2), 50–59.

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., & Ward, T. G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *84*(3), 570–573. https://doi.org/10.3181/00379727-84-20714

Schmitz, H., Wigand, R., & Heinrich, W. (1983). Worldwide epidemiology of human adenovirus infections. *American Journal of Epidemiology*, *117*(4), 455–466. https://doi.org/10.1093/oxfordjournals.aje.a113563

Scott, M. K., Chommanard, C., Lu, X., Appelgate, D., Grenz, L., Schneider, E., Gerber, S. I., Erdman, D. D., & Thomas, A. (2016). Human Adenovirus Associated with Severe Respiratory Infection, Oregon, USA, 2013-2014. *Emerging Infectious Diseases*, *22*(6), 1044–1051. PubMed. https://doi.org/10.3201/eid2206.151898

Shayakhmetov, D. M., Gaggar, A., Ni, S., Li, Z.-Y., & Lieber, A. (2005). Adenovirus binding to blood factors results in liver cell infection and hepatotoxicity. *Journal of Virology*, *79*(12), 7478–7491. PubMed. https://doi.org/10.1128/JVI.79.12.7478-7491.2005

Videla, C., Carballal, G., & Kajon, A. (1999). Genomic analysis of adenovirus isolated from Argentinian children with acute lower respiratory infections. *Journal of Clinical Virology : The Official Publication of the Pan American Society for Clinical Virology*, *14*(1), 67–71. https://doi.org/10.1016/s1386-6532(99)00033-5

Wang, S.-L., Chi, C.-Y., Kuo, P.-H., Tsai, H.-P., Wang, S.-M., Liu, C.-C., Su, I.-J., & Wang, J.-R. (2013). High-Incidence of Human Adenoviral Co-Infections in Taiwan. *PLOS ONE*, *8*(9), e75208. https://doi.org/10.1371/journal.pone.0075208

Wood, D. J. (1988). Adenovirus gastroenteritis. *British Medical Journal (Clinical Research Ed.)*, *296*(6617), 229–230. PubMed. https://doi.org/10.1136/bmj.296.6617.229-a

Yao, L., Wang, C., Wei, T., Wang, H., Ma, F., & Zheng, L. (2019). Human adenovirus among hospitalized children with respiratory tract infections in Beijing, China, 2017–2018. *Virology Journal*, *16*(1), 78. https://doi.org/10.1186/s12985-019-1185-x

Zhang, J., Kang, J., Dehghan, S., Sridhar, S., Lau, S. K. P., Ou, J., Woo, P. C. Y., Zhang, Q., & Seto, D. (2019). A Survey of Recent Adenoviral Respiratory Pathogens in Hong Kong Reveals Emergent and Recombinant Human Adenovirus Type 4 (HAdV-E4) Circulating in Civilian Populations. *Viruses*, *11*(2). https://doi.org/10.3390/v11020129

Supplemental Material

Table S1. Universal primers for detecting, typing, and sequencing HAdV. Forward and backward primer sequences, genome positions, and the predicted PCR product (34 cycles denatured at 94^{o}C for 30s, annealed at 52^{o}C for 30s, and elongated at 72^{o}C for 100s) sizes were deducted for the three major capsid genes.

1640

Gene	Position	Length (bp)	Primer	Primer sequence	Position[a]	PCR product
Penton base	13904 - 15538	1635	Penton-F	5'-CTATCAGAACGACCACAGCAACTT-3'	14152 - 14175	1253 bp
			Penton-R	5'-TCCCGTGATCTGTGAGAGCRG-3'	15384 - 15404	
Hexon	18422 - 21256	2835	HVR-F	5'-CAGGATGCTTCGGAGTACCTGAG-3'	18473 - 18495	1685 bp
			HVR-R	5'-TTTCTGAAGTTCCACTCGTAGGTGTA-3'	20132 - 20157	
Fiber	31301 - 32260	960 (1278[b])	Fiber-F	5'-CCCTCTTCCCAACTCTGGTA-3'	31180 - 31199	1153 bp (1519 bp[b])
			Fiber-R	5'-GGGGAGGCAAAATAACTACTCG-3'	32311 - 32332	
		1746[c]	Fiber-CR	5'-GAGGTGGCAGGTTGAATACTAG-3'	32311 - 32332	2027 bp[c]

[a] Positions are based on the reference genome HAdV-B3 (GenBank acc. no. DQ099432).
[b] Reference based on the length of HAdV-E4 fiber gene.
[c] Reference based on the length of the PCR product of HAdV-C5 fiber gene.

1641

1642

1643

1644
1645

1646 Part 3. genomic analysis of a human adenovirus type 4 strain associated with an outbreak

1647

1648

Genomic Analysis of a Human Adenovirus Type 4 Strain Associated with an Outbreak

June Kang[1a], Daniel Negrón[1a,2], Jie Li[3], Quanyi Wang[3], Qiwei Zhang[4,5], Donald Seto[1*]

[1]Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

[2]Noblis, Inc., 2002 Edmund Halley Dr, Reston, VA 20191

[3]Beijing Centers for Disease Prevention and Control, No.16 Hepingli Middle Road, Beijing 100013, China

[4]Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, Guangdong 510515, China

[5]Guangzhou Key Laboratory of Virology, Institute of Medical Microbiology, Jinan University, Guangzhou, Guangdong 510632, China

[a]These authors contributed equally.

*Corresponding author: Donald Seto. Email: dseto@gmu.edu. Address: 10900 University Blvd, Manassas, VA 20110.

Keywords: Adenovirus; HAdV-E4, Infectious disease; Human pathogen; Taxonomy;

Short running title:

Abstract

Human adenovirus type 4 (HAdV-E4) primarily affected military populations historically, specifically U.S. military basic training cohorts. A recent large civilian outbreak occurred in Beijing, China, in a university environment. Genome sequence analysis showed this strain to be similar to recent circulating strains, bearing a NF-1 motif for optimal viral replication, unlike the prototype strain, indicating possible host adaptation of this virus.

Dear Editor,

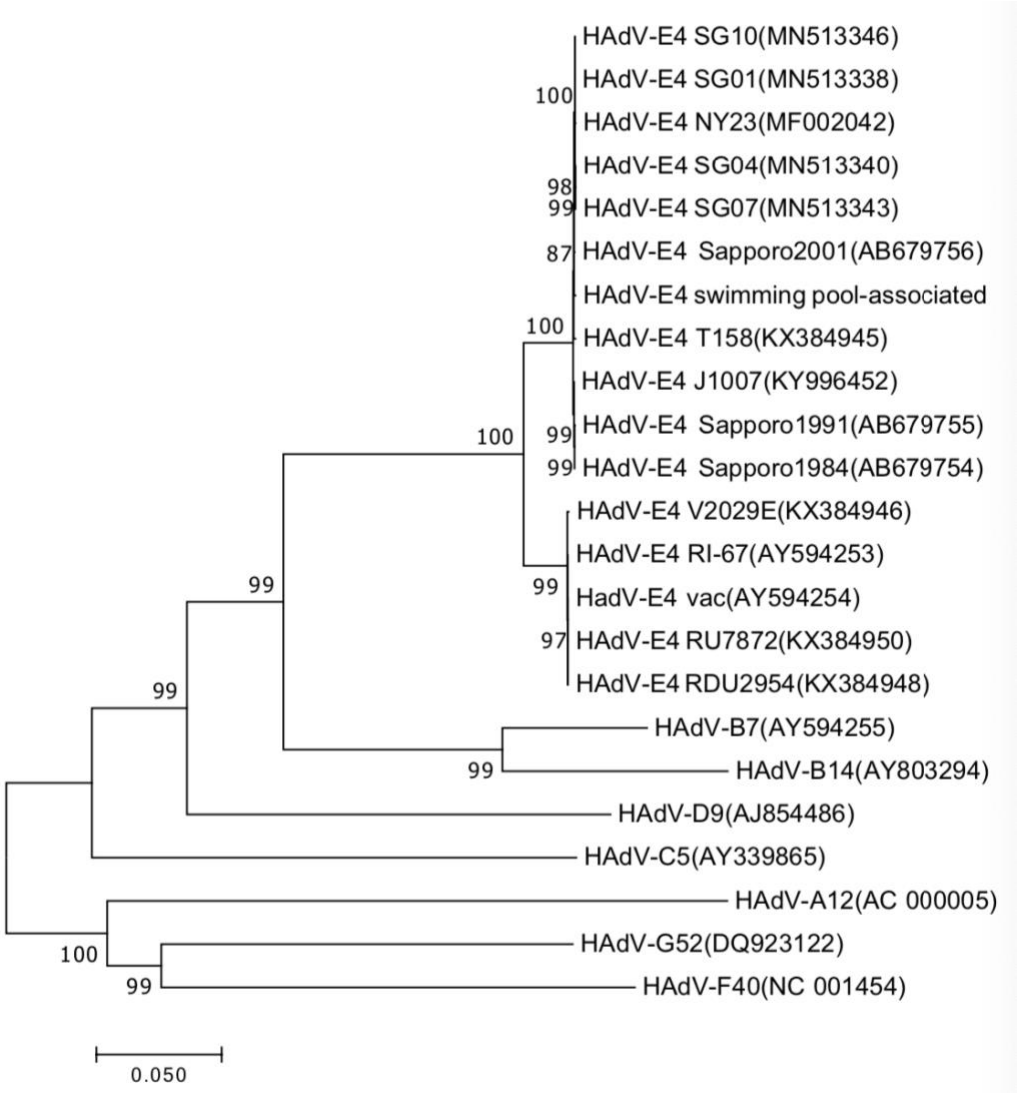
Human adenovirus E4 (HAdV-E4) has been recognized as a prevalent agent causing acute respiratory disease (ARD) and conjunctivitis with severity ranging from ‘pink eye’ to epidemic keratoconjunctivitis (EKC) (Rowe et al. 1953; Hilleman and Werner 1954; Aoki et al. 1982; D’Angelo et al. 1979; Tsuzuki-Wang et al. 1997). Historically, HAdV-E4 appeared limited to U.S. military training populations. The HAdV-E4 prototype was first identified in a military outbreak at the U.S. Army training installation Fort Leonard Wood (MO) in 1952 (Rowe et al. 1953). It continued to cause respiratory health issues in trainees until 1965, when the military eventually developed a vaccine to control this pathogen (Top et al. 1971). The vaccine lowered the HAdV-E4 morbidity rate from 98% to essentially 0% before it was discontinued in 1996 (Gaydos and Gaydos 1995; Dudding et al. 1973; Kajon et al. 2007; Blasiole et al. 2004). HAdV-E4 reemerged after the vaccine supplies were exhausted in 1999, but was confined to the military recruits, with only a few strains appearing in the general public (Cersovsky et al. 1999; Barraza et al. 1999; Kajon et al. 2007). In 2013, the Haidian District Centers for Disease Control and Prevention (HCDC) identified an outbreak of pharyngoconjunctival fever (PCF) at the University of Beijing in China. HAdV-E4 was isolated from 55 students and staff members, with symptoms at onset (Li et al. 2018). It was later confirmed that fifty patients swam in the same swimming pool two weeks before this outbreak (Li et al. 2018).

With this swimming pool-associated outbreak and other reports, it appears that HAdV-E4 infections has shifted toward dense civilian populations (Coleman et al. 2019; Zhang et al. 2019). These include reports from Singapore (Coleman et al. 2019), Hong Kong (Zhang et al., 2019), and New York (Kajon et al., 2018). To analyze this Beijing isolate, a total of 52 HAdV-E4 nucleotide sequences were collected from NCBI and collaborators. The inverted terminal repeats (ITR) were extracted and inspected using BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Additionally, multiple sequence alignments for the ITR and whole genomes were generated using Clustal Omega with default parameters (https://www.ebi.ac.uk/Tools/msa/clustalo/). Phylogenetic analysis for the whole genomes and the ITRs were performed in Mega, version 7.0.26 (https://www.megasoftware.net) to generate maximum parsimony and neighbor-joining trees with 1000 bootstraps. Furthermore, an *in silico* restriction enzyme analysis using six enzymes (EcoRV, BamHI, SmaI, DraI, XhoI, and EcoRI) were generated in python (Cock et al., 2009) and plotted using various packages in RStudio version 3.5.1 (R Core Team, 2018).[2]

The swimming pool-associated strain is identified as the HAdV-E4a1 variant, based on the reference restriction patterns in the literature (Kajon et al., 2018). Variation maps facilitated comparison of substitution and indel mutations across samples which describe the changes over evolutionary time. IQ-TREE generated a maximum likelihood tree based on MAFFT alignment of selected HAdV-E4a1 strains using default parameters

[2] R script packages: dplyr (Wickham et al. 2018), tidyr (Wickham and Henry 2019), tidyselect (Henry & Wickham, 2018), ggplot2 (Wickham 2016), readr(Wickham et al. 2018), tidyverse (Wickham 2017).

with the indication of HKY+F as the substitution model based on the Bayesian information criterion. BEAST version 1.10.4 (Suchard et al. 2018; http://beast.community) was used to perform Bayesian phylogenetic analysis on the same alignment, with date tip sampling from a uniform distribution with an uncertainty of 10 years. The chain length was set to 1×10^8 to ensure an effective sampling size (ESS) over 300. A tree prior, of constant size with an uncorrelated relaxed clock, was set to estimate the most recent common ancestor (tMRCA). A chronogram was generated using Bayesian analysis for HAdV-E4a1 types with 95% confidence intervals of the node age (data not shown). The chronogram and neighbor joining tree revealed that the swimming pool-associated strain is more similar to the Hong Kong strains (unpublished) than the New York strains.


HAdV-E4 SG10(MN513346)
HAdV-E4 SG01(MN513338)
HAdV-E4 NY23(MF002042)
HAdV-E4 SG04(MN513340)
HAdV-E4 SG07(MN513343)
HAdV-E4 Sapporo2001(AB679756)
HAdV-E4 swimming pool-associated
HAdV-E4 T158(KX384945)
HAdV-E4 J1007(KY996452)
HAdV-E4 Sapporo1991(AB679755)
HAdV-E4 Sapporo1984(AB679754)
HAdV-E4 V2029E(KX384946)
HAdV-E4 RI-67(AY594253)
HadV-E4 vac(AY594254)
HAdV-E4 RU7872(KX384950)
HAdV-E4 RDU2954(KX384948)
HAdV-B7(AY594255)
HAdV-B14(AY803294)
HAdV-D9(AJ854486)
HAdV-C5(AY339865)
HAdV-A12(AC 000005)
HAdV-G52(DQ923122)
HAdV-F40(NC 001454)
100
98
99
87
100
100
99
99
99
99
97
99
99
100
99
0.050

1734

Figure 3.3.1. Neighbor-joining phylogenetic relationships of HAdV-E4 isolates. Whole genome sequences from selected HAdV-E4 isolates were aligned using Clustal Omega with default parameters (https://www.ebi.ac.uk/Tools/msa/clustalo/) and analyzed using the neighbor-joining algorithm in MEGA version 7.0.26 (https://www.megasoftware.net). The iterations were run with default parameters and 1000 bootstrap replicates. HAdV-E4 species are naturally divergent into two separate subclades.

A divergence between the prototype-like isolates and the host-adapted clade was shown in both the phylogenetic (Figure 3.3.1) and restriction enzyme analyses. The prototype-like clade resembles the prototype and mostly consists of strains isolated prior to the 2000s. The host-adapted clade is comprised of sequences collected after the 2000s, including the swimming pool-associated strain. There are some fundamental differences between these two clades. The inverted terminal repeats (ITRs) length is ~115 bp and ~210 bp for the prototype-like clade and host-adapted clade, respectively. Multiple sequence alignment indicates that the ITR nucleotide are highly different at bases 115 to 180 for prototype-like clade and host-adapted clade. However, the bases, after 180, are highly similar for both clades. It is likely the host-adapted ITRs were the result of a recombinant event that occurred between the prototype-like HAdV-E4 and other human adenovirus types. This adaptation may grant the host-adapted clade with better infectivity. This is also confirmed by the neighbor-joining phylogeny tree in which some of the host-adapted ITRs were more related to the B species, than the other prototype-like sequences (data not shown).

The core origin and host nuclear factor III (NF-III) had very minor base changes compared to the adjacent host nuclear factor I (NF-I). The NF-I for the prototype-like clade resembled the simian adenovirus type E, as shown in the previous study (Dehghan, Seto, Liu, et al., 2013). The ITR of the host-adapted clade had more resemblance to the human adenoviruses type B and C, as shown in the Hong Kong (Zhang et al., 2019) and the Singapore (Coleman et al., 2019) isolates.

The regulatory regions prior to the start of the E1A gene (1- ~550 bp) were evaluated using multiple sequence alignments. The TATA box, CAAT signal site, and enhancer (~302 - ~310 bp) were identical for all HAdV-E4 isolates. Nonetheless, the enhancer (~209 - ~217 bp) had two base substitutions out of the nine-nucleotide sequence. The promoter of E1A had four base substitutions out of the forty nucleotides. The effect of these single nucleotide point mutations remains unknown. These changes may or may not affect the replication rate and the expression of the E1A gene, which is used in early transcription to mediate the replication of the viral genome.

Other HAdV-E4 cases were reported at different regions in the civilian facilities. In 2006, an outbreak of acute respiratory disease (ARD) happened to 15 elderly residents at the Hebrew Rehabilitation Center (HRC; Boston, MA). Four of the patients were confirmed to have HAdV-E4 induced infection, three patients had negative results, and seven were suspected cases. Three of the patients from 4 confirmed cases eventually died from

secondary complications from ARD (Kandel et al., 2010). Isolation from two of the four confirmed cultures were identified to be the HAdV-E4a1 variant. During 2011–2015, 36 cases of HAdV-E4 infection were identified in New York with 27 of these cases sampled from college settings. In October 2015, eight cases of the HAdV-E4a1 variant were identified in the same college at Thompkins county, NY (Kajon et al., 2018). Two cases of HAdV-E4a1 genome type were also identified at the same college a year before (Oct and Dec 2014). In addition to these minor outbreaks, HAdV-E4 were also sampled in other places: Hong Kong and Singapore. Out of the 102 adenovirus sequences randomly picked from patient files and sequenced in Hong Kong, 10 of these sequences were HAdV-E4 and seven were genome typed as HAdV-E4a1 variant (data not shown). A HAdV study in Singapore sequenced 533 HAdV-positive clinical samples collected between 2012 – 2018 showed that 15.2% of these cases were caused by HAdV-E4 infection (Coleman et al., 2019).

The Beijing swimming pool-associated outbreak is the largest and first recorded HAdV-E4 outbreak that occurred in the civilian population. In contrast, numerous cases of HAdV-E4 outbreaks had occurred in the military setting during the postvaccination era. For example, 194 U.S. Army recruits, from Fort Benning, Georgia, were hospitalized between April 23 - May 6, 2000, due to HAdV-E4 infection (DuVernoy et al., 2000). From April 23 – May 13, 1995, 70 unvaccinated basic trainees at Fort Jackson, South Carolina were hospitalized due to ARD. Six of the seven were serologically sampled as adenovirus type 4 (Barraza et al., 1999). Furthermore, between October – November

1998, a prospective study of ARD at the Fort Jackson, South Carolina resulted in 114/678 trainees hospitalized. 72% of the isolates, out of 97 specimens, were serotyped as HAdV-E4 (Cersovsky et al., 1999). Compared to the military epidemic, the general public had never experienced such a large-scale outbreak until now.

HAdV-E4 may have been spread from military to civilian populations and could cause more epidemics in densely populated areas in the future. The swimming pool-associated strain and other cases in recent years show a more frequent appearance of HAdV-E4 induced illness. This adaptation in ITR sequences is likely contributing to the virulence and universality of HAdV-E4 circulating in the general population. The reemergence of HAdV-E4 pathogen can be insidious, as it had never left the military population even with the presence of vaccine. The spread to civilian populations in the last decade was subtle and unstoppable. The appearance of the Beijing swimming pool-associate outbreak maybe the first one of many to come. More attention should be paid by the medical community to enhance biosurveillance efforts as it could cause severe damages to public health.

Acknowledgment

We thank Qiwei “Kevin” Zhang, The University of Hong Kong, for assistance in obtaining the sequencing data of HAdV-E4 swimming pool-associated strain. We would

also like to thank Noblis, Inc. for lending their computer cluster to support the Bayesian analysis.

References

Aoki, K, M Kato, H Ohtsuka, K Ishii, N Nakazono, and H Sawada. 1982. "Clinical and Aetiological Study of Adenoviral Conjunctivitis, with Special Reference to Adenovirus Types 4 and 19 Infections." *British Journal of Ophthalmology* 66 (12): 776. https://doi.org/10.1136/bjo.66.12.776.

Barraza, Emilio Margáin, Sharon L. Ludwig, Joel C. Gaydos, and John F. Brundage. 1999. "Reemergence of Adenovirus Type 4 Acute Respiratory Disease in Military Trainees: Report of an Outbreak during a Lapse in Vaccination." *The Journal of Infectious Diseases* 179 6: 1531–33.

Blasiole, Daniel A., David Metzgar, Luke T. Daum, Margaret A. K. Ryan, Jianguo Wu, Christopher Wills, Charles T. Le, et al. 2004. "Molecular Analysis of Adenovirus Isolates from Vaccinated and Unvaccinated Young Adults." *Journal of Clinical Microbiology* 42 (4): 1686–93. https://doi.org/10.1128/jcm.42.4.1686-1693.2004.

Cersovsky, Steven, Binn LN, Mitchell FD, Kolavic SA, Polyak CS, Jose Sanchez Jr, and Bester WT. 1999. "A Prospective Study of Acute Respiratory Diseases (ARD) in a Cohort of US Army Basic Trainees, October-November 1998, Fort Jackson, South Carolina." *Medical Surveillance Monthly Report* 5: 10, 12.

Cock, Peter J. A., Tiago Antao, Jeffrey T. Chang, Brad A. Chapman, Cymon J. Cox, Andrew Dalke, Iddo Friedberg, et al. 2009. "Biopython: Freely Available Python

Tools for Computational Molecular Biology and Bioinformatics." *Bioinformatics* 25 (11): 1422–23. https://doi.org/10.1093/bioinformatics/btp163.

Coleman, Kristen K, Chui Ching Wong, Jayanthi Jayakumar, Tham T Nguyen, Abigail W L Wong, Su Yadana, Koh C Thoon, et al. 2019. "Adenoviral Infections in Singapore: Should New Antiviral Therapies and Vaccines Be Adopted?" *The Journal of Infectious Diseases*, no. jiz489 (September). https://doi.org/10.1093/infdis/jiz489.

D'Angelo, L. J., J. C. Hierholzer, R. A. Keenlyside, L. J. Anderson, and W. J. Martone. 1979. "Pharyngoconjunctival Fever Caused by Adenovirus Type 4: Report of a Swimming Pool-Related Outbreak with Recovery of Virus from Pool Water." *The Journal of Infectious Diseases* 140 (1): 42–47. https://doi.org/10.1093/infdis/140.1.42.

Dehghan, Shoaleh, Jason Seto, Elizabeth B. Liu, Michael P. Walsh, David W. Dyer, James Chodosh, and Donald Seto. 2013. "Computational Analysis of Four Human Adenovirus Type 4 Genomes Reveals Molecular Evolution through Two Interspecies Recombination Events." *Virology* 443 (2): 197–207. https://doi.org/10.1016/j.virol.2013.05.014.

Dudding, B. A., F. H. Jr Top, P. E. Winter, E. L. Buescher, T. H. Lamson, and A. Leibovitz. 1973. "Acute Respiratory Disease in Military Trainees: The Adenovirus Surveillance Program, 1966-1971." *American Journal of Epidemiology* 97 (3): 187–98. https://doi.org/10.1093/oxfordjournals.aje.a121499.

DuVernoy, TS, RL Coldren, BH Feighner, NN Jordan, JL Sanchez, BJ Alsip, RS Gonsalez, RW Townsend, LN Binn, and S Williams. 2000. “Investigation of an Acute Respiratory Disease Outbreak Due to Adenovirus Type 4 among Recruits, Fort Benning, Georgia, April-May 2000.” *US Army Med Surveill Monthly Rep* 6: 2–7.

