The ancient cardioprotective mechanisms of ACE2 bestow SARS-CoV-2 with a wide host range

One sentence summary: A conserved mechanism essential for ACE2 catalytic activity is exploited by SARS-CoV-2 binding, allowing the virus to infect a wide range of species.

Authors

Gianni M. Castiglione[1], Lingli Zhou[1], Zhenhua Xu[1], Zachary Neiman[1], Chien-Fu Hung[2], and Elia J. Duh[1]*.

Affiliations

[1]Department of Ophthalmology, [2]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

*Correspondence to: eduh@jhmi.edu

Abstract

SARS-CoV-2 infects a broader range of mammalian species than previously anticipated, suggesting there may be additional unknown hosts wherein the virus can evolve and potentially circumvent effective vaccines. We find that SARS-CoV-2 gains a wide host range by binding ACE2 sites essential for ACE2 carboxypeptidase activity. Six mutations found only in rodent species immune to SARS-CoV-2 are sufficient to abolish viral binding to human and dog ACE2. This is achieved through context-dependent mutational effects (intramolecular epistasis) conserved despite ACE2 sequence divergence between species. Across mammals, this epistasis generates sequence-function diversity, but through structures all bound by SARS-CoV-2. Mutational trajectories to the mouse conformation not bound by SARS-CoV-2 are blocked, by single mutations functionally deleterious in isolation, but compensatory in combination, explaining why human polymorphisms at these sites are virtually non-existent. Closed to humans, this path was opened to rodents *via* permissive cardiovascular phenotypes and ancient increases to ACE2 activity, serendipitously granting SARS-CoV-2 immunity. This reveals how ancient evolutionary trajectories are linked with unprecedented phenotypes such as COVID-19 and suggests extreme caution should be taken to monitor and prevent emerging animal reservoirs of SARS-CoV-2.

MAIN TEXT

Introduction

SARS-coronavirus-2 (SARS-CoV-2) is thought to have evolved from an ancestor

(RaTG13) isolated from the Chinese horseshoe bat (*Rhinolophus sinicus*) (*1*), potentially fusing

36 with viruses from an intermediate source (e.g. pangolin; *Manis javanica*) (*2*). These zoonotic

37 origins ultimately led to the evolution of a virus that is highly transmissible among humans,

38 causing an unprecedented public health emergency (*3, 4*). A wide phylogenetic range of

39 mammalian species have been demonstrated to be susceptible to SARS-CoV-2 infection,

40 including non-human primates, dogs, cats, ferrets, hamsters and minks (*5-12*). Characterizing the

41 entire host range of SARS-CoV-2 is important for identifying the risks of anthroponosis and

42 zoonosis between humans and other species, which can pose a major health risk by forming

43 novel viral reservoirs where new mutations can evolve (*13, 14*). A recent example is mink

44 farms, where infection by humans led to the evolution of novel viral strains, which have since

45 reinfected human populations (*11*). Attempts to predict the host range of SARS-CoV-2 have

46 greatly underestimated the extent to which SARS-CoV-2 can infect certain non-primate species,

47 especially non-felid carnivores including mink, dog and ferret (*5-11, 15*). These predictive

48 methods depend on comparative sequence analyses of angiotensin-converting enzyme 2 (ACE2)

49 – the cellular receptor for SARS-CoV-2 – and scoring based on sequence and structural

50 homology to the human ACE2 viral binding interface (*15-17*). The difficulty of predicting the

51 SARS-CoV-2 host range through these methods proves that the virus can bind a wide-range of

52 ACE2 orthologs despite extensive sequence divergence. This opens the possibility for the

53 existence of additional unknown mammalian hosts.

