Tracing adaptive pathways in a proofreading-deficient coronavirus

By

Kevin Whittle Graepel

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Microbe-Host Interactions
May 31, 2019
Nashville, Tennessee

Approved:

James E. Cassat, M.D., Ph.D.

Manuel J. Ascano, Jr., Ph.D.

Seth R. Bordenstein, Ph.D.

Suman R. Das, Ph.D.

Mark R. Denison, M.D.

To my family, past, present, and future.

ACKNOWLEDGEMENTS

To get it out of the way, the work described in this dissertation would not be possible without the following funding sources: United States Public Health Service awards R01-AI108197 (M.R.D), T32-GM007347 (Vanderbilt MSTP), and F30-AI129229 (K.W.G), all from the National Institutes of Health.

Sincere thanks to all my mentors past, without whom my life's trajectory ends somewhere else entirely. To Nicole Snyder, my first research mentor, for your example of tireless dedication and advocacy, for pushing me to think critically and solve problems independently, and for showing me that productive scientists can still have fun. To Leslie North, for convincing me to expand my life experience before starting medical school and for writing a knock-out recommendation letter. To Bob Collins, for donating time, advice, and a phone call to a complete stranger. To Barney Graham, for taking a chance on an inexperienced student, for encouraging him to seize responsibility and independence, for treating that him like a colleague, and for showing him that humble people can achieve massive success. To members of the Vaccine Research Center—Kaitlyn, Azad, Syed, Tracy, Allison, Patrick, Sung-han, Erez, Masaru, Jason, Gordon, Hadi, Ivelin, Annie, Faust, Anu, Doug, Jie, and especially Man—for mentorship, friendship, and rides to the hospital. To Kayvon Modjarrad, my reciprocal mentor, for taking me (not too) seriously, for solid music recommendations, and for the worst round of bar trivia I've ever played.
Graduate school is canonically compared with, well, hell. Special thanks due to the lab team that makes every day, even the bad days, a whole lot better. In no particular order: Clint, for surprise carbon dioxide sublimations and for stepping up when I needed you

most. Nicole, for letting me know what you think without saying a word. Brett, for advice on fantasy football and home ownership. Maria, for hipster gags and shared suffering. Jennifer, for taking on the mantle of "the weird one" and serving as the consultant geek. Litton, for social planning and a well-earned pizza party. Erica, for forgiving me after the meatball incident. Jim, for an open door and deep insight. Xiaotao, for steadfast support, institutional memory, and a great pair of lab hands. Laura, for gluten-free pastries and a sense of porpoise. Tia, for technical expertise and for tolerating incomplete instructions. Thayer, for lessons about the challenges of mentorship. Andrea, for irreverence and a no-nonsense approach to organization. Further thanks to the broader communities of VI4, PMI, MHI, MPI, and other alphabet soup acronyms I've forgotten, for keeping me humble and broadening my horizons.

Distinctive thanks are due to Xiaotao Lu and Brett Case, for the interminable viral passage experiment that is the basis for this dissertation research.

If graduate school is hell, then graduate school plus medical school must be double-hell. I have, fortunately, been surrounded by a remarkable group of stellar scientists, clinicians, and people. To Shawn, John, Josh, Katherine, Eileen, Steph, Sumeeth, Alex, Joey, David, and Gabby, for taking this crazy journey with me, for serving as collective guinea pigs, and for modeling lives worth living. To all students in the Vanderbilt MSTP, for pushing me beyond my perceived limits. To Terry Dermody, Jim Bills, Larry Swift, and Melissa Krasnove for building a program worth joining, and to Chris Williams, Megan Williams, Lourdes Estrada, Sally York, Ambra Pozzi, Bryn Sierra, and Danny Winder for making it one worth sticking with. To members of Gabbe College, class of 2017, for open-

mindedness, raw honesty, and great snacks, and to Ban Allos, Amy Fleming, and Matt Miller for guiding us through two challenging years. To all members of the VUSM class of 2017, for showing me that residency is going to be ok.

It's easy to convince yourself that your work is worth doing, but it's harder to convince everyone else. Many thanks to my thesis committee—Manny Ascano, Seth Bordenstein, Jim Cassat, Suman Das, Earl Ruley—for challenging me on my ideas, for suggesting critical experiments, and for treating me, always, like someone worth listening to.

For graduate research, they told me to pick the mentor, not the project. Endless thanks due to Mark Denison, for four years of enthusiastic support, for independence, for life advice, for always finding meaning in negative results, for your example of the physician-scientist career, for stealing puns out from under my nose, and for caring so much about my success.

Nothing is possible without the support of your friends. Vital thanks to Kevin, Steve, Dan, Greg, Danny, Jamie, Jacob, Corey, Bill, Zach, Otey, Sam, Hansel, Pete, Jack, Jenna, Jessi, Katie, Dana, Courtney, Stef, Jamie, Mary, Issy, Billy, Keith, Andrew, and more—you all own a small piece of this achievement. Similarly, to family. Mabel, George, Maxine, Allen, David, Jayne, Dave, Jenny, Ric, Randy, Kim, Jerry, Christine, Molly, Andrew, Aimee, Brandon, Kerrin, Brian, Danielle, Christopher, Suzanne, Nora Frances, Vaughan, Douglas, Sarah, Alexander, Christina, Cameron, Allison.

This journey never belonged to me, alone. Boundless thanks to my wife, Selby, for giving me the freedom to choose the right program, for waiting as long as you did, for being the

best mom, for your constant faith in me, for croquet parties, and for delicious, arbitrarily spiced meals. Thanks also to Flora and Sadie, for making it hard to leave in the morning but easy to come home.

Above all, deepest thanks to my parents, for everything.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1: BACKGROUND AND LITERATURE REVIEW

Introduction

Valentine's Day celebrations in 2003 did not stop the World Health Organization from releasing its Weekly Epidemiological Record (World Health Organization, 2003). The bulk of that week's issue was unremarkable—an update on the global status of immunization safety—but five sentences were devoted to several hundred cases of an acute respiratory syndrome in China's Guangdong Province. The document is written in dry, scientific language, and it does not set off any alarm bells. In fact, nothing suggests that these were the first rumblings of a worldwide threat. Over the ensuing weeks, case reports continued to climb—1000 by March 27, 3000 by April 14, 7000 by May 8—and the WHO's reporting became steadily grimmer. The disease had earned an ominous moniker, severe acute respiratory syndrome (SARS). Global, national, and local health officials scrambled to contain, identify, and trace the unknown pathogen causing SARS, and the world waited anxiously. Was this the pandemic we had long feared?

Without knowing the cause of SARS, containment and quarantine were the primary weapons for its control. In a cruel twist, these strategies meant that those tasked with caring for the sick, healthcare workers, were one of the most highly infected groups. By the time the outbreak was stopped, in early June 2003, SARS had reached every inhabited continent except South America and caused 8096 cases and 774 deaths. The infectious agent had also been identified—SARS was caused by a coronavirus (Drosten et al., 2003; Ksiazek et al., 2003; Peiris et al., 2003b; Rota, 2003).

In and of itself, isolating a coronavirus that caused human respiratory illness was unremarkable.

As early as the 1960's, coronaviruses were known causes of the "common cold" but were of no unusual threat to the world's population (Mäkelä et al., 1998; McIntosh et al., 1973; 1970). In fact, before SARS, members of this single-stranded RNA virus family were much more important economically than clinically. The first identified coronavirus, infectious bronchitis virus (IBV), causes devastating respiratory disease in commercial chickens (Cavanagh, 2007). Transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV), predictably, cause severe diarrhea in pigs. PEDV alone may cause between $900 million and $1.8 billion in annual economic losses in the United States (Annamalai et al., 2015). Cattle are not exempt from coronaviral illnesses, either—newborn calves are susceptible to both enteric and respiratory disease from bovine coronavirus infections. Adult cattle are susceptible to shipping fever and the colorfully-named "winter dysentery with hemorrhagic diarrhea syndrome" (Saif, 2010). Murine hepatitis virus (MHV), the focus of this dissertation, is a scourge of animal research facilities across the globe (National Research Council (US), 2000). Coronaviruses have also been found in a broad range of other non-commercial mammalian and avian species (Brister et al., 2014; Leibowitz et al., 2011). The emergence of a highly pathogenic human coronavirus changed the paradigm. Rapidly transmitted, terrifyingly lethal, and nearly untreatable, SARS-CoV ignited a new era of research into coronavirus biology, ecology, and countermeasures.

In the 15 years since SARS made its mark on the public memory and the global health network, intensive study of SARS-CoV and its relatives has revealed fascinating aspects of coronaviral biology. Coronaviruses carry some of the largest single-strand RNA genomes in existence and encode an array of enzymes rarely found in RNA viral systems (Gorbalenya et al., 2006; Smith et al., 2014; Snijder et al., 2016; 2003). Worldwide networks of global health officials, epidemiologists, zoologists, and virologists have identified numerous coronaviruses poised for

emergence into humans (Becker et al., 2008; Ge et al., 2013; Lau et al., 2005; Li et al., 2005; Poon et al., 2005), and collaborative efforts between coronavirologists, chemists, structural biologists, and pharmaceutical companies have explored diverse strategies for countermeasures against CoV disease (Modjarrad, 2016; Rabaan et al., 2017). In this time, as well, three new human coronaviruses have been identified (van der Hoek et al., 2004; Woo et al., 2004; Zaki et al., 2012), two of which cause mild respiratory disease. The third, MERS-CoV, emerged from bat and camel populations to cause the Middle East respiratory syndrome and remains a public health threat 6 years after its identification (Arabi et al., 2017; Zaki et al., 2012).

In this dissertation, I describe my contribution to this compendium of work. I focus on an enzyme rarely found in RNA viral biology: a proofreading exoribonuclease. In particular, I examine the importance of this enzyme for coronavirus replication, and I assess the evolutionary consequences of its disruption. Along the way, I have adapted experimental methods from other viral systems for the detailed study of coronaviruses.

Coronavirus genome organization and replication cycle

The phylogeny of coronaviruses was recently updated to accommodate newly discovered viral species (Gorbalenya et al., 2017). The family *Coronaviridae* remains classified within the order *Nidovirales*, suborder *Cornidoviridae*, and is now divided into two subfamilies: *Orthocoronavirinae* and *Alphaletovirinae.* All coronaviruses discussed in this dissertation are *Orthocoronavirinae*, divided into four genera: *Alphacoronavirus*, *Betacoronavirus*, *Deltacoronavirus*, and *Gammacoronavirus.* Of the six coronaviruses infecting humans, two are alphacoronaviruses (HCoV-229E and HCoV-NL63), and four are betacoronaviruses (SARS-

CoV, MERS-CoV, HCoV-OC43, and HCoV-HKU1). Murine hepatitis virus, the focus of this dissertation, is also a betacoronavirus. The recently described microhyla letovirus 1 (MleV) is the sole member of the *Alphaletovirinae* subfamily (Bukhari et al., 2018). From this point forward, "coronavirus" will refer only to the *Orthocoronavirinae*, as little is known about MleV.

Coronavirus genomes are single strands of positive-sense RNA (that is, directly translatable by the ribosome) that range from 26-32 kilobases in length (Brister et al., 2014) (Figure 1). Like cellular mRNA, coronavirus genomes have 5′ caps and 3′ poly(A)-tails, and the termini fold into conserved RNA secondary structures essential for replication (Goebel et al., 2007; 2004a; 2004b; Masters, 2006). Coronaviruses encode 16 nonstructural proteins (nsps), a minimum of four structural proteins, and a species-dependent array of accessory proteins (Perlman and Netland, 2009). In general, nonstructural proteins are involved in modifying cell membranes, cleaving viral proteins, and replicating the genome, while structural and accessory proteins promote viral assembly and dissemination. Nonstructural proteins are translated from two overlapping open-reading frames (ORFs) covering the first two-thirds of the genome. Structural and accessory proteins are encoded in separate ORFs in the final third of the genome and are translated from a nested set of subgenomic mRNA species (sgmRNA). Each sgmRNA contains the same 5′ and 3′ terminal sequences as the full genome, including the cap and poly(A)-tail. Although most sgmRNAs contain multiple protein-coding regions, only the 5′-most ORF is translated (Masters, 2006). Transcriptional regulatory sequences (TRS) precede each ORF.

Coronavirus replication begins when the spike fusion glycoprotein (hereafter, spike protein) interacts with its cellular receptor (Figure 2). Host receptor distribution is a critical determinant

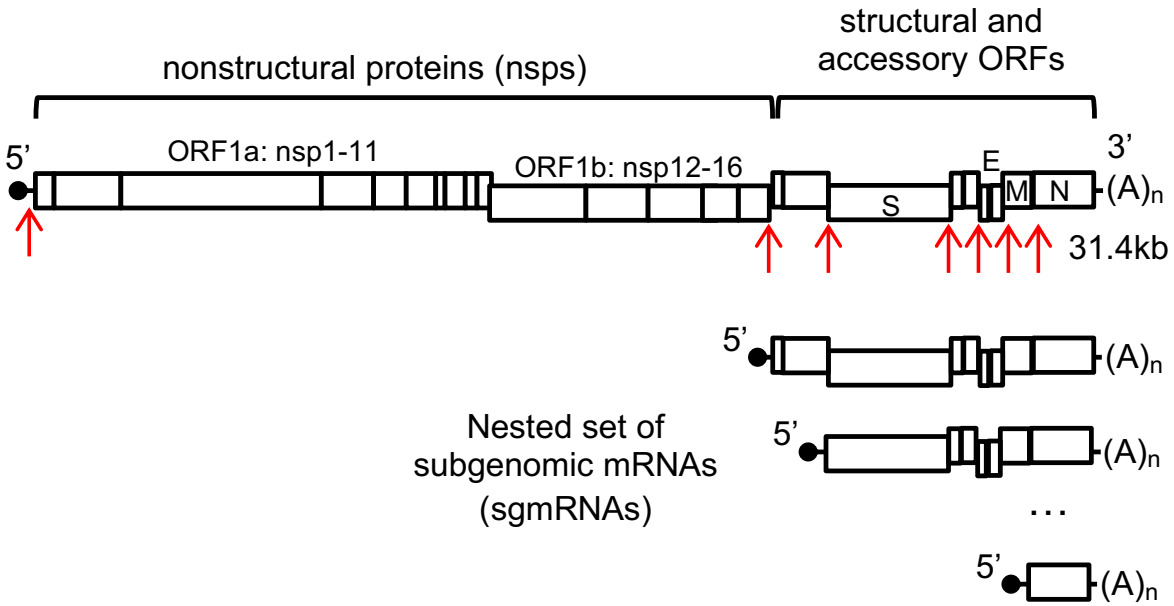


Figure 1. Murine hepatitis virus (MHV) genome organization. The MHV genome is a 31.4 kilobase, capped (dark circle), and polyadenylated positive-sense RNA molecule. The first two thirds of the genome encode 16 nonstructural proteins translated as a single polyprotein with a ribosomal frameshift. The final one-third encodes the structural proteins (spike – S, envelope – E, matrix – M, and nucleocapsid – N) and several accessory proteins. Transcriptional regulatory sequences (red arrows) at the 5′ end and upstream of the structural and accessory ORFs drive transcription of a nested set of subgenomic mRNAs. Other coronavirus genomes differ from MHV primarily in the number and identity of accessory proteins.

of viral tropism, and structural variation in receptors between species is one of the key challenges to developing animal models for human coronavirus infections (Agrawal et al., 2015; van Doremalen et al., 2014; Zhao et al., 2014). The receptors for MHV and SARS-CoV are carcinoembryonic antigen-related cell adhesion molecule-1 (CEACAM-1) and angiotensin-converting enzyme-2 (ACE2), respectively (Dveksler et al., 1993; Li et al., 2003). Dipeptidyl peptidase 4 (DPP4), in concert with CEACAM-5, permits MERS-CoV entry (Chan et al., 2016; Raj et al., 2013). Receptor engagement triggers the spike protein to fuse viral and cellular membranes, releasing nucleoprotein-coated genomes into the cytoplasm. Host ribosomes immediately translate the first open-reading frame, ORF1, which covers the first two-thirds of the genome. Approximately 60% of translation events terminate at around nucleotide 13,000, yielding polyprotein 1a (pp1a). In the remaining 40%, the ribosome slips backwards by one nucleotide and continues to synthesize polyprotein in a new frame (a -1 ribosomal frameshift) to yield pp1ab (Brierley et al., 1987; 1989; Irigoyen et al., 2016). Virus-encoded proteases cleave individual nsps from the polyprotein (Masters, 2006). Once matured, nsps 3, 4, and 6 convert the host endoplasmic reticulum into a network of double-membrane vesicles and convoluted membranes (Knoops et al., 2008; Oudshoorn et al., 2017). The remaining nsps then assemble into the replicase-transcriptase complex (RTC) and initiate genome replication and sgmRNA transcription. The RTC will be discussed in greater detail in subsequent sections.

Schematically, genome replication is straightforward (Figure 3A,B). Full-length negative-sense RNAs are generated and serve as the template for genomic RNA production. The mechanism of sgmRNA transcription remains controversial, but data to date support a model of discontinuous transcription (Figure 3C,D) (Sawicki et al., 2007). During negative-strand synthesis, the RTC

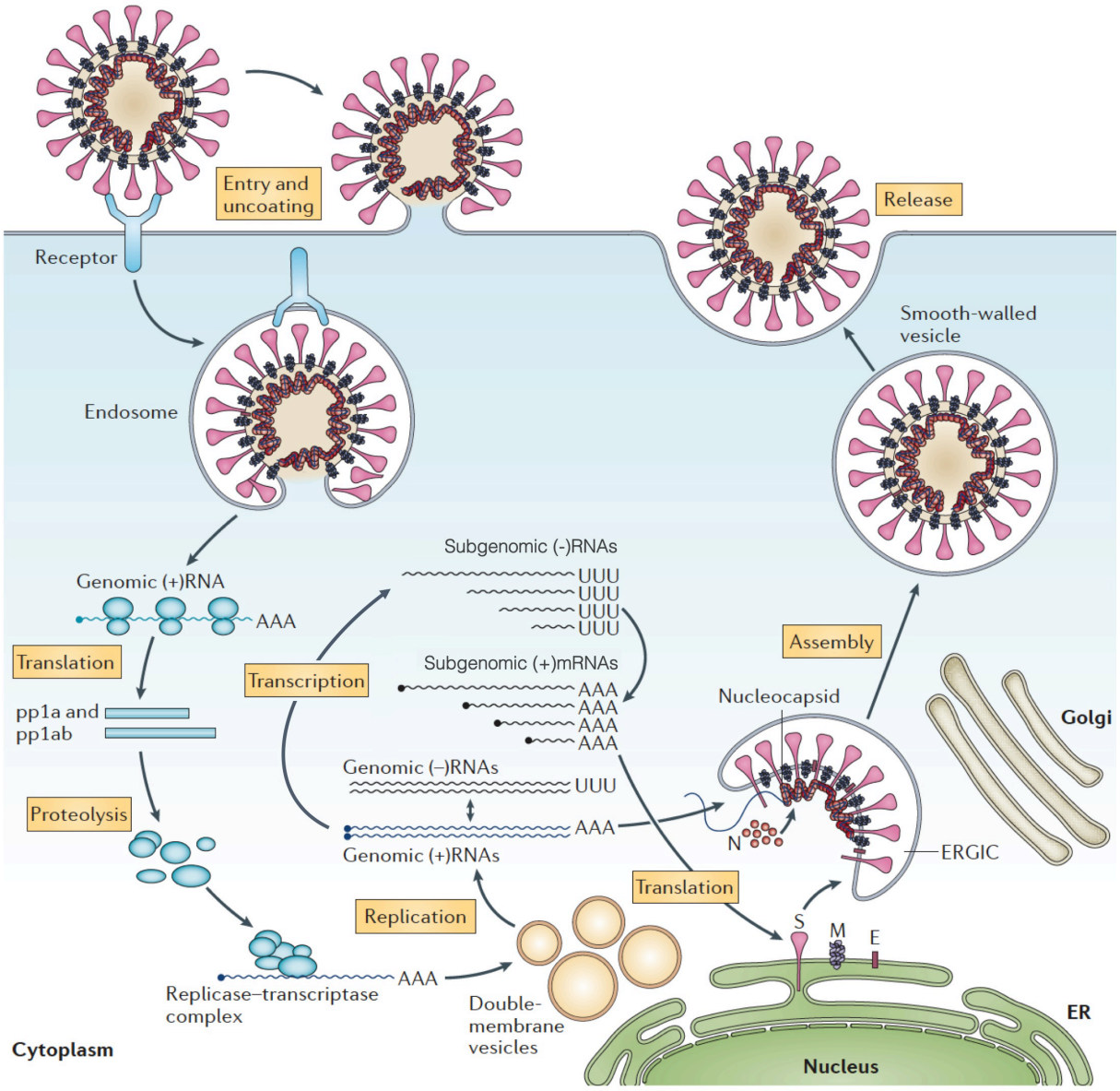


Figure 2. Coronavirus replication cycle. Coronavirus spike proteins facilitate viral entry and virion uncoating. Genomic RNA is immediately recognized and translated by host ribosomes in the cytoplasm, and the resulting polypeptide is cleaved by viral proteases into mature nonstructural proteins (nsps). The replicase-transcriptase complex assembles on virus-induced double-membrane vesicles, where they replicate genomic RNA and transcribe subgenomic mRNAs (sgmRNAs). Coronavirus structural proteins are translated from the sgmRNAs and assemble into full virions in the endoplasmic reticulum golgi-intermediate complex (ERGIC). Progeny virions are trafficked and released by non-lytic exocytic pathways. This figure is reproduced from Case, 2018 and de Wit et al., 2016, with permission.

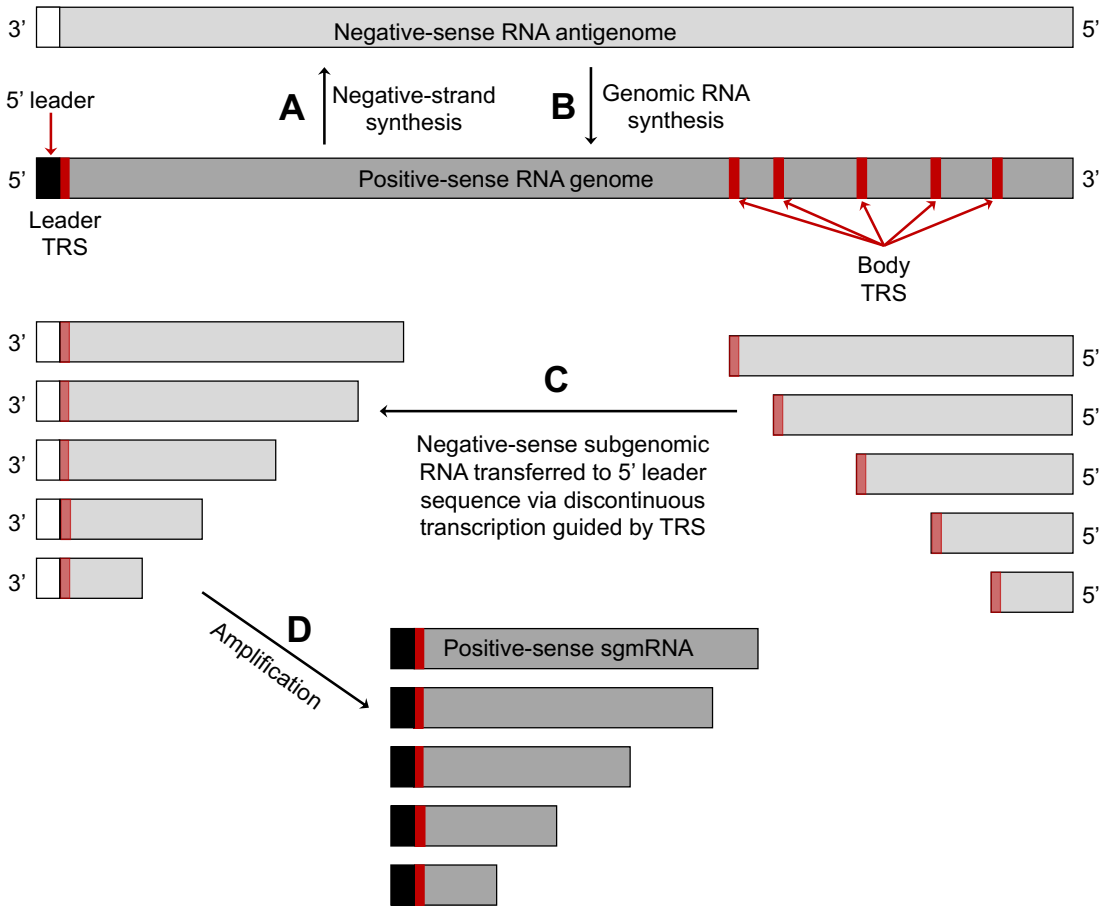


Figure 3. Coronavirus genomic replication and subgenomic mRNA transcription. The coronavirus replicase-transcriptase complex (RTC) generates a full-length antigenome (A) that serves as the template for synthesis of positive-sense RNA genomes (B). During negative-strand synthesis, the RTC jumps between transcriptional regulatory sequences (TRS) to generate subgenomic RNA species (C) that are templates for subgenomic mRNA synthesis (D).

may leap between a TRS and a homologous sequence at the 5′ end of the genome, fusing the ORF with the 5′ untranslated region. The RTC then uses these negative-strand templates to generate positive-sense sgmRNAs that are translated to form the structural and accessory proteins. The frequency with which the individual sgmRNAs are generated determines the yield of individual proteins, which has important implications for genome assembly and packaging. Once structural proteins are generated, virion assembly begins. Within the endoplasmic reticulum golgi-intermediate complex (ERGIC), coronavirus genomes are coated with protective layer of the viral nucleoprotein (N) and encapsulated within a cell-derived lipid envelope decorated with the spike protein (S) (Perlman and Netland, 2009) (Figure 2). In negative-stain microscopy, spike proteins form a characteristic halo of flare-like projections that resembles a solar corona—hence, the name coronavirus (Estola, 1973). The matrix (M) and envelope (E) proteins are also embedded in the viral envelope, and progeny virions are released by cellular secretory pathways.

Human coronaviruses: origins, epidemiology, pathogenesis, clinical manifestations, and countermeasures

Origins. Six coronaviruses are known to infect humans, and all are proposed to be of zoonotic origin. Molecular clock analyses and epidemiological studies suggest that coronavirus emergence into humans occurred recently by evolutionary timescales. HCoV-NL63 emerged the earliest between 1190-1449 CE (Huynh et al., 2012), followed by HCoV-229E between 1501-1883 CE (Pfefferle et al., 2009), and HCoV-OC43 between 1866-1918 CE (Vijgen et al., 2005b). The emergence of HCoV-HKU1 has not been well-studied, but available evidence suggests that it may have entered humans in the 1950's (Al-Khannaq et al., 2016a). SARS-CoV emerged in

late 2002 and was traced to civet cats in animal markets in Guanzhou province, China (de Wit et al., 2016; Guan et al., 2003). MERS-CoV was first identified from a patient in Saudi Arabia in 2012, and later cases were subsequently linked to dromedary camels, where MERS-CoV-like viruses have circulated for at least 30 years (Cui et al., 2018; de Wit et al., 2016; Müller et al., 2014). HCoV-229E, HCoV-NL63, SARS-CoV, and MERS-CoV precursors likely originated in bats, while HCoV-OC43 and HCoV-HKU1 likely arose from rodents (Corman et al., 2018; Cui et al., 2018; Forni et al., 2016a).

Epidemiology. The four endemic human coronaviruses, HCoV-229E, -OC43, -HKU1, and -NL63, are distributed worldwide (Arden et al., 2005; Bastien et al., 2005; Ebihara et al., 2005; Fouchier et al., 2004; Gerna et al., 2006; Kin et al., 2016; Kuypers et al., 2007; Moës et al., 2005; Pyrc et al., 2007; Sloots et al., 2006; Vabret et al., 2006a). CoVs are estimated to cause 10-30% of mild respiratory infections ("common colds"), but these figures may be skewed by frequent coinfection with other respiratory pathogens like respiratory syncytial virus, human metapneumovirus, rhinovirus, and adenovirus (Dominguez et al., 2009; Gaunt et al., 2010; Kuypers et al., 2007; McIntosh et al., 1970). HCoV epidemics display marked seasonality, with peaks of incidence in the winter and early spring and few or no infections detected in the summer (Gaunt et al., 2010; Hodinka, 2016). Individual HCoV species cause epidemics every 2-3 years, and reinfection is common even in the presence of pre-existing immunity (Bradburne and Somerset, 1972; Bradburne et al., 1967; McIntosh et al., 1970; Reed, 1984). All ages are susceptible, but elderly populations may be at greater risk for severe disease requiring hospitalization (Falsey et al., 1997; 2002). Volunteer studies have demonstrated that transmission is predominantly through respiratory droplets (Bradburne and Somerset, 1972; Bradburne et al., 1967; Reed, 1984).

No cases of SARS-CoV have been reported since 2004, but the mechanisms of introduction, spread, and disease remain important for potential animal-to-human transmission and disease. Although the virus appears to have been introduced into humans though animal contact, the clustering of cases is more consistent with human-to-human spread than with thousands of independent animal exposures. Indeed, nosocomial transmission was a common feature of the outbreak, likely related to the fact that peak viral shedding occurs when respiratory symptoms are most severe (Cheng et al., 2004; Peiris et al., 2003a). These observations are consistent with droplet exposure as the primary mode of SARS-CoV transmission. Fomite transmission has also been implicated, as SARS-CoV RNA was found on hospital surfaces (Booth et al., 2005; Xiao et al., 2017a). Subsequent studies have demonstrated that viable SARS-CoV can persist on surfaces for a week at room temperature (Chan et al., 2011; Geller et al., 2012; Rabenau et al., 2004). SARS-CoV RNA can be detected in stool, and fecal-oral transmission may have occurred in the setting of profuse diarrhea (Booth et al., 2005; Xiao et al., 2017a). Aerosol transmission is implicated in several "superspreading" events, in which a small number of infected individuals transmitted infection to a much larger number of persons (Yu et al., 2004).

Since its identification in 2012, MERS-CoV has continued to emerge within and spread outside of the Arabian Peninsula. As of January 27, 2019, the World Health organization has recorded 2,279 laboratory-confirmed cases of MERS-CoV from 27 countries, all of which are linked to exposures in the Arabian Peninsula (approximately 80% in Saudi Arabia). As with SARS-CoV, the majority of reported MERS-CoV cases are linked to nosocomial transmission of respiratory droplets (Arabi et al., 2017). However, in contrast to the SARS-CoV epidemic, no sustained human-to-human transmission has been observed. Instead, as many as 40% of primary cases are linked to exposure to dromedary camels (Alraddadi et al., 2016). Camels carry great cultural

significance in the Middle East, hindering efforts to contain MERS-CoV. Higher levels of MERS-CoV RNA are detected in lower respiratory tract samples and can persist in some patients for more than a month (Memish et al., 2014; Song et al., 2015). MERS-CoV is also frequently detected in blood and stool and can survive at least 48 hours on environmental surfaces (Kim et al., 2016; van Doremalen et al., 2013). Most outbreaks of MERS-CoV are relatively small, but at least one superspreading event resulted in 186 cases and 36 deaths in Korea (Ki, 2015).

Pathogenesis of SARS and MERS. Severe disease in SARS and MERS likely results from direct virologic damage and subsequent immunopathology. SARS-CoV and MERS-CoV replicate to high titers in respiratory epithelial cells early during infection, ravaging the pulmonary lining (Channappanavar et al., 2016; Chu et al., 2004; Oh et al., 2016; Peiris et al., 2003a). Both viruses encode multiple innate immune antagonists that delay and dysregulate host responses to infection (reviewed in (Channappanavar and Perlman, 2017)). High virus titers eventually initiate an inflammatory cascade that leads to a vicious cycle of cytokine elaboration, vascular leakage, and inflammatory cell infiltration (Channappanavar et al., 2016; Gralinski and Baric, 2015; Totura and Baric, 2012). Overexuberant immune responses impair viral clearance and lead to acute respiratory distress syndrome, the primary cause of death from severe CoV disease (Arabi et al., 2017). Infection of distant organs has been observed for SARS-CoV (Gu et al., 2005) but not for MERS-CoV. Recapitulation of human clinical features in animal models of MERS-CoV infection remains challenging, but promising new models are in development.

Clinical manifestations. *Endemic human coronaviruses*. The clinical presentations of endemic HCoV-OC43 and HCoV-229E are indistinguishable from the "common cold" symptoms of other respiratory viruses. The average incubation period is 2-4 days, with symptoms typically lasting

4-7 days. Rhinorrhea, cough, sore throat, malaise, and headache are the most common symptoms. Fever occurs in up to 60% of cases (Bradburne and Somerset, 1972; Bradburne et al., 1967; McIntosh et al., 1970; Reed, 1984). Clinical manifestations of HCoV-HKU1 and HCoV-NL63 are presumed to be similar but have not been well-studied, though HCoV-NL63 has been associated with croup in children (van der Hoek et al., 2005). HCoV infections are linked to episodes of wheezing and asthma attacks in both adults and children (McIntosh et al., 1973; Nicholson et al., 1993). Severe lower respiratory tract infections have been reported across various demographics but may be more likely in older adults and in immunocompromised patients. Whereas coronaviruses are known causes of severe diarrhea in animals (Annamalai et al., 2015; Saif, 2010), evidence is mixed on whether HCoVs contribute to enteric disease (Esper et al., 2010; Jevšnik et al., 2016; 2013). Similarly, although coronaviruses are important models for neurological diseases like multiple sclerosis (Bailey et al., 1949; Perlman et al., 1987), their role in similar disease processes in humans remains unproven. Diagnostic testing for endemic HCoVs is available (e.g. BioFire FilmArray Respiratory Panel; www.biofiredx.com/filmarrayrp), but its clinical value remains to be determined.

Severe acute respiratory syndrome. The prodrome of SARS-CoV infection lasts three to seven days and is non-specific, consisting of fever, malaise, headache, and myalgias with a notable lack of upper respiratory symptoms. Respiratory symptoms develop during the second week of infection, including nonproductive cough progressing to dyspnea, chest pain, and, in approximately 20% of cases, acute respiratory distress syndrome. As many as 25% of SARS patients required mechanical ventilation in the intensive care unit. Risk factors for severe disease included diabetes, cancer, cardiovascular disease, and chronic obstructive pulmonary disease. Extrapulmonary manifestations including gastrointestinal symptoms (nausea, vomiting, diarrhea)

were common, and approximately 7% of patients developed acute kidney injury (Booth et al., 2003; Leung et al., 2004; Peiris et al., 2003a). Several patients developed neuromuscular symptoms, though it is not clear whether these derived specifically from SARS-CoV infection or were sequelae of chronic illness (Tsai et al., 2004). The case fatality rate from SARS-CoV infection during the 2003 outbreak was 9.6%. The case definition for SARS is provided in Table 1 (World Health Organization, 2009).

Middle East respiratory syndrome. The clinical presentations of reported MERS cases were thoroughly collated and presented in Arabi et al., 2017. Briefly, the incubation period of MERS-CoV is between 2-14 days and typically presents with nonspecific clinical features typical of acute respiratory illnesses, including low-grade fever, rhinorrhea, sore throat, and myalgia. Severe disease is characterized by the acute respiratory distress syndrome with multilobe involvement visible by radiography. Risk factors for severe disease include age >50 years and comorbidities like obesity, diabetes, COPD, end-stage renal disease, cancer, and immunosuppression. Extrapulmonary manifestations are common in severe MERS disease and include gastrointestinal symptoms (nausea, vomiting, diarrhea), acute kidney injury, and encephalitis-like neurological findings. Laboratory analyses typically detect leukopenia and lymphopenia with occasional thrombocytopenia, anemia, and aminotransferase elevations. The case fatality rate remains at 35%, though the true incidence of MERS-CoV infection is likely underestimated by existing data due to a potentially high incidence of asymptomatic infection. The case definition for MERS is provided in Table 2 (World Health Organization, 2017; 2018).

Countermeasures. Coronavirus infections of humans are acute and self-limited. No effective antivirals are approved to treat human coronaviruses. The nucleoside analogue ribavirin was

Table 1. Clinical case definition of severe acute respiratory syndrome. World Health Organization, 2009.

	Definition
Clinical data	• Documented fever **AND** • Symptoms of lower respiratory tract illness (cough, shortness of breath, difficult breathing) **AND** • Radiographic or autopsy evidence consistent with pneumonia or acute respiratory distress syndrome • No alternative diagnosis fully explaining the illness
Laboratory data	Diagnosis must be confirmed with one or both of: • Detection of viral RNA by RT-PCR from two separate samples **OR** virus culture from any specimen • Detection of rise in antibody titer

Table 2. Clinical case definition of Middle East respiratory syndrome. World Health Organization, 2017 and 2018.

	Definition
Confirmed Case	Laboratory confirmation of infection irrespective of clinical signs and symptoms
Probable case	Three definitions: • Acute febrile respiratory illness with clinical, radiographic, or histopathologic evidence of pneumonia or acute respiratory distress syndrome **AND** • Direct epidemiologic link with confirmed MERS case **AND** • Testing for MERS-CoV is unavailable, negative on inadequate specimen, or inconclusive **OR** • Acute febrile respiratory illness with clinical, radiographic, or histopathologic evidence of pneumonia or acute respiratory distress syndrome **AND** • Person resides in or traveled to Middle East or country with known recent human/dromedary infections **AND** • Testing for MERS-CoV is inconclusive **OR** • Acute febrile respiratory illness of any severity **AND** • Direct epidemiologic link with confirmed MERS case **AND** • Testing for MERS-CoV is inconclusive
Laboratory data	• Detection of viral nucleic acid by RT-PCR for at least two specific genomic targets **OR** one genomic target with sequencing of a second target • Evidence of seroconversion by acute and convalescent antibody titers taken at least 14 days apart

used extensively and empirically during the 2003 SARS-CoV outbreak, but ribavirin-associated adverse events outweighed any measurable clinical benefits (Booth et al., 2003; Chiou et al., 2005; Muller et al., 2007; Stockman et al., 2006). The ineffectiveness of ribavirin has been attributed both to its effects on host immune responses and to the SARS-CoV proofreading exoribonuclease (Barnard et al., 2006; Ferron et al., 2018; Smith et al., 2013). However, recent studies in our lab have identified a nucleoside analogue, remdesivir, with potent activity against epidemic and zoonotic coronaviruses that is proceeding towards clinical trials (Agostini et al., 2018; Sheahan et al., 2017). Non-specific immune modulators including systemic corticosteroids may be associated with increased mortality in SARS-CoV and MERS-CoV and are thus not recommended unless indicated for another clinical condition (Arabi et al., 2017; Auyeung et al., 2005; Stockman et al., 2006). However, host-directed therapies intended to promote antiviral immunity or prevent inflammatory damage are under consideration (Zumla et al., 2015). Meta-analysis of observational studies suggests that human convalescent plasma may reduce SARS mortality (Mair-Jenkins et al., 2014), but the use of blood products has not been well-studied in MERS due, in part, to the challenge of harvesting large quantities of human hyperimmune serum (Arabi et al., 2016). Several polyclonal and monoclonal antibody preparations have shown positive results against SARS- and MERS-CoV in animal studies (Beigel et al., 2018; Luke et al., 2016; Pascal et al., 2015; Traggiai et al., 2004; Wang et al., 2018; Zhu et al., 2007), but high cost and the potential for antibody-enhanced disease may limit their clinical utility (Wang et al., 2014).

Challenges for development of effective vaccines targeted against HCoV-OC43, -229E, -HKU1, and -NL63 include the fact that infections are rarely life-threatening and reinfection is the rule, even in the presence of natural immunity from previous infections (Bradburne and Somerset,

1972; Bradburne et al., 1967; McIntosh et al., 1970; Reed, 1984). The durability of immunity to SARS-CoV and MERS-CoV is poorly understood. Nevertheless, effective vaccines for SARS-CoV and MERS-CoV are highly desirable but not yet available. A potential vaccine target is the viral spike protein, which recognizes and penetrates host cells. In MERS-CoV-infected patients, the production of anti-spike IgG is correlated with survival, and anti-spike antibody levels inversely correlate with viral load in the respiratory tract (Corman et al., 2016a). Comprehensive discussion of existing vaccine development efforts for CoVs is beyond the scope of this dissertation and has been reviewed elsewhere (Choi et al., 2017; Rabaan et al., 2017; Schindewolf and Menachery, 2019). Briefly, candidate vaccine platforms include delivery of recombinant spike protein immunogens (Li et al., 2013; Pallesen et al., 2017); delivery of spike-encoding nucleic acid (Kim et al., 2014; Muthumani et al., 2015; Song et al., 2013; Wang et al., 2015a); and attenuation of live virus by disrupting critical protein functions or genetic architectures (Almazan et al., 2013; Fett et al., 2013; Graham et al., 2012; 2018; Lamirande et al., 2008; Menachery et al., 2018; 2014). Of particular relevance to this dissertation is the observation that a live, recombinant SARS-CoV mutant lacking exoribonuclease activity is attenuated and protective (Graham et al., 2012; Menachery et al., 2018). Nevertheless, despite promising results in preclinical studies, no vaccines have yet been approved for human coronaviruses. Thus, approaches for rapid development of stably attenuated live viruses or broadly immunogenic and cross-protective protein immunogens continues to be a key area for future research.

Reservoirs for emerging coronaviruses

Within a few months after SARS-CoV was identified, the epidemic had been traced to wet

markets in Guangzhou province, China. Virus was detected in captive civet cats but not in wild populations, indicating that these animals were incidental hosts rather than true reservoirs for SARS-CoV (Guan et al., 2003; Kan et al., 2005; Wang et al., 2006). Subsequent field studies implicated Chinese horseshoe bats as the likely origin of SARS-CoV (Lau et al., 2005; Li et al., 2005). Since then, the search for prepandemic coronaviruses has focused on bats.

Bats are a common starting point in any search for potentially zoonotic viruses, as bats are excellent hosts for evolving viruses. In part, this is a numbers game – with bats constituting more than 20% of all mammalian species, we would expect to find many different viruses among them. However, several other properties contribute to the viral diversity maintained within bat populations (reviewed in (Calisher et al., 2006)). Bats are exposed to a wide variety of viruses as they traverse wide geographic regions, with some individual animals migrating more than 800 miles annually. Viral infections in bats are commonly asymptomatic and persistent, with virus shedding sometimes continuing for several months (Baker et al., 2013). Within massive and dense communities of bats (e.g. several million bats roosting at ~300 animals per square foot), viruses likely transmit rapidly and frequently. Together, these factors provide extensive evolutionary time to viruses and allow frequent shuffling of genetic elements that may drastically alter tropism and pathogenicity.

Studies of bat populations since 2003 have determined that viruses phylogenetically related to SARS-CoV (SARSr-CoV) are distributed across Europe, Africa, and Southeast Asia (Cui et al., 2018; de Wit et al., 2016). In one cave in Yunnan province, China, researchers found all the genetic elements needed to form SARS-CoV (Hu et al., 2017; Wang et al., 2016). Two SARSr-CoV are capable of infecting human cells without adaptation and are immediately poised for

emergence (Ge et al., 2013; Menachery et al., 2016), while several more are on the precipice (Cui et al., 2018; Hu et al., 2017; Menachery et al., 2015). Numerous MERS-CoV-related viruses (MERSr-CoV) have been identified in bats (Annan et al., 2013; Corman et al., 2014a; 2014b; Lau et al., 2013; Yang et al., 2014), though none are an immediate progenitor of circulating dromedary and human viruses (Cui et al., 2018). Viruses related to endemic HCoV-229E and HCoV-NL63 have also been identified in bats (Corman et al., 2015; Drexler et al., 2010). The preponderance of CoV sequences in bat populations suggests that bats are likely the major natural reservoirs for coronaviruses.

The focus on bats as coronaviral reservoirs is valuable but neglects the larger trends in coronavirus emergence. Contact between humans and bats is relatively rare compared with our exposure to other coronavirus-susceptible species such as dogs, cats, pigs, chickens, turkeys, rodents, and camels (Brister et al., 2014; Leibowitz et al., 2011). Accordingly, bat exposures have not been implicated directly in transmission of any endemic or zoonotic coronaviruses. Instead, the evidence indicates that coronaviruses enter humans after amplification (and perhaps evolution) within domesticated, intermediate hosts: civet cats for SARS-CoV, dromedary camels for MERS-CoV, and cattle for HCoV-OC43. In a mechanism paralleling that of MERS-CoV, precursors to HCoV-229E evolved towards the human genotype within camels before entering human populations (Corman et al., 2016b; 2015). Curiously, no host-specific coronaviruses have been found in any primate species, further suggesting that exposure to domesticated animals may be essential for human acquisition of coronaviruses (Corman et al., 2018). In addition, the financial costs of field studies of bats has stifled examinations of other potential reservoir hosts, such as rodents. Recent studies have identified rodent coronaviruses closely related to HCoV-OC43 and HCoV-HKU1 (Lau et al., 2015; Wang et al., 2015b), and a relative of MERS-CoV

was identified in European hedgehogs (Corman et al., 2014b). Thus, while bats may be a genuine incubator for potentially pandemic coronaviruses, predictive models for coronavirus emergence should be bolstered by data regarding other potential reservoirs and zoonotic sources.

SARS-CoV and MERS-CoV have clearly established the potential for zoonotic coronaviruses to cause outbreaks, but coronaviruses already circulating in humans could be the source of the next pandemic. Coronaviruses continue to evolve within their human hosts. SARS-CoV isolates from early in the outbreak have complete ORF8 coding sequences, but most viral genomes from middle- and late-epidemic patients have a 29-nucleotide deletion in ORF8 (Chinese SARS Molecular Epidemiology Consortium, 2004). Some genomes isolated in the very late stages of the epidemic lack ORF8 entirely, indicating that SARS-CoV was evolving even during the few months of the outbreak. Deletions potentially reflecting adaptation to the human host (or the release of selective pressure in a reservoir) have also been identified in MERS-CoV isolated in Jordan and Korea (Lamers et al., 2016; Payne et al., 2018; Xie et al., 2017). The impact of MERS-CoV deletions on viral fitness is unclear, but the SARS-CoV ORF8 deletions are attenuating in multiple model systems (Muth et al., 2018). Extensive positive selection has been detected in the nonstructural proteins of MERS-CoV during four years of circulation (Forni et al., 2016b), strongly suggesting adaptation to the human host. Endemic human coronaviruses also evolve over time, demonstrating that coronaviruses are not static in human populations (Al-Khannaq et al., 2016b; Gerna et al., 2006; St-Jean et al., 2004; Vabret et al., 2006b; Vijgen et al., 2005a). Thus, the next pandemic human coronavirus could emerge not from an animal, but from our own lungs.

Evolution of RNA viruses

The remarkable diversity of coronaviruses worldwide is a small example of a much larger truth: RNA viruses are the most diverse group of biological entities on the planet (Holmes, 2016). RNA viruses replicate with some of the highest mutation rates known, contributing to their enormous and notorious adaptability (Duffy et al., 2008; Sanjuán et al., 2010). High mutation rates are both a blessing and a curse for RNA viruses, promoting rapid replication and adaptation but also constraining genome size and complexity. In this section, I will discuss the sources and importance of mutation rates in RNA viruses and establish general principles that determine the outcomes of viral evolution.

How do viruses generate diversity? Genomic mutations are the ultimate source of all viral diversity. A key driver of mutation rates in RNA viruses is the intrinsic fidelity of their replication. Replication fidelity describes the accuracy with which nucleic acids are copied relative to the template strand. In cellular DNA replication, fidelity is regulated at multiple levels: selection and incorporation of correct bases by the polymerase, detection and editing of mismatches during replication by proofreading exonucleases, and repair of errors after replication. The confluence of these mechanisms means that, on average, DNA replication errors only occur at a rate of 10^{-11} to 10^{-9} mutations per nucleotide per replication cycle (approximately one error per billion bases copied) (Smith et al., 2014). In contrast, RNA virus fidelity is regulated almost exclusively by RNA-dependent RNA polymerases (RdRp) or retroviral reverse transcriptases (RT). Although the intrinsic fidelity of RdRps does not differ significantly from DNA polymerases (Arnold and Cameron, 2004), RNA viruses typically lack proofreading or post-replicative repair mechanisms and thus have mutation rates of approximately 10^{-6} to 10^{-4}

mutations per nucleotide per cell infection (Arnold and Cameron, 2004; Sanjuán et al., 2010; Steinhauer et al., 1992). Other viral proteins, such as the chikungunya helicase, West Nile virus N7-methyltransferase, and yellow fever envelope and pre-membrane proteins, can modulate viral fidelity but do not drastically change mutation rates (Collins et al., 2018; Stapleford et al., 2015; Van Slyke et al., 2015). A key exception is the proofreading exoribonuclease encoded by coronaviruses, which is discussed in detail in later sections.

Viral replication enzymes do not operate in isolated biochemical systems, however, and thus their error rates also depend upon the cellular microenvironment. Reactive oxygen and nitrogen species produced in response to infection can damage viral nucleic acids, disrupting the interactions between Watson-Crick base pairs that are critical for accurate nucleotide selection (Seronello et al., 2011; Yoshitake et al., 2004). In addition, misincorporation frequencies can be altered by disrupting cellular nucleotide homeostasis (Diamond et al., 2002; Smith et al., 2013; Stapleford et al., 2015). Host-encoded protein families can directly mutate RNA viruses. Apolipoprotein B mRNA-editing catalytic peptide-like enzymes (APOBECs) convert cytosines to uracils in viral DNA from hepatitis B virus, papillomaviruses, and herpesviruses (Sanjuán et al., 2010). APOBECs are also packaged within HIV-1 virions and cause 98% of mutations in HIV-1 *in vivo*, a >40-fold increase over the virus-encoded RT (Cuevas et al., 2015). Adenosine deaminases (ADARs) convert adenosines to inosines in double-stranded RNA. In subsequent replication events, inosines pair with guanosines rather than thymines, causing A-to-G transitions. ADAR editing has been demonstrated in a variety of RNA and DNA viruses (Tomaselli et al., 2015). Additional undetermined host factors may affect mutation rates across cell types. For example, vesicular stomatitis virus (VSV) has higher mutation rates in mammalian cells compared to insect cells (Combe and Sanjuán, 2014), and the distribution of

mutation types in cucumber mosaic virus varies depending on the plant host (Pita et al., 2007). The precise mechanisms for these results have not yet been determined. Whether mutagenesis driven by the cellular microenvironment is pro- or antiviral remains controversial and may depend upon the particular model under study (Sanjuán and Domingo-Calap, 2016; Tomaselli et al., 2015).

Once mutations generate initial variability in viral populations, recombination amplifies genetic diversity by shuffling genetic material. In RNA viruses, recombination occurs when the viral polymerase switches between different templates present within the same cell. By transferring large regions of viral sequence between templates, recombination can readily access evolutionary innovations that would be unattainable by mutation alone. Recombination is important within individual populations, where it can purge deleterious mutations or combine beneficial ones (Bentley and Evans, 2018; Sanjuán et al., 2010; Xiao et al., 2016). It is also important for global viral ecology, where large-scale transfer of whole genes (e.g. accessory proteins) or protein functional domains (e.g. spike protein receptor-binding motifs) facilitates new strain and species emergence (Arenas et al., 2018; Bentley and Evans, 2018; Mandary and Poh, 2018). Recombination occurs frequently in coronaviruses in tissue culture (Lai et al., 1985; Masters, 2006), and hallmarks of recombination are frequently found in coronavirus ecological surveillance (Su et al., 2016). Recombination likely facilitated the emergence of both SARS-CoV and MERS-CoV and has promoted vaccine failure for infectious bronchitis virus and porcine epidemic diarrhea virus (Chen et al., 2017; Feng et al., 2018; Forni et al., 2016a; Hon et al., 2008; Sabir et al., 2016). RNA-dependent RNA polymerases drive recombination in enteroviruses (Bentley and Evans, 2018), but the mechanism of coronavirus recombination is not fully understood.

Why is viral diversity important? High mutation rates in RNA viruses have profound biological consequences. Under logarithmic growth, RNA viruses can theoretically make one mutation at every site in the genome during every replication cycle (Lauring and Andino, 2010). Consequently, RNA viruses do not exist as collections of identical clones but as populations of closely-related but genetically distinct variants (mutant swarm) that interact to determine viral phenotypes. The importance of genetic diversity for virulence was elegantly demonstrated in poliovirus (Vignuzzi et al., 2006) (Figure 4). Wild-type poliovirus invades the central nervous system of mice to cause lethal paralysis, but a high-fidelity poliovirus with reduced genetic diversity was unable to reach the spinal cord. Treatment of the high-fidelity variant with a mutagenic compound restored wild-type-like diversity and neurovirulence. No consensus-level mutations were identified in the spinal cord, indicating that low-frequency variants cooperated to permit neural tropism. Several additional studies have demonstrated cooperation in measles virus, influenza virus, HIV-1, and coxsackievirus B3 (Bordería et al., 2015; 2010; Shirogane et al., 2012; Xue et al., 2016), and high-fidelity variants of diverse RNA viruses frequently show restricted tropism and virulence, suggesting that cooperation is likely a widespread phenomenon in viral mutant swarms (Arias et al., 2008; Arnold et al., 2005; Campagnola et al., 2015; Cheung et al., 2014; Coffey et al., 2011; Fitzsimmons et al., 2018; Liu et al., 2013; Meng and Kwang, 2014; Pfeiffer and Kirkegaard, 2003; 2005; Rai et al., 2017; Sadeghipour et al., 2013; Sadeghipour and McMinn, 2013; Sierra et al., 2007; Vignuzzi et al., 2006; 2008; Xie et al., 2014; Zeng et al., 2014; 2013). The mechanisms underlying cooperation are poorly understood and likely highly variable, but they may require multiple genomes to be present inside individual cells (coinfection). Whether coinfection occurs due to multiple virions infecting the same cell or a single virion containing multiple genomes will likely depend upon the virus and host factors

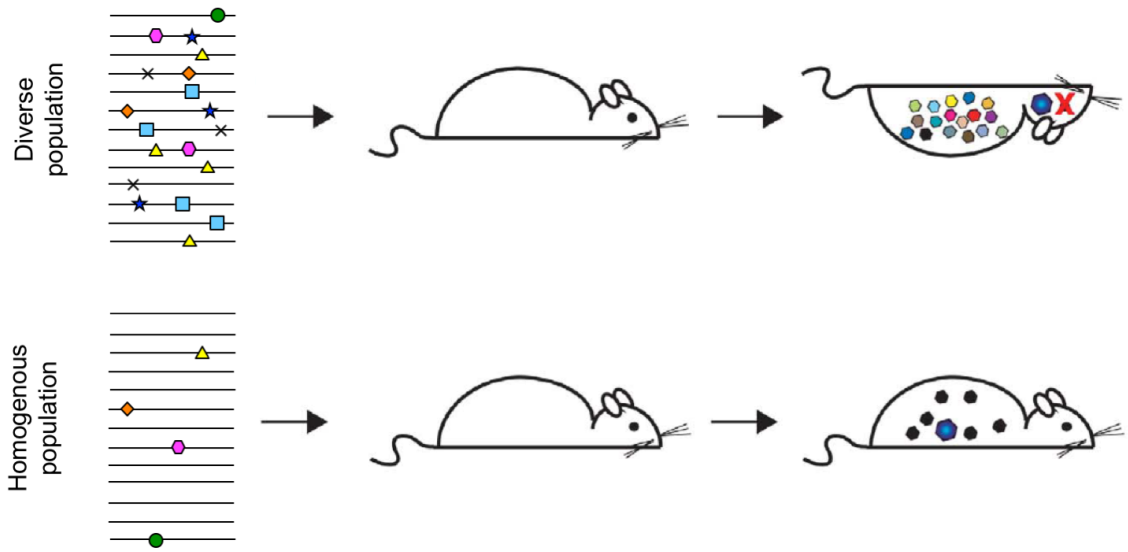


Figure 4. Genetic diversity is important for virulence. Results from Vignuzzi et al. 2006 are presented. Mice inoculated with a diverse population of poliovirus succumb to paralytic disease, while those infected with a homogenous population survive. Figure modified from Lauring and Andino, 2010 under a Creative Commons Attributions License.

(Bordería et al., 2015; Chen et al., 2015; Feng et al., 2013; Shirogane et al., 2012).

High mutation rates also facilitate viral evolution. The standing genetic diversity within mutant swarms represents an enormous reservoir of potentially adaptive mutations. For example, mutations conferring resistance of HIV-1 and hepatitis C virus to antiviral compounds can be found in treatment-naïve patients, likely contributing to the rapid development of antiviral resistance (Jabara et al., 2014; 2011; Lech et al., 1996; Mzingwane et al., 2016; Takeda et al., 2017; Telele et al., 2018). High mutation rates also allow RNA viruses to establish cell- and tissue-specific subpopulations within hosts (Koel et al., 2019; Riemersma et al., 2018; Xiao et al., 2017b). At global scales, high mutation rates allow influenza virus to circumvent vaccine-induced immunity and drive rapid reversion of the live-attenuated oral poliovirus vaccine (OPV), which is now the greatest challenge to global poliovirus eradication (Kew, 2012; Minor, 2009). Finally, MERS-CoV emergence into humans was likely facilitated by two point mutations in a bat coronavirus spike protein (Yang et al., 2015).

Although high mutation rates permit rapid adaptation, “more mutagenic” does not necessarily mean “more adaptable” (Elena and Sanjuán, 2005; Furió et al., 2005). RNA viruses are proposed to replicate close to their error threshold, beyond which deleterious mutations accumulate and disrupt essential genetic information. RNA viruses are therefore always close to “error catastrophe” (Figure 5) (Domingo et al., 2012; Holmes, 2003). Most mutations in RNA viruses are detrimental (Acevedo et al., 2014; Elena and Moya, 1999; Malpica et al., 2002; Peris et al., 2010; Sanjuán et al., 2004a; Theys et al., 2018; Visher et al., 2016), so even marginal increases in mutation rates are expected to reduce viral fitness. Accordingly, altered-fidelity variants with increased mutation rates are generally attenuated (Campagnola et al., 2015; Eckerle et al., 2010;

2007; Gnädig et al., 2012; Kautz et al., 2018; Korboukh et al., 2014; Rai et al., 2017; Riemersma et al., 2018; Rozen-Gagnon et al., 2014; Sexton et al., 2016; Smith et al., 2013; 2015). Therapeutic exploitation of error thresholds with mutagenic compounds is termed lethal mutagenesis (Bull et al., 2007; Crotty et al., 2001; 2002; Moreno et al., 2011; 2012; Smith et al., 2013). The error threshold is thought to constrain RNA genome lengths, as increasing genome size without increasing replication fidelity could lead to error catastrophe (Bradwell et al., 2013). Thus, RNA virus genomes and replicases are evolutionarily fine-tuned to maximize genetic diversity (and concomitantly, replication speed) while maintaining sufficient genetic stability.

What determines evolutionary outcomes? In the simplest view of viral evolution, natural selection drives a population towards high fitness by the canonical "survival of the fittest." Simply defined, a virus' fitness is its ability to replicate in a given environment and is determined by the aggregate effects of all mutations in its genome (its genotype) (Domingo and Holland, 1997; Wargo and Kurath, 2012). Most mutations in RNA viruses are detrimental or lethal, but others have minimal effects on viral fitness and are considered neutral. Relatively few mutations provide an adaptive advantage (Acevedo et al., 2014; Elena and Moya, 1999; Malpica et al., 2002; Peris et al., 2010; Sanjuán et al., 2004a; Theys et al., 2018; Visher et al., 2016). In this paradigm, the outcome of viral evolution is determined by the fitness of individuals within the population (deterministic evolution) (Figure 6). Beneficial mutations rise to fixation by positive selection, detrimental mutations are purged by negative selection, and neutral mutations are ignored by selection. These effects can be visualized as a rugged fitness landscape, with peaks of high fitness and valleys of low fitness, that determine the evolutionary trajectory of a viral population (Figure 7).

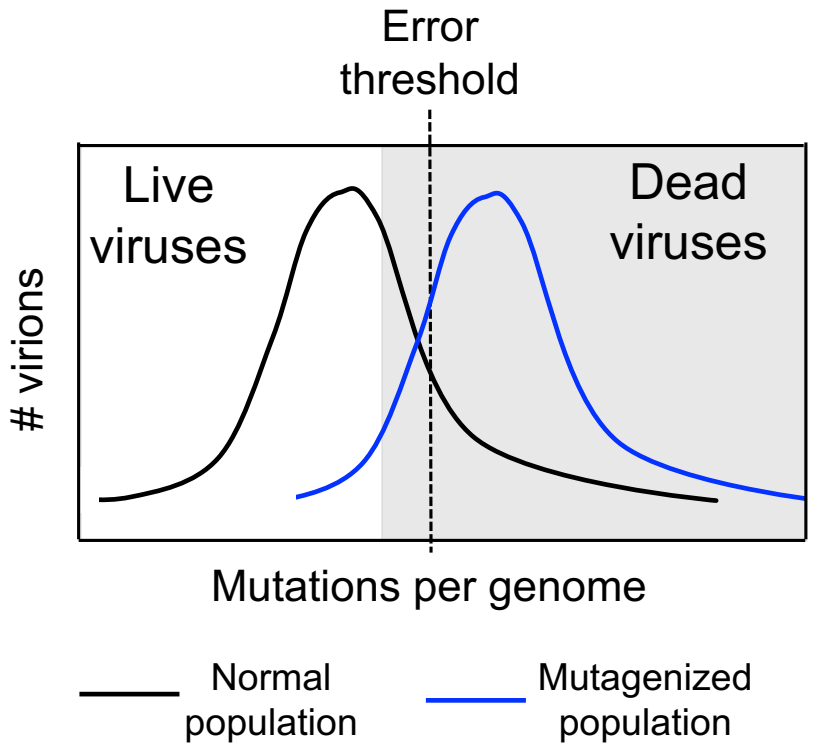


Figure 5. Model for a viral mutant swarm on the brink of error catastrophe. Black and blue lines represent histograms of mutations per genome in a normal or mutagenized population, respectively. The majority of variants in the normal population are live (white), but some proportion are nonviable due to random mutations (gray). An increase in the mutational load (mutagenized population) shifts the majority of variants over the error threshold, where the number of mutations per genome is high enough to extinguish most variants.

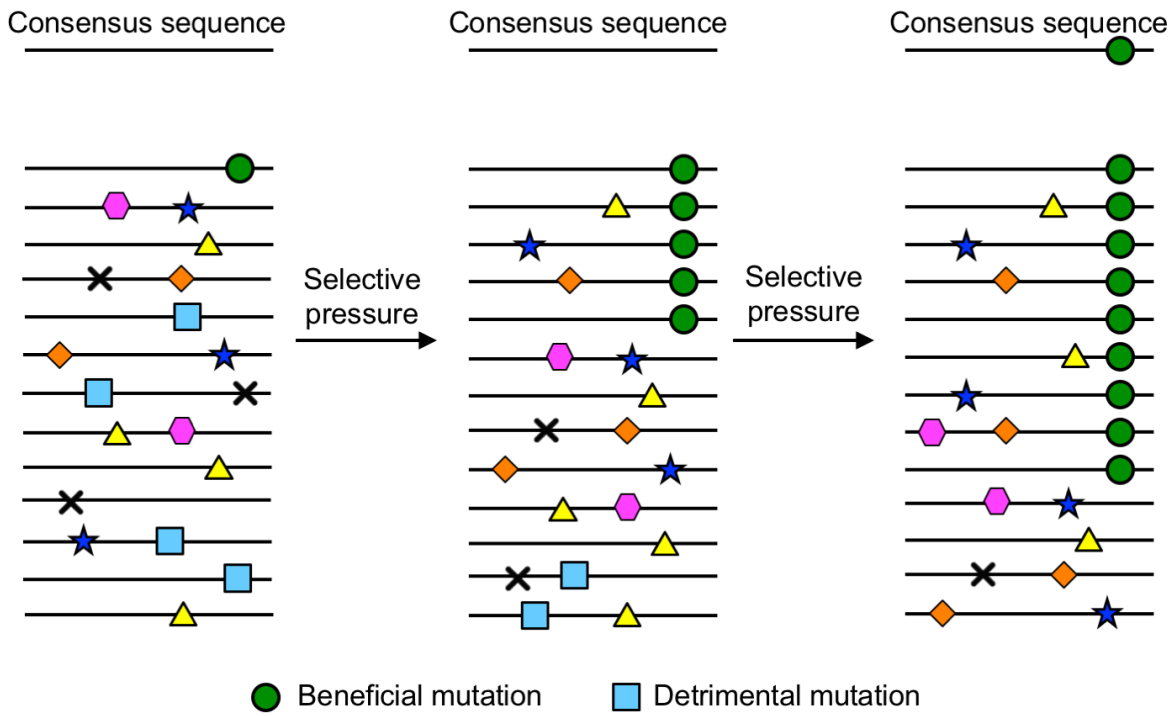


Figure 6. Deterministic evolution by natural selection. A diverse mutant swarm is propagated during multiple rounds of infection. A beneficial mutation (green circle) increases in frequency with each round of replication because it provides an adaptive advantage, while deleterious mutations (blue squares) are lost. New deleterious mutations are generated in each round of replication by chance. The beneficial mutation is not detected at consensus level until it is in greater than 50% of genomes. Horizontal lines indicate individual genomes, and shapes represent mutations.

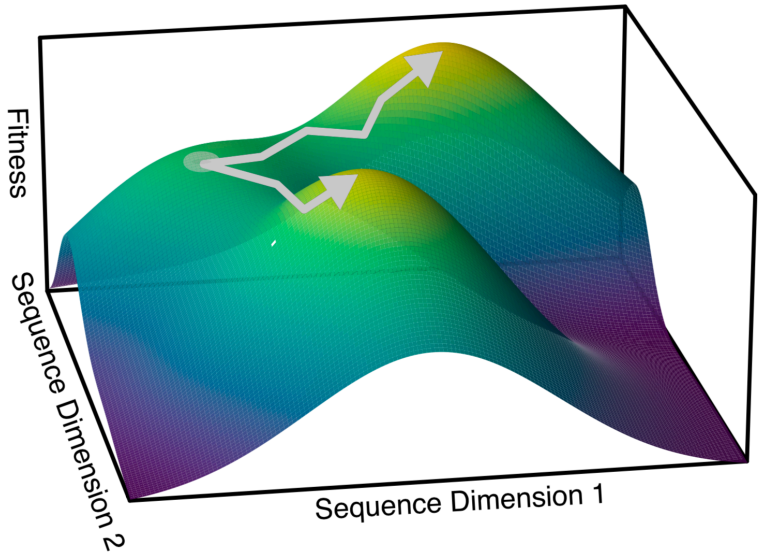


Figure 7. Genotype-fitness landscape. From a starting point (gray dot), a virus explores the genetic landscape (x-y plane) through mutations. Natural selection drives evolution for increased fitness (z-axis) towards local fitness maxima (gray lines). Different trajectories vary in the achievable fitness. Figure from Dolan et al., 2018, with permission.

Natural selection is driven by differences in viral fitness, but fitness is not simply the sum of individual mutational effects. Instead, units of selection are complicated by the effects of epistasis and cooperation. Epistasis occurs when the phenotypic effect of a mutation depends upon other mutations in the genome. Epistatic interactions may make the genome inhospitable to the vast majority of mutations (Elena et al., 2010; Sanjuán et al., 2004b; Van den Bergh et al., 2018). However, epistatic interactions can promote increases in viral fitness. Pauly et al. demonstrated epistasis in influenza virus during the evolution of drug resistance. Two mutations, PB1 T123A and PA T97I, confer 10- and 5-fold reductions in sensitivity to 5-fluorouracil, respectively, while the two in combination yield a 54-fold reduction (Pauly and Lauring, 2015; Pauly et al., 2017a). At larger scales, epistasis makes evolutionary trajectories dependent upon history. For example, the impact of mutations on receptor-binding and antigenicity of the influenza virus hemagglutinin is strain-dependent (Das et al., 2013; Koel et al., 2018; Nakajima et al., 2005). Thus, groups of mutations within individual genomes can be units of selection. Complicating the picture still further is the role of cooperation in determining the phenotype of a mutant swarm (described above), making the entire population a target of natural selection (Bordería et al., 2010; 2015; Shirogane et al., 2012; Xue et al., 2016). Cooperation allows detrimental mutations to be maintained at higher-than-predicted levels due to genetic piggy-backing. Regardless of the scale at which viral fitness is determined—individual mutants, groups of mutations, or collections of genomes—the principles of deterministic evolution still apply.

Though natural selection is important for viral evolution, viral populations are also subject to the stochastic effects of random genetic drift (probabilistic evolution) (Figure 8). All finite populations encounter bottlenecks during their evolutionary history. Bottlenecks cause a sharp reduction in the size of a population and can lead to the fixation of neutral or deleterious

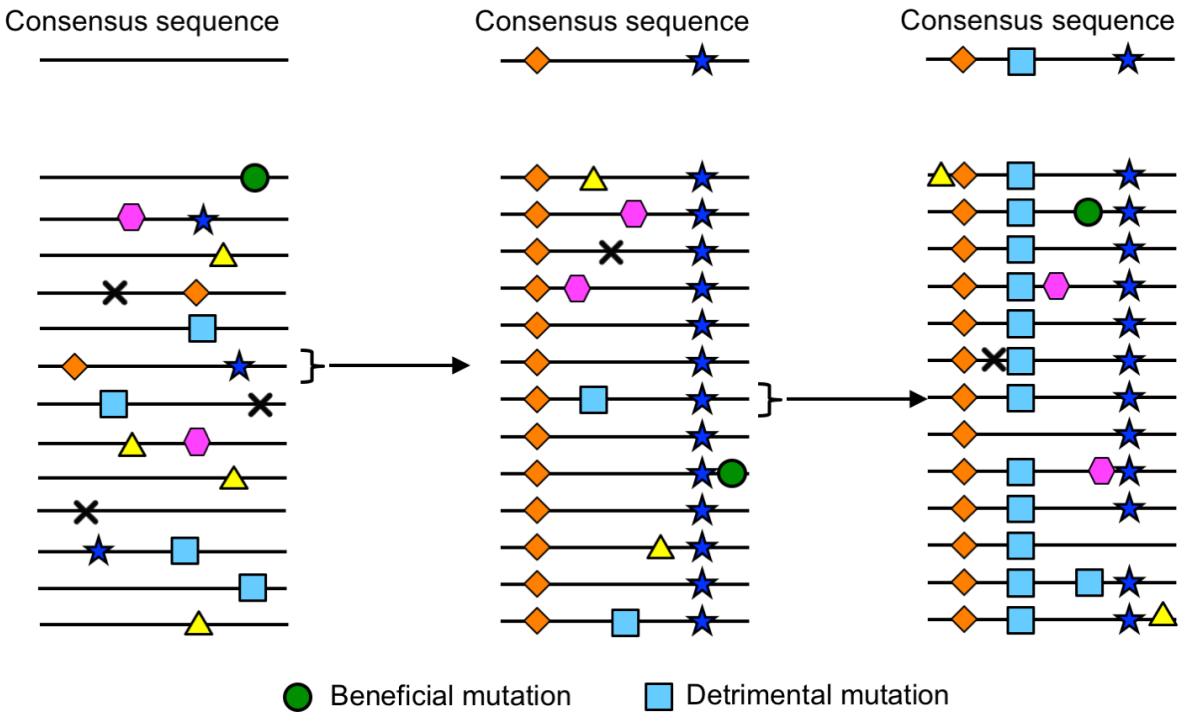


Figure 8. Probabilistic evolution by random genetic drift. A diverse mutant swarm contains beneficial (green circle), detrimental (cyan square), and neutral mutations (other shapes). When only a single genome is replicated during a stringent bottleneck, detrimental and neutral mutations can be fixed at the expense of beneficial mutants. Horizontal lines indicate genomes.

mutations at the expense of beneficial mutations. The relative contributions of deterministic and probabilistic evolution depend upon the population size. The smaller the population, the more the outcome is governed by random chance. In general, bottlenecks slow the rate of adaptation by reducing the likelihood that beneficial mutations are carried into the next generation (Sanjuán et al., 2005). In extreme cases like plaque-to-plaque passage, where only a single infectious unit initiates each generation, mutations accumulate rapidly and cause error catastrophe, an effect known as Muller's ratchet (Duarte et al., 1992). Viruses typically replicate to very high titers within their hosts, but sharp reductions in population size frequently occur during inter-host and cross-species transmission (McCrone and Lauring, 2017). High mutation rates allow viruses to restore genetic diversity and recover fitness after genetic bottlenecks, potentially explaining why high-fidelity mutants are typically attenuated in animal models (Xiao et al., 2017b).

Why study viral evolution? Predictive models for infectious disease emergence are one of the most important goals of biomedical science (Holmes, 2013). Emerging zoonotic infections will always be a risk at the limits of human civilization, which are constantly changing due to climate change, population dynamics, and economic incentives (Geoghegan and Holmes, 2017). Predicting viral evolution remains a significant challenge at multiple scales, from global evolutionary dynamics, to inter-host transmission, to interhost evolution, and to single cells (Dolan et al., 2018a; 2018b; Geoghegan and Holmes, 2017). Thus, understanding the processes that dictate viral evolution are important for anticipating future outbreaks of zoonotic disease, including coronaviruses. Further, disrupting evolutionary dynamics is a well-described approach for developing viral countermeasures. Most altered-fidelity variants described to date are attenuated *in vivo* and protect against reinfection, highlighting their potential utility as live-attenuated vaccines (Graham et al., 2012; Kautz and Forrester, 2018; Korboukh et al., 2014; Lee

et al., 2016; Smith et al., 2014; Vignuzzi et al., 2008; Weeks et al., 2012). Similarly, altering genetic recombination affects evolutionary trajectories and virulence (Graham et al., 2018; Kempf et al., 2016; Xiao et al., 2016; 2017b), and recoding genomes to change genotype-fitness landscapes attenuates poliovirus, coxsackievirus B3, influenza virus, and respiratory syncytial virus (Lauring et al., 2012; Le Nouën et al., 2017; Moratorio et al., 2017; Stobart et al., 2016). These studies underscore the importance of understanding the molecular mechanisms by which RNA viruses generate diversity and the selective forces that drive RNA viral evolution.

Coronaviruses encode a proofreading exoribonuclease

A paradigm of RNA virus biology is error-prone genomic replication due the lack of proofreading or post-replicative repair mechanisms (Sanjuán et al., 2010; Steinhauer et al., 1992). Mutation rates in RNA viruses are inversely correlated with genome length, suggesting that low-fidelity replication constrains genome size (Bradwell et al., 2013). Under this paradigm, the maintenance of large coronavirus genomes requires a corresponding increase in replication fidelity. The first indication that coronaviruses might replicate with higher fidelity than other RNA viruses was the discovery that SARS-CoV encodes a 3′-to-5′ exoribonuclease (ExoN) (Figure 9) (Snijder et al., 2003). The SARS-CoV ExoN is grouped with the DE-D-Dh superfamily of ExoNs that detect and excise nucleotide mismatches during DNA replication (proofreading) (Snijder et al., 2003; Zuo and Deutscher, 2001). A bonafide proofreading activity would be the first of its kind in an RNA virus, and the corresponding increase in replication fidelity could have allowed expansion of large coronavirus genomes. Consistent with this hypothesis, ExoN domains are found in all large members of the *Nidovirus* order (>20 kilobases) but not in those with smaller genomes (Gorbalenya et al., 2006; Lauber et al., 2013).

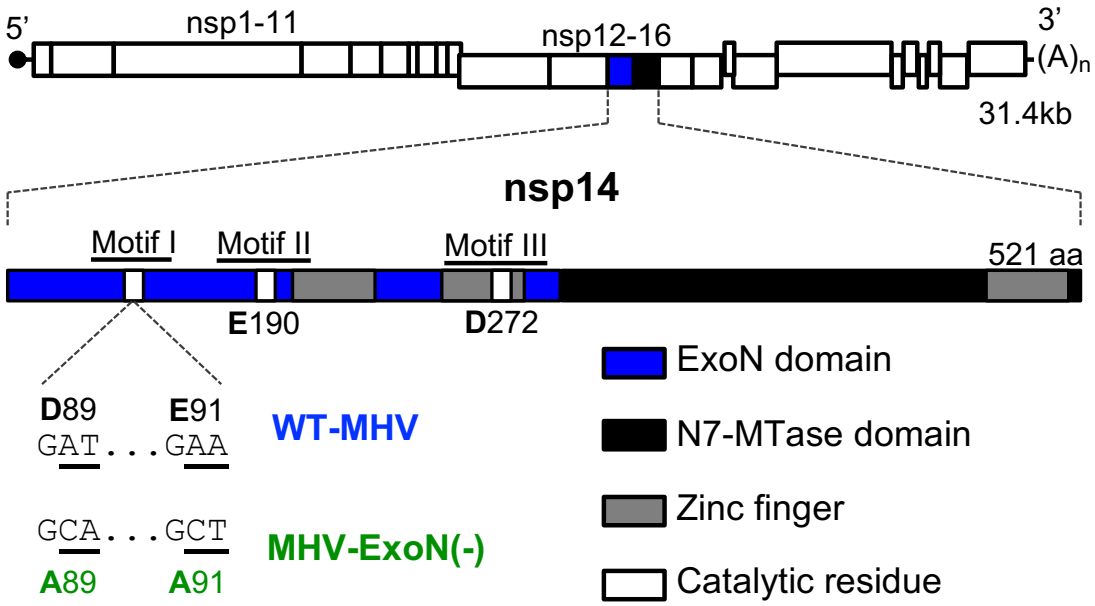


Figure 9. Schematic of MHV-nsp14-ExoN. MHV-nsp14 encodes an exoribonuclease (blue) and an N7-methyltransferase (black) and has 3 zinc fingers (gray boxes) predicted from the solved SARS-CoV nsp10/14 crystal structure (PDB 5C8U, Ma et al. 2015). Catalytic residues for ExoN are marked with white boxes, and the engineered mutations for MHV-ExoN(-) are shown below the genome.

The coronavirus ExoN is encoded within the N-terminal half of nonstructural protein 14 (Figure 9) (Minskaia et al., 2006; Snijder et al., 2003). Bacterially expressed SARS-CoV nsp14 can remove 3′ mismatched nucleotides, including ribavirin 5′-monophosphate, from double-stranded DNA templates (Bouvet et al., 2012; Chen et al., 2007; Ferron et al., 2018; Minskaia et al., 2006). Coronavirus ExoN activity depends on conserved magnesium-coordinating acidic amino acids in three motifs that together constitute the active site (DE-E-D) (Ferron et al., 2018; Ma et al., 2015) (Figure 9). Alanine substitution of CoV motif I DE residues (DE→AA) reduces biochemical ExoN activity in SARS-CoV (Bouvet et al., 2012; Ferron et al., 2018; Ma et al., 2015; Minskaia et al., 2006). Murine hepatitis virus A59 (MHV) and SARS-CoV lacking ExoN activity [ExoN(-)] display defective replication and RNA synthesis, 8- to 20-fold greater mutation frequencies than wild-type, and high susceptibility to nucleoside analogues (Eckerle et al., 2007; 2010; Sexton et al., 2016; Smith et al., 2013; 2015). Recombinant variants of HCoV-229E and transmissible gastroenteritis virus (TGEV) lacking ExoN activity [ExoN(-)] are not viable (Becares et al., 2016; Minskaia et al., 2006). Thus, all data to date support the hypothesis that nsp14-ExoN is the first known proofreading enzyme encoded by an RNA virus and demonstrate the critical importance of this enzyme for coronaviral replication.

The coronavirus replicase-transcriptase complex

The coronavirus replicase-transcriptase complex (RTC) is an intricate, multi-subunit molecular machine reminiscent of DNA replication complexes (Smith et al., 2014) (Figure 10). The keystone of the complex is the RNA-dependent RNA polymerase in nsp12 that assembles the growing RNA strand. The crystal structure of nsp12-RdRp has not been solved, but computational modeling suggests that the catalytic core of the RdRp folds into the canonical

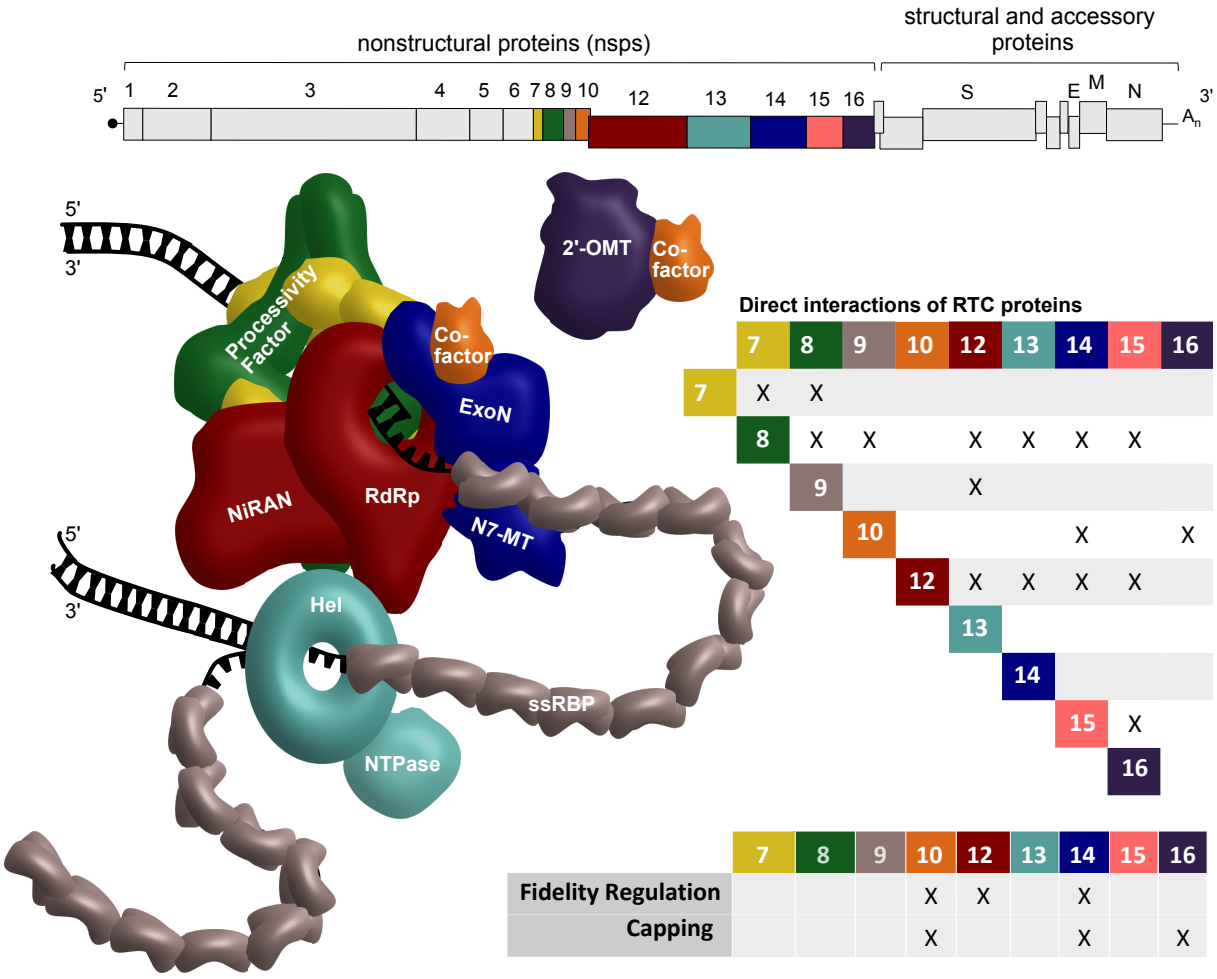


Figure 10. Model for the multi-subunit coronavirus replicase-transcriptase complex (RTC). Coronaviruses encode a large RTC anchored by the nsp12 RNA-dependent RNA polymerase (RdRp) with an associated nucleotidylation domain (NiRAN). The nsp13 helicase (Hel, with attached NTPase) and nsp7+8 processivity factor directly interact with nsp12. The nsp14-exoribonuclease (ExoN) and its nsp10 co-factor increase replication fidelity. Nsp10 also binds and regulates the 2′-*O*-methyltransferase (2′-OMT) in nsp16. Nsp9 is a single-stranded RNA binding protein. Nsp15 is an endoribonuclease but may not be associated directly with the RTC. This model is derived from the theoretical, biochemical, and virological evidence cited in the text and is used with permission from Sexton, 2017.

fingers-palm-thumb arrangement characteristic of viral RdRps (Sexton et al., 2016; Xu et al., 2003). The CoV RdRp regulates fidelity by mechanisms similar to those in other RNA viruses but has lower overall fidelity (Ferron et al., 2018; Sexton et al., 2016). An accessory domain is encoded in the N-terminal half of nsp12, fused to the RdRp core. This domain is highly conserved among all nidoviruses, suggesting an important role in nidoviral biology. The role of this domain is poorly understood, in part because it has low predicted structural homology to known proteins and has therefore been difficult to model with high confidence. However, an essential nucleotidylation activity is encoded by this domain, now termed the nidovirus RdRp-associated nucleotidylation domain (NiRAN) (Lehmann et al., 2015), and a small piece of the NiRAN domain is important for localizing nsp12 to replication sites (Brockway et al., 2003).