Gaydos, C. A., and J. C. Gaydos. 1995. “Adenovirus Vaccines in the U.S. Military.” *Military Medicine* 160 (6): 300–304.

Henry, Lionel, and Hadley Wickham. 2018. *Tidyselect: Select from a Set of Strings*. R package version 0.2.5. https://CRAN.R-project.org/package=tidyselect.

Hilleman, M. R., and J. H. Werner. 1954. “Recovery of New Agent from Patients with Acute Respiratory Illness.” *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* 85 (1): 183–88. https://doi.org/10.3181/00379727-85-20825.

Kajon, Adriana E., Daryl M. Lamson, Camden R. Bair, Xiaoyan Lu, Marie L. Landry, Marilyn Menegus, Dean D. Erdman, and Kirsten St. George. 2018. “Adenovirus Type 4 Respiratory Infections among Civilian Adults, Northeastern United States, 2011–20151.” *Emerging Infectious Diseases* 24 (2): 201–9. https://doi.org/10.3201/eid2402.171407.

Kajon, Adriana E., Jennifer M. Moseley, David Metzgar, Huo-Shu Huong, Aya Wadleigh, Margaret A. K. Ryan, and Kevin L. Russell. 2007. “Major Articles-Molecular Epidemiology of Adenovirus Type 4 Infections in US Military Recruits

in the Postvaccination Era (1997–2003).” *The Journal of Infectious Diseases* 196 (1): 67–75. https://doi.org/10.1086/518442.

Kandel, Ruth, Erika D’Agata, Xiaoyan Lu, DrPH Erdman, and MD, MPH, Michael. 2010. “Outbreak of Adenovirus Type 4 Infection in a Long-Term Care Facility for the Elderly.” *Infection Control and Hospital Epidemiology: The Official Journal of the Society of Hospital Epidemiologists of America* 31 (July): 755–57. https://doi.org/10.1086/653612.

Li, Jie, Xiaoyan Lu, Yamin Sun, Changying Lin, Feng Li, Yang Yang, Zhichao Liang, et al. 2018. “A Swimming Pool-Associated Outbreak of Pharyngoconjunctival Fever Caused by Human Adenovirus Type 4 in Beijing, China.” *International Journal of Infectious Diseases* 75 (October): 89–91. https://doi.org/10.1016/j.ijid.2018.08.009.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rowe, W. P., R. J. Huebner, L. K. Gilmore, R. H. Parrott, and T. G. Ward. 1953. “Isolation of a Cytopathogenic Agent from Human Adenoids Undergoing Spontaneous Degeneration in Tissue Culture.” *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* 84 (3): 570–73. https://doi.org/10.3181/00379727-84-20714.

Suchard, Marc A, Philippe Lemey, Guy Baele, Daniel L Ayres, Alexei J Drummond, and Andrew Rambaut. 2018. “Bayesian Phylogenetic and Phylodynamic Data

Integration Using BEAST 1.10." *Virus Evolution* 4 (vey016). https://doi.org/10.1093/ve/vey016.

Top, Franklin H., Jr., Richard A. Grossman, Peter J. Bartelloni, Herbert E. Segal, Burton A. Dudding, Philip K. Russell, and Edward L. Buescher. 1971. "Immunization with Live Types 7 and 4 Adenovirus Vaccines. I. Safety, Infectivity, Antigenicity, and Potency of Adenovirus Type 7 Vaccine in Humans." *The Journal of Infectious Diseases* 124 (2): 148–54. https://doi.org/10.1093/infdis/124.2.148.

Tsuzuki-Wang, Liping, Koki Aoki, Kazumi Isobe, Shia Shiao, Kazunori Toba, Nobuyoshi Kobayashi, Yuzo Noguchi, and Shigeaki Ohno. 1997. "Genome Analysis of Adenovirus Type 4 Strains Isolated from Acute Conjunctivitis in Japan." *Japanese Journal of Ophthalmology* 41 (5): 308–11. https://doi.org/10.1016/S0021-5155(97)00067-1.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org.

———. 2017. *Tidyverse: Easily Install and Load the "Tidyverse."* R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. *Dplyr: A Grammar of Data Manipulation*. R package version 0.7.8. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Lionel Henry. 2019. *Tidyr: Tidy Messy Data*. R package version 1.0.0. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Jim Hester, and Romain François. 2018. *Readr: Read Rectangular Text Data*. R package version 1.3.1. https://CRAN.R-project.org/package=readr.

Zhang, Jing, June Kang, Shoaleh Dehghan, Siddharth Sridhar, Susanna K. P. Lau, Junxian Ou, Patrick C. Y. Woo, Qiwei Zhang, and Donald Seto. 2019. "A Survey of Recent Adenoviral Respiratory Pathogens in Hong Kong Reveals Emergent and Recombinant Human Adenovirus Type 4 (HAdV-E4) Circulating in Civilian Populations." *Viruses* 11 (2). https://doi.org/10.3390/v11020129.

CHAPTER 4: Genomic and Bioinformatic Analysis of SAdV-21

Simian Adenovirus type 21 (SAdV-21), a chimpanzee virus, is included in a clade phylogenetically with members of human adenovirus species B. It is a model for analyzing the putative zoonotic or anthroponosis origins of emergent HAdV pathogens. SAdV-21 presents more genome percent identity similarity to human adenoviruses than simian adenoviruses, and, again, this provides a unique opportunity to examine the possibility of zoonotic origins. Genome recombination and phylogenetic analysis are used as tools to examine the lateral gene transfer within SAdV-21 genome.

1946 Genomics-based examination of simian adenovirus type 21 that clusters within the
1947 HAdV-B species.
1948

Genomics-based Examination of simian adenovirus type 21 that clusters within the HAdV-B species.

June Kang[1], Shoaleh Dehghan[1,2], and Donald Seto[1*]

[1]Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

[2]Chemistry Department, American University, Washington, D.C. 20016, USA

*Corresponding author. Mailing address: 10900 University Blvd., MSN 5B3, Department of Bioinformatics and Computational Biology, George Mason University, Manassas, VA 20110. Phone: (703) 993-8403. Fax: (703)993-8401. E-mail: dseto@gmu.edu.

Keywords: Adenovirus; SAdV-21; HAdV-B; Zoonosis; Anthroponosis; Amphizoonosis; Cross-species; Phylogenetics; Recombination.

Short running title: Computational analysis of SAdV-21 identifies a possible zoonotic or anthroponosis origin.

Abstract

Simian adenovirus type 21 (SAdV-21) is a chimpanzee virus that naturally clusters with the human adenovirus (HAdV) species B clade. SAdV-21 is the only non-human virus that is included in this clade. The genome of SAdV-21 is a recombination of HAdV-B21, SAdV-27.1, and SAdV-35.1. The multiple recombination events between HAdVs and SAdVs suggest it is a cross-species zoonotic or anthroponotic transition.

Dear editor,

Adenoviruses are single-stranded DNA viruses that were first isolated as respiratory pathogens from humans, and have been isolated and identified from other primate hosts including chimpanzees, gorillas, etc (Harrach et al., 2011; Rowe et al., 1953a). Human adenoviruses (HAdVs) can cause diverse pathogens in the respiratory, ocular, hepatic, renal, gastrointestinal, and cardiac systems, and have been associated with metabolic disease and obesity as well (Arnold et al., 2010; Lion, 2014; Yamada et al., 2012). Human adenoviruses are subdivided into seven different species from A to G (Harrach et al., 2011), parsed originally with respect to biological similarity.

Human and simian adenoviruses (SAdV) are classified separately by ICTV Adenovirus Working Group (https://talk.ictvonline.org). However, studies have shown that HAdV and SAdV are phylogenomically-related, and similar, and may cross species barriers in infecting hosts (Kang et al., 2020; Roy et al., 2004, 2006). Notably, HAdV-E4, which causes acute respiratory outbreaks, is a simian-like adenovirus that has been isolated from the military populations since 1954 (Coleman et al., 2019; Hilleman & Werner, 1954; Li et al., 2018; Rowe et al., 1953a). Another exception is SAdV-21, isolated from feces of a chimpanzee with respiratory illness, shown to have high similarity with HAdV-B species. SAdV-21 has been examined using bioinformatics tools to understand better its relationship to HAdVs, and to investigate whether it is a product of zoonosis.

Genome recombination analysis. The whole genome sequence and the three capsid genes hexon, fiber, and penton base of SAdV-21 was divided into 500 bp segments for detailed analysis. Genome fragments were Blast-analyzed to determine degrees of similarity between the sequences. A multiple sequence alignment (MSA) was built from selected sequences (best matches from Blast) along with the query sequences using Clustal Omega with default parameters (https://www.ebi.ac.uk/Tools/msa/clustalo/). MSA was also used to analyze the genome recombination events using Simplot bootscan (Lole et al., 1999), with default parameters: window size (1000 nucleotides [nt]), step size (20 nt), replicates used (n=100), gap stripping (on), distance model (Kimura), and tree model (neighbor-joining). Bootscan utilizes phylogenetic algorithms to analyze the possible recombination sites in the genome with the same parameter and altered window size (500 nt). A genome alignment visualization of the recombination events can also be viewed using multi-zPicture (https://zpicture.dcode.org/multiz.php).

Phylogenetic analysis. A multiple sequence alignment of nucleotide from 17 adenovirus species B isolates along with other HAdV outgroups was constructed using the web-accessible software tool MAFFT (https://mafft.cbrc.jp/alignment/server/). A whole genome maximum likelihood tree with 1000 bootstrap replica was generated using MEGA X 10.1 (https://www.megasoftware.net) software with the default parameter. The capsid genes for hexon, penton base, and fiber were extracted and blasted

(https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the closest match with the highest percent sequence identity.

GenBank accession numbers. Adenovirus genomes were retrieved from Genbank and are as follows: SAdV-21 (AC_000010); SAdV-35.1 (FJ025912); SAdV-35.2 (FJ025910); SAdV-27.1 (FJ025909); SAdV-28.2 (FJ025915); SAdV-32 (FJ025911); SAdV-47 (FJ025929); HAdV-76 (KF633445); HAdV-B21 (AY601633); HAdV-B50 (AY737798); HAdV-B16 (AY601636); HAdV-B68 (JN860678); HAdV-B7 (AY594255); HAdV-B3 (AY599834); HAdV-B66 (JN860676); HAdV-B34 (AY737797); HAdV-B35 (AY128640); HAdV-B11 (AY163756); HAdV-B79 (LC177352); HAdV-B14 (AY803294); HAdV-B55 (FJ643676); HAdV-B77 (KF268328); HAdV-B78 (KT970440); HAdV-E4 (AY594253); HAdV-D9 (AJ854486); HAdV-C1 (AF534906); HAdV-A18 (GU191019); HAdV-F40 (NC_001454); and HAdV-G52 (DQ923122).

SAdV-21 is the explicit non-human SAdV that clades globally phylogenetically with HAdV-B species (Figure 4.2). The genome of SAdV-21 has a length of 35,524 base pairs and a GC content of 51.16%. SAdV-21 clusters with the SAdV-35.1, SAdV-35.2, and HAdV-B76 (Figure 4.2). Notably, HAdV-B76, SAdV-35.1, and SAdV-35.2 are the same type of virus isolated from three different hosts(Dehghan, Seto, Jones, et al., 2013). SAdV-35.1 and SAdV-35.2 were originally identified from chimpanzees and bonobos (Roy et al., 2009), and HAdV-B76 were isolated from human host. HAdV-B76, SAdV-35.1, and SAdV-35.2 are a recombination of HAdV-B21, HAdV-B16, and SAdV-

21(Dehghan et al., 2019). SAdV-21 shares a 95.27 % genome similarity with HAdV-B76, SAdV-35.1, and SAdV-35.2. And the genome of SAdV-21 showed a similarity of 93.86 to HAdV-B21.

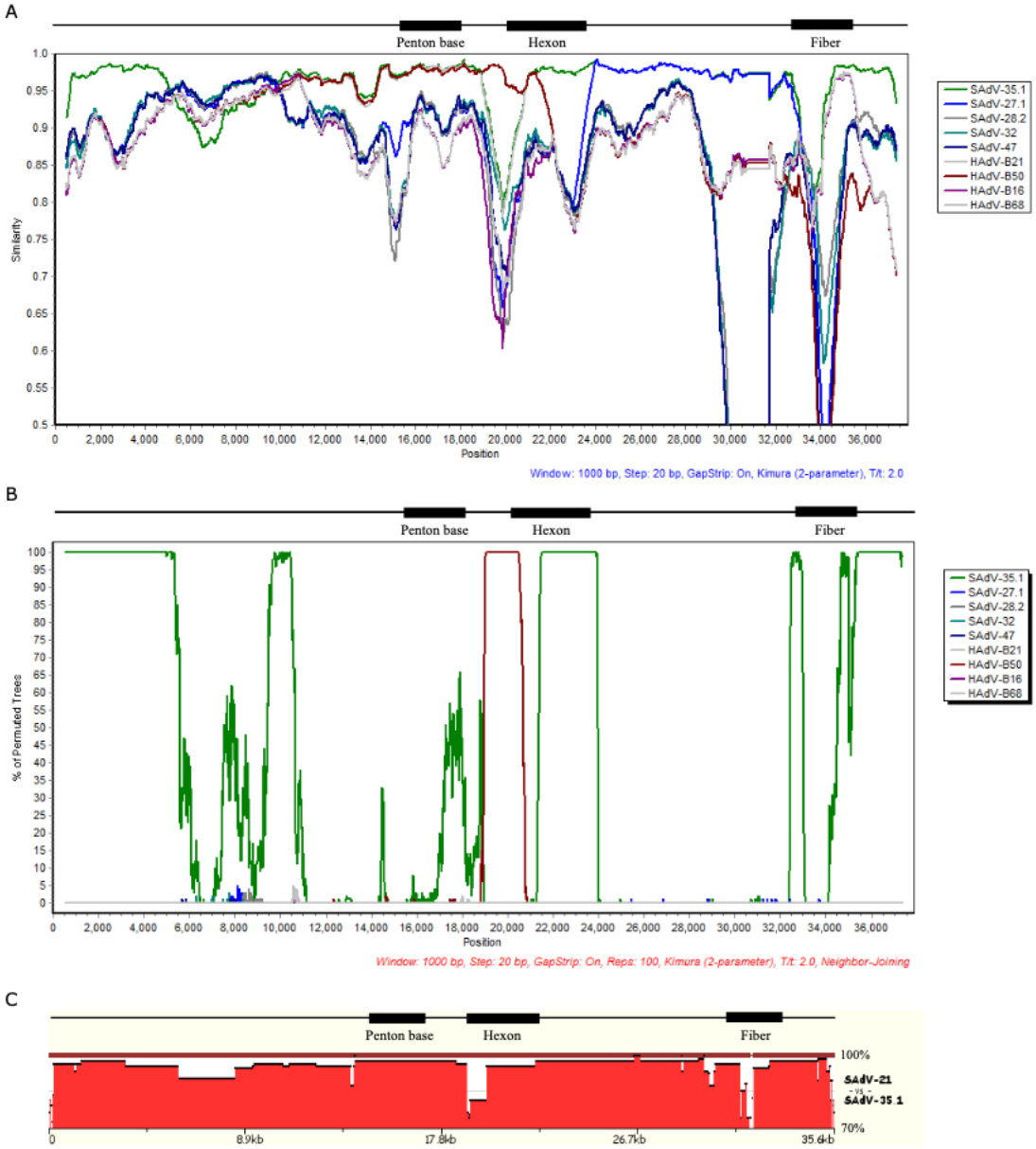
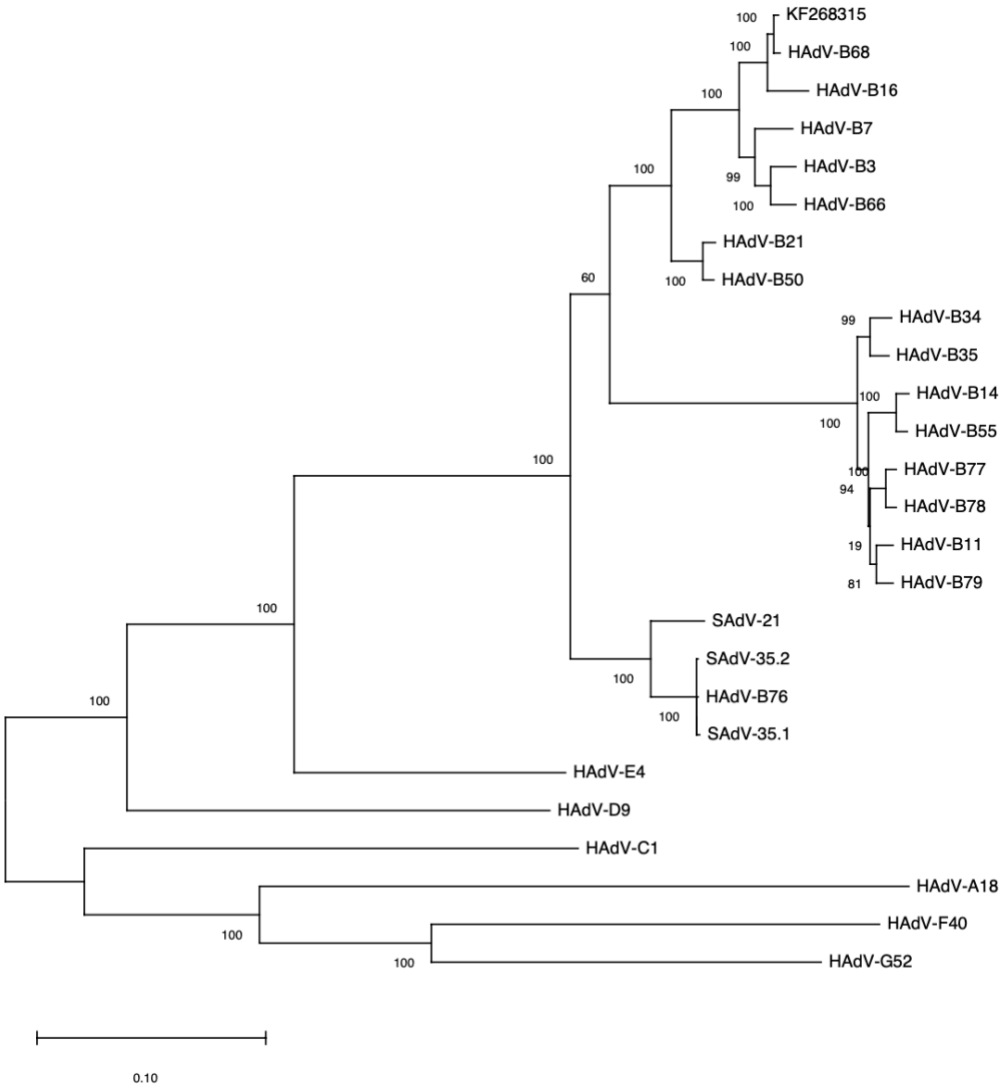


Figure 4.1. Whole genome sequence recombination analysis of SAdV-21. The sequence of SAdV-21 (AC_00010) was analyzed for homologous recombination with both SAdV and HAdV genome references using Simplot software tool (https://sray.med.som.jhmi.edu/SCRoftware/simplot/). Simplot analysis (A) showed a

high similarity between SAdV-21 and HAdV-B21. Bootscan analysis of the whole genome (B shows a lateral gene transfer of hexon gene from the HAdV-B50 and HAdV-B21. Recombination events occurred at the “changes in peaks” in the bootscan graph. Z-picture software (https://zpicture.dcode.org) was used to show the dynamic alignment between SAdV-21 and SAdV-35.1.

The penton base of HAdV-B21 and SAdV-21 have a sequence identity of 98%. SAdV-21 and SAdV-35.1 shares a 95% sequence identity for fiber. The first three hundred bases of fiber contain variations, and the remainder is identical to SAdV-35.1. The SAdV-21 penton base is a recombination of HAdV-B21 HVR1 (98.33%) and HAdV-B76 RGD (93.57%). The highest identity for SAdV-21 hexon, which is 94.41%, is with a SAdV isolated from chimpanzee feces (GenBank: KM659156.1). The hexon loop 1 and loop 2 sequence identity between SAdV-21 and KM659156.1 is 98.43% and 97.56%, respectively. SAdV-21 contains novel recombination from both HAdV and SAdV species and is possibly a result of cross-species zoonotic transmission between humans and chimpanzees in contact. A previous study also has shown HAdV and SAdV are phylogenomically related and may cross species barriers in infecting hosts (Kang et al., 2020; Roy et al., 2004, 2006).



2072

2073 Figure 4.2. Whole genome phylogenomic analysis of HAdV species B with selected
2074 HAdVs representing the span of species A-G. A multiple sequence alignment of
2075 nucleotide sequences was constructed using the MAFFT software with default parameters
2076 (https://mafft.cbrc.jp/alignment/server/). Maximum-likelihood tree with 1000 bootstrap
2077 replicates was generated using MEGA X 10.1 with a Tamura-Nei model
2078 (https://www.megasoftware.net) . Bootstrap values are based on a percentage of 1000

replicates, with values above 80 noted as robust. The scale bar indicates the unit of nucleotide substitutions per site.

SAdV-21 was originally the only SAdV that was a member of the species B clade of human adenoviruses, and now share a most recent common ancestor with HAdV-B76, SAdV-35.1, and SAdV-35.2. SAdV-21 is a recombination of HAdV-B21, SAdV-27.1, and HAdV-B50. This novel recombination suggests a possible zoonotic or anthroponosis origin for SAdV-21. As more simian adenoviruses are isolated, this will aid in the understanding of the molecular evolution of cross-species infection among primates.

Acknowledgments

References

Arnold, J., Jánoska, M., Kajon, A. E., Metzgar, D., Hudson, N. R., Torres, S., Harrach, B., Seto, D., Chodosh, J., & Jones, M. S. (2010). Genomic characterization of human adenovirus 36, a putative obesity agent. *Virus Research*, *149*(2), 152–161. https://doi.org/10.1016/j.virusres.2010.01.011

Coleman, K. K., Wong, C. C., Jayakumar, J., Nguyen, T. T., Wong, A. W. L., Yadana, S., Thoon, K. C., Chan, K. P., Low, J. G., Kalimuddin, S., Dehghan, S., Kang, J., Shamsaddini, A., Seto, D., Su, Y. C. F., & Gray, G. C. (2019). Adenoviral Infections in Singapore: Should New Antiviral Therapies and Vaccines Be

Adopted? *The Journal of Infectious Diseases*, *jiz489*. https://doi.org/10.1093/infdis/jiz489

Dehghan, S., Seto, J., Jones, M. S., Dyer, D. W., Chodosh, J., & Seto, D. (2013). Simian adenovirus type 35 has a recombinant genome comprising human and simian adenovirus sequences, which predicts its potential emergence as a human respiratory pathogen. *Virology*, *447*(1–2), 265–273. https://doi.org/10.1016/j.virol.2013.09.009

Dehghan, S., Seto, J., Liu, E. B., Ismail, A. M., Madupu, R., Heim, A., Jones, M. S., Dyer, D. W., Chodosh, J., & Seto, D. (2019). A Zoonotic Adenoviral Human Pathogen Emerged through Genomic Recombination among Human and Nonhuman Simian Hosts. *Journal of Virology*, *93*(18). https://doi.org/10.1128/JVI.00564-19

Harrach, B., Benkö, M., Both, G. W., Brown, M., Davison, A. J., Echavarria, M., Hess, M., Jones, M. S., Kajon, A., Lehmkuhl, H. D., Mautner, V., Mittal, S. K., & Wadell, G. (2011). *Virus Taxonomy: Classification and Nomenclature of Viruses*. 125–141.