54   Infection by SARS-CoV-2 is mediated by the binding of the viral S-protein receptor

55 binding domain (RBD) to ACE2 (*1, 18*), displaying a nanomolar affinity higher than that of

56 SARS-CoV-1 (*19*). Previously, it was identified that a single mutation to ACE2 was sufficient to

57 decrease SARS-CoV-1 RBD binding by nearly 90% (*20*). By contrast, structural analyses have

58 identified that the SARS-CoV-2 RBD targets multiple binding 'hotspots' within the ACE2

59   ectodomain, presenting a diffuse and multifaceted binding strategy (*21-23*). While many sites

60   within the viral binding domain of ACE2 can modulate binding of the SARS-CoV-2 S protein

61   RBD (*24*), it remains unclear how amino acid interactions between these sites facilitate viral

62   binding and infection. The ability of the SARS-CoV-2 S protein RBD to bind divergent

63   sequences may be therefore be explained by ACE2 intramolecular epistasis, a form of context

64   dependence, where amino acids have different functional effects depending on residues at other

65   sites (*25*). Although intramolecular epistasis can limit the number of possible amino acids

66   combinations that will confer a given function (thus making evolution 'predictable'; (*26, 27*)), it

67   can also give rise to novel compensatory interactions following an 'original' mutation (*28-31*),

68   ultimately generating completely different amino acid combinations converging on the same

69   structural/functional 'solution', making sequence-function evolution unpredictable (*32-36*). If

70   pervasive intramolecular epistasis exists within the viral binding interface of ACE2, then this

71   may explain why ACE2 sequence homology alone has not been a reliable predictor of viral

72   binding and infection.

73   In mammals, ACE2 evolved to serve an essential reno- and cardio-protective role *in vivo*,

74   mediated through its regulation of the renin-angiotensin-aldosterone system (RAAS) in

75   conjunction with ACE. The signaling peptide Angiotensin-II (Ang-II) generated by ACE—a

76   major clinical target for hypertension—stimulates vasoconstriction, inflammation, and fibrosis

77   responses through the Ang-II/AT$_1$R axis (*37, 38*). ACE2 carboxypeptidase activity counteracts

78   these effects through conversion of Ang-II to Ang-(1-7), a peptide which induces vasodilatory,

79   anti-inflammatory, and anti-fibrotic effects through MAS signaling (*38, 39*). Loss of ACE2 in

80   mice worsens cardiac dysfunction in obesity, increases diabetic kidney dysfunction, increases

81   mortality rates after myocardial infarction, and can have severe effects on cardiac contractility

82   (*37, 39-41*). This protective role of ACE2 is largely mediated through its enzymatic processing

83   of Ang II to Ang-(1-7), where exogenous delivery of either ACE2 or Ang-(1-7) can protect

84   against pathogenic features of multiple cardiovascular and kidney diseases (*37, 41, 42*). Given

85   the myriad protective roles of ACE2 enzymatic activity, it may be expected that ACE2 function

86   is highly conserved across species. However, mice display ~ 50% higher ACE2 activity relative

87   to humans (*43*), and are insensitive to the vasodilatory effects of Ang-(1-7) (*39*), suggesting

88   increased ACE2 activity in rodents may have evolved to serve non-vasodilatory protective

89   functions. These observations suggest that minor differences in ACE2 function can drive major

90   physiological differences between species, as seen in the evolution of other protein systems (*27,*

91   *31, 44*). Organismal sensitivity to mutations affecting ACE2 activity may be especially

92   pronounced due to the X-chromosome location of the *ACE2* gene (*45*), where even heterozygous

93   *ACE2* knockout females (-/x) display increased susceptibility to heart and kidney injury (*39*).

94   If natural variation across species has evolved to shape ACE2 function, then the sequence

95   recognized by SARS-CoV-2 may have a pleiotropic role in mediating ACE2 function. We

96   therefore investigated whether SARS-CoV-2 binding to ACE2 could be abolished *without*

97   disrupting ACE2 enzyme function. To expedite this, we took an evolutionary approach that

98   leveraged the natural sequence variation found in mouse ACE2, which SARS-CoV-2 is unable to

99   bind (*1, 18*). Here we identify a specific combination of mutations unique to rodents which fully

100   abolishes RBD binding when inserted into human and dog ACE2, but which in isolation,

101   significantly decreases ACE2 enzyme activity. These detrimental intermediates would likely

102   severely compromise the cardio- and reno-protective functions of ACE2 activity, explaining why

103   these mutations are rare across mammalian species.

