

# Evaluation of a recombination-resistant coronavirus as a broadly applicable, rapidly implementable vaccine platform

Rachel L. Graham <sup>1</sup>, Damon J. Deming<sup>1,2,3</sup>, Meagan E. Deming<sup>2,4</sup>, Boyd L. Yount<sup>1</sup> & Ralph S. Baric<sup>1,2</sup>

Emerging and re-emerging zoonotic viral diseases are major threats to global health, economic stability, and national security. Vaccines are key for reducing coronaviral disease burden; however, the utility of live-attenuated vaccines is limited by risks of reversion or repair. Because of their history of emergence events due to their prevalence in zoonotic pools, designing live-attenuated coronavirus vaccines that can be rapidly and broadly implemented is essential for outbreak preparedness. Here, we show that coronaviruses with completely rewired transcription regulatory networks (TRNs) are effective vaccines against SARS-CoV. The TRN-rewired viruses are attenuated and protect against lethal SARS-CoV challenge. While a 3-nt rewired TRN reverts via second-site mutation upon serial passage, a 7-nt rewired TRN is more stable, suggesting that a more extensively rewired TRN might be essential for avoiding growth selection. In summary, rewiring the TRN is a feasible strategy for limiting reversion in an effective live-attenuated coronavirus vaccine candidate that is potentially portable across the Nidovirales order.

---

<sup>1</sup>Department of Epidemiology, The University of North Carolina at Chapel Hill, 2107 McGavran-Greenberg, CB 7435, Chapel Hill, NC 27599, USA.

<sup>2</sup>Department of Microbiology and Immunology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>3</sup>Present address: Food and Drug Administration, 10933 New Hampshire Avenue, Bldg 22, Rm 6170, Silver Spring, MD 20993, USA. <sup>4</sup>Present address: University of Maryland Medical Center, Department of Medicine, Division of Infectious Disease, Institute of Human Virology, 725 West Lombard Street, Room 211A, Baltimore, MD 21201, USA. Correspondence and requests for materials should be addressed to R.S.B. (email: [rbaric@ad.unc.edu](mailto:rbaric@ad.unc.edu))

Emerging and re-emerging zoonotic viral diseases are major threats to global human health, economic stability, and national security<sup>1-6</sup>. The incidence of human zoonotic disease is estimated to surpass 1 billion cases per year, with novel emerging infectious diseases accruing hundreds of billions of dollars in economic losses<sup>7,8</sup>, losses that are greatly magnified when new emerging viruses such as coronaviruses (CoVs) devastate economically critical livestock populations across the globe. With the continued encroachment of human populations into animal habitats and our close contact with domesticated animals, zoonoses will continue to increase as the human and livestock population numbers and density expand over the next century. In fact, a recent study recognized that the majority of emerging infectious disease events have origins in wildlife<sup>3,8</sup>, underscoring the importance of developing broadly applicable strategies for vaccine design for virus families that are harbored within extensive zoonotic pools.

Vaccines are well established in their capacity to reduce viral disease burden. Live-attenuated vaccines, because they can elicit balanced innate and adaptive—and often lifelong—protective immune responses, including lactogenic immunity, are ideal candidates for vaccine development in humans and animals<sup>1</sup>. However, their utility as broadly applicable vaccine platforms has long been limited by risks of reversion of attenuated vaccine strains to virulence, largely because the stability of the attenuation cannot be clearly evaluated or assured.

The emergence of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle-East Respiratory Syndrome Coronavirus (MERS-CoV) in the 21st century emphasizes the threat of pandemic viral infections originating from cross-species transmission events<sup>9-11</sup>. These highly pathogenic variants are prime models for the development of broad-based strategies for evaluating live virus vaccines for the Nidovirales order. Coronaviruses (CoVs) all reproduce with conserved replication strategies, emphasizing the strength and rapidly adaptable potential of a vaccine design platform that takes advantage of this biology. CoVs replicate and transcribe subgenomic RNAs (sgRNAs) via a discontinuous transcription mechanism mediated by transcription regulatory sequences (TRSs), a series of conserved nucleotide sequences positioned near the 5′-end of the genome and at several locations immediately 5′ of each downstream open reading frame (ORF) (Fig. 1). These TRSs regulate a transcription attenuation program via base-pairing interactions between the leader TRS and body TRSs that results in the production of sgRNAs, from which downstream ORFs are translated. Within the TRS, a 6- to 8-nt core sequence (ACGAAC for SARS-CoV) guides base-pairing and duplex formation between nascent RNA and the leader TRS<sup>12-14</sup>. While this collective TRS arrangement, the transcription regulatory network (TRN), is conserved, our laboratory has shown that it is possible to rewire the guide sequence of the SARS-CoV TRN and produce infectious virus<sup>13</sup>. We also showed that recombination between this rewired TRN virus and wild-type (WT) SARS-CoV was not

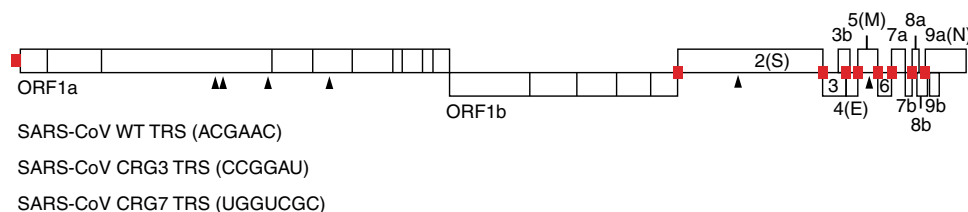
viable, indicating that recombination-mediated reversion of a CoV vaccine platform featuring a rewired TRN is highly unlikely.

Based on conservation of the TRN biology across CoVs, this report further explores the feasibility of the development of a stably attenuated vaccine platform featuring a completely rewired TRN as a candidate strategy for a broadly applicable, rapidly implementable CoV vaccine platform that is highly resistant to recombination repair and stably attenuated in both young and highly vulnerable mouse models of human disease.