Hilleman, M. R., & Werner, J. H. (1954). Recovery of new agent from patients with acute respiratory illness. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *85*(1), 183–188. https://doi.org/10.3181/00379727-85-20825

Kang, J., Ismail, A. M., Dehghan, S., Rajaiya, J., Allard, M. W., Lim, H. C., Dyer, D. W., Chodosh, J., & Seto, D. (2020). Genomics-based re-examination of the taxonomy

and phylogeny of human and simian Mastadenoviruses: An evolving whole genomes approach, revealing putative zoonosis, anthroponosis, and amphizoonosis. *Cladistics*, *36*(4), 358–373. https://doi.org/10.1111/cla.12422

Li, J., Lu, X., Sun, Y., Lin, C., Li, F., Yang, Y., Liang, Z., Jia, L., Chen, L., Jiang, B., & Wang, Q. (2018). A swimming pool-associated outbreak of pharyngoconjunctival fever caused by human adenovirus type 4 in Beijing, China. *International Journal of Infectious Diseases*, *75*, 89–91. https://doi.org/10.1016/j.ijid.2018.08.009

Lion, T. (2014). Adenovirus infections in immunocompetent and immunocompromised patients. *Clinical Microbiology Reviews*, *27*(3), 441–462. https://doi.org/10.1128/CMR.00116-13

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., & 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), 152–160.

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., & Ward, T. G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *84*(3), 570–573. https://doi.org/10.3181/00379727-84-20714

Roy, S., Gao, G., Lu, Y., Zhou, X., Lock, M., Calcedo, R., & Wilson, J. M. (2004). Characterization of a family of chimpanzee adenoviruses and development of molecular clones for gene transfer vectors. *Human Gene Therapy*, *15*(5), 519–530. https://doi.org/10.1089/10430340460745838

Roy, S., Vandenberghe, L. H., Kryazhimskiy, S., Grant, R., Calcedo, R., Yuan, X., Keough, M., Sandhu, A., Wang, Q., Medina-Jaszek, C. A., Plotkin, J. B., & Wilson, J. M. (2009). Isolation and Characterization of Adenoviruses Persistently Shed from the Gastrointestinal Tract of Non-Human Primates. *PLOS Pathogens*, *5*(7), e1000503. https://doi.org/10.1371/journal.ppat.1000503

Roy, S., Zhi, Y., Kobinger, G. P., Figueredo, J., Calcedo, R., Miller, J. R., Feldmann, H., & Wilson, J. M. (2006). Generation of an adenoviral vaccine vector based on simian adenovirus 21. *The Journal of General Virology*, *87*(Pt 9), 2477–2485. https://doi.org/10.1099/vir.0.81989-0

Yamada, T., Hara, K., & Kadowaki, T. (2012). Association of adenovirus 36 infection with obesity and metabolic markers in humans: A meta-analysis of observational studies. *PloS One*, *7*(7), e42031. https://doi.org/10.1371/journal.pone.0042031

CHAPTER 5: Phylogenetic Analysis of Human and Simian Adenoviruses

The assessment of the taxonomical and, presumably, biological relationships of human and simian adenoviruses using phylogenetic approaches is explored. A set of 167 primate adenoviruses (PrAdV) were analyzed using the maximum parsimony algorithm and phylogenetic networking. As a result, adenoviruses are reclassified and characterized based on genome sequence data, including the genome identity and single-nucleotide polymorphism changes in each clade. Moreover, the "most parsimonious" phylogenetic trees of hexon, penton base, fiber, DNA polymerase, E4ORF6, and E1A regions are presented, parsed with 1000 bootstrap iterations. This comprehensive study is important for understanding the basis of PrAdV speciation and is useful as a reference for other PrAdV studies.

Genome-based re-examination of the taxonomy and phylogeny of human and simian *Mastadenoviruses*: An evolving whole genome approach, revealing putative zoonosis, anthroponosis, and amphizoonosis.

Genomics-based re-examination of the taxonomy and phylogeny of *human* and *simian Mastadenoviruses*: An evolving whole genomes approach, revealing putative zoonosis, anthroponosis, and amphizoonosis.

June Kang[1a], Ashrafali Mohamed Ismail[2a], Shoaleh Dehghan[1,3], Jaya Rajaiya[2], Marc W. Allard[4], Haw Chuan Lim[5], David W. Dyer[6], James Chodosh[2], and Donald Seto[1*]

[1]Bioinformatics and Computational Biology Program, School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

[2]Department of Ophthalmology, Howe Laboratory, Massachusetts Eye and Ear, Harvard Medical School, Boston, Massachusetts 02114, USA

[3]Chemistry Department, American University, Washington, D.C. 20016, USA

[4]Division of Microbiology (HFS-710), Center for Food Safety & Applied Nutrition, US Food & Drug Administration, College Park, Maryland 20740, USA

[5]Department of Biology, George Mason University Manassas, Virginia 20110, USA

[6]Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, US

[a] These authors contributed equally.

*Corresponding author: dseto@gmu.edu.

2198 Keywords: Adenovirus; Infectious disease; Human pathogen; Cross-species transmission;
2199 Zoonosis; Anthroponosis; Amphizoonosis; Phylogenetic network; Taxonomy;
2200 Speciation; Character analysis
2201
2202 Short running title: Taxonomy and genome-based phylogeny of human and simian
2203 adenoviruses comprising the genus *Mastadenovirus*.
2204
2205

Abstract

With the advent of high-resolution and cost-effective genomics and bioinformatics tools and methods contributing to a large database of both human (HAdV) and simian (SAdV) adenoviruses, a genomics-based re-evaluation of their taxonomy is sensible. Interest in these particular adenoviruses is growing in part due to the applications of both in gene transfer protocols, including gene therapy and vaccines, as well in oncolytic protocols. In particular, the re-evaluation of SAdVs as appropriate vectors in human is important as zoonosis precludes the assumption that human immune systems may be naïve to these vectors. Additionally, adenoviruses are a model organism system for understanding viral pathogen emergence through zoonosis and anthroponosis, particularly among the primate species, along with recombination, host adaptation, and selection, as evidenced by one long-standing human respiratory pathogen HAdV-4 and a recent re-evaluation of another, HAdV-76. The latter reflects the insights on amphizoonosis, defined as infections in both directions amongst host species including “other than human”, that are possible with the growing database of nonhuman adenovirus genomes. HAdV-76 is a recombinant that has been isolated from human, chimpanzee, and bonobo hosts. On-going and potential impacts of adenoviruses on public health and translational medicine drive this evaluation of 174 whole genome sequences from HAdVs and SAdVs archived in GenBank. The conclusion is that rather than separate HAdV and SAdV phylogenetic trees, a single, intertwined tree is adequate with all HAdVs and SAdVs forming sister clades. Therefore, a single designation of “primate adenovirus” (PrAdV) superceding both HAdV and

SAdV is proposed, or alternatively, keeping HAdV for human adenovirus but expanding the SAdV nomenclature officially to include host species as in ChAdV for chimpanzee adenovirus, GoAdV for gorilla adenovirus, BoAdV for bonobo adenovirus, and *ad libitum*.

INTRODUCTION

Adenoviruses are non-enveloped, icosahedral viruses with linear double-stranded DNA genomes. These viruses are found in all vertebrates examined to date, spanning fishes to amphibians to reptiles to birds to mammals (Harrach et al., 2011). The family *Adenoviridae* is subdivided into five genera into which human adenoviruses (HAdV) and simian adenoviruses (SAdV) comprise the genus *Mastadenovirus* (Harrach et al., 2011). Within the historical taxonomic scheme based on presumed hosts and biology (and speculation), arbitrarily separate HAdV and SAdV species were recognized, with the HAdV serotypes distributed across seven species, HAdV species A-G, and SAdV types distributed into at least three species, SAdV species A-C arbitrarily noted in the literature (Chen et al., 2011b; Chiu et al., 2013; Malouli et al., 2014; Roy et al., 2012), or more, SAdV species A-G (Panto et al., 2015), all of which may or may not be recognized formally by the ICTV Adenovirus Working Group (Harrach et al., 2011). Recently, additional SAdV species D-I have been proposed to accommodate newly sequenced genomes (Dadakova et al., 2017). Oddly, a 2019 report of a global examination of selected adenoviruses noted the HAdV species but left SAdV entries as “unnamed” with respect to species (Harrach et al., 2019). Nevertheless, the shared, intertwined phylogenetic relationships between all SAdVs and HAdVs are noted in the speculation that several "members of three HAdV species most probably originated from [Old World Monkey] AdVs", from one member of the ICTV (Panto et al., 2015). This reinforces the

analyses presented in this report and supports the proposal that there should be no distinction between primate hosts amongst these clades or species, and, more importantly, the potential for zoonosis is substantial and worrisome (Bailey et al., 2018).

In this genomics era, HAdV identification and typing have evolved from the low-resolution, restrictive, and data-limited serological endeavors based solely on two antigenic epitopes (Aoki et al., 2011) to a broader, host-independent DNA sequence- and genome-based classification (Seto et al., 2011; Seto et al., 2013) that has allowed much more comprehensive and expansive views, with higher resolution understanding of their relationships providing important implications for their taxonomy. These implications include recognizing the branching of HAdV and SAdV genomes from a single phylogenetic tree with viruses of both groups forming subclades comprising the seven ICTV-recognized "HAdV" species and several novel but phylogenetically related species that could potentially include both SAdVs and HAdVs. That is, unlike the past taxonomic scheme of separating HAdV and SAdV species, the genomes of both "simian" and "human" AdV genotypes comingle, intertwine, and form clades and subclades together in one phylogenetic tree. Therefore, rather than segregating arbitrarily into HAdV species and SAdV species, we propose the designation of "primate adenovirus" (PrAdV) to recognize that these viruses encode highly similar genomes that apparently may interchange hosts. As the term "simian" in the past referred to only "nonhuman" SAdVs, for consistency with the historical literature, "simian" AdVs will refer to "nonhuman" simian AdVs in this report.

The timely re-evaluation of HAdV and SAdV taxonomy and phylogeny is of academic as well as clinical, epidemiological, and public health importance. HAdVs were among the first human viral respiratory pathogens to be isolated and characterized (Hilleman and Werner, 1954; Rowe et al., 1953). They are linked to a range of diseases in the respiratory, ocular, renal, hepatic, cardiac, and gastrointestinal systems, and are implicated in a metabolic disease, obesity (Arnold et al., 2010; Lion, 2014; Yamada et al., 2012). There are currently 103 Genbank-recognized genotypes of HAdVs (http://hadvwg.gmu.edu/), evaluated by a Human Adenovirus Working Group. Rather than being simply a basic biological research interest, SAdVs are increasingly recognized as important to human health from both an epidemiological perspective, e.g., zoonosis (Dehghan et al., 2013b; Purkayastha et al., 2005a; Purkayastha et al., 2005b), and a clinical applications perspective, e.g., gene delivery vectors (Graham and Prevec, 1992; Roy et al., 2004; Roy et al., 2006) and as oncolytic agents (Doronin and Shayakhmetov, 2012; Larson et al., 2015). Sequence recombination within the separate groups of HAdV and SAdV genomes has been reported as an important mechanism of adenovirus evolution, leading to novel and emergent viruses and pathogens (Hoppe et al., 2015b; Robinson et al., 2011; Robinson et al., 2013; Walsh et al., 2009; Walsh et al., 2010). Surprisingly, genome recombination between HAdVs and SAdVs has also been reported, suggesting zoonosis and anthroponosis (Dehghan et al., 2013a; Dehghan et al., 2019), along with complementing reports of cross-species viral transmissions between humans and other simians that elicited neutralizing antibodies in both (Chen et al., 2011b; Chiu et

al., 2013; Ersching et al., 2010; Hoppe et al., 2015a; Pauly et al., 2015; Roy et al., 2009; Wevers et al., 2011; Xiang et al., 2006). To provide a high-resolution and more comprehensive view of the phylogenetic relationships and taxonomy of HAdVs and SAdVs, this report employs a broader examination of all published HAdV and SAdV genomes to yield most parsimonious phylogenomic trees. This analysis yields a revised, broad taxonomic view of an intertwined phylogenetic tree with implications for human health, not only because these viruses are putative emergent and re-emergent pathogens (Chen et al., 2011b; Chiu et al., 2013; Ersching et al., 2010; Hoppe et al., 2015a; Pauly et al., 2015; Roy et al., 2009; Wevers et al., 2011; Xiang et al., 2006), but also because of their use as potential gene delivery and gene therapy vectors for medical therapies (Graham and Prevec, 1992; Roy et al., 2004; Roy et al., 2006). The latter is important as SAdVs have been presumed to be attractive, alternative vectors that bypass pre-existing human prior exposure and resultant immunity. It should be noted that due to the lack of whole genome data for many nonhuman and nonsimian adenoviruses, the exploration of cross-species transmission may be incomplete; therefore, this report is focused on primate hosts. Intriguingly, to date, partial hexon sequences suggest cross-species transmissions may have occurred between human and bat, and cat hosts, as noted in a recent literature survey of zoonosis, anthroponosis, and amphizoonosis (Borkenhagen et al., 2019).

MATERIALS AND METHODS

Genomes

174 adenovirus whole genome sequences were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/). Among these, 98 genomes were from viruses infecting human hosts, 69 were isolated from simian hosts, and seven were from non-human and non-simian hosts, selected as representing outgroups for phylogeny analysis. Note, some genomes were deposited by researchers for patent application purposes. One example is SAdV-6, which apparently is a modified vector (genomic regions deleted and/or laboratory-derived chimeric) and therefore excluded from this study. All viruses were collected and reported in the literature between 1951 and 2017, with sequence data deposits relatively recent (1984-2007). Additional information associated with these viruses is described in Table S1.

Genome sequence processing

Genome sequence alignments of the 174 strains were generated using tools in the software Multiple Alignment using Fast Fourier Transform package (MAFFT version 7), with default parameters (https://mafft.cbrc.jp/alignment/server/). MAFFT was selected for all alignments, including gene sequences, due to its speed and its capacity to handle a large amount of data. The whole genome alignment was additionally, manually inspected

for accuracy using BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html). To provide a more detailed analysis, eight genes and DNA sequences were selected, extracted, and analyzed in addition to the whole genome. These include the major capsid proteins used for recognizing and characterizing HAdV genotypes: penton base, hexon, and fiber. Additionally, the E1A, DNA polymerase, and E4ORF6 coding regions were sampled. The variable regions, loops 1 and 2, comprising one antigenic epitope, "epsilon", are contained in the hexon gene, while the second antigenic epitope, "gamma", is contained in the fiber gene. For individual gene processing, the genome coordinates of the start and end of each gene were mapped through GenBank records. Genes of interest were extracted from the whole genome alignment based its genome position in BioEdit. The sequences of each set of genes were aligned in MAFFT using default parameters. From this, both fasta and Nexus output alignment formats were saved. The fasta sequence alignment was visualized and inspected in BioEdit, and the Nexus alignment was used in PAUP to generate the most parsimonious tree. Other alignment software, such as ClustalW2, were also used to confirm alignments produced by MAFFT. Trees generated from either MAFFT and ClustalW2 alignments were not significantly distinct.

Outgroup Choices

Representative genomes were chosen as outgroups from the other genera in *Adenoviridae* for context, as the focus is on primate AdVs. Two AdVs were picked from each of the non-*Mastadenovirus* genera: *Aviadenovirus*, *Atadenovirus*, and *Siaadenovirus*. These were isolates from mammalian-distinct species including turkey, junglefowl, lizard,

snake, frog, and penguin, respectively. A fourth genus, *Ichtadenovirus*, is not represented as the archetype (fish) genome only contains partial CDS sequences. As an additional outgroup reference, a closely-related mammalian but nonprimate-hosted adenovirus, the northern treeshrew adenovirus (TAV), was used in the analyses. It forms a clade within the *Mastadenovirus* genus.

Phylogenetic analysis

Phylogenetic relationships amongst the sequences were determined using a most parsimonious tree algorithm for the whole genomes and select individual genes as implemented by "Phylogenetic Analysis Using Parsimony" (PAUP, version 4.0a build156; http://paup.sc.fsu.edu/) with 1,000 bootstrap replicates. All bootstrap analyses first were resampled with a full heuristic search. A random stepwise addition with 10 replicates was applied. The whole genome phylogenetic tree was also reconfirmed using the software "Tree analysis using New Technology" (TNT, version 1.5; http://www.lillo.org.ar/phylogeny/tnt/) (Nixon, 1999). A full search with Ratchet, tree drifting, and tree fusing set in default were used with the minimum length set as five times. Phylogenetic trees were annotated and manipulated in "Treegraph2" (version 2.13.0-748 beta; http://treegraph.bioinfweb.info/). Nodes and tip labels were edited in "Inkscape" (version 0.91; https://inkscape.org/en/).

Phylogenetic Network

A phylogenetic networks analysis based on default parameters was constructed using "SplitsTree4" (version 2.14.6; http://www.splitstree.org/) (Huson and Bryant, 2006). This was performed by submitting a "characters block" to NeighborNet and EqualAngle tools implementing a neighbor-net network and equal angle algorithm. A model comparison test was performed in jModelTest-2.1.10 to select the best-fit model of nucleotide substitution. The likelihood score with the best base tree search was computed and the model with the lowest Bayesian Information Criterion score (delta=0) was considered to best fit the nucleotide substitution pattern. Consequently, the GTR+I+G nucleotide substitution model was chosen for the characters transformations. Phylogenetic network analysis was utilized to account for the horizontal gene transfer, hybridization, and recombination using the realistic model. Nodes and edges were reformatted to illustrate a more accurate visualization.

Percent identity clustering

Genome sequence identity matrices for each lineage were generated using BioEdit. To analyze the identity matrices data, an R script was compiled to evaluate the percent identity clustering using the package 'gplots' (R Core Team 2017; https://cran.r-project.org/web/packages/gplots/index.html). Heat maps, reflecting sequence identities, for each designated PrAdV species were created to visualize the lineages in phylogenetic

trees. The lowest percent identity between two human adenovirus genomes was set as the cutoff threshold for species discrimination.

Comparative genome analysis

Detailed analyses of the sequences at the nucleotide ("character") level were performed using PAUP for the 167 query whole genome sequences after the most parsimonious tree was generated. In the context of the very large dataset generated, to define lineages and top gene changes that vary between subgroups of the phylogenetic tree of PrAdV that have CI = 1 were examined in more detail (Table S2). Conserved single nucleotide polymorphisms (SNP) across the genomes within the designated PrAdV-A, B, C, D, E, and H phylogenetic lineages ("species") were interpreted. Other PrAdV subclades and character changes with CI < 1 were included in Table S3 for further analysis.

Non-gap phylogenetic character changes were manually identified using the alignments in BioEdit. The Expasy translate tool allowed each nucleotide character identified to be translated into its corresponding protein sequence (http://web.expasy.org/translate/), which enabled the identification of point mutations based on nucleotide changes. Nucleotide and amino acid were then realigned in Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) to locate their positions and to confirm the changes in the genes.

RESULTS

Phylogenetic analysis of the whole genomes

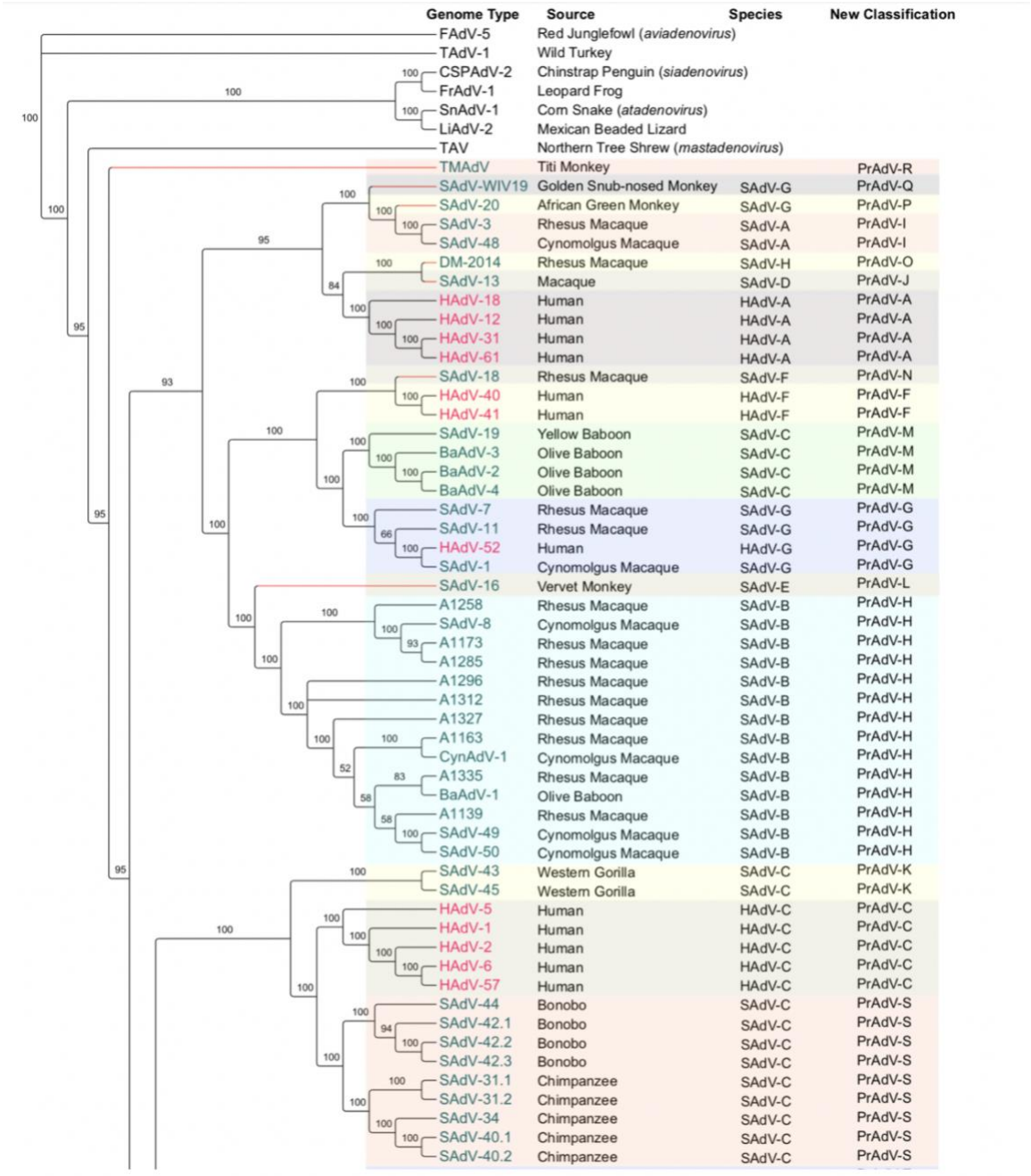
A most parsimonious phylogenetic analysis was performed on available whole genome sequences as well as on select gene/coding sequences of 174 HAdV and SAdV isolates collected and reported between 1951 and the present. These genome sequences include representatives of all HAdV and SAdV genotypes, along with seven outgroup representatives for reference. *In toto*, these genomes range in size from 25,000 to 45,000 nucleotides, with the average genome length amongst *Mastadenovirus* viruses being ~35,000 nucleotides, reconfirming earlier analyses with an incomplete set of whole genome data (Seto et al., 2010). The average G+C content is about 56%, which is distinctive for the *Mastadenovirus* genus, with characteristic percentages reflecting each of the seven HAdV species, and again reconfirming earlier observations (Seto et al., 2010).