104

Results

We investigated ACE2-RBD binding as well as ACE2 enzymatic function across a range

of boreoeutherian mammals either susceptible or immune to SARS-CoV-2 infection [human

(*Homo sapiens*), dog (*Canis lupus familiaris*), and pangolin (*Manis javanica*); *vs.* mouse, (*Mus*

musculus) and Chinese horseshoe bat (*Rhinolophus sinicus*), respectively] (*1, 7, 9, 18*). Using

flow cytometry, we found a trend of significantly stronger association of the SARS-CoV-2 RBD

with both human and pangolin ACE2 relative to that of CoV-1, consistent with previous studies

(*19*) (Fig. 1A-D). Notably, CoV-1 and -2 RBD association was strongest with human ACE2. We

also found evidence that the RBD of both SARS-CoV-1 and -2 S protein could not bind mouse

nor bat ACE2 (Fig. 1B-D), consistent with previous studies and reaffirming pangolin as a likely

intermediate between bat and human (*2, 7, 9, 18, 46*). To complement these binding assays, we

next characterized the ability of ACE2 orthologs to enable SARS-CoV-2 S pseudovirus entry

into cells. We transfected ACE2 orthologs into human cells (HEK293T) and exposed these cells

to pseudotyped MLV particles containing the SARS-CoV-2 S protein. Consistent with our flow

cytometry data and with previous studies, this assay recapitulated the host range of WT SARS-

CoV-2 (*18*), displaying significant infection of cells expressing human, dog, and pangolin ACE2,

but not that of mouse or bat (Fig. 1E). To investigate whether this variation in ACE2-SARS-

CoV-2 S binding is mirrored by functional variation in ACE2 enzyme activity, we measured the

carboxypeptidase activity of ACE2 orthologs *in vitro* using a well-characterized fluorometric

biochemical assay (*43*) (Materials and Methods; Fig. S1). Strikingly, we found significant

variation in ACE2 hydrolysis rates across all mammalian species, suggesting tuning of

enzymatic function across evolutionary history (Fig. 1F-G), perhaps in response to interspecies

RAAS variation. Interesting, dog, bat and pangolin all displayed low ACE2 hydrolysis relative to

129   human and mouse. Mice displayed the highest ACE2 hydrolysis rates, consistent with previous

130   reports (*43*). This demonstrates that ACE2 function varies considerably across species, and