Several additional proteins assist the CoV nsp12-RdRp in maintaining efficient and accurate RNA synthesis. Multimers of nsps7 and 8 assemble into a ring-like structure that increase nsp12's RNA-binding efficiency and processivity (Subissi et al., 2014; Velthuis et al., 2012; Xiao et al., 2012). Nsp8 may also synthesize primers to initiate RNA synthesis (Imbert et al., 2006). A helicase encoded by nsp13 unwinds double-stranded RNA, and nsp9 may protect the single-stranded RNA from reannealing or degrading (Adedeji et al., 2012; Egloff et al., 2004; Smith et al., 2014). The proofreading nsp14-ExoN can assemble into a complex with nsps7, 8, 10, and 12 without disrupting polymerization or excision (Subissi et al., 2014). Nsp10 is a nonenzymatic cofactor that orients the catalytic DE-E-D motifs to stimulate nsp14-ExoN activity (Donaldson et al., 2007; Ferron et al., 2018; Subissi et al., 2014). The model of an RTC composed of nsps7-14 is supported by interaction studies and by genetic analyses showing strong purifying selection against mutations in these proteins, suggesting that they are highly coevolved and interdependent (Brunn et al., 2007; Forni et al., 2016b; Sawicki et al., 2005).

Components of the RTC have additional functions beyond RNA synthesis. Viral mRNAs are capped by the sequential action of two enzymes: the N7-methyltransferase domain in the C-terminal half of nsp14 and the nsp16 2′-*O*-methyltransferase (Bouvet et al., 2010; Case et al., 2016; Chen et al., 2009; Decroly et al., 2008; Sevajol et al., 2014). In addition to its role in regulating fidelity via nsp14-ExoN, nsp10 also promotes capping through its interaction with nsp16 (Bouvet et al., 2014; Decroly et al., 2011). The interaction surface is similar between nsp10-nsp14 and nsp10-16, though the former is substantially larger (Decroly et al., 2011; Ferron et al., 2018; Lugari et al., 2010; Ma et al., 2015). Additionally, nsp14 has been implicated in evading innate immune responses, perhaps by degrading used or damaged viral RNA to prevent host recognition (Becares et al., 2016; Case et al., 2017). Finally, nsp15 encodes a uridylate-specific endoribonuclease involved in evasion of host immune responses, but it is not clear whether this protein should be considered part of the coronavirus RTC (Deng et al., 2017; Kindler et al., 2017; Snijder et al., 2003).

Summary

Faithful replication of large and complex coronavirus genomes requires correspondingly large and complex replication machinery, with the first known proofreading activity in the RNA world. Individual components of the coronavirus RTC have been studied in detail, revealing conceptually new strategies for live-attenuation of coronavirus vaccines (Graham et al., 2012; Menachery et al., 2014; 2018; Schindewolf and Menachery, 2019). However, studies examining complexes of coronavirus replicase proteins remain relatively scarce. Further, the coronavirus RTC likely lies in close proximity to most, if not all, coronavirus-encoded proteins (Brockway et al., 2004; 2003; Brunn et al., 2007; Kuo et al., 2016; Sawicki et al., 2005; V'kovski et al., 2019),

but to date, models of coronavirus replication have largely ignored potential interactions outside of nsps7-16. Unraveling the connections between coronavirus replicase proteins (and potential collaborators across the genome) is essential for understanding their evolutionary dynamics and may reveal new tactics to counter emerging coronavirus infections.

In Chapter 2, I demonstrate that experimental evolution of MHV lacking proofreading activity compensates for the loss of ExoN activity in part through a high-fidelity polymerase. In Chapter 3, I describe experiments to identify fidelity-regulating proteins outside of nsp12-RdRp and nsp14-ExoN. In Chapter 4, I tackle the remarkable observation that despite the critical importance of ExoN activity for viral fitness, ExoN(-) coronaviruses do not revert in tissue culture or in animals. In Chapter 5, I describe two new tools for detailed analysis of coronavirus diversity and fitness. In Chapter 6, I summarize the materials and methods used in this dissertation. Finally, in Chapter 7, I examine the broader implications of this work and outline avenues for continuing investigation.

CHAPTER 2: A PROOFREADING-DEFICIENT CORONAVIRUS ADAPTS FOR INCREASED FIDELITY AND FITNESS OVER LONG-TERM PASSAGE WITHOUT REVERSION OF EXORIBONUCLEASE-INACTIVATING MUTATIONS

Introduction

At the beginning of this dissertation research, the coronavirus exoribonuclease (ExoN) had been established as a critical regulator of efficient and accurate replication in both murine hepatitis virus (MHV) and SARS-CoV (Eckerle et al., 2007; 2010; Graham et al., 2012; Smith et al., 2013). Yet, despite the fitness costs associated with disruption of ExoN motif I, neither MHV-ExoN(-) or SARS-CoV-ExoN(-) had reverted under any experimental conditions. MHV-ExoN(-) and SARS-CoV-ExoN(-) motif I were genetically stable during 17 and 20 passages in cell culture, respectively, and during treatment with RNA mutagens (Eckerle et al., 2007; 2010; Sexton et al., 2016; Smith et al., 2013; 2015). We also had not detected motif I reversion in SARS-CoV-ExoN(-) after eight acute passages in aged BALB/c mice and 60 days of persistent infection in immunodeficient Rag-/- mice (Graham et al., 2012). To test the stability of the ExoN(-) genotype and phenotype, we subjected MHV-ExoN(-) to long-term passage (250 passages over one year; P250). We demonstrate that MHV-ExoN(-) did not extinguish during passage and adapted for increased replication. MHV-ExoN(-) concurrently evolved reduced susceptibility to multiple nucleoside and base analogues, consistent with selection for increased replication fidelity. Importantly, the ExoN-inactivating substitutions did not revert. The evolved mutations in MHV-ExoN(-) nsp14 and nsp12, which encodes the RdRp, accounted for only part of the increased nucleoside analogue resistance of MHV-ExoN(-) P250, implicating multiple replicase proteins in adaptation for viral fitness. These data support the proposed link between

CoV fidelity and fitness, demonstrate the surprising stability of the ExoN-inactivating substitutions, and identify additional proteins outside of nsp12 and nsp14 that may contribute to CoV fidelity regulation.

Coauthor contributions

Xiaotao Lu and Brett Case completed the P250 passage series. Xiaotao cloned and recovered all recombinant viruses except for the ExoN(-) silent reference, which Nicole Sexton generated. Clint Smith and Xiaotao obtained the full-length sequence data for the P250 viruses. Clint Smith and I performed at least one independent experiment for each panel, with few exceptions: Clint performed all replication curves, and I performed the competitive fitness and specific infectivity assays. Tia Hughes prepared PrimerID libraries.

Results

Long-term passage of WT-MHV and MHV-ExoN(-). We serially passaged WT-MHV and MHV-ExoN(-) in delayed brain tumor (DBT) cells 250 times (P250). Virus from each passage was harvested once 50-100% of the monolayer was involved in syncytia, which occurred between 8 and 24 hours post-infection. Passage conditions varied for WT-MHV and MHV-ExoN(-) due to differences in replication kinetics between the two viruses. We stopped passage at P250 after observing reduced syncytia formation in MHV-ExoN(-)-infected flasks, likely resulting from a mutation in the MHV-ExoN(-) P250 spike protein cleavage site (discussed below).

MHV-ExoN(-) and WT-MHV replicate with identical kinetics following 250 passages. MHV-ExoN(-) has a significant replication defect relative to WT-MHV (Eckerle et al., 2007).

We first tested whether replication of MHV-ExoN(-) P250 was affected by long-term passage by examining replication at two different multiplicities of infection (MOI). At both MOI = 1 and MOI = 0.01 PFU/cell, MHV-ExoN(-) P3 replication was delayed by ~2 hours and peak titer reduced by ~1 $\log_{10}$ relative to WT-MHV P3 (Figure 11A and B), consistent with our previous studies (Eckerle et al., 2007). By P250, both viruses replicated with identical kinetics (Figure 11, A and B, dotted lines). This represented a ~1 $\log_{10}$ increase in peak replication for WT-MHV and a ~2 $\log_{10}$ increase for MHV-ExoN(-), compared with the respective parental virus. At MOI = 0.01 PFU/cell, we also measured replication of MHV-ExoN(-) at P10, P50, P100, and P160. Replication kinetics gradually increased over passage, reaching P250-like levels by P100 (Figure 11B). To determine whether the increased replication of MHV-ExoN(-) P250 was affected by the presence of potential defective viral genomes or by some other population-based phenomenon, both WT-MHV P250 and MHV-ExoN(-) P250 were plaque purified three times (Figure 12). The plaque-purified viruses replicated indistinguishably from the parent populations (Figure 11C). Together, these data demonstrate that WT-MHV and MHV-ExoN(-) populations adapted for increased replication and that either individual genomes or those derived from a single virus plaque encoded the adaptive changes required by the total population.

MHV-ExoN(-) accumulated 8-fold more mutations than WT-MHV but did not revert ExoN-inactivating substitutions. To determine whether the increased replication of MHV-ExoN(-) P250 resulted from primary reversion of ExoN(-) motif I, we sequenced nsp14 from infected-cell total RNA. MHV-ExoN(-) P250 retained the motif I DE→AA substitutions, demonstrating that primary reversion of ExoN(-) motif I did not occur. To identify potentially adaptive consensus mutations, we performed full-genome di-deoxy sequencing of MHV-ExoN(-) P250 and WT-MHV P250. Within WT-MHV P250, we identified 23 mutations, of which 17

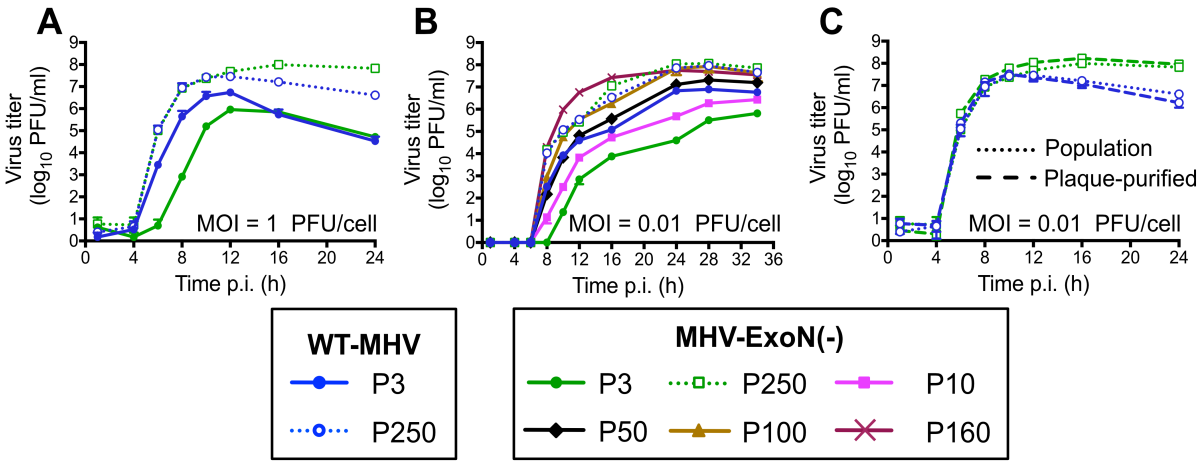


Figure 11. MHV-ExoN(-) evolved increased replicative capacity over long-term passage. Replication kinetics were examined for the indicated viruses at MOI = 1 PFU/cell (*A*) and MOI = 0.01 PFU/cell (*B*). (*C*) Replication kinetics of plaque-purified WT-MHV P250 and MHV-ExoN(-) P250 in parallel with the full population (MOI = 0.01 PFU/cell). Supernatants were collected at the indicated times post infection, and titers were determined by plaque assay. Data for *A*-*C* represent mean and standard deviation of n = 3.

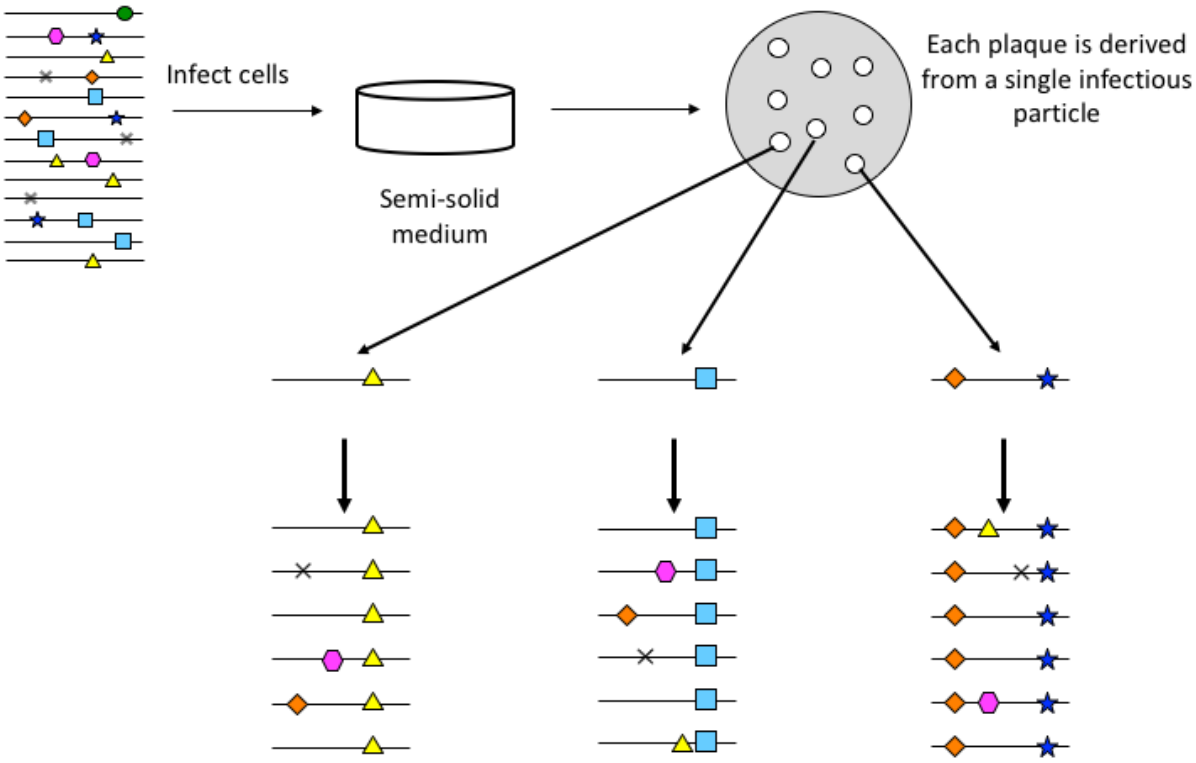


Figure 12. Plaque purification of viral populations. A viral population contains distinct but closely-related genomes. When viruses are cultured under semi-solid media, they are unable to disperse through the supernatant and can only infect adjacent cells. Over multiple replication cycles, they form distinct plaques derived, in principle, from a single infectious particle. Propagating virus from individual plaques yields distinct populations. Horizontal lines represent individual genomes, and shapes represent mutations.

were nonsynonymous (NS) (Figure 13A). In contrast, MHV-ExoN(-) P250 had 171 total mutations (74 NS) (Figure 13B). The full-genome sequences have been deposited in GenBank, and the mutations for both viruses are listed in Appendix A.1 and A.2. We identified only one mutation shared by both viruses (nsp1 A146T), though it was present in approximately 50% of the WT-MHV P250 population by di-deoxy sequencing. Both viruses deleted most of the hemagglutinin esterase (HE). In MHV-A59, HE mRNA is not transcribed *in vitro* (Luytjes et al., 1988; Yokomori and Lai, 1991; Yokomori et al., 1991), and HE protein expression is detrimental to MHV-A59 fitness in cell culture (Lissenberg et al., 2005). WT-MHV P250 also deleted ORF 4a, which is dispensable for MHV replication in cell culture (Gadlage et al., 2008). The C-terminal region of ns2 within MHV-ExoN(-) P250 was truncated and fused to HE with a -1 frameshift. Ns2 is a phosphodiesterase (PDE) that protects viral RNA by degrading 2′→5′ oligoadenylate, the activating factor for cellular RNase L (Li and Weiss, 2016; Zhang et al., 2013; Zhao et al., 2012). The portion of ns2 deleted in MHV-ExoN(-) P250 lies outside the PDE catalytic domain, in a region of unknown function. C-terminally truncated ns2 retains enzymatic activity (Sui et al., 2016), but whether this specific deletion and fusion disrupts PDE activity remains to be tested. Nevertheless, ns2 is dispensable for MHV replication in immortalized cells (Schwarz et al., 1990; Zhao et al., 2011). Details about the deletion sites are provided in Figure 14. Within proteins predicted to be part of the replicase-transcriptase complex (nsp7-16 and nucleocapsid) (Smith et al., 2014), WT-MHV P250 had only one NS change, located in the nsp13-helicase (Figure 13A and Appendix A.1). In contrast, MHV-ExoN(-) P250 had 17 NS changes within this region (Figure 13B and Appendix A.2).

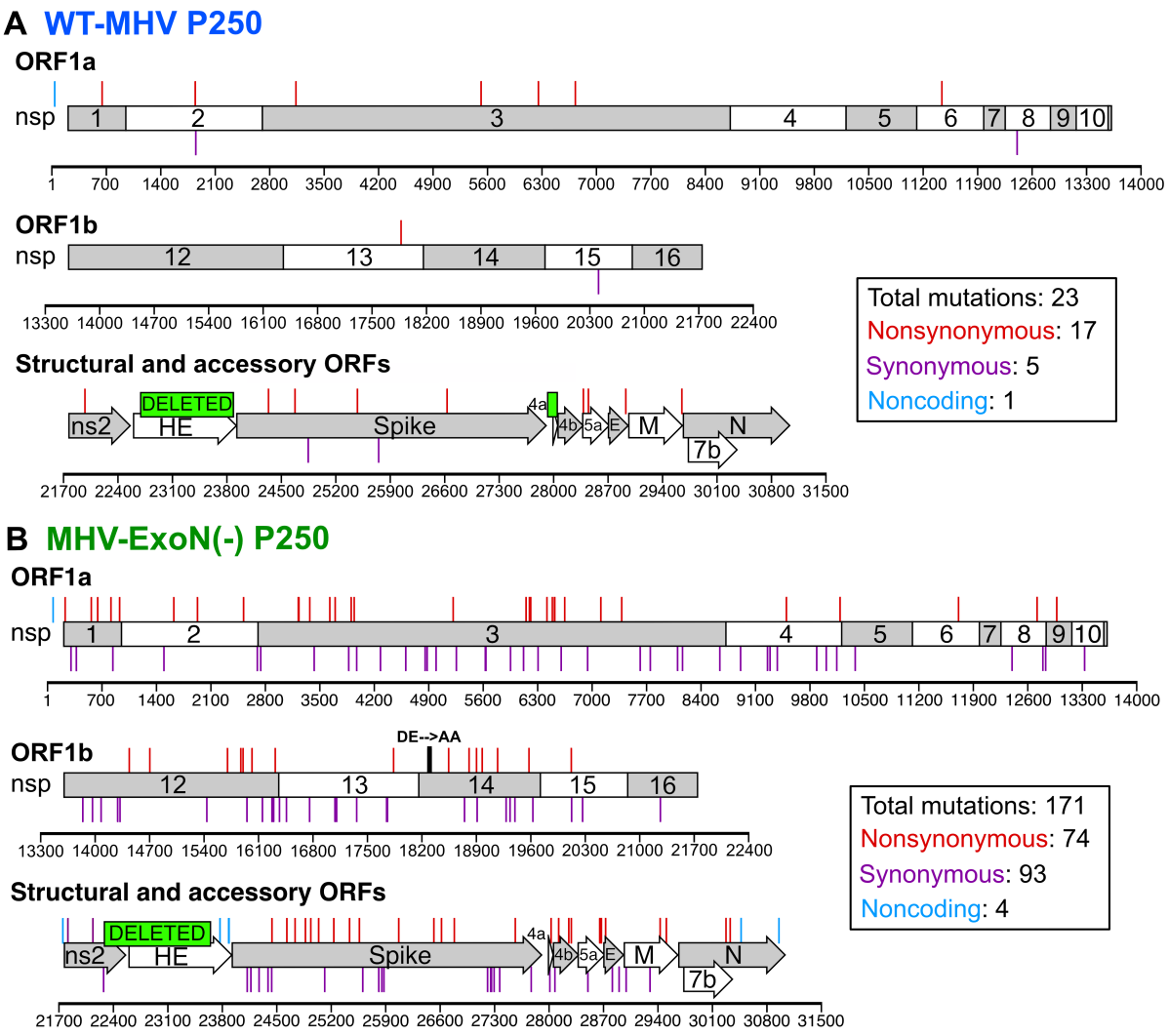


Figure 13. Mutations within P250 viruses. Mutations present at >50% by di-deoxy sequencing at passage 250 in WT-MHV (*A*) and MHV-ExoN(-) (*B*). Nonsynonymous mutations (red), noncoding mutations (cyan), and deletions (green boxes) are plotted above the schematic, and synonymous mutations (purple) are below.

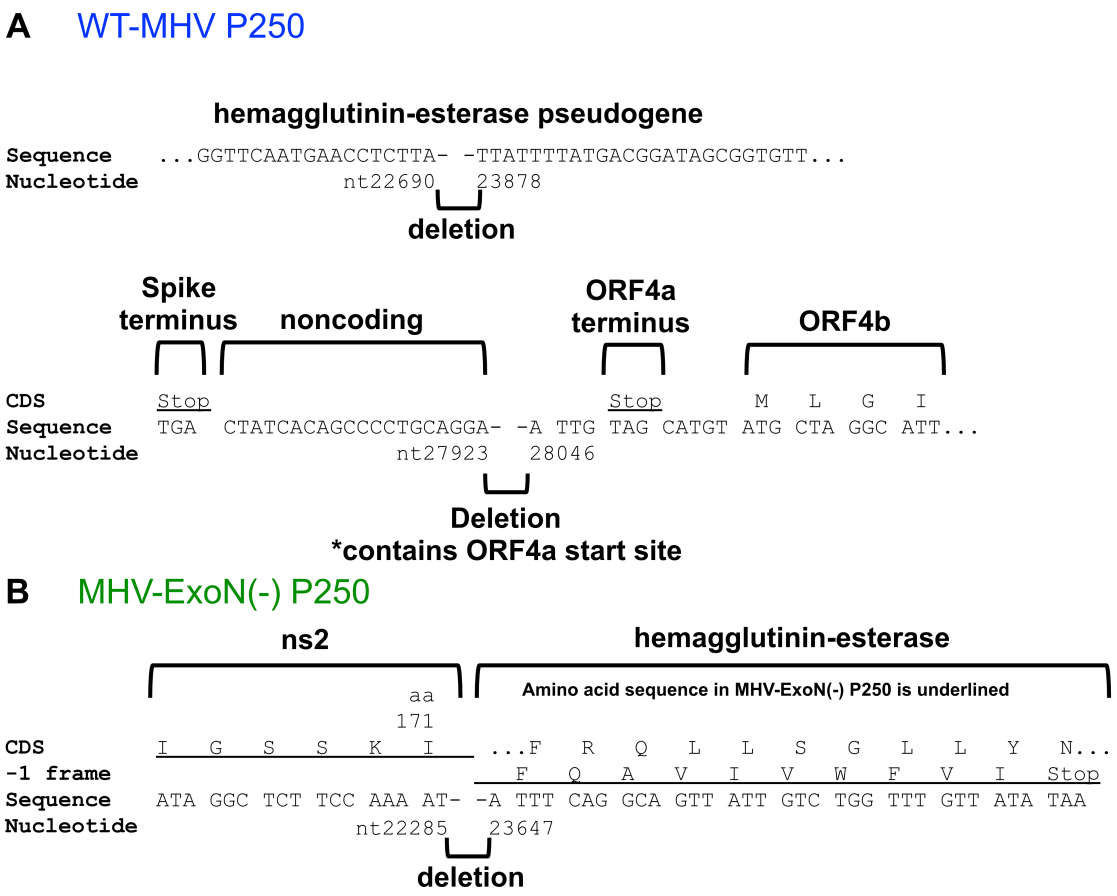


Figure 14. Deleted regions within WT-MHV P250 and MHV-ExoN(-) P250.

MHV-ExoN(-) P250 displays increased genomic RNA accumulation and increased resistance to 5-fluorouracil. Coronaviruses lacking ExoN consistently display defects in RNA synthesis relative to WT (Eckerle et al., 2007; Minskaia et al., 2006; Smith et al., 2013). To determine whether the increased replication of MHV-ExoN(-) P250 was associated with restored genomic RNA (gRNA) production, we measured gRNA accumulation over time using two-step real-time quantitative PCR (Smith et al., 2013; 2015). MHV-ExoN(-) P250 accumulated similar levels of gRNA to WT-MHV P3 and WT-MHV P250 at early time points, while gRNA levels for MHV-ExoN(-) P3 were ~1 $\log_{10}$ lower (Figure 15A). MHV-ExoN(-) P250 gRNA levels fell below those of WT-MHV and WT-MHV P250 after 8 hours and were similar to those of MHV-ExoN(-) P3 at 10 hours post-infection. Normalizing to the gRNA abundance at four hours for each virus demonstrated that the rates of gRNA accumulation were similar for all four viruses (Figure 15B). These data suggest that the increased replication of P250 viruses relative to WT-MHV is not fully accounted for by increased RNA synthesis. In addition to RNA synthesis defects, ExoN(-) CoVs have up-to 20-fold increased mutation frequencies and profoundly increased sensitivity to nucleoside and base analogues relative to WT CoVs (Eckerle et al., 2007; 2010; Graham et al., 2012; Sexton et al., 2016; Smith et al., 2013). To determine whether nucleoside analogue sensitivity of MHV-ExoN(-) was altered by long-term passage, we treated cells infected with parental and passaged viruses with the base analogue, 5-fluorouracil (5-FU). 5-FU is converted intracellularly into a nucleoside analogue that incorporates into growing RNA strands and causes A:G and U:C mutations. For simplicity, I hereafter refer to 5-FU as a nucleoside analogue. Incorporation of 5-FU is increased in the absence of ExoN activity (Smith et al., 2013). All viruses displayed a concentration-dependent decrease in viral titer but differed greatly in their susceptibility to 5-FU (Figure 15C). At 120μM, WT-MHV P3 titers were reduced

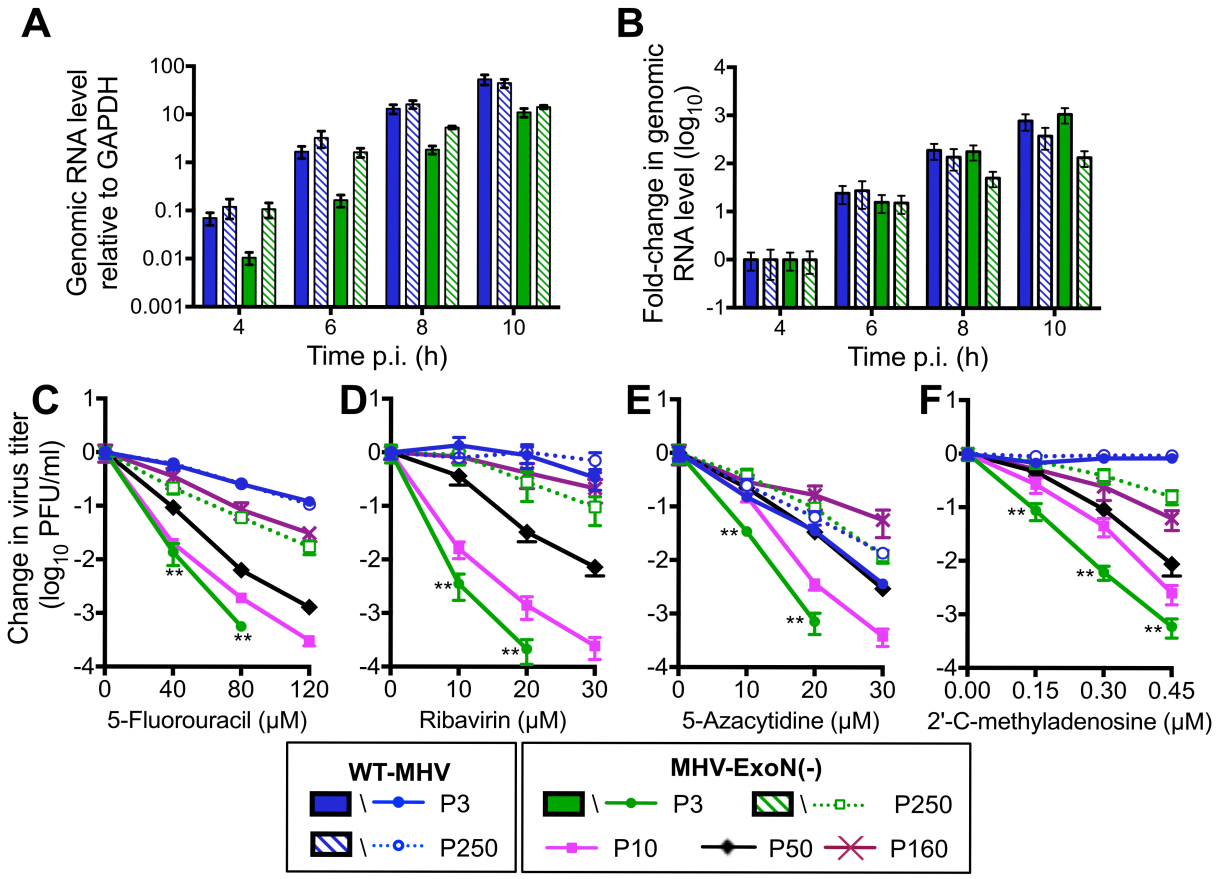


Figure 15. MHV-ExoN(-) evolved WT-like genomic RNA accumulation and increased resistance to multiple nucleoside analogues over passage. (*A*) Cells were infected with the indicated viruses at MOI = 1 PFU/cell, and intracellular RNA was harvested using TRIzol at the indicated times post infection. MHV genomic RNA was detected using SYBR green and primers directed to nsp10, and values were normalized to intracellular GAPDH. (*B*) Same data as in panel (*A*) normalized to RNA level for each virus at 4 hours post infection. Data represent mean and standard error for n = 9 (3 triplicate experiments). (*C-F*) Sensitivity of passaged viruses to nucleoside analogues at MOI = 0.01 PFU/cell. Cells were treated with the indicated concentrations of 5-FU (*C*), RBV (*D*), AZC (*E*), or CMeA (*F*) for 30 minutes prior to infection, and supernatants were harvested at 24 hours post-infection and titered by plaque assay. Data represent change in titer relative to untreated control and are plotted as mean and standard error of n = 6 (two triplicate experiments). For panel C-F, statistical significance for change in titer of MHV-ExoN(-) P3 relative to MHV-ExoN(-) P250 was determined using Mann-Whitney test (**P<0.01).

by ~1 $\log_{10}$, while MHV-ExoN(-) P3 titers were undetectable (> 5 $\log_{10}$-fold reduction). WT-MHV 5-FU sensitivity was not altered by passage. MHV-ExoN(-) P250 was less susceptible than MHV-ExoN(-) P3 to 5-FU treatment, with only a ~1.5 $\log_{10}$ decrease in titer at 120 μM. MHV-ExoN(-) P250 remained more sensitive to 5-FU than WT-MHV, suggesting that WT-like resistance requires an intact ExoN. These data demonstrate that MHV-ExoN(-) P3 evolved resistance to 5-FU through mutations outside of ExoN(-) motif I.

Spike mutations in MHV-ExoN(-) P250 do not increase resistance to 5-FU. Bacteriophage ϕX174 acquired resistance to 5-FU by delaying cell lysis, thereby reducing the number of replication cycles in which 5-FU can be incorporated (Pereira-Gómez and Sanjuán, 2014). MHV-ExoN(-) P250 had multiple mutations in the spike glycoprotein, including one in the spike furin cleavage site that reduced syncytia formation. To test whether the spike mutations manifested in resistance to 5-FU, we cloned the spike gene from MHV-ExoN(-) P250 into the isogenic MHV-ExoN(-) background. The recombinant virus demonstrated intermediate replication kinetics between MHV-ExoN(-) P3 and MHV-ExoN(-) P250 (Figure 16A) and did not form syncytia. Spike-P250 also increased the specific infectivity of viral particles (Figure 16B). However, the MHV-ExoN(-) P250 spike did not affect sensitivity of the recombinant virus to 5-FU (Figure 16C). Thus, any adaptive increase in 5-FU resistance must be located elsewhere in the genome.

MHV-ExoN(-) passage resulted in unique mutations in nsp12 and nsp14. To date, three proteins have been shown to alter CoV sensitivity to 5-FU: nsp12-RdRp, nsp14-ExoN, and nsp10, which stimulates ExoN activity (Sexton et al., 2016; Smith et al., 2014; 2015). Neither

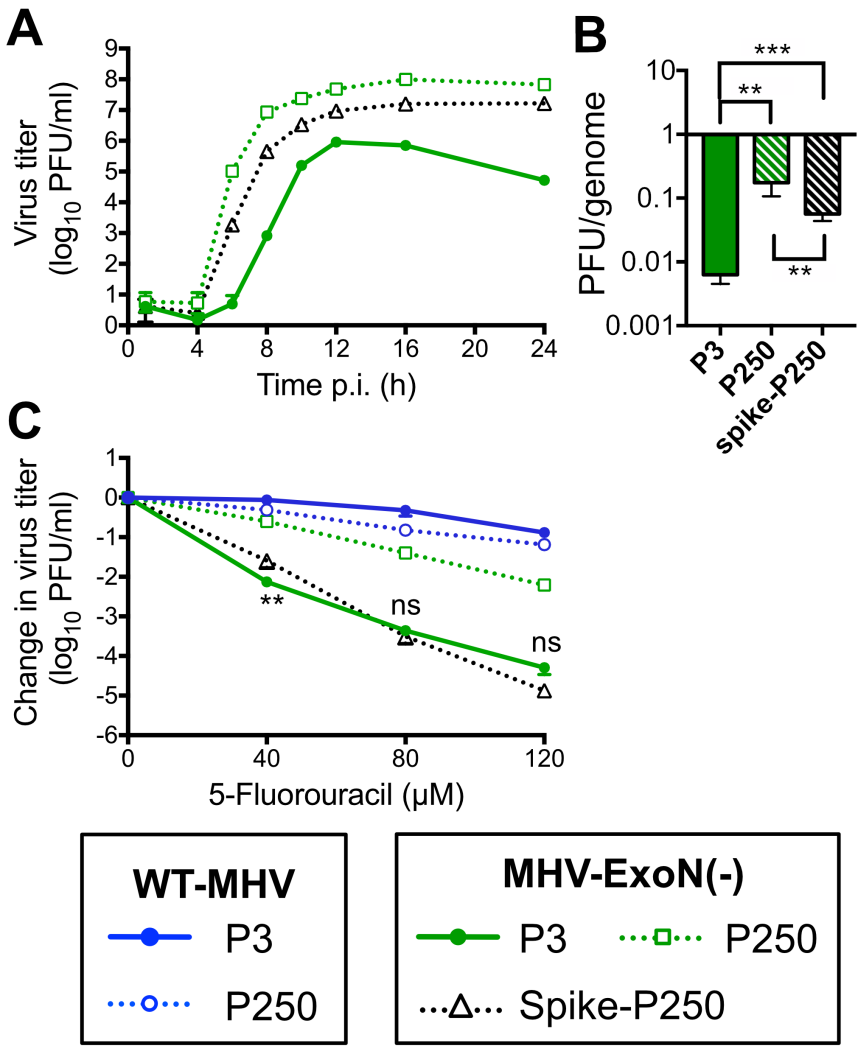


Figure 16. Mutations in the spike envelope protein from MHV-ExoN(-) P250 increase replicative capacity but do not affect sensitivity to 5-fluorouracil. (*A*) Replication kinetics of indicated viruses (MOI = 0.01 PFU/cell) plotted as mean and standard deviation with n = 3. (*B*) Specific infectivity of indicated viruses 12 hours post-infection (MOI = 1 PFU/cell). Data represent mean and standard error of n = 6 (two triplicate experiments). (*C*) Sensitivity of indicated viruses to 5-fluorouracil at MOI = 0.01 PFU/cell, as described in Fig. 4. Data represent mean and standard error of n = 6 (two triplicate experiments). For panel B, statistical significance was determined using one-way ANOVA. For Panel C, statistical significance for change in titer of MHV-ExoN(-) spike-P250 relative to MHV-ExoN(-) P3 was determined using Mann-Whitney test (**P<0.01, ***P<0.001, ns = not significant).

WT-MHV nor MHV-ExoN(-) P250 contained a NS mutation in nsp10, and WT-MHV P250 had no mutations within either nsp12 or nsp14. In contrast, MHV-ExoN(-) P250 had 7 NS mutations in nsp12 and 6 NS mutations in nsp14 (Figure 13 and Figure 17), none of which have been described previously *in vitro* or in viable viruses. Within nsp12, six mutations were in the predicted RdRp fingers, palm, and thumb domains (Figure 17A) (Ng et al., 2008). Four residues (H709, F766, S776, and M814) can be visualized on a Phyre^2-modeled structure of the MHV-nsp12 RdRp, while the remainder lie outside the modeled core RdRp (Figure 17A) (Sexton et al., 2016). One mutation, M288T, lies in the CoV-specific domain, which is conserved among nidoviruses. This domain has been implicated in membrane targeting in MHV-A59 (Brockway et al., 2003) and performs an essential nucleotidylation activity in the arterivirus, equine arteritis virus (Lehmann et al., 2015). However, M288T is not predicted to catalyze nucleotidylation. Within nsp14, 4 NS mutations were identified in the ExoN domain, and 2 NS mutations were in the C-terminal N7-methyltransferase domain (Figure 17B). We next modeled the structure of MHV nsp14 using Phyre^2 software (Kelley et al., 2015), resulting in highest-probability similarity to the SARS-CoV nsp14-nsp10 complex (PDB: 5C8S) (Ma et al., 2015) with high-confidence (i.e. the calculated probability of true homology between the structures) of 100% for residues 3-519 of MHV-nsp14. The model predicts that five mutations are located close to surface of the protein (Figure 17B). All three modeled zinc finger domains contain one NS mutation (F216Y, Y248H, L473I). Two mutations, D128E and F216Y, are located near the interface between nsp10 and nsp14, though neither site has previously been implicated in nsp10-nsp14 interaction (Bouvet et al., 2014; Donaldson et al., 2007; Smith et al., 2015). One NS mutation resulted in a D272E substitution in ExoN motif III, a metal-coordinating active site residue. We previously reported that alanine substitution of D272 results in an ExoN(-) phenotype (Eckerle et al., 2007), but the viability or phenotype of a D272E substitution was not

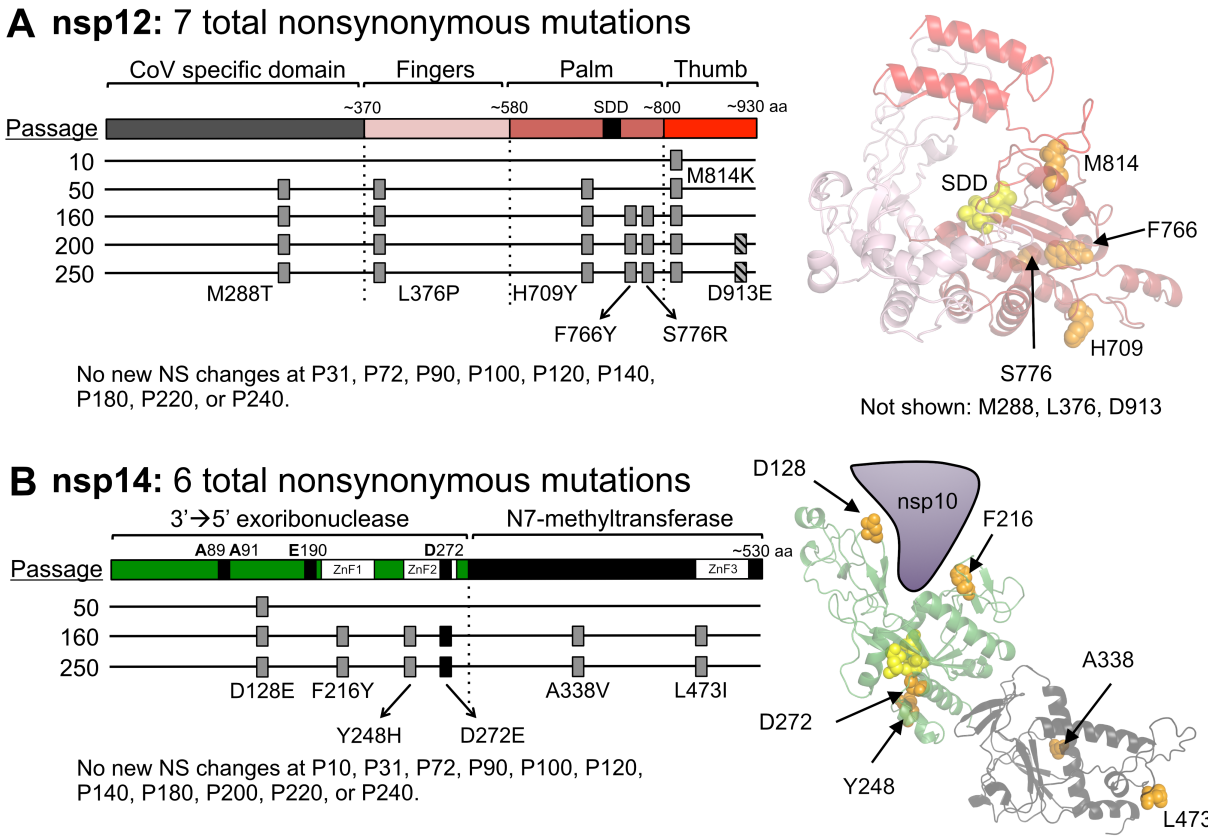


Figure 17. The timing of fixation of mutations in nsp12-RdRp and nsp14-ExoN within MHV-ExoN(-). (*A*) A schematic of nsp12-RdRp with the CoV-specific region and the canonical fingers, palm, and thumb domains of RdRps is shown. The nsp12-RdRp coding region was sequenced at the indicated passage, and the nonsynonymous changes are plotted; gray boxes indicate consensus changes, and hatched boxes indicate variants present at <50% of the population by di-deoxy sequencing. At right, mutations are marked in orange on a Phyre[2]-modeled structure of MHV-nsp12, with the active site residues marked in yellow (Sexton et al., 2016). RdRp domains are colored according to the linear schematic. M288T, L376P, and D913E lie outside the modeled region and thus are not marked. (*B*) A schematic of nsp14 with the ExoN and N7-methyltransferase domains is shown, with mutation plotting as in panel A. The black box denotes a mutation to ExoN motif III. At right, mutations are marked in orange on a Phyre[2]-modeled structure of the MHV-nsp14/nsp10 complex. Domains are colored according to the linear schematic.

tested in that study. These data suggest that a network of residues evolved to regulate nsp12 and nsp14 activity or stability in the ExoN(-) background.

Fixed mutations in nsp12 and nsp14 in MHV-ExoN(-) P250 directly correlate with increased resistance to multiple nucleoside analogues. To determine approximately when the mutations in nsp12 and nsp14 arose, we performed di-deoxy sequencing across these protein-coding regions roughly every 20 passages (P10, 31, 50, 72, 90, 100, 120, 140, 160, 180, 200, 220, 240). By this method, we detected consensus NS mutations at P10, P50, and P160 for nsp12, and at P50 and P160 for nsp14 (Figure 17). Both nsp12 and nsp14 carried their full complement of P250 consensus mutations by P160, except for a minority variant (D913E) in nsp12 maintained at <50% of the population between P200 and P250. These passage levels correlated with increased replication of MHV-ExoN(-) (Figure 11B) and with decreasing sensitivity to 5-FU (Figure 15C). Neither replication nor 5-FU sensitivity of MHV-ExoN(-) changed substantially between P160 and P250. To determine whether MHV-ExoN(-) evolved increased resistance to multiple nucleoside analogues, we treated virus-infected cells with three additional analogues that are substrates for viral RdRps: ribavirin (RBV), a guanine analogue that inhibits viral replication through multiple mechanisms, including mutagenesis and inhibition of purine biosynthesis (Crotty et al., 2002); 5-azacytidine (AZC), an RNA mutagen (Pathak and Temin, 1992); and 2'-C-methyladenosine (CMeA), which is proposed to incorporate in viral RNA and terminate nascent transcripts (Carroll et al., 2003). As with 5-FU, we observed dose-dependent sensitivity to RBV, AZC, and CMeA in all MHV-ExoN(-) viruses that decreased with increasing passage number (Figure 15D-F). Except against AZC, MHV-ExoN(-) sensitivity did not change between P160 and P250. Together, these data demonstrate that MHV-ExoN(-) evolved increased resistance to multiple nucleoside analogues that correlated with the length of

passage and the acquisition of mutations in nsp12 and nsp14. Importantly, this occurred in the absence of specific mutagenic selection and without reversion of ExoN motif I. This increased general selectivity towards all four classes of nucleotide strongly supports an overall increase in fidelity in MHV-ExoN(-) P250.

Mutations in nsp12 partially account for increased resistance of MHV-ExoN(-) P250 to multiple nucleoside analogues. We hypothesized that mutations in MHV-ExoN(-) P250 nsp12 and nsp14 were the most likely to impact replication and nucleoside analogue sensitivity based on their enzymatic activities and temporal association with phenotypic changes. To test this hypothesis, we engineered recombinant MHV-ExoN(-) to encode the P250 nsp12 and nsp14 sequences, alone and together. Expression of nsp12-P250 and nsp14-P250, alone or in combination, altered replication kinetics of MHV-ExoN(-) without affecting peak titers (Figure 18A) and increased gRNA levels above those of MHV-ExoN(-) P3 (Figure 18B). Nsp12-P250 had a greater effect than nsp14-P250 on the sensitivity of MHV-ExoN(-) to all analogues tested, and the combination of nsp12- and nsp14-P250 did not increase resistance above nsp12-P250 alone (Figure 18C-E). None of the recombinant viruses recapitulated the resistance phenotypes of the MHV-ExoN(-) P250 population. Together, these data demonstrate that nsp12-P250 mutations only partially account for the nucleoside analogue resistance of MHV-ExoN(-) P250, and that adaptations in nsp12-P250 mask those in nsp14-P250. We also can conclude that the nsp14-P250 D272E active site mutation does not correct the defect caused by the motif I DE→AA substitutions.

Nsp12-P250 is a high-fidelity polymerase. Mutations in nsp12-P250 increased resistance of MHV-ExoN(-) to multiple nucleoside analogues (Figure 18). Although resistance to a single

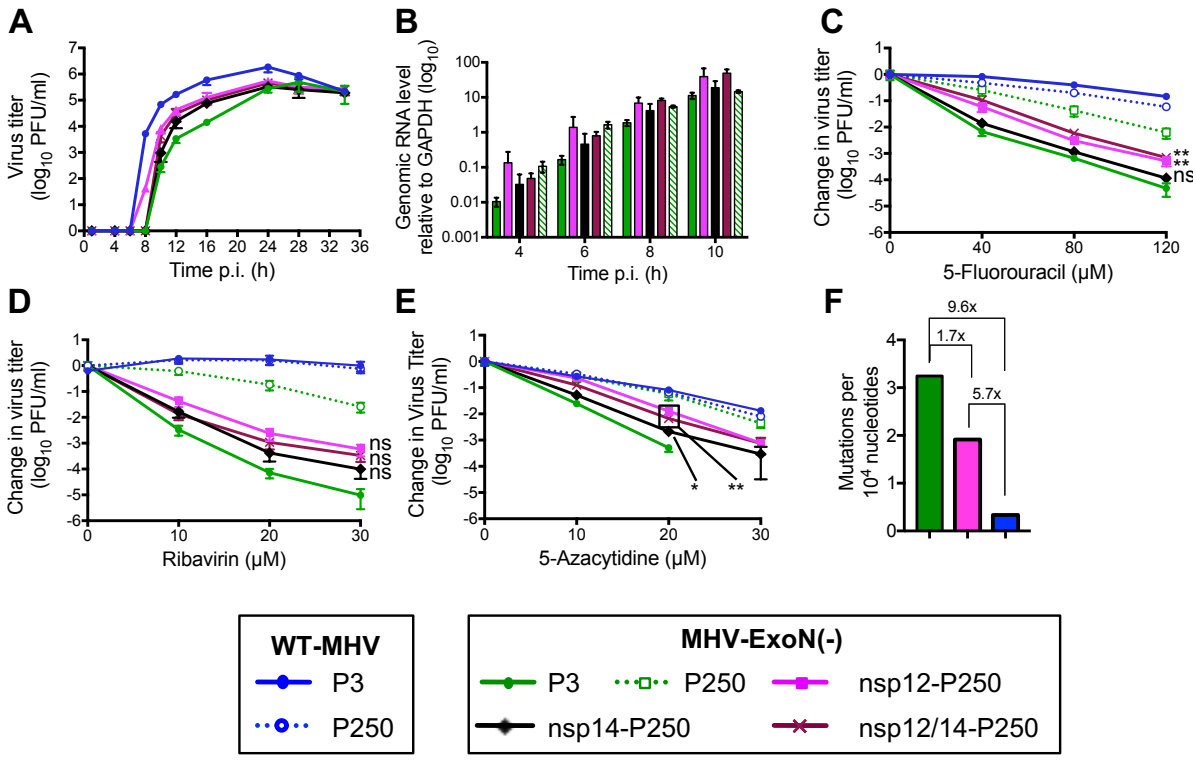


Figure 18. Mutations in nsp12-RdRp and nsp14-ExoN from MHV-ExoN(-) P250 incompletely increase resistance to nucleoside analogues through high-fidelity replication. (*A*) Replication kinetics of recombinant P250 viruses (MOI = 0.01 PFU/cell) plotted as mean and standard deviation with n = 3. (*B*) Genomic RNA accumulation relative to intracellular GAPDH, as described in Fig. 4. Data represent mean and standard error for n = 6-9 (2-3 triplicate experiments). (*C-E*) Sensitivity of recombinant MHV-ExoN(-) viruses to 5-FU (*C*), ribavirin (*D*), and 5-azacytidine (*E*) at MOI = 0.01 PFU/cell, as described in Fig. 4. Data represent mean and standard error of n = 6. (*F*) Mutation frequencies were measured using PrimerID. For panels C-E, statistical significance for change in titer of swapped viruses relative to MHV-ExoN(-) P3 at highest drug concentration tolerated was determined using Mann-Whitney test (*P<0.05, **<0.01, ns = not significant). Boxed points have the same P value.

nucleoside analogue can evolve without increasing overall fidelity, resistance to multiple nucleoside analogues strongly suggests a broadly increased capacity to discriminate nucleotides, or high-fidelity replication (Arias et al., 2008; Sierra et al., 2007; Smith et al., 2015; Zeng et al., 2013). To determine whether nsp12-P250 is a bona fide high-fidelity RdRp, I measured mutation frequencies of WT-MHV, MHV-ExoN(-) P3, and MHV-ExoN(-) nsp12-P250 by PrimerID (Chapter 5). In this strategy, individual genomic RNA templates are tagged with unique barcode sequences (PrimerIDs) during reverse transcription. During analysis, these barcodes are used to generate consensus sequences for individual cDNA species, correcting for PCR and base-calling artifacts commonly introduced during deep sequencing experiments (Zhou et al., 2015). We isolated supernatant genomic RNA from each virus after approximately two rounds of replication and generated libraries with duplicate reactions containing 50,000 genomic templates. We obtained >2 million raw reads that condensed to >40,000 consensus sequences representing 40.8-43.1% of input templates. The mutation frequencies for WT-MHV and MHV-ExoN(-) P3 were 3.38×10^{-5} and 3.24×10^{-4} mutations per nucleotide sequenced, respectively, a difference of 9.6-fold and consistent with previous findings (Eckerle et al., 2007; Sexton et al., 2016). The mutation frequency of ExoN(-) nsp12-250 was 1.92×10^{-4}, which is 1.7-fold lower than MHV-ExoN(-) P3 and 5.7-fold higher than WT-MHV, demonstrating that nsp12-250 is a high-fidelity polymerase (Figure 18F).

Resistance to nucleoside analogues correlates with MHV-ExoN(-) fitness. I hypothesized that mutations in nsp12 and nsp14 provided a fitness advantage to MHV-ExoN(-) P250. I competed the recombinant viruses with a reference MHV-ExoN(-) virus (P1 stock) containing 10 silent mutations in the nsp2 coding region. I detected mutant and reference viruses from the mixed infection by real-time quantitative PCR using dual-labeled probes specific for each virus. MHV-

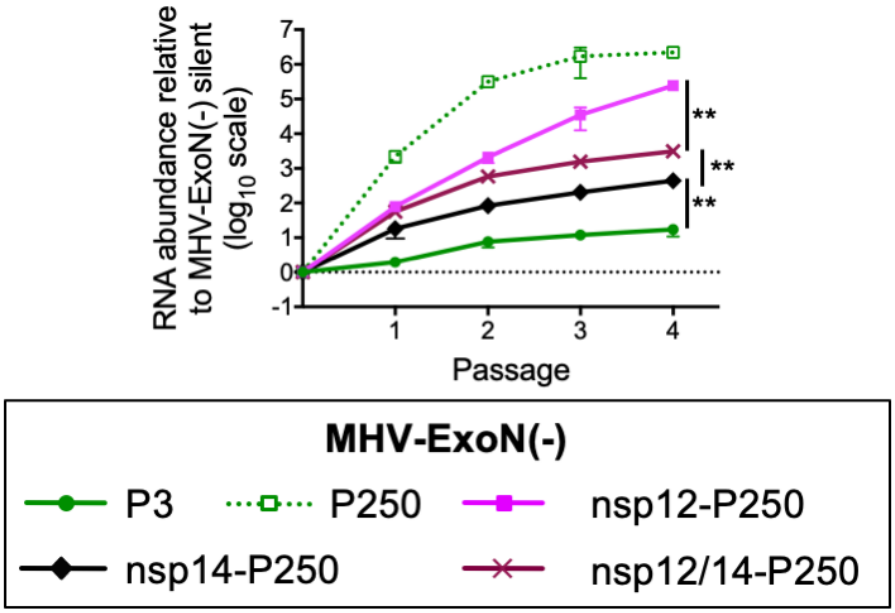


Figure 19. Mutations in nsp12-RdRp and nsp14-ExoN from MHV-ExoN(-) P250 increase the fitness of MHV-ExoN(-). Recombinant viruses were competed against a reference MHV-ExoN(-) containing 10 silent mutations within nsp2. The ratio of viral genome copies relative to the MHV-ExoN(-) reference is plotted. Data represent mean and standard error of n = 6. MHV-ExoN(-) P250 data set contains 4 replicates at passage 3 and a single replicate at passage 4 due to undetectable levels of MHV-ExoN(-) silent. Statistical significance for the indicated comparisons was determined using Mann-Whitney test (**P<0.01).

ExoN(-) P3 showed a modest fitness advantage over the reference P1 MHV-ExoN(-) silent (Figure 19, solid green). MHV-ExoN(-) P250 profoundly outcompeted MHV-ExoN(-) silent, with >1000-fold more MHV-ExoN(-) P250 genomes present at the end of passage 1 (Figure 19, dotted green). MHV-ExoN(-) nsp12-P250 had greater relative fitness than MHV-ExoN(-) nsp14-P250, and MHV-ExoN(-) nsp12/14-P250 was intermediate between the single recombinants, implicating a complex evolutionary interaction between these two proteins. The measured fitness correlated with the patterns of nucleoside analogue resistance and RNA synthesis associated with mutations in nsp12 and nsp14, suggesting a link between the evolution of these phenotypes. The result also confirms that nsp12 and nsp14 are important but not sufficient to account for the significantly increased fitness of MHV-ExoN(-) P250 relative to MHV-ExoN(-) P3.

Discussion

This chapter describes the experimental adaptive evolution of WT-MHV and MHV-ExoN(-) during long-term passage in cell culture. WT-MHV evolved increased replication kinetics over 250 passages, with few consensus mutations arising in the WT-MHV P250 genome. In contrast, MHV-ExoN(-) accumulated 8-fold more mutations than WT-MHV, none of which occurred at the ExoN-inactivating substitutions (Figure 13). Nevertheless, MHV-ExoN(-) P250 demonstrated increased replication kinetics and fitness, as compared to MHV-ExoN(-) P3 (Figure 11 and Figure 19). Our previous studies demonstrate that ExoN-mediated proofreading is required for CoV fitness (Graham et al., 2012; Sexton et al., 2016; Smith et al., 2013). Thus, MHV-ExoN(-) was likely under selective pressure for restoration of high-fidelity replication or for tolerance of the increased mutational load. Consistent with this hypothesis, MHV-ExoN(-) P250 exhibited increased resistance to multiple nucleoside analogues, a phenotype strongly

associated with high-fidelity viruses (Arias et al., 2008; Sierra et al., 2007; Smith et al., 2015; Zeng et al., 2013). Our results raise two important questions. In the face of selection for increased fidelity, why didn't MHV-ExoN(-) revert? Can MHV replicase proteins mediate high-fidelity replication without ExoN proofreading?

In the face of selective pressure for increased fidelity, why didn't MHV-ExoN(-) revert? Although our data suggest that MHV-ExoN(-) was under selective pressure for increased fidelity, we detected no primary reversion at the DE→AA substitutions in MHV-ExoN(-) at any passage tested. These data are consistent with and significantly extend previous studies reporting genotypic stability of ExoN(-) motif I in MHV and SARS-CoV (Eckerle et al., 2007; 2010; Graham et al., 2012; Sexton et al., 2016; Smith et al., 2013). Complete reversion within ExoN(-) motif I to DE would require four nucleotide changes. This likely represents a high genetic barrier to reversion, especially given that fitness can be increased by mutations outside of nsp14-ExoN (Figure 19) (Eckerle et al., 2010). Single and double nucleotide changes within motif I could restore acidic charge to individual residues (e.g. motif I EA, AD, ED, etc). However, the active site compositions of DEDDh exonucleases, such as the Klenow fragment, are so stringent that even conservative mutations (D-to-E or E-to-D) reduce ExoN activity by >96% (Derbyshire et al., 1991). Thus, intermediate amino acid changes may not have a selective advantage compared to motif I AA, limiting the evolutionary pathways to reversion. However, nsp14-P250 had detectable effects on RBV and AZC resistance as well as the competitive fitness of MHV-ExoN(-) (Figure 18 and Figure 19), demonstrating a modest capacity for fitness adaptation in nsp14 outside of the catalytic residues. Whether these mutations resulted from genetic drift or positive selection remains unclear. Nevertheless, our data show that MHV-ExoN(-) can adapt for increased fitness without fully restoring exoribonuclease activity. While some mutations in

MHV-ExoN(-) P250 likely confer DBT cell-specific selective advantages, others may represent generalizable strategies for overcoming ExoN(-) defects in other cell types and in other coronaviruses. Thus, understanding the mechanisms by which MHV-ExoN(-) P250 compensated for ExoN activity could allow recovery of ExoN(-) variants of other CoVs, such as transmissible gastroenteritis virus and human CoV 229E, which to date have been nonviable as ExoN(-) recombinants (Becares et al., 2016; Minskaia et al., 2006).

Can MHV replicase proteins mediate high-fidelity replication without ExoN proofreading?

MHV-ExoN(-) P250 exhibits increased resistance to four nucleoside analogues after passage (Figure 15). Although resistance to a single nucleoside analogue can evolve without increasing overall fidelity, resistance to multiple nucleoside analogues strongly suggests a broadly increased capacity to discriminate nucleotides, or high-fidelity replication (Arias et al., 2008; Sierra et al., 2007; Smith et al., 2015; Zeng et al., 2013). Increased-fidelity variants in RNA viruses have most frequently been mapped to RdRps (Cameron et al., 2016; Coffey et al., 2011; Pfeiffer and Kirkegaard, 2003; Vignuzzi et al., 2008). Thus, if increased fidelity contributes to nucleoside analogue resistance in MHV-ExoN(-) P250, the most likely protein involved would be nsp12-P250. Four findings are consistent with the hypothesis that mutations within nsp12-P250 increase RdRp fidelity. First, nonsynonymous mutations to nsp12 arose in the low-fidelity MHV-ExoN(-) but not in the presence of proofreading (WT-MHV). Second, five of the mutations lie in or near structural motifs important for fidelity regulation in other RdRps. Amino acid substitutions in the fingers and palm domains have been repeatedly shown to affect viral RdRp fidelity (Campagnola et al., 2015; Pfeiffer and Kirkegaard, 2003), and we have recently reported a fingers mutation (nsp12-V553I) that likely increases the fidelity of the MHV RdRp (Sexton et al., 2016). Our modeled structure predicts that nsp12-P250 contains three mutations in the palm

domain and one in the fingers domain, with the M814K thumb domain mutation lying near the palm (Figure 17A). Third, exchange of nsp12-P250 alone into the background of MHV-ExoN(-) reduced the susceptibility of MHV-ExoN(-) to three different nucleoside analogues (Figure 18). Finally, nsp12-P250 reduced the mutation frequency of MHV-ExoN(-) P3 by 1.7-fold, respectively, consistent with high-fidelity RdRps in other viruses (Figure 18) (Smith et al., 2014). Thus, all data support the hypothesis that nsp12-P250 is a high-fidelity RdRp.

Importantly, nsp12-P250 only partially accounts for the MHV-ExoN(-) P250 nucleoside analogue resistance phenotype (Figure 18), suggesting a possible limit to the compensation achievable by mutating the RdRp alone. Further, the effects of mutations in nsp12-P250 and nsp14-P250 are not additive and may be antagonistic when isolated from the whole passaged virus (Figure 18), indicating that the relationships between nsp12- and nsp14-P250 mutations are likely evolutionarily linked with those in other MHV proteins. In fact, a substantial component of the evolved resistance to nucleoside analogues cannot be explained by nsp12-P250 and nsp14-P250, alone or together. In support of this hypothesis, we identified several nonsynonymous mutations in other replicase proteins, such as nsps 8, 9, 13, and 15. SARS-CoV nsp8 and nsp13 have functional interactions with nsp12, acting as a primase/processivity factor (Imbert et al., 2006; Subissi et al., 2014) and a helicase/NTPase, respectively (Adedeji et al., 2012). Processivity factors in herpes simplex virus and *Mycobacterium tuberculosis* regulate DNA polymerase fidelity by balancing polymerase extension and exonuclease activity (Chaudhuri et al., 2003; Gu et al., 2016), and helicases in chikungunya virus and foot-and-mouth disease virus can evolve to increase fidelity (Stapleford et al., 2015) and alter the frequency of ribavirin-induced mutations (Agudo et al., 2016), respectively. SARS-CoV nsp9 has RNA-binding activities and is proposed to participate in the multi-protein replicase complex (Egloff et al.,

2004; Smith et al., 2014), and MHV nsp15 is a uridylate-specific endoribonuclease (Bhardwaj et al., 2004). Both could plausibly be involved in modulating polymerase activity. Additionally, it remains possible that evolution for increased fidelity could involve proteins outside the canonical replication complex (nsps7-16), including those in the structural and accessory cassette. Thus, while immediate studies will focus on testing whether replicase proteins nsp8, 9, 13, and 15 regulate fidelity, it is exciting to consider the possibility that this virus-directed discovery approach will reveal novel interactions between multiple MHV proteins.

Conclusions. The proofreading activity of the nsp14 exoribonuclease is a critical determinant of CoV replication, fidelity, and fitness. I show that CoVs also have the capacity to compensate for loss of ExoN activity through a network of mutations in nsp12, nsp14, and elsewhere in the genome. Thus, while nsp14-ExoN appears to play a dominant role in CoV replication fidelity, its activity is likely closely tied to a highly evolved network of proteins. The demonstrated co-adaptation for replication, competitive fitness, and likely increased fidelity within MHV-ExoN(-) supports the hypothesis that these roles are linked functionally and evolutionarily. It will be interesting to test whether evolution in other cell types derived from different species or with different innate immune environments would result in similar adaptive strategies.

CHAPTER 3: IDENTIFICATION OF FIDELITY-ALTERING MUTATIONS OUTSIDE OF THE MHV POLYMERASE AND EXORIBONUCLEASE

Introduction

Prior to this dissertation research, the proofreading exoribonuclease (ExoN) had been identified as a critical determinant of viral fitness (Eckerle et al., 2007; 2010; Graham et al., 2012; Smith et al., 2013; 2015). In the previous chapter, I demonstrated that a murine hepatitis virus lacking ExoN activity [MHV-ExoN(-), hereafter ExoN(-)] can adapt during long-term passage for increased replication and fitness without reverting the ExoN-inactivating mutations. Passage-adapted ExoN(-) mutants also demonstrated increasing resistance to nucleoside analogues that is explained only partially by a high-fidelity RNA-dependent RNA polymerase (nsp12-250). These data suggested that high-fidelity coronaviral replication is mediated by the interplay of multiple replicase proteins and supported the proposed link between coronavirus fidelity and fitness.

In this chapter, I describe my efforts to identify proteins outside of the nsp12-RdRp that contributed to the high-fidelity phenotype of ExoN(-) P250. As with experiments in Chapter 2, the general approach was to engineer recombinant ExoN(-) viruses containing passage-selected sequences of various proteins, alone and in combination, and screen for high-fidelity replication using nucleoside analogues. Using a large panel of recombinant viruses, I show that the high-fidelity nsp12-RdRp is the only replicase protein that adapted to increase nucleoside analogue resistance in ExoN(-) P250. My data also suggests that structural and accessory protein mutations may contribute to nucleoside analogue resistance but indicates that the full phenotype of ExoN(-) P250 depends upon mutations across the genome that may be inseparable.

Co-author contributions

I performed all molecular biology and all experiments described in this chapter with few exceptions. Xiaotao Lu was instrumental in cloning recombinant MHV fragments and assisted in virus recovery. Thayer Taft performed multiple replicates of 5-fluorouracil sensitivity assays reported in Figure 27.

Results

Selection of viruses for study. The experiments described in Chapter 2 focused on coding sequences from ExoN(-) P250. However, I chose to use coding sequences from ExoN(-) P160 in subsequent experiments for several reasons. First, ExoN(-) P160 and P250 have nearly indistinguishable replication and nucleoside analogue resistance (Figure 11, Figure 15, and reproduced in Figure 20). Genomic RNA accumulation was also similar for both viruses (Figure 20C). Second, the nsp12 and nsp14 sequences of ExoN(-) P160 and P250 are identical (Figure 17 and Appendices A.2, A.3). Third, the P160 consensus sequence has 20 fewer nonsynonymous mutations affecting four fewer proteins than P250, simplifying the matrix of testable interactions (Table 3 and Appenfix A.3). Finally, ExoN(-) P250 is challenging to use experimentally because it does not form distinct syncytia, while ExoN(-) P160 induces visible cytopathology.

The infectious clone of ExoN(-) P160 cannot be recovered. Viral populations are complex, and the consensus sequence is not necessarily reflective of the population structure. I intended to use ExoN(-) P160 as a control in all experiments in this chapter, but I did not want minority variants in the passaged population to cloud my results. To circumvent this potential problem, I attempted to generate a full-length infectious clone of ExoN(-) P160 based on the consensus

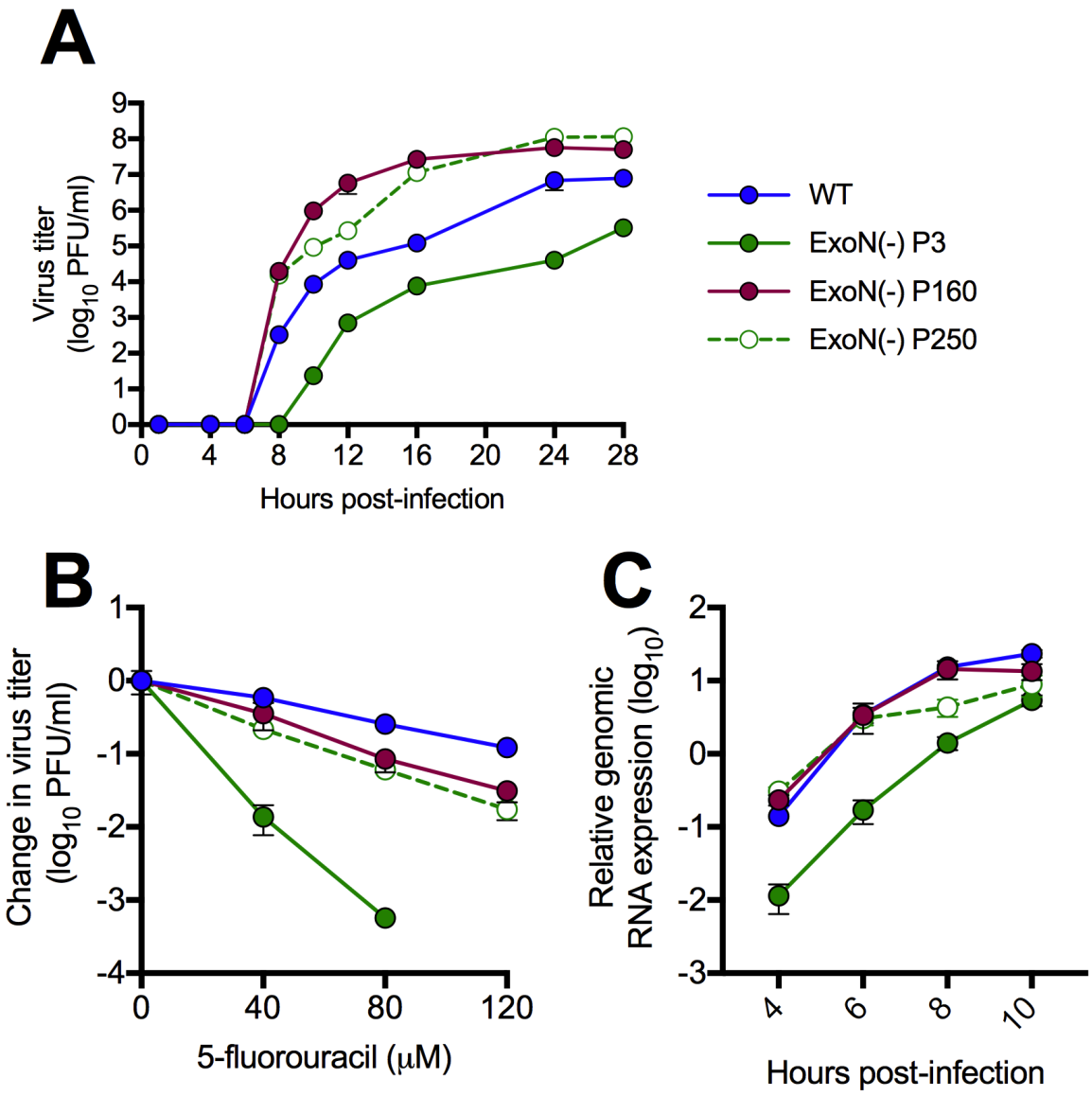


Figure 20. Phenotypes of ExoN(-) P160 and ExoN(-) P250 are nearly identical. ExoN(-) P160 and ExoN(-) P250 have nearly identical replication kinetics (*A*), sensitivity to 5-fluorouracil (*B*), and accumulation of genomic RNA (*C*). Note: Data in (*A*) and (*B*) are the same as presented in Figure 11 and Figure 15. Data in (*C*) represent mean and standard deviation of n = 3. MOI was 0.01 PFU/cell in panels (*A*) and (*B*) and 1 PFUcell in panel (*C*).

Table 3. Comparison of ExoN(-) P160 and ExoN(-) P250 mutations.

ExoN(-) passage level	P160	P250
Total # mutations	125	171
Nonsynonymous	54	74
Synonymous	61	90
Noncoding	7	7
# proteins with nonsynonymous mutations	15	19

sequence (Appendix A.3). Xiaotao Lu used ExoN(-) P160 RNA to clone fragments C, D, and E from the MHV reverse genetics system, and I ordered fragments A, B, F, and G from a commercial supplier. ExoN(-) P160 sequence has a deletion of nucleotides 22286-23646. Not knowing the importance of this deletion for viral fitness, I generated two versions of the fragments. One set generates the native ExoN(-) P160 sequence, lacking nucleotides 22285-23646 (ExoN(-) P160Δ in Figure 21). In the other, I inserted wild-type nucleotides 22285-23646 (ExoN(-) P160ic in Figure 21). All fragments were confirmed by sequence analysis and restriction digests. Due to the high replicative capacity of ExoN(-) P160 (Figure 20), I expected that these viruses would be readily recoverable. However, I could not recover either ExoN(-) P160ic or ExoN(-) P160Δ despite six and three attempts, respectively, including protocol modifications intended for recovery of highly debilitated viruses. I did observe small foci of cytopathic effect (punctate syncytia) within 18-24 hours after plating, indicating that spike protein subgenomic mRNAs are transcribed and translated, but no viral particles were released. In general, the rate and yield of virus recovery correlates with their replicative fitness, and in my hands, even debilitated viruses (e.g. ExoN(-)-like) recover in <3 attempts. Thus, the inability to recover these viruses is more likely due to an intrinsic property of ExoN(-) P160 than to technical errors.

Passaged ExoN(-) viruses are marginally restricted in BHK-R cells. Viral fitness is dependent upon the specific selective environment. Accordingly, mutations may be beneficial in one environment but maladaptive in others, and tissue culture-specific adaptations occur frequently in experimental evolution studies with viruses, including MHV (Baric et al., 1997; Iketani et al., 2018). I hypothesized that some ExoN(-) P160 mutations were DBT-specific and thus restricted ExoN(-) P160 replication in our recovery cell line, baby hamster kidney cells

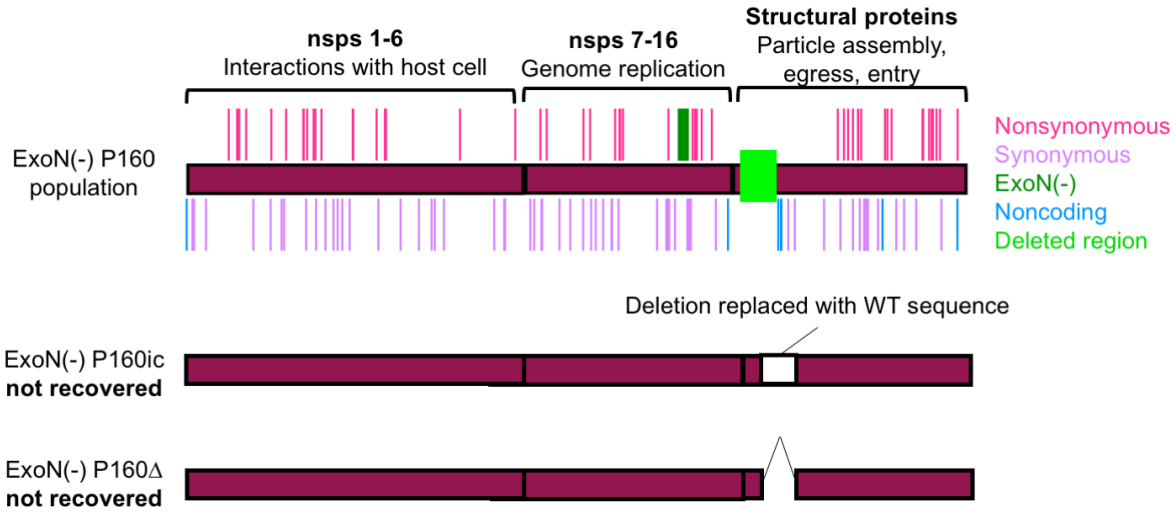


Figure 21. Schematic of attempted recombinant ExoN(-) P160 infectious clones. Linear schematic of ExoN(-) P160 population with mutations marked (top). ExoN(-) P160 sequence is denoted in maroon, and wild-type sequence is in white.

(BHK-R). To test this hypothesis, I measured replication of ExoN(-) P3, P160, and P250 in both DBT9 and BHK-R cells. As expected given their low passage number, WT P3 and ExoN(-) P3 titers at 24 hours post-infection were identical in both cell lines (Figure 22). In contrast, ExoN(-) P160 and ExoN(-) P250 titers were reduced by 0.5 $\log_{10}$ and 1 $\log_{10}$, respectively, indicating that some DBT9-specific adaptation occurred. However, both passaged ExoN(-) viruses replicated to more than 10^7 PFU/mL in both cell lines, well above that achieved by ExoN(-) P3, suggesting that cell-type-specific adaptations do not restrict ExoN(-) P160 infectious clone recovery. Consistent with these results, co-plating of electroporated BHK-R cells with DBT9s did not yield infectious virus (not shown). Given these challenges, I used the ExoN(-) P160 passaged population as a control in all subsequent experiments.

Replicase mutations outside of nsp12 and nsp14 do not affect nucleoside analogue sensitivity. In Chapter 2, I determined that mutations in nsp12-250 and nsp14-250 were likely evolutionarily linked with mutations in other MHV proteins. Intuitively, the most likely fidelity-regulating proteins would be the remaining components of the proposed replicase machinery: nsps 7, 8, 9, 10, 13, 15, and 16 (Smith et al., 2014). Of these, only nsp8, nsp9, and nsp13 had nonsynonymous mutations in ExoN(-) P160. In SARS-CoV, nsp8 acts as a primase and processivity factor (Imbert et al., 2006; Subissi et al., 2014), nsp13 is a helicase (Adedeji et al., 2012), and nsp9 binds to single-stranded RNA (Egloff et al., 2004). Processivity factors in herpes simplex virus and *Mycobacterium tuberculosis* regulate DNA polymerase fidelity by balancing polymerase extension and exonuclease activity (Chaudhuri et al., 2003; Gu et al., 2016), and helicases in chikungunya virus and foot-and-mouth disease virus can influence fidelity and susceptibility to nucleoside analogues (Agudo et al., 2016; Stapleford et al., 2015). Based on their known biochemical activities and association with the replicase-transcriptase, I selected

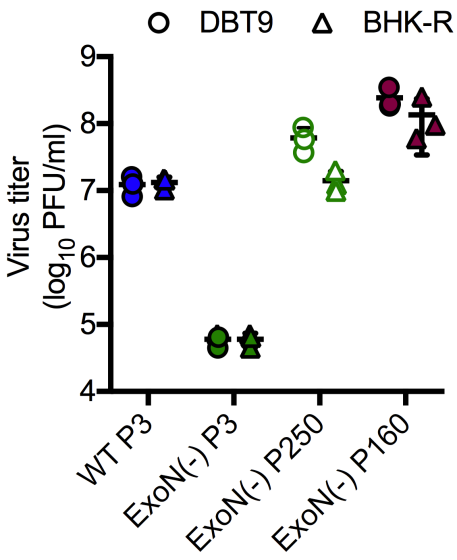


Figure 22. Viruses passaged in DBT9 cells are marginally restricted in BHK-R cells.
Subconfluent monolayers of DBT9 cells or BHK-R cells were infected with the indicated viruses at MOI = 0.01 PFU/cell. Supernatants were harvested and titered after 24 hours.

these three mutants for initial testing. I cloned and recovered recombinant ExoN(-) viruses containing each of these three mutants individually (Figure 23). None of these variants had significant effects on the sensitivity of ExoN(-) to 5-fluorouracil, ribavirin, 5-azacytidine, or 2′-*C*-methyladenosine (Figure 24), indicating that the mutations in nsp8, 9, and 13 had no independent influence on replication fidelity. However, I suspected that they might be epistatically linked with nsp12-250 and nsp14-250 (that is, the effects of mutations in nsp8, 9, and 13 might depend upon interactions with nsp12-250 and nsp14-250). I cloned and recovered each mutation separately with nsp12- and nsp14-250, as well as a recombinant virus containing all five P160 replicase protein sequences (Figure 23 and Figure 26). However, no combination of replicase mutations increased the 5-fluorouracil sensitivity beyond that of nsp12-250 alone (Figure 25), indicating that the high-fidelity nsp12-RdRp is the sole ExoN(-) P160 replicase protein that evolved to resist nucleoside analogues. Interestingly, and in contrast to experiments in Chapter 2 (Figure 18), nsp12-250 was sufficient for the resistance phenotype of ExoN(-) P160 to 5-azacytidine and 2′-*C*-methyladenosine (Figure 24).

Mutations in the structural and accessory proteins may affect 5-fluorouracil sensitivity. Although I considered nsp8, 9, and 13 to be the most likely candidates for fidelity regulation in ExoN(-) P160, proteins outside of the canonical replication complex could be important fidelity regulators (Collins et al., 2018). To identify potentially contributory mutations outside of the predicted replicase, I designed viruses to measure the collective effects of mutations in two additional genome regions: nsps1-6, which have various functions including interferon antagonism, polyprotein cleavage, and host membrane modification [ExoN(-) nsp1-6-160]; and the structural and accessory proteins, which primarily mediate virus assembly and dissemination

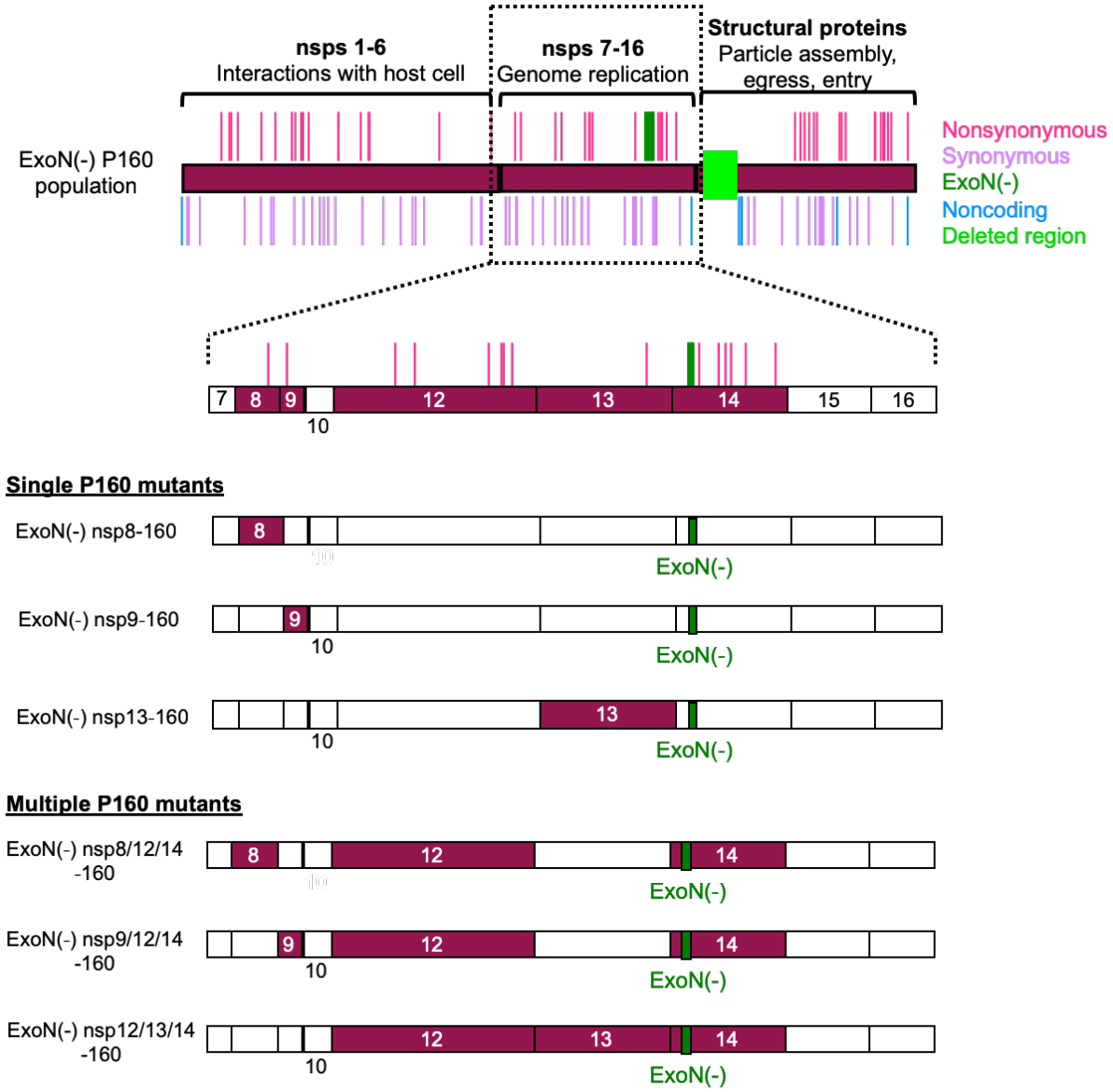


Figure 23. Schematic of P160 replicase mutants. Linear schematic of ExoN(-) P160 population with mutations marked (top). The region containing components of the RTC is marked and expanded. Proteins containing nonsynonymous mutations from ExoN(-) P160 are filled in maroon and those without are white.