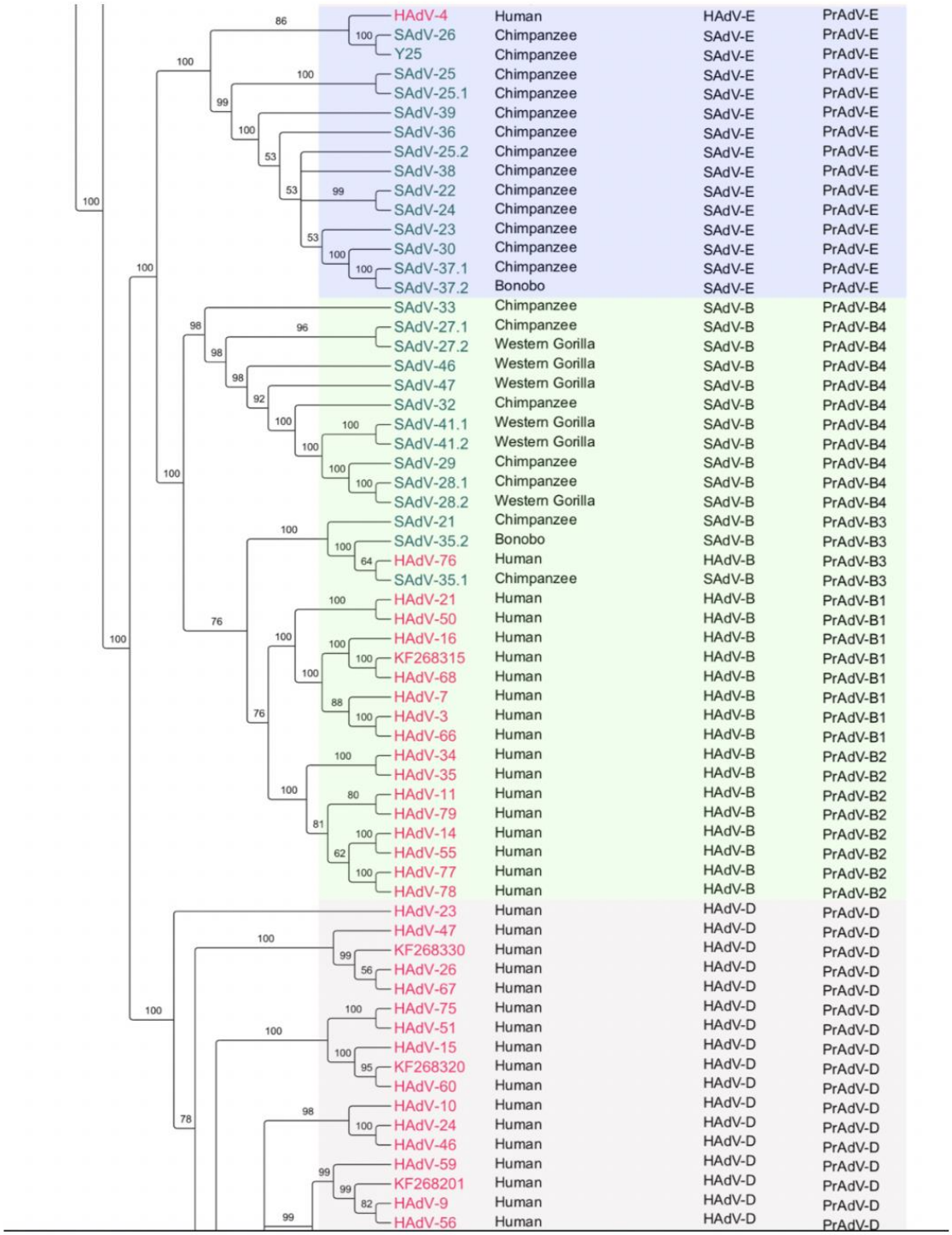
In this parsimony search, 68,414 characters were resampled within each replicate. Within this parsimony analysis, all characters have equal weight and are of the type ‘unord’. In summary, of the 68,414 total characters, 20,493 characters are constant, 9,598 variable characters are parsimony-uninformative, and 38,323 characters are parsimony-informative. The latter class of characters were then used to generate the lineages in the

most parsimonious phylogenetic tree, shown in Figure 5.1. Detailed information of the genomes are presented in Table S1, including GenBank accession numbers. Within this single phylogenetic tree, it should be noted that among clades comprising solely HAdV genotypes or SAdV genotypes, there are clades containing both. All clades branch closely together and given additional novel genomes, further iterations may show even closer phylogenetic relationships, as reported in the past, for example the inclusion of SAdV-21 into HAdV species B.

An additional phylogenetic analysis using TNT software reconfirmed the most parsimonious tree results, as shown in Supplemental Figures S1 and S2. The TNT analyses provided similar Consistency Index (CI) and Retention Index (RI) values, and resulted in highly similar lineages in both trees. A RI value of 0.782 was generated from TNT, versus 0.7809 for PAUP, with a reference of "1" representing the scenario in which the characters fit the tree perfectly, in a range from 0 to 1 (Farris, 1989). For trees generated in TNT with equal weight, this indicates a moderate "fit for character" measure. TNT analysis yielded a CI of 0.289, with "1" representing no homoplasy, in a range from 0 to 1 (Kluge and Farris, 1969), which is consistent with the PAUP analysis at 0.2884. The CI value may indicate homoplasy in alignments, as the MAFFT tool generated reproducible alignments, which were additionally re-evaluated and confirmed manually, and were supported by the Mauve results. The total branch lengths generated using TNT and PAUP were also similar. The total length reported in TNT is 348,219, which is very similar to 348,940, generated in PAUP. TNT showed a superior heuristic

search than PAUP. Total branch length differences suggest fluctuations in evolutionary time estimation between nodes.




HAdV-4 Human HAdV-E PrAdV-E
SAdV-26 Chimpanzee SAdV-E PrAdV-E
Y25 Chimpanzee SAdV-E PrAdV-E
SAdV-25 Chimpanzee SAdV-E PrAdV-E
SAdV-25.1 Chimpanzee SAdV-E PrAdV-E
SAdV-39 Chimpanzee SAdV-E PrAdV-E
SAdV-36 Chimpanzee SAdV-E PrAdV-E
SAdV-25.2 Chimpanzee SAdV-E PrAdV-E
SAdV-38 Chimpanzee SAdV-E PrAdV-E
SAdV-22 Chimpanzee SAdV-E PrAdV-E
SAdV-24 Chimpanzee SAdV-E PrAdV-E
SAdV-23 Chimpanzee SAdV-E PrAdV-E
SAdV-30 Chimpanzee SAdV-E PrAdV-E
SAdV-37.1 Chimpanzee SAdV-E PrAdV-E
SAdV-37.2 Bonobo SAdV-E PrAdV-E
SAdV-33 Chimpanzee SAdV-B PrAdV-B4
SAdV-27.1 Chimpanzee SAdV-B PrAdV-B4
SAdV-27.2 Western Gorilla SAdV-B PrAdV-B4
SAdV-46 Western Gorilla SAdV-B PrAdV-B4
SAdV-47 Western Gorilla SAdV-B PrAdV-B4
SAdV-32 Chimpanzee SAdV-B PrAdV-B4
SAdV-41.1 Western Gorilla SAdV-B PrAdV-B4
SAdV-41.2 Western Gorilla SAdV-B PrAdV-B4
SAdV-29 Chimpanzee SAdV-B PrAdV-B4
SAdV-28.1 Chimpanzee SAdV-B PrAdV-B4
SAdV-28.2 Western Gorilla SAdV-B PrAdV-B4
SAdV-21 Chimpanzee SAdV-B PrAdV-B3
SAdV-35.2 Bonobo SAdV-B PrAdV-B3
HAdV-76 Human HAdV-B PrAdV-B3
SAdV-35.1 Chimpanzee SAdV-B PrAdV-B3
HAdV-21 Human HAdV-B PrAdV-B1
HAdV-50 Human HAdV-B PrAdV-B1
HAdV-16 Human HAdV-B PrAdV-B1
KF268315 Human HAdV-B PrAdV-B1
HAdV-68 Human HAdV-B PrAdV-B1
HAdV-7 Human HAdV-B PrAdV-B1
HAdV-3 Human HAdV-B PrAdV-B1
HAdV-66 Human HAdV-B PrAdV-B1
HAdV-34 Human HAdV-B PrAdV-B2
HAdV-35 Human HAdV-B PrAdV-B2
HAdV-11 Human HAdV-B PrAdV-B2
HAdV-79 Human HAdV-B PrAdV-B2
HAdV-14 Human HAdV-B PrAdV-B2
HAdV-55 Human HAdV-B PrAdV-B2
HAdV-77 Human HAdV-B PrAdV-B2
HAdV-78 Human HAdV-B PrAdV-B2
HAdV-23 Human HAdV-D PrAdV-D
HAdV-47 Human HAdV-D PrAdV-D
KF268330 Human HAdV-D PrAdV-D
HAdV-26 Human HAdV-D PrAdV-D
HAdV-67 Human HAdV-D PrAdV-D
HAdV-75 Human HAdV-D PrAdV-D
HAdV-51 Human HAdV-D PrAdV-D
HAdV-15 Human HAdV-D PrAdV-D
KF268320 Human HAdV-D PrAdV-D
HAdV-60 Human HAdV-D PrAdV-D
HAdV-10 Human HAdV-D PrAdV-D
HAdV-24 Human HAdV-D PrAdV-D
HAdV-46 Human HAdV-D PrAdV-D
HAdV-59 Human HAdV-D PrAdV-D
KF268201 Human HAdV-D PrAdV-D
HAdV-9 Human HAdV-D PrAdV-D
HAdV-56 Human HAdV-D PrAdV-D

2461

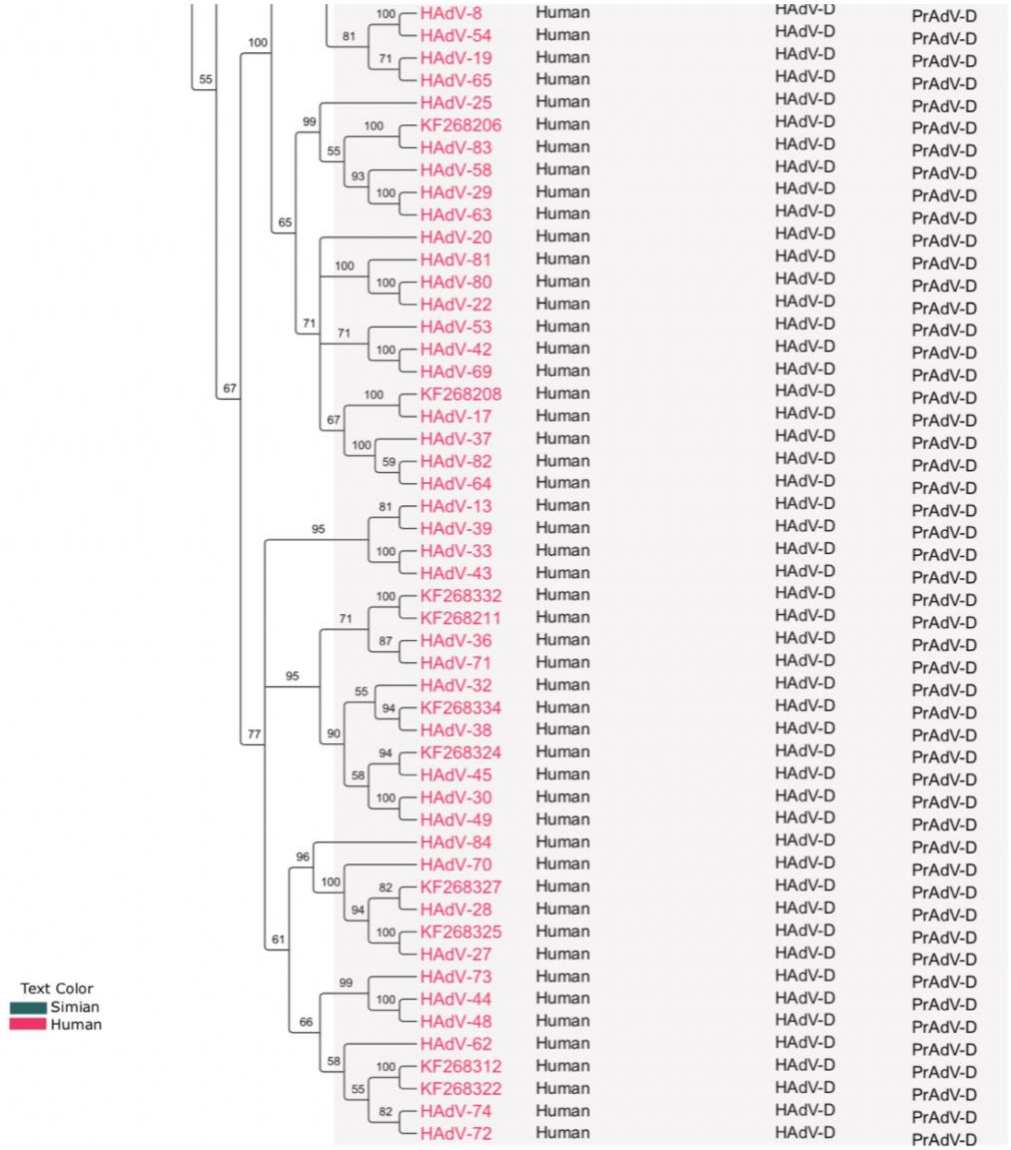


Figure 5.1. Maximum parsimony phylogenetic relationships of *Mastadenovirus* viruses. Whole genome sequences from 174 human, simian, and outgroup adenoviruses were aligned using MAFFT version 7 software, with default parameters (https://mafft.cbrc.jp/alignment/server/), and analyzed using the maximum parsimony algorithm embedded in the PAUP software (http://paup.sc.fsu.edu/). The iterations were

run with default parameters and 1000 bootstrap replicates. The clades are indicated in colored boxes in which the amaranth- and mauve-colored viruses as the leaves represent adenovirus genomes sampled from simians and humans, respectively. The current species designations of the human and simian adenovirus species are noted, followed by the proposed new designation “primate adenovirus species” (PrAdV). This classification recognizes the intertwined genomes and close phylogenetic distances. Bootstrap values are shown, with 80% noted as robust.

Nineteen distinct OTUs (Operational Taxonomic Units), clades, or lineages (to be referred to as “species” in accordance with the adenovirus literature) are classified from the whole genome most parsimonious phylogenetic tree analysis. These species are proposed to be renamed “Primate Adenovirus” species (PrAdV-A to S) to acknowledge their close relationships branching from a single tree and to reflect the intermingling of viral lineages regardless of host. This tree includes the established and formally-recognized human adenovirus (HAdV-A to G) species along with previously proposed and/or informally recognized various SAdV species that are noted in the literature (Chen et al., 2011b; Chiu et al., 2013; Dadakova et al., 2017; Harrach et al., 2011; Malouli et al., 2014; Panto et al., 2015; Roy et al., 2012), to be renamed “primate adenoviruses” with PrAdV-A to –G for the former and PrAdV-H to -S for the latter. Again, as noted in Figure 5.1, each of these species designations may include both HAdVs and SAdVs within the same clade, rather than the previous arbitrary separation and designation of species for each (Chen et al., 2011b; Chiu et al., 2013; Dadakova et al., 2017; Harrach et

al., 2011; Malouli et al., 2014; Panto et al., 2015; Roy et al., 2012). This observation is consistent with recent reports of and analyses incorporating additional SAdV genomes (Hoppe et al., 2015a; Hoppe et al., 2015b; Pauly et al., 2015; Roy et al., 2009; Wevers et al., 2010; Wevers et al., 2011).

Several major clades in the whole genome tree showed a more significant lineage. For example, species PrAdV-D is distinct and separated from all other AdV species. PrAdV-B clade comprises HAdV-B and SAdV-B isolates. This supports the hypothesis that the distinction between the HAdVs and SAdVs are trivial. The PrAdV-B clade is further divided into four subclades: PrAdV-B1, PrAdV-B2, PrAdV-B3 and PrAdV-B4. The lineages of these most parsimonious phylogenetic trees are consistent with previous analyses of smaller whole genome data sets and for individual HAdV genes.

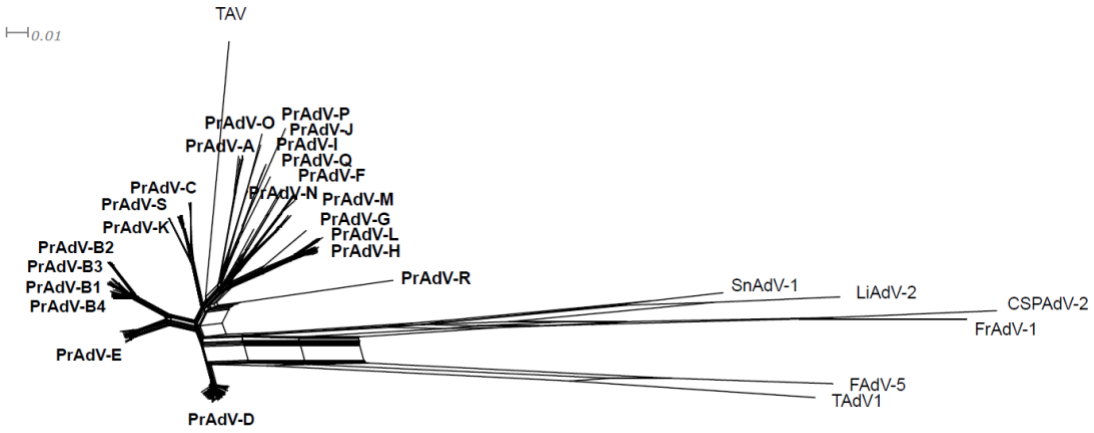


Figure 5.2. Phylogenetic network analysis of *Mastadenovirus* genomes. 174 whole genome sequences that were available from GenBank were collected and analyzed. A

phylogenetic network was constructed using SplitsTree software with default parameters applied to character blocks using the GTR+I+G algorithm model (http://www.splitstree.org/). These genomes represented all human and simian adenoviruses contained in the genus *Mastadenovirus*, along with six non-*Mastadenovirus* outgroup viruses and one closely-related nonprimate adenovirus (treeshrew, TAV). As shown in this network, the proposed primate adenovirus species (PrAdV) segregate into 19 species or clades, with one comprising subclades (PRAdV-B1 to B4). The ingroup for the adenovirus clades is in bold for each cluster. Provided is a scale bar that represents the number of weighted splits (bipartisans).

PrAdV-E comprises HAdV-4 and all SAdVs within the current HAdV-E species clade. PrAdV-C contains HAdV-C and a part of the SAdV-C subclade. PrAdV-G contains HAdV-52 and several SAdV-G strains. SAdV-43 and SAdV-45 form a separate clade, Finally, PrAdV-K is a clade of SAdV-43 and SAdV-45, leaving the following that do not cluster with any other genotypes and are separated into individual clades: SAdV-16, SAdV-18, SAdV-13, DM-2014, SAdV-20, SAdV-WIV19, and TMAdV. Note, changes to the existing parameters in PAUP generate the same results. The topology of the most parsimonious phylogenetic tree computed from the whole genome data was similar to the topologies of most parsimonious trees obtained for each individual gene and DNA sequence sets (hexon, penton base, fiber, DNA polymerase, E4ORF6, E1A, and hexon loop 1 and hexon loop 2). That is, no significant changes in the clades were evident. Phylogenetic network analysis (Neighbor-net algorithm; SplitsTree4, v2.14.6) illustrated

a similar relationship with multiple reticulate events at the outgroup regions and PrAdV-R. In general, both the phylogenetic network and the most parsimonious tree produce similar clades and node connections, as shown in Figure 5.2. This supports the parsing of species within human and simian adenoviruses.

Percent identity matrix analysis

Percent identity matrices of whole genome sequences were generated to determine the level of sequence similarities supporting the phylogenetic clades. To be consistent with the literature, pairs of HAdV genomes were compared to yield the minimum threshold observed for the separation of these HAdVs. An arbitrary threshold cutoff of 78.8% was determined and designated to be the minimum for two HAdVs belonging within a species clade, as shown for PrAdV-A in Table 5.1; the next minimum value observed was 79.1%, for a pair of HAdVs in the PRAdV-B clade. This designated threshold value parses the whole genome phylogenetic tree into 19 distinct clades or species. As shown in Figure 5.3A, the average percent identity separating each AdV from its nearest neighbor within each cluster is above the threshold value.

Table 5.1. Genome percent identity matrix summary table for 19 distinct primate adenovirus species noting the number of samples. The orange rectangle highlights the cutoff value for speciation amongst these genomes. The minimum for each primate adenovirus species is above the cutoff.

Type	# Isolates	Maximum	Minimum	Average	SD
PrAdV-A	4	0.976	0.788	0.845	0.033
PrAdV-B	31	0.997	0.791	0.862	0.010
PrAdV-C	5	0.977	0.934	0.948	0.008
PrAdV-D	68	0.983	0.875	0.923	0.006
PrAdV-E	15	0.999	0.869	0.922	0.013
PrAdV-F	2	0.856	0.856	0.856	-
PrAdV-G	4	0.954	0.826	0.879	0.031
PrAdV-H	14	0.989	0.854	0.916	0.018
PrAdV-I	2	0.913	0.913	0.913	-
PrAdV-J	1	-	-	-	-
PrAdV-K	2	0.97	0.97	0.970	-
PrAdV-L	1	-	-	-	-
PrAdV-M	4	1	0.86	0.902	0.027
PrAdV-N	1	-	-	-	-
PrAdV-O	1	-	-	-	-
PrAdV-P	1	-	-	-	-
PrAdV-Q	1	-	-	-	-
PrAdV-R	1	-	-	-	-
PrAdV-S	9	0.996	0.902	0.936	0.003

Genomes within each individual clade identify with themselves and are highly correlated, i.e., they have high sequence similarities to each other. For PrAdV-D, a total of 68 genomes were used in the percent identity matrix analysis, resulting in a minimal percent identity of 87.5%. Analysis of the 15 PrAdV-E genomes reveals a minimal percent identity of 86.9%. Average percent identities for PrAdV-E and PrAdV-D are 92.2%

±0.13% and 92.3% ±0.6%, respectively. The PrAdV-B heat map showed a clear separation among the four groups which directly correlate with the B1, B2, B3, and B4 in the whole genome tree. As a control, genomes have a percent identity of 1 when run against themselves. The average for individual clades is not skewed by this control in the matrix. The standard deviation shows that the error is in the acceptable range.