131   suggests that ACE2 sequence variation may reflect diversification of ACE2 enzymatic function.
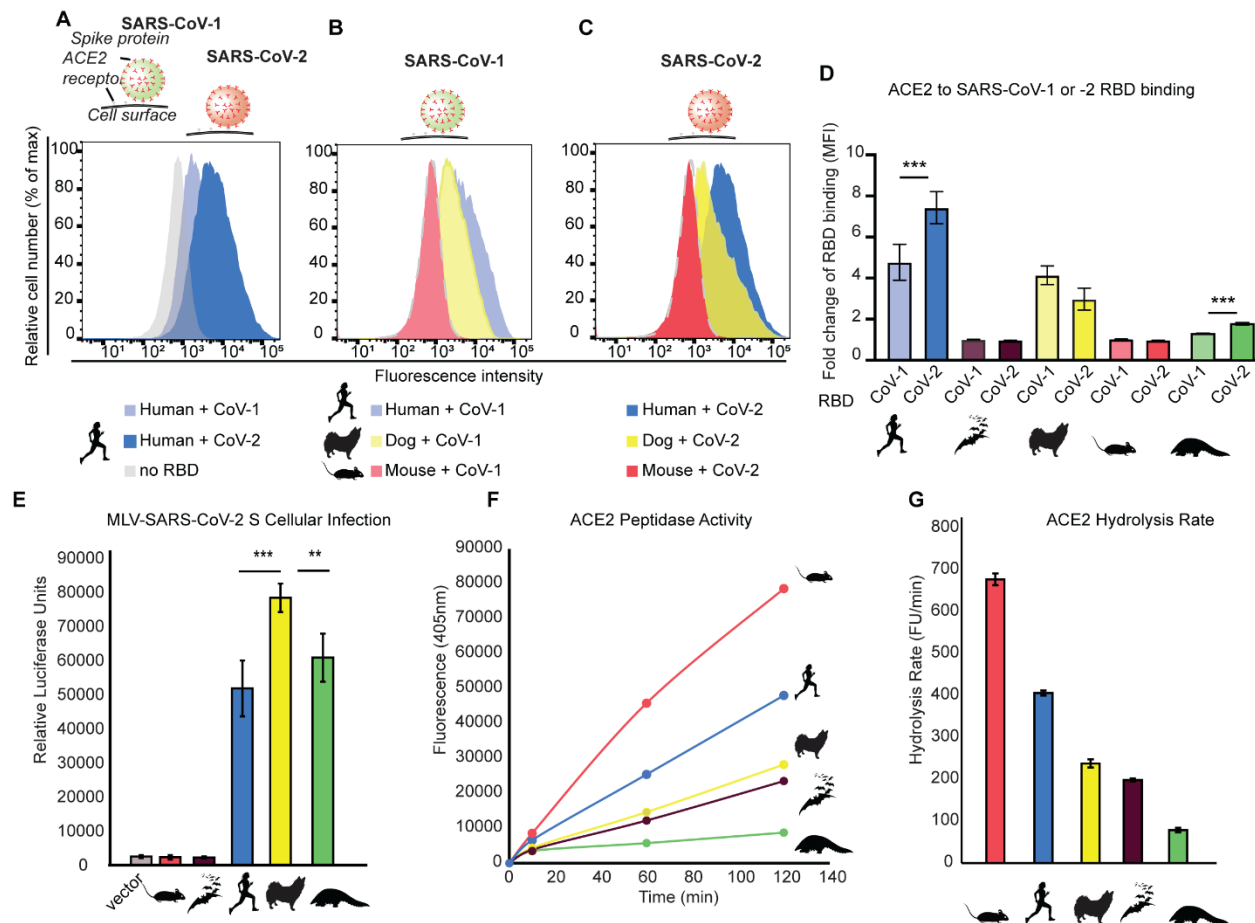

Fig 1. Natural variation in SARS-CoV-2 binding is mirrored by diversity in ACE2 enzyme activities. (A-D) Flow cytometry was used to quantify SARS-CoV-1 and CoV-2 RBD-Fc association with human cells (HEK293T) expressing ACE2-eGFP orthologs from various mammalian species (mean fluorescence intensity; MFI (*19*)). (E) SARS-CoV-2 S pseudovirus infection of HEK293T cells expressing ACE2 orthologs was quantified using a luciferase reporter system. (F) The carboxypeptidase activity of ACE2 was quantified using a fluorometric peptide incubated with solubilized HEK293T cells transfected with ACE2 orthologs. (G) Hydrolysis rate of ACE2 orthologs (Fluorescence units/minute).

141   To test this, we first took a bioinformatic approach. We reasoned that since ACE2 has

142   evolved under pressures related to its enzymatic processing of Ang-II/Ang-(1-7), signatures of

143   natural selection at sites within the viral binding interface may reflect a potential role of those

144   sites in mediating ACE2 catalytic activity (*47*). We therefore searched for shifts in *ACE2*

145   mutational rates (d_N/d_S) beyond what may be expected from neutral evolutionary pressures alone

146   (*48, 49*). To do this, we constructed a phylogenetic dataset representing full-length mammalian

147   *ACE2* sequences (Fig. S2; Table S1) and conducted a codon-based statistical phylogenetic

148   analysis (PAML, HyPhy). We found significant evidence that mammalian ACE2 is under

149   positive selection, outperforming models assuming neutral molecular evolution (Table S2). This

150   provides evidence that mammalian ACE2 functional diversification is a product of natural

151   selection, rather than neutral processes.