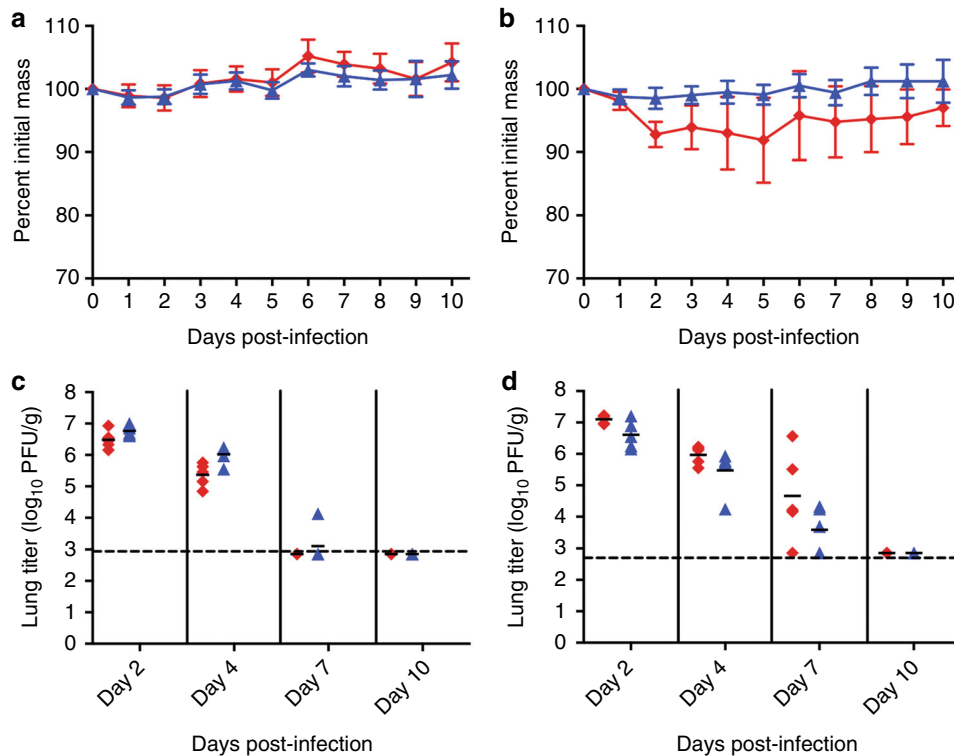
## Results

**The 3-nt TRN mutant is attenuated for virulence.** In a previous study in our laboratory, we demonstrated that the SARS-CoV TRN could be reprogrammed, provided the individual TRSs were replaced with matching sequences<sup>13</sup>. The rewired TRN replaced the conserved 6-nt TRS with a 6-nt cassette that is not used in any other characterized CoVs, encoding a net change of 3 nts (ACGAAC to CCGGAU). Our previous work showed that this rewired TRN was refractory to recombination with WT genomes<sup>13</sup>. Therefore, we tested its replication and pathogenesis in young and aged BALB/c mice. Consistent with earlier reports, WT SARS-CoV only caused weight loss in aged animals. In contrast, CRG3 replicated but caused no weight loss in young (10-week-old) mice (Fig. 2a) and minimal weight loss in aged (12-month-old) BALB/c mice (Fig. 2b) (young mice:  $P = 0.25$  for titer, Wilcoxon test,  $P = 0.37$  for weight loss, Mann-Whitney test; old mice:  $P = 0.25$  for titer,  $P < 0.001$  for weight loss). Both WT and recombinant viruses replicated to high titers that were detectable on days 2 and 4 post-infection (p.i.) in both young and aged animals, with titers beginning to clear by day 7 p.i. (Fig. 2c, d), as is usually observed in mice infected with SARS-CoV<sup>1</sup>.

**Vaccination with CRG3 protects mice against challenge.** The high replication titer and low or absent virulence in young and aged mouse models paired with the inherently recombination-refractory genome suggested that the CRG3 virus would be an ideal vaccine candidate. To test CRG3's efficacy in protecting against homologous and heterologous challenge, the virus was administered to mice in a single-dose vaccination, alongside viral replicon particles (VRPs) expressing the viral Spike attachment protein (VRP-S) as a control<sup>15</sup>. On day 22 post-vaccination, mice were then challenged with either mouse-adapted SARS-CoV (MA15—homologous challenge) or SARS-CoV expressing the Spike gene from the Himalayan palm civet (*Paguma larvata*) strain HCSZ6103 (heterologous challenge). Mice were then observed for morbidity and mortality, and surviving animals were euthanized on day 4 p.i. Upon homologous challenge, CRG3 vaccination was protective against weight loss and mortality in young and aged mice, with no detectable viral titer in the lungs at 4 days p.i., in contrast to PBS and VRP-S subcutaneous vaccination (Fig. 3a-d, Supplementary Fig. 1A-B). Importantly, against heterologous challenge, while both CRG3 and VRP-S vaccines



**Fig. 1** Schematic of the rewired TRN SARS-CoV mutants. The SARS-CoV genome is depicted, with open reading frames (ORFs) indicated. The locations of mouse-adapted mutations are denoted by black triangles. The location of each characterized TRS is denoted by a red box. The specific TRS core sequences are listed underneath the genome



**Fig. 2** In vivo characterization of the SARS-CoV CRG3 mutant. **a, b** Weight loss is depicted as percent initial mass in **(a)** young and **(b)** aged mice. **c, d** Viral replication in the lungs is depicted as titer ( $\log_{10}$  PFU/g) in **(c)** young and **(d)** aged BALB/c mice. Viruses: WT, red diamonds; CRG3, blue triangles. Dashed lines: limits of detection. Error bars depict standard deviation

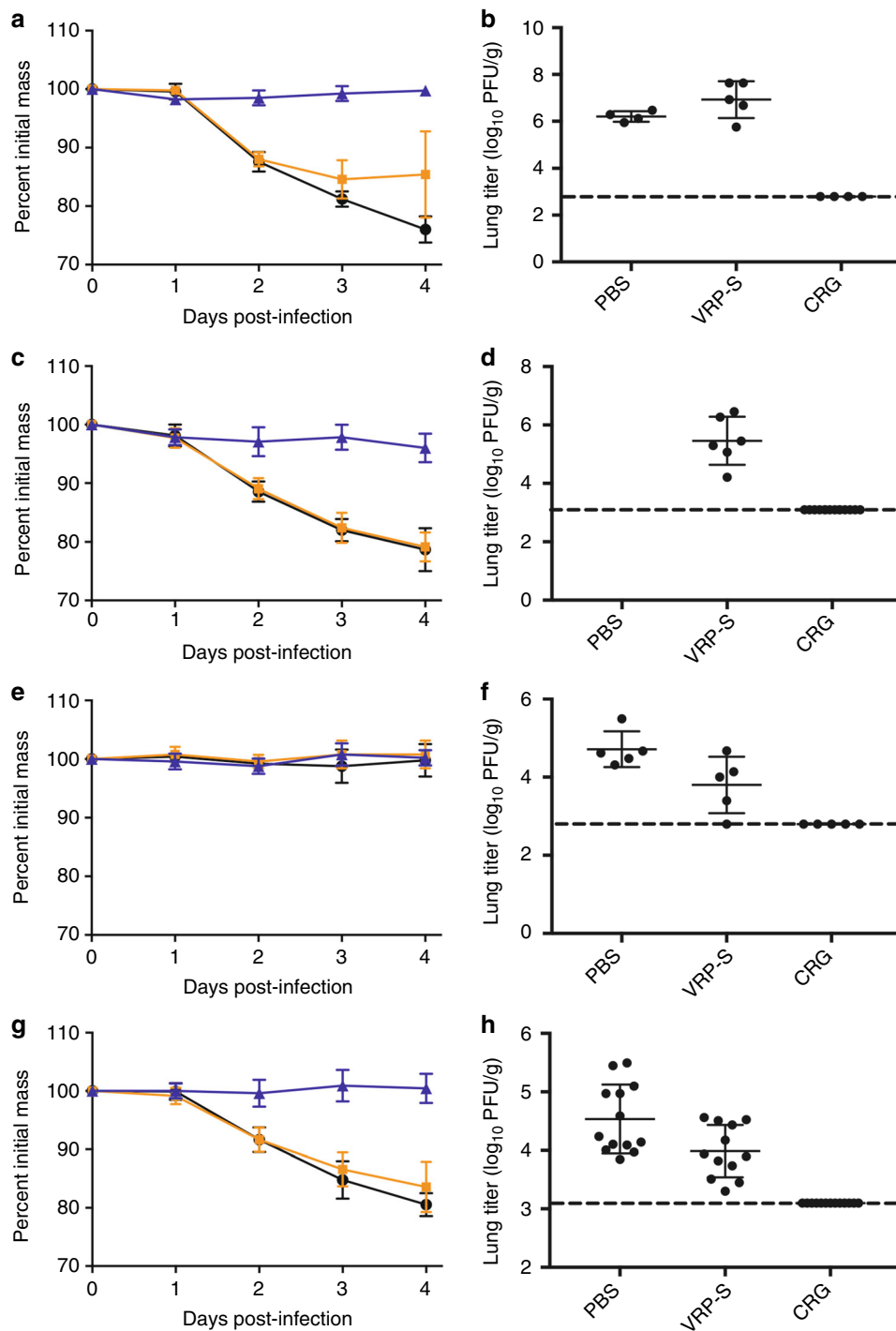
were protective in young mice, only CRG3 was protective in aged mice against replication and mortality, with no detectable titer in the lungs at 4 days p.i. (Fig. 3e–h, Supplementary Fig. 1C–D). CRG vaccination induced SARS-CoV neutralizing antibody titers approaching  $4 \log_{10}$  in both young and aged mice, nearly 2-fold higher than the titers induced by VRP-S vaccination ( $P < 0.0001$ , two-way ANOVA) (Supplementary Fig. 1E). Moreover, CRG vaccination largely protected aged mice from the extensive inflammatory cell infiltration, tissue damage, perivascular cuffing, edema, and septal thickening observed in PBS- and VRP-S-vaccinated mice (Supplementary Fig. 2).