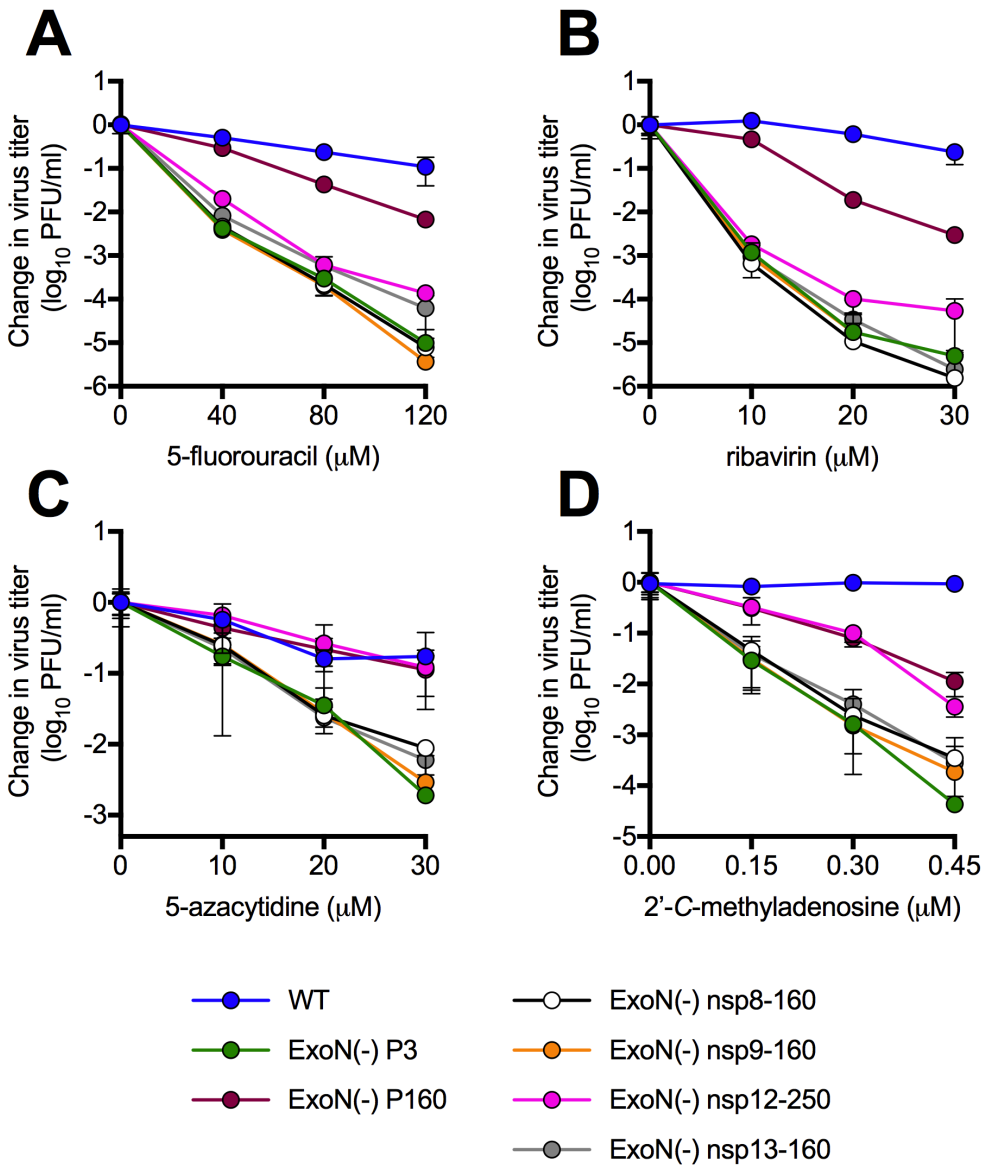


Figure 24. P160 mutations in nsp8, nsp9, and nsp13 do not affect sensitivity to nucleoside analogues. Cells were incubated with 5-fluorouracil (*A*), ribavirin (*B*), 5-azacytidine (*C*), or 2'-*C*-methyladenosine (*D*) for 30 minutes prior to infection with the indicated virus at 0.01 PFU/cell. Supernatants were harvested and titered by plaque assay at 24 hours post-infection. Data represent mean and standard deviation of n = 6. Statistical significance of individual replicase mutant relative to ExoN(-) was calculated using two-way ANOVA with Dunnett's multiple comparisons testing. No comparisons were statistically significant.

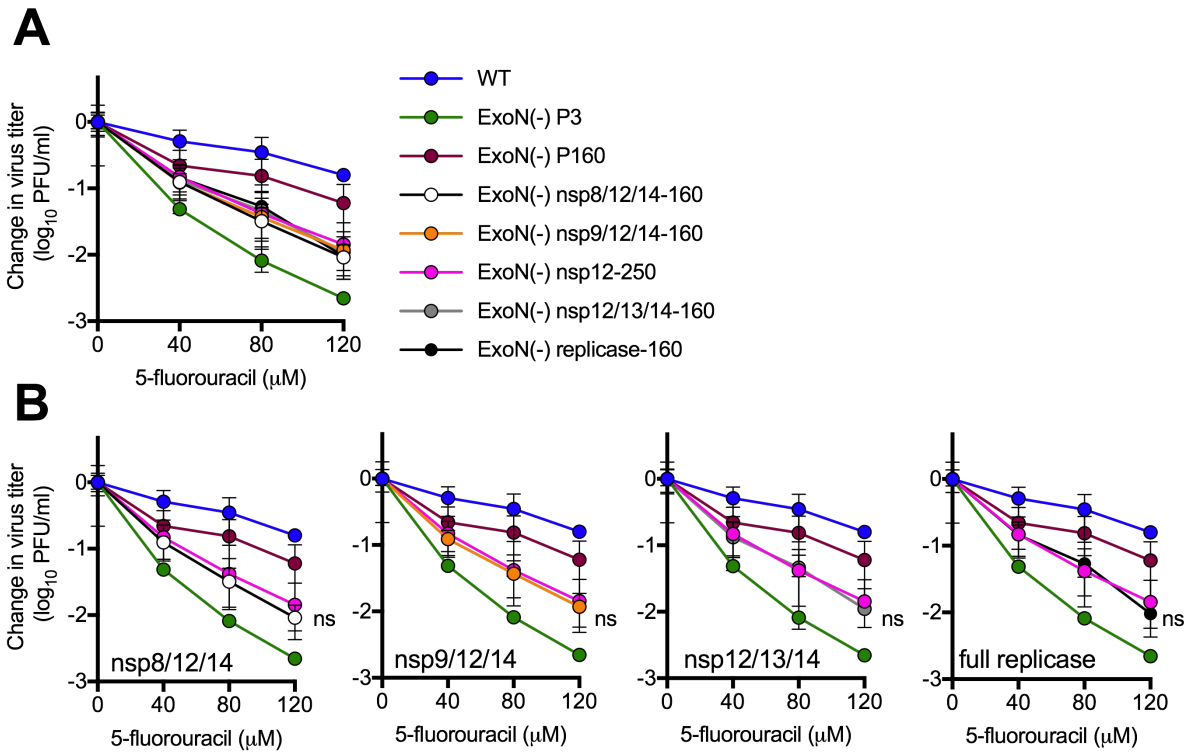


Figure 25. Nsp12-250 is the primary determinant of 5-fluorouracil sensitivity in the P160-passaged replicase. Cells were incubated with 5-fluorouracil for 30 minutes prior to infection with the indicated viruses at MOI = 0.01 PFU/cell. Supernatants were harvested and titered by plaque assay at 24 hours post-infection. Panel (*A*) contains all data. Individual replicase mutant viruses are separated for ease of visualization in panel (*B*). Data for WT, ExoN(-) P3, ExoN(-) P160, ExoN(-) nsp12-250 are identical in all 3 panels of (*B*). Data represent mean and standard deviation of n = 6. Statistical significance of individual replicase mutant relative to ExoN(-) was calculated using two-way ANOVA with Dunnett's multiple comparisons testing. Only marked comparisons were performed. ns = not significant.

[ExoN(-) structural-160] (Figure 26) (Perlman and Netland, 2009). I could not recover any virus containing P160 sequences in nsps1-6, suggesting that this region may have prevented recovery of ExoN(-) P160 infectious clones. However, I did readily recover ExoN(-) structural-160, which appears to increase resistance to 5-fluorouracil relative to ExoN(-) P3 (Figure 27). This difference was not statistically significant.

Plaque clones of ExoN(-) P160 have identical 5-fluorouracil sensitivity to the full population. A foundational principle of viral dynamics is that groups of variants in viral populations cooperate and complement each other. Thus, the behavior of a viral population depends upon the full spectrum of variants found within it (Bordería et al., 2015; Dolan et al., 2018b; Domingo et al., 2012; Vignuzzi et al., 2006). Population structures could influence nucleoside analogue resistance by a number of mechanisms, including absorption of the mutational load, complementation between distinct genomes, or recombination to purge deleterious mutations. To test for large-scale population effects, I plaque purified three clones of ExoN(-) P160 three times (Figure 12). Each clone is derived, theoretically, from a single infectious unit. During stock generation, each clone should develop a population structure that is distinct from the original passaged population. All three clones had identical 5-fluorouracil sensitivity to the passaged population (Figure 28), indicating that either individual genomes or those derived from a single virus plaque encoded the adaptive changes required for resisting nucleoside analogues.

Discussion

RNA viral susceptibility to mutagenic nucleoside analogues is determined by multiple factors, including polymerase fidelity, mutational robustness, and, in the case of coronaviruses,

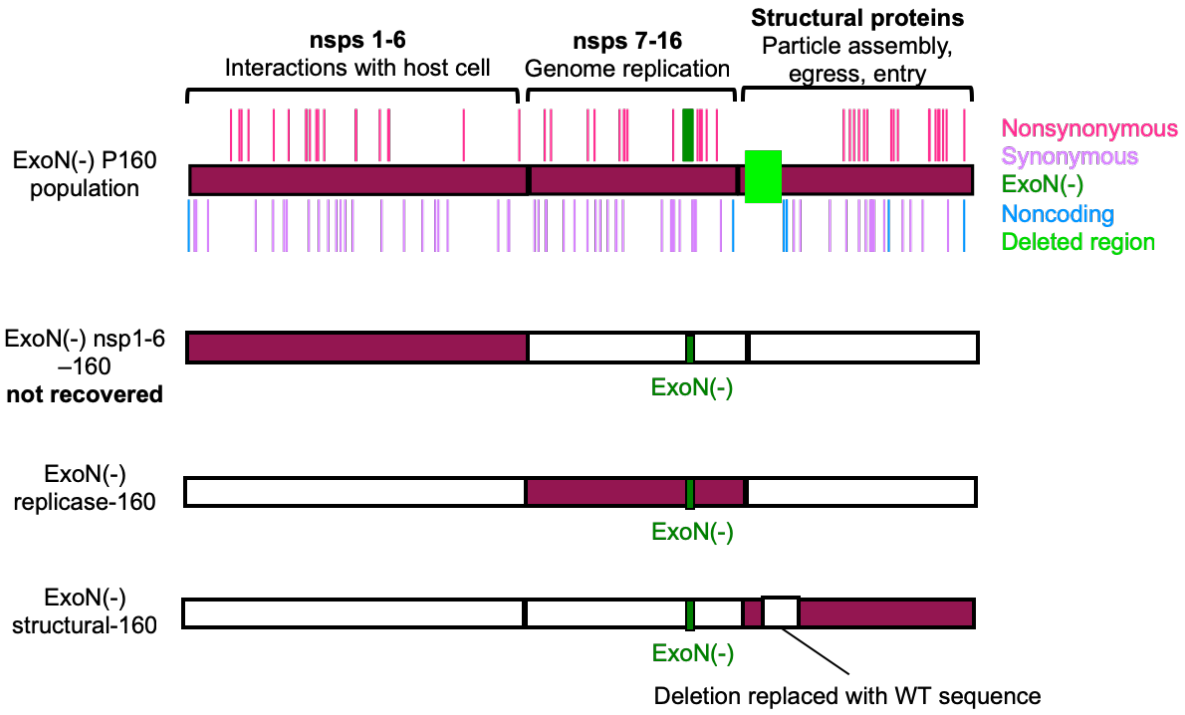


Figure 26. Schematic of P160 regional swaps. Linear schematic of ExoN(-) P160 population with mutations marked (top). Proteins containing nonsynonymous mutations from ExoN(-) P160 are filled in maroon, and those without are white.

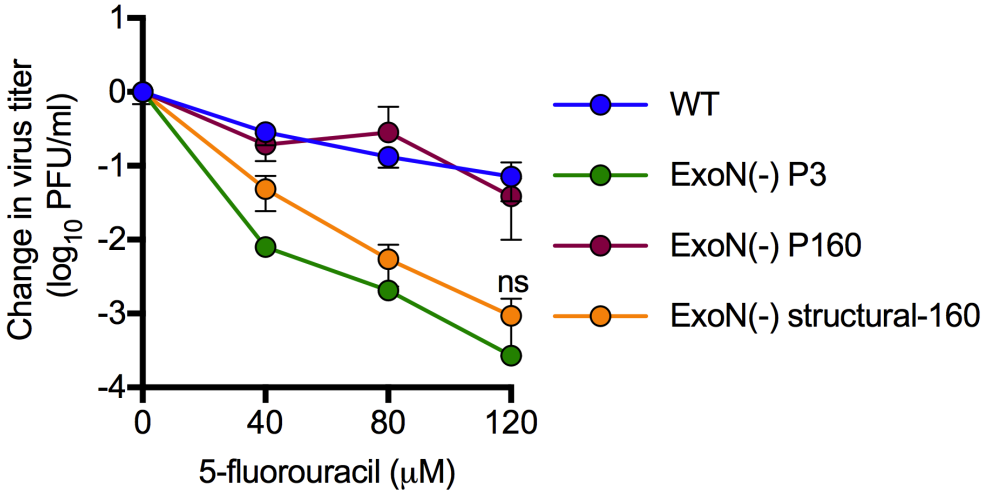


Figure 27. Passage-acquired mutations in structural and accessory proteins may contribute to 5-fluorouracil resistance. Cells were incubated with 5-fluorouracil for 30 minutes prior to infection with the indicated viruses at MOI = 0.01 PFU/cell. Supernatants were harvested and titered by plaque assay at 24 hours post-infection.

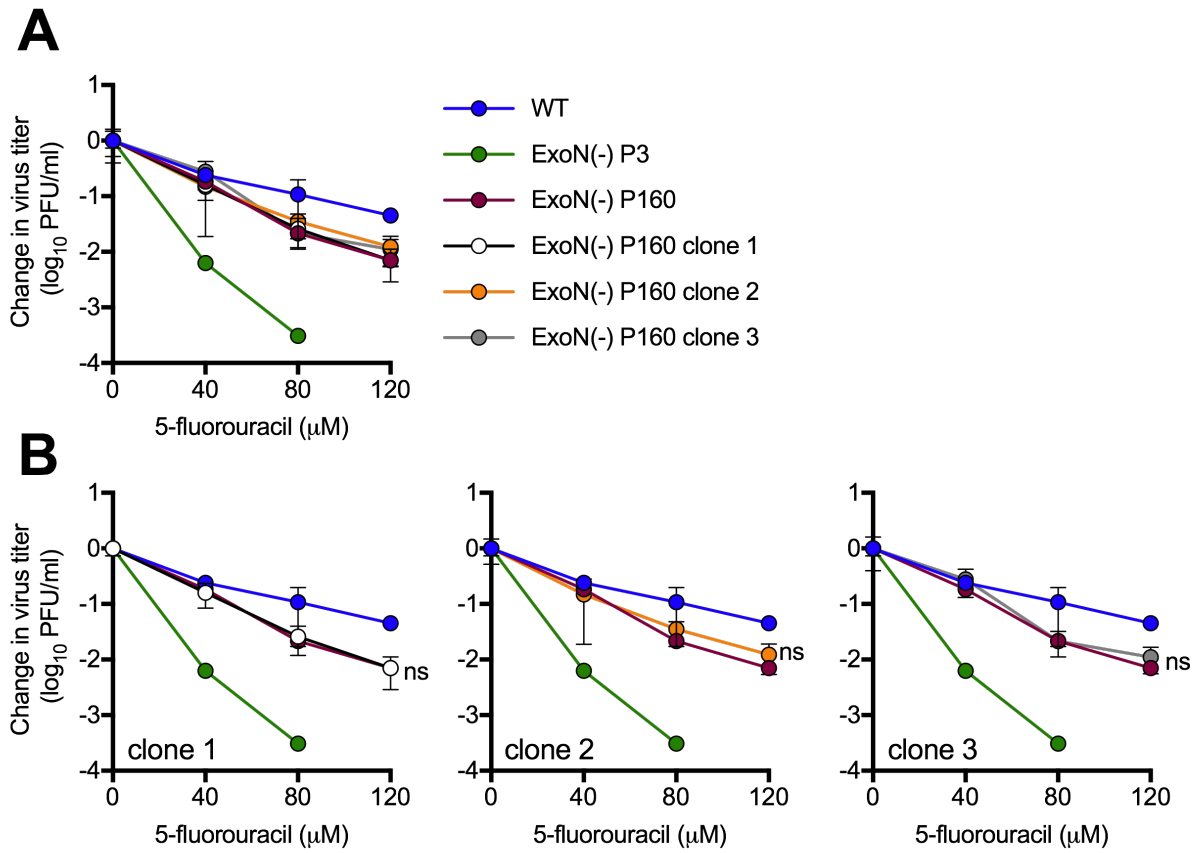


Figure 28. Plaque-purified ExoN(-) P160 clones have identical 5-fluorouracil sensitivity to the full population. Cells were incubated with 5-fluorouracil for 30 minutes prior to infection with the indicated viruses at MOI = 0.01 PFU/cell. Supernatants were harvested and titered by plaque assay at 24 hours post-infection. Panel (*A*) contains all data. Individual replicase mutant viruses are separated for ease of visualization in panel (*B*). Data for WT, ExoN(-) P3, ExoN(-) P160, ExoN(-) nsp12-250 are identical in all 3 panels of (*B*). Data represent mean and standard deviation of n = 3. Statistical significance of individual replicase mutant relative to ExoN(-) was calculated using two-way ANOVA with Dunnett's multiple comparisons testing. Only marked comparisons were performed. ns = not significant.

replicative proofreading. Accordingly, disruption of proofreading in SARS-CoV and MHV leaves them highly susceptible to nucleoside analogues. In Chapter 2, I demonstrated that proofreading-deficient MHV is capable of restoring high-level resistance to nucleoside analogues without primary reversion of ExoN(-). Only part of the resistance phenotype of passaged MHV-ExoN(-) could be accounted for by a high-fidelity polymerase, raising important questions about which other proteins or mechanisms mediate high-level nucleoside analogue resistance. In this chapter, I used a large panel of recombinant MHV containing passage-acquired sequences to examine potential sources of nucleoside analogue resistance in ExoN(-) P160. My results revealed a number of surprising findings. First, despite the high replication capacity of ExoN(-) P160, I could not recover its infectious clone, raising key questions about how it may interact with different cell types. Second, no replicase-transcriptase complex (RTC) proteins outside of nsp12-RdRp had any effect upon nucleoside analogue resistance, but mutations within the structural and accessory proteins increased resistance to 5-fluorouracil, indicating that the ExoN(-) P160 phenotype is complex and probably multifactorial. Finally, the high-level resistance of ExoN(-) P160 to nucleoside analogues did not depend upon a specific population architecture, strongly suggesting that the ExoN(-) P160 nucleoside analogue resistance phenotype likely depends upon a complex coevolution of mutations across the genome that may not be isolable.

Why can't an infectious clone of ExoN(-) P160 be recovered? In general, the replication kinetics of MHV variants correlate well with the ease of their recovery by reverse genetics. Given that ExoN(-) P160 replicates faster and to higher titers than wild-type (Figure 20), I expected that recovery of an infectious clone would be straightforward. However, despite evidence of viral protein synthesis (small puncta of cytopathic effect), I was unable to recover

viable virus from ExoN(-) P160ic or ExoN(-) P160Δ. The reasons for this restriction are unclear. The simplest explanation is that the full-genome sequence of ExoN(-) P160 lacks critical adaptive mutations or contains lethal artifacts. Comparing the full-genome sequences of plaque-purified clones (Figure 28) to that of ExoN(-) P160 could reveal permissive or restrictive mutations. More interestingly, however, is the notion that long-term adaptation to DBT9 cells prevents efficient recovery in alternative cell lines. ExoN(-) P160 replication was not dramatically reduced in BHK-R cells (Figure 22), but mechanisms that subtly restrict experimental infections could be lethal in the setting of virus recovery. In the MHV reverse genetics system, full-length RNA genomes are electroporated into BHK-R cells along with nucleocapsid-encoding mRNA (Yount et al., 2002). In bypassing spike-mediated entry steps, the first event in the life cycle is translation of ORF1a (Figure 2). Thus, the first proteins that interact with the cellular microenvironment are nsps1-6, which establish a permissive environment for RNA synthesis. DBT9-specific adaptations in nsps1-6 could disrupt formation of replication organelles or other undefined structures and processes inside BHK-R cells. Consistent with this hypothesis, I could not recover a recombinant virus containing P160 sequence from nsps1-6 [ExoN(-) nsp1-6-160, Figure 26]. If true, ExoN(-) P160 mutations could be scalpels for fine dissection of the virus-host interactions important for establishing MHV replication.

Why does ExoN(-) P160 tolerate so many nonsynonymous mutations in the RTC? The predicted coronavirus replicase-transcriptase complex (RTC) contains, at minimum, nsps7-14 and may include nsp15 and nsp16 (Figure 10) (Smith et al., 2014). Many of these proteins have multiple domains that perform distinct functions during replication (Snijder et al., 2016). In order to maintain their extensive interactions, RTC proteins must co-evolve in a cooperative manner. Mutations in these proteins are likely under strong purifying selection to preserve both their

essential individual functions and their interactions with the rest of the complex. Put another way, RTC proteins are so important for coronaviral fitness that mutations are unlikely to fixate by chance. Accordingly, phylogenetic analysis of lineage C *Betacoronaviruses* (including MERS-CoV) found evidence of strong negative selection in nsps 7, 8, 9, 12, 13, and 14 of ecological coronaviruses (Forni et al., 2016b). Over long-term passage, only one mutation was found in the replicase of WT-MHV, suggesting that most mutations were not tolerated (Figure 13, Appendix A.1). Thus, it seems likely that the nonsynonymous mutations in the ExoN(-) P160 RTC are important for viral fitness. Mutations in nsp12-250 clearly increase replication fidelity, RNA synthesis, and fitness in the ExoN(-) background (Figures 18, 19, 24, 25, and 27), but the effects of other replicase mutations remains unclear. To date, all fidelity-altering mutations in MHV have affected 5-fluorouracil sensitivity (Graepel et al., 2017; Sexton et al., 2016; Smith et al., 2013; 2015). No combination of P160 replicase mutations increased resistance to 5-fluorouracil beyond that of nsp12-250 alone, strongly suggesting that they are not important for regulating fidelity in the absence of proofreading. However, these mutations are likely still important for viral fitness, perhaps by modulating interactions between nsp12-250 and nsp14-250, which may be antagonistic without additional mutations (Figure 19). Such effects could be revealed by replication analyses and studies of RNA synthesis. Further, competitive fitness assays could aid in determining whether these mutants are indeed the result of positive selection or are artifacts of genetic drift. Finally, these mutations could mediate interactions with proteins outside of the canonical RTC, explaining why their effects were not detected as isolated recombinants.

Do structural or accessory proteins regulate coronavirus fidelity? The paradigm of low-fidelity replication in RNA viruses focuses on RdRps as the primary regulators of error rates. By

encoding a multi-subunit replicase complex that includes a co-factor-driven, proofreading exoribonuclease, coronaviruses have already broken the paradigm. My data suggest that, yet again, coronaviruses may be breaking the mold. Indeed, my results indicate that mutations in structural and/or accessory proteins may contribute to nucleoside analogue resistance in ExoN(-) P160 (Figure 27). We previously demonstrated that spike protein mutations, and the associated increase in replication kinetics, did not alter 5-fluorouracil sensitivity (Figure 16). However, we found single nonsynonymous mutations in both the envelope (E) and matrix (M) proteins of ExoN(-) P160, implicating these proteins in resistance to 5-fluorouracil. It is tempting, though perhaps far-fetched, to speculate that coronavirus structural proteins are direct fidelity regulators. Indeed, coronavirus structural proteins cannot be produced until after the replication machinery is assembled, as they are translated from subgenomic mRNAs (Figure 2 and Figure 3). Further, E and M are not closely associated with the RTC during MHV replication (V'kovski et al., 2019). However, the pre-membrane/membrane and envelope proteins of yellow fever virus partially control genetic diversity, though by an undefined mechanism (Collins et al., 2018). One potential explanation relates to the kinetics of replication. For example, mutations in the lysis protein of bacteriophage ϕX174 limit the accumulation of 5-fluorouracil-induced mutations by reducing the total number of replication cycles required for population expansion (Pereira-Gómez and Sanjuán, 2014). If a similar mechanism is at play in ExoN(-) structural-160, then these mutations should resist other mutagenic nucleoside analogues, such as 5-azacytidine or ribavirin, but not a non-obligate chain terminator like 2′-*C*-methyladenosine. Deep sequencing experiments with PrimerID (Chapter 5) would also be useful to determine whether the 5-fluorouracil resistance of ExoN(-) structural-160 is driven by high-fidelity replication or an alternative mechanism. Regardless of whether structural and accessory proteins resist 5-fluorouracil by altering fidelity

or some other mechanism, examining their contributions to nucleoside analogue resistance will be important for understanding how coronaviruses can evolve in response to antiviral drugs.

Which mechanisms other than increased fidelity might account for MHV-ExoN(-) P250 nucleoside analogue resistance? Genomic mutations in RNA viruses are most commonly detrimental or lethal (Elena and Moya, 1999; Malpica et al., 2002; Peris et al., 2010; Sanjuán et al., 2004a; Visher et al., 2016). One strategy to prevent extinction by mutational load is to increase replication fidelity, as discussed above. An alternative strategy is to become more mutationally robust. Mutational robustness describes a virus' capacity to buffer the fitness effects of mutations. In the setting of low-fidelity replication, as in MHV-ExoN(-), increased mutational robustness could have provided a selective advantage (Goldhill et al., 2014; Montville et al., 2005; Sanjuán et al., 2007). Selection for increased robustness could explain the ~90 synonymous changes in MHV-ExoN(-) P160. Synonymous changes can alter codons to reduce the probability of non-conservative amino acid changes (Lauring et al., 2012). Alternatively, the increased population size of MHV-ExoN(-) P160 could promote robustness by a "safety-in-numbers" effect, allowing efficient purging of low-fitness mutants while maintaining population fitness (Lauring et al., 2013). Large populations also increase the likelihood of co-infection, allowing complementation between viral genomes. Although increased replication conferred by mutations in spike did not alter 5-fluorouracil resistance (Figure 16), a study with poliovirus suggests that mutagenized populations have elevated coinfection frequencies (Aguilera et al., 2017). Thus, complementation may contribute to MHV-ExoN(-) P160 nucleoside analogue resistance. Conflicting evidence exists regarding whether mutational robustness itself affects the sensitivity to RNA mutagens (Graci et al., 2012; Lauring et al., 2012; O'Dea et al., 2010); nevertheless, the robustness of MHV-ExoN(-) P160 merits further investigation.

Conclusions. Coronaviruses are among the most genetically complex RNA viruses known, so it should come as no surprise that the interactions of mutations across the genome are equally complex. Despite the abundance of experimental data about coronavirus biology, our models for coronavirus replication probably lack critical details about how proteins across the genome interact during infection and about how genetic elements, including synonymous mutations, contribute to diverse phenotypes. My data indicate that even a phenotype as seemingly straightforward as resistance to nucleoside analogues, thought to be primarily driven by replication fidelity, may be multifactorial, and perhaps an accidental consequence of adaptation for other phenotypes. Accordingly, my data indicates that the high-level fitness of ExoN(-) P160 is driven by complex genetic and functional interactions between proteins across the entire genome that may not be isolable. Genetic and biochemical testing of the rich mutational resource of ExoN(-) P160 could reveal new insights about diverse areas of coronaviral biology and inform countermeasures for endemic and emerging coronaviruses.

CHAPTER 4: FITNESS BARRIERS LIMIT REVERSION OF A PROOFREADING-DEFICIENT CORONAVIRUS

Introduction

Prior to the start of this dissertation, a decade of work had demonstrated critical roles for the coronavirus ExoN in efficient and accurate RNA synthesis, fitness, and virulence. Given the defects associated with disruption of ExoN, ExoN(-) coronaviruses should be under strong selective pressure for primary reversion. In line with this expectation, ExoN motif III mutants of SARS-CoV and MHV rapidly and repeatedly revert ((Eckerle et al., 2010) and unpublished observations). In contrast, we have never detected partial or complete reversion of ExoN motif I mutants (ExoN-AA) in SARS-CoV or MHV during 10 years of study and hundreds of experiments. More specifically, we have not detected consensus or minority variants of any kind at the motif I AA codons in either virus strain during acute infections and prolonged passage in tissue culture and following treatment with multiple nucleoside analogues (Agostini et al., 2018; Case et al., 2017; Eckerle et al., 2007; 2010; Graepel et al., 2017; Sexton et al., 2016; Smith et al., 2013; 2015). SARS-CoV-ExoN-AA also is stable during acute and persistent animal infections in immunocompetent and immune-compromised mice (Graham et al., 2012). The lack of primary reversion is not due simply to reduced adaptive capacity, as both SARS-CoV- and MHV-ExoN-AA can adapt for increased replication (Eckerle et al., 2010; Graepel et al., 2017). Most strikingly, long-term passage of MHV-ExoN-AA (250 passages, P250) yielded a highly fit population that had directly compensated for defective proofreading through evolution of a likely high-fidelity RdRp (Graepel et al., 2017). Yet, where primary reversion would have required just four total consensus mutations, MHV-ExoN-AA-P250 contained more than 170.

In this chapter, I describe experiments to determine whether specific genetic or fitness barriers prevent primary reversion of ExoN motif I AA. To this end, I identified and engineered viable genetic pathways towards ExoN-AA motif I reversion in MHV (hereafter, ExoN-AA). My results show that partial reversion did not confer a selective advantage compared to ExoN-AA. Further, ExoN-AA adapted within 10 passages for greater fitness than any of the intermediate revertants. Finally, restoration of WT-ExoN-DE in the setting of passage-selected mutations in the nsp12 RNA-dependent RNA polymerase (RdRp) and nsp14-ExoN exacted profound fitness costs. Together, these data are the first observation of an ExoN(-) CoV genotype-fitness landscape and identify multiple genetic barriers underlying ExoN(-) motif I stability in MHV. Further, they suggest extensive coevolution between MHV replicase proteins during adaptation and reveal potential strategies for stabilizing ExoN mutant CoVs.

Co-author contributions

I performed all molecular biology and recovered all recombinant viruses described in this chapter, with two exceptions. Nicole Sexton generated ExoN(-) nsp12-M814K, and Xiaotao Lu recovered ExoN(+) nsp12-250 and isolated the ExoN(-) P10 plaque clone. Maria Agostini performed 5-fluorouracil sensitivity experiments in Figure 31 and Figure 32 and completed plaque assays for additional experiments.

Results

Primary reversion of ExoN(-) motif I. MHV-ExoN(-), hereafter ExoN-AA, contains two engineered substitutions in each codon of motif I, such that complete reversion to WT-ExoN-DE requires mutations to all four sites (Figure 29A). Viral mutation rates in the absence of

proofreading range from 10^{-4} to 10^{-6} mutations per nucleotide per round of replication (μ) (Sanjuán et al., 2010). Assuming an ExoN-AA mutation rate of 10^{-4} μ and accounting for codon degeneracy, the probability of restoring the native amino acid sequence in a single round of replication is 10^{-18}. Only rarely do ExoN-AA titers exceed 10^6 PFU/mL, so it is unlikely that ExoN-AA could navigate this genetic barrier in a single infectious cycle. Thus, we hypothesized that ExoN-AA reversion, if possible, would proceed incrementally. To identify potential pathways towards ExoN-AA reversion, we examined the possible single-nucleotide substitutions surrounding A89 and A91 (Figure 29B). Three mutations are synonymous, and five mutations yield amino acids unlikely to coordinate with the positively-charged metals required for ExoN catalysis (glycine, valine, proline, threonine, and serine) (Chen et al., 2007; Ma et al., 2015; Minskaia et al., 2006; Steitz and Steitz, 1993). One mutation per site can restore the acidic charge (i.e. AA-to-ED) but not the native amino acid. These variants have not been tested in a CoV ExoN, but biochemical studies of *E. coli* DNA polymerase I ExoN mutants suggest that these conservative substitutions would not restore WT-like ExoN activity (Derbyshire et al., 1991). We predicted stepwise pathways to ExoN-AA→DE reversion based on restoration of acidic charge followed by reversion to native amino acids (Figure 29C). We engineered and recovered variants in ExoN-AA requiring three mutations (3nt; ExoN-AD, ExoN-EA), two mutations (2nt; ExoN-DA, ExoN-ED, ExoN-AE), or one mutation (1nt; ExoN-DD, ExoN-EE) for reversion to WT-ExoN-DE (Table 4). We will hereafter refer to these mutants as intermediate revertants. All intermediate revertants generated viable progeny during recovery, demonstrating that reversion to WT-ExoN-DE along these pathways is theoretically possible. The 3nt and 2nt mutants were genetically stable during recovery, as confirmed by dideoxy sequencing. However, both 1nt mutants (ExoN-DD and ExoN-EE) reverted to WT-ExoN-DE during three independent

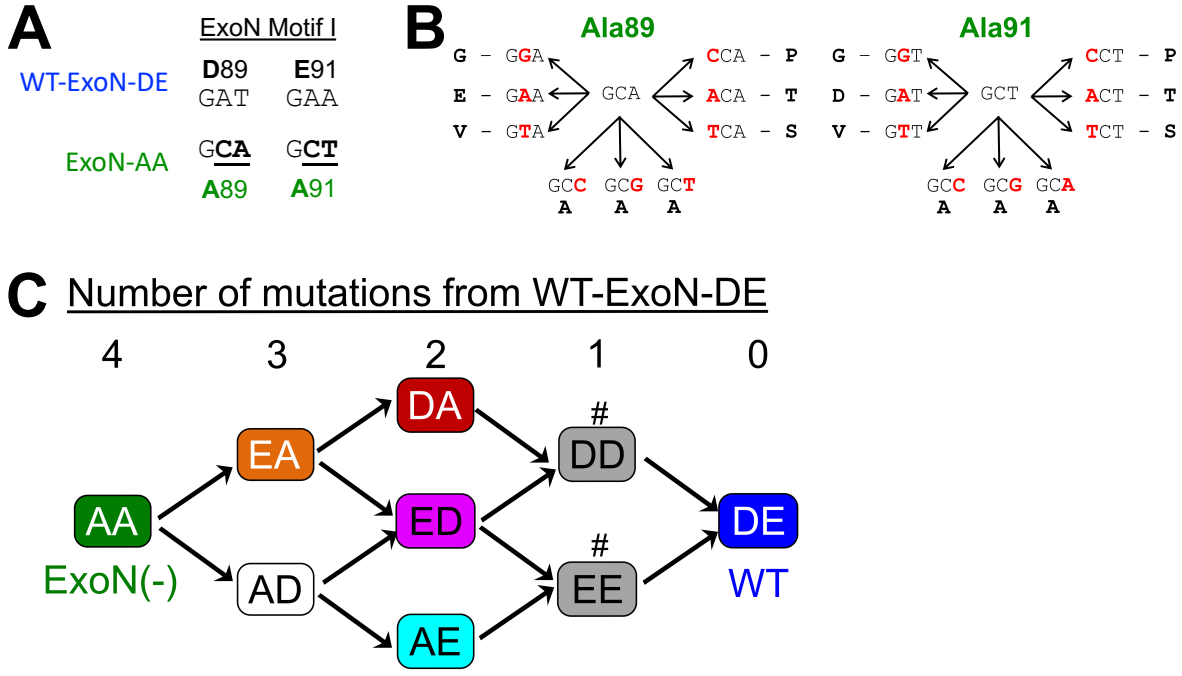


Figure 29. Sequence landscape around ExoN-AA motif I. (*A*) ExoN motif I nucleotide sequences. (*B*) Landscape of single-nucleotide substitutions within ExoN-AA motif I. (*C*) Predicted pathways to reversion of ExoN-AA. Variants marked with # reverted to WT during three independent recovery attempts.

recovery attempts, suggesting that these two variants are less fit than WT-ExoN-DE and demonstrating that reversion by 1nt mutation is readily accessible. To test whether the 3nt or 2nt mutants would revert more rapidly than ExoN-AA (4nt), we passaged three lineages of each mutant 10 times at multiplicities of infection (MOI) of 0.5 and 0.01 PFU/cell. We harvested supernatants and screened for reversion by visual inspection of plaque phenotypes at each passage. WT-ExoN-DE and WT-like viruses produce uniform, large plaques, while ExoN-AA-like viruses yield small, variably-sized plaques (Eckerle et al., 2007). When we observed mixed plaque phenotypes, we sequenced three large plaques from each lineage to confirm reversion. The 3nt (ExoN-AD and ExoN-EA) and 2nt (ExoN-DA and ExoN-ED) intermediate revertants showed no evidence of reversion over 10 passages at either MOI (Table 4). In contrast, the 2nt ExoN-AE contained WT-revertants by P2 in all lineages at MOI = 0.5 PFU/cell and by P8 in one lineage at MOI = 0.01 PFU/cell. Once observed, WT-revertants dominated the ExoN-AE population for the remaining passages. These data indicate that at least one 2nt mutation pathway can lead to full reversion in tissue culture. The probability of ExoN-AE arising during a single infectious cycle of ExoN-AA is low but theoretically achievable ($\sim 10^{-9}$), so ExoN-AA could conceivably revert within just two infectious cycles. However, complete reversion has never been observed even during prolonged passage or persistent infections, suggesting that additional barriers to the replication, fitness, or maintenance of intermediate revertants exist.

Partial reversion of MHV-ExoN(-) motif I does not confer a selective advantage. Because the intermediate revertants are viable as recombinants but are not found in ExoN-AA populations, we hypothesized that they confer no selective advantage over ExoN-AA (Eckerle et al., 2007; Sexton et al., 2016; Smith et al., 2013). To test this hypothesis, we first analyzed replication of the 3nt and 2nt intermediate revertants (Figure 30A). All variants achieved similar

Table 4. Recovery and passage of intermediate revertants.

n.d.: not done.
[a]Bolded nucleotides must mutate to reach WT-ExoN-DE genotype.
[b]Recovered viruses were subjected to 10 passages at the indicated MOI. Samples were screened for wild-type revertants by plaque assay, and revertant lineages were sequence-confirmed.

Virus	# of mutations to WT-ExoN-DE	Motif I sequence[a]	# of reverted lineages by passage 10[b]	
			MOI = 0.01	MOI = 0.5
ExoN-AA	4	G**CA**...G**CT**	0/3	0/3
ExoN-AD	3	G**CA**...GA**T**	0/3	0/3
ExoN-EA	3	GA**A**...G**CT**	0/3	0/3
ExoN-DA	2	GAT...G**CT**	0/3	0/3
ExoN-AE	2	G**CA**...GAA	1/3 (by P8)	3/3 (by P2)
ExoN-ED	2	GA**A**...GA**T**	0/3	0/3
ExoN-EE	1	GA**A**...GAA	n.d.	n.d.
ExoN-DD	1	GAT...GA**T**	n.d.	n.d.
WT-ExoN-DE	0	GAT...GAA	n.d.	n.d.

peak titers to ExoN-AA, but detailed examination of their kinetics suggested a potential delay of up to 1.5 hours for all intermediate revertants compared to ExoN-AA. Of note, ExoN-AE was the most delayed, and we detected WT-ExoN-DE revertants in two of three replicates, suggesting increased selective pressure against this variant. We next measured the competitive fitness of each intermediate revertant relative to a recombinant ExoN-AA containing seven silent mutations in the nsp2 coding region (ExoN-AA-reference). Intermediate revertants were mixed with an equal titer of ExoN-AA-reference at a combined MOI = 0.05 PFU/cell and passaged four times. The ratio of each intermediate revertant to ExoN-AA-reference was quantified at each passage by RT-qPCR, and the change in ratio over time was used to calculate their relative fitness. WT-ExoN-DE was significantly more fit than ExoN-AA, whereas the intermediate revertants (ExoN-AD, -EA, -DA, and -ED) had no increased fitness relative to ExoN-AA (Figure 30B). The apparent increased fitness of ExoN-AE resulted from all lineages reverting to WT-ExoN-DE during the experiment. Finally, our previous studies have shown that adaptation of ExoN-AA includes partial compensation for the replication fidelity defect, as measured by reduced susceptibility to the mutagen 5-fluorouracil (5-FU) (Case et al., 2016; 2017; Graepel et al., 2017; Sexton et al., 2016; Smith et al., 2013; 2015). None of the intermediate variants demonstrated statistically significant differences in 5-FU sensitivity as compared to ExoN-AA (Figure 30C). Thus, with the exception of the ExoN-AE→DE revertants, no 3nt and 2nt intermediate genotypes along our predicted pathway demonstrated an advantage in replication, fitness, or fidelity that would favor their maintenance or expansion in the viral population. Thus, natural selection is unlikely to drive ExoN-AA down these pathways towards reversion.

Secondary adaptations outside of ExoN-AA motif I increase fitness along alternative pathways. Although we did not find fitness advantages to intermediate revertants, we also did

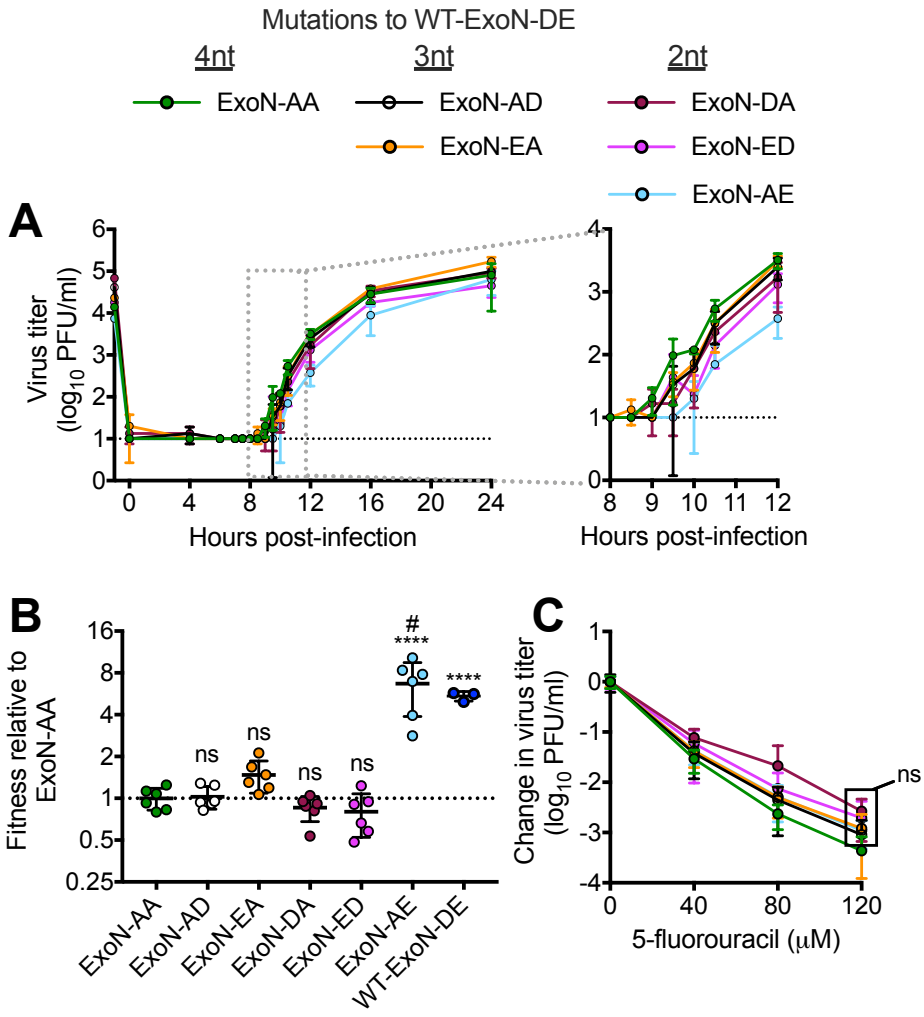


Figure 30. Intermediate revertants of ExoN-AA motif I do not have selective advantages. (*A*) Replication kinetics at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) Competitive fitness of each variant relative to ExoN-AA. Viruses were competed with a tagged ExoN-AA-reference strain, and relative fitness was normalized to the mean of ExoN-AA. (*C*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell. Statistical significance of each variant relative to ExoN-AA was determined by one-way ANOVA with multiple comparisons (Panel *D*) two-way ANOVA with Dunnett's multiple comparisons (panel *C*). ****p<0.0001; ns = not significant. Data in (*B*) and (*C*) represent mean ± SD of n = 6. Boxed values have the same significance. [#]All lineages of ExoN-AE reverted to WT-ExoN-DE during the experiment.

not identify profound fitness costs that would drive their immediate loss from populations. We have previously demonstrated that during 250 passages (P250), ExoN-AA can adapt for increased replication, fitness, and fidelity via secondary mutations outside of motif I (Graepel et al., 2017). We tested whether secondary adaptive mutations could exceed the fitness of ExoN-AA and its intermediate revertants. To examine the early adaptation of ExoN-AA, we studied passage 10 from the P250 passage series (Figure 31). ExoN-AA-P10 retains the ExoN-AA motif I genotype but has increased replication and reduced susceptibility to 5-FU, altogether manifesting in greater relative fitness (Figure 31) (Graepel et al., 2017). We identified only six total mutations within ExoN-AA-P10 by dideoxy sequencing (Table 5), indicating that rapid adaptation of and compensation for ExoN-AA requires relatively few genetic changes at the consensus level. To test whether interactions between multiple mutations or population level effects contribute to ExoN-AA-P10 fitness, we isolated a plaque-purified clone of ExoN-AA-P10. The clone replicated to higher titers than the ExoN-AA-P10 population but had identical 5-FU sensitivity and relative fitness (Figure 31), indicating that genomes derived from a single virus plaque encode the adaptive changes required by the total population. Together, these data demonstrate that mutations outside of ExoN(-) motif I can confer greater fitness advantages than intermediate revertants even at early passages. These early adaptive mutations likely reduce the selective pressure for motif I reversion and place the intermediate revertants at a selective disadvantage.

Adaptive mutations in nsp12 and nsp14 that increase ExoN-AA fitness confer significant fitness costs to WT-ExoN-DE. Mutational fitness effects are highly dependent upon the genetic background (Das et al., 2013; Koel et al., 2018; Nakajima et al., 2005). In addition to reducing selective pressure for reversion, mutations conferring increased fitness to ExoN-AA might also

Table 5. Mutations in ExoN(-) P10. Data derived from dideoxy sequencing.

[a]Mutation present at approximately 50% of population.
[b]MHV HE is not transcribed in tissue culture.
[c]Amino acid numbers designate positions within cleaved nsps, not the polyprotein.

Mutation number	Nucleotide change	Protein	Amino acid change[c]
1	G2520A[a]	nsp2	D524N
2	A3080G[a]	nsp3	silent
3	T16017A	nsp12	M814K
4	A17836G[a]	nsp13	I492M
5	G22673A[a]	HE[b]	noncoding
6	A29298C[a]	M	silent

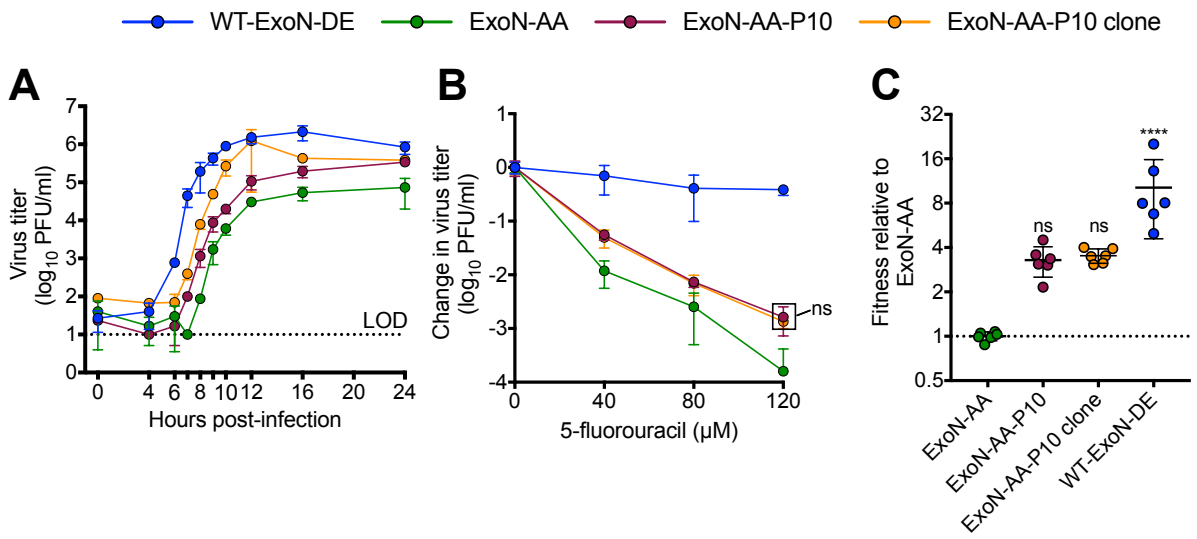


Figure 31. ExoN-AA adapts for increased fitness within 10 passages. (*A*) Replication kinetics of indicated viruses at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell. (*C*) Competitive fitness of individual recombinants relative to ExoN-AA. Viruses were competed with a tagged ExoN-AA-reference strain, and relative fitness was normalized to the mean of ExoN-AA. Statistical significance of each virus relative to ExoN-AA was determined by two-way ANOVA with Dunnett's multiple comparisons (panel *B*) or by one-way ANOVA with multiple comparisons (Panel *C*). ****p < 0.0001, ns = not significant. LOD = limit of detection. Data in (*B*) and (*C*) represent mean ± SD of n = 6. Boxed values have the same significance.

reduce the benefits of motif I reversion. We previously reported that long-term passage of ExoN-AA selects for secondary adaptive mutations in the nsp12 RdRp and nsp14 (nsp12-P250 and nsp14-P250) (Graepel et al., 2017). Nsp12-P250 contains 7 nonsynonymous mutations that partially compensate for defective proofreading and increase ExoN-AA fitness. Nsp14-P250 contains 6 nonsynonymous mutations, including a conservative D-to-E substitution in ExoN motif III, and increases ExoN-AA fitness without compensating for defective proofreading. To test whether the fitness effects of passage-associated mutations in nsp12-P250 and nsp14-P250 depend upon the ExoN-AA genotype, we engineered a WT motif I (ExoN-DE) into viruses containing nsp12-P250 and nsp14-P250, alone and together, and analyzed replication, 5-FU sensitivity, and competitive fitness. Compared to WT-ExoN-DE, both ExoN-DE-nsp12-P250 and ExoN-DE-nsp14-P250 displayed delayed and decreased replication (Figure 32A). In 5-FU sensitivity assays, ExoN-DE-nsp14-P250 was indistinguishable from WT-ExoN-DE, while both variants containing nsp12-P250 (ExoN-DE-nsp12-P250 and ExoN-DE-nsp12/14-P250) were significantly more sensitive to 5-FU (Figure 32B). Finally, the nsp12-P250 and nsp14-P250 mutations significantly decreased fitness relative to WT-ExoN-DE (Figure 32C). We detected no statistical differences between the specific infectivity of WT-ExoN-DE and any of the nsp12-P250 and nsp14-P250 variants in isolated infections (Figure 32D). Thus, mutations in nsp12 and nsp14 that arose in the ExoN-AA background were detrimental to replication, mutagen sensitivity, and competitive fitness in the presence of a fully-reverted ExoN-DE. These results support the conclusion that the adaptive pathways available to ExoN-AA may stabilize the ExoN-AA genotype, reducing both the selective pressure for, and the potential benefits of, primary reversion.

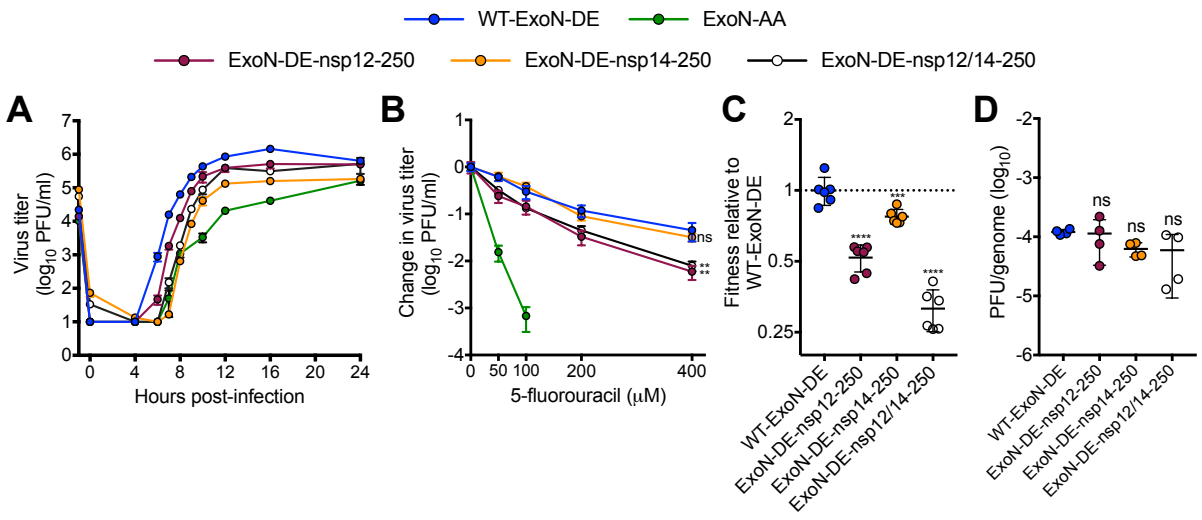


Figure 32. Mutations that increase ExoN-AA fitness are detrimental in the presence of WT-ExoN-DE. (*A*) Replication kinetics of indicated viruses at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell, mean ± SD of n = 6. (*C*) Competitive fitness of individual recombinants relative to WT-ExoN-DE. Viruses were competed with a tagged WT-ExoN-DE reference strain, and relative fitness was normalized to the mean of WT-ExoN-DE, mean ± SD of n = 6. (*D*) Specific infectivity (genomes per PFU) from isolated infections, mean ± SD of n = 4.. Statistical significance of each virus relative to WT-ExoN-DE was determined with two-way ANOVA with Dunnett's multiple comparisons test (panel *B*) or by ordinary one-way ANOVA with Dunnett's multiple comparisons test (panels *C* and *D*). **p < 0.01; ***p < 0.001; ****p < 0.0001; ns = not significant.

Discussion

In this study, we demonstrate that the stability of the ExoN(-) motif I genotype in MHV (ExoN-AA) is a consequence of the limitations and opportunities of the genetic landscape it explores during replication (Figure 33). Our results support a model in which the viable adaptive pathways leading to direct reversion of motif I from AA-to-DE are relatively flat on a fitness landscapes, as intermediate revertants remain phenotypically ExoN(-) and confer no fitness advantage over ExoN-AA. In contrast, at least one alternative adaptive pathway is readily accessible and imparts immediate fitness gains over ExoN-AA. We propose that even minimal alternative pathway adaptive fitness gains reduce the likelihood and benefits of motif I reversion, until eventually the changing genetic background renders reversion detrimental. These data and this model suggest that selection during replication favors immediate, incremental fitness gains along the most accessible pathway rather than dramatic fitness increases across a larger genetic barrier.

Our results also extend existing studies of CoV ExoN motif I. Motif I AA→DE mutations in the SARS-CoV nsp14-ExoN dramatically reduce nuclease activity in biochemical assays, but no study has examined the contributions of each residue independently (Bouvet et al., 2012; Ma et al., 2015). Intermediate revertants of ExoN-AA did not display consistent or statistical differences in replication, 5-FU sensitivity, or competitive fitness relative to ExoN-AA, suggesting that they remain phenotypically ExoN(-) during infection and supporting previous studies that motif I DE is essential for WT ExoN function. Given these results, we were surprised to observe repeated reversion of the ExoN-AE but not the other two 2nt variants, ExoN-DA and ExoN-ED. One potential explanation is that the specific mutational bias of ExoN-AE makes the

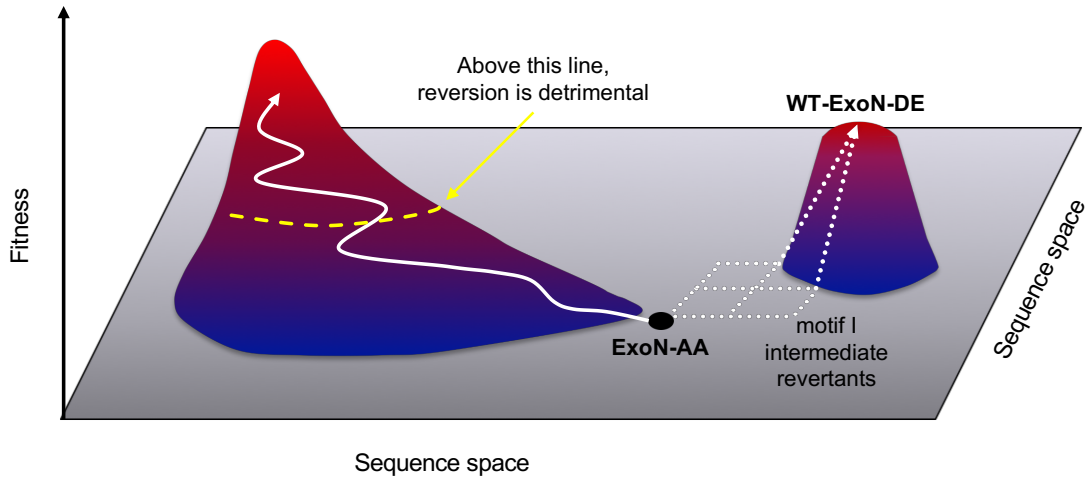


Figure 33. Model for the *in vitro* evolution of MHV-ExoN-AA. MHV-ExoN-AA(black dot) is a low-fitness variant. Reversion to WT-ExoN-DE would dramatically increase fitness but can only be achieved by traversing a flat landscape and climbing a steep fitness cliff (dotted white arrows). However, secondary mutations that incrementally increase fitness are more accessible (solid white arrow). Eventually, the genetic background changes enough that reversion becomes detrimental (dotted yellow line).

revertant mutations more accessible than in ExoN-DA or ExoN-ED. Alternatively, if ExoN-AE has profound replication or fitness defects, selection could drive primary reversion more quickly away from this genotype. Consistent with this hypothesis, ExoN-AE reverted more quickly at a higher MOI, where natural selection acts more efficiently on a larger population size (Table 1) (Dolan et al., 2018b). Biochemical studies would be valuable to understand how nsp14-ExoN-AE differs from the other intermediate revertants, with an eye towards understanding the catalytic constraints and functional interactions with nascent CoV RNA.

CoV replication proceeds through the concerted action of multiple proteins proposed to resemble DNA replication holoenzymes (Smith et al., 2014). Due to the extensive interactions between CoV proteins, they must coevolve in a highly cooperative manner to maintain their essential functions. Consistent with this hypothesis, the fitness effects of mutations in nsp12-P250 and nsp14-P250 differ based on the motif I genotype; they are beneficial in ExoN-AA but detrimental in the WT-ExoN-DE background. In previous studies, it has been difficult to determine whether the fitness defects in ExoN(-) CoVs are directly linked to low-fidelity replication or through some other mechanism. Our data suggest that the proofreading function of nsp14-ExoN can be uncoupled from its more general role in replication (Figure 32), providing an opportunity to examine additional roles for this essential protein. Nsp12-P250 also will be an important tool for understanding the relationship between RdRp fidelity and ExoN proofreading during CoV replication and for studying replication complex assembly and interactions. Our studies suggest that compensatory mutations identified through long-term passage could stabilize the ExoN-AA genotype. In particular, the high-fidelity nsp12-P250 could reduce the probability of reversion by reducing mutational sampling within motif I (Arnold et al., 2005), and both nsp12-P250 and nsp14-P250 render the MHV genome inhospitable to a WT-ExoN-DE.

Together, these studies argue that experimental evolution can generate reagents to define critical interactions involved in CoV replication and can identify new strategies for stabilizing attenuated CoVs.

In this chapter, I provide the first glimpse of a coronavirus genotype-fitness landscape and establish the genetic basis of ExoN(-) motif I stability in MHV. I also demonstrate that compensatory mutations derived from experimental evolution are highly dependent upon the genetic background, suggesting that they may be useful for stabilizing live-attenuated coronavirus vaccines. The high conservation of nsp14-ExoN across diverse coronaviruses suggests that they may be subject to similar genetic barriers, supporting the proposed strategy of ExoN-inactivation as a broadly applicable platform for attenuation of emerging coronaviruses.

CHAPTER 5: NEW METHODS FOR MEASURING MURINE HEPATITIS VIRUS FIDELITY AND FITNESS

Introduction

Mutation rates and viral fitness are two key parameters that influence viral evolutionary dynamics. Recent technological advances have greatly enhanced our ability to measure mutation rates and viral fitness with high sensitivity and accuracy (Dolan et al., 2018a). Most of the new tools are developed for model viruses like poliovirus and influenza due to the ease of genome manipulation and the sheer number of investigators working on these pathogens. By comparison, the tools available for studying coronavirus evolution are limited in scope and sensitivity. In this chapter, I discuss the development of two new tools for murine hepatitis virus: a next-generation sequencing pipeline for highly accurate measurement of mutation rates and a competitive fitness assay for MHV variants.

Coauthor contributions

I designed all primer sets for PrimerID and performed all analysis. Tia Hughes optimized library preparation and generated all libraries reported in this dissertation. I received bioinformatic software and guidance from Shuntai Zhou (University of North Carolina Chapel Hill), Adam Lauring and Matthew Pauly (University of Michigan), Jennifer Gribble (Vanderbilt University, Denison lab), and Ben Reisman (Vanderbilt University, Bachmann lab). Jim Chappell helped me to design the competition experiments. I performed all benchwork and analysis for competitive fitness assays.

Highly accurate measurement of coronaviral mutation frequencies and rates by deep sequencing

The notorious adaptability of RNA viruses is driven by their capacity to generate enormous genetic diversity through mutation and recombination. Accurate measurements of mutation rates and biases in RNA viruses is essential for understanding and anticipating their evolutionary dynamics. The primary viral determinant of mutation rates, replication fidelity, can be measured in a variety of ways. Biochemical experiments with expressed proteins are the gold standard for quantifying polymerase fidelity and have been instrumental in identifying the molecular mechanisms of nucleotide (mis)incorporation (Arnold and Cameron, 2004; Arnold et al., 2005; Ferron et al., 2018; Moustafa et al., 2014). Biochemical data for coronavirus replication is, to date, relatively limited. The SARS-CoV RNA-dependent RNA polymerase is unstable when expressed alone and requires, at minimum, nsp7 and nsp8 for stable and processive polymerization (Subissi et al., 2014; Velthuis et al., 2010). Recent biochemical data demonstrates that the SARS-CoV RdRp (in complex with nsp7 and nsp8) has lower intrinsic fidelity than the dengue virus RdRp (isolated protein), probably indicating that the coronavirus proofreading ExoN relieves the RdRp of constraints on fidelity (Ferron et al., 2018). Consistent with this hypothesis, MHV RdRp fidelity effects are masked by intact proofreading, so biochemical studies should ideally incorporate all components of the purported coronavirus replicase-transcriptase complex (Chapter 1, Figure 10) (Sexton et al., 2016; Smith et al., 2014). To date, no such biochemical system exists. Still, biochemical systems, by design, exclude the complex milieu of viral and cellular factors that may influence fidelity during infections, potentially obscuring details about replication in living cells.

Nucleoside analogues (artificial compounds structurally similar to natural nucleosides) can also

be used to measure replication fidelity. Resistance to one or more nucleoside analogues typically correlates with higher overall fidelity (for example, (Arribas et al., 2011; Graepel et al., 2017; Pfeiffer and Kirkegaard, 2003; Sexton et al., 2016; Stapleford et al., 2015). However, mutations conferring nucleoside analogue resistance can decrease fidelity in the absence of drug, suggesting that mechanisms for excluding unnatural nucleosides are not identical to those preventing misincorporation of natural nucleosides (Arias et al., 2008; Sexton et al., 2016; Sierra et al., 2007). Further, the results of nucleoside analogue sensitivity experiments are, at best, semi-quantitative (e.g. virus A is 5-fold less sensitive than virus B to drug X at concentration Y), making it difficult to extrapolate to conclusions about replication errors in the absence of drug. Still, nucleoside analogue experiments have been massively informative for understanding fidelity regulation in coronaviruses and have contributed key findings to the body of work supporting the proofreading function of nsp14-ExoN (Graepel et al., 2017; Sexton et al., 2016; Smith et al., 2013; 2015).

Whereas biochemical studies and nucleoside analogue sensitivities examine fidelity under "non-normal" conditions, fidelity can also be investigated during "normal" infection cycles, in the presence of all viral and cellular factors that contribute to replication errors. Results from such studies are typically reported as either mutation rates or mutation frequencies (Bradwell et al., 2013; Peris et al., 2010). The difference between rates and frequencies is subtle but important. Mutation rates describe the number of errors incorporated during a discrete unit of time (e.g. mutations per infection cycle); that is, the rate at which the replicase makes mistakes. By contrast, a mutation frequency is the number of mutations identified per nucleotide sequenced. Frequency measurements, therefore, reflect both the rate at which a mutation is generated and its ability to persist in the population. Thus, mutation frequencies can be skewed by individual

fitness effects and genetic drift. Still, mutation rates and frequencies are both valid measurements of replication fidelity in experimental settings.

A common strategy for measuring mutation rates is to use a Luria-Delbrück fluctuation test (Luria and Delbrück, 1943). In this method, parallel cultures are assessed for a scorable phenotype after a defined period of growth. Scorable phenotypes include, for example, escape from neutralizing antibodies or sensitivity to guanidine-hydrochloride (Furió et al., 2005; Pfeiffer and Kirkegaard, 2003), and should be accessible by just a single nucleotide change. The ratio of resistant-to-sensitive variants corresponds to the mutation rate. However, these typically sample just a few sites or mutational classes (e.g. A→G, G→C, etc) and may not reflect the rate of all possible mutation types. Recently, Pauly et al. reported a fluctuation test for influenza based on reversion to fluorescence of virally-encoded green fluorescent protein (GFP) mutants that can interrogate all 12 mutation classes (Pauly et al., 2017b). This type of assay is potentially feasible in coronaviruses but would require significant genetic engineering. The stability of the required viruses would be difficult to predict given high rates of recombination in coronaviruses that rapidly delete non-essential genes (Chinese SARS Molecular Epidemiology Consortium, 2004; Graepel et al., 2017; Makino et al., 1986).

Genome sequencing studies can also be used to measure viral mutation rates and frequencies. One strategy is to clone and dideoxy sequence large numbers of individual genomes to determine changes in mutation accumulation (Gnädig et al., 2012; Rozen-Gagnon et al., 2014). However, this method has a small dynamic range unless thousands of clones are analyzed, which quickly becomes labor-prohibitive (Sexton, 2017). Next-generation sequencing ("deep sequencing") strategies are less labor-intensive and more sensitive for measuring mutation frequencies.

However, artifacts incorporated and propagated during all steps of sample preparation and sequencing occur at a rate of ~1-2%, confounding data interpretation (Dohm et al., 2008). New techniques for sample preparation, including circular resequencing (cir-seq) and duplex sequencing, have been developed to computationally identify and remove erroneous bases (Acevedo et al., 2014; Lou et al., 2013; Schmitt et al., 2012). However, cir-seq and duplex sequencing do not enrich for virus-specific sequence and require large quantities of RNA, which may be limiting for some viral variants. For instance, ExoN-inactivation greatly reduces genomic RNA production (Eckerle et al., 2007; Graepel et al., 2017; Smith et al., 2013). A separate strategy, termed PrimerID, offers similar error correction but requires vastly less total RNA (Figure 34). In this strategy, individual RNA genomes are tagged with a unique barcode (PrimerID) during reverse transcription, allowing reconstruction of consensus sequences derived from the same initial template during analysis. Genome-specific primers allow selective amplification and sequencing of viral RNA even from contaminated mixtures, and robust data can be obtained with as few as 370 initial templates (Zhou et al., 2015). For these reasons, I chose to develop a PrimerID-based sequencing assay for MHV.

Amplicon selection and primer design. Mutation rates can be measured as the frequency of "nonsense mutational targets" (NSMT) in sequencing data (Acevedo et al., 2014; Cuevas et al., 2015; 2009; Pauly et al., 2017b). NSMT are specific mutations that are lethal for viral replication and therefore should be eliminated by negative selection. Put another way, lethal mutations should enter the population according to the mutation rate and should be removed immediately. Thus, the frequency of NSMTs in a population approximates the mutation rate. NSMT include, but are not limited to, mutations that introduce premature stop codons into RNA virus open-reading frames (Peris et al., 2010; Sanjuán et al., 2004a; Visher et al., 2016). Eighteen of the 61

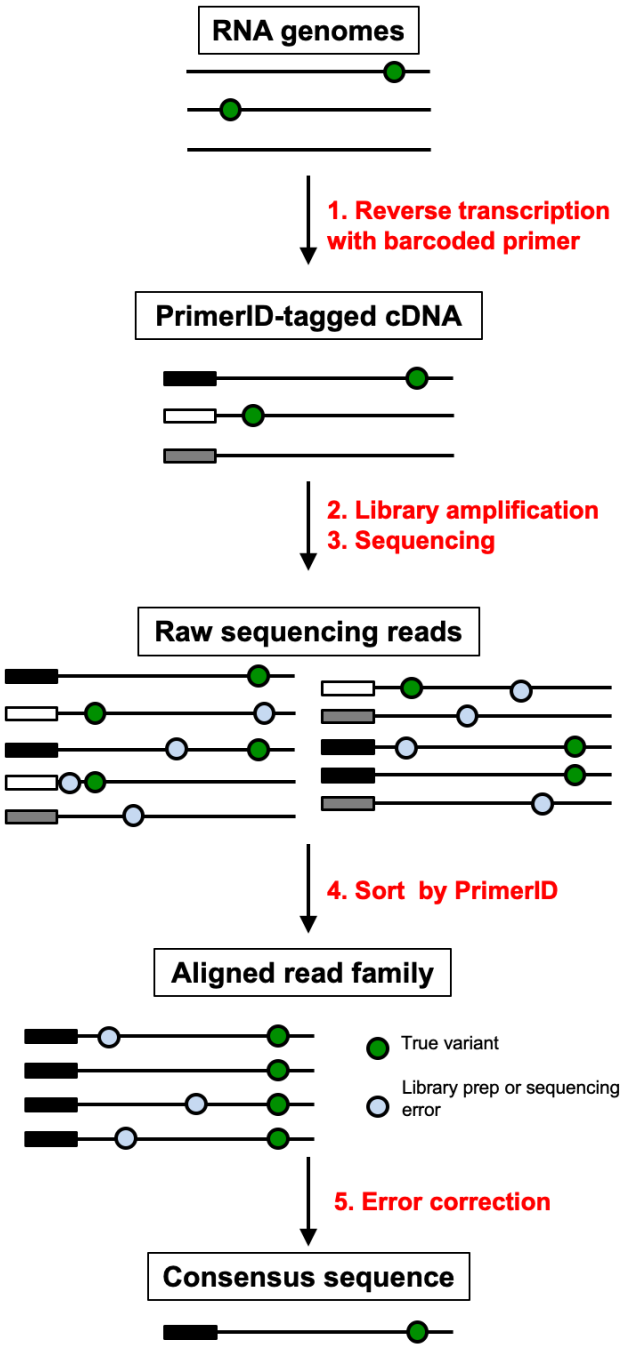


Figure 34. High-resolution measurement of mutation frequencies using PrimerID. Genomic RNA templates are tagged with unique barcodes (PrimerIDs) during reverse transcription (1). Errors are introduced during library preparation and sequencing (2-3), but PrimerIDs allow construction of read families derived from the same original template (4). Mutations present in all reads are likely true variants, and those present in few templates are discarded (5).

sense codons are just a single mutation away from becoming a stop codon (Table 6). With sufficient coverage, deep sequencing experiments can use NSMT to measure mutation rates for 8 mutational classes (Acevedo et al., 2014; Cuevas et al., 2009; 2015; Pauly et al., 2017b). I used a custom R script, generously provided by Adam Lauring (https://github.com/lauringlab/NGS_mutation_rate_assay), to identify a 423 base region in ORF1ab of MHV-A59 (nts 20814-21265, covering portions of nsp15 and nsp16) with a balanced concentration of 63 total NSMT (MHV20814; Table 6). I used the IDT Oligoanalyzer (www.idtdna.com/calc/analyzer) to design gene-specific primers with little predicted secondary structure or cross-reactivity. Ten degenerate nucleotides (N^{10}) in the reverse transcription primer serve as the primerID, allowing specific labeling of up to 1,048,576 (4^{10}) unique templates (Figure 35). The forward primer for the first round of PCR is virus-specific and contains Illumina adaptor sequences and four degenerate nucleotides to increase sequencing diversity; the reverse primer adds an Illumina adaptor. A second round of PCR appends index barcodes for multiplex sequencing. See Appendix B.1. for the library preparation protocol, including primer sequences.

Example PrimerID experiment. We performed PrimerID on 100,000 copies of virion RNA from WT-MHV and ExoN(-) P3. The libraries were sequenced on the Illumina MiSeq with 2x250 paired-end chemistry along with 50% bacterial DNA. Sequencing was performed by the Vanderbilt Technologies for Advanced Genomics (VANTAGE). Consensus sequences based on PrimerIDs were obtained using the web portal operated by the Swanstrom lab at UNC Chapel Hill, who developed this method (https://tcs-dr-dept-tcs.cloudapps.unc.edu/TCS/) (Zhou et al., 2015). Jennifer Gribble assisted with data analysis, and additional code was provided by Adam Lauring (https://github.com/lauringlab/NGS_mutation_rate_assay) (Pauly et al., 2017b). We obtained >2 million reads per sample that condensed to >40,000 consensus sequences,

Table 6. Nonsense mutational targets. Eighteen of the 61 sense codons are one mutation away from becoming premature stop codons, representing eight mutational classes. Four mutational classes (N-to-C and A-to-G) cannot reach premature stop codons. The MHV20814 amplicon contains 63 total NSMT.

NSMT type	NSMT Codons	Abundance in MHV20814 amplicon
U-to-A and U-to-G	UUA, UAU	8
U-to-A	UUG, UGU	10
C-to-A, C-to-G	UCA, UAC	10
C-to-A	UGC, UCG	5
C-to-U	CAA, CAG, CGA	6
G-to-A	UGG	7
A-to-U	AAA, AAG, AGA	6
G-to-U	GAA, GAG, GGA	11

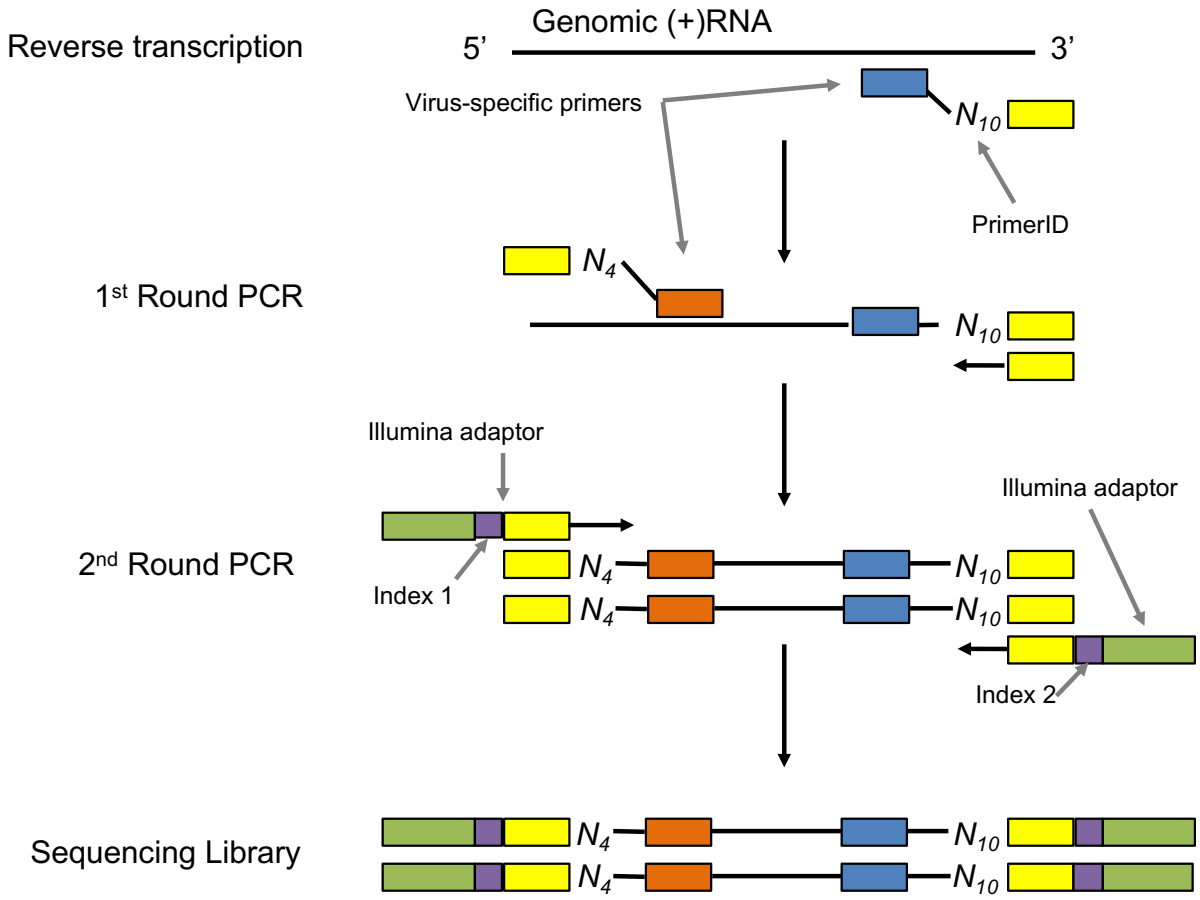


Figure 35. Library preparation for PrimerID sequencing. Genomic RNA is reverse transcribed with a gene-specific primer containing a 10-nucleotide PrimerID. A second gene-specific primer is used to amplify the region of interest, and then Illumina sequencing adaptors are added during a second round of PCR.

representing 40.8% and 43.1% template recovery for WT-MHV and ExoN(-) P3 (Table 7). The mutation frequency of ExoN(-) P3 was 9.6-fold higher than that of WT, consistent with published results (Table 7 and Figure 36A) (Eckerle et al., 2007; 2010; Sexton et al., 2016). I also calculated mutation frequencies for individual mutational classes (Figure 36A, right). G-to-A mutations were not dramatically overrepresented in these samples, as was expected due to the known mutational biases of reverse transcriptases (Cuevas et al., 2015; Holtz and Mansky, 2013; Pauly et al., 2017b). Neither virus had any C-to-G mutations at this coverage depth, recapitulating results from (Sexton et al., 2016). All other mutation types except for G-to-U were detected at greater frequency in ExoN(-) P3 than WT-MHV, with the largest difference detected for U-to-A substitutions. I also used the NSMT method to estimate the overall mutation rate for each virus (Table 7 and Figure 36B). For the eight mutational classes represented, the overall mutation rates for ExoN(-) P3 and WT-MHV were 2.02×10^{-5} and 5.33×10^{-6}, respectively, a difference of 3.8-fold and close to previous estimates (Eckerle et al., 2007). The mutation rates for WT-MHV and ExoN(-) P3 were lower than mutation frequencies by 6- and 16-fold, respectively. The accumulation of mutations to frequencies above those of the NSMT mutations indicates the action of selection on newly generated variants, which may confound frequency measurements. I also calculated the ratio of mutation rates and frequencies of ExoN(-) P3 to WT-MHV for each class (Figure 36C). The general shape of the distribution was similar between mutation rates and frequencies, but in most mutational classes, the ratio of mutation frequencies was greater than the ratio of mutation rates, further suggesting the action of selection. Without multiple replicates, it is difficult to assess the reproducibility and significance of these findings.

Challenges and considerations. The most important limitation of PrimerID the inability to correct for errors during reverse transcription. Mutation rates in reverse transcriptases are similar

Table 7. Summary statistics from PrimerID experiment. Consensus threshold refers to the number of reads containing a given PrimerID required to build a reliable read family and is empirically determined during analysis. Mutation frequencies and rates are reported as mutations per nucleotide sequenced and mutations per nucleotide per infectious cycle, respectively.