152   To focus our analysis on ACE2 sites responsible for mediating SARS-CoV-2 binding, we took

153   an evolutionary approach leveraging natural sequence variation found in mouse ACE2, which

154   SARS-CoV-2 is unable to bind (*1, 18*). We identified a set of six residues unique to mouse

155   ACE2 that were of particular interest due to their proximity to viral RBD residues implicated in

156   the ACE2-RBD structure (Fig. 2A). These sites contact opposite ends of the viral RBD,

157   facilitating binding through a combination of hydrogen bonds (Q24, K353), as well as van der

158   Waals forces within a hydrophobic pocket of ACE2 (L79, M82, Y83, P84) (*21, 23*). These

159   ACE2 residues are located outside the ACE2 active site mediating catalysis of Ang-II/Ang-(1-7)

160   (Fig. 2A; yellow), suggesting that any role of these sites in mediating ACE2 function would

161   likely be through indirect effects modulating the protein structure, as seen in other protein

162   systems (*47, 50*). Through phylogenetic methods (Table S2; M8; Bayes empirical Bayes), we

163   found that several of these sites deviated significantly from neutral expectations (sites 24 ,79,

164   dN/dS>1; site 353, dN/dS<1), suggesting potential functional roles (Fig.2B). Interestingly, these

165   sites are less variable in primates relative to other mammalian lineages (Fig. 2B; (*16*)),

166   displaying significantly decreased *ACE2* evolutionary rates relative to bats (Chiroptera) and

167   rodents (Table S3; (CmD; dN/dS). These evolutionary signatures suggest that the RBD binding

168   interface of ACE2 may play a role in the functional diversification of ACE2 catalytic activity,

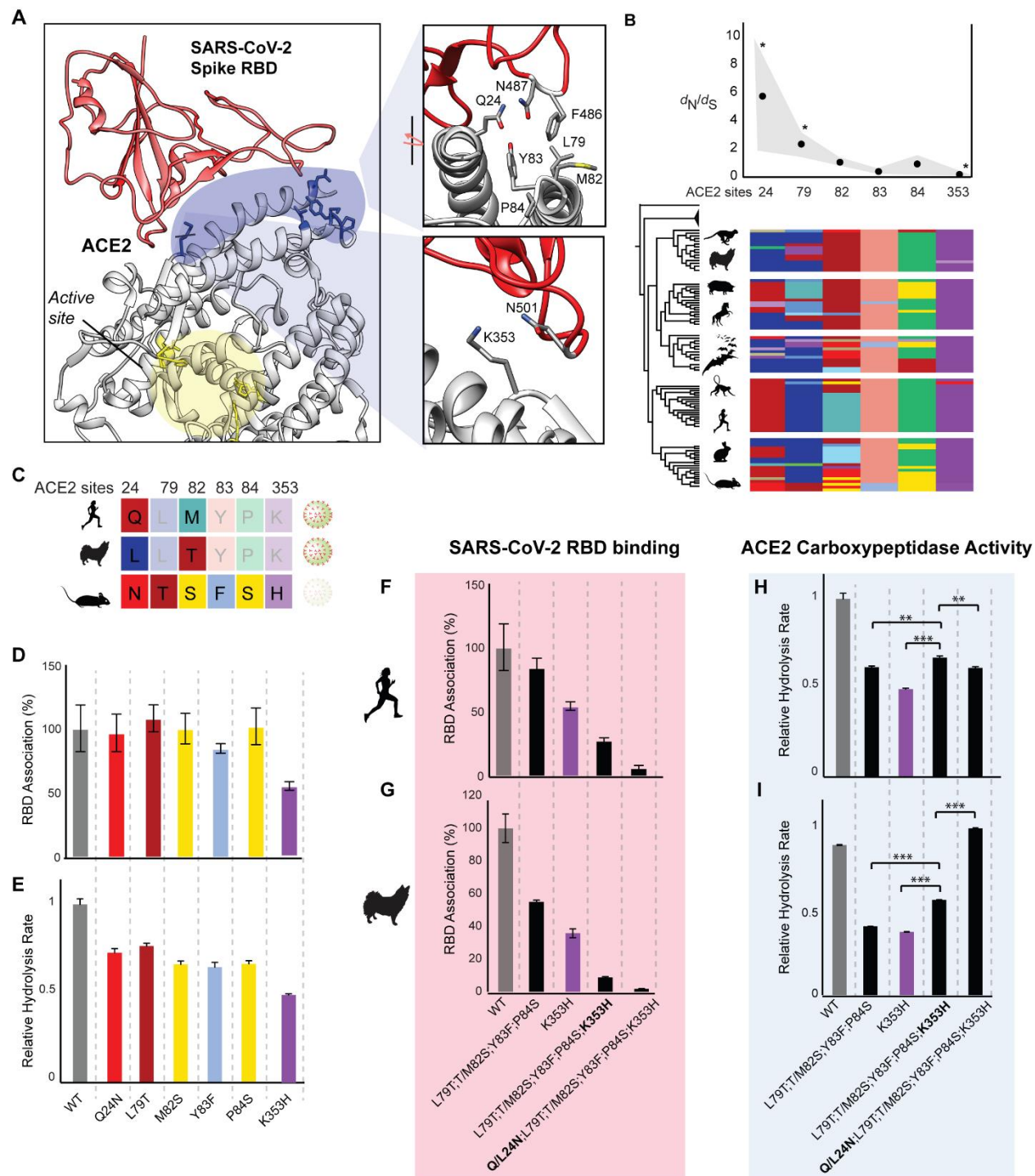169   with extant residues in primates being strongly constrained.