**CRG3 reverts with pandemic-associated mutations.** A live-attenuated vaccine candidate should demonstrate phenotypic stability in infected host populations. Therefore, to test its resistance to reversion to virulence, CRG3 was subjected to five independent serial passages in parallel with WT SARS-CoV in aged (14-month-old) BALB/c mice. Over the course of six 4-day passages, all CRG3 passaged viruses acquired a virulent phenotype, with weight losses and percent survival curves mirroring the kinetic rates of the emergence of virulence seen with WT passaged viruses (Supplementary Fig. 3). To attempt to identify the genotypic causes of this phenotypic reversion to virulence, the genomes of 5 independent plaque isolates of the passage 6 viruses (WT P6 and CRG3 P6) were submitted to Sanger sequencing. Surprisingly, and contrary to expectations, no mutations in either the Spike or the Membrane protein—both of which were targets for mutational selection in 3 separate adaptations of SARS-CoV Urbani to mice (Fig. 4a)—were identified in WT or CRG3 revertants. Instead, 4 of 5 CRG3 P6 isolates showed evidence of often large deletions in the accessory ORFs 7b, 8a, and 8b (Fig. 4a, b). These deletions, often in-frame, were reminiscent of several incidences of host range-associated deletions identified in human isolates from the 2003 SARS-CoV epidemic, including a

29-nt deletion in ORF8 relative to the Himalayan palm civet strain<sup>16</sup> and 82-nt<sup>17</sup> and 386-nt<sup>18</sup> deletions in this same region. In contrast, deletions in accessory ORFs were rare in virulent WT revertants. Rather, scattered nonsynonymous mutations were identified in ORF1a, ORF3a, ORF8a, and ORF9. Notably, only 1 WT revertant exhibited a small deletion in ORF7b. Collectively, these findings demonstrated that this region of the SARS-CoV genome is inherently unstable when subjected to replication pressure, especially in aged animals, and indicated that a TRN-rewired virus would require additional stabilizing mutations to be feasible as a vaccine candidate.

#### **An attenuated SARS-CoV mutant with a 7-nt TRN replacement.**

The demonstration of second-site reversion to virulence upon passage of CRG3 suggested that a more ideal TRN-rewired vaccine candidate would be one that would be less likely to phenotypically revert in vivo. CoV TRS networks are finely tuned and regulate the expression of both highly abundant and low-frequency mRNA transcripts; thus, small changes in the TRS might subtly alter the regulation of subgenomic transcripts<sup>13</sup>. Moreover, the exact regulatory milieu around each TRS is uncertain beyond the recognition that both, up- and downstream sequences can influence mRNA expression efficiency, perhaps in a TRN sequence-specific manner<sup>19,20</sup>. We hypothesized that more extensive remodeling of the TRS network may better disrupt the finely tuned network of abundant and low-frequency transcripts, altering the natural regulation of global gene expression, and thereby leading to decreased virulence. In addition, it was possible that the virulence observed in CRG3 P6 isolates was due, at least in part, to mouse adaptation during passage. Therefore, to evaluate the effects of more extensive TRN rewiring in a lethal pathogenic model and to further stabilize the TRN against recombination repair and mutation selection, a TRN mutant was constructed that featured the following: (1) the set of

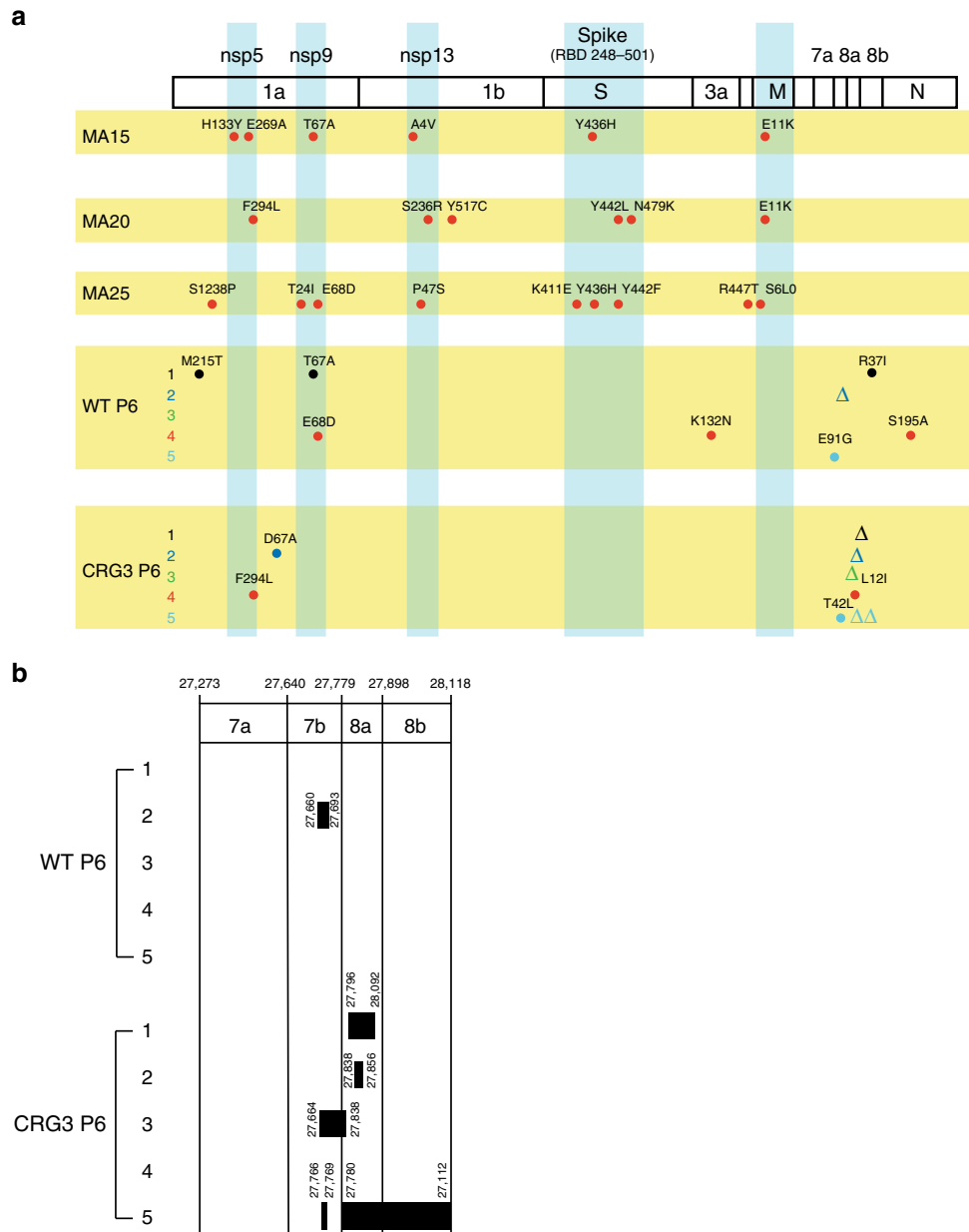


**Fig. 3** Vaccination of young and aged mice with SARS-CoV CRG3. (**a, c, e, g**) Weight loss upon homologous (**a, c**) and heterologous (**e, g**) challenge is depicted as percent initial mass in young (**a, e**) and aged (**c, g**) vaccinated BALB/c mice. **b, d, f, h** Viral replication in the lungs upon homologous (**b, d**) and heterologous (**f, h**) challenge is depicted as titer ( $\log_{10}$  PFU/g) in (**b, f**) young and (**d, h**) aged vaccinated BALB/c mice. Vaccinations: black circles, PBS vaccination; orange squares, VRP-S vaccination; blue triangles, CRG3 vaccination. Dashed lines: limits of detection. Error bars depict standard deviation

6 mouse-adapted mutations present in the SARS-MA15 virus (see Fig. 4a); and (2) a newly designed TRN consisting of a 7-nt complete replacement of the 9 SARS-CoV TRS loci (ACGAAC to UGGUCGC – CRG7-MA – Fig. 1). These replacements yielded infectious CRG7-MA virus that was capable of replicating to titers equivalent to those of the mouse-adapted WT background virus, SARS-MA15, though with a pinpoint plaque phenotype that remained stable after tissue culture passage. CRG3-MA (9 CRG3 replacements in the virulent SARS-MA15 backbone) was also

generated as a control TRN-rewired virus. northern blot analysis of CRG7-MA versus CRG3 and SARS-MA15 revealed that the virus produced the expected bands at similar proportions to WT virus, although some additional low-abundance transcripts were also noted in CRG7-MA (Supplementary Fig. 4).