Virus	Number of raw reads	Consensus threshold	Total consensus sequences	Template recovery	Mutation frequency	Mutation rate
WT-MHV	2,129,516	7	40,837	40.8%	3.34×10^{-5}	5.33×10^{-6}
ExoN(-) P3	2,617,306	8	43,087	43.1%	3.24×10^{-4}	2.02×10^{-5}

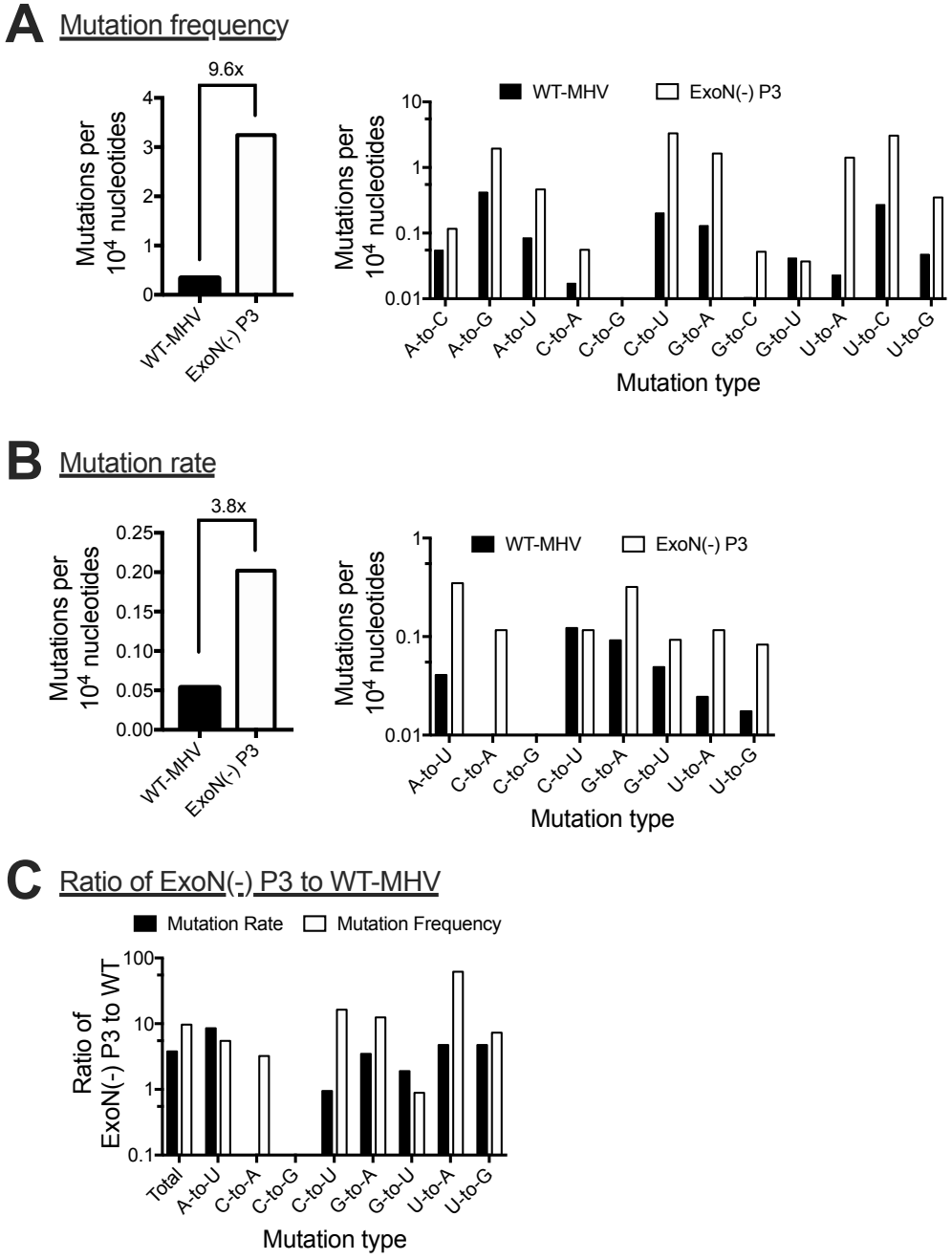


Figure 36. MHV mutation frequencies and rates by PrimerID sequencing. (*A*) Mutation frequencies were measured by comparing the total number of variants detected with the total number of nucleotides sequenced. (*B*) Mutation rates were calculated as the frequency of nonsense mutations, which should not propagate through multiple replication cycles. (*C*) The ratio of ExoN(-) P3 to WT-MHV was calculated for individual mutational classes using both mutation rates and mutation frequencies.

to those of RNA-dependent RNA polymerases, representing a theoretical lower-limit to the detectable mutation rates (Sanjuán et al., 2010). However, the mutational spectrum of reverse transcriptases is well-described and biased towards G-to-A substitutions (Cuevas et al., 2015; Holtz and Mansky, 2013; Pauly et al., 2017b). In the above experiment, G-to-A substitutions were not over-represented relative to other mutation types (Figure 36), suggesting either that G-to-A substitutions are relatively low in coronaviruses at baseline or that the sequencing depth was too shallow. In addition, the data above demonstrates that selection can confound estimates of mutation frequency (Figure 36C), particularly within mutational classes, so mutation rates are preferred for these comparisons. Limiting the number of replication cycles should reduce the influence of selection, and increasing coverage depth should improve estimates of mutation rates by detecting very low-frequency variants. Illumina sequencing chips have a finite capacity, so there is a trade-off between sample number and data obtained: the more samples in a given run, the less data obtained per sample. Importantly, each library must be diluted with fragmented bacterial DNA to ensure sufficient sequence diversity on the Illumina chip, which is essential for focusing the optical detector. We added 50% bacterial DNA to each sequencing run discussed in this dissertation. To ensure sufficient coverage, we have multiplexed no more than four samples. Subsequent experiments to determine the optimal number of samples and percentage of bacterial DNA would be warranted to minimize costs. During data interpretation, it is important to bear in mind that PrimerID examines only a small segment of a genome. Mutation rates of retro- and flaviviruses vary depending upon RNA secondary structures and local sequence contexts (Geller et al., 2015; 2016; Pathak and Temin, 1992), so PrimerID results using MHV20814 may not be generalizable to the whole genome. It would be worthwhile to sequence PrimerID-tagged libraries derived from multiple genome regions. However, the trade-off between sample number and coverage depth should be considered carefully in designing such an experiment.

Development of a quantitative-PCR-based competitive fitness assay for MHV

Although mutation rates determine the genetic diversity of a viral population, mutational fitness effects define the peaks, ridges, and valleys of genotype-fitness landscapes. As with replication fidelity, viral fitness can be measured in numerous ways. The simplest method is to measure a particular phenotype (e.g. growth rate) for multiple viruses in isolation and convert them to relative fitness values (for example, (Peris et al., 2010)). While informative, these studies often mask subtle (and not-so-subtle) fitness differences between variants that, in a mixed population, would affect evolutionary outcomes. For instance, recombinant MHV chimeras containing HCoV proteases replicate identically to wild-type MHV in isolation. However, in a mixed culture with wild-type MHV, the chimeras rapidly disappear from the population (Stobart et al., 2013). Similarly, altered fidelity variants often replicate identically to their wild-type counterparts but are rapidly outcompeted in mixed cultures (Coffey et al., 2011; Dapp et al., 2013; Fitzsimmons et al., 2018). Thus, direct competition experiments are preferred for viral fitness measurements.

The key challenge for competition assays is to discriminate between variants in a mixed infection. Without genetic modifications, dideoxy or deep sequencing methods can be used to detect the frequency of individual variants within a population (Acevedo et al., 2014; Agostini et al., 2018). However, chromatographic traces from Sanger sequencing experiments are imprecise and have a narrow dynamic range, and deep sequencing experiments are expensive and may require development of new library preparation methods. Other experimentalists distinguish variants based on genetic tags that can be identified without sequencing, including restriction sites or monoclonal antibody-resistance mutations (Furió et al., 2005; Stobart et al., 2013). More

recently, the field has started using real-time quantitative polymerase chain reaction (qPCR) to distinguish competitors (Bordería et al., 2015; Carrasco et al., 2007; Lauring et al., 2012; Moratorio et al., 2017). qPCR is a highly sensitive technique for monitoring the amplification of DNA molecules in real-time, rather than at the end, as in conventional PCR. qPCR assays have wider dynamic ranges than other methods because they allow sensitive and specific detection of variants at very low frequencies within a sample. Additionally, the high-throughput of qPCR-based assays can reveal the topography of large landscapes (Moratorio et al., 2017). These advantages have made qPCR-based competition assays the field standard for several years.

Quantitative PCR relies on fluorescent labels to monitor the amplification of double-stranded DNA over multiple cycles of PCR. Two common methods for detection of PCR products in real-time are Taqman probes and intercalating dyes. Taqman probes are sequence-specific oligonucleotides dual-labeled with a fluorescent reporter and a quencher. At rest, the quencher prevents detection of fluorescence. During PCR, the polymerase digests probes bound to DNA templates, breaking the reporter-quencher proximity and allowing detection of fluorescence. Taqman probes are highly specific, so detecting distinct variants requires two specific probes (selective detection). The other method for DNA detection in qPCR is through intercalating dyes, which fluoresce only when bound to double-stranded DNA. The most commonly used dye is SYBR green. SYBR assays are simpler to design than Taqman assays because they require primers but no probe. SYBR green will interact with any double-stranded DNA present, so assays must only amplify their intended target (selective amplification).

Quantitative PCR assays based on selective detection of competitors are prone to amplification bias. At the start of this dissertation research, Clint Smith and Nicole Sexton had

generated reference viruses containing 10 synonymous substitutions in a 22-nucleotide region of nsp2 ("tagged RNA", Figure 37A). I designed a Taqman assay with a probe complementary to the tag, flanked by PCR primers (Figure 37B; "reference assay"). Importantly, this assay amplifies both tagged and untagged genomes but only emits fluorescence when bound to tagged RNA. I also designed a Taqman assay upstream that detects all genomes in the sample, serving as the denominator for calculating relative frequency (Figure 37B, "competitor assay"). Both assays were highly efficient and specific for their targets (Figure 37C). However, I worried that because the reference assay amplifies all genomes, regardless of tag, it might susceptible to amplification bias. Amplification bias can occur when two targets are present in starkly different abundance. The more abundant target can continue to amplify and consume reagents, thereby hindering amplification of the scarce target (Figure 38). Using mixtures of tagged and untagged RNA standards, I measured assay performance over a 10-$\log_{10}$ range of known input ratios (10^5:1 to 1:10^5 untagged:tagged). Beyond ratios of 100:1 or 1:100, the experimentally-derived ratios diverge drastically from the expected results (Figure 39), indicating that assays based on selective detection can be dramatically skewed by amplification bias.

qPCR assays based on selective amplification of competitors do not suffer from amplification bias. In single-plex assays (that is, one assay per tube), amplification bias can only occur when both targets are amplified by the same set of primers. To circumvent this limitation, I designed two assays that selectively amplify their targets. My reference virus contains seven consecutive synonymous substitutions within nsp2 (nucleotides 1301-1307; Figure 40A). The reverse primer (also the reverse transcription primer) is the same for both assays, and the forward primers discriminate the tagged and untagged RNA (Figure 40B). I chose to use SYBR chemistry to avoid potential interactions between primers and a Taqman

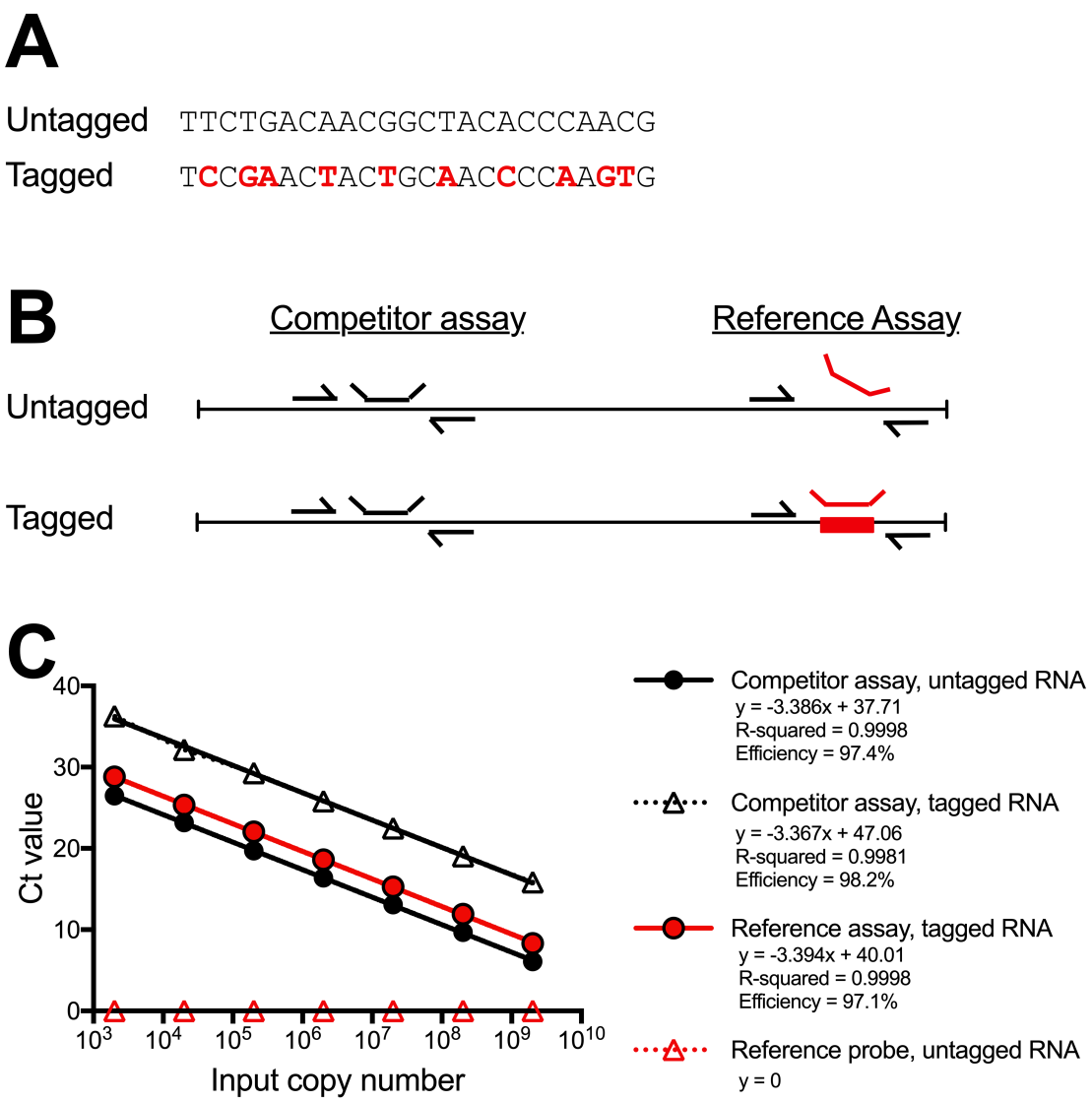


Figure 37. Taqman probes selectively detect marked reference genomes. (*A*) The genome is tagged with 10 synonymous substitutions within nsp2. (*B*) Schematic of competitor and reference assays. The competitor assay detects all genomes in a sample, and the reference assay detects only the tagged RNA. (*C*) Both assays are highly specific and efficient.

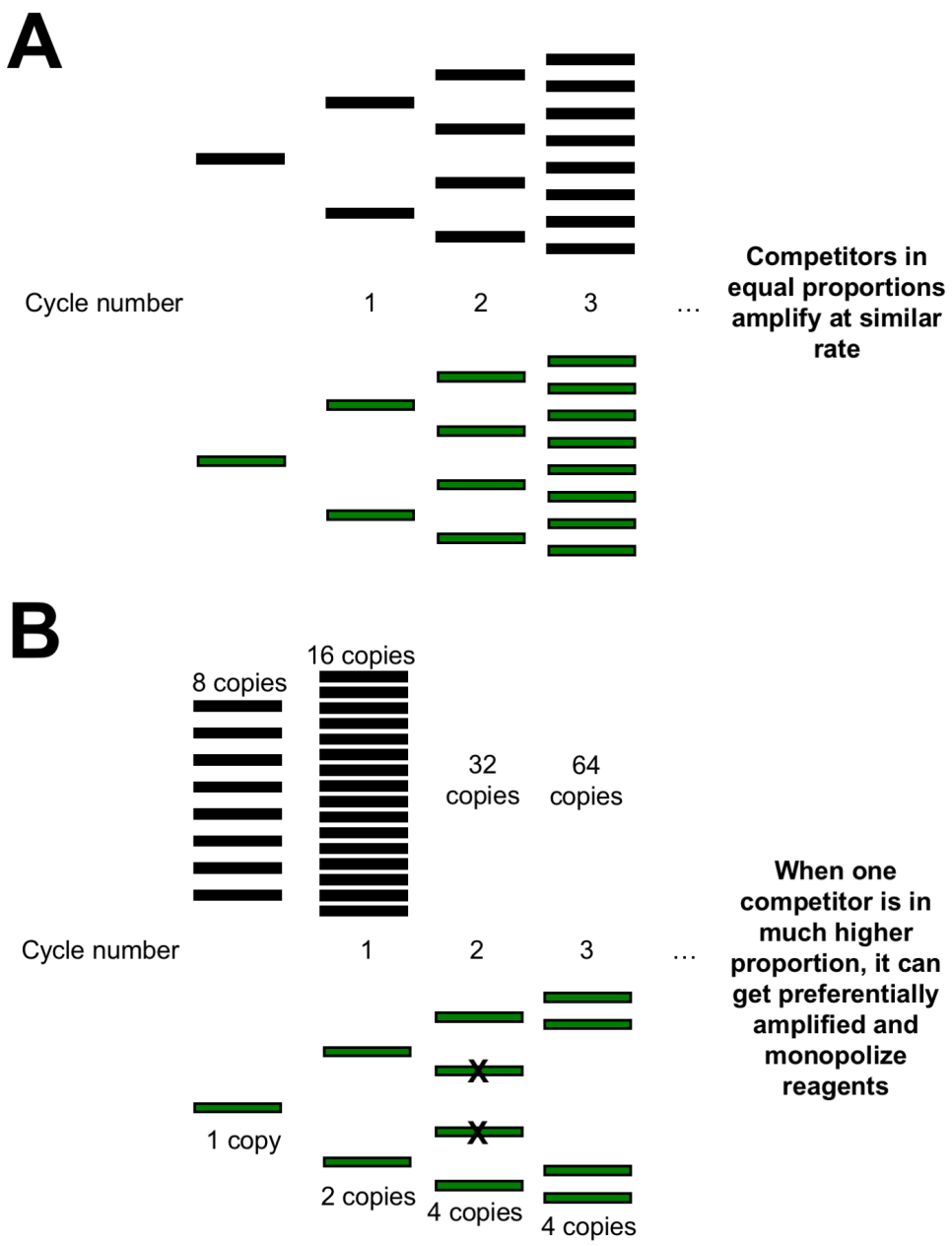


Figure 38. Amplification bias can skew the results of qPCR. (*A*) Competitors at equal proportions in the sample amplify at similar rates. (*B*) The more abundant competitor can be preferentially amplified and monopolize reagents. X indicates DNA copies not amplified in a given round of replication.

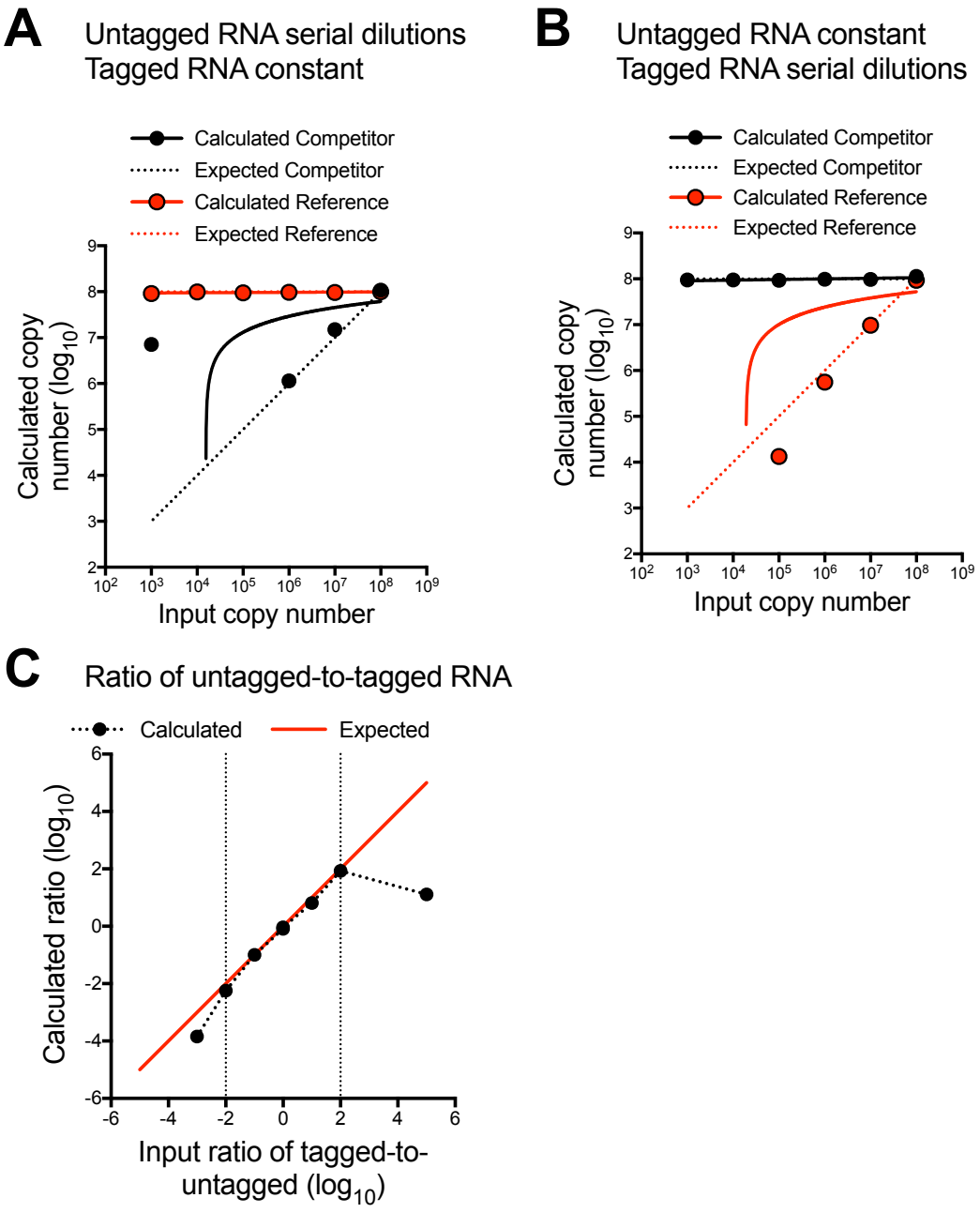


Figure 39. A qPCR assay based on selective detection is prone to amplification bias. (*A*) 1 x 10^8 copies of tagged RNA were added to serially-diluted samples of untagged RNA. RT-qPCR was performed with both competitor and reference assays, and copy numbers of each RNA were calculated and plotted. (*B)* Performed as in (*A*), except that 1 x 10^8 copies of untagged RNA were added to serially-diluted samples of tagged RNA. (*C*) Data from (*A*) and (*B*) were converted to ratios of tagged-to-untagged RNA.

probe. Both assays are specific for their intended target across a 5-$\log_{10}$ range of concentrations (Figure 40C) and show no evidence of amplification bias across an 8-$\log_{10}$ range of ratios (Figure 41). Thus, I chose to use these SYBR-based assays (and reference viruses) for my competitive fitness experiments.

Experimental design. All experimental viruses are co-cultured with the genetically tagged reference virus at a known multiplicity of infection (MOI) and passaged four times (Figure 42). The ratio of competitor-to-reference is measured at each passage, and the change in relative frequency of competitor and references over time reflects their relative fitness. Note that this yields indirect comparisons of fitness between experimental viruses. That is, instead of direct competition between virus A and virus B, the immediate comparisons will be "fitness of virus A relative to reference" vs "fitness of virus B relative to reference."

Example competition assay. To determine whether the variant ExoN AD is more fit than ExoN(-) P3, I competed each virus against the ExoN(-) reference virus (Chapter 4, Figure 19). I passaged each lineage at a total MOI of 0.1 PFU/cell four times. The relative frequency of each competitor decreased over passage (Figure 43A), suggesting either that the genetic tag is not selectively neutral, or more likely, that the stock reference virus contains additional mutations. The results were consistent across all three individual lineages for each virus. I performed linear regressions of each lineage (Figure 43B) and converted the slopes to relative fitness (relative fitness = 10^{slope}) (Figure 43C). To account for potential fitness-altering mutations in the ExoN(-) reference stock, I then normalized all values to the mean relative fitness of ExoN(-) P3 (Figure 43D). This step reduces the influence of stock effects for the reference viruses and permits direct comparison of relative fitness values from different experiments.

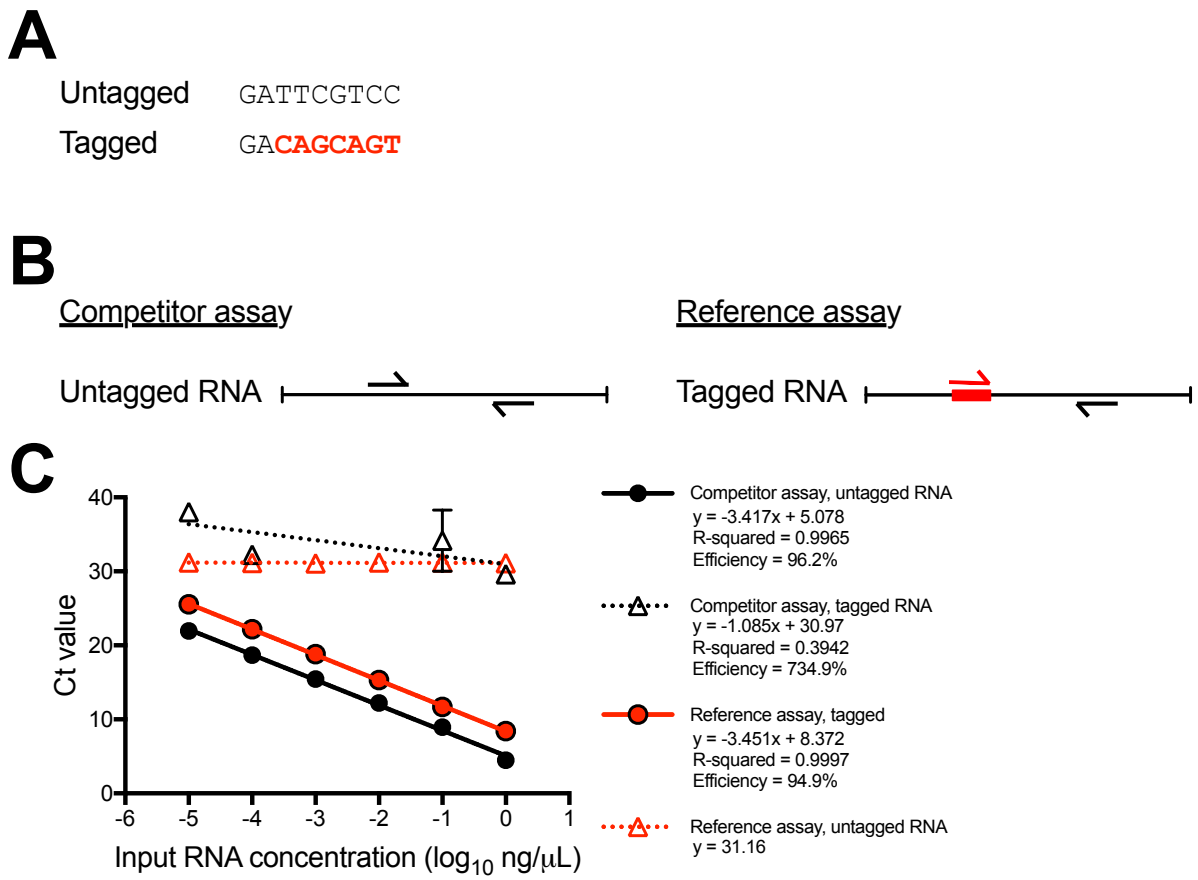


Figure 40. SYBR green-based assays selectively amplify their intended targets with high efficiency. (*A*) The genome is tagged with 7 consecutive synonymous substitutions within nsp2. (*B*) Schematic of competitor and reference assays. Both assays use the same reverse primer, and forward primers discriminate tagged and untagged RNA. (*C*) Both assays are highly specific and efficient. All data points in open triangles (assays against their nonspecific target) were within 1 Ct of the no-template control.

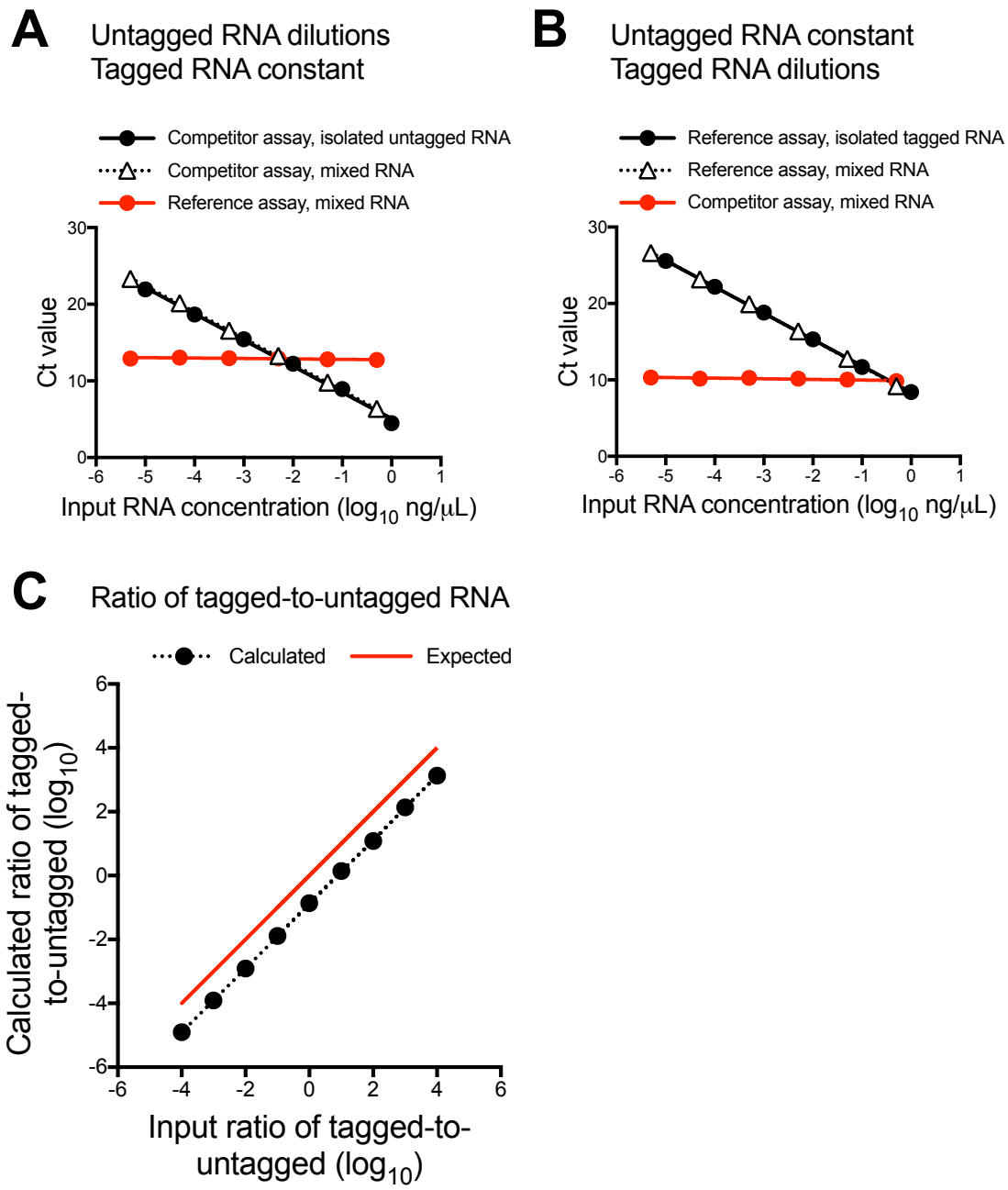


Figure 41. A qPCR assay based on selective amplification does not suffer from amplification bias. (*A*) 0.05 nanograms of tagged RNA were added to serially-diluted samples of untagged RNA. RT-qPCR was performed with both competitor and reference assays. (*B*) Performed as in (*A*), except that 0.05 nanograms of untagged RNA were added to serially-diluted samples of tagged RNA. (*C*) Data from (*A*) and (*B*) were converted to ratios of tagged-to-untagged RNA.

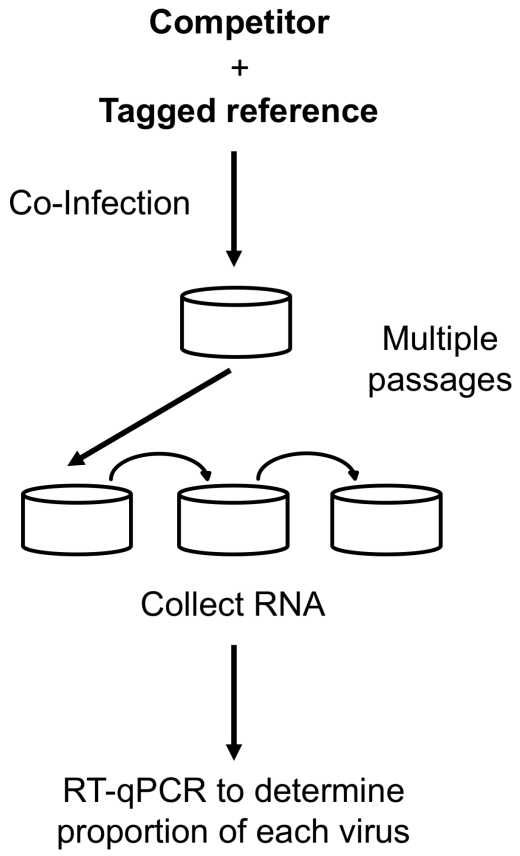


Figure 42. Experimental schematic of competitive fitness assay.

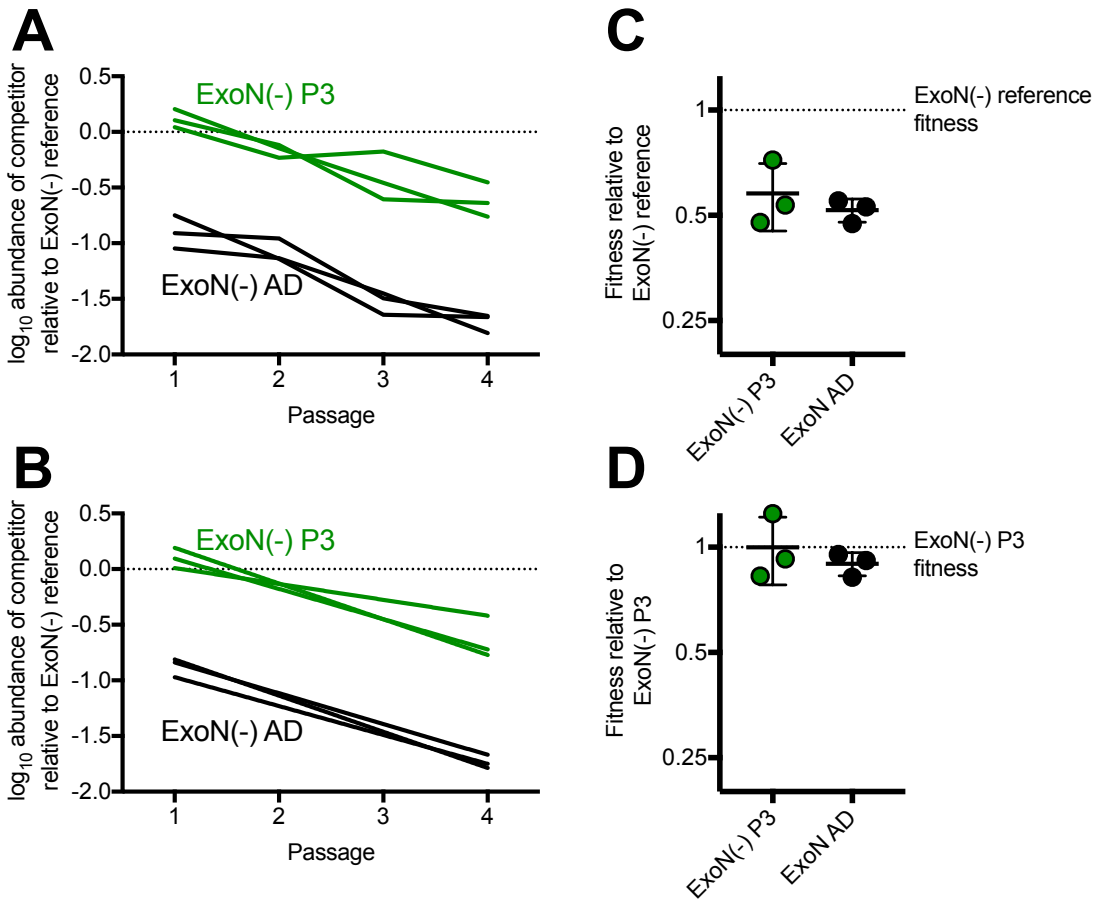


Figure 43. Example data from competitive fitness experiment. ExoN(-) P3 and ExoN AD were competed against an ExoN(-) reference containing 7 synonymous substitutions and passaged four times. (*A*) Ratio of competitor-to-reference for individual lineages at each passage. (*B*) Linear regressions of each lineage. (*C*) Fitness of each competitor relative to the reference were plotted, and then (*D*) normalized to mean of ExoN(-) P3. Dotted lines in panels (*A*) and (*B*) represent 1:1 ratio.

Challenges and considerations. The outcomes of viral evolution depend on both viral fitness and the random stochastic effects of genetic drift. To minimize the effects of genetic drift, experiments should be performed with large population sizes (minimum total MOI of 0.1 PFU per cell; MOI = 0.05 PFU/cell each for competitor and reference). Synchronizing infections can also reduce genetic drift by ensuring that virions enter cells simultaneously (Voronin et al., 2009). Additionally, because viral fitness varies with experimental conditions, technical consistency is essential for reproducibility. I take the following measures to ensure consistent conditions across passages and across experiments: (1) I count cells before seeding plates, delivering 1 x 10^5 cells per well; (2) I infect cells exactly 24 hours after plating; (3) I incubate virus and cells for 30 minutes on ice to synchronize infections; (4) I titer all samples after each passage to maintain the same input MOI. I also perform a control in every experiment to account for variability in stocks of reference viruses (e.g. compete ExoN(-) P3 versus the ExoN(-) reference). I then normalize all relative fitness values to the control. Finally, because this assay measures RNA content, viruses that generate large quantities of defective particles can yield artificially high fitness values, so this assay should be used only with viruses of similar specific infectivity.

Discussion

With their multi-subunit replicase complex and proofreading exoribonuclease, coronaviruses are a unique RNA viral system for studying the relationships between replication fidelity, fitness, and evolution. However, the techniques for studying coronavirus evolution are less extensive than those available for simpler RNA viruses. In this chapter, I extend the toolkit for coronavirus evolution with new methods to measure murine hepatitis virus replication fidelity and fitness.

PrimerID sequencing circumvents the interpretive challenges of nucleoside analogue sensitivities and reduces artifacts that confound conventional deep sequencing approaches, yielding sensitive and accurate measurements of mutation frequencies and rates. PrimerID will allow us to discriminate subtle fidelity phenotypes, as well as identify differences in mutational biases that could change the accessible sequence landscapes. The competitive fitness assay described in this chapter can reveal the contours of these sequence landscapes, allowing us to trace, and potentially predict, coronavirus evolutionary trajectories. By varying experimental conditions and selective pressures, we can examine evolutionary pathways leading to antiviral drug resistance or escape from immune responses. Further, fitness experiments can reveal interactive networks between coronaviral proteins, as discussed in Chapter 4. In combination, PrimerID and competitive fitness assays will allow dissection of the fundamental relationships between coronavirus replication fidelity and fitness, informing the design of countermeasures against emerging and endemic coronaviruses.

CHAPTER 6: MATERIALS AND METHODS

Cell culture

Delayed brain tumor (DBT-9) cells (Chen and Baric, 1996) and baby hamster kidney 21 cells expressing the MHV receptor (BHK-R) (Yount et al., 2002) were maintained at 37°C in Dulbecco's modified Eagle Medium (DMEM, Gibco) supplemented with 10% serum (HyClone FetalClone II, GE Healthcare or Fetal Bovine Serum, Invitrogen), 100 U/mL penicillin and streptomycin (Gibco), and 0.25 μM amphotericin B (Corning). BHK-R cells were further supplemented with 0.8 mg/mL G418 selection antibiotic (Gibco). The infectious clone of the murine hepatitis virus strain A59 (MHV-A59; GenBank accession number AY910861) was used as the template for all recombinant viruses in this dissertation (Yount et al., 2002).

Determination of viral titer by plaque assay

Virus samples were serially diluted and inoculated on subconfluent DBT-9 cell monolayers in either 6- or 12-well format. Cells were overlaid with a 1% agar in DMEM and incubated overnight at 37°C. Plates were fixed with 4% formaldehyde and agar plugs were removed. The number of plaques per well was counted by hand and used to calculate titer. For plaque assays of viruses containing the spike protein from MHV-ExoN(-) P250, which does not form syncytia, plaques were visualized with neutral red (Sigma #N6264, diluted 1:10 in PBS containing calcium and magnesium). Neutral red was added 24 hours after plating and incubated for an additional 3-8 hours before formaldehyde fixation.

Recombinant virus recovery

The MHV-A59 reverse genetics system divides the genome into 7 fragments, A through G, maintained on separate plasmids. Plasmids are digested using pairs of restriction enzymes yielding unique termini, and digested fragments are purified by gel extraction. Cut fragments are pooled in equimolar amounts (total DNA content 1.5 μg) and ligated using T4 DNA ligase (NEB) to generate a linear, double-stranded product containing a T7 promoter. Full-length RNA genomes are *in vitro* transcribed using mMessage mMachine T7 Transcription kit (Invitrogen) and electroporated into BHK-R cells. Cells are monitored for signs of cytopathic effect, and recombinants are harvested by freeze-thaw, aliquoted, and sequenced as described below.

Molecular cloning

With few exceptions noted below, site-directed mutatgenesis in MHV genome fragments was performed using "round the horn" PCR (originally described in (Ho et al., 1989)). Briefly, adjacent primers containing the mutation of interest were 5′-phosphorylated using T4 polynucleotide kinase (NEB, M0201S) using the buffer from the T4 DNA ligase, which contains ATP (M0202S). PCR was performed on a plasmid template using the Q5 High-fidelity 2x Master Mix (NEB, M0492L), with primers at final concentration of 500nM. The linear amplification product was purified using the Promega Wizard SV Gel and PCR Clean-up System (Promega Corporation, A9282), and 4 μL was ligated at 16°C overnight with the T4 DNA ligase (NEB M0202S). After transformation into chemically-competent Top10 *E. coli* (lab-derived) and expansion in liquid culture, the MHV segment of each plasmid was sequenced. For larger regions (e.g. nsp12 and nsp14 swapped viruses, Chapter 2), the region of interest was amplified by RT-PCR from viral RNA and ligated into the desired vector by Gibson Assembly (Gibson et

al., 2009) using the InFusion HD Cloning kit (Takara Bio USA, Inc, Mountain View, CA). Several fragments intended for the ExoN(-) P160 infectious clone were synthesized commercially by GenScript, Inc.

Generation of experimental stocks

Experimental stocks were generated by infecting a subconfluent 150 cm^2 flask of DBT-9 cells at MOI of 0.01 PFU/cell. Flasks were frozen at -80°C when monolayers were fully involved, approximately 20-28 hours post-infection depending on the variant. After thawing, the supernatant was clarified by centrifugation at 4,000 x g (Sorvall RC 3B Plus; HA-6000A rotor) for 10 min at 4°C. For intermediate revertants, stocks were generated in serum-free DMEM and processed as above before being concentrated roughly 10-fold by centrifugation at 4,000 x g using Amicon Ultra-15 Centrifugal Filter Units, 100kDa (EMD Millipore, UFC910008). The virus titer of each stock was determined by plaque assay using DBT-9 cells as described above.

Plaque purification of viral populations

DBT-9 cells were infected with serial dilutions of virus and overlaid with 1% agar in DMEM. Single plaques were isolated with glass Pasteur pipettes, resuspended in PBS containing calcium and magnesium, and inoculated onto fresh DBT-9s. This process was completed 3 times before experimental stocks were generated, as above.

Sequencing of virus stocks

RNA was purified from the harvested TRIzol samples according to the manufacturer's protocol and reverse transcribed (RT) using SuperScript III (Invitrogen) as described previously (Eckerle

et al., 2007). Full-genome di-deoxy sequencing was performed for both WT-MHV P250 and MHV-ExoN(-) P250 using 12 overlapping amplicons approximately three kilobases in length. Two microliters of RT product were used for each PCR reaction (Smith et al., 2013). Di-deoxy sequencing was performed by Genhunter Corporation (Nashville, TN) and GENEWIZ (South Plainfield, NJ). Sequence analysis was performed using MacVector version 14 (MacVector, Inc.; Apex, North Carolina) using the MHV-A59 infectious clone reference genome (GenBank accession no. AY910861).

Experimental evolution of viruses

The infectious cDNA clone for MHV-A59 and the recovery of MHV-ExoN(-) are described previously (Eckerle et al., 2007; Yount et al., 2002). Long-term passage (P250; Chapter 2) was initiated by infecting subconfluent monolayers of DBT-9 cells in 25 cm^2 flasks with either wild-type MHV-A59 or MHV-ExoN(-) at a MOI of approximately 0.1 PFU/cell. One lineage of each virus was passaged for a total of 250 passages (P250). Supernatant was harvested at each passage and stored at -80°C. Total RNA was harvested for most passages using 1 ml of TRIzol Reagent (Ambion) per 25 cm^2 flask and stored at -80°C. Intermediate revertants of MHV-ExoN(-) (Chapter 5) were passaged 10 times on subconfluent DBT-9 cell monolayers in 24-well plates at an estimated MOI of either 0.01 or 0.5 PFU/cell. Supernatants were harvested at 24 and 20 hours post-infection for MOI = 0.01 and 0.5 PFU/cell, respectively.

Replication kinetics

Viral replication kinetics in DBT-9 cells were determined at indicated MOIs as described previously (Smith et al., 2015). Experiments in Chapter 5 were synchronized by adding virus to cells on ice and incubating for 30 minutes before transferring to the 37°C incubator. Supernatant

(300 μL) was harvested at the indicated time points, and the virus titer was determined by plaque assay.

Genomic RNA accumulation kinetics

The accumulation of genomic RNA at MOI = 1 PFU/cell was measured by two-step real-time quantitative RT-PCR (RT-qPCR). RNA was harvested from infected monolayers at the indicated times using TRIzol (Invitrogen). RNA was reverse-transcribed using SuperScript III and random hexamers (Invitrogen), and cDNA derived from intracellular viral RNA was measured using primers directed to nsp10. Values were normalized to levels of the endogenous control glyceraldehyde-3-phosphate dehydrogenase (GAPDH). No mutations within the primer binding sites emerged in either P250 population. The primers and amplification conditions are the same as reported previously (Smith et al., 2013), except that the RT product was diluted 1:10 prior to use. Samples were plated in technical duplicate to minimize well-to-well variation. Data are presented as $2^{-\Delta Ct}$, where ΔCT denotes $Ct_{nsp10} - Ct_{GAPDH}$.

Determination of specific infectivity

Subconfluent monolayers of DBT-9 cells in 24-well plates were infected with the indicated virus at MOI = 1 PFU/cell, and supernatant was harvested at 12 hours post-infection. The levels of genomic RNA in supernatant were measured using one-step real-time quantitative RT-PCR (RT-qPCR) on TRIzol-extracted RNA as described previously (Sexton et al., 2016). Briefly, genomic RNA was detected with a 5' 6-carboxyfluorescein (FAM) and 3' black hole quencher 1 (BHQ-1) labeled probe targeting nsp2 (Biosearch Technologies, Petaluma, CA), and RNA copy number was calculated by reference to an RNA standard derived from the MHV A fragment. Samples

were plated in technical duplicate to minimize well-to-well variation. Titers were determined by plaque assay in DBT-9 cells, and specific infectivity was calculated as PFU per supernatant genomic RNA copy.

Nucleoside and base analogue sensitivity assays

Nucleoside and base analogue sensitivity assays were performed as described (Graepel et al., 2017; Smith et al., 2013) 5-azacytidine (AZC), 5-fluorouracil (5-FU), and ribavirin (RBV) were purchased from Sigma (product numbers A2385, F6627, and R9644, respectively), and stock solutions were prepared in dimethyl sulfoxide (DMSO). 2'-C-methyladenosine (CMeA) was received from Gilead Sciences, Inc (Foster City, CA). Sensitivity assays were performed in 24-well plates at MOI = 0.01 PFU/cell. Cells were incubated with drug for 30 minutes prior to infection. Supernatants were harvested at 24 hours post-infection, and titers were determined by plaque assay.

Measurement of mutation frequency using PrimerID

A custom R script, generously provided by Adam Lauring (https://github.com/lauringlab/NGS_mutation_rate_assay) was used to identify a 452-base region in ORF1ab of MHV-A59 (nts 20814-21265, covering portions of nsp15 and nsp16) with the highest concentration of nonsense mutational targets. Supernatant RNA genomes were isolated from DBT-9 monolayers infected with wild-type, ExoN(-), or ExoN(-) nsp12-250 (MOI of 0.01 PFU/cell) using QIAamp viral RNA mini kit. Supernatant genomes were quantified by one-step RT-qPCR by reference to an RNA standard derived from the MHV A fragment, as described previously (Sexton et al., 2016). PrimerID libraries were prepared from 5 x 10^4 genome copies using SuperScript III (Invitrogen)

and reverse transcription primer MHV20184-cDNA (5′-GTGACTGGAGTTCAGACGTGTGC-TCTTCCGATCTNNNNNNNNNNCCAATGTTCTTAGTAAGAGGGTCGTACATA-3′), which contains N^{10} degenerate barcode sequences. Reaction mixtures containing only RNA template, cDNA primer, and dNTPs were incubated 65°C for 5 minutes, followed by addition of remaining reagents on ice. Reverse transcription proceeded at 50°C and 55°C for one hour each, followed by 15 minute incubation at 70°C to inactivate enzyme. Reactions were treated with RNaseH for 20 minutes at 37°C and cleaned up with Agencourt RNAClean XP beads (1.2:1 bead:cDNA ratio). The resulting PrimerID-tagged cDNAs were amplified by PCR using KAPA2G Robust Hotstart (Kapa Biosystems, Woburn, MA) and forward primer MHV20184-PCR1 (5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGAAGGTCATGACT-TTCTATCCTCGTTTGC-3′) and reverse primer PID-PCR1-R (5′-GTGACTGGAGTTCAGAC-GTGTGCTC-3′). The PCR cycling protocol was initial denaturation at 95°C for 1 minute followed by 25 cycles of 95°C for 15 seconds, 58°C for 1 minute, and 72°C for 30 seconds, with a final extension at 72°C for 3 minutes. PCR1 products were purified using Agencourt AMPure XP beads (1.2:1 bead:PCR1 ratio). A second round of PCR was performed to add indexed Illumina adaptors. The forward primer was 5′-AATGATACGGCGACCACCGAGATCTACAC-[index]ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′ and the reverse primer was 5′-CAAGCAGAAGACGGCATACGAGAT[index]GTGACTGGAGTTCAGACGTGTGCTCTTC CGATC-3′. PCR was performed using Kapa HiFi PCR Hotstart (Kapa Biosystems, Woburn, MA) with an initial denaturation of 95°C for 2 minutes followed by 25 cycles of 98°C for 20 seconds, 63°C for 15 seconds, and 72°C for 30 seconds, with a final extension at 72°C for 3 minutes. PCR2 products were gel-extracted using the MinElute Gel Extraction Kit (Qiagen).

Libraries were quantified by Qubit and pooled at 5nM each. We used 250-bp paired-end multiplex Illumina MiSeq (San Diego, CA) to sequence the constructed libraries. The pooled libraries made up half of the DNA input for the sequencing run, with the remaining DNA composed of bacterial genome library. Sequencing was performed by the Vanderbilt Technologies for Advanced Genomics core facility (VANTAGE).

Phyre^2-modeling of MHV-nsp14

The MHV nsp14 structure was modeled with the Phyre^2 online program (Kelley et al., 2015) using nsp14 residues 3-519, corresponding to residues 6056-6573 of the ORF1ab polyprotein. The model was analyzed using the Pymol Molecular Graphics System, Version (Schrödinger, LLC).

Taqman-based competitive fitness assays

Experimental viruses were competed with an MHV-ExoN(-) virus harboring 10 silent mutations in the probe-binding region within nsp2. Subconfluent DBT-9 monolayers in 24-well plates were coinfected at a total MOI of 0.01 PFU/cell with competitor and reference viruses at a 1:1 ratio and passaged 4 times. For each passage, supernatants were harvested at 24 hours. RNA was extracted from 100µL of supernatant using 900µL of TRIzol reagent and PureLink RNA Mini Kit columns (Thermo Scientific, Waltham, MA), and 150µL was used to infect fresh cells in a 24-well plate (total MOI estimated at 1 PFU/cell). The proportion of each virus was determined by real-time RT-qPCR from the infection supernatant using two Taqman probes with different fluorescent dyes in separate reactions. Competitor viruses were detected with the same probe used in specific infectivity analyses (Eckerle et al., 2007). Reference viruses were detected by a

probe targeting the same region but with 10 silent mutations (5'-TCCGAACTACTGCAACCC-CAAGTG-3') and labeled with 5' Quasar 670 and 3' black hole quencher 2 (BHQ-2) (Biosearch Technologies, Petaluma, CA). RNA copy number was calculated by reference to an RNA standard generated by *in vitro* transcription of the corresponding MHV A fragment, and relative RNA abundance was calculated as the ratio of competitor genomes to reference genomes.

SYBR-based competitive fitness assays

ExoN(-) and ExoN(+) reference viruses were marked with 7 consecutive silent mutations within nsp2 (wild-type: 1301-TTCGTCC-1307; silent: 1301-CAGCAGC-1307) by round the horn PCR, as described above. Competitions were performed in triplicate on DBT-9 cells in 12-well plates, plated at a density of 1 x 10^5 cells per well 24 hours prior to infection. Cells were infected at a total MOI of 0.1 PFU/cell (MOI = 0.05 PFU/cell each for competitor and reference virus). Supernatants were harvested 15 and 16 hours post-infection for experiments with ExoN(+) reference and ExoN(-) reference, respectively, and passaged 4 times. Samples were titered between all passages to maintain total MOI of 0.1 PFU/cell. RNA was extracted from 70µL of supernatant using QIAamp 96 virus QIAcube HT kit on the QIAcube HT System (Qiagen). Each RNA sample was analyzed by one-step RT-qPCR with two SYBR Green assays. Reference viruses were detected with forward primer SS-qPCR-Sil-F (5′-CTATGCTGTATACGGACAGC AGT-3′; 200nM final) and reverse primer SS-qPCR-R2 (5′-GGTGTCACCACAACAATCCAC-3′, 200nM final). Competitors were detected with forward primer SS-qPCR-WT-F (5′-CTATG-CTGTATACGGATTCGTCC-3′, 450nM final) and reverse primer SS-qPCR-R2 (5′-GGTGTC-ACCACAACAATCCAC-3′, 450nM final). RNA samples were diluted 1:100 prior to RT-qPCR with *Power* SYBR Green RNA-to-Ct 1-step kit (Applied Biosystems) according to the

manufacturer's protocol. Duplicate wells were averaged, and values were excluded from subsequent analysis if the duplicate wells differed by > 0.5 Ct. The relative abundance of competitor and reference were determined by subtracting Ct thresholds ($\Delta Ct_{competitor} = Ct_{competitor} - Ct_{reference}$) and converted to reflect the fold-change in ratio ($\Delta ratio = 2^{-\Delta Ct\ competitor}$). The $\log_{10}\Delta ratio$ was plotted against passage number, and the change in $\log_{10}\Delta ratio$ (i.e. slope of linear regression) is the relative fitness. Note that regressions were fit only through P1-P4, as slight deviations in 1:1 ratio in the input (P0) can skew the slope.

Statistical analysis

GraphPad Prism 6 (La Jolla, CA) was used to perform statistical tests. Only the comparisons shown [e.g. ns or asterisk(s)] within the figure or legend were performed. In many cases the data were normalized to untreated controls. This was performed using GraphPad Prism 6. The number of replicate samples is denoted within each figure legend.

Accession numbers

Full-length genome sequences for WT-MHV P250, MHV-ExoN(-) P250 can be downloaded from the GenBank (accession numbers MF618252, and MF618253, respectively).

CHAPTER 7: SUMMARY AND FUTURE DIRECTIONS

Introduction

Coronaviruses cause human illnesses ranging from the common cold to severe and lethal disease, and they are potent drivers of economic losses in commercial animals (Annamalai et al., 2015; Cavanagh, 2007; de Wit et al., 2016; Saif, 2010). Since 2002, two coronaviruses (SARS-CoV and MERS-CoV) have emerged as zoonoses with pandemic potential, and close relatives of SARS- and MERS-CoV circulate in animal populations around the globe (Drosten et al., 2003; Ksiazek et al., 2003; Li et al., 2005; Menachery et al., 2015; 2016; Zaki et al., 2012). With no approved coronavirus vaccines or antivirals available, our ability to respond rapidly to the next highly pathogenic coronavirus depends upon intensive study of conserved coronavirus genetic elements and proteins. Coronaviruses have considerably larger genomes than other RNA viruses, encoding a correspondingly intricate replication machine distinguished by an exoribonuclease (ExoN) (Smith et al., 2014) (Figure 10). At the start of this dissertation research, the coronavirus ExoN had been revealed as the first proofreading ExoN encoded by an RNA virus, establishing a new paradigm for RNA virus replication fidelity. ExoN is highly conserved across all known coronaviruses and is critical for efficient viral replication, genome maintenance, and innate immune evasion (Case et al., 2017; Eckerle et al., 2007; Gorbalenya et al., 2006). Accordingly, disruption of ExoN activity yields mutator viruses with diminished fitness and virulence, suggesting that ExoN could be a viable target for coronavirus countermeasures (Graham et al., 2012; Menachery et al., 2018). In this dissertation, I have used experimental evolution and reverse genetics to examine the complex protein interactions involved in coronavirus replication and to understand how the unique proofreading function of coronaviruses affects their evolution. In this chapter, I summarize the key findings and implications of my dissertation research.

High-fidelity replication is important for coronavirus fitness

High RNA virus mutation rates are proposed to constrain genome size and complexity (Bradwell et al., 2013; Bull et al., 2007). According to this model, maintenance of large coronavirus genomes requires a commensurate reduction in mutation rates to prevent error catastrophe, suggesting that high-fidelity replication is an essential component of coronaviral fitness. Consistent with this hypothesis, coronaviruses are the first known RNA viruses to encode a proofreading ExoN, and disruption of proofreading activity impairs both replication fidelity and fitness (Eckerle et al., 2007; 2010; Graepel et al., 2017; Graham et al., 2012; Minskaia et al., 2006; Sexton et al., 2016). Despite this correlation, it has not been clear whether fitness defects in ExoN(-) coronaviruses derive specifically from altered fidelity or from some other function of nsp14-ExoN. Support for the proposed link between fidelity and fitness came during the course of this dissertation, when Nicole Sexton discovered that a high-fidelity RdRp variant (nsp12-V553I), engineered through homology modeling with picornaviral RdRps, increases the fitness of MHV-ExoN(-) (Sexton et al., 2016). The research in this dissertation extends her observations by demonstrating that MHV-ExoN(-) is under selective pressure for increased fidelity during experimental evolution (Chapter 2). Passaged MHV-ExoN(-) P250 displays near wild-type fidelity, as measured by nucleoside analogue sensitivities, and surpasses wild-type fitness (Figure 11, Figure 15, Figure 18). The high-fidelity phenotype of MHV-ExoN(-) P250 is partially mediated by nsp12-250, which reduces mutation frequencies of MHV-ExoN(-) by 1.7-fold, similar to known high-fidelity RdRps (Figure 18) (Smith et al., 2014). Further, nsp12-250 independently increased the competitive fitness of MHV-ExoN(-), directly correlating polymerase selectivity and viral fitness. I have not yet determined whether or how individual mutations in nsp12-250 (7 total nonsynonymous substitutions) increase replication fidelity, and I

cannot exclude the possibility that one or more were fixed by genetic drift. However, I favor the hypothesis that the nsp12-250 mutations form an interdependent network. Fidelity regulation in RdRps is a dynamic process driven by coordinated movements across the whole enzyme (Korboukh et al., 2014; Korneeva and Cameron, 2007; Moustafa et al., 2014). This delicate process is easily disrupted, so it is unlikely that individual nonsynonymous substitutions were fixed by chance. Indeed, MERS-CoV nsp12-RdRp is subject to strong purifying selection, suggesting that coronavirus RdRps are relatively intolerant to mutations (Forni et al., 2016b). Instead, I propose that at least a subset of the seven mutations in nsp12-250 cooperate to drive high-fidelity replication. In MHV-ExoN(-), nsp12-RdRp is clearly undergoing rapid adaptive evolution, which tends to enrich for epistatically interacting substitutions (Gong and Bloom, 2014). Similar networks of epistatic interactions affecting replication fidelity have been identified in the RdRps of influenza virus and the oral poliovirus vaccine strain (Liu et al., 2015; Pauly et al., 2017a). With mutations distributed across the fingers, palm, and thumb domains (Figure 17), biochemical or next-generation sequencing studies of nsp12-250 would reveal important insights into the long-range dynamics of fidelity regulation by RdRps.

Coronavirus replicase proteins are tightly co-evolved and cooperate to optimize replication kinetics and fidelity

The predicted coronavirus replicase-transcriptase complex (RTC) contains, at minimum, 7 nonstructural proteins that interact extensively to maintain efficient and accurate RNA synthesis (Figure 10) (Smith et al., 2014). Mutations in nsps7-14 are likely under strong purifying selection to preserve both their individual and collective functions, ensuring that RTC proteins evolve cooperatively (Forni et al., 2016b). In this dissertation, we observed very little evidence of adaptive evolution between wild-type replicase proteins (Figure 13, Appendix A.1). In

contrast, the selective pressures provoked by altered proofreading drove extensive evolution within the RTC (Figure 13, Appendix A.2). Recombinant viruses containing subsets of P160/P250 replicase proteins revealed complex evolutionary interactions between these mutations. For instance, nsp12-250 and nsp14-250 both independently increased fitness in MHV-ExoN(-), but their fitness effects were not additive (Figure 19). Instead, ExoN(-) nsp12/14-250 was less fit than ExoN(-) nsp12-250, suggesting that other replicase mutations support the interactions between the RdRp and ExoN. In addition, passage-associated phenotypes were lost when placed in a different genetic background. For example, inserting the high-fidelity nsp12-250 polymerase into a proofreading-capable virus did not result in an additive increase in fidelity (Figure 32). Instead, nsp12-250 interfered with normal proofreading, as measured by sensitivity to 5-fluorouracil, and was detrimental to overall replication and competitive fitness (Figure 32). This observation is consistent with Nicole Sexton's work on the likely high-fidelity mutant, nsp12-V553I (Sexton et al., 2016). In that case, nsp12-V553I did not modify proofreading in the wild-type background, but it did delay replication, suggesting that it still interfered with RTC assembly or function. These data suggest that nsp12 and nsp14-ExoN coevolve to optimize distinct properties of viral replication. I favor the hypothesis that the proofreading activity nsp14-ExoN, through interaction with nsp10, relaxes the selective pressure on polymerase nucleotide selectivity, allowing nsp12-RdRp to optimize other parameters, such as replication speed or processivity. Consistent with this hypothesis, the SARS-CoV nsp12-RdRp is lower fidelity than the dengue virus RdRp, a virus with a genome one-third of its size (Ferron et al., 2018). Under this model, there is a trade-off between replication speed and fidelity—the slower the polymerase proceeds, the more time it has to discriminate incoming nucleotides (Elena and Sanjuán, 2005; Fitzsimmons et al., 2018; Pandey et al., 1996; Regoes et al., 2013). Thus, increasing polymerase fidelity would disrupt the kinetic balance between

extension and excision during replication. In line with this expectation, both nsp12-V553I and nsp12-250 delay replication by the wild-type RTC (Figure 32) (Sexton et al., 2016). Additionally, one or more mutations in nsp12-250 could dysregulate physical interactions with wild-type nsp14. For example, if nsp14-ExoN(-) disrupts replication by binding to, but not excising, nucleotide mismatches, then mutations that disrupt the nsp12-nsp14 interaction could have been advantageous. Whether and how passage-associated mutations in nsp8, nsp9, and nsp13 fit into this model is an open question, but it is clear that they do not modify fidelity directly (Chapter 3). However, the recombinant viruses described in Chapter 3 will be useful tools for investigating replication complex assembly and function. The ability to test these hypotheses in replicating viruses or *in vitro* reconstituted biochemical systems provides a unique opportunity to map the coronavirus replication machinery.

nsp14-ExoN is more than just the first proofreading enzyme in an RNA virus

The experiments in this dissertation have focused on the proofreading activity of nsp14-ExoN, and appropriately so, given the novelty of this enzymatic activity in the RNA world and its implications for coronavirus evolution and emergence. Proofreading allows coronaviruses to encode numerous innate immune antagonists, to replicate faithfully despite changes to nucleotide pools, and to resist antiviral nucleoside analogues (Graepel et al., 2017; Lauber et al., 2013; Sexton et al., 2016; Smith et al., 2013). However, I have not intended to imply that nsp14-ExoN's role in coronavirus biology is limited to preventing error catastrophe. Indeed, the close association of nsp14 with other components of the replicase complex suggests that this protein may be intimately involved in regulating replication at multiple levels. This hypothesis is supported by abundant evidence from this dissertation and other studies. For instance, although

catalytic mutants of MHV-nsp14-ExoN can be recovered, full deletions of nsp14 prevent virus recovery (unpublished observations, Denison lab), indicating that the presence of the ExoN domain itself is important for viability. The ExoN domain is linked by a flexible hinge region to an N7-methyltransferase domain (N7-MT) involved in capping RNA molecules (Ferron et al., 2018). The catalytic sites of nsp14-ExoN and nsp14-N7-MT are functionally distinct and physically independent, but ExoN and N7-MT activities require the presence of both domains (Bouvet et al., 2010; Case et al., 2016; Chen et al., 2009; Ferron et al., 2018). The two domains interact through an extensive hydrophobic surface proximal to the hinge, which may act as a molecular switch to spatially and temporally regulate capping of nascent viral RNAs (Ferron et al., 2018; Ma et al., 2015). Coronavirus ExoN activity has been implicated in evasion of innate immune responses independently of viral capping. In the *Alphacoronavirus* transmissible gastroenteritis virus (TGEV), an ExoN zinc finger mutant triggered a reduced antiviral response relative to wild-type TGEV by generating lower levels of double-stranded RNA during infection (Becares et al., 2016). ExoN catalytic mutants of TGEV were not viable, so their effects on innate immunity could not be measured. However, catalytic mutants in MHV are highly sensitive to interferon treatment, through an undefined mechanism (Case, 2018; Case et al., 2017). These studies have clearly demonstrated that nsp14-ExoN has important roles in replication beyond fidelity regulation. The data presented in this dissertation extends this understanding by identifying mutations in nsp14 that modify replication and fitness without affecting proofreading (Figure 18, Figure 19, Figure 32). The stark difference in fitness effects of nsp14-250 mutations between the ExoN(-) and ExoN(+) backgrounds provides an exciting opportunity to examine how proofreading is linked to other properties of the coronavirus RTC, including speed, processivity, and assembly. Together, my work and existing studies suggest a model for a multifunctional nsp14, in which the ExoN domain interacts with numerous viral, and potentially

host, proteins to regulate multiple steps of the coronavirus life cycle. Indeed, the high structural flexibility of nsp14 has been compared to intrinsically disordered proteins (IDPs) (Ferron et al., 2018). IDPs function as hubs in protein interaction networks, coordinating signaling cascades by interacting with numerous partners in multiple combinations (Dyson, 2016). Like IDPs, nsp14 appears to be a critical nucleus of the coronavirus replication machinery (Imbert et al., 2008; Smith et al., 2014; Subissi et al., 2014; V'kovski et al., 2019). Understanding the various roles of nsp14 in MHV will allow us to understand and overcome the obstacles to recovery of ExoN(-) mutants in other coronaviruses and may reveal new strategies for therapeutic targeting of nsp14.

MHV-ExoN(-) P250 is a sandbox for studying coronavirus biology

In this dissertation, I show that MHV-ExoN(-) is capable of adapting for high-level fitness without primary reversion, displaying dramatic increases in replication, RNA synthesis, and replication fidelity. I traced small contributions to these phenotypes to individual, adapted proteins, but I could never recapitulate the full-genome phenotype using recombinant viruses (Chapters 2,3). These data suggested that multiple proteins contribute additively and/or collaboratively to various adaptive phenotypes, and in many cases, may involve long-range interactions across the entire genome. Some mutations, including those in nsp12-RdRp, directly compensated for ExoN(-) activities and may represent generalizable strategies for overcoming ExoN(-) defects in other cell types and in other coronaviruses (Figure 18, Figure 19). Understanding the mechanisms by which MHV-ExoN(-) P250 compensated for ExoN activity could allow recovery of ExoN(-) variants of other CoVs, such as transmissible gastroenteritis virus, HCoV-229E, and MERS-CoV, which to date have been nonviable as ExoN(-) recombinants (Becares et al., 2016; Minskaia et al., 2006). More broadly, I envision MHV-

ExoN(-) P250 at the center of a wheel, with spokes representing different avenues of potential research. I have focused primarily on the adaptive changes related to replication fidelity, but mutations in MHV-ExoN(-) P250 could conceivably affect any step in the replication cycle. For example, do spike mutations influence entry, assembly, egress, or environmental stability of MHV? Could the nsp2 mutations reveal the function(s) of this conserved but dispensable protein? Do mutations in nsp4 and nsp6 affect the formation of double membrane vesicles, and do mutations in nsp3 modify proteolytic cleavage? Could the lone mutation in the nsp12-NiRAN domain of MHV-ExoN(-) P250 peel back the curtain on this essential but enigmatic enzyme? The questions are limitless, and we have only begun to scratch the surface.

Experimental evolution reveals important insights for coronavirus vaccine development

Early strategies for vaccine development were based on a fundamental observation: pathogenic viruses tend to lose their virulence when passaged experimentally in foreign hosts, such as cultured cells or embryonated eggs. Several life-saving vaccines still in use today were derived from prolonged passage experiments, including the oral poliovirus vaccine and the yellow fever virus vaccine. However, long-term passage has lost favor in the vaccine development world for two related reasons. First, because live-attenuated vaccines are able to replicate, they always carry a risk for reversion to virulence. The oral poliovirus vaccine, for instance, can evolve virulence with just two mutations, and these vaccine-derived polioviruses are the Achilles heel of poliovirus eradication efforts (Kew, 2012). Second, it can be difficult to define the mechanism of attenuation in passaged viruses, raising safety concerns and hampering efforts to predict or prevent reversion (Bull, 2015). With the emergence of new technologies, vaccine development efforts have shifted focus to rational attenuation using genetic engineering. However, the

research in this dissertation argues that experimental evolution can still contribute key insights for vaccine development. Using the toolkit of evolutionary biology, I defined genetic impediments to ExoN(-) reversion that strongly support the potential genetic stability of an ExoN(-) coronavirus vaccine. I also characterized a more accessible evolutionary pathway and demonstrated that passage-adapted protein sequences can constrain the reversibility of attenuating mutations. In their own right, these insights will almost certainly inform strategies to prevent and treat coronavirus infections, while simultaneously developing new reagents for characterizing the critical interactions involved in coronavirus replication. Unfortunately, the experiments described in this dissertation may not be feasible for all viruses, especially those subject to gain-of-function or dual-use-of-concern restrictions. However, the lessons learned from experiments in MHV should be applicable to closely-related viruses like SARS-CoV and MERS-CoV. Thus, long-term passage studies should complement rational genetic engineering in the pursuit of safe and effective vaccine candidates.

Concluding remarks

I remember the fear surrounding the SARS epidemic—the images of healthcare workers in masks on the front of newspapers, the suspicious glances at anyone who coughed, the stories of heroes and victims from across the world. In the preceding pages, I have described my humble contributions to the body of knowledge about coronavirus biology and evolution. I hope that this work has opened doors to illuminate the intricate mechanisms underlying coronavirus replication, evolution, and emergence, and I hope that these insights will aid in developing broadly effective countermeasures for coronavirus disease. With focused study and a little bit of luck, I hope that the next time a coronavirus emerges, we will be able to tell a different story.

Post-script

The story of coronaviruses, at its core, is one of evolution. The incessant swirl of mutation and recombination in coronavirus ecology has, on at least six occasions, and without any foresight or intention, allowed coronaviruses to jump from animals to humans. It seems fitting, then, that the foundational experiment igniting this entire dissertation was undertaken without any agenda or expectations. It started with just a simple question: what will the virus do? This dissertation serves as a reminder, for me at least, of the importance of humility and of curiosity. When Xiaotao and Brett began passaging MHV-ExoN(-), they had no idea that it would reveal so much about coronavirus replication or inform vaccine development efforts for emerging coronaviruses. They just wanted to see what would happen. So, if I may leave one final message, it would be this: nature will tell you what is important, but only if you let it.

APPENDIX A. Mutations in passaged viruses

A.1 Mutations in WT-MHV P250.