170

171 **Fig 2. Amino acid interactions mediating ACE2 activity govern SARS-CoV-2 binding to human**
172 **and dog ACE2** (**A**) SARS-CoV-2 gains cellular entry through the viral spike protein receptor binding
173 domain (RBD, red), which targets binding hotspots on the ACE2 receptor (blue) distal to the ACE2 active
174 site (yellow) [6M17; (*23*)]. (**B**) Statistical phylogenetic analyses (dN/dS averages [dots] and ranges
175 [grey]; PAML, HyPhy) of an alignment of mammalian ACE2 coding sequences reveals positive and
176 negative selection (*) on RBD binding hotspots. Alignment of ACE2 residues across boreoeutherian
177 mammals is shown. (**C**) SARS-CoV-2 is unable to bind mouse ACE2, which displays unique amino acid
178 residues at positions within the RBD binding interface relative to other mammals. (**D**) Flow cytometry
179 was used to quantify RBD association with human cells (HEK293T) expressing WT and mutant ACE2.
180 (**E**) The effect of ACE2 mutations on ACE2 hydrolysis rates was measured using a fluorometric peptide
181 (**F-G**) Interaction effects between amino acids on opposite ends of the viral binding interface determine
182 RBD binding (F-G) and ACE2 activity (H-I) in both human (F, H) and dog ACE2 (G, I).

183   We next experimentally characterized the role of these sites on both SARS-CoV-1 and

184 CoV-2 binding and infection, as well as their effect on ACE2 catalytic activity. We did this by

185 substituting the rare residues found in mouse ACE2 into human and dog ACE2 (Fig. 2C). First,

186 we characterized the role of single mutations on the binding of the SARS-CoV-1/2 S RBD to

187 human WT and mutant ACE2 using co-immunoprecipitation and flow cytometry. In contrast

188 with SARS-CoV-1 (Fig. S3), a single mutation was not sufficient to abolish SARS-CoV-2 RBD

189 binding to ACE2 (K353H; Fig. 2D). Indeed, most single mutations had only minor effects on

190 SARS-CoV-2 RBD binding (Fig. 2D). However, each of the six single mutations displayed

191 considerable effects on human ACE2 hydrolysis activity (Fig 2E). This demonstrates that ACE2

192 residues within the RBD binding interface can indirectly modulate ACE2 activity.

193   Next, we attempted to abolish RBD binding to both human and dog ACE2 by making

194 multiple substitutions at these ACE2 sites. Strikingly, RBD binding to the ACE2 of both species

195 was abolished with just these six substitutions (Fig. 2F-G), despite numerous other sites

196 implicated in mediating interactions with the viral RBD (*21, 23*). Using the MLV-SARS-CoV-2

197 pseudovirus system, we found that these six substitutions in combination also significantly

198 reduced pseudovirus infection of HEK293T cells expressing wild type and mutant human and

199 dog ACE2 (Fig. S4). In both species, abolition of RBD binding depended on mutating sites 24

200   and 82 (Fig. 2F-G), despite displaying different residues in human (Q24; M82) and dog (L24;

201   T82) ACE2 (Fig. 2C). This demonstrates that SARS-CoV-2 utilizes the same combination of

202   ACE2 positions to bind both human and dog ACE2 despite amino acid variation at these sites.

203   This may explain why bioinformatic predictions of SARS-CoV-2 host range based on human

204   ACE2 sequence homology have tended to underestimate the infection risk of species such as

205   dogs, ferrets, and minks (*6, 7, 9-11, 15*).