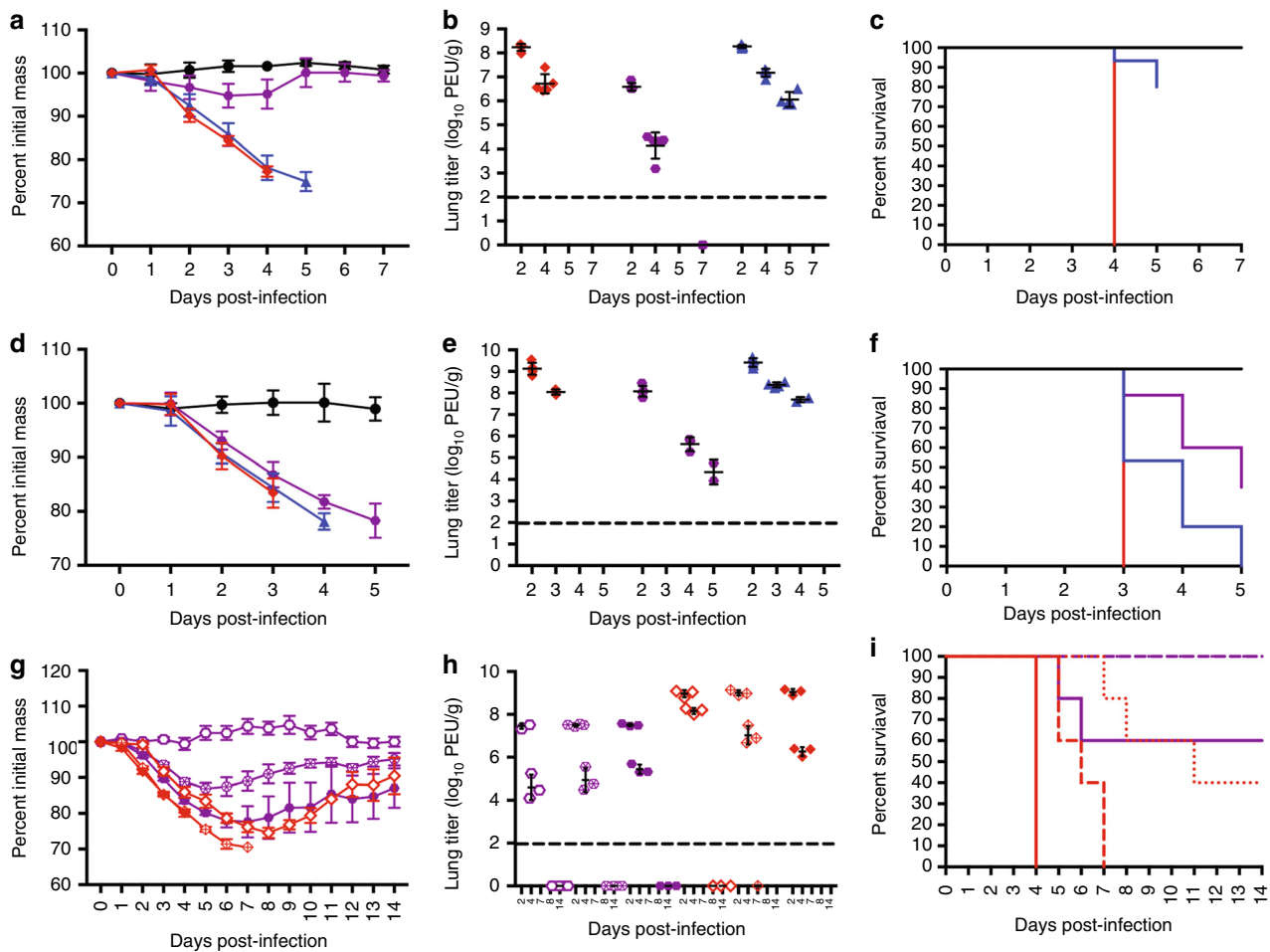
These viruses were then evaluated for replication and virulence in young (10-week-old) and aged (12-month-old) BALB/c mice in comparison with SARS-MA15. In young mice, CRG7-MA was attenuated compared with CRG3-MA and SARS-MA15, causing



**Fig. 4** Accumulation of second-site mutations upon passage of SARS-CoV CRG3. **a** The SARS-CoV genome is depicted at the top. The mutations identified in the 3 published mouse-adapted SARS-CoV genomes (MA15, MA20, and MA25) are shown in the top 3 yellow sections immediately below the genome. The mutations identified via Sanger sequencing in 5 individual plaque isolates of WT SARS-CoV and CRG3 after 6 passages (WT P6 and CRG3 P6, respectively) are depicted in the bottom 2 yellow sections. Nonsynonymous mutations that led to amino acid substitutions are identified by their WT amino acid residue, the residue number in the mature protein, and the substituted residue. **b** The genome bounds of the SARS-CoV ORF 7a, 7b, 8a, and 8b accessory proteins are denoted. Individual rows indicate each of 5 plaque isolates from WT SARS-CoV after 6 passages (WT P6) and CRG3 after 6 passages (CRG3 P6). Deletions are indicated with black boxes in each row, with the bounds of the deletions specified

a weight loss of ~5% ( $P < 0.0001$ ,  $t$  test, on days 2–5 p.i. for CRG7-MA vs. CRG3-MA), though replicating to high titers (approximately  $7 \log_{10}$  PFU for CRG7-MA vs.  $8 \log_{10}$  PFU for CRG3-MA and SARS-MA15) at 2 days p.i. Further, CRG7-MA showed evidence of more rapid clearance kinetics, with a nearly  $3\text{-}\log_{10}$  difference in titer compared with the CRG3-MA and SARS-MA15 controls at 4 days p.i. (Fig. 5a–c) ( $P < 0.001$ ,  $t$  test, for both 2 and 4 days p.i. for CRG3-MA vs CRG7-MA). While morbidity and mortality were more pronounced in old mice, CRG7-MA was still attenuated compared with CRG3-MA and SARS-MA15 in terms of weight loss (less weight loss on days 2–4 p.i.,  $P < 0.05$  on days 2 and 4 p.i.,  $t$  test, CRG3-MA vs. CRG7-MA)

and lung titer ( $8 \log_{10}$  PFU for CRG7-MA vs.  $>9 \log_{10}$  PFU for CRG3-MA and SARS-MA15 on day 2 p.i. ( $P < 0.0001$ ,  $t$  test, CRG3-MA vs. CRG7-MA);  $\sim 2.5\text{-}\log_{10}$  PFU difference in titer on day 4 p.i. ( $P < 0.05$ ,  $t$  test, CRG3-MA vs. CRG7-MA) (Fig. 5d–f). Infection of aged BALB/c mice with  $2 \log_{10}$ ,  $3 \log_{10}$ , or  $4 \log_{10}$  PFU of CRG7-MA or SARS-MA15 emphasized the attenuated virulence phenotype, with the  $2\text{-}\log_{10}$  PFU infection of CRG7-MA causing almost no weight loss over the course of the infection, despite replicating to the same titers as the  $3 \log_{10}$  and  $4 \log_{10}$  PFU infections (Fig. 5g–h). Under identical conditions, SARS-MA15 mice lost more body weight ( $>20\%$ ) and experienced 60% mortality rates. Survival was markedly different