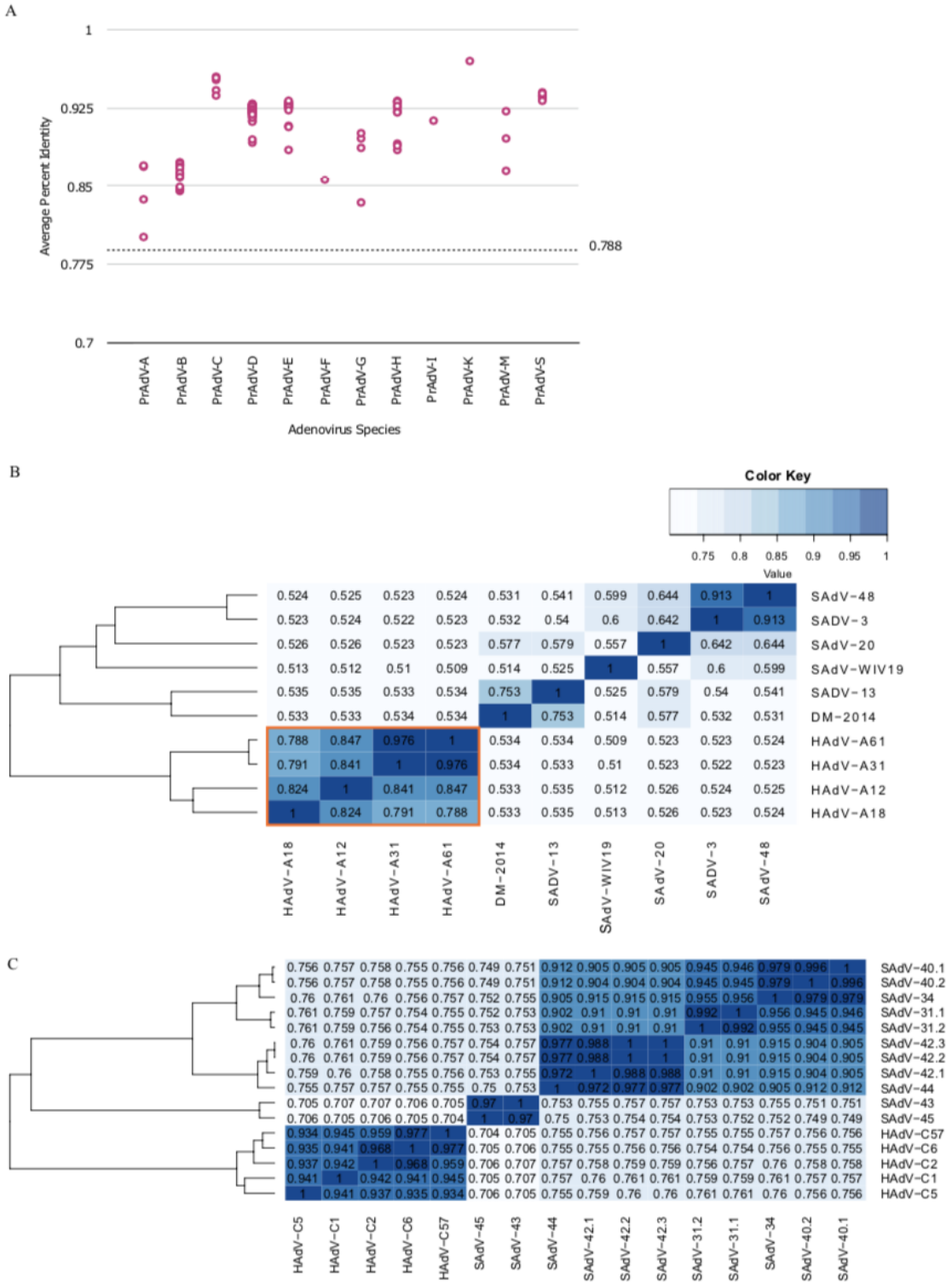

A
1
0.925
0.85
0.775
0.7
Average Percent Identity
0.788
PrAdV-A
PrAdV-B
PrAdV-C
PrAdV-D
PrAdV-E
PrAdV-F
PrAdV-G
PrAdV-H
PrAdV-I
PrAdV-K
PrAdV-M
PrAdV-S
Adenovirus Species
B
Color Key
0.75 0.8 0.85 0.9 0.95 1
Value
SAdV-48
SADV-3
SAdV-20
SAdV-WIV19
SADV-13
DM-2014
HAdV-A61
HAdV-A31
HAdV-A12
HAdV-A18
HAdV-A18
HAdV-A12
HAdV-A31
HAdV-A61
DM-2014
SADV-13
SAdV-WIV19
SAdV-20
SADV-3
SAdV-48
C
SAdV-40.1
SAdV-40.2
SAdV-34
SAdV-31.1
SAdV-31.2
SAdV-42.3
SAdV-42.2
SAdV-42.1
SAdV-44
SAdV-43
SAdV-45
HAdV-C57
HAdV-C6
HAdV-C2
HAdV-C1
HAdV-C5
HAdV-C5
HAdV-C1
HAdV-C2
HAdV-C6
HAdV-C57
SAdV-45
SAdV-43
SAdV-44
SAdV-42.1
SAdV-42.2
SAdV-42.3
SAdV-31.2
SAdV-31.1
SAdV-34
SAdV-40.2
SAdV-40.1

Figure 5.3. Percent identity analysis of human and simian adenovirus genomes. A) Average percent identity analysis is plotted for individual pairs of adenovirus genomes within each primate adenovirus species. The cutoff threshold for separating clades is indicated, at 0.788. B) Genome percent identity matrix is presented as a heatmap for the PrAdV-A, PrAdV-O, PrAdV-J, PrAdV-Q, PrAdV-P, and PrAdV-I pairs. C) Genome percent identity matrix heatmap for the PrAdv-C, PrAdV-K, and PrAdV-S pairs. Heatmap values are presented such that the darker color reflects higher similarity ("Color Key").

Genome-wide substitution analysis

The character changes for whole genome parsimony tree lineages identified 119 important amino acid changes that may define the six commonly seen PrAdV species A total of 18 character informative silent mutations were observed throughout the genome. The genetic variations in all the genes define each PrAdV clade. Single nucleotide polymorphisms (SNPs) are presented in all gene regions except the whole L5 and part of E3 and E4 gene regions due to higher base variations. The most character changes observed are in pIVa2, polymerase, L1 52K, and L4 100K. Hexon sequences also contained a higher number of SNPs at the conserved area C1 and C4 epitope. A base change of G to A at the polymerase and pTP region resulted in two different amino acid change R to H (position 45) and V to I (position 618), respectively (Table 5.2).

Table 5.2. Nonsynonymous SNPs in select genes that may define human adenoviruses comprising distinct PrAdV species. Twenty-one select genes spanning the adenovirus genome were selected and scored for conserved SNPs that are nonsynonymous. Those SNPs that are conserved amongst members of a species indicate a common ancestry and may be indicative of the species. For reference, human adenoviruses traditionally are classified across seven species (HAdV-A to G) and are proposed to be renamed as seven primate species (PrAdV-A to G), along with letters (H-S) to designate additional clades. Font colors green, blue, and orange indicate the polar, hydrophobic, and charged R-groups, respectively. Amino acids with dark font color represent synonymous mutations.

	E1A 28K	E1B 19K	E1B 55K	IX	IVa2	pol	pTP	L1 52K	IIIa	Penton base	VII	V	X	VI	Hexon	Protease	DBP	L4 100K	L4 22K	L4 33K	VIII
PrAdV-A	D256D	T107S	F370C		S350T G286S E133Q L118L	L1089I L1044S D953N S907S C758A A336K R45H*	V618I* V524I H476L M447L R173Q			G390A	A153T		R15R A34S	R16R	T99S R524R S579N	H188N		E539T K594R	R141R	H178L	Y13F S147T
PrAdV-B								L232Y R321R	M368L						T65M Y100F						
PrAdV-C				P19S	E395Q T252I		L94L	R114K T211M	Q311L			K200R	Y14F		G43N F702Y S729A	I83V		K676R			M31I Q53R
PrAdV-D	A74Q		P225A T255S V271A K303Q Q358N		V358A			G327A	A44S S147A S224S E293R	N357T Q425L				K52R N218H	R471K	T97S K110Q	S311S	A259I P261A G273A M302L Y357C Y602F E723V		K172K	R165N S174S
PrAdV-E			C275N		P228I			Y264F R275R	S293N						T888N			T368I			M101L
PrAdV-H					E47A	K554Q F338Y	I587V	M86L V127L K148Q P185A	P29A S293S	C358G				G20A	S5S R46R		P441T F347Y D303N	C362A K363Q G495A L600A	G133A		R51R

* Same base in an overlapping region.

DISCUSSION

The goals of this study are to understand the taxonomic relationships between HAdVs and SAdVs based on whole genome, gene, and DNA sequences; to identify micro-evolutionary differences that may link to the causation of adenovirus variations; and to reaffirm AdVs as an exceptional model system in the genomics era, one in which the origins of emergent human pathogens from zoonotic and anthroponotic sources and their routes in adaptation to a new host can be elucidated. To that end, we present the largest, comprehensive, phylogenetic analysis performed to date of HAdVs and SAdVs circulating worldwide. The trees in this study are very similar to those constructed previously both as preliminary trials and in other studies reported, albeit in small-scale datasets and recovered using different phylogenetic methods (Harrach et al., 2019; Prado-Irwin et al., 2018), with some minor changes in tree topology that do not affect our major conclusions. This larger phylogenomic study of the primate adenoviruses, particularly of 174 whole genomes, allows a more precise and comprehensive understanding of the evolutionary relationship of human and simian adenoviruses.

Recognition and elucidation of virus-host interactions, particularly viral-host adaptive evolution, have been enhanced greatly by genomics technology and databases. Many viruses undergo cross-species transmission (Parrish et al., 2008), resulting in emergent pathogens and novel human diseases. Human Severe Acute Respiratory Syndrome

Coronavirus (SARS-CoV) is an example of an emergent cross-species zoonotic pathogen first identified and characterized using DNA probe hybridization and array technology, coupled with the then-expanding genome and DNA sequence databases (Chen et al., 2011a). Subsequent and continuing analyses have identified bats as a reservoir or natural host (Hu et al., 2017; Li et al., 2005). A second example is the recent emergence of the Middle East Respiratory Syndrome-related Coronavirus (MERS-CoV) virus, from which genomic evidence has suggested origins or non-human reservoirs of bats and camels (Anthony et al., 2017; Chu et al., 2018).

HIV is yet another example of a zoonotically-derived emergent human viral pathogen that has been thoroughly documented and analyzed using genomics and phylogeny analysis tools. The origins of HIV appear to lie in multiple cross-species infections from chimpanzees to human over a period of time, but with only one of these cases leading to the establishment of HIV in humans. This host adaptation was likely the result of genome changes that were selected for optimal virus fitness in the new host. The related human disease, AIDS, and epidemic were apparently driven and amplified by the confluence of social and technological events in 1920 Kinshasa, the Congo (Faria et al., 2014; Keele et al., 2006).

AdVs provide an ideal model system for applying genomics technology, bioinformatics methodology, and resultant sequence databases in studying and understanding virus-host interactions and host adaptation and virus evolution in the context of zoonosis and

anthroponosis. HAdVs have been the subject of countless research inquiries across multiple disciplines and have been well-studied since the first dual simultaneous isolations in 1952 (Hilleman and Werner, 1954; Rowe et al., 1953). Through the years, a vast number of HAdVs, SAdVs, and other AdVs have been isolated, characterized, and genome-sequenced, with the most recent version of the GenBank database yielding 730 complete genomes across a myriad of vertebrate hosts (November 19, 2018). These AdV genomes are represented in the largest number with isolates from human and simian hosts. In particular, the diversity of the SAdVs provide multiple routes of zoonotic entry into new hosts and opportunities for host adaptation, particularly if the resultant human disease is benign or non-life threatening, as observed for most HAdV infections.

Zoonotic transmission of a TMAdV from titi monkeys to a human at the California National Primate Research Center, with subsequent infections and human familial transmission has been reported (Chen et al., 2011b). Another example of primate to human zoonosis was reported in an acute respiratory disease (ARD) outbreak involving baboon AdVs (BaAdV-3), in an olive baboon colony in 1997 at the Texas Biomedical Research Institute (Chiu et al., 2013a). Following that event, neutralizing antibody titers for BaAdV-3 were detected in both baboons and humans, with the additional discovery of neutralizing antibodies amongst the healthy baboons from 1996 to 2003, as well as from staff personnel from 1997 (Chiu et al., 2013a). In non-captive environments, a survey of human sera that tested positive for antibodies against SAdVs may be paired with a sub-Saharan Africa dataset that reported a PCR-based survey of nonhuman primates to

support the hypothesis that cross-species transmissions occur between humans and non-human primates (Wevers et al., 2011; Xiang et al., 2006). Additional reports of cross-species transmission of AdVs include sequence data and/or sero-neutralization data (Ersching et al., 2010; Hoppe et al., 2015a; Pauly et al., 2015; Wang et al., 2018). Again, the importance of these reports is two-fold. First, the bidirectional transmissions and the genome similarities may limit the advantage of using SAdV genomes as vectors for gene therapy and vaccine development. Second, significant, novel and emergent human pathogens may arise.

As an example of a "significant, novel and emergent human pathogen" that has arisen, the highly contagious and "long-established" human respiratory pathogen HAdV-4 serves as a precedent, highlighting zoonosis and host adaptation through genome recombination. This genotype was one of the first HAdVs isolated and characterized as a human pathogen associated with ARD. Its pathology and epidemiology were of such importance that it is one of only two HAdVs for which vaccines have been developed, thrice (1956, 1971, and 2011; https://www.historyofvaccines.org/content/blog/adenovirus-vaccines-reinstated-after-long-absence) (Lyons et al., 2008; Top et al., 1971). Each vaccine pause resulted in the immediate reemergence of HAdV-4 as a major pathogen (Barraza et al., 1999), as the prevalence of HAdV-4 returned to pre-vaccine levels upon discontinuation of the vaccine. Unexpectedly, this significant human pathogen was shown, using genomics and bioinformatics, to be a chimpanzee AdV (Purkayastha et al., 2005a; Purkayastha et al., 2005b). An apparent recent genome recombination event involving a

parental HAdV genome (Dehghan et al., 2013b) and its host protein-binding (NF1) motif that provides for optimized AdV replication in human cells (Hatfield and Hearing, 1991; Kenny et al., 1988; Mul et al., 1990; Nagata et al., 1982; Stillman et al., 1982). may have facilitated its adaptation to humans. Intriguingly, HAdV-4 is still the sole HAdV member of "HAdV" species E (Dehghan et al., 2013b; Purkayastha et al., 2005a), with chimpanzee AdVs comprising the rest of the clade. To reemphasize, it is known as "HAdV-4" despite its high genome similarity to fourteen SAdVs, comprising a clade historically named "HAdV species E" or the proposed PrAdV-E, as shown in Supplemental Figure S1. In more detail, HAdV-4 showed a ~100% genomic identity to SAdV-26, suggesting it is SAdV-26 (or vice versa as perhaps a simian host acquired HAdV-4 through anthroponosis, as it is difficult to ascertain with certainty the direction of gene/genome flow), which again highlights and supports the hypothesis that these HAdVs and SAdVs may be interchangeable with respect to these phylogenetically related hosts. As a model system, this raises the conundrum of whether nomenclature should be modified given high-resolution insights, that is, high-resolution data supercedes "historical entrenchment", and provides a truer perspective for biological, clinical, and public health reality.

Chimpanzee AdVs, for example strain Y25, have been shown to be either antigenically related to HAdV-2 (Hillis et al., 1969) or distinguishable from HAdVs (Hillis et al., 1968). Other chimpanzee AdV isolates, including strains Pan5 and Pan 9, were not neutralized by antisera against HAdVs (Basnight et al., 1971; Rogers et al., 1967).

Complicating the cross-species transmissions are recent genome analysis reports on HAdVs and SAdVs that indicate recombination events may be more common than previously thought, both within AdVs of each species and also between AdVs of different host species. These are significant with respect to public health, as the recombination events do lead to novel, emergent human pathogens, e.g., epidemic keratoconjunctivitis (Walsh et al., 2009) and acute respiratory disease (Walsh et al., 2010). In humans, adenovirus genomes undergo relatively frequent genetic recombination, particularly in the hypervariable regions of the genome, with sequence parameters reported (Lee et al., 2018). Although there are not many reports to date of HAdVs and SAdVs cross-species recombinants in the literature, there is one compelling example. SAdV-35.1 and -35.2 have been isolated from captive populations of chimpanzees and bonobos at two different sites (Dehghan et al., 2013a). The genome sequences show a near 100% identity, hence the non-standard and confusing names. A third, archived genome from 1967, has been recently sequenced and reported to show near identity to the above two genomes as well (manuscript in preparation). As it was isolated from a human host and led to a respiratory illness and fatality, it has been renamed HAdV-76 (Dehghan et al., 2019), rather than "SAdV-35.3" or the original designation "21+16/16". These three genomes contain recombination events from SAdV-21 (chimpanzee), HAdV-21 (human), SAdV-27 (chimpanzee), and HAdV-16 (human) (Dehghan et al., 2013a; Dehghan et al., 2019). This striking example underscores why AdVs may be one of the few, if not the only,

model organisms to explore and understand the phenomenon of “amphizoonosis”, that is, cross-species transmissions occurring in both directions.

A further complication is anthroponosis, which is the transmission of microbe from human to other animal species (Yu et al., 2009). Of concern is transmission to nonprimate host species. Should subsequent amphizoonosis occur, that is, between other nonhuman host species, potential recombinant AdVs may result in novel, emergent human pathogens. Recent reports indicate that bat and cat harbor at least one AdV each that have gene sequences with significant sequence similarities to their HAdV counterparts (Baker et al., 2013; Ongradi, 1999; Ongradi et al., 2019; Phan et al., 2006). Phylogenetic analysis of the fruit bat AdV-1 hexon shows it forms a clade separate from previously sequenced bat AdVs, and clusters with the HAdV clade with a high bootstrap value of 93 (Baker et al., 2013). The bat hexon shared 77% and 90% amino acid similarities with the HAdV counterpart across 58 and 63 amino acids, respectively, of the 900 amino acid hexon protein. The hexon gene of an isolated cat AdV presents with high sequence similarity to HAdV-1 as well (Ongradi, 1999; Ongradi et al., 2019). Independently, and geographically separated, the isolation of a cat AdV in a one-year old child hospitalized for acute gastroenteritis was reported in 2006. Sequence analysis shows 100% amino acid sequence identity in the seven hypervariable regions of the hexon gene and 97% amino acid sequence identity in the fiber gene to the HAdV-1 counterparts (Phan et al., 2006).

The current nomenclature, formally ICTV-approved or informally proposed in reports, is highly confusing and is misleading. In the absence of guidance, suggestions for a universal set of rules go unheard as researchers assign apparently random names. The example alluded to in the previous discussion is indicative of many similar names for different AdVs, isolated from different hosts albeit with near-identical or highly similar genomes. Other examples include SAdV-35.1 (chimpanzee; New Iberia Research Center) and SAdV-35.2 (bonobo; San Diego Zoo); SAdV-37.1 (chimpanzee; New Iberia Research Center) and SAdV-37.2 (bonobo; San Diego Zoo); SAdV-25 and SAdV-25.1; SAdV-27.1 (chimpanzee; New Iberia Research Center); and SAdV-27.2 (gorilla; Atlanta Zoo). In contrast, SAdV-25 and SAdV-25.1 are both chimpanzee viruses. These isolates and the nomenclature confusion reflect the intriguing hypothesis that certain SAdVs and HAdVs are able to infect hosts across species lines, perhaps as multiple events across multiple species hosts over time, and recombine genomes. Although the directionality of transmission is difficult to determine, the result is a novel adenovirus that may be an emergent pathogen. In some, or perhaps most instances, the zoonosis or anthroponosis is short-lived as the virus does not have sufficient fitness to spread significantly in the new host, e.g., TMAdV and BaAdV-3. Perhaps in some instances, conditions allow for a suboptimal grasp on the new host, selecting for a mutation or recombinational event that allows the novel host-adapted virus to thrive in an immune-naïve population, e.g., HAdV-4.

It may be argued that all SAdVs isolated from the primates in captive settings actually originate from human caretakers, as primates in captivity are often bred for generations and may have never been in contact with wild primates. As a counterargument, it should also be noted that, with the exception of HAdV-4 and perhaps HAdV-76, these SAdVs have never been found in human populations, at least the past 67 years since the first recognition and isolation of HAdVs in 1952 (Hilleman and Werner, 1954; Rowe et al., 1953).

Examination of the sequences of 174 isolates suggested that adenoviruses may be parsed into 19 distinct clades. Of these, there may be something fundamentally different about PrAdV-D genomes. First, the PrAdV-D clade is found exclusively in human hosts. Also, PrAdV-D genotypes will only recombine with others within the PrAdV-D clade. This indicates that there might be a barrier between the PrAdV-D clade and other clades which prevents lateral genome transfer. More data are required for a better understanding of these distinctions of the PRAdV-D clade. The average sequence identities for PrAdV-E and PrAdV-D genomes are 0.928 ± 0.032 and 0.924 ± 0.016, respectively. The ~7% difference in genome percent identity are centered in the hypervariable regions that include the penton base, hexon, and fiber genes. Similar results have been observed in other studies with smaller datasets (Robinson et al., 2013). PrAdV-B clade is the second largest clade in the whole genome phylogenetic tree. It can be further divided into four subclades: PrAdV-B1, PrAdV-B2, PrAdV-B3, and PrAdV-B4. While PrAdV-B1, PrAdV-B2, and PrAdV-B4 each contain solely either HAdV or SAdV, PrAdV-B3

contains both, that is, the SAdV-21, SAdV-35.2, SAdV-35.1, and HAdV-76 genotypes. Specifically, SAdV-35.1, SAdV-35.2 and HAdV-76 shared a percent identity of >0.996.

A total of 119 SNPs were identified for six PrAdV clades across four genes examined. These were found to have character informative mutations equal to or greater than the number of SNPs detected in DNA polymerase, which was used as a reference as it is presumably conserved. These four genes, IVa2, L1 52K, hexon C1 and C4 epitopes, and L4 100K may have conserved regions that are important for viral survival. All four genes are essential to produce mature virions, as IVa2 and L1 52K are required for DNA packaging (Zhang et al., 2001); L1 52K interacts with the IVa2 to aid the adenovirus assembly (Ma and Hearing, 2011; Zhang et al., 2001); and L4 100K is a chaperone protein for hexon and essential for cellular infectivity and adenovirus replication (Koyuncu & Dobner, 2009).

A difficulty and a caveat of this study is the potential quality issues in some of the sequences, as genomic sequences were performed by different laboratories, with a range of sequencing platforms and methodologies, and across three decades – some of the strains were sequenced when sequencing technology was still immature. As an example, the first HAdV genome sequences were composites generated in pieces from multiple research groups and included manual radioactive sequencing data. The first HAdV-17 genome data set has been discarded for being of poor quality; in fact, all of the original five HAdV genomes archived in Genbank have been resequenced. Nevertheless,

inaccurate data would contribute to an imprecise result. However, since nearly all of the sequences were performed recently and as there are vast amounts of quality data involved in this study, it is likely that the effects of sequencing errors are insignificant.

A substantial amount of sampling bias is also observed in this study since HAdVs have been studied disproportionally more than other AdVs. Originally, SAdV sequences were few in number and some were poorly sequenced and annotated. Although HAdVs currently outnumber the SAdV sequences available in this study, there are now sufficient sequences to ensure an accurate analysis and conclusions. As more SAdVs are being isolated, even higher quality analyses can be performed. However, it is anticipated that these additional genomes will support and enhance the current findings.

Phylogenetic networks were computed to represent and quantify the uncertainty in the parsimony trees. The reticulate event in the neighbor-net may be due to genetic variations among distinct species. PrAdV-R stands out as an apparent anomaly among the various species comprising the primate adenoviruses as it seemingly separates like the outgroups, which are included for reference. As noted in Figure 5.2, the network analysis displays lines for PrAdV-R and SnAdV-1, indicating past recombination events between the two. The overall layout of the tree and network displayed a high similarity which proves that sampling error and systematic error had little or no effect on the tree construction. Horizontal gene transfer and recombination were also taken into account in the

phylogenetic network. This more realistic approach has proven the validity of using phylogenetic tree lineages for adenovirus speciation.

This study demonstrates that the differences between HAdVs and SAdVs are negligible in terms of phylogeny and taxonomy. Clades or "species" within this phylogenic tree are intermixed with both AdVs as the relationships between HAdVs and SAdVs are complex and entangled, some of which have been revealed with high-resolution analyses. HAdVs and SAdVs do not appear to have evolved separately as would be implied by separate phylogenetic trees as accepted by the ICTV Adenovirus Working Group. These AdVs are from the same origins and utilized cross-species recombination as an evolutionary mechanism to survive and prosper within hosts, presumably also in new hosts with immune-naïve surveillance. Again, the prospect and reality of zoonosis and anthroponosis of AdVs amongst human and other simian hosts is relevant in terms of public health and applications to medicine.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

In memoriam: Professor Clark Tibbetts (2019), who provided the impetus and inspiration for this work. D.S. also thanks Chuck Ellison for generously providing an exceptional, quiet writing retreat during Thanksgiving 2018 at "Maison Jaureguy" in Tardets-

Sorholus, France, and Harold Suarez del Toro and the staff, at Biwa Hotel in Tulum, Mexico during Christmases 2018 and 2019 for the same, as well as “bottomless” cups of coffee. We thank Dr. Patrick Gillevet for discussions and advice on the methods, analyses, and terminology. This project was supported in part by U.S. Public Health Service grants EY013124, EY021558, and EY014104, and a Research to Prevent Blindness Senior Scientific Investigator Award (to J.C.).

References

Anthony, S.J., Gilardi, K., Menachery, V.D., Goldstein, T., Ssebide, B., Mbabazi, R., Navarrete-Macias, I., Liang, E., Wells, H., Hicks, A., Petrosov, A., Byarugaba, D.K., Debbink, K., Dinnon, K.H., Scobey, T., Randell, S.H., Yount, B.L., Cranfield, M., Johnson, C.K., Baric, R.S., Lipkin, W.I., Mazet, J.A., 2017. Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. MBio 8.