206   We reasoned that this could be due to conserved compensatory interactions between

207   ACE2 sites, allowing different residues to ultimately maintain the ACE2 structure bound by

208   SARS-CoV-2. Consistent with this, we observed in both human and dog ACE2 a context-

209   dependence (epistasis) of mutational effects between the distal domains of the RBD binding

210   interface. Specifically, the quadruple mutation (L79T; T/M82S; Y83F; P84S) had a much larger

211   effect on RBD binding when a single mutation on the opposite side of the ACE2 binding cleft

212   were included the genetic background (K353H) (Fig. 2F-G). Epistasis between these distal sites

213   indicates that residues at one end of the ACE2-RBD binding cleft can modulate the structural

214   effects of residues on the other end. We hypothesized that the specific amino acid combinations

215   at these sites therefore originally evolved to maintain the ACE2 structure underpinning ACE2

216   catalytic activity. Consistent with this, we found that the detrimental functional effects of the

217   quadruple mutant (L79T; T/M82S; Y83F; P84S) as well as the K353H mutant were significantly

218   reversed when combined—a phenomenon known as sign epistasis —displaying partial

219   compensation for each other's detrimental effects on ACE2 hydrolysis rates (Fig. H-I). Further

220   evidence of epistasis in this ACE2 domain is seen in the sextuple mutant, where introduction of

221   L24N in dog ACE2 fully rescued ACE2 activity, whereas Q24N decreased activity in human

222   relative to the quintuple mutant (Fig. H-I). This discrepancy is likely attributable to amino acid

223   interactions between site 24 and other sites not investigated. Taken together, these results

224   demonstrate that extensive amino acid interactions within the RBD binding interface of ACE2

225   indirectly mediate ACE2 enzymatic activity. This is achieved through compensatory interactions

226   that modulate the ACE2 structure recognized by SARS-CoV-2. This mechanism is conserved in

227   dog and human ACE2, which are separated by ~90 million years of evolution (*51*), suggesting

228   these structural interactions are a general feature of mammalian ACE2 evolution. Thus,

229   compensatory interactions in this domain appear to have facilitated divergent amino acid

230   combinations to converge and generate the same structure recognized by SARS-CoV-2 (*32-36*).