**Fig. 5** In vivo characterization of the SARS-CoV CRG7-MA mutant. **a, d** Weight loss is depicted as percent initial mass in **(a)** young and **(d)** aged BALB/c mice. **b, e** Viral replication in the lungs is depicted as titer ( $\log_{10}$  PFU/g) in **(b)** young and **(e)** aged BALB/c mice. Black circles: PBS; red diamonds: SARS-MA; blue triangles: CRG3-MA; purple circles: CRG7-MA. **c, f** Mortality is depicted as percent survival in **(c)** young and **(f)** aged mice. Black: PBS; red: SARS-MA; blue: CRG3-MA; purple: CRG7-MA. **g-i** Dose-dependent pathogenicity of SARS-CoV CRG7-MA. Aged BALB/c mice were infected with SARS-CoV CRG7-MA and were evaluated for **(g)** weight loss, depicted as percent initial mass, **(h)** viral replication in the lungs, depicted as titer ( $\log_{10}$  PFU/g), and **(i)** mortality, depicted as percent survival. Red: SARS-MA; purple: CRG7-MA; unfilled symbols:  $2\text{-log}_{10}$  infection; starred symbols:  $3\text{-log}_{10}$  infection; solid symbols:  $4\text{-log}_{10}$  infection. Dashed lines: limits of detection. Error bars depict standard deviation (**a, b, d, e, h**) or standard error of the mean (**g**)

between CRG7-MA and SARS-MA15 infections, with 100% of mice infected with  $2\log_{10}$  or  $3\log_{10}$  PFU and 60% of mice infected with  $4\log_{10}$  PFU surviving the course of CRG7-MA infection. In contrast, only 40% of mice infected with  $2\log_{10}$  PFU of SARS-MA15 survived infection; mice infected with  $3\log_{10}$  or  $4\log_{10}$  PFU survived only until days 6 or 3 p.i. maximum, respectively ( $P < 0.0001$ , log-rank test) (Fig. 5i). When CRG7-MA was serially passaged in young (10-week-old) BALB/c mice, in contrast to the rapid phenotypic reversion observed with CRG3, CRG7-MA remained phenotypically stable with passage, with mice losing a minimal ( $\sim 5\%$ ) amount of their starting weight over the 3-day infections throughout all 4 passages (Supplementary Fig. 5A). Viral titer also remained stable over passage, with  $>5\log_{10}$  PFU detectable at 3 days p.i. in each passage (Supplementary Fig. 5B). Furthermore, when CRG7-MA was serially passaged independently three times for six passages in aged mice, replicating the passage conditions in which CRG3 reverted to virulence, CRG7-MA did not exhibit an increase in virulence, causing  $<20\%$  weight loss in infected animals and  $\sim 60\%$  mortality upon infection of 12-month-old BALB/c mice with  $10^5$  PFU of CRG7-MA post-passage, essentially replicating the mortality shown in aged mice infected with non-passaged CRG7-MA virus and distinct from the 100% mortality caused by wild-type

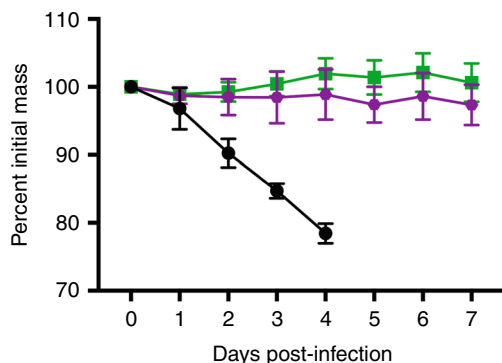
SARS-MA15 infection of aged mice (see Fig. 5i, Supplementary Fig. 5C-D, and Supplementary Table 1).

Finally, CRG7-MA was evaluated for its capacity to protect against lethal SARS-MA15 challenge in aged (12-month-old) mice. Mice were vaccinated with  $2.5\log_{10}$  PFU of either CRG7-MA or ExoN-MA, which we previously demonstrated to be an effective vaccine in the SARS-MA15 backbone in the aged BALB/c mouse model of SARS-CoV pathogenesis<sup>1</sup>. On day 22 post-vaccination, mice were then challenged with a lethal dose ( $5\log_{10}$  PFU) of SARS-MA15 and evaluated for weight loss. Mice lost only minimal amounts ( $\sim 5\%$ ) of their starting weights with both vaccines ( $P = 0.125$  for both viruses versus PBS vaccination, Wilcoxon test), indicating that CRG7-MA and ExoN-MA were equally protective against lethal SARS-MA15 challenge (Fig. 6).

## Discussion

Live-attenuated vaccines remain key players in reducing the global disease burden associated with viral infections in humans, critically important livestock, and companion animals. Historically and contemporarily, live-attenuated vaccines have been used with success to help control measles, mumps, rubella, polio, yellow fever, and chickenpox infections and outbreaks<sup>1,21,22</sup>.





**Fig. 6** Vaccination of aged mice with SARS-CoV CRG7-MA. Weight loss (depicted as percent initial mass) is shown for aged BALB/c mice vaccinated with PBS, ExoN-MA (control attenuated virus known to protect in an aged model) and CRG7-MA and then challenged with a lethal dose of mouse-adapted SARS-CoV. Vaccinations: PBS, black; ExoN-MA, green; CRG7-MA, purple. Error bars depict standard deviation

However, live-attenuated vaccines are also associated with the risk of reversion by either mutation- or recombination-driven processes, which can cause dangerous outbreaks in unvaccinated populations, including animals<sup>22</sup>. For example, highly pathogenic porcine epidemic diarrhea virus (PEDV) strains emerged in China in 2012, circumventing existing vaccines, and RNA recombination events between wild-type and live-attenuated PEDV and between avian infectious bronchitis virus (IBV) strains have seeded new outbreaks<sup>23–26</sup>. Therefore, measures are needed to stabilize live-attenuated vaccines against reversion under selective pressure, particularly for viruses like CoVs, which employ recombination as a standard feature of their replication cycle, as incidental recombination events in the context of a co-infection could unintentionally introduce alleles with enhanced virulence into an attenuated vaccine genome, with consequences that could be difficult to predict<sup>13,27,28</sup>. Several groups, including our own, have developed novel strategies in fidelity regulation and control that attenuate RNA virus pathogenesis, and in the case of CoVs, the development of mutants that prevent reversion repair to virulence<sup>1,29</sup>.

Recombination repair is a well-characterized process essential for genome evolution in many biological systems and plays critical roles in spread, virulence, and pathogenesis<sup>30</sup>. In general, strategies to engineer recombination-resistant RNA viruses have been limited to CoVs and, to a lesser extent, enteroviruses<sup>13,31</sup>. In the context of a CoV infection, recombination occurs when the viral RNA-dependent RNA polymerase (RdRp) switches templates during nascent RNA synthesis, using the nascent RNA itself as a primer once the RdRp reassociates with the template genome. The likelihood of a recombination event occurring is regulated by several viral factors, including replication rates, RNA secondary structure, genome size, and the nature of the replicase/transcriptase protein complex; however, the heritability of a recombination, once it has occurred, is mediated by the replication fitness of the resulting progeny genome<sup>13,32–35</sup>. Several reports have described the emergence of CoVs with enhanced virulence that, upon genome analysis, clearly originated from recombination events of related viruses in avian and mammalian hosts, including SARS-CoV<sup>27,28,36–38</sup>. Furthermore, reservoir species, such as bats, from which the prevailing evidence suggests that both potentially lethal human CoVs, SARS-CoV and MERS-CoV, emerged, have been shown to harbor multiple CoV species, and recombinant RNA genomes have been frequently identified within colonies and among individual infected animals,

increasing the likelihood of recombination-driven alterations in species specificity and virulence<sup>16,27,39,40</sup>.