Sequencing starts at nucleotide 21 and ends after nucleotide 31279.												
Double asterisk (**) denotes mixed nucleotides at approximately 50%-50% prevelance in the population.												
Nucleotide Change in Genome						Codon		Amino Acid Change in Polyprotein					
Mutation	Position	nsp	Change			Original	New	Type	Original	Position: polyprotein (cleaved)	Mutation	Notes	nsp Boundary (nt)
1	36	N/A	A	→	C	noncoding region							
2	645	1	G	→	A	GCC	ACC	Coding	Ala	146 (146)	Thr	Mixed; A>G	nsp1 (210-950)
3	1844	2	T	→	G	AGT	AGG	Coding	Ser	545 (298)	Arg		nsp2 (951-2705)
4**	1850	2	C	→	Y	CTC	CTY	Silent				Mixed; C≅T	
5	3139	3	A	→	G	GAA	GGA	Coding	Glu	977 (145)	Gly	Mixed; A>G	nsp3 (2706-8720)
6	5521	3	T	→	C	ATC	ACC	Coding	Ile	1771 (939)	Thr		
7	6257	3	T	→	G	AGT	AGG	Coding	Ser	2016 (1184)	Arg	Mixed; G>T	
8	6733	3	C	→	A	ACT	AAT	Coding	Thr	2175 (1343)	Asn	Mixed; A>C	
9	11448	6	G	→	A	GAT	AAT	Coding	Asp	3747 (111)	Asn	Mixed; A>G	nsp6 (11118-11978)
10	12416	8	T	→	C	GCT	GCC	Silent					nsp8 (12255-12836)
11	17871	13	G	→	T	AGT	ATT	Coding	Ser	5888 (504)	Ile		nsp13 (16361-18160)
12	20410	15	T	→	C	TCT	TCC	Silent					nsp15 (19724-20845)

Nucleotide Change in Genome						Codon		Amino Acid Change in Accessory/Structural Protein					
Mutation	Position	Protein	Change			Original	New	Type	Original	Position	Mutation	Notes	Protein Boundary (nt)
13	21971	ns2	A	→	T	CAA	CAT	Coding	Gln	67	His		ns2 (21771-22556)
Deletion: 22690-23878			The majority of was HE deleted. [HE: 22602-23921]										
14	24332	Spike	C	→	A	ACT	AAT	Coding	Thr	135	Asn	Mixed; A>C	Spike (23929-27903)
15	24673	Spike	A	→	C	ACA	CCA	Coding	Thr	249	Pro		
16	24843	Spike	C	→	T	GTC	GTT	Silent					
17	25476	Spike	T	→	A	AAT	AAA	Coding	Asn	516	Lys		
18	25749	Spike	C	→	A	ACC	ACA	Silent					
19	26630	Spike	G	→	T	CGT	CTT	Coding	Arg	901	Leu		
Deletion: 27924-28045			The majority of 4a was deleted. [4a: 27993-28052]										
20	28382	5a	C	→	T	CCA	CTA	Coding	Pro	3	Leu		5a (28375-28713)
21	28447	5a	A	→	C	ATT	CTT	Coding	Ile	25	Leu		
22	28928	E	C	→	A	CTG	ATG	Coding	Leu	75	Met	Mixed; A>C	E (28706-28957)
23	29650	M	C	→	T	ACC	ATC	Coding	Thr	228	Ile		M (28968-29654)

A.2 Mutations in MHV-ExoN(-) P250

Sequencing starts at nucleotide 21 and ends after nucleotide 31275.													
Double asterisk (**) denotes mixed nucleotides at approximately 50%-50% prevelance in the population.													
Nucleotide Change in Genome						Codon		Amino Acid Change in Polyprotein					
Mutation	Position	nsp	Change			Original	New	Type	Original	Position: polyprotein (cleaved)	Mutation	Notes	nsp Boundary (nt)
1	72	N/A	T	→	A	noncoding region							
2	227	1	A	→	C	AAA	AAC	Coding	Lys	6 (6)	Asn		
3	302	1	T	→	A	CCT	CCA	Silent					
4	371	1	G	→	A	TTG	TTA	Silent					
5	563	1	T	→	A	AAT	AAA	Coding	Asn	118 (118)	Lys		nsp1 (210-950)
6	645	1	G	→	A	GCC	ACC	Coding	Ala	146 (146)	Thr		
7	816	1	T	→	C	TCC	CCC	Coding	Ser	203 (203)	Pro		
8	839	1	T	→	C	GTT	GTC	Silent					
9	927	1	G	→	A	GCT	ACT	Coding	Ala	240 (240)	Thr		
10	1496	2	T	→	A	GGT	GGA	Silent					
11**	1623	2	A	→	W	ATG	WTG	Coding	Met	472 (225)	Xxx	Mixed; A≅T	
12	1927	2	T	→	C	GTA	GCA	Coding	Val	573 (326)	Ala		nsp2 (951-2705)
13	2520	2	G	→	A	GAT	AAT	Coding	Asp	771 (524)	Asn		
14	2696	2	T	→	C	TTT	TTC	Silent				Mixed; C>T	
15	2741	3	C	→	T	GTC	GTT	Silent					
16	3223	3	A	→	T	GAG	GTG	Coding	Glu	1005 (173)	Val		
17	3231	3	G	→	A	GCT	ACT	Coding	Ala	1008 (176)	Thr		
18	3371	3	T	→	A	TTT	TTA	Coding	Phe	1054 (222)	Leu		
19	3428	3	G	→	A	GCG	GCA	Silent					
20	3628	3	T	→	C	TTG	TCG	Coding	Leu	1140 (308)	Ser		
21	3697	3	A	→	T	AAG	ATG	Coding	Lys	1163 (331)	Met		
22	3872	3	T	→	C	TCT	TCC	Silent					
23	3903	3	T	→	A	TTG	ATG	Coding	Leu	1232 (400)	Met	Mixed; A>T	
24	3942	3	G	→	A	GTG	ATG	Coding	Val	1245 (413)	Met		
25	3974	3	G	→	A	ACG	ACA	Silent					
26	4280	3	T	→	A	GTT	GTA	Silent				Mixed; A>T	
27	4604	3	T	→	C	AGT	AGC	Silent				Mixed; C>T	
28	4853	3	C	→	T	GAC	GAT	Silent					
29	4880	3	T	→	C	GTT	GTC	Silent					
30	4994	3	T	→	C	TTT	TTC	Silent				Mixed; C>T	
31	5214	3	G	→	A	GAG	AAG	Coding	Glu	1669 (837)	Lys	Mixed; A>G	
32	5255	3	T	→	C	GAT	GAC	Silent					
33	5627	3	T	→	A	GTT	GTA	Silent					
34	5636	3	T	→	A	GTT	GTA	Silent				Mixed; A>T	
35	5951	3	T	→	C	TTT	TTC	Silent					nsp3 (2706-8720)
36	6119	3	G	→	A	TTG	TTA	Silent					
37	6151	3	C	→	T	GCT	GTT	Coding	Ala	1981 (1149)	Val		
38	6194	3	A	→	C	AAA	AAC	Coding	Lys	1995 (1163)	Asn		
39	6211	3	C	→	A	ACA	AAA	Coding	Thr	2001 (1169)	Lys		
40	6305	3	T	→	A	CTT	CTA	Silent					
41**	6420	3	C	→	Y	CCT	YCT	Coding	Pro	2071 (1239)	Xxx	Mixed; T≅C	
42	6489	3	A	→	G	AAG	GAG	Coding	Lys	2094 (1262)	Glu		
43	6516	3	A	→	G	AGT	GGT	Coding	Ser	2103 (1271)	Gly		
44	6520	3	T	→	A	GTG	GAG	Coding	Val	2104 (1272)	Glu		
45	6602	3	T	→	G	GTT	GTG	Silent					
46	6648	3	A	→	T	ACT	TCT	Coding	Thr	2147 (1315)	Ser		
47	6942	3	T	→	A	TTT	ATT	Coding	Phe	2245 (1413)	Ile	Mixed; A>G	
48	7113	3	G	→	A	GCT	ACT	Coding	Ala	2302 (1470)	Thr		
49	7381	3	T	→	C	ATT	ACT	Coding	Ile	2391 (1559)	Thr	Mixed; C>T	
50	7619	3	T	→	C	TGT	TGC	Silent				Mixed; C>T	
51	7748	3	T	→	C	GCT	GCC	Silent					
52	8099	3	C	→	T	GCC	GCT	Silent				Mixed; T>C	
53	8162	3	T	→	C	ACT	ACC	Silent				Mixed; C>T	
54	8642	3	C	→	A	GGC	GGA	Silent				Mixed; A>G	
55	8909	3	C	→	T	AAC	AAT	Silent				Mixed; T>C	
56**	9254	3	T	→	Y	TCT	TCY	Silent				Mixed; C≅T	
57	9287	3	T	→	C	TAT	TAC	Silent				Mixed; C>T	
58	9380	3	C	→	T	TGC	TGT	Silent					
59	9498	3	T	→	A	TTT	ATT	Coding	Phe	3097 (260)	Ile		nsp4 (8721-10208)
60	9887	4	A	→	T	GCA	GCT	Silent					
61	10010	4	T	→	A	TCT	TCA	Silent					
62	10145	4	T	→	C	CAT	CAC	Silent					
63	10188	4	G	→	A	GTT	ATT	Coding	Val	3327 (490)	Ile		

64	10382	5	T	→	C	CTT	CTC	Silent					nsp5 (10209-11117)
65	11710	6	T	→	A	TTC	TAC	Coding	Phe	3834 (198)	Tyr		nsp6 (11118-11978)
66**	12398	8	G	→	R	GAG	GAR	Silent				Mixed; A=G	nsp8 (12255-12836)
67	12720	8	T	→	C	TTT	CTT	Coding	Phe	4171 (248)	Leu		
68	12794	8	T	→	A	ATT	ATA	Silent					
69	12830	8	T	→	A	GTT	GTA	Silent					
70	12973	9	G	→	A	TGT	TAT	Coding	Cys	4255 (46)	Tyr		nsp9 (12837-13166)
71	13331	10	T	→	C	ATT	ATC	Silent				Mixed; C>T	nsp10 (13167-13577)
72	13841	12	T	→	C	TTG	CTG	Silent					nsp12 (13601-16360)
73	13966	12	A	→	T	GCA	GCT	Silent					
74	14077	12	G	→	A	GAG	GAA	Silent				Mixed; A>G	
75	14290	12	T	→	C	TAT	TAC	Silent					
76	14320	12	T	→	C	TGT	TGC	Silent					
77	14439	12	T	→	C	ATG	ACG	Coding	Met	4744 (288)	Thr		
78	14703	12	T	→	C	CTT	CCT	Coding	Leu	4832 (376)	Pro		
79	15436	12	T	→	A	CGT	CGA	Silent					
80	15701	12	C	→	T	CAC	TAC	Coding	His	5165 (709)	Tyr		
81	15873	12	T	→	A	TTT	TAT	Coding	Phe	5222 (766)	Tyr		
82	15902	12	A	→	C	AGT	CGT	Coding	Ser	5232 (776)	Arg		
83	15952	12	G	→	A	GAG	GAA	Silent					
84	16017	12	T	→	A	ATG	AAG	Coding	Met	5270 (814)	Lys		
85	16150	12	T	→	A	CTT	CTA	Silent				Mixed; A>T	
86	16273	12	C	→	T	TAC	TAT	Silent					
87	16294	12	T	→	C	TGT	TGC	Silent				Mixed; C>T	
88**	16315	12	T	→	W	GAT	GAW	Coding	Asp	5369 (913)	Xxx	Mixed; C=T	
89	16369	13	T	→	C	GGT	GGC	Silent				Mixed; C>T	nsp13 (16361-18160)
90	16459	13	T	→	C	CAT	CAC	Silent					
91	16756	13	G	→	A	TTG	TTA	Silent					
92	17083	13	A	→	G	GTA	GTG	Silent					
93	17104	13	C	→	T	AGC	AGT	Silent					
94	17362	13	T	→	A	CCT	CCA	Silent					
95	17746	13	T	→	C	GCT	GCC	Silent					
96	17758	13	T	→	C	AAT	AAC	Silent				Mixed; C>T	
97	17836	13	A	→	G	ATA	ATG	Coding	Ile	5876 (492)	Met		
98	18298	14	T	→	A	TCT	TCA	Silent					nsp14 (18161-19723)
	18426	14	A	→	C	GAT	GCT	Coding	Asp	6073 (89)	Ala	ExoN-(engineered)	
	18427	14	T	→	A	GAT	GAA	Coding					
	18432	14	A	→	C	GAA	GCA	Coding	Glu	6075 (91)	Ala		
	18433	14	A	→	T	GAA	GAT	Coding					
99	18544	14	T	→	A	GAT	GAA	Coding	Asp	6112 (128)	Glu		
100	18748	14	T	→	C	TAT	TAC	Silent					
101	18807	14	T	→	A	TTT	TAT	Coding	Phe	6200 (216)	Tyr		
102	18902	14	T	→	C	TAT	CAT	Coding	Tyr	6232 (248)	His		
103	18910	14	A	→	G	GGA	GGG	Silent					
104	18976	14	T	→	A	GAT	GAA	Coding	Asp	6256 (272)	Glu		
105	19173	14	C	→	T	GCC	GTC	Coding	Ala	6322 (338)	Val		
106	19282	14	T	→	A	GGT	GGA	Silent					
107	19333	14	T	→	C	GTT	GTC	Silent					
108	19396	14	T	→	C	AGT	AGC	Silent					
109	19577	14	T	→	A	TTA	ATA	Coding	Leu	6457 (473)	Ile		
110	19627	14	T	→	A	CTT	CTA	Silent					
111	20120	15	A	→	G	AAT	GAT	Coding	Asn	6638 (133)	Asp	Mixed; G>A	nsp15 (19724-20845)
112	20125	15	C	→	T	GGC	GGT	Silent					
113	20266	15	T	→	C	GAT	GAC	Silent					
114	21265	16	G	→	C	GGG	GGC	Silent					nsp16 (20846-21742)
115	21751	N/A	T	→	C	noncoding region							

Nucleotide Change in Genome						Codon		Amino Acid Change in Accessory/Structural Protein					
Mutation	Position	Protein	Change			Original	New	Type	Original	Position	Mutation	Notes	Protein Boundary (nt)
116	21816	ns2	G	→	A	GCC	ACC	Coding	Ala	16	Thr		ns2 (21771-22556)
117	22137	ns2	T	→	A	TGG	AGG	Coding	Trp	123	Arg		
118	22274	ns2	T	→	C	GGT	GGC	Silent				Mixed; C>T	
Deletion: 22285-23647			ns2 was truncated and the majority of was HE deleted. [HE: 22602-23921]										
119	23771	HE	A	→	T	HE is a non-functional gene with the MHV-A59 reverse genetics system.							HE (22602-23921)
120	23882	HE	T	→	A								
121	23887	HE	T	→	A								
122	24120	S	G	→	A	TTG	TTA	Silent					Spike (23929-27903)
123	24168	S	T	→	C	GCT	GCC	Silent					
124	24273	S	G	→	A	ACG	ACA	Silent					
125	24387	S	C	→	T	TGC	TGT	Silent					
126	24435	S	T	→	C	ACT	ACC	Silent					
127	24438	S	T	→	A	AAT	AAA	Coding	Asn	170	Lys		
128	24630	S	T	→	A	GAT	GAA	Coding	Asp	234	Glu		
129	24736	S	T	→	A	TTT	ATT	Coding	Phe	270	Ile		
130	24869	S	A	→	G	CAA	CGA	Coding	Gln	314	Arg		
131	24938	S	C	→	A	GCT	GAT	Coding	Ala	337	Asp		
132	25039	S	T	→	G	TTT	GTT	Coding	Phe	371	Val		
133	25116	S	C	→	T	CCC	CCT	Silent					
134	25233	S	T	→	A	AAT	AAA	Coding	Asn	435	Lys		
135	25436	S	T	→	C	GTG	GCG	Coding	Val	503	Ala		
136	25562	S	A	→	G	GAT	GGT	Coding	Asp	545	Gly		
137	25605	S	T	→	C	AAT	AAC	Silent					
138	25812	S	T	→	C	TAT	TAC	Silent					
139	25851	S	A	→	T	GCA	GCT	Silent					
140	25878	S	T	→	C	GCT	GCC	Silent					
141	26068	S	A	→	G	AGG	GGG	Coding	Arg	714	Gly	Furin site	
142	26520	S	T	→	A	AGT	AGA	Coding	Ser	864	Arg		
143	26618	S	G	→	A	GGT	GAT	Coding	Gly	897	Asp		
144	26782	S	G	→	A	GTT	ATT	Coding	Val	952	Ile		
145	27210	S	C	→	T	GGC	GGT	Silent					
146	27249	S	T	→	A	CCT	CCA	Silent				Mixed; A>T	
147	27261	S	T	→	C	TAT	TAC	Silent					
148	27294	S	T	→	C	TTT	TTC	Silent					
149	27366	S	T	→	C	TTT	TTC	Silent					
150	27566	S	T	→	C	ATT	ACT	Coding	Ile	1213	Thr		
151	27771	S	G	→	A	GTG	GTA	Silent					
152	28010	4a	A	→	G	CTA	CTG	Silent					4a (27993-28052)
153	28027	4a	T	→	A	CTG	CAG	Coding	Leu	12	Gln		
154	28075	4b	A	→	T	CTA	CTT	Silent					4b (28058-28378)
155	28124	4b	A	→	T	ATA	TTA	Coding	Ile	23	Leu		
156	28254	4b	T	→	A	CTT	CAT	Coding	Leu	66	His	Mixed; A>G	
157	28286	4b	T	→	A	TTT	ATT	Coding	Phe	77	Ile		
158	28500	5a	T	→	C	GGT	GGC	Silent					5a (28375-28713)
159	28652	5a	T	→	A	TTA	TAA	Coding	Leu	93	***		
160	28673	5a	T	→	A	TTT	TAT	Coding	Phe	100	Tyr		
161	28727	E	G	→	A	GAC	AAC	Coding	Asp	8	Asn		E (28706-28957)
162	28813	E	T	→	A	TCT	TCA	Silent					
163	28900	E	T	→	A	CTT	CTA	Silent				Mixed; A>T	
164	28991	M	A	→	T	CCA	CCT	Silent					M (28968-29654)
165	29298	M	A	→	C	AGG	CGG	Silent					
166	29428	M	A	→	T	CAC	CTC	Coding	His	154	Leu		
167	29508	M	T	→	C	TCA	CCA	Coding	Ser	181	Pro		
168	30275	N	T	→	A	TCT	ACT	Coding	Ser	203	Thr	Mixed; A>T	N (29669-31033)
169	30329	N	A	→	T	AAC	TAC	Coding	Asn	221	Tyr	Mixed; T>A	
170	30469	N/A	T	→	A	ATT	ATA	Silent	noncoding region				
171	30955	N/A	A	→	G	CCA	CCG	Silent					

A.3 Mutations in MHV-ExoN(-) P160

Sequencing starts at nucleotide 21 and ends after nucleotide 31292.													
Double asterisk (**) denotes mixed nucleotides at approximately 50%-50% prevalance in the population.													
Nucleotide Change in Genome						Codon		Amino Acid Change in Polyprotein					
Mutation	Position	nsp	Change			Original	New	Type	Original	Position	Mutation	Notes	nsp Boundary (nt)
1	72	N/A	T	→	A	noncoding region							
2	227	1	A	→	C	AAA	AAC	Coding	Lys	6	Asn		nsp1 (210-950)
3	302	1	T	→	A	CCT	CCA	Silent					
4	371	1	G	→	A	TTG	TTA	Silent					
5	563	1	T	→	A	AAT	AAA	Coding	Asn	118	Lys		
6	645	1	G	→	A	GCC	ACC	Coding	Ala	146	Thr		
7	839	1	T	→	C	GTT	GTC	Silent					
8	927	1	G	→	A	GCT	ACT	Coding	Ala	240	Thr		
9	1927	2	T	→	C	GTA	GCA	Coding	Val	573	Ala	Mixed; C>T	nsp2 (951-2705)
10	2520	2	G	→	A	GAT	AAT	Coding	Asp	771	Asn		
11	2741	3	C	→	T	GTC	GTT	Silent					nsp3 (2706-8720)
12	3223	3	A	→	T	GAG	GTG	Coding	Glu	1005	Val	Mixed; T>A	
13	3231	3	G	→	A	GCT	ACT	Coding	Ala	1008	Thr		
14	3371	3	T	→	A	TTT	TTA	Coding	Phe	1054	Leu		
15	3428	3	G	→	A	GCG	GCA	Silent					
16	3628	3	T	→	C	TTG	TCG	Coding	Leu	1140	Ser		
17	3697	3	A	→	T	AAG	ATG	Coding	Lys	1163	Met		
18	3872	3	T	→	C	TCT	TCC	Silent					
19	3942	3	G	→	A	GTG	ATG	Coding	Val	1245	Met		
20	3974	3	G	→	A	ACG	ACA	Silent					
21	4853	3	C	→	T	GAC	GAT	Silent					
22	5255	3	T	→	C	GAT	GAC	Silent					
23	5627	3	T	→	A	GTT	GTA	Silent					
24	5951	3	T	→	C	TTT	TTC	Silent					
25	6119	3	G	→	A	TTG	TTA	Silent					
26	6151	3	C	→	T	GCT	GTT	Coding	Ala	1981	Val		
27	6194	3	A	→	C	AAA	AAC	Coding	Lys	1995	Asn		
28	6211	3	C	→	A	ACA	AAA	Coding	Thr	2001	Lys		
29	6305	3	T	→	A	CTT	CTA	Silent					
30	6489	3	A	→	G	AAG	GAG	Coding	Lys	2094	Glu		
31	6516	3	A	→	G	AGT	GGT	Coding	Ser	2103	Gly		
32	6520	3	T	→	A	GTG	GAG	Coding	Val	2104	Glu		
33	6602	3	T	→	G	GTT	GTG	Silent					
34	7748	3	T	→	C	GCT	GCC	Silent					
35	8642	3	C	→	G	GGC	GGG	Silent					
36	9380	4	C	→	T	TGC	TGT	Silent					nsp4 (8721-10208)
37	9498	4	T	→	A	TTT	ATT	Coding	Phe	3097	Ile		
38	9887	4	A	→	T	GCA	GCT	Silent					
39	10010	4	T	→	A	TCT	TCA	Silent					
40	10382	5	T	→	C	CTT	CTC	Silent					nsp5 (10209-11117)
41	11710	6	T	→	A	TTC	TAC	Coding	Phe	3834	Tyr		nsp6 (11118-11978)
42	12398	8	G	→	A	GAG	GAA	Silent					nsp8 (12255-12836)
43	12720	8	T	→	C	TTT	CTT	Coding	Phe	4171	Leu	Mixed T=C	
44	12794	8	T	→	A	ATT	ATA	Silent					
45	12830	8	T	→	A	GTT	GTA	Silent					
46	12973	9	G	→	A	TGT	TAT	Coding	Cys	4255	Tyr		nsp9 (12837-13166)
47	13841	12	T	→	C	TTG	CTG	Silent					nsp12 (13601-16360)
48	13966	12	A	→	T	GCA	GCT	Silent					
49	14290	12	T	→	C	TAT	TAC	Silent					
50	14320	12	T	→	C	TGT	TGC	Silent					
51	14439	12	T	→	C	ATG	ACG	Coding	Met	4744	Thr		
52	14703	12	T	→	C	CTT	CCT	Coding	Leu	4832	Pro		
53	15004	12	T	→	C	TTT	TTC	Silent				Mixed; T=C	
54	15436	12	T	→	A	CGT	CGA	Silent				Mixed; T=A	
55	15701	12	C	→	T	CAC	TAC	Coding	His	5165	Tyr		
56	15873	12	T	→	A	TTT	TAT	Coding	Phe	5222	Tyr		
57	15902	12	A	→	C	AGT	CGT	Coding	Ser	5232	Arg		
58	15952	12	G	→	A	GAG	GAA	Silent					
59	16017	12	T	→	A	ATG	AAG	Coding	Met	5270	Lys		
60	16273	12	C	→	T	TAC	TAT	Silent				Mixed; T>C	

61	16459	13	T	→	C	CAT	CAC	Silent					nsp13 (16361-18160)
62	16756	13	G	→	A	TTG	TTA	Silent					
63	17083	13	A	→	G	GTA	GTG	Silent					
64	17104	13	C	→	T	AGC	AGT	Silent				Mixed; T>C	
65	17362	13	T	→	A	CCT	CCA	Silent					
66	17836	13	A	→	G	ATA	ATG	Coding	Ile	5876	Met		
	18426	14	A	→	C	GAT	GCT	Coding	Asp	6073	Ala	ExoN-(engineered)	nsp14 (18161-19723)
	18427	14	T	→	A	GAT	GAA	Coding					
	18432	14	A	→	C	GAA	GCA	Coding	Glu	6075	Ala		
	18433	14	A	→	T	GAA	GAT	Coding					
67	18544	14	T	→	A	GAT	GAA	Coding	Asp	6112	Glu		
68	18807	14	T	→	A	TTT	TAT	Coding	Phe	6200	Tyr		
69	18902	14	T	→	C	TAT	CAT	Coding	Tyr	6232	His		
70	18910	14	A	→	G	GGA	GGG	Silent					
71	18976	14	T	→	A	GAT	GAA	Coding	Asp	6256	Glu		
72	19173	14	C	→	T	GCC	GTC	Coding	Ala	6322	Val		
73	19282	14	T	→	A	GGT	GGA	Silent				Mixed; A=T	
74	19333	14	T	→	C	GTT	GTC	Silent					
75	19396	14	T	→	C	AGT	AGC	Silent					
76	19577	14	T	→	A	TTA	ATA	Coding	Leu	6457	Ile		
77	19627	14	T	→	A	CTT	CTA	Silent					
78	20125	15	C	→	T	GGC	GGT	Silent					nsp15 (19724-20845)
79	20200	15	A	→	T	GTA	GTT	Silent					
80	20266	15	T	→	C	GAT	GAC	Silent					
81	21265	16	G	→	C	GGG	GGC	Silent					nsp16 (20846-21742)
82	21751	N/A	T	→	C	noncoding region							

Nucleotide Change in Genome						Codon		Amino Acid Change in Accessory/Structural Protein					
Mutation	Position	Protein	Change			Original	New	Type	Original	Position	Mutation	Notes	Protein Boundary (nt)
83	21816	ns2	G	→	A	GCC	ACC	Coding	Ala	16	Thr		ns2 (21771-22556)
84	22137	ns2	T	→	A	TGG	AGG	Coding	Trp	123	Arg		
Deletion: 22286-23646			ns2 was truncated and the majority of was HE deleted. [HE: 22602-23921]										
85	23404	HE	C	→	T	HE is a non-functional gene with the MHV-A59 reverse genetics system.							HE (22602-23921)
86	23522	HE	G	→	A								
87	23771	HE	A	→	T								
88	23882	HE	T	→	A								
89	23887	HE	T	→	A								
90	24168	S	T	→	C	GCT	GCC	Silent					Spike (23929-27903)
91	24435	S	T	→	C	ACT	ACC	Silent					
92	24630	S	T	→	A	GAT	GAA	Coding	Asp	234	Glu		
93	24869	S	A	→	G	CAA	CGA	Coding	Gln	314	Arg		
94	25039	S	T	→	G	TTT	GTT	Coding	Phe	371	Val		
95	25233	S	T	→	A	AAT	AAA	Coding	Asn	435	Lys		
96	25436	S	T	→	C	GTG	GCG	Coding	Val	503	Ala		
97	25562	S	A	→	G	GAT	GGT	Coding	Asp	545	Gly		
98	25605	S	T	→	C	AAT	AAC	Silent					
99	25851	S	A	→	T	GCA	GCT	Silent					
100	26259	S	A	→	T	GCA	GCT	Silent					
101	26520	S	T	→	A	AGT	AGA	Coding	Ser	864	Arg		
102	26618	S	G	→	A	GGT	GAT	Coding	Gly	897	Asp		
103	26766	S	G	→	A	GTG	GTA	Silent					
104	26782	S	G	→	A	GTT	ATT	Coding	Val	952	Ile		
105	27036	S	T	→	A	GCT	GCA	Silent					
106	27210	S	C	→	T	GGC	GGT	Silent					
107	27261	S	T	→	C	TAT	TAC	Silent					
108	27294	S	T	→	C	TTT	TTC	Silent					
109	27366	S	T	→	C	TTT	TTC	Silent					
110	27771	S	G	→	A	GTG	GTA	Silent					
111	27948	N/A	A	→	G	noncoding region							
112	28027	4a	T	→	G	CTG	CGT	Coding	Leu	12	Arg		4a (27993-28052)
113	28028	4a	G	→	T								
114	28048	4a	T	→	A	TTG	TAG	***	Leu	19	Stop		
115	28286	4b	T	→	A	TTT	ATT	Coding	Phe	77	Ile		4b (28058-28378)
116	28391	5a	C	→	T	ACA	ATA	Coding	Thr	6	Ile		5a (28375-28713)
117	28436	5a	A	→	G	GAC	GGC	Coding	Asp	21	Gly		
118	28500	5a	T	→	C	GGT	GGC	Silent					
119	28586	5a	G	→	A	GGT	GAT	Coding	Gly	71	Asp		
120	28727	E	G	→	A	GAC	AAC	Coding	Asp	8	Asn		E (28706-28957)
121	28813	E	T	→	A	TCT	TCA	Silent					
122	29298	M	A	→	C	AGG	CGG	Silent					M (28968-29654)
123	29428	M	A	→	T	CAC	CTC	Coding	His	154	Leu		
124	30310	N	T	→	C	AAT	AAC	Silent					N (29669-31033)
125	30955	N/A	A	→	G	CCA	CCG	Silent	noncoding region				

APPENDIX B. New protocols

B1. PrimerID for measuring mutation frequencies in murine hepatitis virus

Reagents and kits:

1. QIAamp viral RNA mini kit (Qiagen #52904)
2. Nuclease-free H2O
3. Mixed dNTPs at 10mM each (any supplier)
4. Superscript III First-Strand Synthesis System (Invitrogen #18080051) or your preferred reverse transcriptase
5. RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen #10777019)
6. Ribonuclease H (Invitrogen #18021071)
7. Agencourt RNAclean XP beads (Beckman Coulter #A63987)
8. Agencourt AMPure beads (Beckman Coulter #A63881)
9. Dynamag-2 Magnetic Bead Rack (Invitrogen #12321D)
10. 100% ethanol
11. Kapa2G Robust PCR Kit (Kapa Biosystems #07960867001)
12. Kapa HiFi PCR Kit (Kapa Biosystems #07958846001)
13. MinElute Gel Extraction Kit (Qiagen # 28604)
14. Qubit dsDNA BR Assay Kit (Invitrogen #Q32850)

Primer ID primers.
PrimerID is in bold; indices for multiplex barcoding are underlined.

Primer Name	Purpose	Sequence
MHV20814-cDNA	Reverse transcription	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**NNNN NNNNNN**CCAATGTTCTTAGTAAGAGGGTCGTACATA
MHV20814-PCR1	PCR1, forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNA GAAGGTCATGACTTTCTATCCTCGTTTGC
PID-PCR-1	PCR1 reverse	GTGACTGGAGTTCAGACGTGTGCTC
D501-TATAGCCT	Index, forward	AATGATACGGCGACCACCGAGATCTACACTATAGCCTA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D502-ATAGAGGC	Index, forward	AATGATACGGCGACCACCGAGATCTACACATAGAGGCA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D503-CCTATCCT	Index, forward	AATGATACGGCGACCACCGAGATCTACACCCTATCCTA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D504-GGCTCTGA	Index, forward	AATGATACGGCGACCACCGAGATCTACACGGCTCTGAA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D505-AGGCGAAG	Index, forward	AATGATACGGCGACCACCGAGATCTACACAGGCGAAGA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D506-TAATCTTA	Index, forward	AATGATACGGCGACCACCGAGATCTACACTAATCTTAA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D507-CAGGACGT	Index, forward	AATGATACGGCGACCACCGAGATCTACACCAGGACGTA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D508-GTACTGAC	Index, forward	AATGATACGGCGACCACCGAGATCTACACGTACTGACA CACTCTTTCCCTACACGACGCTCTTCCGATCT
D701-ATTACTCG_R	Index, reverse	CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACT GGAGTTCAGACGTGTGCTCTTCCGATC

Library Preparation and Sequencing.

The starting material for library preparation should be RNA genomes purified from supernatant RNA. I like to use the Qiagen QIAamp viral RNA mini kit because it can be used at the benchtop, unlike TRIzol which requires the fume hood. Quantify genomic RNA copies using the qPCR method of your choice.

Reverse transcription.

This step creates PrimerID-tagged cDNA.

1. Make initial RNA mixture

Reagent	[stock]	Volume (µL)
RNA template	Variable	100,000 genome copies
dNTP mix	10 mL each	3
MHV20814-cDNA	10 µM	1.5
Nuclease free H_2O	N/A	to 39 µL

2. Incubate at 65ºC for 5 minutes to denature RNA template, then transfer to ice for at least 1 minute to anneal cDNA primer.
3. Add the remaining reverse transcription reagents. I typically make a master mix to streamline the process.

Reagent	[stock]	Volume (µL)
5x Buffer	5x	12
DTT	100mL	3
RNaseOUT	40 u/µL	3
SuperScript III	200 u/µL	3
Volume per tube	*N/A*	*21*

4. Incubate in the thermocycler:
    a. 50ºC 1 hour
    b. 55ºC 1 hour
    c. 70ºC 15 minutes (to inactivate SSIII)
5. Add 1 µL of RNase H to each tube and incubate at 37ºC for 20 minutes.
6. Purify cDNA with Agencourt RNAClean XP beads
    a. Resuspend beads, remove an aliquot, and let them equilibrate to room temperature. Vortex immediately before use.
    b. Transfer cDNA reactions to 1.5mL Eppendorf tubes and add beads at a ratio of 1.2:1 (72 µL beads per 60 µL sample).
    c. Mix thoroughly **without vortexing** and incubate at room temperature for 20 minutes.
    d. Place tube on magnetic rack and allow to separate for 5 minutes.
    e. Leaving the tube on the rack, remove supernatant and discard.
    f. Rinse beads three times with 300µL of 80% ethanol.
    g. After third wash, spin tubes briefly with benchtop microfuge and place back on the rack; remove any remaining liquid.
    h. Allow beads to dry until they just turn matte, 5-10 minutes. Over-drying reduces yields.

i. When beads are dry, remove tubes from rack and add 24 µL of nuclease-free H_2O. Resuspend beads by pipetting up and down and incubate 5 minutes.
j. Move tubes back to rack and remove cDNA solution. Use directly in next step.

First-round PCR.

This step amplifies the region of interest.

1. Set up PCR.

Reagent	[stock]	Volume (µL)
Template cDNA		23.5
5x **Buffer A**	5x	10
5x Enhancer 1	5x	10
dNTP mix	10 mM each	1
MHV20814-PCR1	10 µM	2.5
PID-PCR1-R	10 µM	2.5
Kapa2G Robust polymerase	5 U/mL	0.5

2. Incubate in thermocycler:
   a. 95ºC 1 min
   b. 95ºC 15 sec
   c. 58ºC 1 min
   d. 72ºC 30 sec
   e. Repeat (b) – (d) for 25 cycles.
   f. 72ºC 3 min
   g. 4ºC hold
3. Purify PCR1 products with Agencourt AMPureXP beads.
   a. Resuspend beads, remove an aliquot, and let them equilibrate to room temperature. Vortex immediately before use.
   b. Transfer PCR1 solution to 1.5mL Eppendorf tubes and add beads at a ratio of 1.2:1 (60 µL beads per 50 µL sample).
   c. Mix thoroughly **without vortexing** and incubate at room temperature for 20 minutes.
   d. Place tube on magnetic rack and allow to separate for 5 minutes.
   e. Leaving the tube on the rack, remove supernatant and discard.
   f. Rinse beads three times with 300µL of 80% ethanol.
   g. After third wash, spin tubes briefly with benchtop microfuge and place back on the rack; remove any remaining liquid.
   h. Allow beads to dry until they just turn matte, 5-10 minutes. Over-drying reduces yields.
   i. When beads are dry, remove tubes from rack and add 50 µL of nuclease-free H_2O. Resuspend beads by pipetting up and down and incubate 5 minutes.
   j. Move tubes back to rack and remove cDNA solution.

Second-round PCR.

This step adds Illumina adaptors and indexed barcodes for multiplex sequencing.

1. Set up PCR.

Reagent	[stock]	Volume (µL)
PCR1 product		2
5x Kapa HiFi Buffer	5x	5
dNTP mix	10 mM each	1
D50(x) primer	10 µM	1
D701-ATTACTCG_R	10 µM	1
Kapa HiFi polymerase	1 U/mL	0.5
Nuclease-free H_20		14.5

2. Incubate in thermocycler:
    a. 95ºC 2 min
    b. 98ºC 20 sec
    c. 63ºC 15 sec
    d. 72ºC 30 sec
    e. Repeat (b) – (d) for 25 cycles.
    f. 72ºC 3 min
    g. 4ºC hold
3. Run PCR2 products on a 1.2% agarose gel and extract with the Qiagen MinElute gel extraction kit.
    a. Excise DNA fragment and weigh the gel.
    b. Add 3 volumes of buffer QG to 1 volume of gel and incubate at 50ºC, vortexing intermittently, until fully dissolved. If the solution is not yellow, add 10 µL of 3M sodium acetate.
    c. Apply sample to MinElute column, and centrifuge at top speed 1 minute at room temperature. Discard flowthrough.
    d. Add 500 µL buffer QG, and centrifuge at top speed 1 minute at room temperature. Discard flowthrough.
    e. Add 750 µL buffer PE, incubate at room temperature for 5 minutes, then centrifuge at top speed 1 minute at room temperature.
    f. Discard flowthrough and centrifuge at top speed an additional 3 minutes to dry the membrane.
    g. Transfer MinElute column to a fresh 1.5mL Eppendorf tube and add 10 µL buffer EB. Incubate at room temperature for 5 minutes, then centrifuge at top speed for 2 minutes to elute DNA products.
4. Quantify library using the Invitrogen Qubit dsDNA BR assay kit.
    a. Follow manufacturer's instructions.
    b. **Do not use Nanodrop to quantify**; it is not precise enough.
5. Dilute libraries to 5 nM and pool in equal volumes for sequencing.
6. Quality control should include:
    a. qPCR to establish the presence of adaptor sequences
    b. BioAnalyzer to find primer dimers that can interfere with sequencing.

<u>Sequencing.</u>

1. Libraries should be sequenced on the Illumina MiSeq with 2x250 paired-end chemistry.
2. The pooled libraries should make up 50% of the chip, with bacterially-derived DNA ("PhiX") making up the remainder.

Analysis.

Sequencing results should arrive as zipped drives containing fastq files (XX.fastq.gz). Shuntai Zhou (UNC Chapel Hill) wrote the ruby code for creating PrimerID-derived consensus sequences and for combining forward and reverse reads into single reads. The subsequent analysis and code were provided by Matthew Pauly (University of Michigan, Lauring lab; https://github.com/lauringlab/NGS_mutation_rate_assay) (Pauly et al., 2017b).

1. Generate PrimerID consensus sequences. Follow either (a) or (b) below.
    a. The Swanstrom lab developed a web portal that will automatically generate consensus scripts from raw data. https://tcs-dr-dept-tcs.cloudapps.unc.edu/TCS/
        i. Upload sequencing files (name.fastq.gz) to a Dropbox folder.
        ii. Submit link to Dropbox folder along with primer sequences:
            1. MHV20814-cDNA
            2. MHV20814-PCR1
        iii. You will get an email with a link to download the results. This sometimes lands in a spam folder, so if you don't receive them in 24 hours, follow up with Michael Clark (clarkmu@email.unc.edu).

    b. Run the ruby script on your own data.
        i. Unzip the fastq files:

            gunzip file.gz

        ii. Run ruby script:

            ruby TCS_MHV20814.rb {input_directory} {length of primerID}

    c. Regardless of whether you use 1a or 1b, the output will be a new directory within the input directory that contains:
        i. Consensus files (r1.txt and r2.txt, corresponding to r1 and r2 reads from Illumina sequencing)
        ii. Log file (look here for the number of consensus sequences and the consensus threshold)
        iii. PrimerID statistics (identity and usage)

2. Combine the read1 and read2 consensus into a single consensus sequence.
    a. It's easiest to make a new directory for this, organized as follows:
        i. Name_of_directory
            1. Sample1
                a. r1.txt
                b. r2.txt

2. Sample2
    a. r1.txt
    b. r2.txt
3. …etc

b. The script will combine overlapped regions so that they are not counted twice in downstream analysis. The size of the overlap has to be set in the script; for MHV20814, the overlap is 33.

ruby pid_join_two_ends.rb "/name_of_directory"

c. Output will be single .txt files for each sample

3. Convert .txt files from (2c) to .fasta

cp combined.txt combined.fasta

4. Align the combined consensus sequences using bowtie2.
    a. Bowtie2 needs an indexed genome file to use as its reference during alignment. If you don't have one, you'll need to make it. Generate a fasta file containing the region of interest (e.g. nts 20600-21400 of MHV genome for "MHV20814").

bowtie2 build -f MHV20814ref.fasta MHV20814ref

    b. Align reads to your new indexed reference.
        i. -f designates that input is in fasta format rather than fastq
        ii. -x marks the reference file
        iii. -U marked the sequences as unpaired
        iv. -S names the output file

bowtie2 -f -x MHV20814ref -U "combined.fasta" -S "aligned.sam"

5. Use samtools to convert .sam files to .bam files (necessary for subsequent steps).

samtools view -Sb "aligned.sam" > "aligned.bam"

6. Use samtools to create a pileup that lists all consensus sequences in the alignment by base position. (-d 100000 designates depth. Default is 8000, which will be too low oo low)

samtools mpileup -d 100000 "aligned.bam" > "aligned.pileup"

7. Use Lauring lab python script to count the number of each base type at each position in the alignment. This gives the number of A, T, C, G, and N at each position in the sequence.

python counts.py "aligned.pileup" "output.csv"

8. I determine nonsense mutation frequencies manually after opening the output.csv in Excel.

9. Lauring lab script "Mutcount_analysis.R" can be used to determine the frequencies of all mutations in R. The input will be your unmodified output.csv files.

B2. Bioinformatic code for PrimerID analysis

Identifying nonsense mutational targets within an open reading frame

File name: stop_codon_script.R (necessary modifications bolded below). Script prepared by Matthew Pauly (Pauly et al., 2017b)

Note: the script scans along until it finds the first ATG, so make sure the earliest ATG in your genome file is actually your ORF's start codon. Also, it will not recognize the ATG if it is at the start of the sequence, so add at least 10 nucleotides before it.

```
#This script is designed to determine the number of codons in a gene that can
mutate into a premature stop codon.

#function inputs are:
# genomefile - the ".fas" file that you want to analyze
# bases - the number of bases that you want in each window that is analyzed
for codons that can mutate to stop codons
# sortcol - the column in the output that you want to sort by
      # 2- sum of all codons that can mutate to stop codon
      # 3-  codons assessing U2A and U2G
      # 4-  codons assessing U2A
      # 5-  codons assessing C2A and C2G
      # 6-  codons assessing C2A
      # 7-  codons assessing C2U
      # 8-  codons assessing G2A
      # 9-  codons assessing A2U
      # 10- codons assessing G2U

StopCodon<-function(genomefile, bases, sortcol){

genefile<-read.csv("your_genome_file",sep=";")              #reads the
unzipped file

genefile2<-genefile[,]                          #Makes the read-in file
usable for the following manipulation

gene<-toupper(paste(genefile2, collapse=""))               #pastes the read-
in file into a list of the entire genome

start<-(regexpr("ATG", gene)[1])                      #Finds the FIRST start
codon in the seqience that is given

genesplit<-strsplit(gene,split='')        #This and next 2 split string into
list of individual bases

same<-function(x){x}
```

```
singlebase<-sapply(genesplit, same)

len<-length(singlebase)

geneORF<-paste(singlebase[(start:len),], collapse="")    #makes string
starting with the forst ATG

geneORF

########

                        #The following for loop will break the gene into
codons


for (q in 1:length(geneORF)){
  codons<-substring(geneORF[q], seq(1,nchar(geneORF[q])-
1,3),seq(3,nchar(geneORF[q]),3))
}

            #The following replaces all codons that can mutate to a stop
codon with  "1". All other codons replaced with a "0"


lethcodonmut<-function(x){
      y<-gsub("TTA|TAT","1",x)
      y<-gsub("TTG|TGT","1",y)
      y<-gsub("TCA|TAC","1",y)
      y<-gsub("TGC|TCG","1",y)
      y<-gsub("CAA|CAG|CGA","1",y)
      y<-gsub("TGG","1",y)
      y<-gsub("AAA|AAG|AGA","1",y)
      y<-gsub("GAA|GAG|GGA","1",y)
      y<-gsub("...","0",y)
      y
      }

lethcodonmutations<-lethcodonmut(codons)              #changes codons to
character numerals

le<-as.numeric(lethcodonmutations)                    #makes numeric

######

lethcodonmuta<-function(x){
      y<-gsub("TTA|TAT","1",x)
      y<-gsub("TTG|TGT","2",y)
      y<-gsub("TCA|TAC","3",y)
      y<-gsub("TGC|TCG","4",y)
      y<-gsub("CAA|CAG|CGA","5",y)
      y<-gsub("TGG","6",y)
      y<-gsub("AAA|AAG|AGA","7",y)
      y<-gsub("GAA|GAG|GGA","8",y)
      y<-gsub("...","0",y)
      y
      }
```

```
lethcodonmutationsa<-lethcodonmuta(codons)                    #changes codons
to character numerals

lea<-as.numeric(lethcodonmutationsa)                          #makes numeric


#######



cod<-round("bases"/3)                                  #converts the bases
input into number of codons; set "bases" to the amplicon size you want. Make
sure it is evenly divisible by 3.



windows<-matrix(rep(0),nrow=(length(le)-cod), ncol=10)            #makes
vector to populate with scores for windows of codons


                                                  #the following will divide
into windows of 100 codons (300 bases) and sum

colnames(windows)<-c("start base","number of lethal codons","U2A and
U2G","U2A","C2A and C2G","C2A","C2U","G2A","A2U","G2U")

for (q in 1:(length(le)-cod)){
      windows[q,2]<-sum(le[(0+q):(cod-1+q)])
      windows[q,1]<-((3*q)-2)
      windows[q,3]<-sum(lea[(0+q):(cod-1+q)]==1)
      windows[q,4]<-sum(lea[(0+q):(cod-1+q)]==2)
      windows[q,5]<-sum(lea[(0+q):(cod-1+q)]==3)
      windows[q,6]<-sum(lea[(0+q):(cod-1+q)]==4)
      windows[q,7]<-sum(lea[(0+q):(cod-1+q)]==5)
      windows[q,8]<-sum(lea[(0+q):(cod-1+q)]==6)
      windows[q,9]<-sum(lea[(0+q):(cod-1+q)]==7)
      windows[q,10]<-sum(lea[(0+q):(cod-1+q)]==8)

}

windowsort<-windows[order(windows[,sortcol], decreasing=TRUE),]

                  #The following will allow the script to return two things
(The ORF sequence and the output table)

ret<-(head(windowsort, 100))

retu<-list(geneORF,ret)

return(retu)

}
```

Generating consensus reads using the PrimerID

File name: TCS_MHV20814.rb (no modifications should be necessary). Script was written by Shuntai Zhou (Zhou et al., 2015).

```
=begin
TCS Pipeline Version 1.38-07AUG2018
Create Primer ID template consensus sequences from raw MiSeq FASTq file
Input = directory of raw sequences of two ends (R1 and R2 fasta files,
unzipped)
Require parameters:
  list of Primer Sequence of cDNA primer and 1st round PCR forward Primer,
including a tag for the pair name
  ignore the first nucleotide of Primer ID: Yes/No
=end
ver = "1.38-07AUG2018"
#############Patch Note#############
=begin
  1. Improved performace.
=end



#############cDNA gene-specific region and forward primer region needs to be
defined#######

#mutilple cDNA primers, remove the # before use
primers = {}

#change set_name, forward primer sequence and cDNA primer sequence. both
forward primer sequence and cDNA primer sequence should include the entire
sequence, not just biological sequence
#primers["set_name"] = ["forward primer sequence", "cDNA primer sequence"]
#example of primer

primers["MHV20814"] =
["ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGAAGGTCATGACTTTCTATCCTCGTTTGC","GTGAC
TGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNCCAATGTTCTTAGTAAGAGGGTCGTACATA"]

#ignore the first nucleotide of the PID, default value true, remove the #
before use

$ignore_first_nt = true unless defined? $ignore_first_nt

#input file is the directory containing sequences from both ends of one
library
indir = ARGV[0]

####################General Methods

#convert array to hash in a memory-saving way
def array_to_hash(array)
  count = 0
  hash = Hash.new
  (array.length / 2).times do
```

```
    hash[array[count]] = array[count+1]
    count += 2
  end
  return hash
end

#count frequencies of elements in a array.
def count(array)
  hash = Hash.new(0)
  array.each do |element|
    hash[element] +=1
  end
  return hash
end

#calculate consensus cutoff
#error = 0.02
def calculate_cut_off(m)
  n = 0
  if m <= 10
    n = 2
  elsif m <= 8500
    n = -1.24*10**-21*m**6 + 3.53*10**-17*m**5 - 3.90*10**-13*m**4 +
2.12*10**-9*m**3 - 6.06*10**-6*m**2 + 1.80*10**-2*m + 3.15
  else
    n = 0.0079 * m + 9.4869
  end
  n = n.round
  n = 2 if n < 3
  return n
end

=begin
#error rate = 0.01
def calculate_cut_off(m)
  n = 0
  if m <= 10
    n = 2
  else
    n = 1.09*10**-26*m**6 + 7.82*10**-22*m**5 - 1.93*10**-16*m**4 +
1.01*10**-11*m**3 - 2.31*10**-7*m**2 + 0.00645*m + 2.872
  end
  n = n.round
  n = 2 if n < 3
  return n
end
=end

=begin
#error rate = 0.005
def calculate_cut_off(m)
  n = 0
  if m <= 10
    n = 2
  else
    n = -9.59*10**-27*m**6 + 3.27*10**-21*m**5 - 3.05*10**-16*m**4 +
1.2*10**-11*m**3 - 2.19*10**-7*m**2 + 0.004044*m + 2.273
```

```
  end
  n = n.round
  n = 2 if n < 3
  return n
end
=end

#obtain a consensus sequences
def consensus_without_alignment(seq_array,gap_treatment = 1)
  all_length = []
  seq_array.each {|seq| all_length << seq.size}
  length = all_length.mean.round(0)
  consensus_bases = []
  (0..(length-1)).each do |n|
    bases = []
    seq_array.each do |seq|
      bases << seq[n]
    end
    if gap_treatment == 1
      consensus_bases << creat_consensus_base_non_gap(bases)
    else
      consensus_bases << creat_consensus_base_gap(bases)
    end
  end
  consensus_seq = consensus_bases.join('')
end

#create a consensus base call at a position.
def creat_consensus_base_non_gap(base_array_input)
  base_array = Array.new(base_array_input)
  consensus_base = '-'
  number_of_bases = base_array.size
  h = Hash.new(0)
  if base_array.size >0
    base_array.each do |base|
      h[base] += 1
    end
    max_number = h.values.max
    max_list = []
    h.each do |k,v|
      if v == max_number
        max_list << k
      end
    end
    maxi_list_size = max_list.size
    if maxi_list_size == 1
      consensus_base = max_list.shift
    elsif maxi_list_size >= 3
      consensus_base = "N"
    elsif maxi_list_size == 2
      if max_list.include?("A") and max_list.include?("T")
        consensus_base = "W"
      elsif max_list.include?("A") and max_list.include?("C")
        consensus_base = "M"
      elsif max_list.include?("A") and max_list.include?("G")
        consensus_base = "R"
      elsif max_list.include?("T") and max_list.include?("C")
```

```
        consensus_base = "Y"
      elsif max_list.include?("G") and max_list.include?("C")
        consensus_base = "S"
      elsif max_list.include?("T") and max_list.include?("G")
        consensus_base = "K"
      elsif max_list.include?('-')
        max_list.delete('-')
        consensus_base = max_list.shift
      end
    end
  end
  return consensus_base.chr if consensus_base
end


#primer with ambiguities to match
def primer_match (primer = "")
  match = ""
  primer.each_char.each do |base|
    base_array = to_list(base)
    if base_array.size == 1
      match += base_array[0]
    else
      pattern = "[" + base_array.join("|") + "]"
      match += pattern
    end
  end
  return match
end

def to_list(base = "")
  list = []
  case base
  when /[A|T|C|G]/
    list << base
  when "W"
    list = ['A','T']
  when "S"
    list = ['C','G']
  when "M"
    list = ['A','C']
  when 'K'
    list = ['G','C']
  when 'R'
    list = ['A','G']
  when 'Y'
    list = ['C','T']
  when 'B'
    list = ['C','G','T']
  when 'D'
    list = ['A','G','T']
  when 'H'
    list = ['A','C','T']
  when 'V'
    list = ['A','C','G']
  when 'N'
    list = ['A','T','C','G']
```

```
  end
  return list
end

module Enumerable
  def median
    len = self.length
    sorted = self.sort
    median = len % 2 == 1 ? sorted[len/2] : (sorted[len/2 - 1] +
sorted[len/2]).to_f / 2
  end

  def sum
     self.inject(0){|accum, i| accum + i }
  end

  def mean
    self.sum/self.length.to_f
  end

  def sample_variance
    m = self.mean
    sum = self.inject(0){|accum, i| accum + (i-m)**2 }
    sum/(self.length - 1).to_f
  end

  def stdev
    return Math.sqrt(self.sample_variance)
  end

end


#compare PID with sequences which have identical sequences.
#PIDs differ by 1 base will be recognized.
#if PID1 is x time greater than PID2, PID2 will be disgarded

def filter_similar_pid(sequence_hash = {}, cutoff = 10)
  seq = sequence_hash
  uni_seq = seq.values.uniq
  uni_seq_pid = {}
  uni_seq.each do |k|
    seq.each do |name,s|
      name = name[1..-1]
      if k == s
        if uni_seq_pid[k]
          uni_seq_pid[k] << [name.split("_")[0],name.split("_")[1]]
        else
          uni_seq_pid[k] = []
          uni_seq_pid[k] << [name.split("_")[0],name.split("_")[1]]
        end
      end
    end
  end

  dup_pid = []
  uni_seq_pid.values.each do |v|
```

```
    next if v.size == 1
    pid_hash = Hash[v]
    list = pid_hash.keys
    list2 = Array.new(list)
    pairs = []

    list.each do |k|
      list2.delete(k)
      list2.each do |k1|
        pairs << [k,k1]
      end
    end

    pairs.each do |p|
      pid1 = p[0]
      pid2 = p[1]
      if two_pid_x_base_different(pid1,pid2,1)
        n1 = pid_hash[pid1].to_i
        n2 = pid_hash[pid2].to_i
        if n1 >= cutoff * n2
          dup_pid << pid2
          #puts pid1 + "\t" + n1.to_s + "\t" + pid2 + "\t" + n2.to_s
        elsif n2 >= cutoff * n1
          dup_pid << pid1
          #puts pid2 + "\t" + n2.to_s + "\t" + pid1 + "\t" + n1.to_s
        end
      end
    end
  end

  new_seq = {}
  seq.each do |name,s|
    pid = name.split("_")[0][1..-1]
    unless dup_pid.include?(pid)
      new_seq[name] = s
    end
  end
  return new_seq
end

#compare two primer ID sequences. If they differ in x base, return boolean
value "TURE", else, return boolean value "FALSE"
def two_pid_x_base_different(pid1="",pid2="", x=0)
  l = pid1.size
  m = l - x
  n = 0
  if pid1.size != pid2.size
    return false
  else
    (0..(pid1.size - 1)).each do |k|
      if pid1[k] == pid2[k]
        n += 1
      end
    end
    if n >= m
      return true
    else
```

```
      return false
    end
  end
end

def unzip_r(indir, f)
  r_file = indir + "/" + f
  if f =~ /.gz/
    `gzip -d #{r_file}`
    new_f = f.sub ".gz", ""
    r_file = indir + "/" + new_f
  end
  return r_file
end

#####################End of General Methods

#obtain files for two ends for the input directory
indir = ARGV[0]
libname = File.basename(indir)

files = []
Dir.chdir(indir) do
  files = Dir.glob("*")
end
r1_f = ""
r2_f = ""
files.each do |f|
  if f =~ /r1/i
    r1_f = unzip_r(indir, f)
  elsif f =~ /r2/i
    r2_f = unzip_r(indir, f)
  end
end

t = Time.now
#outdir = indir + "/consensus_out" + "_" + t.year.to_s + "_" + t.month.to_s +
"_" + t.day.to_s + "_" + t.hour.to_s + "_" + t.min.to_s
outdir = indir + "/"# + File.basename(indir)
Dir.mkdir(outdir) unless File.directory?(outdir)

temp_out = indir + "/temp_seq"
Dir.mkdir(temp_out) unless File.directory?(temp_out)

primers.each do |setname,primer_pair|
  puts "Processing " + setname
  n_all_seq = 0
  n_filter_r1 = 0
  n_filter_r2 = 0
  n_paired = 0
  forward_primer = primer_pair[0]
  reverse_primer = primer_pair[1]
  if forward_primer.match(/(N+)(\w+)$/)
    $forward_n = $1.size
    $forward_bio_primer = $2
  else
    $forward_n = 0
```

```
  $forward_bio_primer = forward_primer
end
forward_bio_primer_size = $forward_bio_primer.size
forward_starting_number = $forward_n + forward_bio_primer_size
reverse_primer.match(/(N+)(\w+)$/)
reverse_n = $1.size
$reverse_bio_primer = $2
reverse_bio_primer_size = $reverse_bio_primer.size
$ignore_first_nt ? id_l = reverse_n - 1 : id_l = reverse_n
reverse_starting_number = reverse_n + reverse_bio_primer_size

def filter_r2(input_file,id_l=8)
  ref = primer_match($reverse_bio_primer)
  l = ref.size
  count = 0
  sequence_a = []
  sequence_h = {}

  File.open(input_file,'r') do |file|
    file.readlines.collect do |line|
      count +=1
      count_m = count % 4
      if count_m == 1
        line.tr!('@','>')
        sequence_a << line.chomp
      elsif count_m == 2
        sequence_a << line.chomp
      end
    end
  end

  sequence_h = array_to_hash(sequence_a)
  sequence_passed = {}
  $ignore_first_nt ? id_l_for_primer = id_l + 1 : id_l_for_primer = id_l
  sequence_h.each do |name,seq|
    next if seq[1..-2] =~ /N/
    next if seq =~ /A{11}/
    next if seq =~ /T{11}/

    primer = seq[id_l_for_primer,l]
    if primer =~ /#{ref}/
      sequence_passed[name] = seq
    end
  end
  return sequence_passed
end

def filter_r1(input_file)
  ref = primer_match($forward_bio_primer)
  l = ref.size
  count = 0
  sequence_a = []
  sequence_h = {}

  File.open(input_file,'r') do |file|
    file.readlines.collect do |line|
      count +=1
```

```
      count_m = count % 4
      if count_m == 1
        line.tr!('@','>')
        sequence_a << line.chomp
      elsif count_m == 2
        sequence_a << line.chomp
      end
    end
  end

  sequence_h = array_to_hash(sequence_a)
  n = sequence_h.size
  sequence_passed = {}
  sequence_h.each do |name,seq|
    next if seq[1..-2] =~ /N/
    next if seq =~ /A{11}/
    next if seq =~ /T{11}/

    primer = seq[$forward_n,l]
    if primer =~ /#{ref}/
      sequence_passed[name] = seq
    end
  end
  return [sequence_passed,n]
end

puts "Filtering R1...."
r1_temp = filter_r1(r1_f)

filtered_r1_h = r1_temp[0]
n_all_seq = r1_temp[1]
print "The number of raw sequences is #{n_all_seq.to_s}\n"

n_filter_r1 = filtered_r1_h.size
puts "Filtering R2...."
filtered_r2_h = filter_r2(r2_f,id_l)
n_filter_r2 = filtered_r2_h.size

print "R1: #{n_filter_r1}\n"
print "R2: #{n_filter_r2}\n"

puts "Pairing...."
sequence_rtag1 = {}
sequence_rtag2 = {}

filtered_r1_h.each do |k,v|
  k =~ /\s/
  k2 = $`
  sequence_rtag1[k2]= v
end

filtered_r2_h.each do |k,v|
  k =~ /\s/
  k2 = $`
  sequence_rtag2[k2]= v
end
```

```
  keys = sequence_rtag1.keys & sequence_rtag2.keys

  paired_r1 = {}
  paired_r2 = {}

  keys.each do |k|
    paired_r1[k] = sequence_rtag1[k]
    paired_r2[k] = sequence_rtag2[k]
  end

  n_paired = keys.size
  puts "Paired raw sequences are : #{n_paired.to_s}"

  #create a temp file. Temp file contains sequence names, primer ids, and
sequences from two ends
  puts "Create Temp File...."

  temp_file = temp_out + "/temp_file_" + setname
  temp_file_out = File.open(temp_file,'w')

  #building hashes for Primer ID, and two end sequences
  id = {}
  bio_id = {}
  bio_non_id = {}

  $ignore_first_nt ? id_truncate = 1 : id_truncate = 0
  paired_r2.each do |k,r2_seq|
    r1 = paired_r1[k]
    id[k] = r2_seq[id_truncate,id_l]
    bio_id[k] = r2_seq[reverse_starting_number..-2]
    bio_non_id[k] = r1[forward_starting_number..-2]
    temp_file_out.print k+ "\n" + id[k] + "\n" + bio_id[k] +
"\n"+bio_non_id[k] + "\n"
  end
  temp_file_out.close

  #hashes of Primer ID list and Primer ID distribution
  primer_id_list = {}
  primer_id_dis = {}

  puts "Calculate consensus cutoff...."
  #count primer ID
  primer_id_list = id.values
  primer_id_count = count(primer_id_list)
  #Primer ID distribution
  primer_id_dis = count(primer_id_count.values)
  primer_id_in_use = {}

 #calculate distinct_to_raw
  distinct_to_raw = (primer_id_count.size/primer_id_list.size.to_f).round(3)
  #define consensus cutoff
  #in case very little raw sequences, i.e. less than 5 unique PIDs. ignore
this set and move to the next set.
  if primer_id_dis.keys.size < 5
    File.unlink(temp_file)
    next
  end
```

```
max_id = primer_id_dis.keys.sort[-5..-1].mean
n = calculate_cut_off(max_id)
puts "Consensus cutoff is #{n}"
puts "Creating consensus..."

#Pick primer ID over threshold n
primer_id_count_over_n = []
primer_id_count.each do |primer_id,count|
  primer_id_count_over_n << primer_id if count > n
end
nn = primer_id_count_over_n.size
puts "Number of consensus to process: #{nn}"

#output part 1
out_dir_set = outdir + "/" + setname
Dir.mkdir(out_dir_set) unless File.directory?(out_dir_set)
out_dir_consensus = out_dir_set + "/consensus"
Dir.mkdir(out_dir_consensus) unless File.directory?(out_dir_consensus)

outfile_id = out_dir_consensus + "/r2.txt"
outfile_non_id = out_dir_consensus + "/r1.txt"

f1 = File.open(outfile_id,'w')
f2 = File.open(outfile_non_id,'w')

outdir_primer_id = out_dir_set + "/primer_id"
Dir.mkdir(outdir_primer_id) unless File.directory?(outdir_primer_id)

outfile_primer_id_count = outdir_primer_id + "/primer_id_count"
outfile_primer_id_dis = outdir_primer_id + "/primer_id_dis"
outfile_primer_id_in_use = outdir_primer_id + "/primer_id_in_use"

f3 = File.open(outfile_primer_id_count,'w')
f4 = File.open(outfile_primer_id_dis,'w')
f5 = File.open(outfile_primer_id_in_use,'w')

f3.print "Primer ID List and Counts\n\n"
f3.print "Primer ID\tCounts\n"

primer_id_count.each do |k,v|
  f3.print k + "\t" + v.to_s + "\n"
end
f3.close

f4.print "Primer ID Frequence\n\n"
f4.print "Frequence\tCounts\n"
primer_id_dis.keys.sort.each do |c|
  w = primer_id_dis[c]
  f4.print c.to_s + "\t" + w.to_s + "\n"
end
f4.close
#output part 2

#List of sequences with same primer ID over n.Create consensus
id_hash2 = {}
id.each do |name,pid|
  if id_hash2[pid]
```

```
      id_hash2[pid] << name
    else
      id_hash2[pid] = []
      id_hash2[pid] << name
    end
  end
  consensus = {}
  m = 0
  primer_id_count_over_n.each do |primer_id|
    m += 1
    puts "Now processing number #{m}" if m%100 == 0
    seq_with_same_primer_id = id_hash2[primer_id]

    list_id_part = []
    list_non_id_part = []
    seq_with_same_primer_id.each do |seq_name|
      id_part = bio_id[seq_name]
      non_id_part = bio_non_id[seq_name]
      list_id_part << id_part
      list_non_id_part << non_id_part
    end
    #consensus name including the Primer ID and number of raw sequences of
that Primer ID, library name and setname.
    consensus_name = ">" + primer_id + "_" +
seq_with_same_primer_id.size.to_s + "_" + libname + "_" + setname
    consensus_id_part = consensus_without_alignment(list_id_part)
    consensus_non_id_part = consensus_without_alignment(list_non_id_part)
    #consensus name including the Primer ID and number of raw sequences of
that Primer ID
    next if consensus_id_part =~ /[^ATCG]/
    next if consensus_non_id_part =~ /[^ATCG]/
    #get reverse complement sequence of the R2 region
    consensus_id_part.reverse!.tr!('ATCG','TAGC')
    primer_id_in_use[primer_id] = seq_with_same_primer_id.size
    consensus[consensus_name] = [consensus_id_part,consensus_non_id_part]
  end

  consensus_filtered = {}
  r1_consensus = {}
  r2_consensus = {}
  consensus.each do |seq_name,seq|
    r1_consensus[seq_name] = seq[1]
    r2_consensus[seq_name] = seq[0]
  end
  consensus_number_temp = consensus.size

  max_pid_comb = 4**id_l

  if consensus_number_temp < 0.003*max_pid_comb
    puts "Applying PID post consensus filter..."
    r1_consensus_filtered = filter_similar_pid(r1_consensus,10)
    r2_consensus_filtered = filter_similar_pid(r2_consensus,10)
    common_pid = r1_consensus_filtered.keys & r2_consensus_filtered.keys
    common_pid.each do |pid|
      consensus_filtered[pid] =
[r2_consensus_filtered[pid],r1_consensus_filtered[pid]]
    end
```

```
  else
    consensus_filtered = consensus
  end

  n_con = consensus_filtered.size
  puts "Number of consensus sequences:\t" + n_con.to_s
  #output part 2
  consensus_filtered.each do |seq_name,seq|
    f1.print seq_name + "_r2\n" + seq[0] + "\n"
    f2.print seq_name + "_r1\n" + seq[1] + "\n"
  end

  f1.close
  f2.close


  f5.print "Primer ID used to create consensus\n\n"
  f5.print "Primer ID\tCounts\n"
  primer_id_in_use.each do |k,v|
    f5.print k + "\t" + v.to_s + "\n"
  end
  f5.close

  #output log file
  log = out_dir_set + "/log.txt"

  log_f = File.open(log,'w')

  log_f.print "Primer ID pair-end consensus creator Version #{ver}\n\n"

  log_f.print "Primer ID pair-end consensus creator\n\n"

  log_f.print "Runtime: #{t}\n\n"

  log_f.print "Primer set name:\n#{setname}\n\n"

  log_f.puts "Forward primer sequence:\t" + forward_primer
  log_f.puts "Reverse primer sequence:\t" + reverse_primer

  log_f.print "\nNumber of Raw Sequences for each end is: #{n_all_seq}\n\n"

  log_f.print "Number of R1 passed filtered is: #{n_filter_r1}\n\n"

  log_f.print "Number of R2 passed filtered is: #{n_filter_r2}\n\n"

  log_f.print "Number of sequences paired is: #{n_paired}\n\n"

  log_f.print "The consensus threshold is #{n}.\n\n"

  log_f.print "Length of Primer ID is #{id_l.to_s}.\n\n"

  log_f.print "The number of consensus sequences process (including
ambiguities) is #{nn}\n\n"

  log_f.print "The number of consensus sequences is #{n_con}\n\n"

  log_f.print "Distinct Primer ID to raw is #{distinct_to_raw}\n\n"
```

```
  log_f.print "Resampling Parameter is #{(n_con/nn.to_f).round(3)}\n\n"

  if distinct_to_raw > 0.1
    log_f.print "WARNING: NOT ENOUGH RAW SEQUENCES, SAMPLING DEPTH MAY NOT BE
REVEALED!!!"
    print "\t\t\t**************************\nWARNING: NOT ENOUGH RAW
SEQUENCES, SAMPLING DEPTH MAY NOT BE
REVEALED!!!\n\t\t\t**************************\n"
  end

  log_f.close
end

# outdir_tar = outdir + ".tar.gz"

# if File.exists?(outdir_tar)
#   File.unlink(outdir_tar)
# end
# Dir.chdir(indir) {print `tar -czf #{File.basename(outdir_tar)}
#{File.basename(outdir)}`}

# print `rm -rf #{outdir}`
print `rm -rf #{r1_f}`
print `rm -rf #{r2_f}`
print `rm -rf #{temp_out}`
```

Assembling consensus r1 and r2 files into single reads

File name: pid_join_two_ends_33.rb (does not need any modification if you're using MHV20814. If you're using a different amplicon, you will have to change the overlap. Script written by Shuntai Zhou (Zhou et al., 2015).

```
# Script to join Primer ID R1 and R2 sequences
# example:
# ruby pid_join_two_ends.rb target_directory


# target_directory contains library subdirectories, each has one r1 and one
r2 file in it.
# example:
# target_directory
# ├──lib1
# │     lib1_r1.txt
# │     lib1_r2.txt
# ├──lib2
# │     lib1_r2.txt
# │     lib1_r2.txt
# ...

# The overlap number needs to be pre-determined. if overlap == 0, r1 and r2
will be simply joined at the end.
overlap = 33


#############
def fasta_to_hash(infile)
  f=File.open(infile,"r")
  return_hash = {}
  name = ""
  while line = f.gets do
    if line =~ /^\>/
      name = line.chomp
      return_hash[name] = ""
    else
      return_hash[name] += line.chomp.upcase
    end
  end
  f.close
  return return_hash
end

indir = ARGV[0]
if indir =~ /\/$/
  indir = indir[0..-2]
end

outdir = indir + "_combined"
Dir.mkdir(outdir) unless File.directory?(outdir)
libs = Dir[indir + "/*"]

libs.each do |lib|
  outfile = outdir + "/" + File.basename(lib)
```

```
  out = File.open(outfile, "w")
  files = Dir[lib + "/*"]
  r1_file = ""
  r2_file = ""
  files.each do |f|
    if File.basename(f) =~ /r1/
      r1_file = f
      next
    elsif File.basename(f) =~ /r2/
      r2_file = f
      next
    end
  end
  seq1 = fasta_to_hash(r1_file)
  seq2 = fasta_to_hash(r2_file)
  pid_seq1 = {}
  seq1.each do |k,v|
      pid = k.split("_")[0][1..-1]
      pid_seq1[pid] = k
  end
  pid_seq2 = {}
  seq2.each do |k,v|
      pid = k.split("_")[0][1..-1]
      pid_seq2[pid] = k
  end

  pid_seq1.each do |k,v|
      fwd = seq1[v]
      rev = seq2[pid_seq2[k]]
      next unless (fwd and rev)
      combined_seq = ""
      if overlap == 0
        combined_seq = fwd.tr("-","") + rev.tr("-","")
      elsif fwd[-overlap..-1] == rev[0, overlap]
        combined_seq = fwd + rev[overlap..-1]
      else
        next
      end
      out.puts v.gsub(/_r1/, "")
      out.puts combined_seq
  end
  out.close
end
```

Counting the number of each base at every site

File name: counts.py (does not need any modification). Script written by Matthew Pauly (Pauly et al., 2017b).

```
import sys
from collections import Counter
import re

# Create the empty lists to count nucleotides
pos=[]
ref=[]
cov=[]
A=[]

T=[]

G=[]

C=[]
N=[]

with open(sys.argv[1],'r') as pileup:
    for line in pileup:
        line = re.split(r'\t+', line) # split at tab
        pos.append(str(line[1])) # position
        ref.append(str(line[2])) # Reference base
        cov.append(str(line[3])) # Coverage
        counter=Counter(line[4]) # count the occurrence of each character in
the bases string

        A.append(str(counter['A'])) # how many A's

        T.append(str(counter['T'])) # How many T's

        G.append(str(counter['G']))

        C.append(str(counter['C']))
        N.append(str(counter['N']))


total_positions=len(pos)        #how many positions did we look at
with open(sys.argv[2],'w') as out:
    out.write("pos,ref,cov,A,T,G,C,N\n") # Write the header to the out file
    for i in range(0,total_positions) :

out.write(pos[i]+","+ref[i]+","+cov[i]+","+A[i]+","+T[i]+","+G[i]+","+C[i]+",
"+N[i]+"\n") # write each line to the outfile
```

Counting each type of mutation

File name: Mutcount_analysis_MHV20814.R. Script written by Matthew Pauly (Pauly et al., 2017b) and modified by Ben Reisman (Vanderbilt MSTP). Input is the .csv file derived from counts.py.

```
mutcount <- function(A) {
  n <- nrow(A)
  con <- matrix(nrow = c(n), ncol = c(1))

  for (j in 1:n)
  {
    if (A[j, 4] > 1000) {
      con[j, 1] <- c(1)
    }

    if (A[j, 5] > 1000) {
      con[j, 1] <- c(2)
    }
    if (A[j, 6] > 1000) {
      con[j, 1] <- c(3)
    }
    if (A[j, 7] > 1000) {
      con[j, 1] <- c(4)
    }
  }

  mut <- matrix(c(0), nrow = c(n), ncol = c(12))

  for (i in 1:n) {
    if (con[i, 1] == 1) {
      mut[i, 1] <- A[i, 7]
      mut[i, 2] <- A[i, 6]
      mut[i, 3] <- A[i, 5]
    }
    if (con[i, 1] == 2) {
      mut[i, 10] <- A[i, 4]
      mut[i, 11] <- A[i, 7]
      mut[i, 12] <- A[i, 6]
    }
    if (con[i, 1] == 3) {
      mut[i, 7] <- A[i, 4]
      mut[i, 8] <- A[i, 7]
      mut[i, 9] <- A[i, 5]
    }
    if (con[i, 1] == 4) {
      mut[i, 4] <- A[i, 4]
      mut[i, 5] <- A[i, 6]
      mut[i, 6] <- A[i, 5]
    }
  }

  mutnam <-
    c("A2C",
      "A2G",
      "A2T",
```

```
      "C2A",
      "C2G",
      "C2T",
      "G2A",
      "G2C",
      "G2T",
      "T2A",
      "T2C",
      "T2G")

  mutations <- matrix(c(0), nrow = c(1), ncol = c(12))

  colnames(mutations) <- mutnam

  for (i in 1:12) {
    mutations[1, i] <- sum(mut[, i])
  }
  return(mutations)
}



raw_data <- read.csv("Your_file.csv")
mutcount_output <- mutcount(raw_data)
mutcount_output
```

B3. Competitive fitness assay for murine hepatitis virus

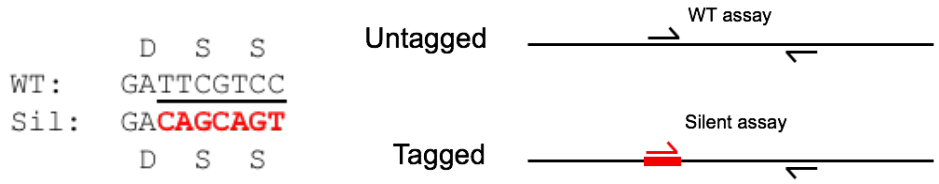
Purpose:
This assay quantifies the competitive fitness of mutant viruses. All experimental viruses are co-cultured with a reference genome tagged with synonymous mutations, allowing discrimination by qPCR. The change in the competitor-to-reference ratio over time reflects their relative fitness. Relevant references for this approach (but not necessarily the chemistry) include:

1. **Carrasco P**, **Daròs JA**, **Agudelo-Romero P**, **Elena SF**. 2007. A real-time RT-PCR assay for quantifying the fitness of tobacco etch virus in competition experiments. Journal of Virological Methods **139**:181–188.
2. **Visher E**, **Whitefield SE**, **McCrone JT**, **Fitzsimmons W**, **Lauring AS**. 2016. The Mutational Robustness of Influenza A Virus. PLoS Pathog **12**:e1005856–25.
3. **Fitzsimmons WJ**, **Woods RJ**, **McCrone JT**, **Woodman A**, **Arnold JJ**, **Yennawar M**, **Evans R**, **Cameron CE**, **Lauring AS**. 2018. A speed–fidelity trade-off determines the mutation rate and virulence of an RNA virus. PLoS Biol **16**:e2006459–20.
4. **Moratorio G**, **Henningsson R**, **Barbezange C**, **Carrau L**, **Bordería AV**, **Blanc H**, **Beaucourt S**, **Poirier EZ**, **Vallet T**, **Boussier J**, **Mounce BC**, **Fontes M**, **Vignuzzi M**. 2017. Attenuation of RNA viruses by redirecting their evolution in sequence space. Nat Microbiol **2**:1–12.
5. **Bordería AV**, **Isakov O**, **Moratorio G**, **Henningsson R**, **Agüera-González S**, **Organtini L**, **Gnädig NF**, **Blanc H**, **Alcover A**, **Hafenstein S**, **Fontes M**, **Shomron N**, **Vignuzzi M**. 2015. Group Selection and Contribution of Minority Variants during Virus Adaptation Determines Virus Fitness and Phenotype. PLoS Pathog **11**:e1004838.

Reagents and kits:

1. Sample collection
    a. Complete DMEM
    b. PBS +/+ (containing calcium and magnesium)
2. RNA extraction:
    a. QIAamp Virus 96 QIAcube HT kit (Qiagen # 57731)
    b. QIAcube HT plasticware (Qiagen # 950067)
    c. Buffer AVL (Qiagen #19073)
    d. cRNA, 1350ug (Qiagen #1017647)
    e. 100% ethanol
3. RT-qPCR
    a. Nuclease-free water
    b. RNaseZap (Sigma #R2020)
    c. Power SYBR green RNA-to-Ct 1-step kit (ThermoFisher #4389986).

Primer Design:
The reference genome is tagged with a run of 7 consecutive synonymous substitutions within nsp2 (nt 1301-1307 for MHV) that permit selective amplification of the competitor (WT, below) and reference viruses. I have designed this for MHV such that the reverse primer (also the reverse transcription primer) is the same for both assays, with the forward primer discriminating each species. Validation data can be provided upon request.



MHV Primers:

SS-WT Assay			
Direction	Primer Name	Primer Sequence	Final [primer]
Forward	SS-qPCR-WT-F	CTATGCTGTATACGGA**TTCGTCC**	450 nM
Reverse	SS-qPCR-R2	GGTGTCACCACAACAATCCAC	450 nM

SS-Sil Assay			
Direction	Primer Name	Primer Sequence	Final [primer]
Forward	SS-qPCR-Sil-F	CTATGCTGTATACGGA**CAGCAGT**	200 nM
Reverse	SS-qPCR-R2	GGTGTCACCACAACAATCCAC	200 nM

I optimized the concentrations by running a matrix for each assay and primer set, evaluating them according to (1) earliest Ct for the target and (2) latest Ct for the opposing target. Data available on request.

I prepared and stored the primer sets as mixed aliquots (i.e. both forward and reverse) at 10x concentration. For SS-WT, each primer is at 4.5µM; for SS-Sil, each primer is at 2µM. Thus, use 2µL of primer mix per 20µL reaction.

Viruses:
I engineered ExoN(+) (i.e. WT) and ExoN(-) viruses containing the synonymous tags that serve as the **reference**. The name "SS" refers to the substitutions at serines—the rest should be self-explanatory.

1. SS-XN(+)
2. SS-XN(-)

Experimental Considerations:

1. Technical and biological consistency is *essential* for this assay. This includes all aspects of the design and execution, from timing, to pipetting (always with clean tips!), to harvesting, to qPCR.
2. Choice of MOI and well size is important! You'll be passaging mixtures of viruses – essentially, imposing a genetic bottleneck. If your population size is too small, then the experimental outcome can be altered by genetic drift. With larger populations, the outcome will depend more upon the fitness of individual variants. Thus, we want to maintain a "higher" MOI in larger plate. The below conditions have worked well for me:
    a. MOI = 0.1 PFU/cell total (MOI = 0.05 PFU/cell each of competitor and reference.
    b. 12-well plates
    c. Triplicate lineages for each experimental virus (i.e. **competitor**)
    d. I make sure that all competitors are present on each plate. If ≤4 viruses, I put them

all on one plate. If >4, I do one well per competitor per plate (i.e. 3 plates total). This reduces the chance of plate-specific bias.

3. We could do this assay as a single-passage experiment, in line with some reports in the literature. However, in my opinion, performing multiple passages better reflects fitness differences over time, while a single-passage is prone to variability based on stock quality, input ratio, etc.
4. Always include an apples-to-apples control. That is, if you're competing against the SS-XN(+), always include a wild-type control. For SS-XN(-), the control is MHV-ExoN(-).

Protocol:

Day -1: Prepare warm base for assay plates.

1. Fill one T150 per experimental plate with dH2O and place in incubator. (Cover the mouth with cellophane before capping to block the paper filter). These take a long time to warm up.
2. I usually keep my T150s in the incubator all the time so they're always ready.

Day 1: Plate cells for the assay

1. Prepare cells as you normally would for splitting.
2. Count cells using the Millipore Scepter 2.0 Handheld Automated Cell counter.
    a. Dilute resuspended cells 1:10 in PBS+/+ (100uL cells + 900uL PBS+/+)
    b. Mix well and measure twice, then average the measurements
3. Dilute cells to 1 x 10^5 cells/mL in complete DMEM.
4. Using a repeater pipette (I like the 10mL size for 12-well plates), add 1 mL cells.
5. Carefully place each plate on its own pre-warmed T150 in the incubator. *Note the time – you will inoculate with virus exactly 24 hours later.*
6. Prepare for day 2:
    a. Set out and label all sample collection tubes.
    b. Prep your calculations. Note: I inoculate 12-well plates with 400uL of virus.

Day 2: Infection

1. Several hours before you start, put 50mL conical tubes containing DMEM and PBS+/+ into the incubator to warm up. If you're behind or forget, you can put them in the bead bath to get them started (but don't overdo it, you don't want them too hot).
2. Complete these steps on ice:
    a. Place cells on ice 10 minutes before inoculation.
    b. Prepare virus dilutions.
    c. Aspirate medium and inoculate cells with virus. Do this one plate at a time.
    d. Transfer excess inoculum to a 1.5mL Eppendorf tube – this will be the "P0" or "input" in the analysis.
    e. Allow virus to adsorb for 30 min on ice. *Do not go longer—D9s don't like the cold very much.*
    f. Rock once after 15 min.
3. Transfer plates to their designated T150 in the 37C incubator.
    a. From this point on, do not remove the plates from their T150.
    b. If you want to be super super rigorous, add one plate at a time to the incubator at ~3-minute intervals. This builds you some buffer time for sample collection later.

4. Incubate at 37C for 30 min, rocking once after 15min.
5. Remove the plates from their incubator, keeping them on their T150s.
6. Wash once with 1.5 mL warm PBS +/+.
    a. Aspirate supernatant.
    b. Add 1.5mL PBS+/+
    c. *If doing more than one plate, repeat steps a,b for each plate before continuing on.*
7. Aspirate PBS+/+ and add back total 1 mL warm DMEM. (One plate at a time.)
8. Transfer back to incubator for the indicated time:
    a. SS-XN(+) reference (and WT-like competitors) – 15 hpi
    b. SS-XN(-) reference (and ExoN(-)-like competitiors) – 16 hpi
9. Set out cells for plaque assay (1.5 x 12-well plate per competitor).

Day 3: Harvest and titer:

1. Turn on refrigerated centrifuge and cool to 4C.
2. At harvest time, transfer all 1mL of supernatant to a 1.5mL Eppendorf tube.
3. Spin samples at 10k x rcf for 5 min at 4C.
4. Transfer clarified supernatant to your long-term storage tube, then remove 100uL into a separate tube.
5. Freeze the long-term and 100uL aliquots.
6. At your convenience, thaw 100uL aliquots and start plaque assay.
    a. For SS-XN(+) experiments, I plate dilutions 3-4-5.
    b. For SS-XN(-) experiments, I plate 2-3-4 for P1 and P2, then reduce to 1-2-3 of P3 and P4.
7. At your desired time, set out cells in 12-well plates to initiate P2, as described in "Day 1" above.

Day 4: Initiate second passage

1. Fixed plaques, pull plugs, and titer.
2. For each sample, calculate the volume of inoculum needed to initiate MOI = 0.1 PFU/cell infection. *Note that we are normalizing the total inoculum. We are not distinguishing the competitor and reference in any way during this step.*
3. Follow the steps above for Day 2, with the following modifications:
    a. Instead of doing virus dilutions as we did for passage 1, I aspirate off medium right after putting the D9s on ice.
    b. I add the amount of DMEM needed to make up 400uL total volume (volume DMEM = 400 - volume virus).
    c. *Note: if volume DMEM < 200, I wait until closer to the inoculation because I worry about drying the cells out.*
4. Set out cells for plaque assay (1.5 x 12-well plate per competitor).

Repeat Days 3-4 until you've completed 4 passages. You do not need to titer the P4 samples, but I do typically set aside the 100uL aliquot just in case I feel I need to continue passage.

Day X: RNA extraction

Because consistency is key for this experiment, I use the QIAcube HT for automated, high-throughput RNA extraction of all experimental samples simultaneously. I start with 70uL of

supernatant per sample, which yields plenty of RNA for qPCR.

Steps:

1. Transfer 70uL of each sample to the S-block in vertical orientation (e.g. A1→H1, A2→H2, etc)
    a. 13 total samples per competitor, including:
        i. Input/inoculum/P0 (however you refer to it)
        ii. 12 passaged samples (3 per passage)
    b. If you have empty wells in your last column, add 70uL H2O (the instrument can't handle empty wells). *You can leave whole columns unused!*
    c. I store the sample tubes on ice while performing this, but I don't put the S-block on ice or in the fridge because the purification chemistry is optimized for room-temperature solutions.
2. Turn on instrument and computer.
3. Launch **QIAcube HT Software**
4. Select the "
5. Click on the **Wizard Hat**:
    a. Designate the number of columns being run.
    b. Click **Next** to go through each of the protocol steps.
        i. Make sure the input volume is 70uL
        ii. Reagent volumes should not need to be changed.
        iii. Set elution volume to 75uL.
6. Load adapters and plasticware:

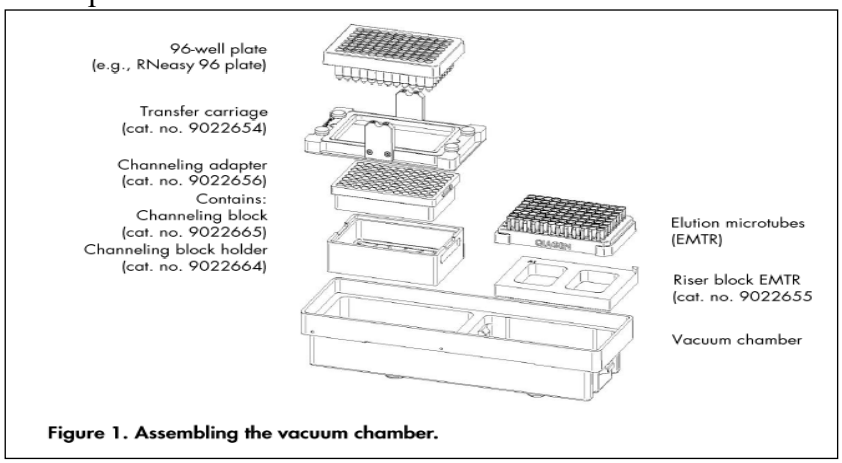


Figure 1. Assembling the vacuum chamber.

    a. Riser Block (Right/Elution side of sink)
    b. Channeling block holder and adapter (Left/Waste side of sink.)
    c. Carriage (must be pushed all the way to the LEFT over the LEFT side of the sink.)
    d. QIAamp Virus 96 plate (taped off if applicable and pushed all the way to the LEFT.) *Put film over any wells you're not using! If you don't, the vacuum will not elute your samples effectively.*
    e. Elution plate (Place on the instrument with the lid and leave on until prompted during pre-run checklist to remove. Plate must be pushed all the way to the LEFT to elute properly.)

f. Tips (with lids on until prompted during pre-run checklist to remove.)
    i. To reset number of tips:
        1. Right click over a tip position.
        2. Then select one of the options depending on what you are loading.

7. Load all buffers
    a. AW1
    b. AW2
    c. AVE
    d. 100% Ethanol
    e. cRNA mix:
        i. Dissolve lyophilized pellet in Buffer AVE.
            1. Two sizes of cRNA: 310ug and 1350ug
                a. 310uL AVE to 310ug
                b. 1350uL AVE to 1350ug
        ii. Add cRNA+AVE to buffer AVL according to below:

# samples	8	16	24	32	40	48	56	64	72	80	88	96
Buffer AVL (mL)	9.23	11.25	13.45	15.64	17.84	20.03	22.23	24.42	26.62	28.81	31.01	33.2
cRNA (uL)	188.4	229.6	274.4	319.2	364	408.8	453.6	498.4	543.2	588	632.8	677.6
Total volume (mL)	9.42	11.48	13.72	15.96	18.2	20.44	22.68	24.92	27.16	29.4	31.64	33.88

8. Gently Load the lysates into the B1 position.
9. Press "**Play**."
10. At the end of the run:
    a. Remove elution plate and cover with lid.
    b. Remove and discard the QIAamp 96 plate (unused columns can be saved and used later) and any empty tip racks.
    c. Removing only the channeling block adapter (not the holder) from the waste sink and either swap for clean one or clean the one removed using DI water to rinse followed by 70% ethanol and allow to dry before next use.)
    d. Trough re-use instructions:
        i. Troughs without salts, AW2, AVE, cRNA mix, and 100% EtOH can have their residual volume used in place of the 20ml of water requested in the waste pail menu for flushing the sink. Once empty, they can be placed upside down on a paper towel until needed again.
        ii. Troughs with salts or enzyme will need to have their residual volume disposed of via your Health and Safety regulations. Once empty, they should be rinsed well, ~3 times, with diH2O and then placed upside down on a paper towel until needed again.
11. When you close the software, it will automatically prompt a clean-up protocol. Follow the instructions, and add the UV step.
12. Spin down elution tubes in their plate briefly to bring total volume to the bottom.
13. Transfer eluted RNA to long-term storage tubes until ready for qPCR.

Day X: RT-qPCR

1. Dilute RNA:

a. Select your dilution according to the expected RNA content of reactions so that you stay within a linear range (for me, I shoot for Ct between 15-25). In my opinion, there's no need to normalize the RNA content across all samples – the relevant output is the ratio between tagged and untagged templates in a given sample, which should not vary with RNA amount.
b. For MHV samples, I use these calculations:
    i. SS-XN(+) and WT-like competitors:
        1. P0 RNA 1:10 (4uL RNA + 36uL H2O)
        2. P1-4 RNA 1:100 (2uL RNA + 198uL H2O
    ii. SS-XN(-) and ExoN(-)-like competitors:
        1. P0 RNA 1:10 (4uL RNA + 36uL H2O)
        2. P1-4 RNA 1:10 (4uL RNA + 36uL H2O)

2. Prepare master mix (same calculations for both SS-WT and SS-Sil assays). Makes enough for one plate.

Reagent	**1x (µL)**	**32x (µL)**
Primer Mix (10x)	2	64
RT (125x)	0.16	5.12
2X Power SYBR	10	320
Total	12.16 per rxn	

3. Add 12.16 uL master mix to wells first.
4. Add 7.84uL diluted RNA per well.
5. Cover plate with an optical film; make sure it's well sealed so that liquid doesn't escape and skew your results.
6. Spin briefly with the plate centrifuge before loading into the StepOne Plus.

Thermocycling conditions:

1. 48C 30min
2. 95C 10min
3. 95C 15sec
4. 60C 1min
5. To (3) x 39
6. Melt curve

	1	2	3	4	5	6	7	8	9	10	11	12
A	A-1	A-1	A-2	A-2	A-3	A-3	A-4	A-4	B-1	B-1	B-2	B-2
B	A-1	A-1	A-2	A-2	A-3	A-3	A-4	A-4	B-1	B-1	B-2	B-2
C	B-3	B-3	B-4	B-4	C-1	C-1	C-2	C-2	C-3	C-3	C-4	C-4
D	B-3	B-3	B-4	B-4	C-1	C-1	C-2	C-2	C-3	C-3	C-4	C-4
E	P0	P0	NTC	NTC								
F	P0	P0	NTC	NTC								
G												
H												

Assay: SS-WT SS-Sil

Data analysis workflow

This assay assumes that fitness differences between variants are fixed. With fixed fitness differences, the proportion of each competitor should change by the same factor with every passage under identical conditions. (This is a big reason that consistency is so critical technically). Thus, the relative fitness describes the rate of change in the proportions of competitor and reference over time.

The general workflow of the analysis is:

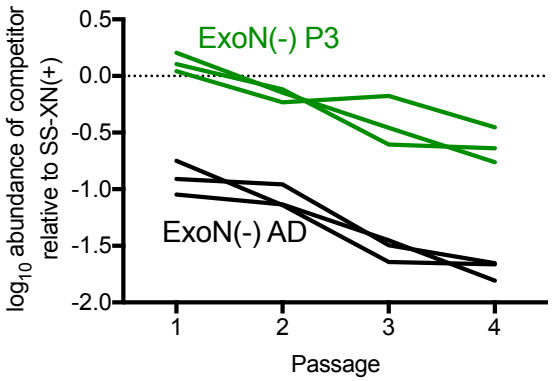
1. Determine the ratio of competitor (SS-WT assay) relative to reference (SS-Sil assay) in each sample. This is performed exactly as though you were measuring MHV genome content relative to GAPDH.
2. Plot those ratios over time.
3. The rate in change of ratio (slope of a linear regression) is related to the relative fitness.
4. Normalize all relative fitness values to the mean of the control (wild-type or MHV-ExoN(-)).

Note that this yields indirect comparisons of fitness between experimental viruses. That is, instead of direct competition between virus A and virus B, the comparisons will be "fitness of virus A relative to reference 1" vs "fitness of virus B relative to reference 1."

Detailed data analysis

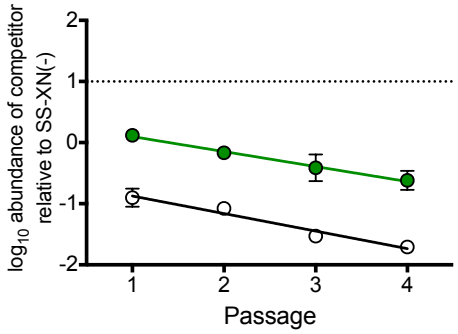
1. Average the Ct values for each assay in each sample.
2. Calculate $\Delta Ct_{competitor}$
    a. **$\Delta Ct_{competitor} = Ct_{competitor} - Ct_{reference.}$**
3. Calculate $\Delta ratio_{competitor}$
    a. PCR doubles DNA content in each cycle, so each Ct represents a 2-fold change. So if $\Delta Ct = 2$, $\Delta ratio = 2^2 = 4$; $\Delta Ct = 5$, $\Delta ratio = 2^5 = 32$)
    b. **$\Delta ratio_{competitor} = 2^{\Delta Ct\text{-}competitor}$**
4. Convert to the $\Delta \log_{10} ratio_{competitor}$
    a. Converting to the $\log_{10}$ allows use of linear regressions in the next step.
    b. $\Delta \log_{10} ratio_{competitor} = \log_{10}(\Delta ratio_{competitor})$
5. Plot $\Delta \log_{10} ratio_{competitor}$ for each replicate against passage number
    a. Plot each lineage (i.e. replicate passage) separately for the purposes of your calculations. The below example shows a table with 2 viruses (AA and AD) with 3 replicates each.
    b. I often leave out the P0 RNA from analysis because:
        i. It's not necessary, and input ratio may not be exact depending on the quality of each stock. In subsequent passages, the samples will have been treated identically, so they are easier to compare.
        ii. Including it tends to skew the linear regressions.