Aoki, K., Benko, M., Davison, A.J., Echavarria, M., Erdman, D.D., Harrach, B., Kajon, A.E., Schnurr, D., Wadell, G., 2011. Toward an integrated human adenovirus designation system that utilizes molecular and serological data and serves both clinical and fundamental virology. Journal of virology 85, 5703-5704.

Arnold, J., Janoska, M., Kajon, A.E., Metzgar, D., Hudson, N.R., Torres, S., Harrach, B., Seto, D., Chodosh, J., Jones, M.S., 2010. Genomic characterization of human adenovirus 36, a putative obesity agent. Virus research 149, 152-161.

Bailey, E.S., Fieldhouse, J.K., Choi, J.Y., Gray, G.C., 2018. A Mini Review of the Zoonotic Threat Potential of Influenza Viruses, Coronaviruses, Adenoviruses, and Enteroviruses. Front Public Health 6, 104.

Baker, K.S., Leggett, R.M., Bexfield, N.H., Alston, M., Daly, G., Todd, S., Tachedjian, M., Holmes, C.E., Crameri, S., Wang, L.F., Heeney, J.L., Suu-Ire, R., Kellam, P., Cunningham, A.A., Wood, J.L., Caccamo, M., Murcia, P.R., 2013. Metagenomic study of the viruses of African straw-coloured fruit bats: detection of a chiropteran poxvirus and isolation of a novel adenovirus. Virology 441, 95-106.

Barraza, E.M., Ludwig, S.L., Gaydos, J.C., Brundage, J.F., 1999. Reemergence of adenovirus type 4 acute respiratory disease in military trainees: report of an outbreak during a lapse in vaccination. The Journal of infectious diseases 179, 1531-1533.

Basnight, M., Jr., Rogers, N.G., Gibbs, C.J., Jr., Gajdusek, D.C., 1971. Characterization of four new adenovirus serotypes isolated from chimpanzee tissue explants. American journal of epidemiology 94, 166-171.

Borkenhagen, L.K., Fieldhouse, J.K., Seto, D., Gray, G.C., 2019. Are adenoviruses zoonotic? A systematic review of the evidence. Emerging microbes & infections 8, 1679-1687.

Chen, E.C., Miller, S.A., DeRisi, J.L., Chiu, C.Y., 2011a. Using a pan-viral microarray assay (Virochip) to screen clinical samples for viral pathogens. J Vis Exp.

Chen, E.C., Yagi, S., Kelly, K.R., Mendoza, S.P., Tarara, R.P., Canfield, D.R., Maninger, N., Rosenthal, A., Spinner, A., Bales, K.L., Schnurr, D.P., Lerche, N.W., Chiu, C.Y.,

2011b. Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. PLoS pathogens 7, e1002155.

Chiu, C.Y., Yagi, S., Lu, X., Yu, G., Chen, E.C., Liu, M., Dick, E.J., Jr., Carey, K.D., Erdman, D.D., Leland, M.M., Patterson, J.L., 2013. A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. MBio 4, e00084.

Chu, D.K.W., Hui, K.P.Y., Perera, R., Miguel, E., Niemeyer, D., Zhao, J., Channappanavar, R., Dudas, G., Oladipo, J.O., Traore, A., Fassi-Fihri, O., Ali, A., Demissie, G.F., Muth, D., Chan, M.C.W., Nicholls, J.M., Meyerholz, D.K., Kuranga, S.A., Mamo, G., Zhou, Z., So, R.T.Y., Hemida, M.G., Webby, R.J., Roger, F., Rambaut, A., Poon, L.L.M., Perlman, S., Drosten, C., Chevalier, V., Peiris, M., 2018. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. Proceedings of the National Academy of Sciences of the United States of America 115, 3144-3149.

Dadakova, E., Chrudimsky, T., Brozova, K., Modry, D., Celer, V., Hrazdilova, K., 2017. New adenoviruses from new primate hosts - growing diversity reveals taxonomic weak points. Mol Phylogenet Evol 107, 305-307.

Dehghan, S., Seto, J., Jones, M.S., Dyer, D.W., Chodosh, J., Seto, D., 2013a. Simian adenovirus type 35 has a recombinant genome comprising human and simian adenovirus sequences, which predicts its potential emergence as a human respiratory pathogen. Virology 447, 265-273.

2917 Dehghan, S., Seto, J., Liu, E.B., Ismail, A.M., Madupu, R., Heim, A., Jones, M.S., Dyer,
2918 D.W., Chodosh, J., Seto, D., 2019. A zoonotic adenoviral human pathogen emerged
2919 through genomic recombination amongst human and nonhuman simian hosts. J Virol.
2920 Dehghan, S., Seto, J., Liu, E.B., Walsh, M.P., Dyer, D.W., Chodosh, J., Seto, D., 2013b.
2921 Computational analysis of four human adenovirus type 4 genomes reveals molecular
2922 evolution through two interspecies recombination events. Virology 443, 197-207.
2923 Doronin, K., Shayakhmetov, D.M., 2012. Construction of targeted and armed oncolytic
2924 adenoviruses. Methods in molecular biology 797, 35-52.
2925 Ersching, J., Hernandez, M.I., Cezarotto, F.S., Ferreira, J.D., Martins, A.B., Switzer,
2926 W.M., Xiang, Z., Ertl, H.C., Zanetti, C.R., Pinto, A.R., 2010. Neutralizing antibodies to
2927 human and simian adenoviruses in humans and New-World monkeys. Virology 407, 1-6.
2928 Faria, N.R., Rambaut, A., Suchard, M.A., Baele, G., Bedford, T., Ward, M.J., Tatem,
2929 A.J., Sousa, J.D., Arinaminpathy, N., Pepin, J., Posada, D., Peeters, M., Pybus, O.G.,
2930 Lemey, P., 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in
2931 human populations. Science (New York, N.Y 346, 56-61.
2932 Farris, J.S., 1989. The Retention Index And The Rescaled Consistency Index. Cladistics
2933 5, 417–419.
2934 Graham, F.L., Prevec, L., 1992. Adenovirus-based expression vectors and recombinant
2935 vaccines. Biotechnology (Reading, Mass 20, 363-390.
2936 Harrach, B., Benkö, M., Both, G.W., Brown, M., Davison, A.J., Echavarria, M., Hess,
2937 M., Jones, M.S., Kajon, A., Lehmkuhl, H.D., Mautner, V., Mittal, S.K., Wadell, G., 2011.
2938 Family Adenoviridae, in: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J.

(Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Press, San Diego, pp. 125-141.

Harrach, B., Tarjan, Z.L., Benko, M., 2019. Adenoviruses across the animal kingdom: a walk in the zoo. FEBS Lett 593, 3660-3673.

Hatfield, L., Hearing, P., 1991. Redundant elements in the adenovirus type 5 inverted terminal repeat promote bidirectional transcription in vitro and are important for virus growth in vivo. Virology 184, 265-276.

Hilleman, M.R., Werner, J.H., 1954. Recovery of new agent from patients with acute respiratory illness. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y 85, 183-188.

Hillis, W.D., Garner, A.C., Hillis, A.I., 1969. A new simian adenovirus serologically related to human adenovirus type 2 and a chimpanzee with "viral hepatitis". American journal of epidemiology 90, 344-353.

Hillis, W.D., Holmes, A.W., Davison, V., 1968. Serologic characterization of adenoviruses isolated from chimpanzees associated with viral hepatitis. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y 129, 366-369.

Hoppe, E., Pauly, M., Gillespie, T.R., Akoua-Koffi, C., Hohmann, G., Fruth, B., Karhemere, S., Madinda, N.F., Mugisha, L., Muyembe, J.J., Todd, A., Petrzelkova, K.J., Gray, M., Robbins, M., Bergl, R.A., Wittig, R.M., Zuberbuhler, K., Boesch, C., Schubert, G., Leendertz, F.H., Ehlers, B., Calvignac-Spencer, S., 2015a. Multiple Cross-Species

Transmission Events of Human Adenoviruses (HAdV) during Hominine Evolution. Molecular biology and evolution 32, 2072-2084.

Hoppe, E., Pauly, M., Robbins, M., Gray, M., Kujirakwinja, D., Nishuli, R., Boji Mungu-Akonkwa, D.D., Leendertz, F.H., Ehlers, B., 2015b. Phylogenomic evidence for recombination of adenoviruses in wild gorillas. The Journal of general virology.

Hu, B., Zeng, L.P., Yang, X.L., Ge, X.Y., Zhang, W., Li, B., Xie, J.Z., Shen, X.R., Zhang, Y.Z., Wang, N., Luo, D.S., Zheng, X.S., Wang, M.N., Daszak, P., Wang, L.F., Cui, J., Shi, Z.L., 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS pathogens 13, e1006698.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23, 254-267.

Keele, B.F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M.L., Bibollet-Ruche, F., Chen, Y., Wain, L.V., Liegeois, F., Loul, S., Ngole, E.M., Bienvenue, Y., Delaporte, E., Brookfield, J.F., Sharp, P.M., Shaw, G.M., Peeters, M., Hahn, B.H., 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science (New York, N.Y 313, 523-526.

Kenny, M.K., Balogh, L.A., Hurwitz, J., 1988. Initiation of adenovirus DNA replication. I. Mechanism of action of a host protein required for replication of adenovirus DNA templates devoid of the terminal protein. The Journal of biological chemistry 263, 9801-9808.

Kluge, A.G., Farris, J.S., 1969. Quantitative Phyletics and the Evolution of Anurans. Systematic Biology 18, 1-32.

Larson, C., Oronsky, B., Scicinski, J., Fanger, G.R., Stirn, M., Oronsky, A., Reid, T.R., 2015. Going viral: a review of replication-selective oncolytic adenoviruses. Oncotarget 6, 19976-19989.

Lee, J.Y., Lee, J.S., Materne, E.C., Rajala, R., Ismail, A.M., Seto, D., Dyer, D.W., Rajaiya, J., Chodosh, J., 2018. Bacterial RecA Protein Promotes Adenoviral Recombination during In Vitro Infection. mSphere 3.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S., Wang, L.F., 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science (New York, N.Y 310, 676-679.

Lion, T., 2014. Adenovirus infections in immunocompetent and immunocompromised patients. Clinical microbiology reviews 27, 441-462.

Lyons, A., Longfield, J., Kuschner, R., Straight, T., Binn, L., Seriwatana, J., Reitstetter, R., Froh, I.B., Craft, D., McNabb, K., Russell, K., Metzgar, D., Liss, A., Sun, X., Towle, A., Sun, W., 2008. A double-blind, placebo-controlled study of the safety and immunogenicity of live, oral type 4 and type 7 adenovirus vaccines in adults. Vaccine 26, 2890-2898.

Ma, H.C., Hearing, P., 2011. Adenovirus structural protein IIIa is involved in the serotype specificity of viral DNA packaging. Journal of virology 85, 7849-7855.

3003 Malouli, D., Howell, G.L., Legasse, A.W., Kahl, C., Axthelm, M.K., Hansen, S.G., Fruh,
3004 K., 2014. Full genome sequence analysis of a novel adenovirus of rhesus macaque origin
3005 indicates a new simian adenovirus type and species. Virol Rep 3-4, 18-29.
3006 Mul, Y.M., Verrijzer, C.P., van der Vliet, P.C., 1990. Transcription factors NFI and
3007 NFIII/oct-1 function independently, employing different mechanisms to enhance
3008 adenovirus DNA replication. Journal of virology 64, 5510-5518.
3009 Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J.H., Hurwitz, J., 1982.
3010 Adenovirus DNA replication in vitro: identification of a host factor that stimulates
3011 synthesis of the preterminal protein-dCMP complex. Proceedings of the National
3012 Academy of Sciences of the United States of America 79, 6438-6442.
3013 Nixon, K.C., 1999. The Parsimony Ratchet, a New Method for Rapid Parsimony
3014 Analysis. Cladistics 15, 407-414.
3015 Ongradi, J., 1999. Identification of a feline adenovirus isolate that replicates in monkey
3016 and human cells in vitro. Am J Vet Res 60, 1463.
3017 Ongradi, J., Chatlynne, L.G., Tarcsai, K.R., Stercz, B., Lakatos, B., Pring-Akerblom, P.,
3018 Gooss, D., Sr., Nagy, K., Ablashi, D.V., 2019. Adenovirus Isolated From a Cat Is Related
3019 to Human Adenovirus 1. Front Microbiol 10, 1430.
3020 Panto, L., Podgorski, II, Janoska, M., Marko, O., Harrach, B., 2015. Taxonomy proposal
3021 for Old World monkey adenoviruses: characterisation of several non-human, non-ape
3022 primate adenovirus lineages. Archives of virology 160, 3165-3177.
3023 Parrish, C.R., Holmes, E.C., Morens, D.M., Park, E.C., Burke, D.S., Calisher, C.H.,
3024 Laughlin, C.A., Saif, L.J., Daszak, P., 2008. Cross-species virus transmission and the

emergence of new epidemic diseases. Microbiology and molecular biology reviews : MMBR 72, 457-470.

Pauly, M., Akoua-Koffi, C., Buchwald, N., Schubert, G., Weiss, S., Couacy-Hymann, E., Anoh, A.E., Mossoun, A., Calvignac-Spencer, S., Leendertz, S.A., Leendertz, F.H., Ehlers, B., 2015. Adenovirus in Rural Cote D`Ivoire: High Diversity and Cross-Species Detection. Ecohealth.

Phan, T.G., Shimizu, H., Nishimura, S., Okitsu, S., Maneekarn, N., Ushijima, H., 2006. Human adenovirus type 1 related to feline adenovirus: evidence of interspecies transmission. Clin Lab 52, 515-518.

Prado-Irwin, S.R., van de Schoot, M., Geneva, A.J., 2018. Detection and phylogenetic analysis of adenoviruses occurring in a single anole species. PeerJ 6, e5521.

Purkayastha, A., Ditty, S.E., Su, J., McGraw, J., Hadfield, T.L., Tibbetts, C., Seto, D., 2005a. Genomic and bioinformatics analysis of HAdV-4, a human adenovirus causing acute respiratory disease: implications for gene therapy and vaccine vector development. Journal of virology 79, 2559-2572.

Purkayastha, A., Su, J., McGraw, J., Ditty, S.E., Hadfield, T.L., Seto, J., Russell, K.L., Tibbetts, C., Seto, D., 2005b. Genomic and bioinformatics analyses of HAdV-4vac and HAdV-7vac, two human adenovirus (HAdV) strains that constituted original prophylaxis against HAdV-related acute respiratory disease, a reemerging epidemic disease. Journal of clinical microbiology 43, 3083-3094.

Robinson, C.M., Singh, G., Henquell, C., Walsh, M.P., Peigue-Lafeuille, H., Seto, D., Jones, M.S., Dyer, D.W., Chodosh, J., 2011. Computational analysis and identification of

an emergent human adenovirus pathogen implicated in a respiratory fatality. Virology 409, 141-147.

Robinson, C.M., Zhou, X., Rajaiya, J., Yousuf, M.A., Singh, G., Deserres, J.J., Walsh, M.P., Wong, S., Seto, D., Dyer, D.W., Chodosh, J., Jones, M.S., 2013. Predicting the next eye pathogen: analysis of a novel adenovirus. MBio 4.

Rogers, N.G., Basnight, M., Gibbs, C.J., Gajdusek, D.C., 1967. Latent viruses in chimpanzees with experimental kuru. Nature 216, 446-449.

Rowe, W.P., Huebner, R.J., Gilmore, L.K., Parrott, R.H., Ward, T.G., 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y 84, 570-573.

Roy, S., Gao, G., Lu, Y., Zhou, X., Lock, M., Calcedo, R., Wilson, J.M., 2004. Characterization of a family of chimpanzee adenoviruses and development of molecular clones for gene transfer vectors. Human gene therapy 15, 519-530.

Roy, S., Sandhu, A., Medina, A., Clawson, D.S., Wilson, J.M., 2012. Adenoviruses in fecal samples from asymptomatic rhesus macaques, United States. Emerging infectious diseases 18, 1081-1088.

Roy, S., Vandenberghe, L.H., Kryazhimskiy, S., Grant, R., Calcedo, R., Yuan, X., Keough, M., Sandhu, A., Wang, Q., Medina-Jaszek, C.A., Plotkin, J.B., Wilson, J.M., 2009. Isolation and characterization of adenoviruses persistently shed from the gastrointestinal tract of non-human primates. PLoS pathogens 5, e1000503.

Roy, S., Zhi, Y., Kobinger, G.P., Figueredo, J., Calcedo, R., Miller, J.R., Feldmann, H., Wilson, J.M., 2006. Generation of an adenoviral vaccine vector based on simian adenovirus 21. The Journal of general virology 87, 2477-2485.

Seto, D., Chodosh, J., Brister, J.R., Jones, M.S., 2011. Using the whole-genome sequence to characterize and name human adenoviruses. Journal of virology 85, 5701-5702.

Seto, D., Jones, M.S., Dyer, D.W., Chodosh, J., 2013. Characterizing, typing, and naming human adenovirus type 55 in the era of whole genome data. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology 58, 741-742.

Seto, J., Walsh, M.P., Mahadevan, P., Zhang, Q., Seto, D., 2010. Applying genomic and bioinformatic resources to human adenovirus genomes for use in vaccine development and for applications in vector development for gene delivery. Viruses 2, 1-26.

Stillman, B.W., Topp, W.C., Engler, J.A., 1982. Conserved sequences at the origin of adenovirus DNA replication. Journal of virology 44, 530-537.

Top, F.H., Jr., Grossman, R.A., Bartelloni, P.J., Segal, H.E., Dudding, B.A., Russell, P.K., Buescher, E.L., 1971. Immunization with live types 7 and 4 adenovirus vaccines. I. Safety, infectivity, antigenicity, and potency of adenovirus type 7 vaccine in humans. The Journal of infectious diseases 124, 148-154.

Walsh, M.P., Chintakuntlawar, A., Robinson, C.M., Madisch, I., Harrach, B., Hudson, N.R., Schnurr, D., Heim, A., Chodosh, J., Seto, D., Jones, M.S., 2009. Evidence of molecular evolution driven by recombination events influencing tropism in a novel human adenovirus that causes epidemic keratoconjunctivitis. PLoS ONE 4, e5635.

Walsh, M.P., Seto, J., Jones, M.S., Chodosh, J., Xu, W., Seto, D., 2010. Computational analysis identifies human adenovirus type 55 as a re-emergent acute respiratory disease pathogen. Journal of clinical microbiology 48, 991-993.

Wang, B., Li, J., Wu, S., Chen, Y., Zhang, Z., Zhai, Y., Guo, Q., Zhang, J., Song, X., Zhao, Z., Hou, L., Chen, W., 2018. Seroepidemiological investigation of HAdV-4 infection among healthy adults in China and in Sierra Leone, West Africa. Emerging microbes & infections 7, 200.

Wevers, D., Leendertz, F.H., Scuda, N., Boesch, C., Robbins, M.M., Head, J., Ludwig, C., Kuhn, J., Ehlers, B., 2010. A novel adenovirus of Western lowland gorillas (Gorilla gorilla gorilla). Virol J 7, 303.

Wevers, D., Metzger, S., Babweteera, F., Bieberbach, M., Boesch, C., Cameron, K., Couacy-Hymann, E., Cranfield, M., Gray, M., Harris, L.A., Head, J., Jeffery, K., Knauf, S., Lankester, F., Leendertz, S.A., Lonsdorf, E., Mugisha, L., Nitsche, A., Reed, P., Robbins, M., Travis, D.A., Zommers, Z., Leendertz, F.H., Ehlers, B., 2011. Novel adenoviruses in wild primates: a high level of genetic diversity and evidence of zoonotic transmissions. Journal of virology 85, 10774-10784.

Xiang, Z., Li, Y., Cun, A., Yang, W., Ellenberg, S., Switzer, W.M., Kalish, M.L., Ertl, H.C., 2006. Chimpanzee adenovirus antibodies in humans, sub-Saharan Africa. Emerging infectious diseases 12, 1596-1599.

Yamada, T., Hara, K., Kadowaki, T., 2012. Association of adenovirus 36 infection with obesity and metabolic markers in humans: a meta-analysis of observational studies. PLoS ONE 7, e42031.

Yu, H., Zhou, Y.J., Li, G.X., Zhang, G.H., Liu, H.L., Yan, L.P., Liao, M., Tong, G.Z., 2009. Further evidence for infection of pigs with human-like H1N1 influenza viruses in China. Virus research 140, 85-90.

Zhang, W., Low, J.A., Christensen, J.B., Imperiale, M.J., 2001. Role for the adenovirus IVa2 protein in packaging of viral DNA. Journal of virology 75, 10446-10454.

Supplemental Material

Figure S1. Most parsimonious phylogenetic tree of the *Mastadenovirus* viruses. Whole genome sequences of 174 adenoviruses were aligned using the "Multiple Alignment" software tool incorporating a" Fast Fourier Transform" (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using the most parsimonious phylogenetic tree algorithm in the TNT software (http://www.lillo.org.ar/phylogeny/tnt/), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans, respectively. The scale bar indicates the branch length. Bootstrap values >50% are displayed, with 80% noted as significant.

Figure S2. Most parsimonious phylogenetic tree of *Mastadenovirus* viruses. Whole genome sequences of 174 adenoviruses were aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default

parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in the PAUP software (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans, respectively. The scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S3. Most parsimonious phylogenetic tree of the adenoviral hexon gene with 1000 bootstrap replicates. Hexon gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S4. Most parsimonious phylogenetic tree of the adenoviral penton base gene with 1000 bootstrap replicates. Penton base gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed

using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S5. Most parsimonious phylogenetic tree of the adenoviral fiber gene with 1000 bootstrap replicates. Fiber gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S6. Most parsimonious phylogenetic tree of the adenoviral DNA polymerase gene with 1000 bootstrap replicates. DNA polymerase gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu),

with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S7. Most parsimonious phylogenetic tree of the adenoviral E4ORF6 gene with 1000 bootstrap replicates. E4ORF6 gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as significant.

Figure S8. Most parsimonious phylogenetic tree phylogeny of the adenoviral E1A gene with 1000 bootstrap replicates. E1A gene sequences of 174 adenoviruses were extracted from the whole genome sequence and aligned using a Multiple Alignment software tool incorporating a Fast Fourier Transform (MAFFT version 7) with default parameters (https://mafft.cbrc.jp/alignment/server/). Post-alignment, the data were parsed using a most parsimonious phylogenetic tree algorithm in PAUP (http://paup.sc.fsu.edu), with default parameters and 1000 bootstrap replicates. Amaranth- and mauve-colored viruses

in the leaves represent the isolates sampled from simians and humans. Scale bar indicates the branch length. Bootstrap values >50% are shown, with 80% noted as robust and significant.

Table S1. Inventory of adenoviruses contained within the genus *Mastadenovirus* used in this study. All genomes, available from GenBank, are noted along with their accession numbers, host from which the virus was isolated, year of isolation, and the country in which the virus was isolated. Simian adenoviruses, in many cases, were isolated from captive populations, including zoos and primate colonies established for medical and scientific research.

Table S2. Variable genes that may define select PrAdV clades, including PrAdV-A, B, C, D, E, and H. Each gene and protein with character informative changes are aligned to locate and confirm the SNPs for each clade. All synonymous and nonsynonymous SNPs are recorded based on the nucleotide and amino acid positions in the gene and gene product. These SNPs are conserved within individual clades which may serve as a potential indicator for a common ancestor.

Table S3. Full list of variable character changes observed that define the lineages in PrAdV phylogenetic tree. SNP variants can be described using variable gene changes for different subgroups along with the CI score.