Thus, engineering elements to render a genome recombination-refractory is an essential step towards ensuring that a live-attenuated vaccine candidate cannot regain virulence during an incidental co-infection with another TRN-compatible CoV genome. SARS-CoV is a highly pathogenic pneumo-enteric pathogen that captures many disease features seen among other CoVs. The TRN conserved sequence (CS) motifs utilized in the CRG3 and CRG7 backgrounds, CCGGAU and UGGUCG, respectively, are both unique sequence motifs when compared with all known CoV genome TRN sequences, greatly reducing and likely eliminating the possibility of recombination with unmodified genomes at the canonical TRS loci. Our previous work demonstrated that introducing mismatched TRSs was lethal for RNA recombinant virus replication<sup>13</sup>. Furthermore, here, we showed that CRG3 TRN replacement did not revert at the primary sites of mutation, indicating that the rewired TRN is stable and not under sufficient selective pressure to revert, even *in vivo*. As coordinated interactions are required for TRN function, TRN reversion is unlikely, given the requirement for nearly simultaneous reversions at multiple sites across the genome. Larger-sized TRS CSs (>7 nts) within the TRN may not prove effective, as evolution appears to have selected for a CoV RNA polymerase that is heavily focused on recognizing a 5- to 7-nt TRS CS within the TRN to regulate subgenomic transcription.

The attenuation resulting from TRN rewiring is most likely attributed to alterations in the viral transcription profile, low-abundance transcripts that either encode or reduce the expression of previously unidentified out-of-frame ORFs, or viral or host factor interactions with the viral genome. The rewired TRN produces obvious novel viral RNA species (Supplementary Fig. 4). These novel RNA species may serve as functional mRNAs, producing noncanonical viral protein products that attenuate viral replication or pathogenesis. Alternatively, these novel RNA species may compete with canonical viral RNA species for replication and transcription, attenuating viral replication and/or pathogenesis due to the altered availability of transcripts encoding bona fide viral virulence factors. Furthermore, the rewired TRN itself may serve as an attenuating factor, as the virus' discontinuous transcription program alters programmed RNA-protein interactions (involving both viral and host proteins) either directly (i.e., by altering the bases required for RNA-protein interactions) or indirectly (i.e., through changes in the RNA's secondary structure that affect the steric availability of RNA-protein interaction sites). Moreover, Di et al., in an arterivirus model, used next-generation sequencing analysis to show that noncanonical transcripts are produced in the course of wild-type infection, indicating that the coding capacity of nidoviruses is actually much larger than what has been characterized using Sanger sequencing and biochemical detection methods. Their findings suggest that attenuation via the TRN may be able to target noncanonical RNA species, which might be able to impact pathogenesis with fewer effects on replication and structural protein production<sup>41</sup>. These fascinating possibilities will be the focus of future studies on the mechanism of TRN-related attenuation.

The hypothesis that alterations in the expression profiles of canonical viral RNA species and protein products attenuate pathogenesis is further strengthened when paired with observations of the types of mutations that were selected upon passage of CRG3 in aged mice: mutation profiles were different if selection occurred in young versus aged populations. Frieman et al. previously demonstrated that mutations in Spike and nonstructural protein 9 (nsp9) were repetitively selected in several independent passages and that these mutations conferred virulence in young

mice, suggesting that Spike–viral receptor interactions and nsp9–replicase protein interactions were most important for virulence<sup>42</sup>. In severely ill SARS-CoV-infected humans (predominantly over 50 years of age), deletions in and around ORF8 were identified, as discussed earlier<sup>16–18,39</sup>. Deletions of different accessory ORFs were the most frequently observed changes after selection in aged animals. Furthermore, in aged mice, nearly any combination of 2 SARS-MA15-identified alleles (shown in Fig. 4a) also conferred increased virulence, most likely reflecting the increased susceptibility of aged mice to lethal outcomes. These data suggest that CoV adaptation to virulence is different in young and aged animals, especially when coupled with variations in virus pathogenic determinants. While immune senescence can enhance virus virulence<sup>43</sup>, our data support the novel hypothesis that virulent zoonotic coronaviruses may emerge more quickly after *in vivo* passage, especially in the aged, where multiple evolutionary pathways exist that can program virus virulence.

The disproportionate identification of mutations in the accessory ORFs following passage of the CRG3 TRN virus in aged animals may reflect the roles these accessory ORFs are hypothesized to play in the modulation of host immunity<sup>44</sup>. Aged animals' immune systems respond differently, and usually more severely, to microbial challenge, with the immune response skewed towards severe innate immune effects and defects in adaptive immunity<sup>45</sup>. However, in the context of CoV infection, this augmented immune response, when presented in the context of accessory ORF deletion, may serve as a virulence fulcrum, with the balance poised between essential immune recognition and severe disease, possibly due to defects in spread and persistence. These potential effects on host immunity argue that a stable vaccine candidate should consider the potential for changes in accessory ORFs.

A well-designed vaccine candidate should include genetic traps that are either independently attenuating or triggered by recombination events. We have previously demonstrated both stable attenuation and protection against lethal challenge with inactivations of the nsp14 exonuclease<sup>1</sup> and the nsp16 2'-O-methyltransferase<sup>46</sup> activities. In addition, Züst et al. demonstrated that a partial deletion of the murine hepatitis virus (MHV) nsp1 replicase protein could protect against homologous and heterologous challenge<sup>47</sup>. Combining alleles that render the virus recombination-refractory, alter RNA replication fidelity, and that result in an altered host immune response could produce stable, reversion-proof, live-attenuated viruses that induce robust neutralizing immunity. Attenuating alleles, coupled with a rewired TRN, are anticipated to increase the stability of the attenuation and to minimize the chances for RNA recombination repair. The strategies reported herein, coupled with the availability of new molecular clones for CoVs that cause severe disease in livestock populations, provide a vehicle for improved live virus vaccine design<sup>48–53</sup>. It may also be possible to “tune” the CoV TRN, with attenuating changes concentrated on virulence alleles, such as accessory genes known to impact pathogenesis. Such alterations have the potential to greater enhance the stability of a vaccine candidate, particularly if considering TRS context outside of the 5- to 7-nt core sequence. Such studies would benefit the most if the RNA structure is also considered, as TRS accessibility is most likely modulated by structural elements in the CoV RNA genome. Such studies will be the focus of future work.

With the continuing identification of zoonotic pools of CoVs that genetically resemble lethal human and animal CoVs, often with only a few percentage points of difference between the zoonotic and lethal human sequences<sup>10,54–57</sup>, the necessity for a rapidly implementable, universal attenuation platform for CoV live-attenuated vaccine design is underscored. In this report, we described the design and implementation of a CoV attenuation

strategy that can be easily and rapidly adapted to any CoV genome. The presence of 8–9 characterized TRSs within any CoV genome, the CSs of which are 6–7 nts each, offers too large and complex a target for primary site reversion to be a likely event. As most single recombination events would decouple TRN expression networks, these recombinants would be lethal. Therefore, this attenuation strategy, when paired with alleles that can resist selection events that lead to second-site reversion, could bring live-attenuated CoV vaccines within the reach of realization in the face of the ever-growing threat of new human and animal CoV-based epidemics.

## Methods

**Viruses and cells.** All virus stocks were propagated in Vero-E6 cells as described in<sup>58</sup>. All virus work was performed in a biological safety cabinet in a biosafety level 3 laboratory.