Table format: XY	X	Group A	Group B	Group C	Group D	Group E	Group F
	Passage	AA	AA	AA	AD	AD	AD
	X	Y	Y	Y	Y	Y	Y
3 Title	1	0.105666933	0.042075690	0.206012753	-0.909105844	-1.047598601	-0.749758689
4 Title	2	-0.117753767	-0.231432369	-0.147777646	-0.958132713	-1.134892408	-1.139695620
5 Title	3	-0.604834090	-0.175715856	-0.456838833	-1.494543639	-1.451948900	-1.643354215
6 Title	4	-0.637881543	-0.452780606	-0.761094511	-1.652480347	-1.807205727	-1.664224690



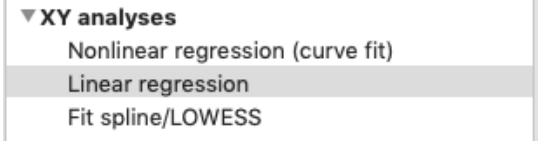
c. You can also plot as mean ± SD (that is, a column for each virus with 3 subcolumns representing individual replicates) to get a sense of how consistent your results are. However, if you try to take the relative fitness based on all the data at once, you won't be able to perform useful statistical tests. Below example is the same data as above but formatted to plot as mean ± SD

Table format: XY	X	Group A			Group B		
	Passage	AA			AD		
	X	A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
3 Title	1	0.105666933	0.042075690	0.206012753	-0.909105844	-1.047598601	-0.749758689
4 Title	2	-0.117753767	-0.231432369	-0.147777646	-0.958132713	-1.134892408	-1.139695620
5 Title	3	-0.604834090	-0.175715856	-0.456838833	-1.494543639	-1.451948900	-1.643354215
6 Title	4	-0.637881543	-0.452780606	-0.761094511	-1.652480347	-1.807205727	-1.664224690

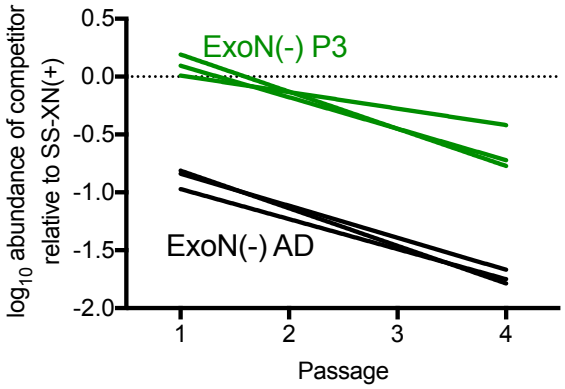


d. **NOTE:** The exact values of $\log_{10}$ratio are *not important to interpretation*. What matters is the slope of the linear regression!

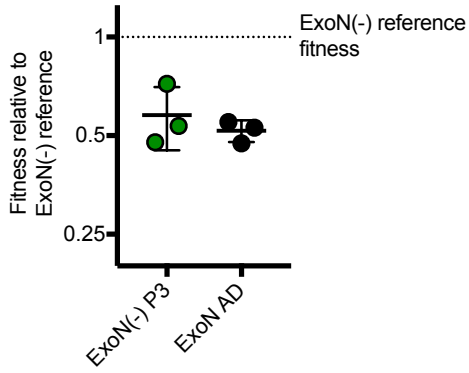
6. Perform linear regression on each individual lineage.
   a. You will not need to change analysis parameters in Prism.



b. The regression will automatically report whether the slope is statistically significant from zero. This sounds exciting but really doesn't matter. All relative fitness values will be normalized to control, and the relevant comparison is between the normalized viral fitnesses.



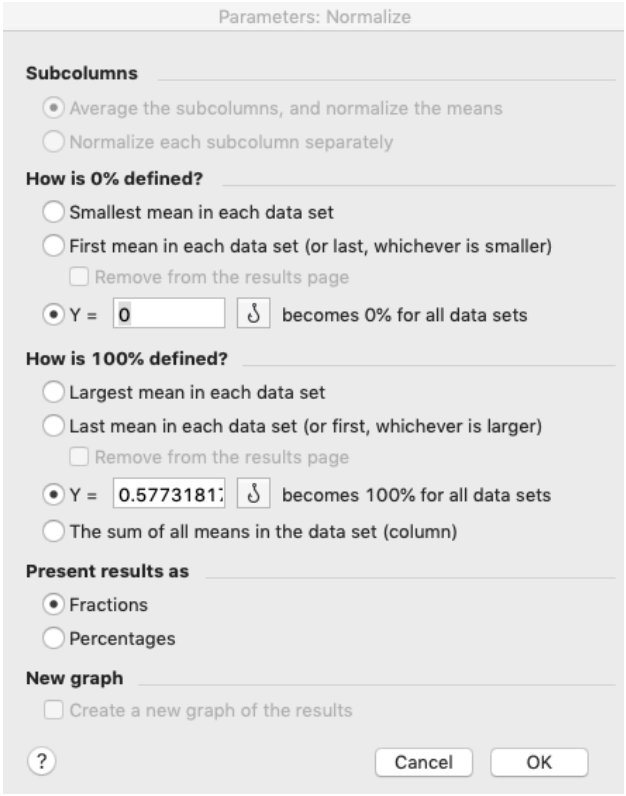
7. Using the calculated slope, convert to relative fitness and plot.
   a. Relative fitness = 10^{slope}

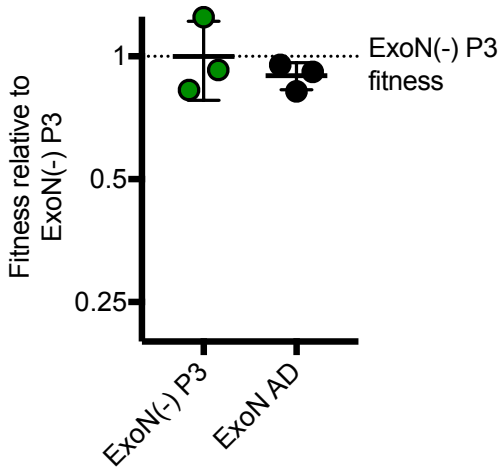


8. Normalize the relative fitness values to the mean of control.
   a. Easiest way to do this is to first plot the raw relative fitness values in Prism:

	Group A	Group B
	AA	AD
	Y	Y
1	0.534810592	0.528810415
2	0.719614656	0.550047253
3	0.477529274	0.473478213

   b. Then, take the mean of your control (in this case, AA) and use it to normalize all your data (including the control).





9. Statistics:
    a. If your experiment is n = 3, use nonparametric tests.
        i. 2 viruses: Mann-Whitney test
        ii. 3+ viruses:
            1. Kruskal-Wallis test
            2. Perform multiple comparisons (that is, compare each experimental virus to your control)
    b. If your experiment is n ≥ 6, you can use parametric tests:
        i. 2 viruses: student's t test

ii. 3+ viruses:
    1. One-way ANOVA
    2. Perform multiple comparisons (that is, compare each experimental virus to your control)

10. Celebrate, then interpret data!

APPENDIX C. Proofreading-deficient coronaviruses adapt for increased fitness over long-term passage without reversion of exoribonuclease-inactivating mutations

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

Proofreading-Deficient Coronaviruses Adapt for Increased Fitness over Long-Term Passage without Reversion of Exoribonuclease-Inactivating Mutations

Kevin W. Graepel,[a,b,c] Xiaotao Lu,[b,c] James Brett Case,[a,b,c] Nicole R. Sexton,[a,b,c] Everett Clinton Smith,[b,c,d] Mark R. Denison[a,b,c]

Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee, USA[a]; Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee, USA[b]; Elizabeth B. Lamb Center for Pediatric Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA[c]; Department of Biology, the University of the South, Sewanee, Tennessee, USA[d]

ABSTRACT The coronavirus (CoV) RNA genome is the largest among the single-stranded positive-sense RNA viruses. CoVs encode a proofreading 3′-to-5′ exoribonuclease within nonstructural protein 14 (nsp14-ExoN) that is responsible for CoV high-fidelity replication. Alanine substitution of ExoN catalytic residues [ExoN(-)] in severe acute respiratory syndrome-associated coronavirus (SARS-CoV) and murine hepatitis virus (MHV) disrupts ExoN activity, yielding viable mutant viruses with defective replication, up to 20-fold-decreased fidelity, and increased susceptibility to nucleoside analogues. To test the stability of the ExoN(-) genotype and phenotype, we passaged MHV-ExoN(-) 250 times in cultured cells (P250), in parallel with wild-type MHV (WT-MHV). Compared to MHV-ExoN(-) P3, MHV-ExoN(-) P250 demonstrated enhanced replication and increased competitive fitness without reversion at the ExoN(-) active site. Furthermore, MHV-ExoN(-) P250 was less susceptible than MHV-ExoN(-) P3 to multiple nucleoside analogues, suggesting that MHV-ExoN(-) was under selection for increased replication fidelity. We subsequently identified novel amino acid changes within the RNA-dependent RNA polymerase and nsp14 of MHV-ExoN(-) P250 that partially accounted for the reduced susceptibility to nucleoside analogues. Our results suggest that increased replication fidelity is selected in ExoN(-) CoVs and that there may be a significant barrier to ExoN(-) reversion. These results also support the hypothesis that high-fidelity replication is linked to CoV fitness and indicate that multiple replicase proteins could compensate for ExoN functions during replication.

IMPORTANCE Uniquely among RNA viruses, CoVs encode a proofreading exoribonuclease (ExoN) in nsp14 that mediates high-fidelity RNA genome replication. Proofreading-deficient CoVs with disrupted ExoN activity [ExoN(-)] either are nonviable or have significant defects in replication, RNA synthesis, fidelity, fitness, and virulence. In this study, we showed that ExoN(-) murine hepatitis virus can adapt during long-term passage for increased replication and fitness without reverting the ExoN-inactivating mutations. Passage-adapted ExoN(-) mutants also demonstrate increasing resistance to nucleoside analogues that is explained only partially by secondary mutations in nsp12 and nsp14. These data suggest that enhanced resistance to nucleoside analogues is mediated by the interplay of multiple replicase proteins and support the proposed link between CoV fidelity and fitness.

KEYWORDS RNA virus, adaptive evolution, competitive fitness, coronavirus, exoribonuclease, plus-strand RNA virus, proofreading, replication fidelity

Received 18 August 2017 **Accepted** 10 October 2017 **Published** 7 November 2017

Citation Graepel KW, Lu X, Case JB, Sexton NR, Smith EC, Denison MR. 2017. Proofreading-deficient coronaviruses adapt for increased fitness over long-term passage without reversion of exoribonuclease-inactivating mutations. mBio 8:e01503-17. https://doi.org/10.1128/mBio.01503-17.

Editor Kanta Subbarao, NIAID, NIH

Address correspondence to Everett Clinton Smith, ecsmith@sewanee.edu, or Mark R. Denison, mark.denison@vanderbilt.edu.

A paradigm of RNA virus biology is error-prone genomic replication due to the lack of proofreading or postreplicative RNA repair mechanisms (1–3). Decreased replication fidelity may constrain RNA genome size and complexity and risks the accumulation of deleterious mutations leading to population extinction (4–7). While genetic diversity allows viral populations to adapt rapidly under selective pressure, many mutations are neutral or detrimental to viral fitness (8–12). Research performed with many RNA viruses supports the hypothesis that the mutation rate of RNA virus replicases has evolved to balance multiple characteristics of the viral population such as genetic diversity, genomic integrity, and virulence. High- or low-fidelity variants are described for many RNA viruses infecting animals, including the coronaviruses (CoVs) murine hepatitis virus (MHV-A59) and severe acute respiratory syndrome-associated coronavirus (SARS-CoV) (13–17), as well as foot-and-mouth disease virus (18–22), poliovirus (23–29), Chikungunya virus (30, 31), influenza virus (32), coxsackievirus B3 (33, 34), and human enterovirus 71 (35–37). Most altered-fidelity variants described to date harbor mutations within the viral RNA-dependent RNA polymerase (RdRp), are attenuated *in vivo*, and protect against reinfection, highlighting their potential utility as live attenuated vaccines (24, 28, 29, 38, 39). Those studies underscored the importance of understanding the molecular mechanisms by which RNA viruses regulate their replication fidelity.

Viruses in the *Coronavirinae* subfamily have large single-stranded positive-sense RNA genomes [(+)ssRNA] (40), ranging between 26 and 32 kb in length (41). CoVs encode a 3′-to-5′ exoribonuclease (ExoN) in the N-terminal half of nonstructural protein 14 (nsp14-ExoN) (42, 43). CoV ExoN activity depends on conserved magnesium-coordinating acidic amino acids in three motifs (DE-E-D) that together constitute the active site (Fig. 1) (44). The CoV ExoN is grouped with the DE-D-Dh superfamily of exonucleases involved in proofreading during prokaryotic and eukaryotic DNA replication (42–46). Alanine substitution of CoV motif I DE residues (DE-to-AA) reduces biochemical ExoN activity in SARS-CoV (44, 46) and human coronavirus 229E (42). MHV-A59 and SARS-CoV lacking ExoN activity [ExoN(-)] have mutation frequencies 8-fold to 20-fold greater than are seen with WT viruses and are highly susceptible to the activity of nucleoside analogues (13–17, 38). Thus, all available data to date support the hypothesis that nsp14-ExoN is the first known proofreading enzyme encoded by an RNA virus.

Despite the critical role of ExoN in virus replication, fidelity, fitness, and virulence, reversion of the ExoN-inactiviting substitutions (Fig. 1) has not been detected following 20 passages in culture, 8 acute passages of SARS-CoV-ExoN(-) in aged BALB/c mice, and 60 days of persistent SARS-CoV-ExoN(-) infection in immunodeficient $Rag^{-/-}$ mice (13, 14, 16, 17, 38). In this study, we sought to determine whether long-term passage of MHV-A59-ExoN(-) (250 passages over 1 year [P250])—here MHV-ExoN(-)—would result in virus extinction, ExoN(-) reversion, or compensation for the loss of proofreading. We demonstrate that MHV-ExoN(-) did not extinguish during passage and adapted for increased replication. MHV-ExoN(-) concurrently evolved reduced susceptibility to multiple nucleoside and base analogues, consistent with selection for increased replication fidelity. Importantly, the ExoN-inactivating substitutions did not revert. The evolved mutations in MHV-ExoN(-) nsp14 and nsp12, which encodes the RdRp, accounted for only part of the increased nucleoside analogue resistance of MHV-ExoN(-) P250, implicating multiple replicase proteins in adaptation for viral fitness. The results of this study support the proposed link between CoV fidelity and fitness, demonstrate the surprising stability of the ExoN-inactivating substitutions, and identify additional proteins outside nsp12 and nsp14 that may contribute to CoV fidelity regulation.

RESULTS

Long-term passage of WT-MHV and MHV-ExoN(-). We serially passaged WT-MHV and MHV-ExoN(-) in delayed brain tumor (DBT) cells 250 times (P250). Virus from each passage was harvested once 50% to 100% of the monolayer was involved in syncytia, which occurred between 8 and 24 hours postinfection (hpi). Passage conditions varied

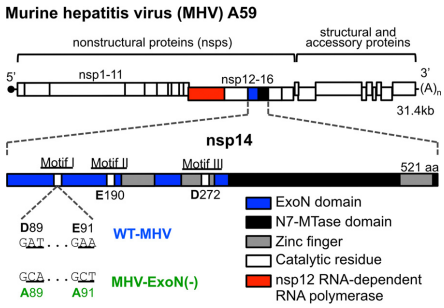


FIG 1 MHV genome organization and nsp14 exoribonuclease motifs. (Top) The MHV genome is a 31.4-kb, capped (dark circle), and polyadenylated positive-sense RNA molecule. The first two-thirds of the genome encode 16 nonstructural proteins translated as a single polyprotein with a ribosomal frameshift. The final one-third encodes the structural and accessory proteins. (Inset) Nsp14 encodes an exoribonuclease (solid blue) and an N7-methyltransferase (hatched blue) and has 3 zinc fingers (gray boxes) predicted from the solved SARS nsp10/14 crystal structure (PDB 5C8U) (44). Catalytic residues for ExoN are marked with white boxes, and the engineered mutations for MHV-ExoN(-) are shown below the genome. The nsp12 RNA-dependent RNA polymerase is highlighted in red.

for WT-MHV and MHV-ExoN(-) due to differences in replication kinetics between the two viruses. We stopped passage at P250 after observing reduced syncytium formation in MHV-ExoN(-)-infected flasks, likely resulting from a mutation in the MHV-ExoN(-) P250 spike protein cleavage site (discussed below).

MHV-ExoN(-) and WT-MHV replicate with identical kinetics following 250 passages. MHV-ExoN(-) has a significant replication defect relative to WT-MHV (14). We first tested whether replication of MHV-ExoN(-) P250 was affected by long-term passage by examining replication at two different multiplicities of infection (MOI). At both MOI = 1 and MOI = 0.01 PFU/cell, MHV-ExoN(-) P3 replication was delayed by ~2 h and the peak titer was reduced by ~1 $\log_{10}$ relative to WT-MHV P3 (Fig. 2A and B), consistent with our previous studies (14). By P250, the two viruses replicated with identical kinetics (Fig. 2A and B, dotted lines). This represented an ~1 $\log_{10}$ increase in peak replication for WT-MHV and an ~2 $\log_{10}$ increase for MHV-ExoN(-), compared with the respective parental viruses. At MOI = 0.01 PFU/cell, we also measured replication of MHV-ExoN(-) at P10, P50, P100, and P160. Replication kinetics gradually increased during the passages, reaching P250-like levels by P100 (Fig. 2B). To determine whether the

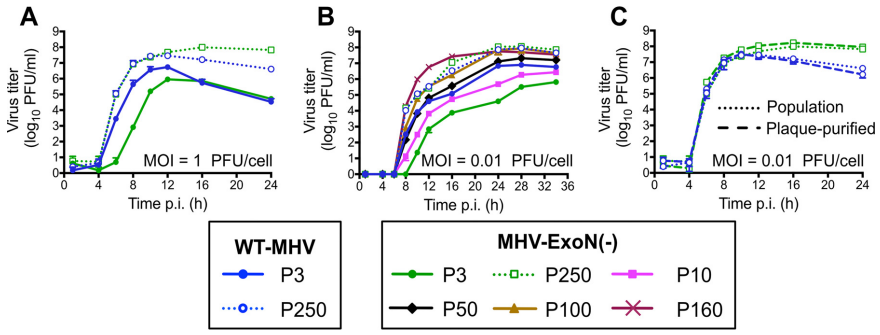


FIG 2 MHV-ExoN(-) evolved increased replicative capacity over long-term passage. Replication kinetics were examined for the indicated viruses at MOI = 1 PFU/cell (A) and MOI = 0.01 PFU/cell (B). (C) Replication kinetics of plaque-purified WT-MHV P250 and MHV-ExoN(-) P250 in parallel with the full population (MOI = 0.01 PFU/cell). Supernatants were collected at the indicated times postinfection, and titers were determined by plaque assay. Data for panels A to C represent means and standard deviations of data from $n = 3$.

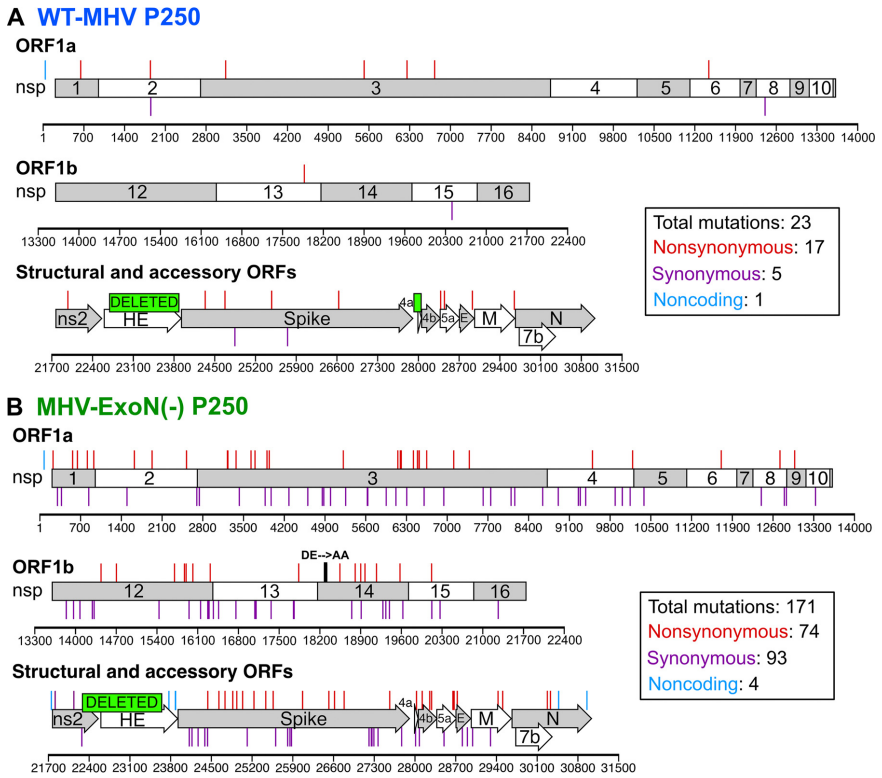


FIG 3 Mutations within P250 viruses. The mutations shown were present at >50% by di-deoxy sequencing at passage 250 in WT-MHV (A) and MHV-ExoN(-) (B). Nonsynonymous mutations (red), noncoding mutations (cyan), and deletions (green boxes) are plotted above the schematic, and synonymous mutations (purple) are plotted below the schematic.

increased replication of MHV-ExoN(-) P250 was affected by the presence of potential defective viral genomes or by some other population-based phenomenon, both WT-MHV P250 and MHV-ExoN(-) P250 were plaque purified three times. The plaque-purified viruses replicated indistinguishably from the parent populations (Fig. 2C). Together, these data demonstrate that WT-MHV and MHV-ExoN(-) populations had adapted for increased replication and that either individual genomes or those derived from a single virus plaque encoded the adaptive changes required by the total population.

MHV-ExoN(-) accumulated 8-fold-more mutations than WT-MHV but did not revert ExoN-inactivating substitutions. To determine whether the increased replication of MHV-ExoN(-) P250 resulted from primary reversion of ExoN(-) motif I, we sequenced nsp14 from infected-cell total RNA. MHV-ExoN(-) P250 retained the motif I DE-to-AA substitutions, demonstrating that primary reversion of ExoN(-) motif I did not occur. To identify potentially adaptive consensus mutations, we performed full-genome di-deoxy sequencing of MHV-ExoN(-) P250 and WT-MHV P250. Within WT-MHV P250, we identified 23 mutations, of which 17 were nonsynonymous (NS) (Fig. 3A). In contrast, MHV-ExoN(-) P250 had 171 total mutations (74 NS) (Fig. 3B). The full-genome sequences have been deposited in GenBank, and the mutations for both viruses are listed in Tables S1 and S2 in the supplemental material. We identified only one mutation shared by both viruses (nsp1 A146T), though it was present in approximately 50% of the

WT-MHV P250 population by di-deoxy sequencing. Both viruses deleted most of the hemagglutinin esterase (HE). In MHV-A59, HE mRNA is not transcribed *in vitro* (47–49), and HE protein expression is detrimental to MHV-A59 fitness in cell culture (50). WT-MHV P250 also deleted open reading frame 4a (ORF4a), which is dispensable for MHV replication in cell culture (51). The C-terminal region of ns2 within MHV-ExoN(-) P250 was truncated and fused to HE with a −1 frameshift. Ns2 is a phosphodiesterase (PDE) that protects viral RNA by degrading 2′-to-5′ oligoadenylate, the activating factor for cellular RNase L (52–54). The portion of ns2 deleted in MHV-ExoN(-) P250 lies outside the PDE catalytic domain, in a region of unknown function. C-terminally truncated ns2 retains enzymatic activity (55), but whether these specific deletions and fusions disrupt PDE activity remains to be tested. Nevertheless, ns2 is dispensable for MHV replication in immortalized cells (56, 57). Details about the deletion sites are provided in Fig. S1 in the supplemental material. Within proteins predicted to be part of the replicase-transcriptase complex (nsp7-16 and nucleocapsid) (39), WT-MHV P250 had only one NS change, located in the nsp13-helicase (Fig. 3A and Table S1). In contrast, MHV-ExoN(-) P250 had 17 NS changes within this region (Fig. 3B and Table S2).

MHV-ExoN(-) P250 displays increased genomic RNA accumulation and increased resistance to 5-fluorouracil. Coronaviruses lacking ExoN consistently display defects in RNA synthesis relative to WT strains (14, 16, 42). To determine whether the increased replication of MHV-ExoN(-) P250 was associated with restored genomic RNA (gRNA) production, we measured gRNA accumulation over time using two-step real-time quantitative PCR (15, 16). MHV-ExoN(-) P250 accumulated levels of gRNA similar to those accumulated by WT-MHV P3 and WT-MHV P250 at early time points, while gRNA levels for MHV-ExoN(-) P3 were ~1 $\log_{10}$ lower (Fig. 4A). MHV-ExoN(-) P250 gRNA levels fell below those of WT-MHV and WT-MHV P250 after 8 h and were similar to those of MHV-ExoN(-) P3 at 10 hpi. Normalizing to the gRNA abundance at 4 h for each virus demonstrated that the rates of gRNA accumulation were similar for all four viruses (Fig. 4B). These data suggest that the increased replication of P250 viruses relative to WT-MHV is not fully accounted for by increased RNA synthesis. In addition to RNA synthesis defects, ExoN(-) CoVs have up to 20-fold-increased mutation frequencies and profoundly increased sensitivity to nucleoside and base analogues relative to WT CoVs (13, 14, 16, 17, 38). To determine whether the nucleoside analogue sensitivity of MHV-ExoN(-) was altered by long-term passage, we treated cells infected with parental and passaged viruses with the base analog, 5-fluorouracil (5-FU). 5-FU is converted intracellularly into a nucleoside analogue that incorporates into growing RNA strands and causes A:G and U:C mutations. For simplicity, we will hereafter refer to 5-FU as a nucleoside analogue. Incorporation of 5-FU is increased in the absence of ExoN activity (16). All viruses displayed a concentration-dependent decrease in viral titer but differed greatly in their levels of susceptibility to 5-FU (Fig. 4C). At 120 μM, WT-MHV P3 titers were reduced by ~1 $\log_{10}$, while MHV-ExoN(-) P3 titers were undetectable (>5 $\log_{10}$-fold reduction). WT-MHV 5-FU sensitivity was not altered by passage. MHV-ExoN(-) P250 was less susceptible than MHV-ExoN(-) P3 to 5-FU treatment, with a decrease in titer of only ~1.5 $\log_{10}$ at 120 μM. MHV-ExoN(-) P250 remained more sensitive to 5-FU than WT-MHV, suggesting that WT-like resistance requires an intact ExoN. These data demonstrate that MHV-ExoN(-) P3 evolved resistance to 5-FU through mutations outside ExoN(-) motif I.

Spike mutations in MHV-ExoN(-) P250 do not increase resistance to 5-FU. Bacteriophage ϕX174 acquired resistance to 5-FU by delaying cell lysis, thereby reducing the number of replication cycles in which 5-FU can be incorporated (58). MHV-ExoN(-) P250 had multiple mutations in the spike glycoprotein, including one in the spike furin cleavage site that reduced syncytium formation. To test whether the spike mutations manifested in resistance to 5-FU, we cloned the spike gene from MHV-ExoN(-) P250 into the isogenic MHV-ExoN(-) background. The recombinant virus demonstrated intermediate replication kinetics between MHV-ExoN(-) P3 and MHV-ExoN(-) P250 (Fig. 5A) and did not form syncytia. Spike-P250 also increased the specific infectivity of viral particles (Fig. 5B). However, the MHV-ExoN(-) P250 spike did not affect

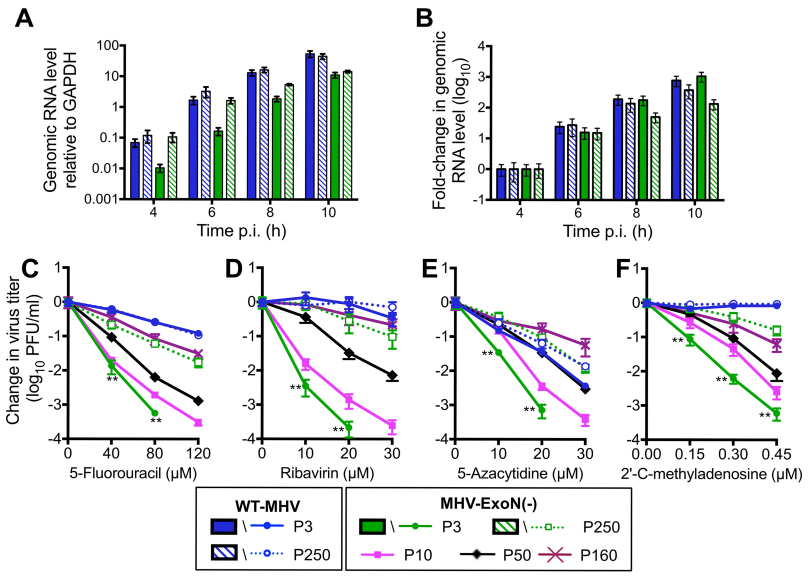


FIG 4 MHV-ExoN(-) evolved WT-like genomic RNA accumulation and increased resistance to multiple nucleoside analogues over the passage. (A) Cells were infected with the indicated viruses at MOI = 1 PFU/cell, and intracellular RNA was harvested using TRIzol at the indicated times postinfection. MHV genomic RNA was detected using SYBR green and primers directed to nsp10, and values were normalized to intracellular GAPDH. (B) Same data as in panel A normalized to the RNA level for each virus at 4 hpi. Data represent means and standard errors of results for $n = 9$ (3 triplicate experiments). (C to F) Sensitivity of passaged viruses to nucleoside analogues at MOI = 0.01 PFU/cell. Cells were treated with the indicated concentrations of 5-FU (C), RBV (D), AZC (E), or CMeA (F) for 30 min prior to infection, supernatants were harvested at 24 hpi, and titers were determined by plaque assay. Data represent changes in titer relative to untreated control results and are plotted as means and standard errors of results from $n = 6$ (two triplicate experiments). For panels C to F, the statistical significance of changes in the titer of MHV-ExoN(-) P3 relative to MHV-ExoN(-) P250 was determined using the Mann-Whitney test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

the sensitivity of the recombinant virus to 5-FU (Fig. 5C). Thus, any adaptive increase in 5-FU resistance must be located elsewhere in the genome.

MHV-ExoN(-) passage resulted in unique mutations in nsp12 and nsp14. To date, three proteins have been shown to alter CoV sensitivity to 5-FU: nsp12-RdRp, nsp14-ExoN, and nsp10 (which stimulates ExoN activity) (15, 17, 39). Neither WT-MHV nor MHV-ExoN(-) P250 contained an NS mutation in nsp10, and WT-MHV P250 had no mutations within either nsp12 or nsp14. In contrast, MHV-ExoN(-) P250 had 7 NS mutations in nsp12 and 6 NS mutations in nsp14 (Fig. 3 and 6), none of which have been described previously *in vitro* or in viable viruses. Within nsp12, six mutations were in the predicted RdRp finger, palm, and thumb domains (Fig. 6A) (59). Four residues (H709, F766, S776, and M814) can be visualized on a Phyre²-modeled structure of the MHV-nsp12 RdRp, while the remaining residues lie outside the modeled core RdRp (Fig. 6A) (17). One mutation, M288T, lies in the CoV-specific domain, which is conserved among nidoviruses. This domain has been implicated in membrane targeting in MHV-A59 (60) and performs an essential nucleotidylation activity in the *Arterivirus* equine arteritis virus (61). However, M288T is not predicted to catalyze nucleotidylation. Within nsp14, 4 NS mutations were identified in the ExoN domain, and 2 NS mutations were in the C-terminal N7-methyltransferase domain (Fig. 6B). We next modeled the structure of MHV nsp14 using Phyre² software (62), resulting in highest-probability similarity to the SARS-CoV nsp14-nsp10 complex (PDB 5C8S) (44) with high confidence (i.e., the calculated probability of true homology between the structures) of 100% for

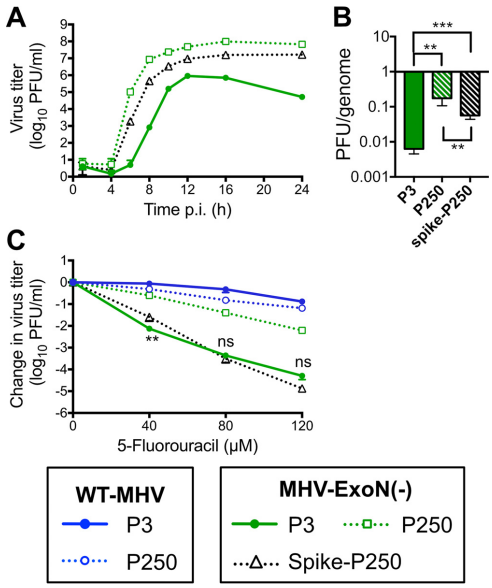


FIG 5 Mutations in the spike envelope protein from MHV-ExoN(-) P250 increase replicative capacity but do not affect sensitivity to 5-fluorouracil. (A) Replication kinetics of indicated viruses (MOI = 0.01 PFU/cell) plotted as means and standard deviations of results determined with $n = 3$. (B) Specific infectivity of indicated viruses 12 hpi (MOI = 1 PFU/cell). Data represent means and standard errors of results from $n = 6$ (two triplicate experiments). (C) Sensitivity of indicated viruses to 5-fluorouracil at MOI = 0.01 PFU/cell, determined as described for Fig. 4. Data represent means and standard errors of results from $n = 6$ (two triplicate experiments). For panel B, the statistical significance was determined using one-way analysis of variance (ANOVA). For panel C, the statistical significance of changes in the titer of MHV-ExoN(-) spike-P250 relative to MHV-ExoN(-) P3 was determined using the Mann-Whitney test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant).

residues 3 to 519 of MHV-nsp14. The model predicts that five mutations are located close to surface of the protein (Fig. 6B). All three modeled zinc finger domains contain one NS mutation (F216Y, Y248H, and L473I). Two mutations, D128E and F216Y, are located near the interface between nsp10 and nsp14, though neither site has previously been implicated in nsp10-nsp14 interactions (15, 63, 64). One NS mutation resulted in a D272E substitution in ExoN motif III, a metal-coordinating active site residue. We previously reported that alanine substitution of D272 results in an ExoN(-) phenotype (14), but the viability or phenotype of a D272E substitution was not tested in that study. These data suggest that a network of residues evolved to regulate nsp12 and nsp14 activity or stability in the ExoN(-) background.

Fixed mutations in nsp12 and nsp14 in MHV-ExoN(-) P250 directly correlate with increased resistance to multiple nucleoside analogues. To determine approximately when the mutations in nsp12 and nsp14 arose, we performed di-deoxy sequencing across these protein-coding regions roughly every 20 passages (P10, P31, P50, P72, P90, P100, P120, P140, P160, P180, P200, P220, and 240). By this method, we detected consensus NS mutations at P10, P50, and P160 for nsp12 and at P50 and P160 for nsp14 (Fig. 6). Both nsp12 and nsp14 carried their full complement of P250 consensus mutations by P160, except for a minority variant (D913E) in nsp12 that was maintained at <50% of the population between P200 and P250. These passage levels correlated with increased replication of MHV-ExoN(-) (Fig. 2B) and with decreasing sensitivity to 5-FU (Fig. 4C). Neither replication nor 5-FU sensitivity of MHV-ExoN(-) changed substantially between P160 and P250. To determine whether MHV-ExoN(-)

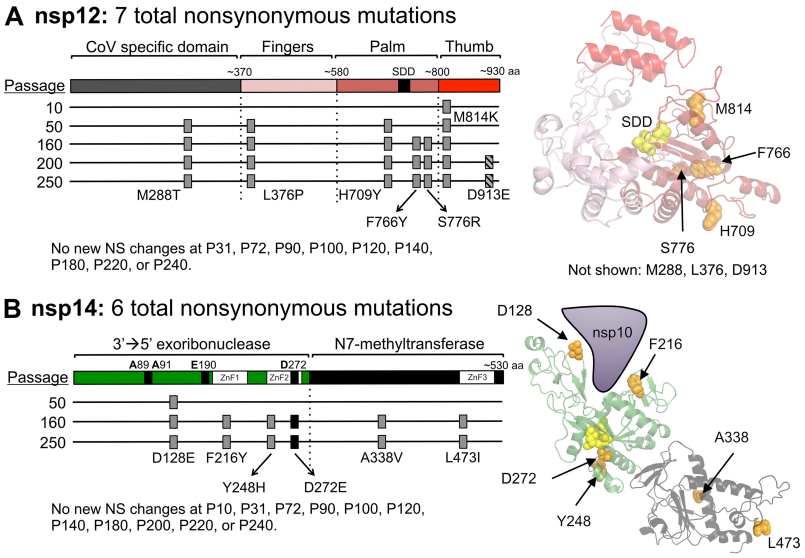


FIG 6 The timing of fixation of mutations in nsp12-RdRp and nsp14-ExoN within MHV-ExoN(-). (A) A schematic of nsp12-RdRp with the CoV-specific region and the canonical finger, palm, and thumb domains of RdRps is shown. The nsp12-RdRp coding region was sequenced at the indicated passage, and the nonsynonymous changes are plotted; gray boxes indicate consensus changes, and hatched boxes indicate variants shown to be present in <50% of the population by di-deoxy sequencing. At right, mutations are marked in orange on a Phyre²-modeled structure of MHV-nsp12, with the active site residues marked in yellow (17). RdRp domains are colored according to the linear schematic. M288T, L376P, and D913E lie outside the modeled region and thus are not marked. (B) A schematic of nsp14 with the ExoN and N7-methyltransferase domains is shown, with mutation plotting depicted as described for panel A. The black box denotes a mutation to ExoN motif III. At right, mutations are marked in orange on a Phyre²-modeled structure of MHV-nsp14. Domains are colored according to the linear schematic.

evolved increased resistance to multiple nucleoside analogues, we treated virus-infected cells with three additional analogues that are substrates for viral RdRps: ribavirin (RBV), a guanine analogue that inhibits viral replication through multiple mechanisms, including mutagenesis and inhibition of purine biosynthesis (65); 5-azacytidine (AZC), an RNA mutagen (66); and 2′-C-methyladenosine (CMeA), which is proposed to incorporate in viral RNA and terminate nascent transcripts (67). As with 5-FU, we observed dose-dependent sensitivity to RBV, AZC, and CMeA in all MHV-ExoN(-) viruses that decreased with increasing passage number (Fig. 4D to F). Except for AZC, MHV-ExoN(-) sensitivity did not change between P160 and P250. Together, these data demonstrate that MHV-ExoN(-) evolved increased resistance to multiple nucleoside analogues that correlated with the length of passage and the acquisition of mutations in nsp12 and nsp14. Importantly, this occurred in the absence of specific mutagenic selection and without reversion of ExoN motif I. This increased general selectivity toward all four classes of nucleotide strongly supports the idea of an overall increase in fidelity in MHV-ExoN(-) P250.

Mutations in nsp12 partially account for increased resistance of MHV-ExoN(-) P250 to multiple nucleoside analogues. We hypothesized that mutations in MHV-ExoN(-) P250 nsp12 and nsp14 were most likely to impact replication and nucleoside analogue sensitivity based on their enzymatic activities and temporal association with phenotypic changes. To test this hypothesis, we engineered recombinant MHV-ExoN(-) to encode the P250 nsp12 and nsp14 sequences, alone and together. Expression of nsp12-P250 and nsp14-P250, alone or in combination, altered replication kinetics of MHV-ExoN(-) without affecting peak titers (Fig. 7A) and increased gRNA levels above those of MHV-ExoN(-) P3 (Fig. 7B). Nsp12-P250 had a greater effect than nsp14-P250 on

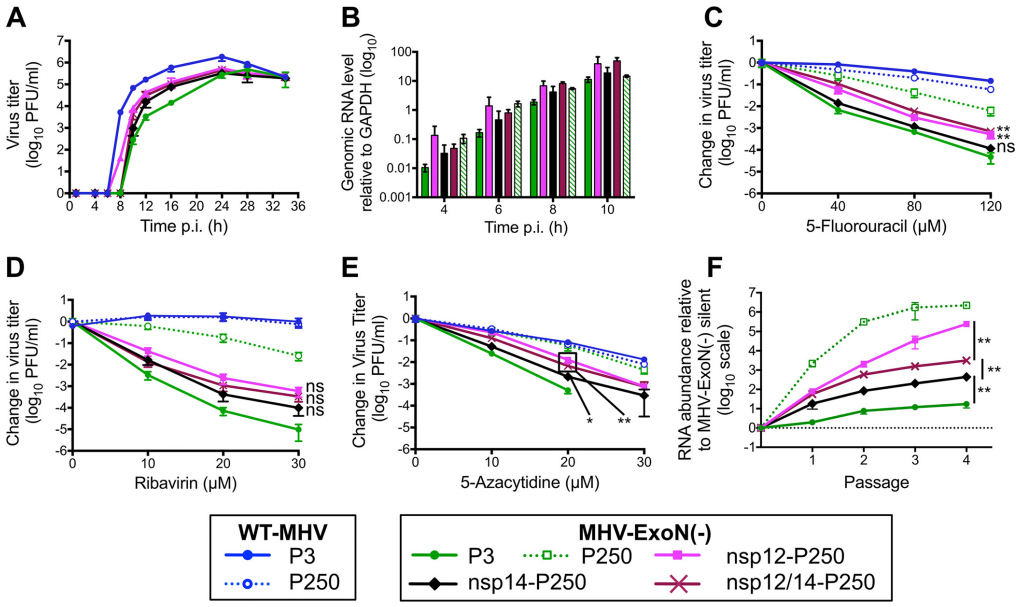


FIG 7 Mutations in nsp12-RdRp and nsp14-ExoN from MHV-ExoN(-) P250 incompletely increase resistance to nucleoside analogues and increase fitness of MHV-ExoN(-). (A) Replication kinetics of recombinant P250 viruses (MOI = 0.01 PFU/cell) plotted as means and standard deviations of results determined with $n = 3$. (B) Genomic RNA accumulation relative to intracellular GAPDH determined as described for Fig. 4. Data represent means and standard errors of results for $n = 6$ to 9 (2 to 3 triplicate experiments). (C to E) Sensitivity of recombinant MHV-ExoN(-) viruses to 5-FU (C), ribavirin (D), and 5-azacytidine (E) at MOI = 0.01 PFU/cell determined as described for Fig. 4. Data represent means and standard errors of results from $n = 6$. (F) Recombinant viruses were competed against a reference MHV-ExoN(-) containing 10 silent mutations within nsp2. The ratio of competitor to reference genomes is plotted. Data represent means and standard errors of results from $n = 6$. The MHV-ExoN(-) P250 data set contained 4 replicates at passage 3 and a single replicate at passage 4 due to undetectable levels of MHV-ExoN(-) (silent). For panels C to E, the statistical significance of changes in the titer of swapped viruses relative to MHV-ExoN(-) P3 at the highest drug concentration tolerated was determined using the Mann-Whitney test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant). For panel F, the statistical significance for the indicated comparisons was determined using the Mann-Whitney test. Boxed points have the same P value.

the sensitivity of MHV-ExoN(-) to all analogues tested, and the combination of nsp12- and nsp14-P250 did not increase resistance above that seen with nsp12-P250 alone (Fig. 7C to E). None of the recombinant viruses recapitulated the resistance phenotypes of the MHV-ExoN(-) P250 population. Together, these data demonstrate that nsp12-P250 mutations account only partially for the nucleoside analogue resistance of MHV-ExoN(-) P250 and that adaptations in nsp12-P250 mask those in nsp14-P250. We also can conclude that the nsp14-P250 D272E active site mutation does not correct the defect caused by the motif I DE-to-AA substitutions.

Resistance to nucleoside analogues correlates with MHV-ExoN(-) fitness. We hypothesized that mutations in nsp12 and nsp14 provided a fitness advantage to MHV-ExoN(-) P250. We competed the recombinant viruses with a reference MHV-ExoN(-) virus (P1 stock) containing 10 silent mutations in the nsp2 coding region. Mutant and reference viruses were detected in the mixed infection by real-time quantitative PCR using dually labeled probes specific for each virus. MHV-ExoN(-) P3 showed a modest fitness advantage over the reference P1 MHV-ExoN(-) silent strain (Fig. 7F, solid green). MHV-ExoN(-) P250 profoundly outcompeted MHV-ExoN(-) silent, with >1,000-fold more MHV-ExoN(-) P250 genomes present at the end of passage 1 (Fig. 7F, dotted green line). MHV-ExoN(-) nsp12-P250 had greater relative fitness than MHV-ExoN(-) nsp14-P250, and MHV-ExoN(-) nsp12/14-P250 was intermediate between the single recombinants, implicating a complex evolutionary interaction between these two proteins. The measured fitness correlated with the patterns of nucleoside analogue

resistance and RNA synthesis associated with mutations in nsp12 and nsp14, suggesting a link between the evolutions of these phenotypes. The result also confirms that nsp12 and nsp14 are important but not sufficient to account for the significantly increased fitness of MHV-ExoN(-) P250 relative to MHV-ExoN(-) P3.

DISCUSSION

In this report, we describe experimental adaptive evolution of WT-MHV and MHV-ExoN(-) during long-term passage in cell culture. WT-MHV evolved increased replication kinetics over 250 passages, with few consensus mutations arising in the WT-MHV P250 genome. In contrast, MHV-ExoN(-) accumulated 8-fold-more mutations than WT-MHV, none of which occurred at the ExoN-inactivating substitutions. Nevertheless, MHV-ExoN(-) P250 demonstrated increased replication kinetics and fitness compared to MHV-ExoN(-) P3 (Fig. 2 and 7). Our previous studies demonstrated that ExoN-mediated proofreading is required for CoV fitness (16, 17, 38). Thus, MHV-ExoN(-) was likely under selective pressure for restoration of high-fidelity replication or for tolerance of the increased mutational load. Consistent with this hypothesis, MHV-ExoN(-) P250 exhibited increased resistance to multiple nucleoside analogues, a phenotype strongly associated with high-fidelity viruses (15, 18, 19, 22). Our results raise several important questions. In the face of selection for increased fidelity, why did MHV-ExoN(-) not revert? Can MHV replicase proteins mediate high-fidelity replication without ExoN proofreading? Which mechanisms other than increased fidelity can compensate for the loss of proofreading?

In the face of selective pressure for increased fidelity, why did MHV-ExoN(-) not revert? Although our data suggest that MHV-ExoN(-) was under selective pressure for increased fidelity, we detected no primary reversion at the DE-to-AA substitutions in MHV-ExoN(-) at any passage tested. These data are consistent with and significantly extend previous studies reporting genotypic stability of ExoN(-) motif I in MHV and SARS-CoV (13, 14, 16, 17, 38). Complete reversion to DE within ExoN(-) motif I would require four nucleotide changes. This likely represents a high genetic barrier to reversion, especially given that fitness can be increased by mutations outside nsp14-ExoN (Fig. 7F) (13). Single and double nucleotide changes within motif I could restore an acidic charge to individual residues (e.g., motif I EA, AD, ED, etc.). However, the active site compositions of DEDDh exonucleases, such as the Klenow fragment, are so stringent that even conservative mutations (D to E or E to D) reduce ExoN activity by >96% (68). Thus, intermediate amino acid changes may not have a selective advantage compared to motif I AA, limiting the evolutionary pathways to reversion. However, nsp14-P250 had detectable effects on RBV and AZC resistance as well as on the competitive fitness of MHV-ExoN(-) (Fig. 7F), demonstrating a modest capacity for fitness adaptation in nsp14 outside the catalytic residues. Whether these mutations resulted from genetic drift or positive selection remains unclear. Nevertheless, our data show that MHV-ExoN(-) can adapt for increased fitness without fully restoring exoribonuclease activity. While some mutations in MHV-ExoN(-) P250 likely confer DBT cell-specific selective advantages, others may represent generalizable strategies for overcoming ExoN(-) defects in other cell types and in other coronaviruses. Thus, understanding the mechanisms by which MHV-ExoN(-) P250 compensated for ExoN activity could allow recovery of ExoN(-) variants of other CoVs, such as transmissible gastroenteritis virus and human CoV 229E, which to date have been nonviable as ExoN(-) recombinants (42, 69).

Can MHV replicase proteins mediate high-fidelity replication without ExoN proofreading? MHV-ExoN(-) P250 exhibits increased resistance to four nucleoside analogues after passage (Fig. 4). Although resistance to a single nucleoside analogue can evolve without increasing overall fidelity, resistance to multiple nucleoside analogues strongly suggests a broadly increased capacity to discriminate nucleotides (15, 18, 19, 22). Increased-fidelity variants in RNA viruses have most frequently been mapped to RdRps (24, 25, 30, 70). Thus, if increased fidelity contributes to nucleoside analogue resistance in MHV-ExoN(-) P250, the most likely protein involved would be nsp12-P250. Three findings are consistent with the hypothesis that mutations within

nsp12-P250 increase RdRp fidelity. First, nonsynonymous mutations to nsp12 arose in the low-fidelity MHV-ExoN(-) strain but not in the presence of proofreading (WT-MHV). Second, five of the mutations lie in or near structural motifs important for fidelity regulation in other RdRps. Amino acid substitutions in the finger and palm domains have been repeatedly shown to affect viral RdRp fidelity (25, 34), and we have recently reported a finger mutation (nsp12-V553I) that likely increases the fidelity of the MHV RdRp (17). Our modeled structure predicts that nsp12-P250 contains three mutations in the palm domain and one in the finger domain, with the M814K thumb domain mutation lying near the palm (Fig. 6A). Third, exchange of nsp12-P250 alone into the background of MHV-ExoN(-) reduced the susceptibility of MHV-ExoN(-) to three different nucleoside analogues (Fig. 7). Thus, all data support the hypothesis that nsp12-P250 is a high-fidelity RdRp. We are actively developing biochemical, phenotypic, and deep sequencing assays to quantify the fidelity of nsp12-P250.

Importantly, nsp12-P250 accounts only partially for the MHV-ExoN(-) P250 nucleoside analogue resistance phenotype (Fig. 7), suggesting a possible limit to the compensation achievable by mutating the RdRp alone. Further, the effects of mutations in nsp12-P250 and nsp14-P250 are not additive and may be antagonistic when they are isolated from the whole passaged virus (Fig. 7), indicating that the relationships between nsp12- and nsp14-P250 mutations are likely evolutionarily linked with those in other MHV proteins. In fact, a substantial component of the evolved resistance to nucleoside analogues cannot be explained by the presence of nsp12-P250 and nsp14-P250, alone or together. In support of this hypothesis, we identified several nonsynonymous mutations in other replicase proteins, such as nsp8, nsp9, nsp13, and nsp15. SARS-CoV nsp8 and nsp13 have functional interactions with nsp12, acting as a primase/processivity factor (71, 72) and a helicase/NTPase, respectively (73). Processivity factors in herpes simplex virus and *Mycobacterium tuberculosis* regulate DNA polymerase fidelity by balancing polymerase extension and exonuclease activity (74, 75), and helicases in Chikungunya virus and foot-and-mouth disease virus can evolve to increase fidelity (76) and alter the frequency of ribavirin-induced mutations (77), respectively. SARS-CoV nsp9 has RNA-binding activities and is proposed to participate in the multiprotein replicase complex (39, 78), and MHV nsp15 is a uridylate-specific endoribonuclease (79). Both could plausibly be involved in modulating polymerase activity. Additionally, it remains possible that evolution for increased fidelity could involve proteins outside the canonical replication complex (nsp7 to nsp16), including those in the structural and accessory cassette. Thus, while the immediate studies will focus on testing whether replicase proteins nsp8, nsp9, nsp13, and nsp15 regulate fidelity, it is exciting to consider the possibility that this virus-directed discovery approach will reveal novel interactions between multiple MHV proteins.

Which mechanisms other than increased fidelity might account for MHV-ExoN(-) P250 nucleoside analog resistance? Genomic mutations in RNA viruses are most commonly detrimental or lethal (8–12). One strategy to prevent extinction by mutational load is to increase replication fidelity, as discussed above. An alternative strategy is to increase mutational robustness. Mutational robustness describes the capacity of a virus to buffer the fitness effects of mutations. In the setting of low-fidelity replication, as in MHV-ExoN(-), increased mutational robustness could have provided a selective advantage (80–82). Selection for increased robustness could explain the ~90 synonymous changes in MHV-ExoN(-) P250. Synonymous changes can alter codons to reduce the probability of nonconservative amino acid changes (83). Alternatively, the increased population size of MHV-ExoN(-) P250 could promote robustness by a "safety-in-numbers" effect, allowing efficient purging of low-fitness mutants while maintaining population fitness (84). Large populations also increase the likelihood of coinfection, allowing complementation between viral genomes. Although increased replication conferred by mutations in spike did not alter 5-FU resistance (Fig. 5), results of a recent study performed with poliovirus suggest that mutagenized populations have elevated coinfection frequencies (85). Thus, complementation may contribute to MHV-ExoN(-) P250 nucleoside analogue resistance. Conflicting evidence exists regarding whether

mutational robustness itself affects the sensitivity to RNA mutagens (83, 86, 87); nevertheless, the robustness of MHV-ExoN(-) P250 merits further investigation.

Conclusions. The proofreading activity of the nsp14 exoribonuclease is a critical determinant of CoV replication, fidelity, and fitness. We showed that CoVs have the capacity to compensate for loss of ExoN activity through a network of mutations in nsp12 and nsp14 and elsewhere in the genome. Thus, while nsp14-ExoN appears to play a dominant role in CoV replication fidelity, its activity is likely closely tied to a highly evolved network of proteins. The demonstrated coadaptation for replication, competitive fitness, and likely increased fidelity within MHV-ExoN(-) supports the hypothesis that these roles are linked functionally and evolutionarily. It will be interesting to test whether evolution in other cell types derived from different species or with different innate immune environments would result in similar adaptive strategies. Genetic and biochemical testing of the rich mutational resource revealed in MHV-ExoN(-) P250 will likely inform the design of countermeasures for endemic and emerging CoVs by defining novel common targets for stable virus attenuation or direct inhibition.

MATERIALS AND METHODS

Cell culture. DBT-9 (delayed brain tumor, murine astrocytoma clone 9) cells were maintained as described previously (88). Baby hamster kidney (BHK) cells stably expressing the MHV-A59 receptor CEACAM1 (BHK-R; 15) were maintained under conditions of selection with 0.8 mg/ml of G418 (Mediatech) as described previously (88).

Long-term passage of virus and stock generation. The infectious cDNA clone for MHV-A59 and the recovery of MHV-ExoN(-) were described previously (14, 88). Long-term passage was initiated by infecting subconfluent monolayers of DBT-9 cells in 25-cm^2 flasks with either wild-type MHV-A59 or MHV-ExoN(-) at an MOI of approximately 0.1 PFU/cell. One lineage of each virus was subjected to a total of 250 passages (P250). Supernatant was harvested at each passage and stored at −80°C. Total RNA was harvested for most passages using 1 ml of TRIzol reagent (Ambion) per 25-cm^2 flask and stored at −80°C. Virus stocks of select intermediate passages were generated by infecting a subconfluent 150-cm^2 flask of DBT-9 cells at an MOI of 0.01 PFU/cell. At approximately 24 hpi, the flask was frozen at −80°C and the supernatant was clarified by centrifugation at 4,000 × *g* (Sorvall RC-3B Plus; HA-6000A rotor) for 10 min at 4°C. The virus titer of each stock was determined by plaque assay using DBT-9 cells as described previously (14, 88). For plaque assays of viruses containing the spike protein from MHV-ExoN(-) P250, which does not form syncytia, plaques were visualized with neutral red (Sigma catalog no. N6264) (dilution at 1:10 in phosphate-buffered saline [PBS] containing calcium and magnesium). Neutral red was added 24 h after plating, and the reaction mixture was incubated for an additional 3 to 8 h before formaldehyde fixation. Plaque purification was performed by infecting DBT cells with serial dilutions of virus and overlaying the cultures with agar. Single plaques were isolated, resuspended in PBS containing calcium and magnesium, and inoculated onto fresh DBTs. This process was completed 3 times before experimental stocks were generated as described above.

Sequencing of virus stocks. Following P250, RNA was purified from the harvested TRIzol samples according to the manufacturer's protocol and reverse transcribed (RT) using SuperScript III (Invitrogen) as described previously (14). Full-genome di-deoxy sequencing was performed for both WT-MHV P250 and MHV-ExoN(-) P250 using 12 overlapping amplicons approximately 3 kb in length. All coding regions were sequenced fully, and, of 31,409 nucleotides, >99% were sequenced for each virus [for WT-MHV P250, 21 to 31,279; for MHV-ExoN(-) P250, 21 to 31,275]. Two microliters of RT product was used for each PCR (16). Di-deoxy sequencing was performed by Genhunter Corporation (Nashville, TN) and Genewiz (South Plainfield, NJ). Sequence analysis was performed using MacVector version 14 (MacVector, Inc., Apex, NC) and the MHV-A59 infectious clone reference genome (GenBank accession number AY910861). The nucleotide sequences of the amplicon and sequencing primers are available upon request. Sequencing of nsp12 and nsp14 from intermediate passages was performed using TRIzol-purified RNA from infected monolayers and the primers listed below. Primers 6M1F (5′-TTTTGGCGAGATGGTAGC-3′) and 7M2R (5′-GGTAAGACAGTTTTAGGTGAG-3′) were used to generate a 3,425 nucleotide amplicon containing all of nsp12. Primers 7M3F (5′-ATGCTTACCAACTATGAGC-3′) and 8M3R (5′-CCGATTTGAATGGCGTAG-3′) were used to generate a 2,713 nucleotide amplicon containing all of nsp14. The PCR conditions for these reactions were the same as those used to generate the amplicons used for full-genome sequencing (16).

Replication and RNA synthesis kinetics. Viral replication kinetics in DBT-9 cells were determined at an MOI of 1 PFU/cell or an MOI of 0.01 PFU/cell as described previously (15). Supernatant (300 μl) was harvested at the indicated time points, and the virus titer was determined by plaque assay. The accumulation of genomic RNA at an MOI of 1 PFU/cell was measured by two-step real-time quantitative RT-PCR. Intracellular RNA was harvested using TRIzol and reverse transcribed using SuperScript III (Invitrogen). The levels of cDNA derived from intracellular positive-sense viral RNA were measured using primers directed to nsp10. Values were normalized to levels of the endogenous control glyceraldehyde-3-phosphate dehydrogenase (GAPDH). No mutations within the primer binding sites emerged in either P250 population. The primers and amplification conditions were the same as reported previously (15),

except that the RT product was diluted 1:10 prior to use. Samples were plated in technical duplicate to minimize well-to-well variation. Data are presented as $2^{-\Delta CT}$, where ΔCT denotes the threshold cycle (C_T) value for the target (nsp10) minus the C_T value for the reference (GAPDH).

Determination of specific infectivity. Subconfluent monolayers of DBT-9 cells in 24-well plates were infected with the indicated virus at an MOI of 1 PFU/cell, and supernatant was harvested at 12 hpi. The levels of genomic RNA in supernatant were measured using one-step real-time quantitative RT-PCR (RT-qPCR) of TRIzol-extracted RNA as described previously (17). Briefly, genomic RNA was detected with a 5′ 6-carboxyfluorescein (FAM)-labeled and 3′ black hole quencher 1 (BHQ-1)-labeled probe targeting nsp2 (Biosearch Technologies, Petaluma, CA), and RNA copy numbers were calculated by reference to an RNA standard derived from the MHV A fragment. Samples were plated in technical duplicate to minimize well-to-well variation. Titers were determined by plaque assay in DBT-9 cells, and specific infectivity was calculated as PFU per supernatant genomic RNA copy.

Nucleoside and base analogue sensitivity assays. 5-azacytidine (AZC), 5-fluorouracil (5-FU), and ribavirin (RBV) were purchased from Sigma (product numbers A2385, F6627, and R9644, respectively), and stock solutions were prepared in dimethyl sulfoxide (DMSO). CMeA (2′-C-methyladenosine) was received from Gilead Sciences (Foster City, CA). Sensitivity assays were performed as described previously (16), except that 24-well plates were used at an MOI of 0.01 PFU/cell. Supernatants were harvested at 24 hpi, and titers were determined by plaque assay.

Phyre² modeling of MHV-nsp14. The MHV nsp14 structure was modeled with the Phyre² online program (62) using nsp14 residues 3 to 519, corresponding to residues 6,056 to 6,573 of the ORF1ab polyprotein. The model was analyzed using the PyMOL Molecular Graphics System, Version 1.3 (Schrödinger, LLC).

Generation of nsp12 and nsp14 swaps. Viruses containing nsp12-P250 or nsp14-P250 or both were generated using the MHV-A59 reverse genetics system (88). RNA from the MHV-ExoN(-) P250 virus was reversed transcribed with SuperScript III (Invitrogen) and used to generate amplicons containing either nsp12 or nsp14. Each amplicon was flanked by 15 bp that overlapped an amplicon generated from the backbone plasmid. Amplicons were inserted into MHV-A59 fragments using an InFusion HD cloning kit (TaKaRa Bio USA, Inc., Mountain View, CA). nsp12 is split across MHV E and F, while nsp14 is contained within MHV F. Reaction mixtures contained 50 ng of vector, 200 ng of insertion, and 2 μl of enzyme and were incubated for 15 min at 50°C. Errors were corrected by site-directed mutagenesis using *Pfu* Turbo polymerase (Agilent, Santa Clara, CA). The nsp12/14-P250 swap was generated through restriction digestion of the individual swaps using BsmBI and StuI followed by gel purification and assembly using T4 DNA ligase (NEB, Ipswich, MA). Viable viruses were constructed and rescued as described previously (88).

Competitive fitness assays. Competitor viruses were competed with an MHV-ExoN(-) virus harboring 10 silent mutations in the probe-binding region within nsp2. Subconfluent DBT-9 monolayers in 24-well plates were coinfected at a total MOI of 0.01 PFU/cell with competitor and reference viruses at a 1:1 ratio and passaged 4 times. For each passage, supernatants were harvested at 24 h. RNA was extracted from 100 μl of supernatant using 900 μl of TRIzol reagent and PureLink RNA minikit columns (Thermo Scientific, Waltham, MA), and 150 μl of supernatant was used to infect fresh cells in a 24-well plate (total MOI estimated at 1 PFU/cell). The proportion of each virus was determined by real-time RT-qPCR from the infection supernatant using two TaqMan probes with different fluorescent dyes in separate reactions. Competitor viruses were detected with the same probe used in specific infectivity analyses (14). Reference viruses were detected by a probe targeting the same region but with 10 silent mutations (5′-TCCGAACTACTGCAACCCCAAGTG-3′) and labeled with 5′ quasar 670 and 3′ black hole quencher 2 (BHQ-2) (Biosearch Technologies, Petaluma, CA). RNA copy numbers were calculated by reference to an RNA standard generated by *in vitro* transcription of the corresponding MHV A fragment, and relative RNA abundances were calculated as ratios of competitor genomes to reference genomes.

Statistical analysis. GraphPad Prism 6 (La Jolla, CA) was used to perform statistical tests. Only the comparisons shown (with statistical significance indicated as "ns" [nonsignificance] or asterisk[s]) within each figure or described in each legend were performed. In many cases, the data were normalized to the results obtained with untreated controls. This was performed using GraphPad Prism 6. The number of replicate samples is denoted within each figure legend.

Accession numbers. Full-length genome sequences for WT-MHV P250 and MHV-ExoN(-) P250 have been deposited in GenBank (accession numbers MF618252 and MF618253, respectively).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.01503-17.

FIG S1, TIF file, 1.7 MB.

TABLE S1, PDF file, 0.1 MB.

TABLE S2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank members of the Denison laboratory for valuable discussions.

This work was supported by United States Public Health Service awards R01-AI108197 (M.R.D.), T32-GM007347 (K.W.G.), F30-AI129229 (K.W.G.), T32-HL07751 (J.B.C.),

T32-AI089554 (N.R.S.), T32-AI095202 (E.C.S.), and F32-AI108102 (E.C.S.), all from the National Institutes of Health.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We declare that we have no conflicts of interest.

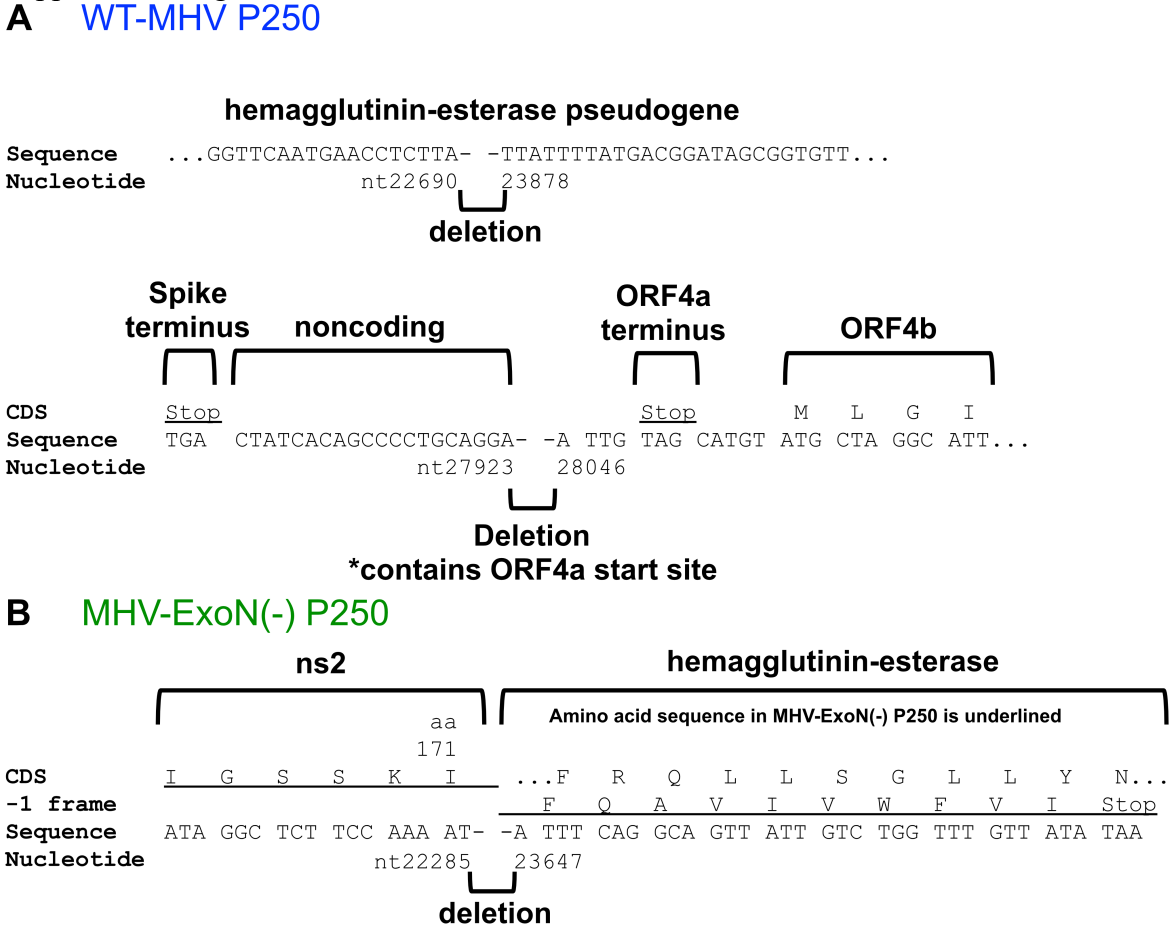
REFERENCES

1. Steinhauer DA, Domingo E, Holland JJ. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 122:281–288. https://doi.org/10.1016/0378-1119(92)90216-C.
2. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. J Virol 84:9733–9748. https://doi.org/10.1128/JVI.00694-10.
3. Drake JW, Holland JJ. 1999. Mutation rates among RNA viruses. Proc Natl Acad Sci U S A 96:13910–13913. https://doi.org/10.1073/pnas.96.24.13910.
4. Bull JJ, Sanjuán R, Wilke CO. 2007. Theory of lethal mutagenesis for viruses. J Virol 81:2930–2939. https://doi.org/10.1128/JVI.01624-06.
5. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev 76:159–216. https://doi.org/10.1128/MMBR.05023-11.
6. Bradwell K, Combe M, Domingo-Calap P, Sanjuán R. 2013. Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Qβ. Genetics 195:243–251. https://doi.org/10.1534/genetics.113.154963.
7. Gago S, Elena SF, Flores R, Sanjuán R. 2009. Extremely high mutation rate of a hammerhead viroid. Science 323:1308–1308. https://doi.org/10.1126/science.1169202.
8. Peris JB, Davis P, Cuevas JM, Nebot MR, Sanjuán R. 2010. Distribution of fitness effects caused by single-nucleotide substitutions in bacteriophage f1. Genetics 185:603–609. https://doi.org/10.1534/genetics.110.115162.
9. Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, García-Arenal F. 2002. The rate and character of spontaneous mutation in an RNA virus. Genetics 162:1505–1511.
10. Elena SF, Moya A. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J Evol Biol 12:1078–1088. https://doi.org/10.1046/j.1420-9101.1999.00110.x.
11. Visher E, Whitefield SE, McCrone JT, Fitzsimmons W, Lauring AS. 2016. The mutational robustness of influenza A virus. PLoS Pathog 12:e1005856. https://doi.org/10.1371/journal.ppat.1005856.
12. Sanjuán R, Moya A, Elena SF. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci U S A 101:8396–8401. https://doi.org/10.1073/pnas.0400146101.
13. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, Lu X, Scherbakova S, Graham RL, Baric RS, Stockwell TB, Spiro DJ, Denison MR. 2010. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog 6:e1000896. https://doi.org/10.1371/journal.ppat.1000896.
14. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J Virol 81:12135–12144. https://doi.org/10.1128/JVI.01296-07.
15. Smith EC, Case JB, Blanc H, Isakov O, Shomron N, Vignuzzi M, Denison MR. 2015. Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. J Virol 89:6418–6426. https://doi.org/10.1128/JVI.00110-15.
16. Smith EC, Blanc H, Surdel MC, Vignuzzi M, Denison MR. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. PLoS Pathog 9:e1003565. https://doi.org/10.1371/journal.ppat.1003565.
17. Sexton NR, Smith EC, Blanc H, Vignuzzi M, Peersen OB, Denison MR. 2016. Homology-based identification of a mutation in the coronavirus RNA-dependent RNA polymerase that confers resistance to multiple mutagens. J Virol 90:7415–7428. https://doi.org/10.1128/JVI.00080-16.
18. Arias A, Arnold JJ, Sierra M, Smidansky ED, Domingo E, Cameron CE. 2008. Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. J Virol 82:12346–12355. https://doi.org/10.1128/JVI.01297-08.
19. Zeng J, Wang H, Xie X, Yang D, Zhou G, Yu L. 2013. An increased replication fidelity mutant of foot-and-mouth disease virus retains fitness in vitro and virulence in vivo. Antiviral Res 100:1–7. https://doi.org/10.1016/j.antiviral.2013.07.008.
20. Zeng J, Wang H, Xie X, Li C, Zhou G, Yang D, Yu L. 2014. Ribavirin-resistant variants of foot-and-mouth disease virus: the effect of restricted quasispecies diversity on viral virulence. J Virol 88:4008–4020. https://doi.org/10.1128/JVI.03594-13.
21. Xie X, Wang H, Zeng J, Li C, Zhou G, Yang D, Yu L. 2014. Foot-and-mouth disease virus low-fidelity polymerase mutants are attenuated. Arch Virol 159:2641–2650. https://doi.org/10.1007/s00705-014-2126-z.
22. Sierra M, Airaksinen A, González-López C, Agudo R, Arias A, Domingo E. 2007. Foot-and-mouth disease virus mutant with decreased sensitivity to ribavirin: implications for error catastrophe. J Virol 81:2012–2024. https://doi.org/10.1128/JVI.01606-06.
23. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439:344–348. https://doi.org/10.1038/nature04388.
24. Vignuzzi M, Wendt E, Andino R. 2008. Engineering attenuated virus vaccines by controlling replication fidelity. Nat Med 14:154–161. https://doi.org/10.1038/nm1726.
25. Pfeiffer JK, Kirkegaard K. 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. Proc Natl Acad Sci U S A 100:7289–7294. https://doi.org/10.1073/pnas.1232294100.
26. Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE. 2005. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J Biol Chem 280:25706–25716. https://doi.org/10.1074/jbc.M503444200.
27. Liu X, Yang X, Lee CA, Moustafa IM, Smidansky ED, Lum D, Arnold JJ, Cameron CE, Boehr DD. 2013. Vaccine-derived mutation in motif D of poliovirus RNA-dependent RNA polymerase lowers nucleotide incorporation fidelity. J Biol Chem 288:32753–32765. https://doi.org/10.1074/jbc.M113.484428.
28. Weeks SA, Lee CA, Zhao Y, Smidansky ED, August A, Arnold JJ, Cameron CE. 2012. A polymerase mechanism-based strategy for viral attenuation and vaccine development. J Biol Chem 287:31618–31622. https://doi.org/10.1074/jbc.C112.401471.
29. Korboukh VK, Lee CA, Acevedo A, Vignuzzi M, Xiao Y, Arnold JJ, Hemperly S, Graci JD, August A, Andino R, Cameron CE. 2014. RNA virus population diversity, an optimum for maximal fitness and virulence. J Biol Chem 289:29531–29544. https://doi.org/10.1074/jbc.M114.592303.
30. Coffey LL, Beeharry Y, Bordería AV, Blanc H, Vignuzzi M. 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. Proc Natl Acad Sci U S A 108:16038–16043. https://doi.org/10.1073/pnas.1111650108.
31. Rozen-Gagnon K, Stapleford KA, Mongelli V, Blanc H, Failloux AB, Saleh MC, Vignuzzi M. 2014. Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. PLoS Pathog 10:e1003877. https://doi.org/10.1371/journal.ppat.1003877.
32. Cheung PPH, Watson SJ, Choy KT, Fun Sia S, Wong DDY, Poon LLM, Kellam P, Guan Y, Malik Peiris JS, Yen HL. 2014. Generation and characterization of influenza A viruses with altered polymerase fidelity. Nat Commun 5:4794. https://doi.org/10.1038/ncomms5794.
33. Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, Gong P, Blanc H, Peersen OB, Vignuzzi M. 2012. Coxsackievirus B3 mutator strains are attenuated in vivo. Proc Natl Acad Sci U S A 109:E2294–E2303. https://doi.org/10.1073/pnas.1204022109.
34. Campagnola G, McDonald S, Beaucourt S, Vignuzzi M, Peersen OB. 2015. Structure-function relationships underlying the replication fidelity of viral RNA-dependent RNA polymerases. J Virol 89:275–286. https://doi.org/10.1128/JVI.01574-14.
35. Sadeghipour S, McMinn PC. 2013. A study of the virulence in mice of high copying fidelity variants of human enterovirus 71. Virus Res 176:265–272. https://doi.org/10.1016/j.virusres.2013.06.019.

36. Sadeghipour S, Bek EJ, McMinn PC. 2013. Ribavirin-resistant mutants of human enterovirus 71 express a high replication fidelity phenotype during growth in cell culture. J Virol 87:1759–1769. https://doi.org/10.1128/JVI.02139-12.
37. Meng T, Kwang J. 2014. Attenuation of human enterovirus 71 high-replication-fidelity variants in AG129 mice. J Virol 88:5803–5815. https://doi.org/10.1128/JVI.00289-14.
38. Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS. 2012. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. Nat Med 18:1820–1826. https://doi.org/10.1038/nm.2972.
39. Smith EC, Sexton NR, Denison MR. 2014. Thinking outside the triangle: replication fidelity of the largest RNA viruses. Annu Rev Virol 1:111–132. https://doi.org/10.1146/annurev-virology-031413-085507.
40. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. Virus Res 117:17–37. https://doi.org/10.1016/j.virusres.2006.01.017.
41. Brister JR, Ako-Adjei D, Bao Y, Blinkova O. 2014. NCBI viral genomes resource. Nucleic Acids Res 43:D571–D577. https://doi.org/10.1093/nar/gku1207.
42. Minskaia E, Hertzig T, Gorbalenya AE, Campanacci V, Cambillau C, Canard B, Ziebuhr J. 2006. Discovery of an RNA virus 3′->5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc Natl Acad Sci U S A 103:5108–5113. https://doi.org/10.1073/pnas.0508200103.
43. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LLM, Guan Y, Rozanov M, Spaan WJM, Gorbalenya AE. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331:991–1004. https://doi.org/10.1016/S0022-2836(03)00865-9.
44. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, Lou Z, Yan L, Zhang R, Rao Z. 2015. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc Natl Acad Sci U S A 112:9436–9441. https://doi.org/10.1073/pnas.1508686112.
45. Zuo Y, Deutscher MP. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res 29:1017–1026. https://doi.org/10.1093/nar/29.5.1017.
46. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. 2012. RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc Natl Acad Sci U S A 109:9372–9377. https://doi.org/10.1073/pnas.1201130109.
47. Yokomori K, Banner LR, Lai MMC. 1991. Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses. Virology 183:647–657. https://doi.org/10.1016/0042-6822(91)90994-M.
48. Luytjes W, Bredenbeek PJ, Noten AF, Horzinek MC, Spaan WJ. 1988. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology 166:415–422. https://doi.org/10.1016/0042-6822(88)90512-0.
49. Yokomori K, Lai MM. 1991. Mouse hepatitis virus S RNA sequence reveals that nonstructural proteins ns4 and ns5a are not essential for murine coronavirus replication. J Virol 65:5605–5608.
50. Lissenberg A, Vrolijk MM, van Vliet ALW, Langereis MA, de Groot-Mijnes JDF, Rottier PJM, de Groot RJ. 2005. Luxury at a cost? Recombinant mouse hepatitis viruses expressing the accessory hemagglutinin esterase protein display reduced fitness in vitro. J Virol 79:15054–15063. https://doi.org/10.1128/JVI.79.24.15054-15063.2005.
51. Gadlage MJ, Graham RL, Denison MR. 2008. Murine coronaviruses encoding nsp2 at different genomic loci have altered replication, protein expression, and localization. J Virol 82:11964–11969. https://doi.org/10.1128/JVI.01126-07.
52. Zhao L, Jha BK, Wu A, Elliott R, Ziebuhr J, Gorbalenya AE, Silverman RH, Weiss SR. 2012. Antagonism of the interferon-induced OAS-RNase L pathway by murine coronavirus ns2 protein is required for virus replication and liver pathology. Cell Host Microbe 11:607–616. https://doi.org/10.1016/j.chom.2012.04.011.
53. Li Y, Weiss SR. 2016. Antagonism of RNase L is required for murine coronavirus replication in Kupffer cells and liver sinusoidal endothelial cells but not in hepatocytes. J Virol 90:9826–9832. https://doi.org/10.1128/JVI.01423-16.
54. Zhang R, Jha BK, Ogden KM, Dong B, Zhao L, Elliott R, Patton JT, Silverman RH, Weiss SR. 2013. Homologous 2′,5′-phosphodiesterases from disparate RNA viruses antagonize antiviral innate immunity. Proc Natl Acad Sci U S A 110:13114–13119. https://doi.org/10.1073/pnas.1306917110.
55. Sui B, Huang J, Jha BK, Yin P, Zhou M, Fu ZF, Silverman RH, Weiss SR, Peng G, Zhao L. 2016. Crystal structure of the mouse hepatitis virus ns2 phosphodiesterase domain that antagonizes RNase L activation. J Gen Virol 97:880–886. https://doi.org/10.1099/jgv.0.000395.
56. Schwarz B, Routledge E, Siddell SG. 1990. Murine coronavirus nonstructural protein ns2 is not essential for virus replication in transformed cells. J Virol 64:4784–4791.
57. Zhao L, Rose KM, Elliott R, Van Rooijen N, Weiss SR. 2011. Cell-type-specific type I interferon antagonism influences organ tropism of murine coronavirus. J Virol 85:10058–10068. https://doi.org/10.1128/JVI.05075-11.
58. Pereira-Gómez M, Sanjuán R. 2014. Delayed lysis confers resistance to the nucleoside analogue 5-fluorouracil and alleviates mutation accumulation in the single-stranded DNA bacteriophage ϕX174. J Virol 88:5042–5049. https://doi.org/10.1128/JVI.02147-13.
59. Ng KKS, Arnold JJ, Cameron CE. 2008. Structure-function relationships among RNA-dependent RNA polymerases. Curr Top Microbiol Immunol 320:137–156. https://doi.org/10.1007/978-3-540-75157-1_7.
60. Brockway SM, Clay CT, Lu XT, Denison MR. 2003. Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J Virol 77:10515–10527. https://doi.org/10.1128/JVI.77.19.10515-10527.2003.
61. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GMC, Ruben M, Overkleeft HS, van Veelen PA, Samborskiy DV, Kravchenko AA, Leontovich AM, Sidorov IA, Snijder EJ, Posthuma CC, Gorbalenya AE. 2015. Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res 43:8416–8434. https://doi.org/10.1093/nar/gkv838.
62. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.
63. Bouvet M, Lugari A, Posthuma CC, Zevenhoven JC, Bernard S, Betzi S, Imbert I, Canard B, Guillemot JC, Lécine P, Pfefferle S, Drosten C, Snijder EJ, Decroly E, Morelli X. 2014. Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. J Biol Chem 289:25783–25796. https://doi.org/10.1074/jbc.M114.577353.
64. Donaldson EF, Sims AC, Graham RL, Denison MR, Baric RS. 2007. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. J Virol 81:6356–6368. https://doi.org/10.1128/JVI.02805-06.
65. Crotty S, Cameron C, Andino R. 2002. Ribavirin's antiviral mechanism of action: lethal mutagenesis? J Mol Med 80:86–95. https://doi.org/10.1007/s00109-001-0308-0.
66. Pathak VK, Temin HM. 1992. 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. J Virol 66:3093–3100.
67. Carroll SS, Tomassini JE, Bosserman M, Getty K, Stahlhut MW, Eldrup AB, Bhat B, Hall D, Simcoe AL, LaFemina R, Rutkowski CA, Wolanski B, Yang Z, Migliaccio G, De Francesco R, Kuo LC, MacCoss M, Olsen DB. 2003. Inhibition of hepatitis C virus RNA replication by 2′-modified nucleoside analogs. J Biol Chem 278:11979–11984. https://doi.org/10.1074/jbc.M210914200.
68. Derbyshire V, Grindley ND, Joyce CM. 1991. The The 3′-5′ exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. EMBO J 10:17–24.
69. Becares M, Pascual-Iglesias A, Nogales A, Sola I, Enjuanes L, Zuñiga S. 2016. Mutagenesis of coronavirus nsp14 reveals its potential role in modulation of the innate immune response. J Virol 90:5399–5414. https://doi.org/10.1128/JVI.03259-15.
70. Cameron CE, Moustafa IM, Arnold JJ. 2016. Fidelity of nucleotide incorporation by the RNA-dependent RNA polymerase from poliovirus DNA replication across taxa, 1st ed. Elsevier Inc., New York, NY.
71. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, Snijder EJ, Canard B, Imbert I. 2014. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci U S A 111:E3900–E3909. https://doi.org/10.1073/pnas.1323705111.
72. Imbert I, Guillemot JC, Bourhis JM, Bussetta C, Coutard B, Egloff MP, Ferron F, Gorbalenya AE, Canard B. 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J 25:4933–4942. https://doi.org/10.1038/sj.emboj.7601368.
73. Adedeji AO, Marchand B, Te Velthuis AJ, Snijder EJ, Weiss S, Eoff RL, Singh K, Sarafianos SG. 2012. Mechanism of nucleic acid unwinding by SARS-CoV helicase. PLoS One 7:e36521. https://doi.org/10.1371/journal.pone.0036521.