CHAPTER 6: Conclusions

The phylogenetic relationships between human (HAdV) and simian adenoviruses (SAdV) have been investigated and clarified using high-resolution genome-based phylogenomic analyses. As a result, a comprehensive overview of the historical and current importance of adenovirus is presented, in particular with respect to the molecular evolution of these viral pathogens, with a re-examination of zoonosis as a basis for one of the oldest HAdVs (Type 4). Chapter 2 reviewed the current status and understanding of human adenoviruses typing in the era of genomics and bioinformatics technology. Chapters 3 and 4 presented examples of putative zoonotic and anthroponosis pathogen transfers between human and non-human simian hosts, as HAdV-E4 and SAdV-21. These studies examined the distribution of recent HAdV-E4 variants in outbreaks that appeared in the civilian population; previously, HAdV-E4 appeared limited to military basic trainee populations. HAdV-E4 variants were examined using data collected from three outbreaks in Singapore, Hong Kong, and Beijing. In the context of data published earlier by this laboratory, the inverted terminal repeats, which contain the host-interacting viral replication functions, was explored. To support the earlier work, the putative host adaptation involving a recombination event that incorporates an HAdV replication motif, the Nuclear Factor 1 binding site (NF-1) was also present in the variants examined. Again, all HAdV respiratory pathogens have this motif that is absent in known SAdVs as

3239 well as the prototype HAdV-E4. It is likely required for optimal adenoviral replication in
3240 human cells, and complements the earlier in vitro studies reported in the literature.
3241 Additionally, SAdV-21, a chimpanzee virus that clusters with the HAdV species B clade,
3242 was analyzed and shown to contain recombinant genomic regions from HAdV-B21,
3243 HAdV-B50, SAdV-27.1, and SAdV-35.1. To further examine the relationships between
3244 HAdV and SAdV, a comparahensive study was conducted to analyze all published
3245 HAdV and SAdV types. Our results demonstrated that HAdVs and SAdVs are
3246 phylogenomically related and may cross-species barriers to infect hosts.

APPENDIX

```
#R code for the nucleotide percent identity heatmap program that is discussed in this
#dissertaion:
#Heatdiagram for taxonomy
install.packages("gplots")
library(gplots)
library(RColorBrewer)

nba <-read.table(
  file = "/Users/junekang/Documents/HK_iden.txt",   ###change filenames here
  header =TRUE,              #add x and y labels
  check.names=FALSE)         #Same name as input

dst <- data.matrix(nba)      #change data into a matrix format

#heatmap
heatmap.2(dst,
          cellnote = dst,              # same data set for cell labels
          #main = "test1",              # heat map title
```

```
3266            notecol="black",              # change font color of cell labels to black
3267            density.info="none",          # turns off density plot inside color legend
3268            trace="none",                 # turns off trace lines inside the heat map
3269            margins =c(9,10),             # widens margins around plot
3270            #col=rev(cm.colors(256)),      #color pallets
3271            col=brewer.pal(8,"Blues"),    # use one color palette defined earlier
3272            #breaks=col_break,            # enable color transition at specified limits
3273            dendrogram="row")             # only draw a row dendrogram
3274
3275    #density graph
3276    m<-mean(dst)
3277    d=density(dst)
3278    plot(d, col="blue", lwd=1, main =NA)   #plot density
3279    polygon(d, col="darkblue")            #add color
3280    abline(v=m, col="red",                #red mean line
3281           lwd=3,                         #how thick the line
3282           lty=3)                         #dashline
3283
3284    #stat
3285    min(dst)   #minimum
3286    max(dst)   #maximum
3287    m          #mean
```

```
3288 sd(dst)    #Standard deviation
3289
3290 #Perl code for extracting fasta files from NCBI database
3291 #Need to have Bioperl installed
3292 #!/usr/bin/perl
3293
3294 use strict;
3295 use warnings;
3296 use Bio::Seq;
3297 use Bio::DB::GenBank;
3298 # File: getgb.pl
3299
3300 # get GI number from command line
3301 my ($id) = @ARGV;
3302
3303 # create a new GenBank handle object
3304 my $gb = new Bio::DB::GenBank();
3305
3306 # Fetch the sequence from GenBank, returning a seq_object
3307 my $seq_obj = $gb->get_Seq_by_id($id);
3308
3309 # print out parts of the sequence
```

```
#print "Id = ", $id, "\n",
#    "Accession = ", $seq_obj->accession_number, "\n",
#    "Description = ", $seq_obj->desc, "\n",
#    "Sequence = ", $seq_obj->seq, "\n";

write_sequence(">tra.fasta",'fasta',$seq_obj);

#Perl code for searching motifs
        #!/usr/bin/perl
        use strict;
        use warnings;

        #Author: June Kang
        #This is used to match motifs in a multiseq fasta file. To use this code, must
        #input the motif patters.
        #After a output file is generated from this program
        #Run the code unique at command line to delete duplications
        #example: uniq report.txt >newfile.txt


        #ITR motif categories (wild)
```

```
3331            my $mp1 = 'AATATACCTTATTTTGGCTGCACGCCAATATGATAATGA';
3332    #CynAdV-1 motif pattern
3333            my $mp2 =
3334    'AATATACCTCAAACTTTTGGTGCGCGTTAATATGCAAATGA'; #SAdV-25.2
3335            #my $mp3 = 'AATATACCTTATTTTGGCTGCACGCCAATATGATAATGA';
3336    #BAdV-1
3337            my $mp4 =
3338    'GAGAGAGGAGGAGTGGGGGTGGCAGGGGGTGGGAAGAAGTGACG';#BAdV-
3339    2,4
3340            my $mp5 =
3341    'AATATAACACCGCAAGATTTTGGCATAGATAAGAAGTTAATGA'; #SAdV-
3342    WIV19
3343            my $mp6 = 'AATATACCTTATTCTGGAAACGTGCCAATATGATAATGA';
3344    #BAdV-3
3345
3346
3347
3348
3349            print "Please enter filename: ";
3350            my $sequence = <STDIN>;
3351
3352            #chomp $sequence;
```

```
3353            unless (open(SEQUENCE, $sequence)){print "Cannot open file
3354    \"$sequence\"\n\n"; exit;}
3355
3356            my $filename ='report.txt';    #open a file before write in
3357            open (my$fh, '>', $filename) or die "Can't open file $filename \n";
3358
3359
3360            my @header = ();     #initilize
3361            my@sequence =();
3362            my $count = 0;
3363            my $n= -1;
3364            while (my $line = <SEQUENCE>){
3365              chomp$line;
3366              if($line =~ /^>/){
3367                $n++;
3368                $header[$n]= $line;
3369                $sequence[$n] = "";
3370              }
3371              else{
3372                next if not @header;
3373                $sequence[$n] .=$line;
3374                #print "$header[$n]\n\n";
```

```
3375                #print "$sequence[$n]\n\n";
3376
3377                if($sequence[$n] =~/$mp1/){
3378                  print $fh "$header[$n] match mp1 $mp1 \n";
3379                }elsif($sequence[$n] =~/$mp2/){
3380                  print $fh "$header[$n] match mp2 $mp2 \n";
3381              #  }elsif($sequence[$n] =~/$mp3/){
3382              #    print $fh "$header[$n] match mp3 $mp3 \n";
3383                }elsif($sequence[$n] =~/$mp4/){
3384                  print $fh "$header[$n] match mp4 $mp4 \n";
3385                }elsif($sequence[$n] =~/$mp5/){
3386                  print $fh "$header[$n] match mp5 $mp5 \n";
3387                }elsif($sequence[$n] =~/$mp6/){
3388                  print $fh "$header[$n] match mp6 $mp6 \n";
3389                }
3390                else{
3391                  print $fh "$header[$n] don't match any known motif \n";
3392
3393                }
3394                }
3395              }
3396
```

```
3397                $count=$n+1;
3398                close SEQUENCE;
3399
3400
3401            exit;
3402            #read the sequence data in and store it
3403            #@sequence = <SEQUENCE>;
3404            #close SEQUENCE;
3405
3406
3407            #do{
3408            #  if ($sequence =~/$mf1/){
3409            #    print
3410            #  }
3411    #}
3412
3413
```

REFERENCES

Aoki, K., Kato, M., Ohtsuka, H., Ishii, K., Nakazono, N., & Sawada, H. (1982). Clinical and aetiological study of adenoviral conjunctivitis, with special reference to adenovirus types 4 and 19 infections. *British Journal of Ophthalmology*, *66*(12), 776. https://doi.org/10.1136/bjo.66.12.776

Arnold, J., Jánoska, M., Kajon, A. E., Metzgar, D., Hudson, N. R., Torres, S., Harrach, B., Seto, D., Chodosh, J., & Jones, M. S. (2010). Genomic characterization of human adenovirus 36, a putative obesity agent. *Virus Research*, *149*(2), 152–161. https://doi.org/10.1016/j.virusres.2010.01.011

Arrand, J. R., & Roberts, R. J. (1979). The nucleotide sequences at the termini of adenovirus-2 DNA. *Journal of Molecular Biology*, *128*(4), 577–594. https://doi.org/10.1016/0022-2836(79)90294-8

Bakay, M., Szalay, K., Beladi, I., Balint, E., Lengyel, A., Adam, E., & Olah, T. (2005). Cross-reactivity between human adenoviruses in delayed-type hypersensitivity. *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, *113*(3), 197–202. https://doi.org/10.1111/j.1600-0463.2005.apm1130307.x

Barraza, E. M., Ludwig, S. L., Gaydos, J. C., & Brundage, J. F. (1999). Reemergence of adenovirus type 4 acute respiratory disease in military trainees: Report of an

outbreak during a lapse in vaccination. *The Journal of Infectious Diseases*, *179 6*, 1531–1533.

Blasiole, D. A., Metzgar, D., Daum, L. T., Ryan, M. A. K., Wu, J., Wills, C., Le, C. T., Freed, N. E., Hansen, C. J., Gray, G. C., & Russell, K. L. (2004). Molecular analysis of adenovirus isolates from vaccinated and unvaccinated young adults. *Journal of Clinical Microbiology*, *42*(4), 1686–1693. https://doi.org/10.1128/jcm.42.4.1686-1693.2004

Carballal, G., Videla, C. M., Espinosa, M. A., Savy, V., Uez, O., Sequeira, M. D., Knez, V., Requeijo, P. V., Posse, C. R., & Miceli, I. (2001). Multicentered study of viral acute lower respiratory infections in children from four cities of Argentina, 1993-1994. *Journal of Medical Virology*, *64*(2), 167–174. https://doi.org/10.1002/jmv.1032

Cersovsky, S., LN, B., FD, M., SA, K., CS, P., Sanchez Jr, J., & WT, B. (1999). A prospective study of acute respiratory diseases (ARD) in a cohort of US Army basic trainees, October-November 1998, Fort Jackson, South Carolina. *Medical Surveillance Monthly Report*, *5*, 10, 12.

Chang, S.-Y., Lee, C.-N., Lin, P.-H., Huang, H.-H., Chang, L.-Y., Ko, W., Chang, S.-F., Lee, P.-I., Huang, L.-M., & Kao, C.-L. (2008). A community-derived outbreak of adenovirus type 3 in children in Taiwan between 2004 and 2005. *Journal of Medical Virology*, *80*(1), 102–112. https://doi.org/10.1002/jmv.21045

Chen, E. C., Yagi, S., Kelly, K. R., Mendoza, S. P., Tarara, R. P., Canfield, D. R., Maninger, N., Rosenthal, A., Spinner, A., Bales, K. L., Schnurr, D. P., Lerche, N.

W., & Chiu, C. Y. (2011a). Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. *PLoS Pathogens*, *7*(7), e1002155. https://doi.org/10.1371/journal.ppat.1002155

Chen, E. C., Yagi, S., Kelly, K. R., Mendoza, S. P., Tarara, R. P., Canfield, D. R., Maninger, N., Rosenthal, A., Spinner, A., Bales, K. L., Schnurr, D. P., Lerche, N. W., & Chiu, C. Y. (2011b). Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. *PLoS Pathogens*, *7*(7), e1002155. https://doi.org/10.1371/journal.ppat.1002155

Chiu, C. Y., Yagi, S., Lu, X., Yu, G., Chen, E. C., Liu, M., Dick, E. J., Carey, K. D., Erdman, D. D., Leland, M. M., & Patterson, J. L. (2013a). A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. *MBio*, *4*(2), e00084. https://doi.org/10.1128/mBio.00084-13

Chiu, C. Y., Yagi, S., Lu, X., Yu, G., Chen, E. C., Liu, M., Dick, E. J. J., Carey, K. D., Erdman, D. D., Leland, M. M., & Patterson, J. L. (2013b). A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. *MBio*, *4*(2), e00084. https://doi.org/10.1128/mBio.00084-13

Chu, W., & Pavan-Langston, E. (1979). Ocular Surfce Manifestations of the Major Viruses. *International Ophthalmology Clinics*, *19*(2). https://journals.lww.com/internat-

ophthalmology/Fulltext/1979/01920/Ocular_Surfce_Manifestations_of_the_Major_Viruses.10.aspx

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*(11), 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

Coleman, K. K., Wong, C. C., Jayakumar, J., Nguyen, T. T., Wong, A. W. L., Yadana, S., Thoon, K. C., Chan, K. P., Low, J. G., Kalimuddin, S., Dehghan, S., Kang, J., Shamsaddini, A., Seto, D., Su, Y. C. F., & Gray, G. C. (2019). Adenoviral Infections in Singapore: Should New Antiviral Therapies and Vaccines Be Adopted? *The Journal of Infectious Diseases*, *jiz489*. https://doi.org/10.1093/infdis/jiz489

Couch, R. B., Chanock, R. M., Cate, T. R., Lang, D. J., Knight, V., & Huebner, R. J. (1963). Immunization with types4 and 7 adenovirus by selective infection of the intestinal tract. *The American Review of Respiratory Disease*, *88*, SUPPL 394-403. https://doi.org/10.1164/arrd.1963.88.3P2.394

D'Angelo, L. J., Hierholzer, J. C., Keenlyside, R. A., Anderson, L. J., & Martone, W. J. (1979). Pharyngoconjunctival fever caused by adenovirus type 4: Report of a swimming pool-related outbreak with recovery of virus from pool water. *The Journal of Infectious Diseases*, *140*(1), 42–47. https://doi.org/10.1093/infdis/140.1.42

Davison, A. J., Benko, M., & Harrach, B. (2003). Genetic content and evolution of adenoviruses. *The Journal of General Virology*, *84*(Pt 11), 2895–2908. https://doi.org/10.1099/vir.0.19497-0

Dehghan, S., Seto, J., Jones, M. S., Dyer, D. W., Chodosh, J., & Seto, D. (2013). Simian adenovirus type 35 has a recombinant genome comprising human and simian adenovirus sequences, which predicts its potential emergence as a human respiratory pathogen. *Virology*, *447*(1–2), 265–273. https://doi.org/10.1016/j.virol.2013.09.009

Dehghan, S., Seto, J., Liu, E. B., Ismail, A. M., Madupu, R., Heim, A., Jones, M. S., Dyer, D. W., Chodosh, J., & Seto, D. (2019). A Zoonotic Adenoviral Human Pathogen Emerged through Genomic Recombination among Human and Nonhuman Simian Hosts. *Journal of Virology*, *93*(18). https://doi.org/10.1128/JVI.00564-19

Dehghan, S., Seto, J., Liu, E. B., Walsh, M. P., Dyer, D. W., Chodosh, J., & Seto, D. (2013). Computational analysis of four human adenovirus type 4 genomes reveals molecular evolution through two interspecies recombination events. *Virology*, *443*(2), 197–207. https://doi.org/10.1016/j.virol.2013.05.014

Dudding, B. A., Top, F. H. J., Winter, P. E., Buescher, E. L., Lamson, T. H., & Leibovitz, A. (1973). Acute respiratory disease in military trainees: The adenovirus surveillance program, 1966-1971. *American Journal of Epidemiology*, *97*(3), 187–198. https://doi.org/10.1093/oxfordjournals.aje.a121499

DuVernoy, T., Coldren, R., Feighner, B., Jordan, N., Sanchez, J., Alsip, B., Gonsalez, R., Townsend, R., Binn, L., & Williams, S. (2000). Investigation of an Acute Respiratory Disease Outbreak due to Adenovirus type 4 among Recruits, Fort Benning, Georgia, April-May 2000. *US Army Med Surveill Monthly Rep*, *6*, 2–7.

ELNIFRO, E., COOPER, R., KLAPPER, P., BAILEY, A., & TULLO, A. (1999). Diagnosis of viral and chlamydial keratoconjunctivitis: Which laboratory test? *The British Journal of Ophthalmology*, *83*(5), 622–627.

Engelmann, I., Madisch, I., Pommer, H., & Heim, A. (2006). An Outbreak of Epidemic Keratoconjunctivitis Caused by a New Intermediate Adenovirus 22/H8 Identified by Molecular Typing. *Clinical Infectious Diseases*, *43*(7), e64–e66. https://doi.org/10.1086/507533

Feldmann, H. (2014). Ebola—A Growing Threat? *New England Journal of Medicine*, *371*(15), 1375–1378. https://doi.org/10.1056/NEJMp1405314

Fierz, W. (2004). Basic problems of serological laboratory diagnosis. *Methods in Molecular Medicine*, *94*, 393–427.

Flint, J. (1999). Organization of the adenoviral genome. In *In Adenoviruses: Basic Biology to Gene Therapy* (pp. 17–30). Landes Company.

Fontaine, M. P., Chabas, D., Fontaine, M., & Brion, A. J. (1963). A Simplified Technique for the Serum-Neutralization Test in Infectious Bronchitis. *Avian Diseases*, *7*(2), 203–206. JSTOR. https://doi.org/10.2307/1588052

Fox, J. P., Hall, C. E., & Cooney, M. K. (1977). The Seattle Virus Watch. VII. Observations of adenovirus infections. *American Journal of Epidemiology*, *105*(4), 362–386.

Friedman, M., & Bennett, C. R. (1957). A Hemagglutination Test for Detection of Adenovirus Antibodies. *Proceedings of the Society for Experimental Biology and Medicine*, *94*(4), 712–717. https://doi.org/10.3181/00379727-94-23062

Garnett, C. T., Talekar, G., Mahr, J. A., Huang, W., Zhang, Y., Ornelles, D. A., & Gooding, L. R. (2009). Latent species C adenoviruses in human tonsil tissues. *Journal of Virology*, *83*(6), 2417–2428. PubMed. https://doi.org/10.1128/JVI.02392-08

Gauger, P. C., & Vincent, A. L. (2014). Serum virus neutralization assay for detection and quantitation of serum-neutralizing antibodies to influenza A virus in swine. *Methods in Molecular Biology (Clifton, N.J.)*, *1161*, 313–324. https://doi.org/10.1007/978-1-4939-0758-8_26

Gaydos, C. A., & Gaydos, J. C. (1995). Adenovirus vaccines in the U.S. military. *Military Medicine*, *160*(6), 300–304.

Gebreyes, W. A., Dupouy-Camet, J., Newport, M. J., Oliveira, C. J. B., Schlesinger, L. S., Saif, Y. M., Kariuki, S., Saif, L. J., Saville, W., Wittum, T., Hoet, A., Quessy, S., Kazwala, R., Tekola, B., Shryock, T., Bisesi, M., Patchanee, P., Boonmar, S., & King, L. J. (2014). The global one health paradigm: Challenges and opportunities for tackling infectious diseases at the human, animal, and

environment interface in low-resource settings. *PLoS Neglected Tropical Diseases*, *8*(11), e3257. https://doi.org/10.1371/journal.pntd.0003257

Gerber, S. I., Erdman, D. D., Pur, S. L., Diaz, P. S., Segreti, J., Kajon, A. E., Belkengren, R. P., & Jones, R. C. (2001). Outbreak of adenovirus genome type 7d2 infection in a pediatric chronic-care facility and tertiary-care hospital. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *32*(5), 694–700. https://doi.org/10.1086/319210

Graham, F. L., & Prevec, L. (1992). Adenovirus-based expression vectors and recombinant vaccines. *Biotechnology (Reading, Mass.)*, *20*, 363–390.

Gray, G. C., McCarthy, T., Lebeck, M. G., Schnurr, D. P., Russell, K. L., Kajon, A. E., Landry, M. L., Leland, D. S., Storch, G. A., Ginocchio, C. C., Robinson, C. C., Demmler, G. J., Saubolle, M. A., Kehl, S. C., Selvarangan, R., Miller, M. B., Chappell, J. D., Zerr, D. M., Kiska, D. L., … Erdman, D. D. (2007). Genotype prevalence and risk factors for severe clinical adenovirus infection, United States 2004-2006. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *45*(9), 1120–1131. https://doi.org/10.1086/522188

Grodzicker, T., Anderson, C., Sharp, P. A., & Sambrook, J. (1974). Conditional Lethal Mutants of Adenovirus 2-Simian Virus 40 Hybrids I. Host Range Mutants of Ad2+ND1. *Journal of Virology*, *13*(6), 1237–1244.

Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J., Guan, Y. J., Butt, K. M., Wong, K. L., Chan, K. W., Lim, W., Shortridge, K. F., Yuen, K. Y., Peiris, J. S. M., & Poon, L. L. M. (2003).

Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China. *Science*, *302*(5643), 276–278. https://doi.org/10.1126/science.1087139

Harrach, B., Benkö, M., Both, G. W., Brown, M., Davison, A. J., Echavarria, M., Hess, M., Jones, M. S., Kajon, A., Lehmkuhl, H. D., Mautner, V., Mittal, S. K., & Wadell, G. (2011). *Virus Taxonomy: Classification and Nomenclature of Viruses*. 125–141.

Hay, R. T., & Russell, W. C. (1989). Recognition mechanisms in the synthesis of animal virus DNA. *Biochemical Journal*, *258*(1), 3–16. https://doi.org/10.1042/bj2580003

Heemskerk, B., Veltrop-Duits, L. A., van Vreeswijk, T., ten Dam, M. M., Heidt, S., Toes, R. E. M., van Tol, M. J. D., & Schilham, M. W. (2003). Extensive cross-reactivity of CD4+ adenovirus-specific T cells: Implications for immunotherapy and gene therapy. *Journal of Virology*, *77*(11), 6562–6566.

Henry, L., & Wickham, H. (2018). *tidyselect: Select from a Set of Strings* [R package version 0.2.5]. https://CRAN.R-project.org/package=tidyselect

Hierholzer, J. C. (1973). Further subgrouping of the human adenoviruses by differential hemagglutination. *The Journal of Infectious Diseases*, *128*(4), 541–550. https://doi.org/10.1093/infdis/128.4.541

Hilleman, M. R., & Werner, J. H. (1954). Recovery of new agent from patients with acute respiratory illness. *Proceedings of the Society for Experimental Biology and*

Medicine. Society for Experimental Biology and Medicine (New York, N.Y.), *85*(1), 183–188. https://doi.org/10.3181/00379727-85-20825

Holz, C. L., Cibulski, S. P., Teixeira, T. F., Batista, H. B. C. R., Dezen, D., Campos, F. S., Varela, A. P. M., & Roehe, P. M. (2010). Serum neutralization with different types and subtypes of bovine herpesvirus 1 and 5. *Pesquisa Veterinária Brasileira*, *30*(7), 515–522. https://doi.org/10.1590/S0100-736X2010000700001

Hong, J.-Y., Lee, H.-J., Piedra, P. A., Choi, E.-H., Park, K.-H., Koh, Y.-Y., & Kim, W.-S. (2001). Lower Respiratory Tract Infections due to Adenovirus in Hospitalized Korean Children: Epidemiology, Clinical Features, and Prognosis. *Clinical Infectious Diseases*, *32*(10), 1423–1429. https://doi.org/10.1086/320146

Houldcroft, C. J., Roy, S., Morfopoulou, S., Margetts, B. K., Depledge, D. P., Cudini, J., Shah, D., Brown, J. R., Romero, E. Y., Williams, R., Cloutman-Green, E., Rao, K., Standing, J. F., Hartley, J. C., & Breuer, J. (2018). Use of Whole-Genome Sequencing of Adenovirus in Immunocompromised Pediatric Patients to Identify Nosocomial Transmission and Mixed-Genotype Infection. *The Journal of Infectious Diseases*, *218*(8), 1261–1271. https://doi.org/10.1093/infdis/jiy323

Huang, G.-H., & Xu, W.-B. (2013). [Recent advance in new types of human adenovirus]. *Bing Du Xue Bao = Chinese Journal of Virology*, *29*(3), 342–348.