**Construction of SARS plasmids and viruses.** TRN3-based plasmids were constructed in<sup>13</sup>. TRN7 mutations were introduced into SARS plasmids A and F using cassettes generated by BioBasic. To generate SARS-F plasmids containing both TRN mutations and the mouse-adapted mutation at nt 2663, the mouse-adapted mutation was cloned into the F TRN plasmids via a PCR and restriction digestion strategy. To generate SARS-F plasmids containing both TRN mutations and the mouse-adapted mutation at nt 2663, the mouse-adapted mutation was cloned into the F TRN plasmids via a PCR and restriction digestion strategy. The plasmids F-BstZ (5'-GGAGGCGCAATTTTGTACCTCTATGCCTTG-3'), R-MAmut (5'-AGCTATCGTCTCCGCTTCTCAACGGTAATAGTACCGTTGTCTG-3'), F-MAmut (5'-AGCTATCGTCTCCAAGCTAAACAACCTCTGGAACAA TGGAAC-3'), and R-Msc (5'-GTGGCTTAGCTACTTCGTTGCTTCCT CAGGC-3') were used to generate 2 amplicons. The resulting amplicons were restriction-digested with *BsmB* I, ligated, and purified, after which the ligated amplicons and the parent vectors were restriction-digested with *BstZ17* I and *Msc* I and ligated. Ligated vectors were transformed into TopTen *E. coli* cells, and the resulting colonies were screened and sequence-verified. Viruses were then constructed as described in ref. <sup>58</sup>.

**Northern blot analysis.** Intracellular RNA was isolated using RiboPure reagents (Ambion, Austin, TX) 12 h post-infection<sup>13</sup>. The mRNA was then isolated using a Qiagen Oligotex mRNA isolation kit, treated with glyoxal, and separated on an agarose gel using NorthernMax-Gly (Ambion). The RNA was then transferred to BrightStar-Plus membrane (Ambion) for 5 h, cross-linked using UV light, pre-hybridized, and probed with an N gene-specific oligonucleotide probe (5'-CTTGACTGCCGCTCTGCT<sup>b</sup>T<sup>b</sup>CCCT<sup>b</sup>CT<sup>b</sup>GC<sup>b</sup>-3'; biotinylated nucleotides are denoted with a superscripted “b”). The blot was hybridized overnight and washed with low- and high-stringency buffers and was then incubated with phosphatase-conjugated streptavidin. The blot was then incubated with CDP-STAR, overlaid with film, and developed.

**Mouse infections with SARS-CoV and mutants.** All experimental protocols involving mice were reviewed and approved by the institutional animal care and use committee at the University of North Carolina, Chapel Hill, NC, USA. The following mice were used: 10-week-old female BALB/c (Charles River Laboratories, Wilmington, MA, USA) and 14-month-old female BALB/c (Harlan Laboratories, Indianapolis, IN, USA). Mice were lightly anesthetized and infected intranasally with varying doses (10<sup>2</sup>–10<sup>6</sup> PFU, depending on the experiment) of SARS-CoV, MA15, or TRN mutants. Mice were weighed daily, and on certain days specified in each experiment, mouse lungs were harvested for virus titer and/or RNA. Serial passages were inoculated as above for passage 1; subsequent passages were inoculated with 50  $\mu$ L of clarified lung homogenate (lungs were homogenized in 1 mL of PBS) from the previous passage. All experiments used a minimum of  $n = 5$  mice per virus per dosage/condition (if applicable) per timepoint. For infection–challenge studies, mice were infected with 10<sup>2</sup>–10<sup>3</sup> PFU of the indicated vaccine virus, weighed for the 7 days following initial infection, and then challenged with a lethal dose (10<sup>6</sup> PFU) of MA15 for the challenge infection.

**Determination of virus titer in infected mouse lungs.** Lungs harvested for virus titer were weighed and homogenized in 1.0 mL of PBS at 6000 rpm for 60 s in a MagnaLyser (Roche, Basel, Switzerland). Virus titers were determined by plaque assay on Vero cells.

**Determination of viral neutralization antibody titers.** Mouse sera were heat-inactivated for 30 min at 55 °C and were then serially diluted to 1:100, 1:200, 1:400, 1:800, and 1:1600 in PBS to a volume of 125  $\mu$ L. Next, 125  $\mu$ L of PBS containing low-concentration SARS-CoV MA15 (40 PFU) or high-concentration SARS-CoV MA15 (240 PFU) was added to each serum dilution. The virus-serum mixtures were incubated at 37 °C for 30 min. Following incubation, virus titers of the



mixtures were determined by plaque assay. Finally, we calculated the 50% plaque reduction neutralization titer (PRNT<sub>50</sub>) values, the serum dilutions at which plaque formation was reduced by 50% relative to that of virus stock not treated with serum.

**Viral genome sequencing.** To determine the sequences of viral genomes present in mouse lungs after passage, plaques were isolated from lung samples as described above. Briefly, individual viral plaques were harvested by collecting the agarose plugs above them using a 200- $\mu$ L pipette tip. Each agarose plug was dropped in 0.5 mL PBS, allowed to diffuse for 24 h at 4 °C, and then applied to ~70% confluent monolayers of Vero-E6 cells in T25 flasks and incubated for 48 h at 37 °C. Infected cell monolayers were then harvested in 1 mL of TRIzol. First-strand cDNA was generated as described in ref. <sup>59</sup>. Amplicons of the viral genomes were generated as described in ref. <sup>1</sup>. Sequence results were analyzed using Geneious R11 (Biomatters, Auckland, New Zealand) and Serial Cloner 2.6.1 (SerialBasics, <http://serialbasics.free.fr/Home/Home.html>).

**Statistical analyses.** Statistical analyses were performed using GraphPad Prism 7 (GraphPad Software, La Jolla, CA, USA). The tests run depended on the experimental design and are specified in the text. Significance was set at  $P < 0.05$ .

### Data availability

The viral genome sequence datasets are available via the NCBI GenBank with accession identifiers MK062179 through MK062184. All other data generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 31 January 2018 Accepted: 19 September 2018

Published online: 29 October 2018

### References

1. Graham, R. L. et al. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* **18**, 1820–1826 (2012).
2. Graham, R. L., Donaldson, E. F. & Baric, R. S. A decade after SARS: strategies for controlling emerging coronaviruses. *Nat. Rev. Microbiol.* **11**, 836–848 (2013).
3. Jones, K. E. et al. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
4. Li, Y. et al. On the origin of smallpox: correlating variola phylogenics with historical smallpox records. *Proc. Natl Acad. Sci. USA* **104**, 15787–15792 (2007).
5. Morens, D. M. & Fauci, A. S. The 1918 influenza pandemic: insights for the 21st century. *J. Infect. Dis.* **195**, 1018–1028 (2007).
6. Sessa, R., Palagiano, C., Scifoni, M. G., di Pietro, M. & Del Piano, M. The major epidemic infections: a gift from the Old World to the New? *Panminerva Med.* **41**, 78–84 (1999).
7. Karesh, W. B. et al. Ecology of zoonoses: natural and unnatural histories. *Lancet* **380**, 1936–1945 (2012).
8. Kreuder Johnson, C. et al. Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci. Rep.* **5**, 14830 (2015).
9. Al-Tawfiq, J. A. et al. Surveillance for emerging respiratory viruses. *Lancet Infect. Dis.* **14**, 992–1000 (2014).
10. Menachery, V. D. et al. SARS-like WIV1-CoV poised for human emergence. *Proc. Natl Acad. Sci. USA* **113**, 3048–3053 (2016).
11. Peiris, J. S., Guan, Y. & Yuen, K. Y. Severe acute respiratory syndrome. *Nat. Med.* **10**, S88–S97 (2004).
12. Sola, I., Moreno, J. L., Zuniga, S., Alonso, S. & Enjuanes, L. Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J. Virol.* **79**, 2506–2516 (2005).
13. Yount, B., Roberts, R. S., Lindesmith, L. & Baric, R. S. Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: engineering a recombination-resistant genome. *Proc. Natl Acad. Sci. USA* **103**, 12546–12551 (2006).
14. Zuniga, S., Sola, I., Alonso, S. & Enjuanes, L. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**, 980–994 (2004).
15. Sheahan, T. et al. Successful vaccination strategies that protect aged mice from lethal challenge from influenza virus and heterologous severe acute respiratory syndrome coronavirus. *J. Virol.* **85**, 217–230 (2011).
16. Graham, R. L. & Baric, R. S. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* **84**, 3134–3146 (2010).