74. Chaudhuri M, Song L, Parris DS. 2003. The herpes simplex virus type 1 DNA polymerase processivity factor increases fidelity without altering pre-steady-state rate constants for polymerization or excision. J Biol Chem 278:8996–9004. https://doi.org/10.1074/jbc.M210023200.
75. Gu S, Li W, Zhang H, Fleming J, Yang W, Wang S, Wei W, Zhou J, Zhu G, Deng J, Hou J, Zhou Y, Lin S, Zhang XE, Bi L. 2016. The β2 clamp in the Mycobacterium tuberculosis DNA polymerase III $\alpha\beta2\varepsilon$ replicase promotes polymerization and reduces exonuclease activity. Sci Rep 6:18418. https://doi.org/10.1038/srep18418.
76. Stapleford KA, Rozen-Gagnon K, Das PK, Saul S, Poirier EZ, Blanc H, Vidalain PO, Merits A, Vignuzzi M. 2015. Viral polymerase-helicase complexes regulate replication fidelity to overcome intracellular nucleotide depletion. J Virol 89:11233–11244. https://doi.org/10.1128/JVI.01553-15.
77. Agudo R, de la Higuera I, Arias A, Grande-Pérez A, Domingo E. 2016. Involvement of a joker mutation in a polymerase-independent lethal mutagenesis escape mechanism. Virology 494:257–266. https://doi.org/10.1016/j.virol.2016.04.023.
78. Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, Snijder EJ, Gorbalenya AE, Cambillau C, Canard B. 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc Natl Acad Sci U S A 101:3792–3796. https://doi.org/10.1073/pnas.0307877101.
79. Bhardwaj K, Guarino L, Kao CC. 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J Virol 78:12218–12224. https://doi.org/10.1128/JVI.78.22.12218-12224.2004.
80. Sanjuán R, Cuevas JM, Furió V, Holmes EC, Moya A. 2007. Selection for robustness in mutagenized RNA viruses. PLoS Genet 3:e93. https://doi.org/10.1371/journal.pgen.0030093.
81. Goldhill D, Lee A, Williams ESCP, Turner PE. 2014. Evolvability and robustness in populations of RNA virus Φ6. Front Microbiol 5:35. https://doi.org/10.3389/fmicb.2014.00035.
82. Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE. 2005. Evolution of mutational robustness in an RNA virus. PLoS Biol 3:e381. https://doi.org/10.1371/journal.pbio.0030381.
83. Lauring AS, Acevedo A, Cooper SB, Andino R. 2012. Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. Cell Host Microbe 12:623–632. https://doi.org/10.1016/j.chom.2012.10.008.
84. Lauring AS, Frydman J, Andino R. 2013. The role of mutational robustness in RNA virus evolution. Nat Rev Microbiol 11:327–336. https://doi.org/10.1038/nrmicro3003.
85. Aguilera ER, Erickson AK, Jesudhasan PR, Robinson CM, Pfeiffer JK. 2017. Plaques formed by mutagenized viral populations have elevated coinfection frequencies. mBio 8:e02020-16. https://doi.org/10.1128/mBio.02020-16.
86. Graci JD, Gnädig NF, Galarraga JE, Castro C, Vignuzzi M, Cameron CE. 2012. Mutational robustness of an RNA virus influences sensitivity to lethal mutagenesis. J Virol 86:2869–2873. https://doi.org/10.1128/JVI.05712-11.
87. O'Dea EB, Keller TE, Wilke CO. 2010. Does mutational robustness inhibit extinction by lethal mutagenesis in viral populations? PLoS Comput Biol 6:e1000811. https://doi.org/10.1371/journal.pcbi.1000811.
88. Yount B, Denison MR, Weiss SR, Baric RS. 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J Virol 76:11065–11078. https://doi.org/10.1128/JVI.76.21.11065-11078.2002.

Supplemental Figure S1.



APPENDIX D. Nelson's Textbook of Pediatrics, 21st edition, Chapter 264: Coronaviruses

The below text will be published in the upcoming 21st edition of the Nelson Textbook of Pediatrics.

Chapter 264

Coronaviruses

Kevin W. Graepel and **Mark R. Denison**

Coronaviruses are increasingly recognized as important human pathogens. They cause up to 15% of common colds and have been implicated in more serious diseases, including croup, asthma exacerbations, bronchiolitis, and pneumonia. Evidence also suggests that coronaviruses may cause enteritis or colitis in neonates and infants and may be underappreciated as agents of meningitis or encephalitis. Four coronaviruses are endemic in humans: human coronaviruses **(HCoVs) 229E**, **OC43**, **NL63**, and **HKU1**. In addition, 2 epidemics of previously unknown coronaviruses caused significant respiratory distress and high mortality rates among infected individuals. The discoveries of **SARS-associated coronavirus (SARS-CoV)** in 2003, the cause of **severe acute respiratory syndrome (SARS)**, and of **Middle East respiratory syndrome coronavirus (MERS-CoV)** in 2012, support the potential for coronaviruses to emerge from animal hosts such as bats and camels and become important human pathogens.

Etiology

Coronaviruses are enveloped viruses of medium to large size (80-220 nm) that possess the largest known single-stranded positive-sense RNA genomes. These viruses encode the protein nsp14-ExoN, which is the first known RNA proofreading enzyme and is likely responsible for the evolution of the large and complex coronavirus genome. Coronaviruses derive their name from the characteristic surface projections of the spike protein, which give a corona or crown-like appearance on negative-stain electron microscopy. Coronaviruses are organized taxonomically by a lettering system based on genomic phylogenetic relationships. Alphacoronaviruses include HCoV-229E and HCoV-NL63. Betacoronaviruses include 4 human pathogens and are commonly divided into 4 lineages, without formal taxonomic recognition. HCoV-OC43 and the HCoV-HKU1 are in lineage A, while SARS-CoV falls in lineage B. Lineages C and D were exclusively comprised of bat coronaviruses until the discovery of MERS-CoV, which aligns with lineage C. Gammacoronaviruses and deltacoronaviruses presently include exclusively nonhuman pathogens.

In 2002-2003, coronaviruses received international attention during the SARS outbreak, which was responsible for more than 800 deaths in 30 countries. SARS-CoV, a novel coronavirus at the time of the epidemic, was found to be the causative agent of SARS. The detection of SARS-like

coronaviruses in a live animal market in the Guangdong province in Southern China, along with serologic evidence of exposure in food handlers in the same market, suggest that these markets may have facilitated the spread of SARS-CoV to humans from an animal reservoir. Subsequent studies identified SARS-like coronaviruses in fecal specimens from asymptomatic Chinese horseshoe bats that are very closely related, but not direct precursors to, SARS-CoV and are capable of infecting human cells. Thus, although bats are a reservoir for SARS-CoV-like precursors, the precise antecedent to SARS-CoV remains to be identified.

In June 2012, another novel coronavirus, **MERS-CoV**, was isolated from a man with acute pneumonia and renal failure in Saudi Arabia. As of March 1, 2017, the WHO has recorded nearly 2000 confirmed cases of MERS, with nearly 700 deaths worldwide (~35% mortality). Dromedary camels are the likely reservoir host for MERS-CoV. **MERS-CoV** differs from SARS in that it seems to be less communicable, although human-to-human transmission has been documented. **MERS-CoV** uses dipeptidyl peptidase 4 and carcinoembryonic antigen-like cell-adhesion molecule 5 as its cellular and co-receptor, respectively; SARS-CoV utilizes ACE-2. With this receptor specificity, **MERS-CoV** can infect cells from several animal lineages, including human, pig, and bat, suggesting the possibility of movement between multiple species.

Epidemiology

Seroprevalence studies have demonstrated that antibodies against 229E and OC43 increase rapidly during early childhood, so that by adulthood 90-100% of persons are seropositive. Although less information is available for HKU1 and NL63, available studies demonstrate similar patterns of seroconversion to these viruses during early childhood. Although some degree of strain-specific protection may be afforded by recent infection, reinfections are common and occur despite the presence of strain-specific antibodies. Attack rates are similar in different age groups. Although infections occur throughout the year, there is a peak during the winter and early spring for each of these HCoVs. In the United States, outbreaks of OC43 and 229E have occurred in 2- to 3-year alternating cycles. Independent studies of viral etiologies of upper and lower respiratory infections during the same period, but from different countries, have confirmed that all known HCoVs have a worldwide distribution. Studies using both viral culture and polymerase chain reaction (PCR) multiplex assays demonstrate that coronaviruses often occur as coinfections with other respiratory viruses, including respiratory syncytial virus, adenovirus, rhinovirus, or human metapneumovirus. Volunteer studies demonstrated that OC43 and 229E are transmitted predominantly through the respiratory route. Droplet spread appears to be most important, although aerosol transmission may also occur.

There have been no identified natural or laboratory-acquired cases of SARS-CoV since 2004, but the mechanisms of introduction, spread, and disease remain important for potential animal-to-human transmission and disease. The primary mode of SARS-CoV transmission occurred through direct or indirect contact of mucous membranes with infectious droplets or fomites. Aerosol transmission was less common, occurring primarily in the setting of endotracheal intubation, bronchoscopy, or treatment with aerosolized medications. Fecal-oral transmission did not appear to be an efficient mode of transmission, but may have occurred because of the profuse diarrhea observed in some patients. The seasonality of SARS-CoV remains unknown. SARS-CoV is not highly infectious, with generally only 2-4 secondary cases resulting from a single infected adult. During the SARS epidemic, a small number of infected individuals, "superspreaders," transmitted infection to a much larger number of persons, but the mechanism

for this high degree of spread remains unknown. In contrast, persons with mild disease, such as children younger than 12 years of age, rarely transmitted the infection to others. Infectivity correlated with disease stage; transmission occurred almost exclusively during symptomatic disease. During the 2003 outbreak, most individuals with SARS-CoV infection were hospitalized within 3-4 days of symptom onset. Consequently, most subsequent infections occurred within hospitals and involved either healthcare workers or other hospitalized patients.

As of March 1, 2017, the World Health organization has recorded 1905 laboratory-confirmed cases of MERS-CoV from 27 countries, all of which are linked to exposures in the Arabian peninsula (approximately 80% in Saudi Arabia). Though the route of transmission between animals and humans is not fully understood, MERS-CoV is proposed to have repeatedly entered the human population through contact with respiratory secretions of dromedary camels and possibly to raw camel products (e.g. unpasteurized milk). Antibodies to MERS-CoV are found in dromedaries throughout the Middle East, and strains identical to human MERS-CoV isolates have been found in camels in Egypt, Oman, Qatar, and Saudi Arabia. These strains do not appear to be highly pathogenic or virulent in camels and have likely circulated within dromedaries for >30 years. Despite well-documented zoonotic transmission, most reported cases are linked human-to-human transmission in healthcare settings, including outbreaks in Jordan, South Korea, and Saudi Arabia in 2015 and 2016. Risk factors for nosocomial MERS-CoV outbreaks include overcrowded emergency departments, delayed diagnosis or isolation, and poor infection control practices. Transmission most likely occurs through respiratory droplets and is thus a greater risk during aerosol-generating procedures. Outside of healthcare settings, human-to-human transmission has been infrequently documented and is primarily associated with close contact within households. No sustained human-to-human transmission has yet been reported.

Pathogenesis of SARS and MERS

Severe disease in SARS and MERS likely results from both direct virologic damage and subsequent immunopathology. Studies with SARS-CoV in human airway epithelial cell cultures indicate that ciliated cells are principal targets for infection, while MERS-CoV preferentially infects bronchial epithelial cells, type I and II pneumocytes, and vascular endothelial cells. Substantial viral loads can be detected in the lower respiratory tract and in blood for both viruses. However, late progression to severe disease appears independent of the quantity and timing of viremia. Thus, excessive host immune responses likely play an important role in the progression to lower respiratory disease and acute respiratory distress syndrome. CoV infections are associated with massive elaboration of inflammatory cytokines and recruitment of inflammatory cells. The roles for inflammatory cells are controversial, with cytotoxic T cells and macrophages implicated variously in immune protection and immunopathology. Recapitulation of human clinical features in animal models of MERS-CoV infection remains challenging, but promising new models are in development.

Clinical Manifestations

Respiratory Infections

Even though up to 50% of respiratory tract infections with OC43 and 229E are asymptomatic, coronaviruses are still responsible for up to 15% of common colds and can cause fatal disease. Cold symptoms caused by HCoVs are indistinguishable from those caused by rhinoviruses and other respiratory viruses. The average incubation period is 2-4 days, with symptoms typically

lasting 4-7 days. Rhinorrhea, cough, sore throat, malaise, and headache are the most common symptoms. Fever occurs in up to 60% of cases. Coronavirus NL63 is a cause of croup in children younger than 3 years of age. Coronavirus infections are linked to episodes of wheezing in asthmatic children, albeit at a lower frequency and severity than observed with rhinovirus and respiratory syncytial virus infections. Lower respiratory tract infections, including bronchiolitis and pneumonia, are also reported in immunocompetent and immunocompromised children and adults. As with respiratory syncytial virus or rhinovirus, coronavirus detection in upper respiratory infections is frequently associated with acute otitis media and can be isolated from middle ear fluid.

Nonrespiratory Sequelae

There is some evidence to support a role for coronaviruses in human gastrointestinal disease, particularly in young children. Coronavirus-like particles have been detected by electron microscopy in the stools of infants with nonbacterial gastroenteritis. In addition, several outbreaks in neonatal intensive care units of gastrointestinal disease characterized by diarrhea, bloody stools, abdominal distention, bilious gastric aspirates, and classic necrotizing enterocolitis have also been associated with the presence of coronavirus-like particles in stools. In older children and adults, coronavirus-like viruses have been observed with similar frequency in symptomatic and asymptomatic individuals, making it difficult to discern if they are pathogenic in the gastrointestinal tract. Coronaviruses are well-known causes of neurologic disease in animals, including demyelinating encephalitis, but their role in causing human neurologic disease remains unclear. They have been detected by culture, in situ hybridization, and reverse transcriptase PCR (RT-PCR) in brain tissue from a few patients with multiple sclerosis. HCoV-OC43 has been detected by RT-PCR in the spinal fluid, nasopharynx, or brain biopsy specimens of two children with acute encephalomyelitis. However, coronavirus RNA has also been recovered from the spinal fluid and brain tissue of adults without neurologic disease.

Severe Acute Respiratory Syndrome–Associated Coronavirus

SARS-CoV infections in teenagers and adults included a viral replication phase and an immunologic phase. During the viral replication phase, there was a progressive increase in viral load that reached its peak during the 2nd week of illness. The appearance of specific antibodies coincided with peak viral replication. The clinical deterioration that typified the 2nd and 3rd week of illness was characterized by a decline in the viral load and evidence of tissue injury, likely from cytokine-mediated immunity. The explanation for milder clinical disease in children younger than 12 years of age has not been determined. Seroepidemiologic studies suggest that asymptomatic SARS-CoV infections were uncommon. The incubation period ranged from 1-14 days, with a median of 4-6 days. The clinical manifestations were nonspecific, most commonly consisting of fever, cough, malaise, coryza, chills or rigors, headache, and myalgia. Coryza was more common in children younger than 12 years of age, whereas systemic symptoms were seen more often in teenagers. Some young children had no respiratory symptoms. Gastrointestinal symptoms, including diarrhea and nausea or vomiting, occurred in up to one third of cases. The clinical course of SARS-CoV infection varied with age. Adults were most severely affected, with initial onset of fever, cough, chills, myalgia, malaise, and headache. Following an initial improvement at the end of the 1st week, fever recurred and respiratory distress developed, with dyspnea, hypoxemia, and diarrhea. These symptoms progressed in 20% of patients to acute respiratory distress syndrome and respiratory failure. Acute renal failure with histologic acute tubular necrosis was present in 6.9% of patients, likely a result of hypoxic kidney damage. Of SARS patients, 28.8% had abnormal urinalysis, with viral genome detectable by quantitative RT-PCR. In contrast, children younger than 12 years of age had a relatively mild nonspecific illness,

with only a minority experiencing significant lower respiratory tract disease and illness typically lasting less than 5 days. There were no deaths or acute respiratory distress syndrome in children younger than 12 years of age from SARS-CoV infection. Adolescents manifested increasing severity in direct correlation to increasing age; respiratory distress and hypoxemia were observed in 10-20% of patients, one-third of whom required ventilator support. The case fatality rate from SARS-CoV infection during the 2003 outbreak was 9.6%. No pediatric deaths were reported. The estimated case fatality rate according to age varied from <1% for those younger than 20 years of age to >50% for those older than 65 years of age.

Middle East Respiratory Syndrome Coronavirus

The incubation period of MERS-CoV is between 2-14 days and typically presents with nonspecific clinical features typical of acute febrile respiratory illnesses, including low-grade fever, rhinorrhea, sore throat, and myalgia. In mildly symptomatic cases, radiographic findings are typically normal. Severe disease is characterized by the acute respiratory distress syndrome with multilobular airspace disease, ground-glass opacities, and occasional pleural effusions on radiography. The median time between hospitalization and ICU transfer for critical illness is 2 days. Risk factors for severe disease include age >50 years and comorbidities like obesity, diabetes, COPD, end-stage renal disease, cancer, and immunosuppression. Specific host genetic risk factors have not been identified. Variation in clinical outcomes does not appear to be explained by viral strain-specific sequence variability. As with SARS, extrapulmonary manifestations are common in severe MERS disease. Gastrointestinal symptoms such as nausea, vomiting, and diarrhea occur in one-third of patients, and acute kidney injury has been documented in half of critically ill patients. Encephalitis-like neurological manifestations have been observed in three cases. Laboratory analyses typically detect leukopenia and lymphopenia with occasional thrombocytopenia, anemia, and aminotransferase elevations. The case fatality rate remains at 35%, though the true incidence of MERS-CoV infection is likely underestimated by existing data. Most patients have been adults, although children as young as 9 months of age have been infected. It is not known whether children are less susceptible to MERS-CoV or present with a different clinical picture.

Diagnosis

In the past, specific diagnostic tests for coronavirus infections were not available in most clinical settings. The use of conserved PCR primers for coronaviruses in multiplex RT-PCR viral diagnostic panels now allows widely available and sensitive detection of the viruses. Virus culture of primary clinical specimens remains a challenge for HCoVs HKU1, OC43, 229E, and NL63, even though both SARS-CoV and MERS-CoV can successfully be grown in culture from respiratory samples. Serodiagnosis with complement fixation, neutralization, hemagglutination inhibition, enzyme immunoassay, and Western blots have been used in the research setting. The diagnosis of SARS-CoV infection can be confirmed by serologic testing, detection of viral RNA using RT-PCR, or isolation of the virus in cell culture. Even though serology for SARS-CoV has sensitivity and specificity approaching 100%, antibodies are not detectable until 10 days after the onset of symptoms, and immunoglobulin G seroconversion may be delayed for up to 4 weeks. In addition, the SARS epidemic resulted in the inclusion of coronavirus-conserved primers in many diagnostic PCR multiplex assays such that coronaviruses may be more readily detected. Diagnosis of MERS-CoV should be guided by clinical features and an epidemiological link. The mainstay for laboratory confirmation of MERS-CoV infection is real-time RT-PCR. Screening should target the region upstream of the envelope gene (upE) followed by confirmation with an assay targeting open reading frame 1a. The best diagnostic sensitivity is achieved from lower

respiratory tract samples collected within the first week of infection, though MERS-CoV RNA can be detected in upper respiratory and blood samples. Alternatively, seroconversion can be documented by screening enzyme-linked immunosorbent assays followed by immunofluorescence microscopy. For all known endemic and emerging HCoVs, respiratory specimens (nasopharyngeal swabs or aspirates) are most likely to be positive, but in a setting of a possible novel coronavirus, serum or stool may be positive.

Treatment and Prevention

Coronavirus infections of humans are acute and self-limited, although persistent infection and shedding occurs in multiple animal models in the setting of minimal or no symptomology. There are no available antiviral agents for clinical use against coronaviruses, although strategies targeting conserved coronavirus proteases and coronavirus polymerases have been shown to block replication of the viruses *in vitro* and are in the drug development pipeline. Thus, treatment of SARS-CoV and MERS-CoV infections is primarily supportive. The role of antiviral and immune-modulating agents remains inconclusive, though several clinical trials are ongoing. Ribavirin was extensively used during the 2003 SARS-CoV outbreak, but is of questionable benefit given its poor in vitro activity against SARS-CoV at clinically relevant concentrations. The identification of the proofreading nsp14-exonuclease in multiple coronaviruses suggests that this activity may be important in resistance to antiviral nucleosides and RNA mutagens such as ribavirin. Systemic corticosteroid therapy may be associated with increased mortality in SARS-CoV and MERS-CoV and is thus not recommended unless indicated for another clinical condition. Meta-analysis of observational studies suggests that human convalescent plasma may reduce SARS mortality; the use of blood products has not been well-studied in MERS. Several monoclonal antibody preparations have shown positive results against SARS- and MERS-CoV in animal studies.

Challenges for development of effective vaccines targeted against OC43, 229E, HKU1, and NL63 include the fact that infections are rarely life-threatening and reinfection is the rule, even in the presence of natural immunity from previous infections. The durability of immunity to SARS-CoV and MERS-CoV is poorly understood. Nevertheless, effective vaccines for SARS-CoV and MERS-CoV are highly desirable but not yet available. A potential vaccine target is the viral spike protein, which could be delivered as a recombinant protein or by viral or DNA vectors. This approach appears to be effective against closely related strains of SARS-CoV but not necessarily early animal or human variants. A SARS-CoV vaccine approach that recently has shown success in animal models used a live recombinant SARS-CoV mutant with inactivated ExoN, demonstrating attenuation and protection in aged, immunocompromised mice. Approaches for rapid development of stably attenuated live viruses or broadly immunogenic and cross-protective protein immunogens continues to be a key area for future research. Although SARS-CoV demonstrated characteristics of symptomatic transmission that made it controllable by public health measures like quarantine, these characteristics cannot be assumed for future novel HCoVs. The recent discovery of MERS-CoV serves as a reminder that coronavirus emergence is both likely and unpredictable, making it important to continue studies of the replication, emergence, and transmission of coronaviruses. Additionally, strategies for rapid recovery, testing, and development of vaccines and neutralizing human monoclonal antibodies may be essential to prevent the high morbidity and mortality associated with previous epidemics.

Bibliography

Alper CM, Winther B, Mandel EM, et al: Rate of concurrent otitis media in upper respiratory tract infections with specific viruses, Arch Otolaryngol Head Neck Surg 135:17–21, 2009.
Arabi YM, Balkhy HH, Hayden FG, et al: Middle East Respiratory Syndrome, N Engl J Med 376(6):584-594, 2017.
Assiri A, McGeer A, Perl TM, et al: Hospital outbreak of Middle East respiratory syndrome coronavirus, N Engl J Med 369:407–416, 2013.
Bialek SR, Allen D, Alvaredo-Ramy R, et al: First confirmed cases of Middle East respiratory syndrome coronavirus (MERS-CoV) infection in the United States, updated information on the epidemiology of MERS-CoV infection, and guidance for the public, clinicians, and public health authorities–May 2014, MMWR Morb Mortal Wkly Rep 63:431–436, 2014.
Bitnun A, Read S, Tellier R, et al: Severe acute respiratory syndrome-associated coronavirus infection in Toronto children: a second look, Pediatrics 123:97–101, 2009.
Breban R, Riou J, Fontanet A: Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk, Lancet 382:694–699, 2013.
Centers for Disease Control and Prevention (CDC): Update on the epidemiology of middle east respiratory syndrome coronavirus (MERS-CoV) infection, and guidance for the public, clinicians, and public health authorities, MMWR Morb Mortal Wkly Rep 64(03):61–62, 2015.
Choi J, Kim M-G, Oh Y-K, Kim YB: Progress of Middle east respiratory syndrome coronavirus vaccines: a patent review, Expert Opin Therapeutic Patents 0(0):1-11, 2017
Dominguez SR, Robinson CC, Holmes KV: Detection of four human coronaviruses in respiratory infections in children: a one-year study in Colorado, J Med Virol 81:1597–1604, 2009.
Gomersall CD, Joynt GM: Middle East respiratory syndrome: new disease, old lessons, Lancet 381:2229–2230, 2013.
Graham RL, Becker MM, Eckerle LD, et al: A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised model of lethal disease, Nat Med 18(12):1820–1826, 2012.
Gralinksi LE, Baric RS: Molecular pathology of coronavirus infections, J Pathol 235(2):185-195, 2015.
Guery B, Poissy J, el Mansouf L, et al: Clinical features and viral diagnosis of two cases of infection with Middle East respiratory syndrome coronavirus: a report of nosocomial transmission, Lancet 381:2265–2272, 2013.
Hui DS: Tracking the transmission and evolution of MERS-CoV, Lancet 382:1962–1964, 2013.
Lessler J, Reich NG, Brookmeyer R, et al: Incubation periods of acute respiratory viral infections: a systematic review, Lancet Infect Dis 9:291–300, 2009.
Memish ZA, Zumla AI, Al-Hakeem RF, et al: Family cluster of Middle East respiratory syndrome coronavirus infections, N Engl J Med 368:2487–2494, 2013.
Memish ZA, Cotten M, Meyer B, et al: Human infection with MERS coronavirus after exposure to infected camels, Saudi Arabia, 2013, Emerg Infect Dis 20(6):1012-1015, 2014.
Perlman S, Zhao J: Human coronavirus EMC is not the same as severe acute respiratory syndrome coronavirus, MBio 4(1):e2–e13, 2013.
Smith EC, Sexton NR, Denison MD: Thinking outside the triangle: Replication fidelity of the largest RNA viruses, Ann Rev Virol 1(1):111-132, 2014.
Stempel HE, Martin ET, Kuypers J, et al: Multiple viral respiratory pathogens in children with bronchiolitis, Acta Paediatr 98:123–126, 2009.
Talbot HK, Crowe JE Jr, Edwards KM, et al: Coronavirus infection and hospitalizations for

acute respiratory illness in young children, J Med Virol 81:853–856, 2009.
van Boheemen S, de Graaf M, Lauber C, et al: Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans, MBio 3(6):e473–e512, 2012.
World Health Organization: Middle East respiratory syndrome coronavirus (MERS-CoV), www.who.int/emergencies/mers-cov/

APPENDIX E. Fitness barriers limit reversion of a proofreading-deficient coronavirus

This manuscript was posted to the BioRxiv.org (https://doi.org/10.1101/618249) on April 26, 2019 and can be accessed at the following web address: https://www.biorxiv.org/content/10.1101/618249v1. It was submitted to *Journal of Virology* on 4/29/2019.

Fitness barriers limit reversion of a proofreading-deficient coronavirus

Kevin W. Graepel[1,3], Maria L. Agostini[1,3], Xiaotao Lu[2], Nicole R. Sexton[4], and Mark R. Denison[1,2,3#]

[1]Department of Pathology, Microbiology, and Immunology; [2]Department of Pediatrics; [3]Vanderbilt Institute for Infection, Immunology and Inflammation (VI4), Vanderbilt University Medical Center, Nashville, TN, USA. [4]Department of Microbiology, Immunology, and Pathology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO, USA.

#Address correspondence to mark.denison@vumc.org

Running head: Barriers to reversion of a debilitated coronavirus

Keywords:
RNA virus, adaptive evolution, competitive fitness, coronavirus, exoribonuclease, plus-strand RNA virus, proofreading, replication fidelity

Abstract word count: 359
Text word count: 4423

ABSTRACT

The 3′-to-5′ exoribonuclease in coronavirus (CoV) nonstructural protein 14 (nsp14-ExoN) mediates RNA proofreading during genome replication. ExoN catalytic residues are arranged in three motifs: I (DE), II (E), III (D). Alanine substitution of the motif I residues (AA-E-D, four nucleotide substitutions) in murine hepatitis virus (MHV) and SARS-CoV yields viable mutants with impaired replication and fitness, increased mutation rates, and attenuated virulence *in vivo*. Despite these impairments, MHV- and SARS-CoV ExoN motif I AA mutants (ExoN-AA) have not reverted at motif I in diverse *in vitro* and *in vivo* environments, suggesting that profound fitness barriers prevent motif I reversion. To test this hypothesis, we engineered MHV-ExoN-AA

with 1, 2 or 3 nucleotide mutations along genetic pathways to AA-to-DE reversion. We show that engineered intermediate revertants were viable but had no increased replication or competitive fitness compared to MHV-ExoN-AA. In contrast, a low passage (P10) MHV-ExoN-AA showed increased replication and competitive fitness without reversion of ExoN-AA. Finally, engineered reversion of ExoN-AA to ExoN-DE in the presence of ExoN-AA passage-adaptive mutations resulted in significant fitness loss. These results demonstrate that while reversion is possible, at least one alternative adaptive pathway is more rapidly advantageous than intermediate revertants and may alter the genetic background to render reversion detrimental to fitness. Our results provide an evolutionary rationale for lack of ExoN-AA reversion, illuminate potential multi-protein replicase interactions and coevolution, and support future studies aimed at stabilizing attenuated CoV ExoN-AA mutants.

IMPORTANCE

Coronaviruses encode an exoribonuclease (ExoN) that is important for viral replication, fitness, and virulence, yet coronaviruses with a defective ExoN (ExoN-AA) have not reverted under diverse experimental conditions. In this study, we identify multiple impediments to MHV-ExoN-AA reversion. We show that ExoN-AA reversion is possible but evolutionarily unfavorable. Instead, compensatory mutations outside of ExoN-AA motif I are more accessible and beneficial than partial reversion. We also show that coevolution between replicase proteins over long-term passage partially compensates for ExoN-AA motif I but renders the virus inhospitable to a reverted ExoN. Our results reveal the evolutionary basis for the genetic stability of ExoN-inactivating mutations, illuminate complex functional and evolutionary relationships between coronavirus replicase proteins, and identify potential mechanisms for stabilization of ExoN-AA coronavirus mutants.

INTRODUCTION

The rapid evolution of RNA viruses represents a significant challenge for preventing, treating, and eradicating RNA viral diseases. High mutation rates in RNA viruses generate extensive opportunities to overcome evolutionary hurdles, such as antiviral drugs, host immunity, or engineered attenuating changes (1). The evolutionary pathways traversed by RNA viruses are shaped by natural selection, which will favor some evolutionary trajectories more than others

based on whether mutations are beneficial, deleterious, or neutral (2). Predicting the likely results of RNA virus evolution is an important step for anticipating viral emergence and for developing escape-resistant antiviral drugs and vaccines (3, 4).

Coronaviruses (CoVs) are a family of positive-sense RNA viruses that cause human illnesses ranging from the common cold to severe and lethal respiratory disease (5). All CoVs encode a proofreading exoribonuclease within nonstructural protein 14 (nsp14-ExoN) that is critical for replication, fidelity, fitness, and virulence, and ExoN-inactivation has been proposed as a strategy for live-attenuated vaccine development (6-15). As members of the DEDDh superfamily of exonucleases, CoV ExoNs hydrolyze nucleotides using four metal-coordinating amino acids arranged in 3 motifs: I (DE), II (E), III (D) (16, 17). Alanine substitution of ExoN motif I (DE-to-AA) disrupts ExoN biochemical activity in both SARS-CoV and human CoV 229E (hCoV-229E) (16, 18, 19). The betacoronaviruses murine hepatitis virus (MHV) and SARS-CoV tolerate disruption of ExoN activity [ExoN(-)] but display mutator phenotypes accompanied by defects in replication, competitive fitness, and evasion of innate immune responses (10, 13, 14). ExoN active site mutants in alphacoronaviruses, including transmissible gastroenteritis virus and hCoV-229E, have yet to be recovered and are proposed to be lethal for replication (19, 20).

Given the critical role of ExoN in CoV biology and the elevated mutation rate, we expected that natural selection would repeatedly drive reversion of the ExoN-inactivating substitutions. In line with this expectation, ExoN motif III mutants of SARS-CoV and MHV rapidly and repeatedly revert ((14) and unpublished observations). In contrast, we have never detected partial or complete reversion of ExoN motif I mutants (ExoN-AA) in SARS-CoV or MHV during 10 years of study and hundreds of experiments. More specifically, we have not detected consensus or minority variants of any kind at the motif I AA codons in either virus strain during acute infections and prolonged passage in tissue culture and following treatment with multiple nucleoside analogues (6-11, 13, 14). SARS-CoV-ExoN-AA also is stable during acute and persistent animal infections in immunocompetent and immune-compromised mice (12). The lack of primary reversion is not due simply to reduced adaptive capacity, as both SARS-CoV- and MHV-ExoN-AA can adapt for increased replication (7, 14). Most strikingly, long-term passage of MHV-ExoN-AA (250 passages, P250) yielded a highly fit population that had directly

compensated for defective proofreading through evolution of a likely high-fidelity RdRp (7). Yet, where primary reversion would have required just four total consensus mutations, MHV-ExoN-AA-P250 contained more than 170.

In this study, we sought to determine whether specific genetic or fitness barriers prevent primary reversion of ExoN motif I AA. To this end, we identified and engineered viable genetic pathways towards ExoN-AA motif I reversion in MHV (hereafter, ExoN-AA). Our results show that partial reversion did not confer a selective advantage compared to ExoN-AA. Further, ExoN-AA adapted within 10 passages for greater fitness than any of the intermediate revertants. Finally, restoration of WT-ExoN-DE in the setting of passage-selected mutations in the nsp12 RNA-dependent RNA polymerase (RdRp) and nsp14-ExoN exacted profound fitness costs. Together, these data are the first observation of an ExoN(-) CoV genotype-fitness landscape and identify multiple genetic barriers underlying ExoN(-) motif I stability in MHV. Further, they suggest extensive coevolution between MHV replicase proteins during adaptation and reveal potential strategies for stabilizing ExoN mutant CoVs.

RESULTS

Primary reversion of ExoN(-) motif I. MHV-ExoN(-), hereafter ExoN-AA, contains two engineered substitutions in each codon of motif I, such that complete reversion to WT-ExoN-DE requires mutations to all four sites (Figure 1A). Viral mutation rates in the absence of proofreading range from 10^{-4} to 10^{-6} mutations per nucleotide per round of replication (μ) (1). Assuming an ExoN-AA mutation rate of 10^{-4} μ and accounting for codon degeneracy, the probability of restoring the native amino acid sequence in a single round of replication is 10^{-18}. Only rarely do ExoN-AA titers exceed 10^6 PFU/mL, so it is unlikely that ExoN-AA could navigate this genetic barrier in a single infectious cycle. Thus, we hypothesized that ExoN-AA reversion, if possible, would proceed incrementally. To identify potential pathways towards ExoN-AA reversion, we examined the possible single-nucleotide substitutions surrounding A89 and A91 (Figure 1B). Three mutations are synonymous, and five mutations yield amino acids unlikely to coordinate with the positively-charged metals required for ExoN catalysis (glycine, valine, proline, threonine, and serine) (16, 19, 21, 22). One mutation per site can restore the acidic charge (i.e. AA-to-ED) but not the native amino acid. These variants have not been tested

in a CoV ExoN, but biochemical studies of *E. coli* DNA polymerase I ExoN mutants suggest that these conservative substitutions would not restore WT-like ExoN activity (23). We predicted stepwise pathways to ExoN-AA→DE reversion based on restoration of acidic charge followed by reversion to native amino acids (Figure 1C). We engineered and recovered variants in ExoN-AA requiring three mutations (3nt; ExoN-AD, ExoN-EA), two mutations (2nt; ExoN-DA, ExoN-ED, ExoN-AE), or one mutation (1nt; ExoN-DD, ExoN-EE) for reversion to WT-ExoN-DE (Table 1). We will hereafter refer to these mutants as intermediate revertants. All intermediate revertants generated viable progeny during recovery, demonstrating that reversion to WT-ExoN-DE along these pathways is theoretically possible. The 3nt and 2nt mutants were genetically stable during recovery, as confirmed by dideoxy sequencing. However, both 1nt mutants (ExoN-DD and ExoN-EE) reverted to WT-ExoN-DE during three independent recovery attempts, suggesting that these two variants are less fit than WT-ExoN-DE and demonstrating that reversion by 1nt mutation is readily accessible. To test whether the 3nt or 2nt mutants would revert more rapidly than ExoN-AA (4nt), we passaged three lineages of each mutant 10 times at multiplicities of infection (MOI) of 0.5 and 0.01 PFU/cell. We harvested supernatants and screened for reversion by visual inspection of plaque phenotypes at each passage. WT-ExoN-DE and WT-like viruses produce uniform, large plaques, while ExoN-AA-like viruses yield small, variably-sized plaques (13). When we observed mixed plaque phenotypes, we sequenced three large plaques from each lineage to confirm reversion. The 3nt (ExoN-AD and ExoN-EA) and 2nt (ExoN-DA and ExoN-ED) intermediate revertants showed no evidence of reversion over 10 passages at either MOI (Table 1). In contrast, the 2nt ExoN-AE contained WT-revertants by P2 in all lineages at MOI = 0.5 PFU/cell and by P8 in one lineage at MOI = 0.01 PFU/cell. Once observed, WT-revertants dominated the ExoN-AE population for the remaining passages. These data indicate that at least one 2nt mutation pathway can lead to full reversion in tissue culture. The probability of ExoN-AE arising during a single infectious cycle of ExoN-AA is low but theoretically achievable ($\sim 10^{-9}$), so ExoN-AA could conceivably revert within just two infectious cycles. However, complete reversion has never been observed even during prolonged passage or persistent infections, suggesting that additional barriers to the replication, fitness, or maintenance of intermediate revertants exist.

Partial reversion of MHV-ExoN(-) motif I does not confer a selective advantage. Because

the intermediate revertants are viable as recombinants but are not found in ExoN-AA populations, we hypothesized that they confer no selective advantage over ExoN-AA (8, 9, 13). To test this hypothesis, we first analyzed replication of the 3nt and 2nt intermediate revertants (Figure 2A). All variants achieved similar peak titers to ExoN-AA, but detailed examination of their kinetics suggested a potential delay of up to 1.5 hours for all intermediate revertants compared to ExoN-AA. Of note, ExoN-AE was the most delayed, and we detected WT-ExoN-DE revertants in two of three replicates, suggesting increased selective pressure against this variant. We next measured the competitive fitness of each intermediate revertant relative to a recombinant ExoN-AA containing seven silent mutations in the nsp2 coding region (ExoN-AA-reference). Intermediate revertants were mixed with an equal titer of ExoN-AA-reference at a combined MOI = 0.05 PFU/cell and passaged four times. The ratio of each intermediate revertant to ExoN-AA-reference was quantified at each passage by RT-qPCR, and the change in ratio over time was used to calculate their relative fitness. WT-ExoN-DE was significantly more fit than ExoN-AA, whereas the intermediate revertants (ExoN-AD, -EA, -DA, and -ED) had no increased fitness relative to ExoN-AA (Figure 2B). The apparent increased fitness of ExoN-AE resulted from all lineages reverting to WT-ExoN-DE during the experiment. Finally, our previous studies have shown that adaptation of ExoN-AA includes partial compensation for the replication fidelity defect, as measured by reduced susceptibility to the mutagen 5-fluorouracil (5-FU) (7-11, 24). None of the intermediate variants demonstrated statistically significant differences in 5-FU sensitivity as compared to ExoN-AA (Figure 2C). Thus, with the exception of the ExoN-AE→DE revertants, no 3nt and 2nt intermediate genotypes along our predicted pathway demonstrated an advantage in replication, fitness, or fidelity that would favor their maintenance or expansion in the viral population. Thus, natural selection is unlikely to drive ExoN-AA down these pathways towards reversion.

Secondary adaptations outside of ExoN-AA motif I increase fitness along alternative pathways. Although we did not find fitness advantages to intermediate revertants, we also did not identify profound fitness costs that would drive their immediate loss from populations. We have previously demonstrated that during 250 passages (P250), ExoN-AA can adapt for increased replication, fitness, and fidelity via secondary mutations outside of motif I (7). We tested whether secondary adaptive mutations could exceed the fitness of ExoN-AA and its

intermediate revertants. To examine the early adaptation of ExoN-AA, we studied passage 10 from the P250 passage series (Figure 3). ExoN-AA-P10 retains the ExoN-AA motif I genotype but has increased replication and reduced susceptibility to 5-FU, altogether manifesting in greater relative fitness (Figure 3) (7). We identified only six total mutations within ExoN-AA-P10 by dideoxy sequencing (Table 2), indicating that rapid adaptation of and compensation for ExoN-AA requires relatively few genetic changes at the consensus level. To test whether interactions between multiple mutations or population level effects contribute to ExoN-AA-P10 fitness, we isolated a plaque-purified clone of ExoN-AA-P10. The clone replicated to higher titers than the ExoN-AA-P10 population but had identical 5-FU sensitivity and relative fitness (Figure 3), indicating that genomes derived from a single virus plaque encode the adaptive changes required by the total population. Together, these data demonstrate that mutations outside of ExoN(-) motif I can confer greater fitness advantages than intermediate revertants even at early passages. These early adaptive mutations likely reduce the selective pressure for motif I reversion and place the intermediate revertants at a selective disadvantage.

Adaptive mutations in nsp12 and nsp14 that increase ExoN-AA fitness confer significant fitness costs to WT-ExoN-DE. Mutational fitness effects are highly dependent upon the genetic background (25-27). In addition to reducing selective pressure for reversion, mutations conferring increased fitness to ExoN-AA might also reduce the benefits of motif I reversion. We previously reported that long-term passage of ExoN-AA selects for secondary adaptive mutations in the nsp12 RdRp and nsp14 (nsp12-P250 and nsp14-P250) (7). Nsp12-P250 contains 7 nonsynonymous mutations that partially compensate for defective proofreading and increase ExoN-AA fitness. Nsp14-P250 contains 6 nonsynonymous mutations, including a conservative D-to-E substitution in ExoN motif III, and increases ExoN-AA fitness without compensating for defective proofreading. To test whether the fitness effects of passage-associated mutations in nsp12-P250 and nsp14-P250 depend upon the ExoN-AA genotype, we engineered a WT motif I (ExoN-DE) into viruses containing nsp12-P250 and nsp14-P250, alone and together, and analyzed replication, 5-FU sensitivity, and competitive fitness. Compared to WT-ExoN-DE, both ExoN-DE-nsp12-P250 and ExoN-DE-nsp14-P250 displayed delayed and decreased replication (Figure 4A). In 5-FU sensitivity assays, ExoN-DE-nsp14-P250 was indistinguishable from WT-ExoN-DE, while both variants containing nsp12-P250 (ExoN-DE-nsp12-P250 and ExoN-DE-nsp12/14-P250) were significantly more sensitive to 5-FU (Figure 4B). Finally, the nsp12-P250

and nsp14-P250 mutations significantly decreased fitness relative to WT-ExoN-DE (Figure 4C). We detected no statistical differences between the specific infectivity of WT-ExoN-DE and any of the nsp12-P250 and nsp14-P250 variants in isolated infections (Figure 4D). Thus, mutations in nsp12 and nsp14 that arose in the ExoN-AA background were detrimental to replication, mutagen sensitivity, and competitive fitness in the presence of a fully-reverted ExoN-DE. These results support the conclusion that the adaptive pathways available to ExoN-AA may stabilize the ExoN-AA genotype, reducing both the selective pressure for, and the potential benefits of, primary reversion.

DISCUSSION

In this study, we demonstrate that the stability of the ExoN(-) motif I genotype in MHV (ExoN-AA) is a consequence of the limitations and opportunities of the genetic landscape it explores during replication (Figure 5). Our results support a model in which the viable adaptive pathways leading to direct reversion of motif I from AA-to-DE are relatively flat on a fitness landscapes, as intermediate revertants remain phenotypically ExoN(-) and confer no fitness advantage over ExoN-AA. In contrast, at least one alternative adaptive pathway is readily accessible and imparts immediate fitness gains over ExoN-AA. We propose that even minimal alternative pathway adaptive fitness gains reduce the likelihood and benefits of motif I reversion, until eventually the changing genetic background renders reversion detrimental. These data and this model suggest that selection during replication favors immediate, incremental fitness gains along the most accessible pathway rather than dramatic fitness increases across a larger genetic barrier.

Our results also extend existing studies of CoV ExoN motif I. Motif I AA→DE mutations in the SARS-CoV nsp14-ExoN dramatically reduce nuclease activity in biochemical assays, but no study has examined the contributions of each residue independently (16, 18). Intermediate revertants of ExoN-AA did not display consistent or statistical differences in replication, 5-FU sensitivity, or competitive fitness relative to ExoN-AA, suggesting that they remain phenotypically ExoN(-) during infection and supporting previous studies that motif I DE is essential for WT ExoN function. Given these results, we were surprised to observe repeated reversion of the ExoN-AE but not the other two 2nt variants, ExoN-DA and ExoN-ED. One potential explanation is that the specific mutational bias of ExoN-AE makes the revertant

mutations more accessible than in ExoN-DA or ExoN-ED. Alternatively, if ExoN-AE has profound replication or fitness defects, selection could drive primary reversion more quickly away from this genotype. Consistent with this hypothesis, ExoN-AE reverted more quickly at a higher MOI, where natural selection acts more efficiently on a larger population size (Table 1) (28). Biochemical studies would be valuable to understand how nsp14-ExoN-AE differs from the other intermediate revertants, with an eye towards understanding the catalytic constraints and functional interactions with nascent CoV RNA.

CoV replication proceeds through the concerted action of multiple proteins proposed to resemble DNA replication holoenzymes (29). Due to the extensive interactions between CoV proteins, they must coevolve in a highly cooperative manner to maintain their essential functions. Consistent with this hypothesis, the fitness effects of mutations in nsp12-P250 and nsp14-P250 differ based on the motif I genotype; they are beneficial in ExoN-AA but detrimental in the WT-ExoN-DE background. In previous studies, it has been difficult to determine whether the fitness defects in ExoN(-) CoVs are directly linked to low-fidelity replication or through some other mechanism. Our data suggest that the proofreading function of nsp14-ExoN can be uncoupled from its more general role in replication (Figure 4), providing an opportunity to examine additional roles for this essential protein. Nsp12-P250 also will be an important tool for understanding the relationship between RdRp fidelity and ExoN proofreading during CoV replication and for studying replication complex assembly and interactions. Our studies suggest that compensatory mutations identified through long-term passage could stabilize the ExoN-AA genotype. In particular, the high-fidelity nsp12-P250 could reduce the probability of reversion by reducing mutational sampling within motif I (30), and both nsp12-P250 and nsp14-P250 render the MHV genome inhospitable to a WT-ExoN-DE. Together, these studies argue that experimental evolution can generate reagents to define critical interactions involved in CoV replication and can identify new strategies for stabilizing attenuated CoVs.

MATERIALS AND METHODS

Cell culture. Delayed brain tumor (DBT-9) cells (31) and baby hamster kidney 21 cells expressing the MHV receptor (BHK-R) (32) were maintained at 37°C in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% serum (HyClone FetalClone II, GE

Healthcare or Fetal Bovine Serum, Invitrogen), 100 U/mL penicillin and streptomycin (Gibco), and 0.25 μM amphotericin B (Corning). BHK-R cells were further supplemented with 0.8 mg/mL G418 selection antibiotic (Gibco). The infectious clone of the murine hepatitis virus strain A59 (MHV-A59; GenBank accession number AY910861) was used as the template for all recombinant viruses .

Determination of viral titer by plaque assay. Virus samples were serially diluted and inoculated on subconfluent DBT-9 cell monolayers in either 6- or 12-well format. Cells were overlaid with 1% agar in DMEM and incubated overnight at 37°C. Plates were fixed with 4% formaldehyde and agar plugs were removed. The number of plaques per well was counted by hand and used to calculate titer (32).

Plaque purification of viral populations. DBT cells were infected with serial dilutions of virus and overlaid with 1% agar in DMEM. Single plaques were isolated with glass Pasteur pipettes, resuspended in PBS containing calcium and magnesium, and inoculated onto fresh DBTs. This process was completed 3 times before generating experimental stocks.

Cloning and recovery of recombinant viruses. Site-directed mutagenesis in MHV genome fragments was performed using "round the horn" PCR (originally described in (33)). Briefly, adjacent primers containing the mutation of interest were 5′-phosphorylated using T4 polynucleotide kinase (NEB, M0201S) using the buffer from the T4 DNA ligase, which contains ATP (M0202S). PCR was performed on a plasmid template using the Q5 High-fidelity 2x Master Mix (NEB, M0492L), with primers at final concentration of 500nM. The linear amplification product was purified using the Promega Wizard SV Gel and PCR Clean-up System (Promega Corporation, A9282), and 4 μL was ligated at 16°C overnight with the T4 DNA ligase (NEB M0202S). After transformation into chemically-competent Top10 *E. coli* (lab-derived) and expansion in liquid culture, the MHV segment of each plasmid was sequenced. Viruses were constructed, rescued, and sequenced as described previously (7, 13, 32). Experimental stocks were generated by infecting a subconfluent 150 cm^2 flask of DBT-9 cells at MOI of 0.01 PFU/cell. Flasks were frozen at -80°C when monolayers were fully involved, approximately 20-28 hours post-infection depending on the variant. After thawing, the supernatant was clarified by

centrifugation at 4,000 x g (Sorvall RC 3B Plus; HA-6000A rotor) for 10 min at 4°C. For intermediate revertants, stocks were generated in serum-free DMEM and processed as above before being concentrated roughly 10-fold by centrifugation at 4,000 x g using Amicon Ultra-15 Centrifugal Filter Units, 100kDa (EMD Millipore, UFC910008). The virus titer of each stock was determined by plaque assay using DBT-9 cells as described above.

Passage of ExoN intermediate revertants. Intermediate revertants of ExoN-AA were passaged 10 times on subconfluent DBT-9 cell monolayers in 24-well plates at an estimated MOI of either 0.01 or 0.5 PFU/cell. Supernatants were harvested at 24 and 20 hours post-infection for MOI = 0.01 and 0.5 PFU/cell, respectively, and screened for WT reversion by plaque assay. At least three WT-like plaques were sequenced for each lineage to confirm motif I reversion.

Replication kinetics. Viral replication kinetics in DBT-9 cells were determined at indicated MOIs as described previously (11). Replicates were synchronized by 30-minute incubation at 4°C before transferring to the 37°C incubator. Supernatant (300 μL) was harvested at the indicated time points and titered by plaque assay.

Determination of specific infectivity. Subconfluent monolayers of DBT-9 cells in 24-well plates were infected with the indicated virus at MOI = 0.05 PFU/cell, and supernatant was harvested at 16 hours post-infection. Genomic RNA in supernatant was quantified using one-step reverse transcription quantitative RT-PCR (RT-qPCR) on TRIzol-extracted RNA as described previously (9). Briefly, genomic RNA was detected with a 5' 6-carboxyfluorescein (FAM) and 3' black hole quencher 1 (BHQ-1) labeled probe targeting nsp2 (Biosearch Technologies, Petaluma, CA), and RNA copy number was calculated by reference to an RNA standard derived from the MHV A fragment. Samples were plated in technical duplicate to minimize well-to-well variation. Titers were determined by plaque assay in DBT-9 cells, and specific infectivity was calculated as PFU per supernatant genomic RNA copy.

5-fluorouracil sensitivity assays. Stock solutions of 5-fluorouracil (Sigma F6627) were prepared in dimethyl sulfoxide (DMSO). Sensitivity assays were performed in 24-well plates at MOI = 0.01 PFU/cell, as previously described (7). Cells were incubated with drug for 30 minutes

prior to infection. Supernatants were harvested at 24 hours post-infection, and titers were determined by plaque assay.

Competitive fitness assays. ExoN-AA-reference and WT-ExoN-DE-reference viruses were marked with 7 consecutive silent mutations within nsp2 (wild-type: 1301-TTCGTCC-1307; reference: 1301-CAGCAGC-1307) by round the horn PCR, as described above. Competitions were performed in triplicate on DBT-9 cells in 12-well plates, plated at a density of 1 x 10^5 cells per well 24 hours prior to infection. Cells were infected at a total MOI of 0.1 PFU/cell (MOI = 0.05 PFU/cell each for competitor and reference virus). Supernatants were harvested 15 and 16 hours post-infection for experiments with ExoN-AA-reference and WT-ExoN-DE-reference, respectively, and passaged 4 times. Samples were titered between all passages to maintain total MOI of 0.1 PFU/cell. RNA was extracted from 70 μL of supernatant using QIAamp 96 virus QIAcube HT kit on the QIAcube HT System (Qiagen). Each RNA sample was analyzed by one-step RT-qPCR with two SYBR Green assays. Reference viruses were detected with forward primer SS-qPCR-Sil-F (5′-CTATGCTGTATACGGACAGCAGT-3′; 200nM final) and reverse primer SS-qPCR-R2 (5′-GGTGTCACCACAACAATCCAC-3′, 200nM final). Competitors were detected with forward primer SS-qPCR-WT-F (5′-CTATGCT-GTATACGGATTCGTCC-3′, 450 nM final) and reverse primer SS-qPCR-R2 (5′-GGTGTCAC-CACAACAATCCAC-3′, 450 nM final). RNA samples were diluted 1:100 prior to RT-qPCR with *Power* SYBR Green RNA-to-Ct 1-step kit (Applied Biosystems) according to the manufacturer's protocol. Duplicate wells were averaged, and values were excluded from subsequent analysis if the duplicate wells differed by > 0.5 Ct. The relative abundance of competitor and reference were determined by subtracting Ct thresholds ($\Delta Ct_{competitor} = Ct_{competitor} - Ct_{reference}$) and converted to reflect the fold-change in ratio ($\Delta ratio = 2^{-\Delta Ct\ competitor}$). The $\log_{10}\Delta ratio$ was plotted against passage number, and the change in $\log_{10}\Delta ratio$ (i.e. slope of linear regression) is the relative fitness. Note that regressions were fit only through P1-P4, as slight deviations in 1:1 ratio in the input (P0) can skew the slope.

Statistical analysis. GraphPad Prism 6 (La Jolla, CA) was used to perform statistical tests. Only the comparisons shown [e.g. ns or asterisk(s)] within the figure or legend were performed. In many cases the data were normalized to untreated controls. This was performed using GraphPad Prism 6. The number of replicate samples is denoted within each figure legend.

ACKNOWLEDGEMENTS

We thank members of the Denison laboratory and Seth Bordenstein for valuable discussions, as well as Andrea Pruijssers for critical review of the manuscript. This work was supported by United States Public Health Service awards R01-AI108197 (M.R.D), T32-GM007347 (K.W.G), F30-AI129229 (K.W.G), T32-AI089554 (N.R.S.), F31-AI133952 (M.L.A.), and T32-AI089554 (M.L.A.) all from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors declare no conflicts of interest.

REFERENCES

1. **Sanjuán R**, **Nebot MR**, **Chirico N**, **Mansky LM**, **Belshaw R**. 2010. Viral mutation rates. J Virol **84**:9733–9748.
2. **Domingo E**, **Sheldon J**, **Perales C**. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev **76**:159–216.
3. **Dolan PT**, **Whitfield ZJ**, **Andino R**. 2018. Mapping the Evolutionary Potential of RNA Viruses. Cell Host and Microbe **23**:435–446.
4. **Stern A**, **Yeh Te M**, **Zinger T**, **Smith M**, **Wright C**, **Ling G**, **Nielsen R**, **Macadam A**, **Andino R**. 2017. The Evolutionary Pathway to Virulence of an RNA Virus. Cell **169**:35–35.e19.
5. **Perlman S**, **Netland J**. 2009. Coronaviruses post-SARS: update on replication and pathogenesis. Nat Rev Microbiol **7**:439–450.
6. **Agostini ML**, **Andres EL**, **Sims AC**, **Graham RL**, **Sheahan TP**, **Lu X**, **Smith EC**, **Case JB**, **Feng JY**, **Jordan R**, **Ray AS**, **Cihlar T**, **Siegel D**, **Mackman RL**, **Clarke MO**, **Baric RS**, **Denison MR**. 2018. Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. MBio **9**:e00221–18–15.
7. **Graepel KW**, **Lu X**, **Case JB**, **Sexton NR**, **Smith EC**, **Denison MR**. 2017. Proofreading-Deficient Coronaviruses Adapt for Increased Fitness over Long-Term Passage without Reversion of Exoribonuclease-Inactivating Mutations. MBio **8**:e01503–17.
8. **Smith EC**, **Blanc H**, **Surdel MC**, **Vignuzzi M**, **Denison MR**. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for

proofreading and potential therapeutics. PLoS Pathog **9**:e1003565.

9. **Sexton NR**, **Smith EC**, **Blanc H**, **Vignuzzi M**, **Peersen OB**, **Denison MR**. 2016. Homology-Based Identification of a Mutation in the Coronavirus RNA-Dependent RNA Polymerase That Confers Resistance to Multiple Mutagens. J Virol **90**:7415–7428.
10. **Case JB**, **Li Y**, **Elliott R**, **Lu X**, **Graepel KW**, **Sexton NR**, **Smith EC**, **Weiss SR**, **Denison MR**. 2017. Murine Hepatitis Virus nsp14 Exoribonuclease Activity Is Required for Resistance to Innate Immunity. J Virol **92**:e01531–17–38.
11. **Smith EC**, **Case JB**, **Blanc H**, **Isakov O**, **Shomron N**, **Vignuzzi M**, **Denison MR**. 2015. Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. J Virol **89**:6418–6426.
12. **Graham RL**, **Becker MM**, **Eckerle LD**, **Bolles M**, **Denison MR**, **Baric RS**. 2012. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. Nat Med **18**:1820–1826.
13. **Eckerle LD**, **Lu X**, **Sperry SM**, **Choi L**, **Denison MR**. 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J Virol **81**:12135–12144.
14. **Eckerle LD**, **Becker MM**, **Halpin RA**, **Li K**, **Venter E**, **Lu X**, **Scherbakova S**, **Graham RL**, **Baric RS**, **Stockwell TB**, **Spiro DJ**, **Denison MR**. 2010. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog **6**:e1000896.
15. **Menachery VD**, **Gralinski LE**, **Mitchell HD**, **Dinnon KH III**, **Leist SR**, **Yount BL Jr.**, **McAnarney ET**, **Graham RL**, **Waters KM**, **Baric RS**. 2018. Combination attenuation offers strategy for live-attenuated coronavirus vaccines. J Virol JVI.00710–18–35.
16. **Ma Y**, **Wu L**, **Shaw N**, **Gao Y**, **Wang J**, **Sun Y**, **Lou Z**, **Yan L**, **Zhang R**, **Rao Z**. 2015. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc Natl Acad Sci USA **112**:9436–9441.
17. **Snijder EJ**, **Bredenbeek PJ**, **Dobbe JC**, **Thiel V**, **Ziebuhr J**, **Poon LLM**, **Guan Y**, **Rozanov M**, **Spaan WJM**, **Gorbalenya AE**. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol **331**:991–1004.
18. **Bouvet M**, **Imbert I**, **Subissi L**, **Gluais L**, **Canard B**, **Decroly E**. 2012. RNA 3'-end

mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc Natl Acad Sci USA **109**:9372–9377.

19. **Minskaia E**, **Hertzig T**, **Gorbalenya AE**, **Campanacci V**, **Cambillau C**, **Canard B**, **Ziebuhr J**. 2006. Discovery of an RNA virus 3“->5” exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc Natl Acad Sci USA **103**:5108–5113.
20. **Becares M**, **Pascual-Iglesias A**, **Nogales A**, **Sola I**, **Enjuanes L**, **Zuñiga S**. 2016. Mutagenesis of Coronavirus nsp14 Reveals Its Potential Role in Modulation of the Innate Immune Response. J Virol **90**:5399–5414.
21. **Steitz TA**, **Steitz JA**. 1993. A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci USA **90**:6498–6502.
22. **Chen P**, **Jiang M**, **Hu T**, **Liu Q**, **Chen XS**, **Guo D**. 2007. Biochemical characterization of exoribonuclease encoded by SARS coronavirus. J Biochem Mol Biol **40**:649–655.
23. **Derbyshire V**, **Grindley ND**, **Joyce CM**. 1991. The 3“-5” exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. EMBO J **10**:17–24.
24. **Case JB**, **Ashbrook AW**, **Dermody TS**, **Denison MR**. 2016. Mutagenesis of S-Adenosyl-l-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. J Virol **90**:7248–7256.
25. **Das SR**, **Hensley SE**, **Ince WL**, **Brooke CB**, **Subba A**, **Delboy MG**, **Russ G**, **Gibbs JS**, **Bennink JR**, **Yewdell JW**. 2013. Defining Influenza A Virus Hemagglutinin Antigenic Drift by Sequential Monoclonal Antibody Selection. Cell Host and Microbe **13**:314–323.
26. **Nakajima K**, **Nobusawa E**, **Nagy A**, **Nakajima S**. 2005. Accumulation of Amino Acid Substitutions Promotes Irreversible Structural Changes in the Hemagglutinin of Human Influenza AH3 Virus during Evolution. J Virol **79**:6472–6477.
27. **Koel BF**, **Burke DF**, **van der Vliet S**, **Bestebroer TM**, **Rimmelzwaan GF**, **Osterhaus ADME**, **Smith DJ**, **Fouchier RAM**. 2018. Epistatic interactions can moderate the antigenic effect of substitutions in hemagglutinin of influenza H3N2 virus 1–16.
28. **Dolan PT**, **Whitfield ZJ**, **Andino R**. 2018. Mechanisms and Concepts in RNA Virus Population Dynamics and Evolution. Annu Rev Virol **5**:annurev–virology–101416–041718–24.

29. **Smith EC**, **Sexton NR**, **Denison MR**. 2014. Thinking Outside the Triangle: Replication Fidelity of the Largest RNA Viruses. Annu Rev Virol **1**:111–132.
30. **Arnold JJ**, **Vignuzzi M**, **Stone JK**, **Andino R**, **Cameron CE**. 2005. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J Biol Chem **280**:25706–25716.
31. **Chen W**, **Baric RS**. 1996. Molecular anatomy of mouse hepatitis virus persistence: coevolution of increased host cell resistance and virus virulence. J Virol **70**:3947–3960.
32. **Yount B**, **Denison MR**, **Weiss SR**, **Baric RS**. 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J Virol **76**:11065–11078.
33. **Ho SN**, **Hunt HD**, **Horton RM**, **Pullen JK**, **Pease LR**. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77**:51–59.

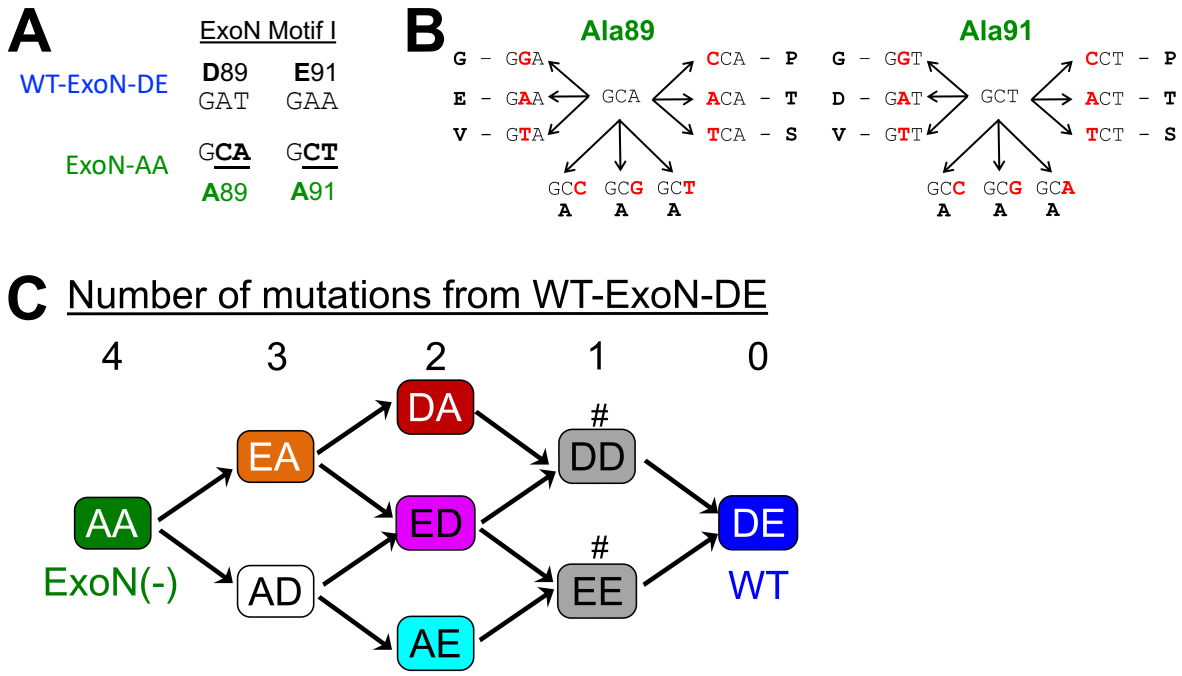


Figure 1. Sequence landscape around ExoN-AA motif I. (*A*) ExoN motif I nucleotide sequences. (*B*) Landscape of single-nucleotide substitutions within ExoN-AA motif I. (*C*) Predicted pathways to reversion of ExoN-AA. Variants marked with # reverted to WT during three independent recovery attempts.

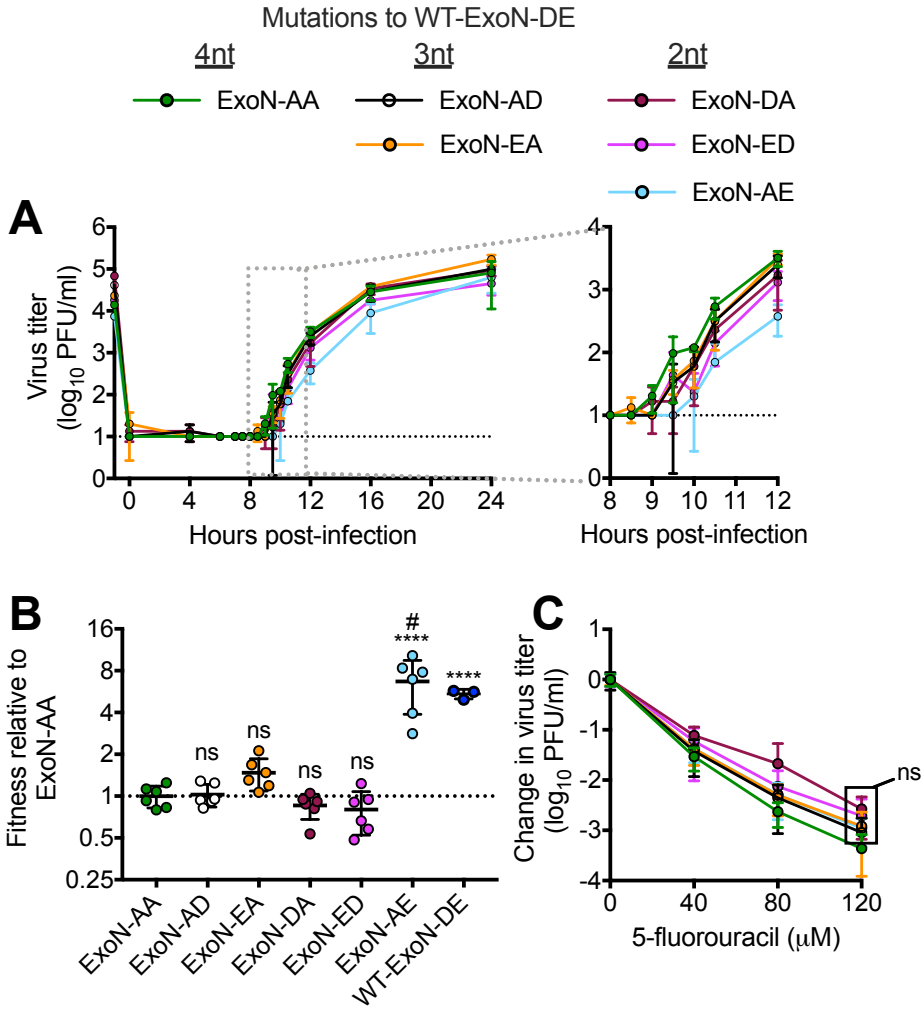


Figure 2. Intermediate revertants of ExoN-AA motif I do not have selective advantages. (*A*) Replication kinetics at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) Competitive fitness of each variant relative to ExoN-AA. Viruses were competed with a tagged ExoN-AA-reference strain, and relative fitness was normalized to the mean of ExoN-AA. (*C*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell. Statistical significance of each variant relative to ExoN-AA was determined by one-way ANOVA with multiple comparisons (Panel *D*) two-way ANOVA with Dunnett's multiple comparisons (panel *C*). ****p<0.0001; ns = not significant. Data in (*B*) and (*C*) represent mean ± SD of n = 6. Boxed values have the same significance. [#]All lineages of ExoN-AE reverted to WT-ExoN-DE during the experiment.

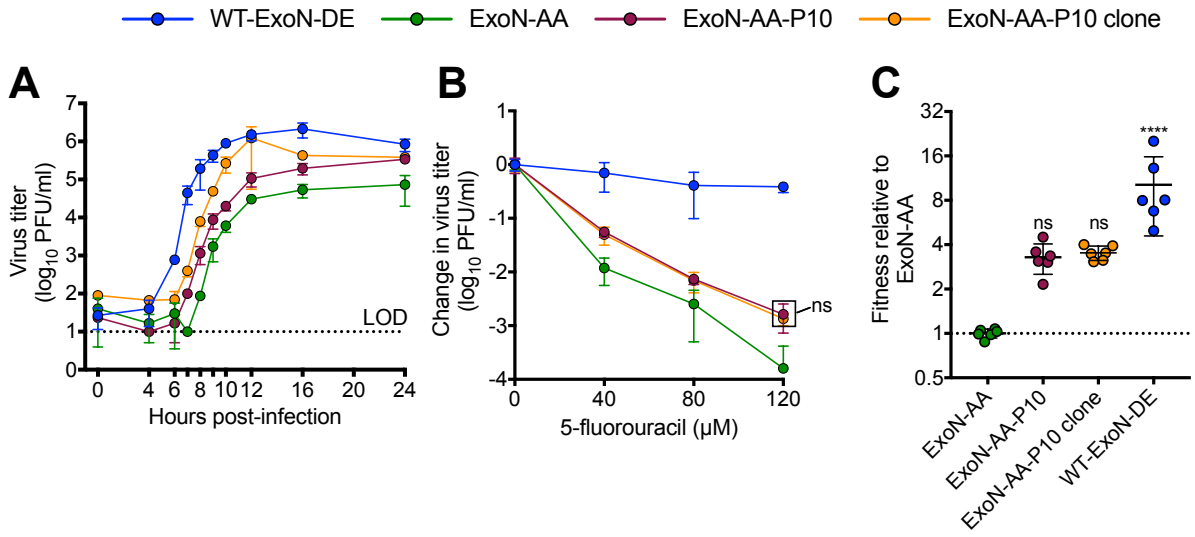


Figure 3. ExoN-AA adapts for increased fitness within 10 passages. (*A*) Replication kinetics of indicated viruses at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell. (*C*) Competitive fitness of individual recombinants relative to ExoN-AA. Viruses were competed with a tagged ExoN-AA-reference strain, and relative fitness was normalized to the mean of ExoN-AA. Statistical significance of each virus relative to ExoN-AA was determined by two-way ANOVA with Dunnett's multiple comparisons (panel *B*) or by one-way ANOVA with multiple comparisons (Panel *C*). ****p < 0.0001, ns = not significant. LOD = limit of detection. Data in (*B*) and (*C*) represent mean ± SD of n = 6. Boxed values have the same significance.

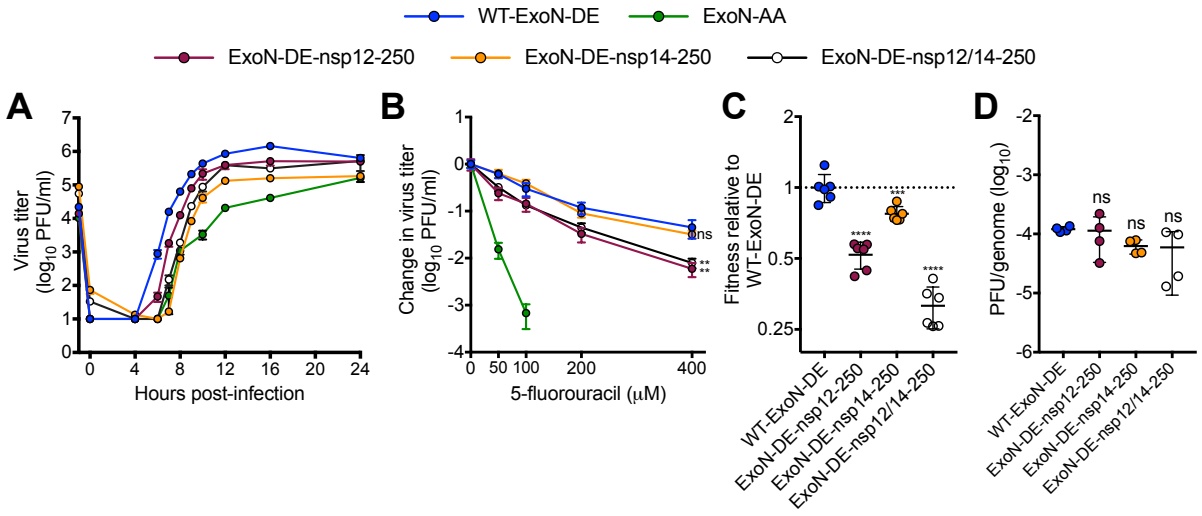


Figure 4. Mutations that increase ExoN-AA fitness are detrimental in the presence of WT-ExoN-DE. (*A*) Replication kinetics of indicated viruses at MOI = 0.01 PFU/cell plotted as mean ± SD of n = 3. (*B*) 5-fluorouracil sensitivity at MOI = 0.01 PFU/cell, mean ± SD of n = 6. (*C*) Competitive fitness of individual recombinants relative to WT-ExoN-DE. Viruses were competed with a tagged WT-ExoN-DE reference strain, and relative fitness was normalized to the mean of WT-ExoN-DE, mean ± SD of n = 6. (*D*) Specific infectivity (genomes per PFU) from isolated infections, mean ± SD of n = 4.. Statistical significance of each virus relative to WT-ExoN-DE was determined with two-way ANOVA with Dunnett's multiple comparisons test (panel *B*) or by ordinary one-way ANOVA with Dunnett's multiple comparisons test (panels *C* and *D*). **p < 0.01; ***p < 0.001; ****p < 0.0001; ns = not significant.

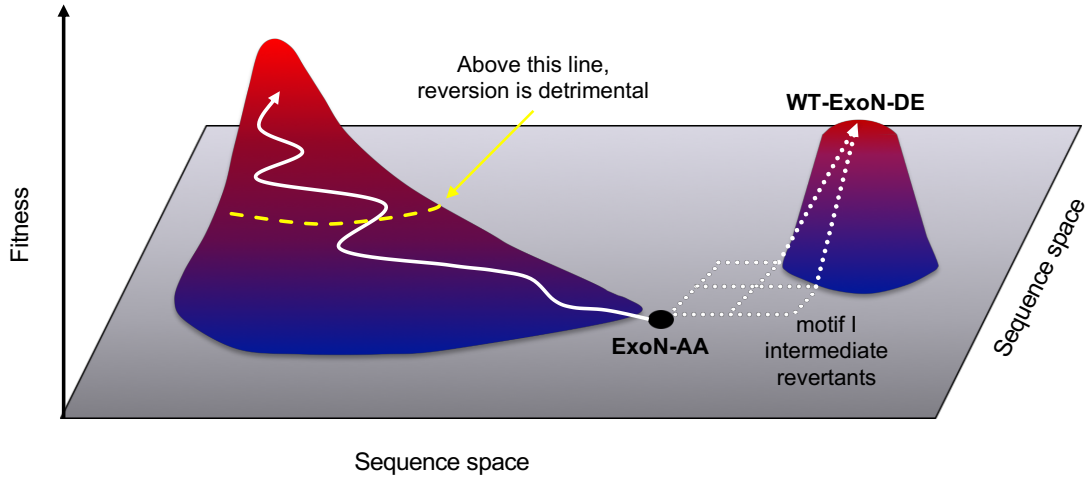


Figure 5. Model for the *in vitro* evolution of MHV-ExoN-AA. MHV-ExoN-AA(black dot) is a low-fitness variant. Reversion to WT-ExoN-DE would dramatically increase fitness but can only be achieved by traversing a flat landscape and climbing a steep fitness cliff (dotted white arrows). However, secondary mutations that incrementally increase fitness are more accessible (solid white arrow). Eventually, the genetic background changes enough that reversion becomes detrimental (dotted yellow line).

Table 1. Recovery and passage of intermediate revertants.

n.d.: not done.
[a]Bolded nucleotides must mutate to reach WT-ExoN-DE genotype.
[b]Recovered viruses were subjected to 10 passages at the indicated MOI. Samples were screened for wild-type revertants by plaque assay, and revertant lineages were sequence-confirmed.

Virus	# of mutations to WT-ExoN-DE	Motif I sequence[a]	# of reverted lineages by passage 10[b]	
			MOI = 0.01	MOI = 0.5
ExoN-AA	4	G**CA**...G**CT**	0/3	0/3
ExoN-AD	3	G**CA**...GA**T**	0/3	0/3
ExoN-EA	3	GA**A**...G**CT**	0/3	0/3
ExoN-DA	2	GAT...G**CT**	0/3	0/3
ExoN-AE	2	G**CA**...GAA	1/3 (by P8)	3/3 (by P2)
ExoN-ED	2	GA**A**...GA**T**	0/3	0/3
ExoN-EE	1	GA**A**...GAA	n.d.	n.d.
ExoN-DD	1	GAT...GA**T**	n.d.	n.d.
WT-ExoN-DE	0	GAT...GAA	n.d.	n.d.

Table 2. Mutations in ExoN(-) P10. Data derived from dideoxy sequencing.

[a]Mutation present at approximately 50% of population.
[b]MHV HE is not transcribed in tissue culture.
[c]Amino acid numbers designate positions within cleaved nsps, not the polyprotein.

Mutation number	Nucleotide change	Protein	Amino acid change[c]
1	G2520A[a]	nsp2	D524N
2	A3080G[a]	nsp3	silent
3	T16017A	nsp12	M814K
4	A17836G[a]	nsp13	I492M
5	G22673A[a]	HE[b]	noncoding
6	A29298C[a]	M	silent

APPENDIX F: Copyright permissions

F.1. de Wit et al., 2016, reproduced in Figure 2.

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jan 15, 2019

This Agreement between Kevin Graepel ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4498930857266
License date	Dec 30, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Microbiology
Licensed Content Title	SARS and MERS: recent insights into emerging coronaviruses
Licensed Content Author	Emmie de Wit, Neeltje van Doremalen, Darryl Falzarano, Vincent J. Munster
Licensed Content Date	Jun 27, 2016
Licensed Content Volume	14
Licensed Content Issue	8
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Analysis of adaptive pathways in a proofreading-deficient coronavirus
Institution name	Vanderbilt University
Expected presentation date	Mar 2019
Portions	Figure 1
Requestor Location	Kevin Graepel 3006 Brightwood Ave NASHVILLE, TN 37212 United States Attn: Kevin Graepel
Billing Type	Invoice
Billing Address	Kevin Graepel 3006 Brightwood Ave

NASHVILLE, TN 37212
United States
Attn: Kevin Graepel

Total 0.00 USD

Terms and Conditions

Springer Nature Terms and Conditions for RightsLink Permissions

Springer Nature Customer Service Centre GmbH (the Licensor) hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

   If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.1

F.2. Dolan et al., 2018, reproduced in Figure 7.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 15, 2019

This Agreement between Kevin Graepel ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4499581263416
License date	Dec 31, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell Host & Microbe
Licensed Content Title	Mapping the Evolutionary Potential of RNA Viruses
Licensed Content Author	Patrick T. Dolan,Zachary J. Whitfield,Raul Andino
Licensed Content Date	Apr 11, 2018
Licensed Content Volume	23
Licensed Content Issue	4
Licensed Content Pages	12
Start Page	435
End Page	446
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1B
Title of your thesis/dissertation	Analysis of adaptive pathways in a proofreading-deficient coronavirus
Publisher of new work	Vanderbilt University
Expected completion date	Mar 2019
Estimated size (number of pages)	1
Requestor Location	Kevin Graepel 3006 Brightwood Ave NASHVILLE, TN 37212 United States Attn: Kevin Graepel
Publisher Tax ID	98-0397604
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.
Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.
17. **For journal authors:** the following clauses are applicable in addition to the above:
Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).
Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).
If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available

version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.
Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.
Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.
Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above:
Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

REFERENCES

Acevedo, A., Brodsky, L., and Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature *505*, 686–690.

Adedeji, A.O., Marchand, B., Velthuis, te, A.J.W., Snijder, E.J., Weiss, S., Eoff, R.L., Singh, K., and Sarafianos, S.G. (2012). Mechanism of Nucleic Acid Unwinding by SARS-CoV Helicase. PLoS ONE *7*, e36521.

Agostini, M.L., Andres, E.L., Sims, A.C., Graham, R.L., Sheahan, T.P., Lu, X., Smith, E.C., Case, J.B., Feng, J.Y., Jordan, R., et al. (2018). Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. MBio *9*, e00221–18–15.

Agrawal, A.S., Garron, T., Tao, X., Peng, B.-H., Wakamiya, M., Chan, T.-S., Couch, R.B., and Tseng, C.-T.K. (2015). Generation of a transgenic mouse model of Middle East respiratory syndrome coronavirus infection and disease. J Virol *89*, 3659–3670.

Agudo, R., la Higuera, de, I., Arias, A., Grande-Pérez, A., and Domingo, E. (2016). Involvement of a joker mutation in a polymerase-independent lethal mutagenesis escape mechanism. Virology *494*, 257–266.

Aguilera, E.R., Erickson, A.K., Jesudhasan, P.R., Robinson, C.M., and Pfeiffer, J.K. (2017). Plaques Formed by Mutagenized Viral Populations Have Elevated Coinfection Frequencies. MBio *8*, e02020–16–12.

Al-Khannaq, M.N., Ng, K.T., Oong, X.Y., Pang, Y.K., Takebe, Y., Chook, J.B., Hanafi, N.S., Kamarulzaman, A., and Tee, K.K. (2016a). Molecular epidemiology and evolutionary histories of human coronavirus OC43 and HKU1 among patients with upper respiratory tract infections in Kuala Lumpur, Malaysia. Virol. J. *13*, 1–12.

Al-Khannaq, M.N., Ng, K.T., Oong, X.Y., Pang, Y.K., Takebe, Y., Chook, J.B., Hanafi, N.S., Kamarulzaman, A., and Tee, K.K. (2016b). Diversity and Evolutionary Histories of Human Coronaviruses NL63 and 229E Associated with Acute Upper Respiratory Tract Symptoms in Kuala Lumpur, Malaysia. Am. J. Trop. Med. Hyg. *94*, 1058–1064.

Almazan, F., DeDiego, M.L., Sola, I., Zuniga, S., Nieto-Torres, J.L., Marquez-Jurado, S., Andres, G., and Enjuanes, L. (2013). Engineering a Replication-Competent, Propagation-Defective Middle East Respiratory Syndrome Coronavirus as a Vaccine Candidate. MBio *4*, 20282–11.

Alraddadi, B.M., Watson, J.T., Almarashi, A., Abedi, G.R., Turkistani, A., Sadran, M., Housa, A., Almazroa, M.A., Alraihan, N., Banjar, A., et al. (2016). Risk Factors for Primary Middle East Respiratory Syndrome Coronavirus Illness in Humans, Saudi Arabia, 2014. Emerging Infect. Dis. *22*, 49–55.

Annamalai, T., Saif, L.J., Lu, Z., and Jung, K. (2015). Age-dependent variation in innate immune responses to porcine epidemic diarrhea virus infection in suckling versus weaned pigs.

Veterinary Immunology and Immunopathology *168*, 193–202.

Annan, A., Baldwin, H.J., Corman, V.M., Klose, S.M., Owusu, M., Nkrumah, E.E., Badu, E.K., Anti, P., Agbenyega, O., Meyer, B., et al. (2013). Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerging Infect. Dis. *19*, 456–459.

Arabi, Y.M., Balkhy, H.H., Hayden, F.G., Bouchama, A., Luke, T., Baillie, J.K., Al-Omari, A., Hajeer, A.H., Senga, M., Denison, M.R., et al. (2017). Middle East Respiratory Syndrome. N. Engl. J. Med. *376*, 584–594.

Arabi, Y.M., Hajeer, A.H., Luke, T., Raviprakash, K., Balkhy, H., Johani, S., Al-Dawood, A., Al-Qahtani, S., Al-Omari, A., Al-Hameed, F., et al. (2016). Feasibility of Using Convalescent Plasma Immunotherapy for MERS-CoV Infection, Saudi Arabia. Emerging Infect. Dis. *22*, 1554–1561.

Arden, K.E., Nissen, M.D., Sloots, T.P., and Mackay, I.M. (2005). New human coronavirus, HCoV-NL63, associated with severe lower respiratory tract disease in Australia. J. Med. Virol. *75*, 455–462.

Arenas, M., Araujo, N.M., Branco, C., Castelhano, N., Castro-Nallar, E., and Pérez-Losada, M. (2018). Mutation and recombination in pathogen evolution: Relevance, methods and controversies. Infect. Genet. Evol. *63*, 295–306.