Jones, M. S., Harrach, B., Ganac, R. D., Gozum, M. M. A., Cruz, W. P. dela, Riedel, B., Pan, C., Delwart, E. L., & Schnurr, D. P. (2007). New Adenovirus Species Found in a Patient Presenting with Gastroenteritis. *Journal of Virology*, *81*(11), 5978–5984. https://doi.org/10.1128/JVI.02650-06

Jones, M. S., Harrach, B., Ganac, R. D., Gozum, M. M. A., dela Cruz, W. P., Riedel, B., Pan, C., Delwart, E. L., & Schnurr, D. P. (2007). New Adenovirus Species Found in a Patient Presenting with Gastroenteritis. *Journal of Virology*, *81*(11), 5978. https://doi.org/10.1128/JVI.02650-06

Kajon, A. E., Lamson, D. M., Bair, C. R., Lu, X., Landry, M. L., Menegus, M., Erdman, D. D., & St. George, K. (2018). Adenovirus Type 4 Respiratory Infections among Civilian Adults, Northeastern United States, 2011–20151. *Emerging Infectious Diseases*, *24*(2), 201–209. https://doi.org/10.3201/eid2402.171407

Kajon, A. E., Moseley, J. M., Metzgar, D., Huong, H.-S., Wadleigh, A., Ryan, M. A. K., & Russell, K. L. (2007). Major Articles-Molecular Epidemiology of Adenovirus Type 4 Infections in US Military Recruits in the Postvaccination Era (1997–2003). *The Journal of Infectious Diseases*, *196*(1), 67–75. https://doi.org/10.1086/518442

Kandel, R., D'Agata, E., Lu, X., Erdman, D., & MD, MPH, Michael. (2010). Outbreak of Adenovirus Type 4 Infection in a Long-Term Care Facility for the Elderly. *Infection Control and Hospital Epidemiology: The Official Journal of the Society of Hospital Epidemiologists of America*, *31*, 755–757. https://doi.org/10.1086/653612

Kang, J., Ismail, A. M., Dehghan, S., Rajaiya, J., Allard, M. W., Lim, H. C., Dyer, D. W., Chodosh, J., & Seto, D. (2020). Genomics-based re-examination of the taxonomy and phylogeny of human and simian Mastadenoviruses: An evolving whole

genomes approach, revealing putative zoonosis, anthroponosis, and amphizoonosis. *Cladistics*, *36*(4), 358–373. https://doi.org/10.1111/cla.12422

Klasse, P. J. (2014). Neutralization of Virus Infectivity by Antibodies: Old Problems in New Perspectives. *Advances in Biology*, *2014*, 157895. https://doi.org/10.1155/2014/157895

Koyuncu, O. O., & Dobner, T. (2009). Arginine methylation of human adenovirus type 5 L4 100-kilodalton protein is required for efficient virus production. *Journal of Virology*, *83*(10), 4778–4790. https://doi.org/10.1128/JVI.02493-08

Larrañaga, C., Kajon, A., Villagra, E., & Avendaño, L. F. (2000). Adenovirus surveillance on children hospitalized for acute lower respiratory infections in Chile (1988-1996). *Journal of Medical Virology*, *60*(3), 342–346.

Lau, S. K. P., Woo, P. C. Y., Li, K. S. M., Huang, Y., Tsoi, H.-W., Wong, B. H. L., Wong, S. S. Y., Leung, S.-Y., Chan, K.-H., & Yuen, K.-Y. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 14040–14045. https://doi.org/10.1073/pnas.0506735102

Li, J., Lu, X., Sun, Y., Lin, C., Li, F., Yang, Y., Liang, Z., Jia, L., Chen, L., Jiang, B., & Wang, Q. (2018). A swimming pool-associated outbreak of pharyngoconjunctival fever caused by human adenovirus type 4 in Beijing, China. *International Journal of Infectious Diseases*, *75*, 89–91. https://doi.org/10.1016/j.ijid.2018.08.009

Lion, T. (2014). Adenovirus infections in immunocompetent and immunocompromised patients. *Clinical Microbiology Reviews*, *27*(3), 441–462. https://doi.org/10.1128/CMR.00116-13

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., & 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), 152–160.

Lu, L., Liu, Q., Du, L., & Jiang, S. (2013). Middle East respiratory syndrome coronavirus (MERS-CoV): Challenges in identifying its source and controlling its spread. *Microbes and Infection*, *15*(8–9), 625–629. https://doi.org/10.1016/j.micinf.2013.06.003

Mahr, J. A., & Gooding, L. R. (1999). Immune evasion by adenoviruses. *Immunological Reviews*, *168*, 121–130.

Malouli, D., Howell, G. L., Legasse, A. W., Kahl, C., Axthelm, M. K., Hansen, S. G., & Früh, K. (2014). Full genome sequence analysis of a novel adenovirus of rhesus macaque origin indicates a new simian adenovirus type and species. *Virology Reports*, *3–4*, 18–29. PubMed. https://doi.org/10.1016/j.virep.2014.10.001

Mautner, V., Williams, J., Sambrook, J., Sharp, P. A., & Grodzicker, T. (1975). The location of the genes coding for hexon and fiber proteins in adenovirus DNA. *Cell*, *5*(1), 93–99.

Mul, Y. M., & Van der Vliet, P. C. (1992). Nuclear factor I enhances adenovirus DNA replication by increasing the stability of a preinitiation complex. *The EMBO Journal*, *11*(2), 751–760. PubMed.

Norrby, E. (1969). The structural and functional diversity of Adenovirus capsid components. *The Journal of General Virology*, *5*(2), 221–236. https://doi.org/10.1099/0022-1317-5-2-221

Olive, M., Eisenlohr, L., Flomenberg, N., Hsu, S., & Flomenberg, P. (2002). The adenovirus capsid protein hexon contains a highly conserved human CD4+ T-cell epitope. *Human Gene Therapy*, *13*(10), 1167–1178. https://doi.org/10.1089/104303402320138952

Paula, D., Oliveira, S., & Fonseca, B. A. L. da. (2004). Dengue: A review of the laboratory tests a clinician must know to achieve a correct diagnosis. *Brazilian Journal of Infectious Diseases*, *8*(6), 390–398. https://doi.org/10.1590/S1413-86702004000600002

Pichla-Gollon, S. L., Drinker, M., Zhou, X., Xue, F., Rux, J. J., Gao, G.-P., Wilson, J. M., Ertl, H. C. J., Burnett, R. M., & Bergelson, J. M. (2007). Structure-Based Identification of a Major Neutralizing Site in an Adenovirus Hexon. *Journal of Virology*, *81*(4), 1680. https://doi.org/10.1128/JVI.02023-06

Purkayastha, A., Ditty, S. E., Su, J., McGraw, J., Hadfield, T. L., Tibbetts, C., & Seto, D. (2005). Genomic and Bioinformatics Analysis of HAdV-4, a Human Adenovirus Causing Acute Respiratory Disease: Implications for Gene Therapy and Vaccine

Vector Development. *Journal of Virology*, *79*(4), 2559. https://doi.org/10.1128/JVI.79.4.2559-2572.2005

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Rajkumar, V., Chiang, C. S. M., Low, J. M., Cui, L., Lin, R. T. P., Tee, N. W. S., Maiwald, M., Chong, C. Y., Thoon, K. C., & Tan, N. W. H. (2015). Risk Factors for Severe Adenovirus Infection in Children during an Outbreak in Singapore. *Annals of the Academy of Medicine, Singapore*, *44*(2), 50–59.

Reddy, P. S., Idamakanti, N., Zakhartchouk, A. N., Baxi, M. K., Lee, J. B., Pyne, C., Babiuk, L. A., & Tikoo, S. K. (1998). Nucleotide sequence, genome organization, and transcription map of bovine adenovirus type 3. *Journal of Virology*, *72*(2), 1394–1402. PubMed. https://doi.org/10.1128/JVI.72.2.1394-1402.1998

Robinson, C. M., Singh, G., Henquell, C., Walsh, M. P., Peigue-Lafeuille, H., Seto, D., Jones, M. S., Dyer, D. W., & Chodosh, J. (2011). Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality. *Virology*, *409*(2), 141–147. https://doi.org/10.1016/j.virol.2010.10.020

Robinson, C. M., Singh, G., Lee, J. Y., Dehghan, S., Rajaiya, J., Liu, E. B., Yousuf, M. A., Betensky, R. A., Jones, M. S., Dyer, D. W., Seto, D., & Chodosh, J. (2013). Molecular evolution of human adenoviruses. *Scientific Reports*, *3*, 1812. https://doi.org/10.1038/srep01812

Rosen, L. (1960). A hemagglutination-inhibition technique for typing adenoviruses. *American Journal of Hygiene*, *71*, 120–128. https://doi.org/10.1093/oxfordjournals.aje.a120085

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., & Ward, T. G. (1953a). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *84*(3), 570–573. https://doi.org/10.3181/00379727-84-20714

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., & Ward, T. G. (1953b). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *84*(3), 570–573.

Roy, S., Gao, G., Lu, Y., Zhou, X., Lock, M., Calcedo, R., & Wilson, J. M. (2004). Characterization of a family of chimpanzee adenoviruses and development of molecular clones for gene transfer vectors. *Human Gene Therapy*, *15*(5), 519–530. https://doi.org/10.1089/10430340460745838

Roy, S., Sandhu, A., Medina, A., Clawson, D. S., & Wilson, J. M. (2012). Adenoviruses in fecal samples from asymptomatic rhesus macaques, United States. *Emerging Infectious Diseases*, *18*(7), 1081–1088. PubMed. https://doi.org/10.3201/eid1807.111665

Roy, S., Vandenberghe, L. H., Kryazhimskiy, S., Grant, R., Calcedo, R., Yuan, X., Keough, M., Sandhu, A., Wang, Q., Medina-Jaszek, C. A., Plotkin, J. B., & Wilson, J. M. (2009). Isolation and Characterization of Adenoviruses Persistently Shed from the Gastrointestinal Tract of Non-Human Primates. *PLOS Pathogens*, *5*(7), e1000503. https://doi.org/10.1371/journal.ppat.1000503

Roy, S., Zhi, Y., Kobinger, G. P., Figueredo, J., Calcedo, R., Miller, J. R., Feldmann, H., & Wilson, J. M. (2006). Generation of an adenoviral vaccine vector based on simian adenovirus 21. *The Journal of General Virology*, *87*(Pt 9), 2477–2485. https://doi.org/10.1099/vir.0.81989-0

Schmitz, H., Wigand, R., & Heinrich, W. (1983a). Worldwide epidemiology of human adenovirus infections. *American Journal of Epidemiology*, *117*(4), 455–466.

Schmitz, H., Wigand, R., & Heinrich, W. (1983b). Worldwide epidemiology of human adenovirus infections. *American Journal of Epidemiology*, *117*(4), 455–466. https://doi.org/10.1093/oxfordjournals.aje.a113563

Scott, M. K., Chommanard, C., Lu, X., Appelgate, D., Grenz, L., Schneider, E., Gerber, S. I., Erdman, D. D., & Thomas, A. (2016). Human Adenovirus Associated with Severe Respiratory Infection, Oregon, USA, 2013-2014. *Emerging Infectious Diseases*, *22*(6), 1044–1051. PubMed. https://doi.org/10.3201/eid2206.151898

Seto, D., Chodosh, J., Brister, J. R., & Jones, M. S. (2011). Using the whole-genome sequence to characterize and name human adenoviruses. *Journal of Virology*, *85*(11), 5701–5702. https://doi.org/10.1128/JVI.00354-11

Seto, D., Jones, M. S., Dyer, D. W., & Chodosh, J. (2013). Characterizing, typing, and naming human adenovirus type 55 in the era of whole genome data. *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, *58*(4), 741–742. https://doi.org/10.1016/j.jcv.2013.09.025

Shayakhmetov, D. M., Gaggar, A., Ni, S., Li, Z.-Y., & Lieber, A. (2005). Adenovirus binding to blood factors results in liver cell infection and hepatotoxicity. *Journal of Virology*, *79*(12), 7478–7491. PubMed. https://doi.org/10.1128/JVI.79.12.7478-7491.2005

Singh, G., Zhou, X., Lee, J. Y., Yousuf, M. A., Ramke, M., Ismail, A. M., Lee, J. S., Robinson, C. M., Seto, D., Dyer, D. W., Jones, M. S., Rajaiya, J., & Chodosh, J. (2015). Recombination of the epsilon determinant and corneal tropism: Human adenovirus species D types 15, 29, 56, and 69. *Virology*, *485*, 452–459. PubMed. https://doi.org/10.1016/j.virol.2015.08.018

Singh-Naz, N., & Rodriguez, W. (1996). Adenoviral infections in children. *Advances in Pediatric Infectious Diseases*, *11*, 365–388.

Smart, J. E., & Stillman, B. W. (1982). Adenovirus terminal protein precursor. Partial amino acid sequence and the site of covalent linkage to virus DNA. *The Journal of Biological Chemistry*, *257*(22), 13499–13506.

Smith, C. A., Woodruff, L. S., Rooney, C., & Kitchingman, G. R. (1998). Extensive cross-reactivity of adenovirus-specific cytotoxic T cells. *Human Gene Therapy*, *9*(10), 1419–1427. https://doi.org/10.1089/hum.1998.9.10-1419

Stillman, B. (1989). Initiation of eukaryotic DNA replication in vitro. *Annual Review of Cell Biology*, *5*, 197–245. https://doi.org/10.1146/annurev.cb.05.110189.001213

Stuiver, M. H., & van der Vliet, P. C. (1990). Adenovirus DNA-binding protein forms a multimeric protein complex with double-stranded DNA and enhances binding of nuclear factor I. *Journal of Virology*, *64*(1), 379–386. https://doi.org/10.1128/JVI.64.1.379-386.1990

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(vey016). https://doi.org/10.1093/ve/vey016

Svensson, L., Wadell, G., Uhnoo, I., Johansson, M., & Von Bonsdorff, C.-H. (1983). Cross-reactivity between Enteric Adenoviruses and Adenovirus Type 4: Analysis of Epitopes by Solid-phase Immune Electron Microscopy. In *Journal of General Virology,* (Vol. 64, Issue 11, pp. 2517–2520). Microbiology Society,.

Swenson, P., Wadell, G., Allard, A., & Hierholzer, J. (2003). *Manual of Clinical Microbiology: Vol. 2.8*. ASM Press.

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1411), 983–989. https://doi.org/10.1098/rstb.2001.0888

Tokunaga, O., Shinagawa, M., & Padmanabhan, R. (1982). Physical mapping of the genome and sequence analysis at the inverted terminal repetition of adenovirus

type 4 DNA. *Gene*, *18*(3), 329–334. https://doi.org/10.1016/0378-1119(82)90171-8

Tolun, A., Aleström, P., & Pettersson, U. (1979). Sequence of inverted terminal repetitions from different adenoviruses: Demonstration of conserved sequences and homology between SA7 termini and SV40 DNA. *Cell*, *17*(3), 705–713. https://doi.org/10.1016/0092-8674(79)90277-0

Top, F. H., Jr., Grossman, R. A., Bartelloni, P. J., Segal, H. E., Dudding, B. A., Russell, P. K., & Buescher, E. L. (1971a). Immunization with Live Types 7 and 4 Adenovirus Vaccines. I. Safety, Infectivity, Antigenicity, and Potency of Adenovirus Type 7 Vaccine in Humans. *The Journal of Infectious Diseases*, *124*(2), 148–154. https://doi.org/10.1093/infdis/124.2.148

Top, F. H., Jr., Grossman, R. A., Bartelloni, P. J., Segal, H. E., Dudding, B. A., Russell, P. K., & Buescher, E. L. (1971b). Immunization with Live Types 7 and 4 Adenovirus Vaccines. I. Safety, Infectivity, Antigenicity, and Potency of Adenovirus Type 7 Vaccine in Humans. *The Journal of Infectious Diseases*, *124*(2), 148–154. https://doi.org/10.1093/infdis/124.2.148

Towner, J. S., Amman, B. R., Sealy, T. K., Carroll, S. A. R., Comer, J. A., Kemp, A., Swanepoel, R., Paddock, C. D., Balinandi, S., Khristova, M. L., Formenty, P. B. H., Albarino, C. G., Miller, D. M., Reed, Z. D., Kayiwa, J. T., Mills, J. N., Cannon, D. L., Greer, P. W., Byaruhanga, E., … Rollin, P. E. (2009). Isolation of Genetically Diverse Marburg Viruses from Egyptian Fruit Bats. *Plos Pathogens*, *5*(7), e1000536. https://doi.org/10.1371/journal.ppat.1000536

Tsurimoto, T., Fairman, M. P., & Stillman, B. (1989). Simian virus 40 DNA replication in vitro: Identification of multiple stages of initiation. *Molecular and Cellular Biology*, *9*(9), 3839–3849. PubMed. https://doi.org/10.1128/mcb.9.9.3839

Tsuzuki-Wang, L., Aoki, K., Isobe, K., Shiao, S., Toba, K., Kobayashi, N., Noguchi, Y., & Ohno, S. (1997). Genome analysis of adenovirus type 4 strains isolated from acute conjunctivitis in Japan. *Japanese Journal of Ophthalmology*, *41*(5), 308–311. https://doi.org/10.1016/S0021-5155(97)00067-1

Uhnoo, I., Wadell, G., Svensson, L., & Johansson, M. E. (1984). Importance of enteric adenoviruses 40 and 41 in acute gastroenteritis in infants and young children. *Journal of Clinical Microbiology*, *20*(3), 365–372.

van der Vliet, P. C., van Driel, W., & Pruijn, G. J. M. (1986). Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. *Nature (London)*, *322*(6080), 656–659.

Videla, C., Carballal, G., & Kajon, A. (1999). Genomic analysis of adenovirus isolated from Argentinian children with acute lower respiratory infections. *Journal of Clinical Virology : The Official Publication of the Pan American Society for Clinical Virology*, *14*(1), 67–71. https://doi.org/10.1016/s1386-6532(99)00033-5

Vincent, T., Harvey, B. G., Hogan, S. M., Bailey, C. J., Crystal, R. G., & Leopold, P. L. (2001). Rapid assessment of adenovirus serum neutralizing antibody titer based on quantitative, morphometric evaluation of capsid binding and intracellular trafficking: Population analysis of adenovirus capsid association with cells is

predictive of adenovirus infectivity. *Journal of Virology*, *75*(3), 1516–1521. PubMed. https://doi.org/10.1128/JVI.75.3.1516-1521.2001

Wadell, G. (1984). Molecular epidemiology of human adenoviruses. *Current Topics in Microbiology and Immunology*, *110*, 191–220.

Walls, T., Shankar, A. G., & Shingadia, D. (2003). Adenovirus: An increasingly important pathogen in paediatric bone marrow transplant patients. *The Lancet. Infectious Diseases*, *3*(2), 79–86.

Walsh, M. P., Chintakuntlawar, A., Robinson, C. M., Madisch, I., Harrach, B., Hudson, N. R., Schnurr, D., Heim, A., Chodosh, J., Seto, D., & Jones, M. S. (2009). Evidence of Molecular Evolution Driven by Recombination Events Influencing Tropism in a Novel Human Adenovirus that Causes Epidemic Keratoconjunctivitis. *PLOS ONE*, *4*(6), e5635. https://doi.org/10.1371/journal.pone.0005635

Walsh, M. P., Seto, J., Jones, M. S., Chodosh, J., Xu, W., & Seto, D. (2010a). Computational analysis identifies human adenovirus type 55 as a re-emergent acute respiratory disease pathogen. *Journal of Clinical Microbiology*, *48*(3), 991–993. https://doi.org/10.1128/JCM.01694-09

Walsh, M. P., Seto, J., Jones, M. S., Chodosh, J., Xu, W., & Seto, D. (2010b). Computational Analysis Identifies Human Adenovirus Type 55 as a Re-Emergent Acute Respiratory Disease Pathogen. *Journal of Clinical Microbiology*, *48*(3), 991–993. https://doi.org/10.1128/JCM.01694-09

Wang, S.-L., Chi, C.-Y., Kuo, P.-H., Tsai, H.-P., Wang, S.-M., Liu, C.-C., Su, I.-J., & Wang, J.-R. (2013). High-Incidence of Human Adenoviral Co-Infections in Taiwan. *PLOS ONE*, *8*(9), e75208. https://doi.org/10.1371/journal.pone.0075208

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org

Wickham, H. (2017). *tidyverse: Easily Install and Load the "Tidyverse"* [R package version 1.2.1]. https://CRAN.R-project.org/package=tidyverse

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *dplyr: A Grammar of Data Manipulation* [R package version 0.7.8]. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2019). *tidyr: Tidy Messy Data* [R package version 1.0.0]. https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & François, R. (2018). *readr: Read Rectangular Text Data* [R package version 1.3.1]. https://CRAN.R-project.org/package=readr

Williams, J., Grodzicker, T., Sharp, P., & Sambrook, J. (1975). Adenovirus recombination: Physical mapping of crossover events. *Cell*, *4*(2), 113–119.

Wood, D. J. (1988). Adenovirus gastroenteritis. *British Medical Journal (Clinical Research Ed.)*, *296*(6617), 229–230. PubMed. https://doi.org/10.1136/bmj.296.6617.229-a

Yamada, T., Hara, K., & Kadowaki, T. (2012). Association of adenovirus 36 infection with obesity and metabolic markers in humans: A meta-analysis of observational studies. *PloS One*, *7*(7), e42031. https://doi.org/10.1371/journal.pone.0042031

Yang, Z., Zhu, Z., Tang, L., Wang, L., Tan, X., Yu, P., Zhang, Y., Tian, X., Wang, J., Zhang, Y., Li, D., & Xu, W. (2009a). Genomic Analyses of Recombinant Adenovirus Type 11a in China. *Journal of Clinical Microbiology*, *47*(10), 3082. https://doi.org/10.1128/JCM.00282-09

Yang, Z., Zhu, Z., Tang, L., Wang, L., Tan, X., Yu, P., Zhang, Y., Tian, X., Wang, J., Zhang, Y., Li, D., & Xu, W. (2009b). Genomic Analyses of Recombinant Adenovirus Type 11a in China. *Journal of Clinical Microbiology*, *47*(10), 3082–3090. https://doi.org/10.1128/JCM.00282-09

Yao, L., Wang, C., Wei, T., Wang, H., Ma, F., & Zheng, L. (2019). Human adenovirus among hospitalized children with respiratory tract infections in Beijing, China, 2017–2018. *Virology Journal*, *16*(1), 78. https://doi.org/10.1186/s12985-019-1185-x

Zhang, J., Kang, J., Dehghan, S., Sridhar, S., Lau, S. K. P., Ou, J., Woo, P. C. Y., Zhang, Q., & Seto, D. (2019). A Survey of Recent Adenoviral Respiratory Pathogens in Hong Kong Reveals Emergent and Recombinant Human Adenovirus Type 4 (HAdV-E4) Circulating in Civilian Populations. *Viruses*, *11*(2). https://doi.org/10.3390/v11020129

Zhao, H., Chen, M., & Pettersson, U. (2014). A new look at adenovirus splicing. *Virology*, *456–457*, 329–341. https://doi.org/10.1016/j.virol.2014.04.006

Zhou, X., Robinson, C. M., Rajaiya, J., Dehghan, S., Seto, D., Jones, M. S., Dyer, D. W., & Chodosh, J. (2012). Analysis of human adenovirus type 19 associated with epidemic keratoconjunctivitis and its reclassification as adenovirus type 64.

Investigative Ophthalmology & Visual Science, *53*(6), 2804–2811. https://doi.org/10.1167/iovs.12-9656

Zhu, Z., Zhang, Y., Xu, S., Yu, P., Tian, X., Wang, L., Liu, Z., Tang, L., Mao, N., Ji, Y., Li, C., Yang, Z., Wang, S., Wang, J., Li, D., & Xu, W. (2009). Outbreak of Acute Respiratory Disease in China Caused by B2 Species of Adenovirus Type 11. *Journal of Clinical Microbiology*, *47*(3), 697. https://doi.org/10.1128/JCM.01769-08

BIOGRAPHY

June Kang graduated from Osbourn Park High School, Manassas, Virginia, in 2012. She received her Bachelor of Science from James Madison University in 2016.