17. Consortium, C. S. M. E. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. *Sci. (New Y., N. Y.)* **303**, 1666–1669 (2004).
18. Tang, J. W. et al. The large 386-nt deletion in SARS-associated coronavirus: evidence for quasispecies? *J. Infect. Dis.* **194**, 808–813 (2006).
19. Mateos-Gomez, P. A., Zuniga, S., Palacio, L., Enjuanes, L. & Sola, I. Gene N proximal and distal RNA motifs regulate coronavirus nucleocapsid mRNA transcription. *J. Virol.* **85**, 8968–8980 (2011).
20. Mateos-Gomez, P. A., Morales, L., Zuniga, S., Enjuanes, L. & Sola, I. Long-distance RNA-RNA interactions in the coronavirus genome form high-order structures promoting discontinuous RNA synthesis during transcription. *J. Virol.* **87**, 177–186 (2013).
21. Lauring, A. S., Jones, J. O. & Andino, R. Rationalizing the development of live attenuated virus vaccines. *Nat. Biotechnol.* **28**, 573–579 (2010).
22. Vignuzzi, M., Wendt, E. & Andino, R. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.* **14**, 154–161 (2008).
23. Feng, K. et al. Epidemiology and characterization of avian infectious bronchitis virus strains circulating in southern China during the period from 2013–2015. *Sci. Rep.* **7**, 6576 (2017).
24. Quinteros, J. A. et al. Full genome analysis of Australian infectious bronchitis viruses suggests frequent recombination events between vaccine strains and multiple phylogenetically distant avian coronaviruses of unknown origin. *Vet. Microbiol.* **197**, 27–38 (2016).
25. Sun, R. Q. et al. Outbreak of porcine epidemic diarrhea in suckling piglets, China. *Emerg. Infect. Dis.* **18**, 161–163 (2012).
26. Zhang, Y. et al. Complete genome sequence and recombination analysis of infectious bronchitis virus attenuated vaccine strain H120. *Virus Genes* **41**, 377–388 (2010).
27. Wang, L. et al. Discovery and genetic analysis of novel coronaviruses in least horseshoe bats in southwestern China. *Emerg. Microbes Infect.* **6**, e14 (2017).
28. Wang, L., Junker, D. & Collisson, E. W. Evidence of natural recombination within the S1 gene of infectious bronchitis virus. *Virology* **192**, 710–716 (1993).
29. Menachery, V. D., et al. Combination attenuation offers strategy for live-attenuated coronavirus vaccines. *J. Virol.* **92**, e00710–18 (2018).
30. Xiao, Y. et al. RNA recombination enhances adaptability and is required for virus spread and virulence. *Cell Host Microbe* **19**, 493–503 (2016).
31. Runckel, C., Westesson, O., Andino, R. & DeRisi, J. L. Identification and manipulation of the molecular determinants influencing poliovirus recombination. *PLoS Pathog.* **9**, e1003164 (2013).
32. Cheng, C. P., Servienu, E. & Nagy, P. D. Suppression of viral RNA recombination by a host exoribonuclease. *J. Virol.* **80**, 2631–2640 (2006).
33. Figlerowicz, M., Nagy, P. D. & Bujarski, J. J. A mutation in the putative RNA polymerase gene inhibits nonhomologous, but not homologous, genetic recombination in an RNA virus. *Proc. Natl Acad. Sci. USA* **94**, 2073–2078 (1997).
34. Nagy, P. D., Pogany, J. & Simon, A. E. RNA elements required for RNA recombination function as replication enhancers in vitro and in vivo in a plus-strand RNA virus. *EMBO J.* **18**, 5653–5665 (1999).
35. Nagy, P. D. & Simon, A. E. New insights into the mechanisms of RNA recombination. *Virology* **235**, 1–9 (1997).
36. Terada, Y. et al. Emergence of pathogenic coronaviruses in cats by homologous recombination between feline and canine coronaviruses. *PLoS ONE* **9**, e106534 (2014).
37. Truyen, U., Parrish, C. R., Harder, T. C. & Kaaden, O. R. There is nothing permanent except change. *Emerg. New Virus Dis. Vet. Microbiol.* **43**, 103–122 (1995).
38. Hu, B. et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* **13**, e1006698 (2017).
39. Tang, X. C. et al. Prevalence and genetic diversity of coronaviruses in bats from China. *J. Virol.* **80**, 7481–7490 (2006).
40. Tao, Y., et al. Surveillance of Bat coronaviruses in Kenya identifies relatives of human coronaviruses NL63 and 229E and their recombination history. *J. Virol.* **91**, e01953–16 (2017).
41. Di, H. et al. Expanded subgenomic mRNA transcriptome and coding capacity of a nidovirus. *Proc. Natl Acad. Sci. USA* **114**, E8895–E8904 (2017).
42. Frieman, M. et al. Molecular determinants of severe acute respiratory syndrome coronavirus pathogenesis and virulence in young and aged mouse models of human disease. *J. Virol.* **86**, 884–897 (2012).
43. Goldstein, D. R. Aging, imbalanced inflammation and viral infection. *Virulence* **1**, 295–298 (2010).
44. Totura, A. L. & Baric, R. S. SARS coronavirus pathogenesis: host innate immune responses and viral antagonism of interferon. *Curr. Opin. Virol.* **2**, 264–275 (2012).
45. Rockx, B. et al. Early upregulation of acute respiratory distress syndrome-associated cytokines promotes lethal disease in an aged-mouse model of severe acute respiratory syndrome coronavirus infection. *J. Virol.* **83**, 7062–7074 (2009).

46. Menachery, V. D. et al. Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-O-methyltransferase activity. *J. Virol.* **88**, 4251–4264 (2014).
47. Zust, R. et al. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. *PLoS Pathog.* **3**, e109 (2007).
48. Beall, A. et al. Characterization of a pathogenic full-length cDNA clone and transmission model for porcine epidemic diarrhea virus strain PC22A. *MBio* **7**, e01451 (2016). 01415.
49. Gonzalez, J. M., Penzes, Z., Almazan, F., Calvo, E. & Enjuanes, L. Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by insertion of an intron. *J. Virol.* **76**, 4655–4661 (2002).
50. Meulenber, J. J., Bos-de Ruijter, J. N., van de Graaf, R., Wensvoort, G. & Moormann, R. J. Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. *J. Virol.* **72**, 380–387 (1998).
51. Youn, S., Leibowitz, J. L. & Collisson, E. W. In vitro assembled, recombinant infectious bronchitis viruses demonstrate that the 5a open reading frame is not essential for replication. *Virology* **332**, 206–215 (2005).
52. Zhou, H. et al. Identification of a novel recombinant virulent avian infectious bronchitis virus. *Vet. Microbiol.* **199**, 120–127 (2017).
53. Zhou, Y. S. et al. Establishment of reverse genetics system for infectious bronchitis virus attenuated vaccine strain H120. *Vet. Microbiol.* **162**, 53–61 (2013).
54. Ge, X. Y. et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535–538 (2013).
55. Ge, X. Y. et al. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Viol. Sin.* **31**, 31–40 (2016).
56. Yang, X. L. et al. Isolation and characterization of a novel Bat coronavirus closely related to the direct progenitor of severe acute respiratory syndrome coronavirus. *J. Virol.* **90**, 3253–3256 (2015).
57. Menachery, V. D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med* **21**, 1508–1513 (2015).
58. Yount, B. et al. Reverse genetics with a full length infectious cDNA of the severe acute respiratory syndrome coronavirus. *Proc. Natl Acad. Sci. USA* **100**, 12995–13000 (2003).
59. Becker, M. M. et al. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc. Natl Acad. Sci. USA* **105**, 19944–19949 (2008).

## Acknowledgements

This research was supported by NIH NIAID grants U19-AI107810, R01-AI108197 and U54-AI057157 to R.S.B. and 5F32AI080148 to R.L.G.

## Author contributions

R.L.G. designed and performed experiments, analyzed data and wrote and edited the paper. D.J.D., M.E.D. and B.L.Y. designed and performed experiments and analyzed data. R.S.B. analyzed data and wrote and edited the paper.

## Additional information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s42003-018-0175-7>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018