Arias, A., Arnold, J.J., Sierra, M., Smidansky, E.D., Domingo, E., and Cameron, C.E. (2008). Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. J Virol *82*, 12346–12355.

Arnold, J.J., and Cameron, C.E. (2004). Poliovirus RNA-Dependent RNA Polymerase (3D pol): Pre-Steady-State Kinetic Analysis of Ribonucleotide Incorporation in the Presence of Mg 2+ †. Biochemistry *43*, 5126–5137.

Arnold, J.J., Vignuzzi, M., Stone, J.K., Andino, R., and Cameron, C.E. (2005). Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J. Biol. Chem. *280*, 25706–25716.

Arribas, M., Cabanillas, L., and Lázaro, E. (2011). Identification of mutations conferring 5-azacytidine resistance in bacteriophage Qβ. Virology *417*, 343–352.

Auyeung, T.W., Lee, J.S.W., Lai, W.K., Choi, C.H., Lee, H.K., Lee, J.S., Li, P.C., Lok, K.H., Ng, Y.Y., Wong, W.M., et al. (2005). The use of corticosteroid as treatment in SARS was associated with adverse outcomes: a retrospective cohort study. J. Infect. *51*, 98–102.

Bailey, O.T., Pappenheimer, M.D., Cheever, F.S., and Daniels, J.B. (1949). A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin: II. Pathology. J. Exp. Med. *90*, 195–212.

Baker, M.L., Schountz, T., and Wang, L.F. (2013). Antiviral immune responses of bats: a

review. Zoonoses Public Health *60*, 104–116.

Baric, R.S., Yount, B., Hensley, L., Peel, S.A., and Chen, W. (1997). Episodic evolution mediates interspecies transfer of a murine coronavirus. J Virol *71*, 1946–1955.

Barnard, D.L., Day, C.W., Bailey, K., Heiner, M., Montgomery, R., Lauridsen, L., Winslow, S., Hoopes, J., Li, J.K.-K., Lee, J., et al. (2006). Enhancement of the infectivity of SARS-CoV in BALB/c mice by IMP dehydrogenase inhibitors, including ribavirin. Antiviral Res. *71*, 53–63.

Bastien, N., Robinson, J.L., Tse, A., Lee, B.E., Hart, L., and Li, Y. (2005). Human Coronavirus NL-63 Infections in Children: a 1-Year Study. J. Clin. Microbiol. *43*, 4567–4573.

Becares, M., Pascual-Iglesias, A., Nogales, A., Sola, I., Enjuanes, L., and Zuñiga, S. (2016). Mutagenesis of Coronavirus nsp14 Reveals Its Potential Role in Modulation of the Innate Immune Response. J Virol *90*, 5399–5414.

Becker, M.M., Graham, R.L., Donaldson, E.F., Rockx, B., Sims, A.C., Sheahan, T., Pickles, R.J., Corti, D., Johnston, R.E., Baric, R.S., et al. (2008). Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. Proc. Natl. Acad. Sci. U.S.a. *105*, 19944–19949.

Beigel, J.H., Voell, J., Kumar, P., Raviprakash, K., Wu, H., Jiao, J.-A., Sullivan, E., Luke, T., and Davey, R.T., Jr (2018). Safety and tolerability of a novel, polyclonal human anti-MERS coronavirus antibody produced from transchromosomic cattle: a phase 1 randomised, double-blind, single-dose-escalation study. Lancet Infect Dis *18*, 410–418.

Bentley, K., and Evans, D.J. (2018). Mechanisms and consequences of positive-strand RNA virus recombination. Journal of General Virology *3*, 326–12.

Bhardwaj, K., Guarino, L., and Kao, C.C. (2004). The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J Virol *78*, 12218–12224.

Booth, C.M., Matukas, L.M., Tomlinson, G.A., Rachlis, A.R., Rose, D.B., Dwosh, H.A., Walmsley, S.L., Mazzulli, T., Avendano, M., Derkach, P., et al. (2003). Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. Jama *289*, 2801–2809.

Booth, T.F., Kournikakis, B., Bastien, N., Ho, J., Kobasa, D., Stadnyk, L., Li, Y., Spence, M., Paton, S., Henry, B., et al. (2005). Detection of airborne severe acute respiratory syndrome (SARS) coronavirus and environmental contamination in SARS outbreak units. J. Infect. Dis. *191*, 1472–1477.

Bordería, A.V., Isakov, O., Moratorio, G., Henningsson, R., Agüera-González, S., Organtini, L., Gnädig, N.F., Blanc, H., Alcover, A., Hafenstein, S., et al. (2015). Group Selection and Contribution of Minority Variants during Virus Adaptation Determines Virus Fitness and Phenotype. PLoS Pathog *11*, e1004838.

Bordería, A.V., Lorenzo-Redondo, R., Pernas, M., Casado, C., Alvaro, T., Domingo, E., and López-Galíndez, C. (2010). Initial Fitness Recovery of HIV-1 Is Associated with Quasispecies Heterogeneity and Can Occur without Modifications in the Consensus Sequence. PLoS ONE *5*, e10319–8.

Bouvet, M., Debarnot, C., Imbert, I., Selisko, B., Snijder, E.J., Canard, B., and Decroly, E. (2010). In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog *6*, e1000863.

Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012). RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc. Natl. Acad. Sci. U.S.a. *109*, 9372–9377.

Bouvet, M., Lugari, A., Posthuma, C.C., Zevenhoven, J.C., Bernard, S., Betzi, S., Imbert, I., Canard, B., Guillemot, J.-C., Lécine, P., et al. (2014). Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. J. Biol. Chem. *289*, 25783–25796.

Bradburne, A.F., and Somerset, B.A. (1972). Coronavirus antibody titres in sera of healthy adults and experimentally infected volunteers. Epidemiol. Infect. *70*, 235–244.

Bradburne, A.F., Bynoe, M.L., and Tyrrell, D.A. (1967). Effects of a "new" human respiratory virus in volunteers. British Medical Journal *3*, 767–769.

Bradwell, K., Combe, M., Domingo-Calap, P., and Sanjuán, R. (2013). Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Qβ. Genetics *195*, 243–251.

Brierley, I., Boursnell, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D., and Inglis, S.C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. Embo J. *6*, 3779–3785.

Brierley, I., Digard, P., and Inglis, S.C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. Cell *57*, 537–547.

Brister, J.R., Ako-Adjei, D., Bao, Y., and Blinkova, O. (2014). NCBI Viral Genomes Resource. Nucleic Acids Res. *43*, gku1207–D577.

Brockway, S.M., Lu, X.T., Peters, T.R., Dermody, T.S., and Denison, M.R. (2004). Intracellular Localization and Protein Interactions of the Gene 1 Protein p28 during Mouse Hepatitis Virus Replication. J Virol *78*, 11551–11562.

Brockway, S.M., Clay, C.T., Lu, X.T., and Denison, M.R. (2003). Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J Virol *77*, 10515–10527.

Brunn, von, A., Teepe, C., Simpson, J.C., Pepperkok, R., Friedel, C.C., Zimmer, R., Roberts, R., Baric, R., and Haas, J. (2007). Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. PLoS ONE *2*, e459.

Bukhari, K., Mulley, G., Gulyaeva, A.A., Zhao, L., Shu, G., Jiang, J., and Neuman, B.W. (2018). Description and initial characterization of metatranscriptomic nidovirus-like genomes from the proposed new family Abyssoviridae, and from a sister group to the Coronavirinae, the proposed genus Alphaletovirus. Virology *524*, 160–171.

Bull, J.J. (2015). Evolutionary reversion of live viral vaccines: Can genetic engineering subdue it? Virus Evol *1*, vev005–vev010.

Bull, J.J., Sanjuán, R., and Wilke, C.O. (2007). Theory of lethal mutagenesis for viruses. J Virol *81*, 2930–2939.

Calisher, C.H., Childs, J.E., Field, H.E., Holmes, K.V., and Schountz, T. (2006). Bats: Important Reservoir Hosts of Emerging Viruses. Clinical Microbiology Reviews *19*, 531–545.

Cameron, C.E., Moustafa, I.M., and Arnold, J.J. (2016). Fidelity of Nucleotide Incorporation by the RNA-Dependent RNA Polymerase from Poliovirus (Elsevier Inc.).

Campagnola, G., McDonald, S., Beaucourt, S., Vignuzzi, M., and Peersen, O.B. (2015). Structure-Function Relationships Underlying the Replication Fidelity of Viral RNA-Dependent RNA Polymerases. J Virol *89*, 275–286.

Carrasco, P., Daròs, J.A., Agudelo-Romero, P., and Elena, S.F. (2007). A real-time RT-PCR assay for quantifying the fitness of tobacco etch virus in competition experiments. Journal of Virological Methods *139*, 181–188.

Carroll, S.S., Tomassini, J.E., Bosserman, M., Getty, K., Stahlhut, M.W., Eldrup, A.B., Bhat, B., Hall, D., Simcoe, A.L., LaFemina, R., et al. (2003). Inhibition of hepatitis C virus RNA replication by 2'-modified nucleoside analogs. J. Biol. Chem. *278*, 11979–11984.

Case, J.B. (2018). Roles of the Coronavirus 3“-to-5” Exoribonuclease and N7-Methyltransferase in Counteracting Innate Immunity.

Case, J.B., Ashbrook, A.W., Dermody, T.S., and Denison, M.R. (2016). Mutagenesis of S-Adenosyl-l-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. J Virol *90*, 7248–7256.

Case, J.B., Li, Y., Elliott, R., Lu, X., Graepel, K.W., Sexton, N.R., Smith, E.C., Weiss, S.R., and Denison, M.R. (2017). Murine Hepatitis Virus nsp14 Exoribonuclease Activity Is Required for Resistance to Innate Immunity. J Virol *92*, e01531–17–38.

Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. Veterinary Research *38*, 281–297.

Chan, C.-M., Chu, H., Wang, Y., Wong, B.H.-Y., Zhao, X., Zhou, J., Yang, D., Leung, S.P., Chan, J.F.-W., Yeung, M.-L., et al. (2016). Carcinoembryonic Antigen-Related Cell Adhesion Molecule 5 Is an Important Surface Attachment Factor That Facilitates Entry of Middle East Respiratory Syndrome Coronavirus. J Virol *90*, 9114–9127.

Chan, K.H., Peiris, J.S.M., Lam, S.Y., Poon, L.L.M., Yuen, K.Y., and Seto, W.H. (2011). The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. Advances in Virology *2011*, 1–7.

Channappanavar, R., and Perlman, S. (2017). Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. Semin Immunopathol *2*, 1–11.

Channappanavar, R., Fehr, A.R., Vijay, R., Mack, M., Zhao, J., Meyerholz, D.K., and Perlman, S. (2016). Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal Pneumonia in SARS-CoV-Infected Mice. Cell Host and Microbe *19*, 181–193.

Chaudhuri, M., Chaudhuri, M., Song, L., and Parris, D.S. (2003). The herpes simplex virus type 1 DNA polymerase processivity factor increases fidelity without altering pre-steady-state rate constants for polymerization or excision. J. Biol. Chem. *278*, 8996–9004.

Chen, N., Li, S., Zhou, R., Zhu, M., He, S., Ye, M., Huang, Y., Li, S., Zhu, C., Xia, P., et al. (2017). Two novel porcine epidemic diarrhea virus (PEDV) recombinants from a natural recombinant and distinct subtypes of PEDV variants. Virus Research *242*, 90–95.

Chen, P., Jiang, M., Hu, T., Liu, Q., Chen, X.S., and Guo, D. (2007). Biochemical characterization of exoribonuclease encoded by SARS coronavirus. J. Biochem. Mol. Biol. *40*, 649–655.

Chen, W., and Baric, R.S. (1996). Molecular anatomy of mouse hepatitis virus persistence: coevolution of increased host cell resistance and virus virulence. J Virol *70*, 3947–3960.

Chen, Y.-H., Du, W., Hagemeijer, M.C., Takvorian, P.M., Pau, C., Cali, A., Brantner, C.A., Stempinski, E.S., Connelly, P.S., Ma, H.-C., et al. (2015). Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses. Cell *160*, 619–630.

Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. U.S.a. *106*, 3484–3489.

Cheng, P.K., Wong, D.A., Tong, L.K., Ip, S.-M., Lo, A.C., Lau, C.-S., Yeung, E.Y., and Lim, W.W. (2004). Viral shedding patterns of coronavirus in patients with probable severe acute respiratory syndrome. The Lancet *363*, 1699–1700.

Cheung, P.P.H., Watson, S.J., Choy, K.-T., Fun Sia, S., Wong, D.D.Y., Poon, L.L.M., Kellam, P., Guan, Y., Malik Peiris, J.S., and Yen, H.-L. (2014). Generation and characterization of influenza A viruses with altered polymerase fidelity. Nat Commun *5*, 4794.

Chinese SARS Molecular Epidemiology Consortium (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science *303*, 1666–1669.

Chiou, H.-E., Liu, C.-L., Buttrey, M.J., Kuo, H.-P., Liu, H.-W., Kuo, H.-T., and Lu, Y.-T. (2005). Adverse effects of ribavirin and outcome in severe acute respiratory syndrome: experience in two medical centers. Chest *128*, 263–272.

Choi, J., Kim, M.-G., Oh, Y.-K., Kim, Y.B., Choi, J., Kim, M.-G., Oh, Y.-K., and Kim, Y.B. (2017). Progress of Middle East respiratory syndrome coronavirus vaccines: a patent review. Expert Opinion on Therapeutic Patents *00*, 1–11.

Chu, C.-M., Poon, L.L.M., Cheng, V.C.C., Chan, K.-S., Hung, I.F.N., Wong, M.M.L., Chan, K.-H., Leung, W.-S., Tang, B.S.F., Chan, V.L., et al. (2004). Initial viral load and the outcomes of SARS. Cmaj *171*, 1349–1352.

Coffey, L.L., Beeharry, Y., Bordería, A.V., Blanc, H., and Vignuzzi, M. (2011). Arbovirus high fidelity variant loses fitness in mosquitoes and mice. Proc. Natl. Acad. Sci. U.S.a. *108*, 16038–16043.

Collins, N.D., Beck, A.S., Widen, S.G., Wood, T.G., Higgs, S., and Barrett, A.D.T. (2018). Structural and Nonstructural Genes Contribute to the Genetic Diversity of RNA Viruses. MBio *9*, 1225–13.

Combe, M., and Sanjuán, R. (2014). Variation in RNA virus mutation rates across host cells. PLoS Pathog *10*, e1003855.

Corman, V.M., Albarrak, A.M., Omrani, A.S., Albarrak, M.M., Farah, M.E., Almasri, M., Muth, D., Sieberg, A., Meyer, B., Assiri, A.M., et al. (2016a). Viral Shedding and Antibody Response in 37 Patients With Middle East Respiratory Syndrome Coronavirus Infection. Clinid *62*, 477–483.

Corman, V.M., Eckerle, I., Memish, Z.A., Liljander, A.M., Dijkman, R., Jonsdottir, H., Juma Ngeiywa, K.J.Z., Kamau, E., Younan, M., Masri, Al, M., et al. (2016b). Link of a ubiquitous human coronavirus to dromedary camels. Proc. Natl. Acad. Sci. U.S.a. 201604472–13.

Corman, V.M., Muth, D., Niemeyer, D., and Drosten, C. (2018). Hosts and Sources of Endemic Human Coronaviruses. *100*, 1–26.

Corman, V.M., Baldwin, H.J., Tateno, A.F., Zerbinati, R.M., Annan, A., Owusu, M., Nkrumah, E.E., Maganga, G.D., Oppong, S., Adu-Sarkodie, Y., et al. (2015). Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. J Virol *89*, 11858–11870.

Corman, V.M., Ithete, N.L., Richards, L.R., Schoeman, M.C., Preiser, W., Drosten, C., and Drexler, J.F. (2014a). Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. J Virol *88*, 11297–11303.

Corman, V.M., Kallies, R., Philipps, H., Göpner, G., Müller, M.A., Eckerle, I., Brünink, S., Drosten, C., and Drexler, J.F. (2014b). Characterization of a novel betacoronavirus related to middle East respiratory syndrome coronavirus in European hedgehogs. J Virol *88*, 717–724.

Crotty, S., Cameron, C.E., and Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. Proc. Natl. Acad. Sci. U.S.a. *98*, 6895–6900.

Crotty, S., Cameron, C., and Andino, R. (2002). Ribavirin's antiviral mechanism of action: lethal

mutagenesis? J. Mol. Med. *80*, 86–95.

Cuevas, J.M., Geller, R., Garijo, R., López-Aldeguer, J., and Sanjuán, R. (2015). Extremely High Mutation Rate of HIV-1 In Vivo. PLoS Biol. *13*, e1002251.

Cuevas, J.M., González-Candelas, F., Moya, A., and Sanjuán, R. (2009). Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo. J Virol *83*, 5760–5764.

Cui, J., Li, F., and Shi, Z.-L. (2018). Origin and evolution of pathogenic coronaviruses. Nat. Rev. Microbiol. *362*, 1–12.

Dapp, M.J., Heineman, R.H., and Mansky, L.M. (2013). Interrelationship between HIV-1 Fitness and Mutation Rate. J. Mol. Biol. *425*, 41–53.

Das, S.R., Hensley, S.E., Ince, W.L., Brooke, C.B., Subba, A., Delboy, M.G., Russ, G., Gibbs, J.S., Bennink, J.R., and Yewdell, J.W. (2013). Defining Influenza A Virus Hemagglutinin Antigenic Drift by Sequential Monoclonal Antibody Selection. Cell Host and Microbe *13*, 314–323.

de Wit, E., van Doremalen, N., Falzarano, D., and Munster, V.J. (2016). SARS and MERS: recent insights into emerging coronaviruses. Nat. Rev. Microbiol. 1–12.

Decroly, E., Debarnot, C., Ferron, F., Bouvet, M., Coutard, B., Imbert, I., Gluais, L., Papageorgiou, N., Sharff, A., Bricogne, G., et al. (2011). Crystal Structure and Functional Analysis of the SARS-Coronavirus RNA Cap 2′-O-Methyltransferase nsp10/nsp16 Complex. PLoS Pathog *7*, e1002059.

Decroly, E., Imbert, I., Coutard, B., Bouvet, M., Selisko, B., Alvarez, K., Gorbalenya, A.E., Snijder, E.J., and Canard, B. (2008). Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. J Virol *82*, 8071–8084.

Deng, X., Hackbart, M., Mettelman, R.C., O'Brien, A., Mielech, A.M., Yi, G., Kao, C.C., and Baker, S.C. (2017). Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. Proc. Natl. Acad. Sci. U.S.a. *114*, E4251–E4260.

Derbyshire, V., Grindley, N.D., and Joyce, C.M. (1991). The 3"-5" exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. Embo J. *10*, 17–24.

Diamond, M.S., Zachariah, M., and Harris, E. (2002). Mycophenolic Acid Inhibits Dengue Virus Infection by Preventing Replication of Viral RNA. Virology *304*, 211–221.

Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. *36*, e105–e105.

Dolan, P.T., Whitfield, Z.J., and Andino, R. (2018a). Mapping the Evolutionary Potential of RNA Viruses. Cell Host and Microbe *23*, 435–446.

Dolan, P.T., Whitfield, Z.J., and Andino, R. (2018b). Mechanisms and Concepts in RNA Virus Population Dynamics and Evolution. Annu Rev Virol *5*, annurev–virology–101416–041718–24.

Domingo, E., and Holland, J.J. (1997). RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. *51*, 151–178.

Domingo, E., Sheldon, J., and Perales, C. (2012). Viral quasispecies evolution. Microbiol. Mol. Biol. Rev. *76*, 159–216.

Dominguez, S.R., Robinson, C.C., and Holmes, K.V. (2009). Detection of four human coronaviruses in respiratory infections in children: A one-year study in Colorado. J. Med. Virol. *81*, 1597–1604.

Donaldson, E.F., Sims, A.C., Graham, R.L., Denison, M.R., and Baric, R.S. (2007). Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. J Virol *81*, 6356–6368.

Drexler, J.F., Gloza-Rausch, F., Glende, J., Corman, V.M., Muth, D., Goettsche, M., Seebens, A., Niedrig, M., Pfefferle, S., Yordanov, S., et al. (2010). Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. J Virol *84*, 11336–11349.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.-R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A.M., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. *348*, 1967–1976.

Duarte, E., Clarke, D., Moya, A., Domingo, E., and Holland, J. (1992). Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc. Natl. Acad. Sci. U.S.a. *89*, 6015–6019.

Duffy, S., Shackelton, L.A., and Holmes, E.C. (2008). Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. *9*, 267–276.

Dveksler, G.S., Dieffenbach, C.W., Cardellichio, C.B., McCuaig, K., Pensiero, M.N., Jiang, G.S., Beauchemin, N., and Holmes, K.V. (1993). Several members of the mouse carcinoembryonic antigen-related glycoprotein family are functional receptors for the coronavirus mouse hepatitis virus-A59. J Virol *67*, 1–8.

Dyson, H.J. (2016). Making Sense of Intrinsically Disordered Proteins. Biophysj *110*, 1013–1016.

Ebihara, T., Endo, R., Ma, X., Ishiguro, N., and Kikuta, H. (2005). Detection of human coronavirus NL63 in young children with bronchiolitis. J. Med. Virol. *75*, 463–465.

Eckerle, L.D., Becker, M.M., Halpin, R.A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R.L., Baric, R.S., Stockwell, T.B., et al. (2010). Infidelity of SARS-CoV Nsp14-exonuclease

mutant virus replication is revealed by complete genome sequencing. PLoS Pathog *6*, e1000896.

Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., and Denison, M.R. (2007). High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J Virol *81*, 12135–12144.

Egloff, M.-P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E.J., Gorbalenya, A.E., Cambillau, C., and Canard, B. (2004). The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. U.S.a. *101*, 3792–3796.

Elena, S.F., and Moya, A. (1999). Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J. Evol. Biol. *12*, 1078–1088.

Elena, S.F., and Sanjuán, R. (2005). Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol *79*, 11555–11558.

Elena, S.F., Solé, R.V., and Sardanyés, J. (2010). Simple genomes, complex interactions: epistasis in RNA virus. Chaos *20*, 026106.

Esper, F., Ou, Z., and Huang, Y.T. (2010). Human coronaviruses are uncommon in patients with gastrointestinal illness. J. Clin. Virol. *48*, 131–133.

Estola, T. (1973). Coronaviruses, a new group of animal RNA viruses. Avian Dis. *14*, 330–336.

Falsey, A.R., McCann, R.M., Hall, W.J., Criddle, M.M., Formica, M.A., Wycoff, D., and Kolassa, J.E. (1997). The “Common Cold” in Frail Older Persons: Impact of Rhinovirus and Coronavirus in a Senior Daycare Center. Journal of the American Geriatrics Society *45*, 706–711.

Falsey, A.R., Walsh, E.E., and Hayden, F.G. (2002). Rhinovirus and coronavirus infection-associated hospitalizations among older adults. J. Infect. Dis. *185*, 1338–1341.

Feng, K.Y., Chen, T., Zhang, X., Shao, G.M., Cao, Y., Chen, D.K., Lin, W.C., Chen, F., and Xie, Q.M. (2018). Molecular characteristic and pathogenicity analysis of a virulent recombinant avain infectious bronchitis virus isolated in China. Poult. Sci. *97*, 3519–3531.

Feng, Z., Hensley, L., McKnight, K.L., Hu, F., Madden, V., Ping, L., Jeong, S.-H., Walker, C., Lanford, R.E., and Lemon, S.M. (2013). A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. Nature *496*, 367–371.

Ferron, F., Subissi, L., Silveira De Morais, A.T., Le, N.T.T., Sevajol, M., Gluais, L., Decroly, E., Vonrhein, C., Bricogne, G., Canard, B., et al. (2018). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. Proc. Natl. Acad. Sci. U.S.a. *115*, E162–E171.

Fett, C., Dediego, M.L., Regla-Nava, J.A., Enjuanes, L., and Perlman, S. (2013). Complete protection against severe acute respiratory syndrome coronavirus-mediated lethal respiratory

disease in aged mice by immunization with a mouse-adapted virus lacking E protein. J Virol *87*, 6551–6559.

Fitzsimmons, W.J., Woods, R.J., McCrone, J.T., Woodman, A., Arnold, J.J., Yennawar, M., Evans, R., Cameron, C.E., and Lauring, A.S. (2018). A speed–fidelity trade-off determines the mutation rate and virulence of an RNA virus. PLoS Biol. *16*, e2006459–20.

Forni, D., Cagliani, R., Clerici, M., and Sironi, M. (2016a). Molecular Evolution of Human Coronavirus Genomes. Trends Microbiol. 1–14.

Forni, D., Cagliani, R., Mozzi, A., Pozzoli, U., Al-Daghri, N., Clerici, M., and Sironi, M. (2016b). Extensive Positive Selection Drives the Evolution of Nonstructural Proteins in Lineage C Betacoronaviruses. J Virol *90*, 3627–3639.

Fouchier, R.A.M., Hartwig, N.G., Bestebroer, T.M., Niemeyer, B., de Jong, J.C., Simon, J.H., and Osterhaus, A.D.M.E. (2004). A previously undescribed coronavirus associated with respiratory disease in humans. Proc. Natl. Acad. Sci. U.S.a. *101*, 6212–6216.

Furió, V., Moya, A., and Sanjuán, R. (2005). The cost of replication fidelity in an RNA virus. Proc. Natl. Acad. Sci. U.S.a. *102*, 10233–10237.

Gadlage, M.J., Graham, R.L., and Denison, M.R. (2008). Murine coronaviruses encoding nsp2 at different genomic loci have altered replication, protein expression, and localization. J Virol *82*, 11964–11969.

Gaunt, E.R., Hardie, A., Claas, E.C.J., Simmonds, P., and Templeton, K.E. (2010). Epidemiology and Clinical Presentations of the Four Human Coronaviruses 229E, HKU1, NL63, and OC43 Detected over 3 Years Using a Novel Multiplex Real-Time PCR Method. J. Clin. Microbiol. *48*, 2940–2947.

Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., et al. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature *503*, 535–538.

Geller, C., Varbanov, M., and Duval, R. (2012). Human Coronaviruses: Insights into Environmental Resistance and Its Influence on the Development of New Antiseptic Strategies. Viruses *4*, 3044–3068.

Geller, R., Domingo-Calap, P., Cuevas, J.M., Rossolillo, P., Negroni, M., and Sanjuán, R. (2015). The external domains of the HIV-1 envelope are a mutational cold spot. Nat Commun *6*, 8571.

Geller, R., Estada, Ú., Peris, J.B., Andreu, I., Bou, J.-V., Garijo, R., Cuevas, J.M., Sabariegos, R., Mas, A., and Sanjuán, R. (2016). Highly heterogeneous mutation rates in the hepatitis C virus genome. Nat Microbiol *1*, 16045.

Geoghegan, J.L., and Holmes, E.C. (2017). Predicting virus emergence amid evolutionary noise. Open Biol. *7*, 170189–170189.

Gerna, G., Campanini, G., Rovida, F., Percivalle, E., Sarasini, A., Marchi, A., and Baldanti, F. (2006). Genetic variability of human coronavirus OC43-, 229E-, and NL63-like strains and their association with lower respiratory tract infections of hospitalized infants and immunocompromised patients. J. Med. Virol. *78*, 938–949.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gnädig, N.F., Beaucourt, S., Campagnola, G., Bordería, A.V., Sanz-Ramos, M., Gong, P., Blanc, H., Peersen, O.B., and Vignuzzi, M. (2012). Coxsackievirus B3 mutator strains are attenuated in vivo. Proc. Natl. Acad. Sci. U.S.a. *109*, E2294–E2303.

Goebel, S.J., Miller, T.B., Bennett, C.J., Bernard, K.A., and Masters, P.S. (2007). A Hypervariable Region within the 3' cis-Acting Element of the Murine Coronavirus Genome Is Nonessential for RNA Synthesis but Affects Pathogenesis. J Virol *81*, 1274–1287.

Goebel, S.J., Hsue, B., Dombrowski, T.F., and Masters, P.S. (2004a). Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. J Virol *78*, 669–682.

Goebel, S.J., Taylor, J., and Masters, P.S. (2004b). The 3' cis-acting genomic replication element of the severe acute respiratory syndrome coronavirus can function in the murine coronavirus genome. J Virol *78*, 7846–7851.

Goldhill, D., Lee, A., Williams, E.S.C.P., and Turner, P.E. (2014). Evolvability and robustness in populations of RNA virus Φ6. Frontiers in Microbiology *5*.

Gong, L.I., and Bloom, J.D. (2014). Epistatically Interacting Substitutions Are Enriched during Adaptive Protein Evolution. PLOS Genet *10*, e1004328–10.

Gorbalenya, A.E., brinton, M.A., J, C., de Groot, R.J., Gulyaeva, A., Lauber, C., Neuman, B.W., and Ziebuhr, J. (2017). ICTV Pending Proposal 2017.015S. Reorganization and expansion of the order Nidovirales at the family and sub-order ranks.

Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., and Snijder, E.J. (2006). Nidovirales: evolving the largest RNA virus genome. Virus Research *117*, 17–37.

Graci, J.D., Gnädig, N.F., Galarraga, J.E., Castro, C., Vignuzzi, M., and Cameron, C.E. (2012). Mutational robustness of an RNA virus influences sensitivity to lethal mutagenesis. J Virol *86*, 2869–2873.

Graepel, K.W., Lu, X., Case, J.B., Sexton, N.R., Smith, E.C., and Denison, M.R. (2017). Proofreading-Deficient Coronaviruses Adapt for Increased Fitness over Long-Term Passage without Reversion of Exoribonuclease-Inactivating Mutations. MBio *8*, e01503–e01517.

Graham, R.L., Becker, M.M., Eckerle, L.D., Bolles, M., Denison, M.R., and Baric, R.S. (2012). A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse

model of lethal disease. Nat. Med. *18*, 1820–1826.

Graham, R.L., Deming, D.J., Deming, M.E., Yount, B.L., and Baric, R.S. (2018). Evaluation of a recombination-resistant coronavirus as a broadly applicable, rapidly implementable vaccine platform. Communications Biology *1*, 1–10.

Gralinski, L.E., and Baric, R.S. (2015). Molecular pathology of emerging coronavirus infections. J. Pathol. *235*, 185–195.

Gu, J., Gong, E., Zhang, B., Zheng, J., Gao, Z., Zhong, Y., Zou, W., Zhan, J., Wang, S., Xie, Z., et al. (2005). Multiple organ infection and the pathogenesis of SARS. J. Exp. Med. *202*, 415–424.

Gu, S., Li, W., Zhang, H., Fleming, J., Yang, W., Wang, S., Wei, W., Zhou, J., Zhu, G., Deng, J., et al. (2016). The β2 clamp in the Mycobacterium tuberculosis DNA polymerase III αβ2ε replicase promotes polymerization and reduces exonuclease activity. Sci Rep *6*, 18418.

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., et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science *302*, 276–278.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51–59.

Hodinka, R.L. (2016). Respiratory RNA Viruses. Microbiology Spectrum *4*.

Holmes, E.C. (2003). Error thresholds and the constraints to RNA virus evolution. Trends Microbiol. *11*, 543–546.

Holmes, E.C. (2013). What can we predict about viral evolution and emergence? Curr Opin Virol *3*, 180–184.

Holmes, E.C. (2016). The Expanding Virosphere. Cell Host and Microbe *20*, 279–280.

Holtz, C.M., and Mansky, L.M. (2013). Variation of HIV-1 mutation spectra among cell types. J Virol *87*, 5296–5299.

Hon, C.C., Lam, T.Y., Shi, Z.L., Drummond, A.J., Yip, C.W., Zeng, F., Lam, P.Y., and Leung, F.C.C. (2008). Evidence of the Recombinant Origin of a Bat Severe Acute Respiratory Syndrome (SARS)-Like Coronavirus and Its Implications on the Direct Ancestor of SARS Coronavirus. J Virol *82*, 1819–1826.

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., et al. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog *13*, e1006698–27.

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J.C., Nagel, J., Johnson, J.B., Agnihothram, S., Gates, J.E., et al. (2012). Evidence supporting a zoonotic origin of human

coronavirus strain NL63. J Virol *86*, 12816–12825.

Iketani, S., Shean, R.C., Ferren, M., Makhsous, N., Aquino, D.B., Georges, des, A., Rima, B., Mathieu, C., Porotto, M., Moscona, A., et al. (2018). Viral Entry Properties Required for Fitness in Humans Are Lost through Rapid Genomic Change during Viral Isolation. MBio *9*, 1545–19.

Imbert, I., Guillemot, J.-C., Bourhis, J.-M., Bussetta, C., Coutard, B., Egloff, M.-P., Ferron, F., Gorbalenya, A.E., and Canard, B. (2006). A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. Embo J. *25*, 4933–4942.

Imbert, I., Snijder, E.J., Dimitrova, M., Guillemot, J.-C., Lécine, P., and Canard, B. (2008). The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. Virus Research *133*, 136–148.

Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y.-W., Siddell, S.G., and Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. PLoS Pathog *12*, e1005473.

Jabara, C.B., Hu, F., Mollan, K.R., Williford, S.E., Menezes, P., Yang, Y., Eron, J.J., Fried, M.W., Hudgens, M.G., Jones, C.D., et al. (2014). Hepatitis C Virus (HCV) NS3 sequence diversity and antiviral resistance-associated variant frequency in HCV/HIV coinfection. Antimicrob. Agents Chemother. *58*, 6079–6092.

Jabara, C.B., Jones, C.D., Roach, J., Anderson, J.A., and Swanstrom, R. (2011). Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. Proc. Natl. Acad. Sci. U.S.a. *108*, 20166–20171.

Jevšnik, M., Steyer, A., Pokorn, M., Mrvič, T., Grosek, Š., Strle, F., Lusa, L., and Petrovec, M. (2016). The Role of Human Coronaviruses in Children Hospitalized for Acute Bronchiolitis, Acute Gastroenteritis, and Febrile Seizures: A 2-Year Prospective Study. PLoS ONE *11*, e0155555–15.

Jevšnik, M., Steyer, A., Zrim, T., Pokorn, M., Mrvič, T., Grosek, Š., Strle, F., Lusa, L., and Petrovec, M. (2013). Detection of human coronaviruses in simultaneously collected stool samples and nasopharyngeal swabs from hospitalized children with acute gastroenteritis. Virol. J. *10*, 46.

Kan, B., Wang, M., Jing, H., Xu, H., Jiang, X., Yan, M., Liang, W., Zheng, H., Wan, K., Liu, Q., et al. (2005). Molecular Evolution Analysis and Geographic Investigation of Severe Acute Respiratory Syndrome Coronavirus-Like Virus in Palm Civets at an Animal Market and on Farms. J Virol *79*, 11892–11900.

Kautz, T., and Forrester, N. (2018). RNA Virus Fidelity Mutants: A Useful Tool for Evolutionary Biology or a Complex Challenge? Viruses *10*, 600–617.

Kautz, T.F., Guerbois, M., Khanipov, K., Patterson, E.I., Langsjoen, R.M., Yun, R., Warmbrod, K.L., Fofanov, Y., Weaver, S.C., and Forrester, N.L. (2018). Low-fidelity Venezuelan equine encephalitis virus polymerase mutants to improve live-attenuated vaccine safety and efficacy.

Virus Evol *4*, 721–14.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845–858.

Kempf, B.J., Peersen, O.B., and Barton, D.J. (2016). Poliovirus Polymerase Leu420 Facilitates RNA Recombination and Ribavirin Resistance. J Virol *90*, 8410–8421.

Kew, O. (2012). Reaching the last one per cent: progress and challenges in global polio eradication. Curr Opin Virol *2*, 188–198.

Ki, M. (2015). 2015 MERS outbreak in Korea: hospital-to-hospital transmission. Epidemiol Health *37*, e2015033.

Kim, E., Okada, K., Kenniston, T., Raj, V.S., AlHajri, M.M., Farag, E.A.B.A., Alhajri, F., Osterhaus, A.D.M.E., Haagmans, B.L., and Gambotto, A. (2014). Immunogenicity of an adenoviral-based Middle East Respiratory Syndrome coronavirus vaccine in BALB/c mice. Vaccine *32*, 5975–5982.

Kim, S.Y., Park, S.J., Cho, S.Y., Cha, R.-H., Jee, H.-G., Kim, G., Shin, H.-S., Kim, Y., Jung, Y.M., Yang, J.-S., et al. (2016). Viral RNA in Blood as Indicator of Severe Outcome in Middle East Respiratory Syndrome Coronavirus Infection. Emerging Infect. Dis. *22*, 1813–1816.

Kin, N., Miszczak, F., Diancourt, L., Caro, V., Moutou, F., Vabret, A., and Gouilh, M.A. (2016). Comparative molecular epidemiology of two closely related coronaviruses, bovine coronavirus (BCoV) and human coronavirus OC43 (HCoV-OC43), reveals a different evolutionary pattern. Infect. Genet. Evol. *40*, 186–191.

Kindler, E., Gil-Cruz, C., Spanier, J., Li, Y., Wilhelm, J., Rabouw, H.H., Züst, R., Hwang, M., V'kovski, P., Stalder, H., et al. (2017). Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. PLoS Pathog *13*, e1006195.

Knoops, K., Kikkert, M., Worm, S.H.E.V.D., Zevenhoven-Dobbe, J.C., van der Meer, Y., Koster, A.J., Mommaas, A.M., and Snijder, E.J. (2008). SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol. *6*, e226.

Koel, B.F., Burke, D.F., van der Vliet, S., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., Smith, D.J., and Fouchier, R.A.M. (2018). Epistatic interactions can moderate the antigenic effect of substitutions in hemagglutinin of influenza H3N2 virus. 1–16.

Koel, B.F., van Someren Greve, F., Vigeveno, R.M., Pater, M., Russell, C.A., and de Jong, M.D. (2019). Disparate evolution of virus populations in upper and lower airways of mechanically ventilated patients. bioRxiv http:—biorxiv.org–lookup–doi–10.1101–509901.

Korboukh, V.K., Lee, C.A., Acevedo, A., Vignuzzi, M., Xiao, Y., Arnold, J.J., Hemperly, S., Graci, J.D., August, A., Andino, R., et al. (2014). RNA virus population diversity, an optimum for maximal fitness and virulence. J. Biol. Chem. *289*, 29531–29544.

Korneeva, V.S., and Cameron, C.E. (2007). Structure-function relationships of the viral RNA-dependent RNA polymerase: fidelity, replication speed, and initiation mechanism determined by a residue in the ribose-binding pocket. J. Biol. Chem. *282*, 16135–16145.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., et al. (2003). A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. N. Engl. J. Med. *348*, 1953–1966.

Kuo, L., Koetzner, C.A., and Masters, P.S. (2016). A key role for the carboxy-terminal tail of the murine coronavirus nucleocapsid protein in coordination of genome packaging. Virology *494*, 100–107.

Kuypers, J., Martin, E.T., Heugel, J., Wright, N., Morrow, R., and Englund, J.A. (2007). Clinical Disease in Children Associated With Newly Described Coronavirus Subtypes. Pediatrics *119*, e70–e76.

Lai, M.M., Baric, R.S., Makino, S., Keck, J.G., Egbert, J., Leibowitz, J.L., and Stohlman, S.A. (1985). Recombination between nonsegmented RNA genomes of murine coronaviruses. J Virol *56*, 449–456.

Lamers, M.M., Raj, V.S., Shafei, M., Ali, S.S., Abdallh, S.M., Gazo, M., Nofal, S., Lu, X., Erdman, D.D., Koopmans, M.P., et al. (2016). Deletion Variants of Middle East Respiratory Syndrome Coronavirus from Humans, Jordan, 2015. Emerging Infect. Dis. *22*, 716–719.

Lamirande, E.W., Dediego, M.L., Roberts, A., Jackson, J.P., Alvarez, E., Sheahan, T., Shieh, W.-J., Zaki, S.R., Baric, R., Enjuanes, L., et al. (2008). A live attenuated severe acute respiratory syndrome coronavirus is immunogenic and efficacious in golden Syrian hamsters. J Virol *82*, 7721–7724.

Lau, S.K.P., Li, K.S.M., Tsang, A.K.L., Lam, C.S.F., Ahmed, S., Chen, H., Chan, K.-H., Woo, P.C.Y., and Yuen, K.-Y. (2013). Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J Virol *87*, 8638–8650.

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., and Yuen, K.-Y. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. U.S.a. *102*, 14040–14045.

Lau, S.K.P., Woo, P.C.Y., Li, K.S.M., Tsang, A.K.L., Fan, R.Y.Y., Luk, H.K.H., Cai, J.-P., Chan, K.-H., Zheng, B.-J., Wang, M., et al. (2015). Discovery of a novel coronavirus, China Rattus coronavirus HKU24, from Norway rats supports the murine origin of Betacoronavirus 1 and has implications for the ancestor of Betacoronavirus lineage A. J Virol *89*, 3076–3092.

Lauber, C., Goeman, J.J., del Carmen Parquet, M., Nga, P.T., Snijder, E.J., Morita, K., and Gorbalenya, A.E. (2013). The Footprint of Genome Architecture in the Largest Genome Expansion in RNA Viruses. PLoS Pathog *9*, e1003500.

Lauring, A.S., and Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses. PLoS Pathog *6*, e1001005.

Lauring, A.S., Acevedo, A., Cooper, S.B., and Andino, R. (2012). Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. Cell Host and Microbe *12*, 623–632.

Lauring, A.S., Frydman, J., and Andino, R. (2013). The role of mutational robustness in RNA virus evolution. Nat. Rev. Microbiol. *11*, 327–336.

Le Nouën, C., McCarty, T., Brown, M., Smith, M.L., Lleras, R., Dolan, M.A., Mehedi, M., Yang, L., Luongo, C., Liang, B., et al. (2017). Genetic stability of genome-scale deoptimized RNA virus vaccine candidates under selective pressure. Proc. Natl. Acad. Sci. U.S.a. *114*, E386–E395.

Lech, W.J., Wang, G., Yang, Y.L., Chee, Y., Dorman, K., McCrae, D., Lazzeroni, L.C., Erickson, J.W., Sinsheimer, J.S., and Kaplan, A.H. (1996). In vivo sequence diversity of the protease of human immunodeficiency virus type 1: presence of protease inhibitor-resistant variants in untreated subjects. J Virol *70*, 2038–2043.

Lee, C.A., August, A., Arnold, J.J., and Cameron, C.E. (2016). Polymerase Mechanism-Based Method of Viral Attenuation. Methods Mol. Biol. *1349*, 83–104.

Lehmann, K.C., Gulyaeva, A., Zevenhoven-Dobbe, J.C., Janssen, G.M.C., Ruben, M., Overkleeft, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res. *43*, 8416–8434.

Leibowitz, J., Kaufman, G., and Liu, P. (2011). Coronaviruses: propagation, quantification, storage, and construction of recombinant mouse hepatitis virus. Curr Protoc Microbiol *Chapter 15*, Unit15E.1–15E.1.46.

Leung, G.M., Hedley, A.J., Ho, L.-M., Chau, P., Wong, I.O.L., Thach, T.Q., Ghani, A.C., Donnelly, C.A., Fraser, C., Riley, S., et al. (2004). The epidemiology of severe acute respiratory syndrome in the 2003 Hong Kong epidemic: an analysis of all 1755 patients. Ann Intern Med *141*, 662–673.

Li, J., Ulitzky, L., Silberstein, E., Taylor, D.R., and Viscidi, R. (2013). Immunogenicity and protection efficacy of monomeric and trimeric recombinant SARS coronavirus spike protein subunit vaccine candidates. Viral Immunol. *26*, 126–132.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., et al. (2005). Bats are natural reservoirs of SARS-like coronaviruses. Science *310*, 676–679.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450–454.

Li, Y., and Weiss, S.R. (2016). Antagonism of RNase L is required for murine coronavirus replication in Kupffer cells and liver sinusoidal endothelial cells but not in hepatocytes. J Virol JVI.01423–16–24.

Lissenberg, A., Vrolijk, M.M., van Vliet, A.L.W., Langereis, M.A., de Groot-Mijnes, J.D.F., Rottier, P.J.M., and de Groot, R.J. (2005). Luxury at a Cost? Recombinant Mouse Hepatitis Viruses Expressing the Accessory Hemagglutinin Esterase Protein Display Reduced Fitness In Vitro. J Virol *79*, 15054–15063.

Liu, X., Musser, D., Lee, C., Yang, X., Arnold, J., Cameron, C., and Boehr, D. (2015). Nucleobase but not Sugar Fidelity is Maintained in the Sabin I RNA-Dependent RNA Polymerase. Viruses *7*, 5571–5586.

Liu, X., Yang, X., Lee, C.A., Moustafa, I.M., Smidansky, E.D., Lum, D., Arnold, J.J., Cameron, C.E., and Boehr, D.D. (2013). Vaccine-derived mutation in motif D of poliovirus RNA-dependent RNA polymerase lowers nucleotide incorporation fidelity. J. Biol. Chem. *288*, 32753–32765.

Lou, D.I., Hussmann, J.A., McBee, R.M., Acevedo, A., Andino, R., Press, W.H., and Sawyer, S.L. (2013). High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. Proc. Natl. Acad. Sci. U.S.a. *110*, 19872–19877.

Lugari, A., Betzi, S., Decroly, E., Bonnaud, E., Hermant, A., Guillemot, J.-C., Debarnot, C., Borg, J.-P., Bouvet, M., Canard, B., et al. (2010). Molecular mapping of the RNA Cap 2'-O-methyltransferase activation interface between severe acute respiratory syndrome coronavirus nsp10 and nsp16. J. Biol. Chem. *285*, 33230–33241.

Luke, T., Wu, H., Zhao, J., Channappanavar, R., Coleman, C.M., Jiao, J.-A., Matsushita, H., Liu, Y., Postnikova, E.N., Ork, B.L., et al. (2016). Human polyclonal immunoglobulin G from transchromosomic bovines inhibits MERS-CoV in vivo. Sci Transl Med *8*, 326ra21–326ra21.

Luria, S.E., and Delbrück, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics *28*, 491–511.

Luytjes, W., Bredenbeek, P.J., Noten, A.F., Horzinek, M.C., and Spaan, W.J. (1988). Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology *166*, 415–422.

Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc. Natl. Acad. Sci. U.S.a. *112*, 9436–9441.

Mair-Jenkins, J., Saavedra-Campos, M., Baillie, J.K., Cleary, P., Khaw, F.-M., Lim, W.S., Makki, S., Rooney, K.D., Convalescent Plasma Study Group, Nguyen-Van-Tam, J.S., et al. (2014). The Effectiveness of Convalescent Plasma and Hyperimmune Immunoglobulin for the Treatment of Severe Acute Respiratory Infections of Viral Etiology: A Systematic Review and Exploratory Meta-analysis. J. Infect. Dis. *211*, 80–90.

Makino, S., Keck, J.G., Stohlman, S.A., and Lai, M.M. (1986). High-frequency RNA recombination of murine coronaviruses. J Virol *57*, 729–737.

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W., and García-Arenal, F. (2002). The rate and character of spontaneous mutation in an RNA virus. Genetics *162*, 1505–1511.

Mandary, M., and Poh, C. (2018). Changes in the EV-A71 Genome through Recombination and Spontaneous Mutations: Impact on Virulence. Viruses *10*, 320–321.

Masters, P.S. (2006). The Molecular Biology of Coronaviruses. (Elsevier), pp. 193–292.

Mäkelä, M.J., Puhakka, T., Ruuskanen, O., Leinonen, M., Saikku, P., Kimpimäki, M., Blomqvist, S., Hyypiä, T., and Arstila, P. (1998). Viruses and bacteria in the etiology of the common cold. J. Clin. Microbiol. *36*, 539–542.

McCrone, J.T., and Lauring, A.S. (2017). Genetic bottlenecks in intraspecies virus transmission. Curr Opin Virol *28*, 20–25.

McIntosh, K., Ellis, E.F., Hoffman, L.S., Lybass, T.G., Eller, J.J., and Fulginiti, V.A. (1973). The association of viral and bacterial respiratory infections with exacerbations of wheezing in young asthmatic children. J. Pediatr. *82*, 578–590.

McIntosh, K., Kapikian, A.Z., Turner, H.C., Hartley, J.W., Parrott, R.H., and Chanock, R.M. (1970). Seroepidemiologic studies of coronavirus infection in adults and children. Am. J. Epidemiol. *91*, 585–592.

Memish, Z.A., Assiri, A.M., and Al-Tawfiq, J.A. (2014). Middle East respiratory syndrome coronavirus (MERS-CoV) viral shedding in the respiratory tract: an observational analysis with infection control implications. International Journal of Infectious Diseases *29*, 307–308.

Menachery, V.D., Gralinski, L.E., Mitchell, H.D., Dinnon, K.H., III, Leist, S.R., Yount, B.L., Jr., McAnarney, E.T., Graham, R.L., Waters, K.M., and Baric, R.S. (2018). Combination attenuation offers strategy for live-attenuated coronavirus vaccines. J Virol JVI.00710–18–35.

Menachery, V.D., Yount, B.L., Debbink, K., Agnihothram, S., Gralinski, L.E., Plante, J.A., Graham, R.L., Scobey, T., Ge, X.-Y., Donaldson, E.F., et al. (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat. Med. *21*, 1508–1513.

Menachery, V.D., Yount, B.L., Josset, L., Gralinski, L.E., Scobey, T., Agnihothram, S., Katze, M.G., and Baric, R.S. (2014). Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity. J Virol *88*, 4251–4264.

Menachery, V.D., Yount, B.L., Sims, A.C., Debbink, K., Agnihothram, S.S., Gralinski, L.E., Graham, R.L., Scobey, T., Plante, J.A., Royal, S.R., et al. (2016). SARS-like WIV1-CoV poised for human emergence. Proc. Natl. Acad. Sci. U.S.a. *113*, 201517719–3053.

Meng, T., and Kwang, J. (2014). Attenuation of human enterovirus 71 high-replication-fidelity variants in AG129 mice. J Virol *88*, 5803–5815.

Minor, P. (2009). Vaccine-derived poliovirus (VDPV): Impact on poliomyelitis eradication. Vaccine *27*, 2649–2652.

Minskaia, E., Hertzig, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., and Ziebuhr, J. (2006). Discovery of an RNA virus 3"->5" exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc. Natl. Acad. Sci. U.S.a. *103*, 5108–5113.

Modjarrad, K. (2016). Treatment strategies for Middle East respiratory syndrome coronavirus. J Virus Erad *2*, 1–4.

Moës, E., Vijgen, L., Keyaerts, E., Zlateva, K., Li, S., Maes, P., Pyrc, K., Berkhout, B., van der Hoek, L., and Van Ranst, M. (2005). A novel pancoronavirus RT-PCR assay: frequent detection of human coronavirus NL63 in children hospitalized with respiratory tract infections in Belgium. BMC Infectious Diseases *5*, 629–10.

Montville, R., Froissart, R., Remold, S.K., Tenaillon, O., and Turner, P.E. (2005). Evolution of mutational robustness in an RNA virus. PLoS Biol. *3*, e381.

Moratorio, G., Henningsson, R., Barbezange, C., Carrau, L., Bordería, A.V., Blanc, H., Beaucourt, S., Poirier, E.Z., Vallet, T., Boussier, J., et al. (2017). Attenuation of RNA viruses by redirecting their evolution in sequence space. Nat Microbiol *2*, 1–12.

Moreno, H., Gallego, I., Sevilla, N., la Torre, de, J.C., Domingo, E., and Martín, V. (2011). Ribavirin can be mutagenic for arenaviruses. J Virol *85*, 7246–7255.

Moreno, H., Tejero, H., la Torre, de, J.C., Domingo, E., and Martín, V. (2012). Mutagenesis-Mediated Virus Extinction: Virus-Dependent Effect of Viral Load on Sensitivity to Lethal Defection. PLoS ONE *7*, e32550–18.

Moustafa, I.M., Korboukh, V.K., Arnold, J.J., Smidansky, E.D., Marcotte, L.L., Gohara, D.W., Yang, X., Sánchez-Farrán, M.A., Filman, D., Maranas, J.K., et al. (2014). Structural dynamics as a contributor to error-prone replication by an RNA-dependent RNA polymerase. J. Biol. Chem. *289*, 36229–36248.

Muller, M.P., Dresser, L., Raboud, J., McGeer, A., Rea, E., Richardson, S.E., Mazzulli, T., Loeb, M., Louie, M., Canadian SARS Research Network (2007). Adverse events associated with high-dose ribavirin: evidence from the Toronto outbreak of severe acute respiratory syndrome. Pharmacotherapy *27*, 494–503.

Muth, D., Corman, V.M., Roth, H., Binger, T., Dijkman, R., Gottula, L.T., Gloza-Rausch, F., Balboni, A., Battilani, M., x0010D, D.R., et al. (2018). Attenuation of replication by a 29 nucleotide deletion in SARS- coronavirus acquired during the early stages of human-to-human transmission. Sci Rep *8*, 1–11.

Muthumani, K., Falzarano, D., Reuschel, E.L., Tingey, C., Flingai, S., Villarreal, D.O., Wise, M., Patel, A., Izmirly, A., Aljuaid, A., et al. (2015). A synthetic consensus anti–spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates. Sci Transl Med *7*, 301ra132–301ra132.

Müller, M.A., Corman, V.M., Jores, J., Meyer, B., Younan, M., Liljander, A., Bosch, B.-J., Lattwein, E., Hilali, M., Musa, B.E., et al. (2014). MERS Coronavirus Neutralizing Antibodies in Camels, Eastern Africa, 1983–1997. Emerging Infect. Dis. *20*, 1–3.

Mzingwane, M.L., Tiemessen, C.T., Richter, K.L., Mayaphi, S.H., Hunt, G., and Bowyer, S.M. (2016). Pre-treatment minority HIV-1 drug resistance mutations and long term virological outcomes: is prediction possible? Virol. J. *13*, 1–9.

Nakajima, K., Nobusawa, E., Nagy, A., and Nakajima, S. (2005). Accumulation of Amino Acid Substitutions Promotes Irreversible Structural Changes in the Hemagglutinin of Human Influenza AH3 Virus during Evolution. J Virol *79*, 6472–6477.

National Research Council (US) (2000). Microbial Status and Genetic Evaluation of Mice and Rats: Proceedings of the 1999 US/Japan Conference.

Ng, K.K.S., Arnold, J.J., and Cameron, C.E. (2008). Structure-function relationships among RNA-dependent RNA polymerases. Curr. Top. Microbiol. Immunol. *320*, 137–156.

Nicholson, K.G., Kent, J., and Ireland, D.C. (1993). Respiratory viruses and exacerbations of asthma in adults. Bmj *307*, 982–986.

O'Dea, E.B., Keller, T.E., and Wilke, C.O. (2010). Does Mutational Robustness Inhibit Extinction by Lethal Mutagenesis in Viral Populations? PLoS Comput Biol *6*, e1000811–e1000812.

Oh, M.-D., Park, W.B., Choe, P.G., Choi, S.-J., Kim, J.-I., Chae, J., Park, S.S., Kim, E.-C., Oh, H.S., Kim, E.J., et al. (2016). Viral Load Kinetics of MERS Coronavirus Infection. N. Engl. J. Med. *375*, 1303–1305.

Oudshoorn, D., Rijs, K., Limpens, R.W.A.L., Groen, K., Koster, A.J., Snijder, E.J., Kikkert, M., and Bárcena, M. (2017). Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. MBio *8*, e01658–17–17.

Pallesen, J., Wang, N., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., Cottrell, C.A., Becker, M.M., Wang, L., Shi, W., et al. (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proc. Natl. Acad. Sci. U.S.a. *68*, 201707304–E201707357.

Pandey, V.N., Kaushik, N., Rege, N., Sarafianos, S.G., Yadav, P.N., and Modak, M.J. (1996). Role of methionine 184 of human immunodeficiency virus type-1 reverse transcriptase in the polymerase function and fidelity of DNA synthesis. Biochemistry *35*, 2168–2179.

Pascal, K.E., Coleman, C.M., Mujica, A.O., Kamat, V., Badithe, A., Fairhurst, J., Hunt, C., Strein, J., Berrebi, A., Sisk, J.M., et al. (2015). Pre- and postexposure efficacy of fully human antibodies against Spike protein in a novel humanized mouse model of MERS-CoV infection. Proc. Natl. Acad. Sci. U.S.a. *112*, 8738–8743.

Pathak, V.K., and Temin, H.M. (1992). 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. J Virol *66*, 3093–3100.

Pauly, M.D., and Lauring, A.S. (2015). Effective lethal mutagenesis of influenza virus by three nucleoside analogs. J Virol *89*, 3584–3597.

Pauly, M.D., Lyons, D.M., Fitzsimmons, W.J., and Lauring, A.S. (2017a). Epistatic Interactions within the Influenza A Virus Polymerase Complex Mediate Mutagen Resistance and Replication Fidelity. mSphere *2*, e00323–17–15.

Pauly, M.D., Procario, M.C., and Lauring, A.S. (2017b). A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. Elife *6*, 686.

Payne, D.C., Biggs, H.M., Al-Abdallat, M.M., Alqasrawi, S., Lu, X., Abedi, G.R., Haddadin, A., Iblan, I., Alsanouri, T., Nsour, Al, M., et al. (2018). Multihospital Outbreak of a Middle East Respiratory Syndrome Coronavirus Deletion Variant, Jordan: A Molecular, Serologic, and Epidemiologic Investigation. Open Forum Infect Dis *5*, ofy095.

Peiris, J.S.M., Chu, C.M., Cheng, V.C.C., Chan, K.S., Hung, I.F.N., Poon, L.L.M., Law, K.I., Tang, B.S.F., Hon, T.Y.W., Chan, C.S., et al. (2003a). Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. The Lancet *361*, 1767–1772.

Peiris, J., Lai, S.T., Poon, L., Guan, Y., Yam, L., Lim, W., Nicholls, J., Yee, W., Yan, W.W., Cheung, M.T., et al. (2003b). Coronavirus as a possible cause of severe acute respiratory syndrome. The Lancet *361*, 1319–1325.

Pereira-Gómez, M., and Sanjuán, R. (2014). Delayed lysis confers resistance to the nucleoside analogue 5-fluorouracil and alleviates mutation accumulation in the single-stranded DNA bacteriophage ϕX174. J Virol *88*, 5042–5049.

Peris, J.B., Davis, P., Cuevas, J.M., Nebot, M.R., and Sanjuán, R. (2010). Distribution of Fitness Effects Caused by Single-Nucleotide Substitutions in Bacteriophage f1. Genetics *185*, 603–609.

Perlman, S., Schelper, R., Bolger, E., and Ries, D. (1987). Late onset, symptomatic, demyelinating encephalomyelitis in mice infected with MHV-JHM in the presence of maternal antibody. Microbial Pathogenesis *2*, 185–194.

Perlman, S., and Netland, J. (2009). Coronaviruses post-SARS: update on replication and pathogenesis. Nat. Rev. Microbiol. *7*, 439–450.

Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., Müller, M.A., Annan, A., Vallo, P., Adu-Sarkodie, Y., et al. (2009). Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. Emerging Infect. Dis. *15*, 1377–1384.

Pfeiffer, J.K., and Kirkegaard, K. (2003). A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. Proc. Natl.

Acad. Sci. U.S.a. *100*, 7289–7294.

Pfeiffer, J.K., and Kirkegaard, K. (2005). Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS Pathog *1*, e11.

Pita, J.S., de Miranda, J.R., Schneider, W.L., and Roossinck, M.J. (2007). Environment determines fidelity for an RNA virus replicase. J Virol *81*, 9072–9077.

Poon, L.L.M., Chu, D.K.W., Chan, K.H., Wong, O.K., Ellis, T.M., Leung, Y.H.C., Lau, S.K.P., Woo, P.C.Y., Suen, K.Y., Yuen, K.Y., et al. (2005). Identification of a novel coronavirus in bats. J Virol *79*, 2001–2009.

Pyrc, K., Berkhout, B., and van der Hoek, L. (2007). The Novel Human Coronaviruses NL63 and HKU1. J Virol *81*, 3051–3057.

Rabaan, A.A., Alahmed, S.H., Bazzi, A.M., and Alhani, H.M. (2017). A review of candidate therapies for Middle East respiratory syndrome from a molecular perspective. Journal of Medical Microbiology *66*, 1261–1274.

Rabenau, H.F., Cinatl, J., Morgenstern, B., Bauer, G., Preiser, W., and Doerr, H.W. (2004). Stability and inactivation of SARS coronavirus. Med Microbiol Immunol *194*, 1–6.

Rai, D.K., Diaz-San Segundo, F., Campagnola, G., Keith, A., Schafer, E.A., Kloc, A., de los Santos, T., Peersen, O., and Rieder, E. (2017). Attenuation of Foot-and-Mouth Disease Virus by Engineered Viral Polymerase Fidelity. J Virol *91*.

Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Müller, M.A., Dijkman, R., Muth, D., Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature *495*, 251–254.

Reed, S.E. (1984). The behaviour of recent isolates of human respiratory coronavirus in vitro and in volunteers: evidence of heterogeneity among 229E-related strains. J. Med. Virol. *13*, 179–192.

Regoes, R.R., Hamblin, S., and Tanaka, M.M. (2013). Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. Proc. Biol. Sci. *280*, 20122047–20122047.

Riemersma, K.K., Steiner, C., Singapuri, A., and Coffey, L.L. (2018). Chikungunya virus fidelity variants exhibit differential attenuation and population diversity in cell culture and adult mice. J Virol.

Rota, P.A. (2003). Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. Science *300*, 1394–1399.

Rozen-Gagnon, K., Stapleford, K.A., Mongelli, V., Blanc, H., Failloux, A.-B., Saleh, M.-C., and Vignuzzi, M. (2014). Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. PLoS Pathog *10*, e1003877.

Sabir, J.S.M., Lam, T.T.-Y., Ahmed, M.M.M., Li, L., Shen, Y., Abo-Aba, S.E.M., Qureshi, M.I., Abu-Zeid, M., Zhang, Y., Khiyami, M.A., et al. (2016). Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. Science *351*, 81–84.

Sadeghipour, S., Bek, E.J., and McMinn, P.C. (2013). Ribavirin-resistant mutants of human enterovirus 71 express a high replication fidelity phenotype during growth in cell culture. J Virol.

Sadeghipour, S., and McMinn, P.C. (2013). A study of the virulence in mice of high copying fidelity variants of human enterovirus 71. Virus Research *176*, 265–272.

Saif, L.J. (2010). Bovine Respiratory Coronavirus. Veterinary Clinics of North America: Food Animal Practice *26*, 349–364.

Sanjuán, R., and Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cellular and Molecular Life Sciences 1–16.

Sanjuán, R., Cuevas, J.M., Furió, V., Holmes, E.C., and Moya, A. (2007). Selection for Robustness in Mutagenized RNA Viruses. PLOS Genet *3*, e93.

Sanjuán, R., Cuevas, J.M., Moya, A., and Elena, S.F. (2005). Epistasis and the adaptability of an RNA virus. Genetics *170*, 1001–1008.

Sanjuán, R., Moya, A., and Elena, S.F. (2004a). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. U.S.a. *101*, 8396–8401.

Sanjuán, R., Moya, A., and Elena, S.F. (2004b). The contribution of epistasis to the architecture of fitness in an RNA virus. Proc. Natl. Acad. Sci. U.S.a. *101*, 15376–15379.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. J Virol *84*, 9733–9748.

Sawicki, S.G., Sawicki, D.L., and Siddell, S.G. (2007). A contemporary view of coronavirus transcription. J Virol *81*, 20–29.

Sawicki, S.G., Sawicki, D.L., Younker, D., Meyer, Y., Thiel, V., Stokes, H., and Siddell, S.G. (2005). Functional and genetic analysis of coronavirus replicase-transcriptase proteins. PLoS Pathog *1*, e39.

Schindewolf, C., and Menachery, V. (2019). Middle East Respiratory Syndrome Vaccine Candidates: Cautious Optimism. Viruses *11*, 74–17.

Schmitt, M.W., Kennedy, S.R., Salk, J.J., Fox, E.J., Hiatt, J.B., and Loeb, L.A. (2012). Detection of ultra-rare mutations by next-generation sequencing. Proc. Natl. Acad. Sci. U.S.a. *109*, 14508–14513.

Schwarz, B., Routledge, E., and Siddell, S.G. (1990). Murine coronavirus nonstructural protein ns2 is not essential for virus replication in transformed cells. J Virol *64*, 4784–4791.

Seronello, S., Montanez, J., Presleigh, K., Barlow, M., Park, S.B., and Choi, J. (2011). Ethanol

and reactive species increase basal sequence heterogeneity of hepatitis C virus and produce variants with reduced susceptibility to antivirals. PLoS ONE *6*, e27436.

Sevajol, M., Subissi, L., Decroly, E., Canard, B., and Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. Virus Research *194*, 90–99.

Sexton, N.R. (2017). Murine Hepatitis Virus Regulation of Nucleotide Selectivity and Fidelity by the RNA-dependent RNA Polymerase and the 3′-5′ Exoribonuclease.

Sexton, N.R., Smith, E.C., Blanc, H., Vignuzzi, M., Peersen, O.B., and Denison, M.R. (2016). Homology-Based Identification of a Mutation in the Coronavirus RNA-Dependent RNA Polymerase That Confers Resistance to Multiple Mutagens. J Virol *90*, 7415–7428.

Sheahan, T.P., Sims, A.C., Graham, R.L., Menachery, V.D., Gralinski, L.E., Case, J.B., Leist, S.R., Pyrc, K., Feng, J.Y., Trantcheva, I., et al. (2017). Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. Sci Transl Med *9*, eaal3653.

Shirogane, Y., Watanabe, S., and Yanagi, Y. (2012). Cooperation between different RNA virus genomes produces a new phenotype. Nat Commun *3*, 1235.

Sierra, M., Airaksinen, A., González-López, C., Agudo, R., Arias, A., and Domingo, E. (2007). Foot-and-mouth disease virus mutant with decreased sensitivity to ribavirin: implications for error catastrophe. J Virol *81*, 2012–2024.

Sloots, T., McErlean, P., Speicher, D., Arden, K., Nissen, M., and MacKay, I. (2006). Evidence of human coronavirus HKU1 and human bocavirus in Australian children. Journal of Clinical Virology *35*, 99–102.

Smith, E.C., Blanc, H., Surdel, M.C., Vignuzzi, M., and Denison, M.R. (2013). Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. PLoS Pathog *9*, e1003565.

Smith, E.C., Case, J.B., Blanc, H., Isakov, O., Shomron, N., Vignuzzi, M., and Denison, M.R. (2015). Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. J Virol *89*, 6418–6426.

Smith, E.C., Sexton, N.R., and Denison, M.R. (2014). Thinking Outside the Triangle: Replication Fidelity of the Largest RNA Viruses. Annu Rev Virol *1*, 111–132.

Snijder, E.J., Decroly, E., and Ziebuhr, J. (2016). The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. Adv. Virus Res. *96*, 59–126.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L.M., Guan, Y., Rozanov, M., Spaan, W.J.M., and Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. *331*, 991–1004.

Song, F., Fux, R., Provacia, L.B., Volz, A., Eickmann, M., Becker, S., Osterhaus, A.D.M.E., Haagmans, B.L., and Sutter, G. (2013). Middle East Respiratory Syndrome Coronavirus Spike Protein Delivered by Modified Vaccinia Virus Ankara Efficiently Induces Virus-Neutralizing Antibodies. J Virol *87*, 11950–11954.

Song, J.Y., Cheong, H.J., Choi, M.J., Jeon, J.H., Kang, S.H., Jeong, E.J., Yoon, J.G., Lee, S.N., Kim, S.R., Noh, J.Y., et al. (2015). Viral Shedding and Environmental Cleaning in Middle East Respiratory Syndrome Coronavirus Infection. Infection & Chemotherapy *47*, 252–255.

St-Jean, J.R., Jacomy, H., Desforges, M., Vabret, A., Freymuth, F., and Talbot, P.J. (2004). Human respiratory coronavirus OC43: genetic stability and neuroinvasion. J Virol *78*, 8824–8834.

Stapleford, K.A., Rozen-Gagnon, K., Das, P.K., Saul, S., Poirier, E.Z., Blanc, H., Vidalain, P.-O., Merits, A., and Vignuzzi, M. (2015). Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. J Virol *89*, 11233–11244.

Steinhauer, D.A., Domingo, E., and Holland, J.J. (1992). Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene *122*, 281–288.

Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. U.S.a. *90*, 6498–6502.

Stobart, C.C., Rostad, C.A., Ke, Z., Dillard, R.S., Hampton, C.M., Strauss, J.D., Yi, H., Hotard, A.L., Meng, J., Pickles, R.J., et al. (2016). A live RSV vaccine with engineered thermostability is immunogenic in cotton rats despite high attenuation. Nat Commun *7*, 1–12.

Stobart, C.C., Sexton, N.R., Munjal, H., Lu, X., Molland, K.L., Tomar, S., Mesecar, A.D., and Denison, M.R. (2013). Chimeric exchange of coronavirus nsp5 proteases (3CLpro) identifies common and divergent regulatory determinants of protease activity. J Virol *87*, 12611–12618.

Stockman, L.J., Bellamy, R., and Garner, P. (2006). SARS: systematic review of treatment effects. PLoS Med. *3*, e343.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., Liu, W., Bi, Y., and Gao, G.F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends Microbiol. *24*, 490–502.

Subissi, L., Posthuma, C.C., Collet, A., Zevenhoven-Dobbe, J.C., Gorbalenya, A.E., Decroly, E., Snijder, E.J., Canard, B., and Imbert, I. (2014). One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc. Natl. Acad. Sci. U.S.a. *111*, E3900–E3909.

Sui, B., Huang, J., Jha, B.K., Yin, P., Zhou, M., Fu, Z.F., Silverman, R.H., Weiss, S.R., Peng, G., and Zhao, L. (2016). Crystal structure of the mouse hepatitis virus ns2 phosphodiesterase domain that antagonizes RNase L activation. J. Gen. Virol. *97*, 880–886.

Takeda, H., Ueda, Y., Inuzuka, T., Yamashita, Y., Osaki, Y., Nasu, A., Umeda, M., Takemura,

R., Seno, H., Sekine, A., et al. (2017). Evolution of multi-drug resistant HCV clones from pre-existing resistant-associated variants during direct-acting antiviral therapy determined by third-generation sequencing. Sci Rep *7*, 1–13.

Telele, N.F., Kalu, A.W., Gebre-Selassie, S., Fekade, D., Abdurahman, S., Marrone, G., Neogi, U., Tegbaru, B., and nnerborg, A.S.X. (2018). Pretreatment drug resistance in a large countrywide Ethiopian HIV-1C cohort: a comparison of Sanger and high-throughput sequencing. Sci Rep *8*, 1–10.

Theys, K., Feder, A.F., Gelbart, M., Hartl, M., Stern, A., and Pennings, P.S. (2018). Within-patient mutation frequencies reveal fitness costs of CpG dinucleotides and drastic amino acid changes in HIV. PLOS Genet *14*, e1007420–e1007424.

Tomaselli, S., Galeano, F., Locatelli, F., and Gallo, A. (2015). ADARs and the Balance Game between Virus Infection and Innate Immune Cell Response. Curr Issues Mol Biol *17*, 37–51.

Totura, A.L., and Baric, R.S. (2012). SARS coronavirus pathogenesis: host innate immune responses and viral antagonism of interferon. Curr Opin Virol *2*, 264–275.

Traggiai, E., Becker, S., Subbarao, K., Kolesnikova, L., Uematsu, Y., Gismondo, M.R., Murphy, B.R., Rappuoli, R., and Lanzavecchia, A. (2004). An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat. Med. *10*, 871–875.

Tsai, L.-K., Hsieh, S.-T., Chao, C.-C., Chen, Y.-C., Lin, Y.-H., Chang, S.-C., and Chang, Y.-C. (2004). Neuromuscular disorders in severe acute respiratory syndrome. Arch. Neurol. *61*, 1669–1673.

V'kovski, P., Gerber, M., Kelly, J., Pfaender, S., Ebert, N., Braga Lagache, S., Simillion, C., Portmann, J., Stalder, H., Gaschen, V., et al. (2019). Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling. Elife *8*.

Vabret, A., Dina, J., Gouarin, S., Petitjean, J., Corbet, S., and Freymuth, F. (2006a). Detection of the new human coronavirus HKU1: a report of 6 cases. Clin. Infect. Dis. *42*, 634–639.

Vabret, A., Dina, J., Mourez, T., Gouarin, S., Petitjean, J., van der Werf, S., and Freymuth, F. (2006b). Inter- and intra-variant genetic heterogeneity of human coronavirus OC43 strains in France. J. Gen. Virol. *87*, 3349–3353.

Van den Bergh, B., Swings, T., Fauvart, M., and Michiels, J. (2018). Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. Microbiol. Mol. Biol. Rev. *82*, e00008–e00018.

van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J.M., Wolthers, K.C., Wertheim-van Dillen, P.M.E., Kaandorp, J., Spaargaren, J., and Berkhout, B. (2004). Identification of a new human coronavirus. Nat. Med. *10*, 368–373.

van der Hoek, L., Sure, K., Ihorst, G., Stang, A., Pyrc, K., Jebbink, M.F., Petersen, G., Forster,

J., Berkhout, B., and Überla, K. (2005). Croup Is Associated with the Novel Coronavirus NL63. PLoS Med. *2*, e240–e247.

van Doremalen, N., Bushmaker, T., and Munster, V.J. (2013). Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. Euro Surveill. *18*.

van Doremalen, N., Miazgowicz, K.L., Milne-Price, S., Bushmaker, T., Robertson, S., Scott, D., Kinne, J., McLellan, J.S., Zhu, J., and Munster, V.J. (2014). Host species restriction of Middle East respiratory syndrome coronavirus through its receptor, dipeptidyl peptidase 4. J Virol *88*, 9220–9232.

Van Slyke, G.A., Arnold, J.J., Lugo, A.J., Griesemer, S.B., Moustafa, I.M., Kramer, L.D., Cameron, C.E., and Ciota, A.T. (2015). Sequence-Specific Fidelity Alterations Associated with West Nile Virus Attenuation in Mosquitoes. PLoS Pathog *11*, e1005009.

Velthuis, te, A.J.W., van den Worm, S.H.E., and Snijder, E.J. (2012). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res. *40*, 1737–1747.

Velthuis, te, A.J.W., van den Worm, S.H.E., Sims, A.C., Baric, R.S., Snijder, E.J., and van Hemert, M.J. (2010). Zn(2+) inhibits coronavirus and arterivirus RNA polymerase activity in vitro and zinc ionophores block the replication of these viruses in cell culture. PLoS Pathog *6*, e1001176.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature *439*, 344–348.

Vignuzzi, M., Wendt, E., and Andino, R. (2008). Engineering attenuated virus vaccines by controlling replication fidelity. Nat. Med. *14*, 154–161.

Vijgen, L., Keyaerts, E., Lemey, P., Moës, E., Li, S., Vandamme, A.-M., and Van Ranst, M. (2005a). Circulation of genetically distinct contemporary human coronavirus OC43 strains. Virology *337*, 85–92.

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lemey, P., Vandamme, A.-M., and Van Ranst, M. (2005b). Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. J Virol *79*, 1595–1604.

Visher, E., Whitefield, S.E., McCrone, J.T., Fitzsimmons, W., and Lauring, A.S. (2016). The Mutational Robustness of Influenza A Virus. PLoS Pathog *12*, e1005856–25.

Voronin, Y., Holte, S., Overbaugh, J., and Emerman, M. (2009). Genetic drift of HIV populations in culture. PLOS Genet *5*, e1000431.

Wang, L.-F., Shi, Z., Zhang, S., Field, H., Daszak, P., and Eaton, B.T. (2006). Review of bats

and SARS. Emerging Infect. Dis. *12*, 1834–1840.

Wang, L., Shi, W., Chappell, J.D., Joyce, M.G., Zhang, Y., Kanekiyo, M., Becker, M.M., van Doremalen, N., Fischer, R., Wang, N., et al. (2018). Importance of neutralizing monoclonal antibodies targeting multiple antigenic sites on MERS-CoV Spike to avoid neutralization escape. J Virol JVI.02002–17–46.

Wang, L., Shi, W., Joyce, M.G., Modjarrad, K., Zhang, Y., Leung, K., Lees, C.R., Zhou, T., Yassine, H.M., Kanekiyo, M., et al. (2015a). Evaluation of candidate vaccine approaches for MERS-CoV. Nat Commun *6*, 7712.

Wang, M.-N., Zhang, W., Gao, Y.-T., Hu, B., Ge, X.-Y., Yang, X.-L., Zhang, Y.-Z., and Shi, Z.-L. (2016). Longitudinal surveillance of SARS-like coronaviruses in bats by quantitative real-time PCR. Virol. Sin. *31*, 78–80.

Wang, S.-F., Tseng, S.-P., Yen, C.-H., Yang, J.-Y., Tsao, C.-H., Shen, C.-W., Chen, K.-H., Liu, F.-T., Liu, W.-T., Chen, Y.-M.A., et al. (2014). Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins. Biochem. Biophys. Res. Commun. *451*, 208–214.

Wang, W., Lin, X.-D., Guo, W.-P., Zhou, R.-H., Wang, M.-R., Wang, C.-Q., Ge, S., Mei, S.-H., Li, M.-H., Shi, M., et al. (2015b). Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China. Virology *474*, 19–27.

Wargo, A.R., and Kurath, G. (2012). Viral fitness: definitions, measurement, and current insights. Curr Opin Virol *2*, 538–545.

Weeks, S.A., Lee, C.A., Zhao, Y., Smidansky, E.D., August, A., Arnold, J.J., and Cameron, C.E. (2012). A Polymerase mechanism-based strategy for viral attenuation and vaccine development. J. Biol. Chem. *287*, 31618–31622.

Woo, P.C.Y., Lau, S.K.P., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H.L., Poon, R.W.S., Cai, J.J., Luk, W.K., et al. (2004). Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. J Virol *79*, 884–895.

World Health Organization (2003). Acute respiratory syndrome, China. Wkly Epidemiol Rec *78*, 41.

World Health Organization (2009). Case definitions for the 4 diseases requiring notification to WHO in all circumstances under the International Health Regulations (2005). *84*, 52–56.

World Health Organization (2017). Middle East respiratory syndrome Case definition for reporting to WHO, Interim case definition, 26 July 2017.

World Health Organization (2018). Laboratory Testing for Middle East Respiratory Syndrome Coronavirus, Interim guidance, January 2018.

Xiao, S., Li, Y., Wong, T.-W., and Hui, D.S.C. (2017a). Role of fomites in SARS transmission

during the largest hospital outbreak in Hong Kong. PLoS ONE *12*, e0181558.

Xiao, Y., Ma, Q., Restle, T., Shang, W., Svergun, D.I., Ponnusamy, R., Sczakiel, G., and Hilgenfeld, R. (2012). Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. J Virol *86*, 4444–4454.

Xiao, Y., Dolan, P.T., Goldstein, E.F., Li, M., Farkov, M., Brodsky, L., and Andino, R. (2017b). Poliovirus intrahost evolution is required to overcome tissue-specific innate immune responses. Nat Commun *8*, 1–12.

Xiao, Y., Rouzine, I.M., Bianco, S., Acevedo, A., Goldstein, E.F., Farkov, M., Brodsky, L., and Andino, R. (2016). RNA Recombination Enhances Adaptability and Is Required for Virus Spread and Virulence. Cell Host and Microbe *19*, 493–503.

Xie, Q., Cao, Y., Su, J., Wu, J., Wu, X., Wan, C., He, M., Ke, C., Zhang, B., and Zhao, W. (2017). Two deletion variants of Middle East respiratory syndrome coronavirus found in a patient with characteristic symptoms. Arch. Virol. *9*, 1–5.

Xie, X., Wang, H., Zeng, J., Li, C., Zhou, G., Yang, D., and Yu, L. (2014). Foot-and-mouth disease virus low-fidelity polymerase mutants are attenuated. Arch. Virol. *159*, 2641–2650.

Xu, X., Liu, Y., Weiss, S., Arnold, E., Sarafianos, S.G., and Ding, J. (2003). Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. Nucleic Acids Res. *31*, 7117–7130.

Xue, K.S., Hooper, K.A., Ollodart, A.R., Dingens, A.S., and Bloom, J.D. (2016). Cooperation between distinct viral variants promotes growth of H3N2 influenza in cell culture. Elife *5*, e13974.

Yang, L., Wu, Z., Ren, X., Yang, F., Zhang, J., He, G., Dong, J., Sun, L., Zhu, Y., Zhang, S., et al. (2014). MERS-related betacoronavirus in Vespertilio superans bats, China. Emerging Infect. Dis. *20*, 1260–1262.

Yang, Y., Liu, C., Du, L., Jiang, S., Shi, Z., Baric, R.S., and Li, F. (2015). Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. J Virol *89*, 9119–9123.

Yokomori, K., and Lai, M.M. (1991). Mouse hepatitis virus S RNA sequence reveals that nonstructural proteins ns4 and ns5a are not essential for murine coronavirus replication. J Virol *65*, 5605–5608.

Yokomori, K., Banner, L.R., and Lai, M.M.C. (1991). Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses. Virology *183*, 647–657.

Yoshitake, J., Akaike, T., Akuta, T., Tamura, F., Ogura, T., Esumi, H., and Maeda, H. (2004). Nitric oxide as an endogenous mutagen for Sendai virus without antiviral activity. J Virol *78*, 8709–8719.

Yount, B., Denison, M.R., Weiss, S.R., and Baric, R.S. (2002). Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J Virol *76*, 11065–11078.

Yu, I.T.S., Li, Y., Wong, T.-W., Tam, W., Chan, A.T., Lee, J.H.W., Leung, D.Y.C., and Ho, T. (2004). Evidence of Airborne Transmission of the Severe Acute Respiratory Syndrome Virus. N. Engl. J. Med. *350*, 1731–1739.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D.M.E., and Fouchier, R.A.M. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. *367*, 1814–1820.

Zeng, J., Wang, H., Xie, X., Li, C., Zhou, G., Yang, D., and Yu, L. (2014). Ribavirin-resistant variants of foot-and-mouth disease virus: the effect of restricted quasispecies diversity on viral virulence. J Virol *88*, 4008–4020.

Zeng, J., Wang, H., Xie, X., Yang, D., Zhou, G., and Yu, L. (2013). An increased replication fidelity mutant of foot-and-mouth disease virus retains fitness in vitro and virulence in vivo. Antiviral Res. *100*, 1–7.

Zhang, R., Jha, B.K., Ogden, K.M., Dong, B., Zhao, L., Elliott, R., Patton, J.T., Silverman, R.H., and Weiss, S.R. (2013). Homologous 2“,5-”phosphodiesterases from disparate RNA viruses antagonize antiviral innate immunity. Proc. Natl. Acad. Sci. U.S.a. *110*, 13114–13119.

Zhao, J., Li, K., Wohlford-Lenane, C., Agnihothram, S.S., Fett, C., Zhao, J., Gale, M.J., Baric, R.S., Enjuanes, L., Gallagher, T., et al. (2014). Rapid generation of a mouse model for Middle East respiratory syndrome. Proc. Natl. Acad. Sci. U.S.a. *111*, 4970–4975.

Zhao, L., Jha, B.K., Wu, A., Elliott, R., Ziebuhr, J., Gorbalenya, A.E., Silverman, R.H., and Weiss, S.R. (2012). Antagonism of the Interferon-Induced OAS-RNase L Pathway by Murine Coronavirus ns2 Protein Is Required for Virus Replication and Liver Pathology. Cell Host and Microbe *11*, 607–616.

Zhao, L., Rose, K.M., Elliott, R., Van Rooijen, N., and Weiss, S.R. (2011). Cell-type-specific type I interferon antagonism influences organ tropism of murine coronavirus. J Virol *85*, 10058–10068.

Zhou, S., Jones, C., Mieczkowski, P., and Swanstrom, R. (2015). Primer ID Validates Template Sampling Depth and Greatly Reduces the Error Rate of Next-Generation Sequencing of HIV-1 Genomic RNA Populations. J Virol *89*, 8540–8555.

Zhu, Z., Chakraborti, S., He, Y., Roberts, A., Sheahan, T., Xiao, X., Hensley, L.E., Prabakaran, P., Rockx, B., Sidorov, I.A., et al. (2007). Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. Proc. Natl. Acad. Sci. U.S.a. *104*, 12123–12128.

Zumla, A., Azhar, E.I., Arabi, Y., Alotaibi, B., Rao, M., McCloskey, B., Petersen, E., and Maeurer, M. (2015). Host-directed therapies for improving poor treatment outcomes associated with the middle east respiratory syndrome coronavirus infections. Int. J. Infect. Dis. *40*, 71–74.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. *29*, 1017–1026.