231   Hence, our evidence strongly suggests that the unpredictable and wide host range of SARS-CoV-
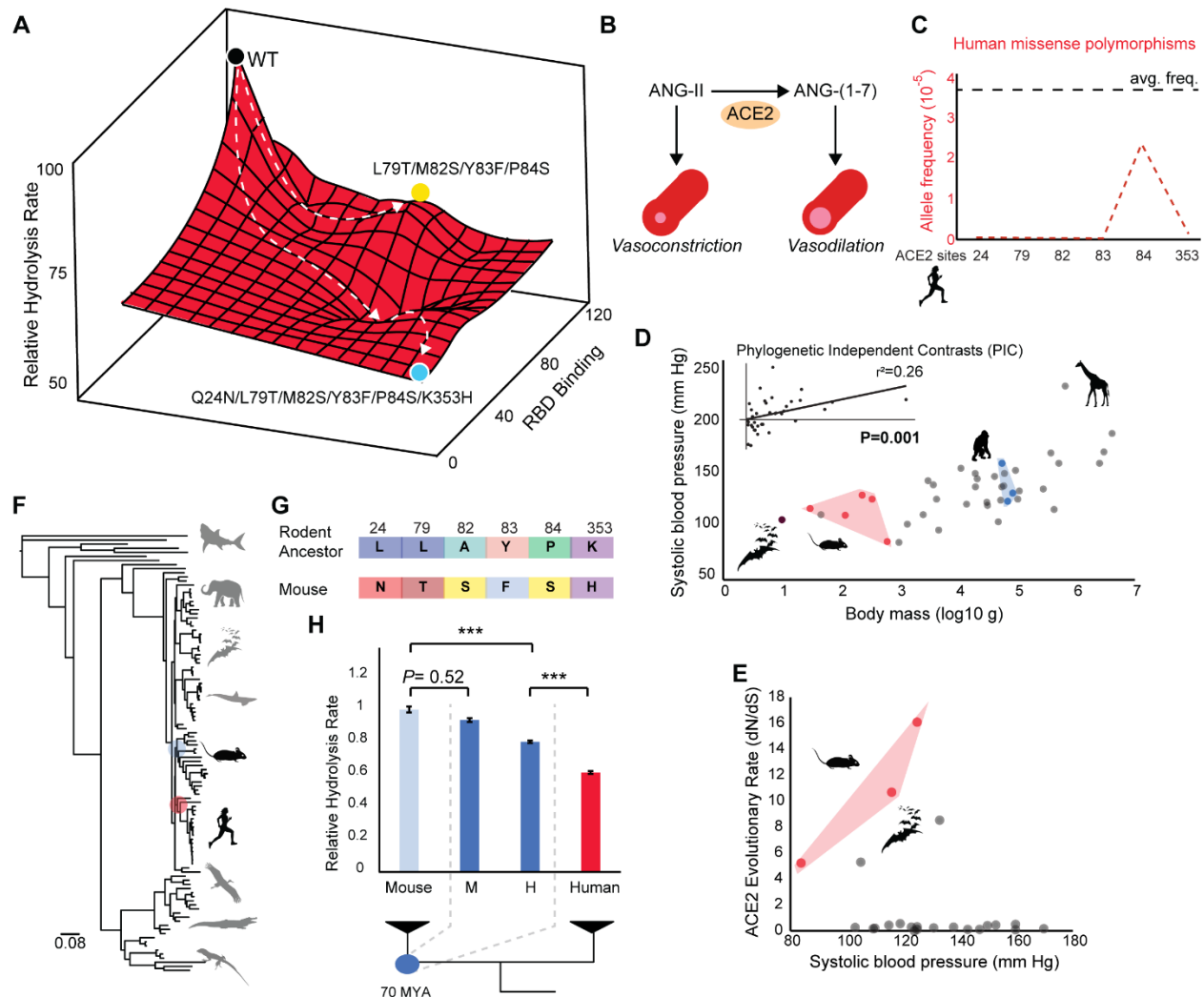
232   2 is likely due to this ancient epistasis of ACE2.

233

Fig 3. Human polymorphisms required for SARS-CoV-2 immunity are blocked by evolutionary valleys and cardiovascular constraints. (**A**) Substitution of mouse residues into human ACE2 abolishes ACE2-RBD binding, but with large trade-offs on ACE2 hydrolytic activity. Sign-epistasis in human and dog ACE2 activity creates a jagged landscape of intermediate valleys that block this evolutionary trajectory. (**B**) Hydrolysis of ANG-II by ACE2 generates ANG-(1-7), a vasodilatory peptide (*39*). (**C**) Allele frequencies for human missense polymorphisms at these positions are lower than the protein-wide average frequency at a given ACE2 site (black line). (**D**)Systolic blood pressure divergence between rodents (red) and primates (blue) is correlated with differences in body size (*P*=0.001; phylogenetically independent least squares linear regression). (**E**) Rodent and bats species display higher ACE2 evolutionary rates relative to other species (aBSREL). (**F**) Phylogeny of jawed vertebrates used in the ancestral reconstruction of *ACE2* from the last common ancestor of rodents (RodAnc). (**G**) RodAnc ACE2 displays none of the rare sequence variants unique to the RBD binding domain of mice (**H**) Evolution of high ACE2 activity in the rodent ancestor. The C-terminus of mouse (M) or human (H) was used.

250   Why did these substitutions serendipitously blocking SARS-CoV-2 binding only evolve

251   in some rodent species? The compensatory interactions inherent to intramolecular epistasis can

252   both open and close evolutionary trajectories depending on how accessible favorable mutational

253   combinations are within evolutionary space (*31, 52-54*). In humans and dogs, this trajectory

254   appears to have been closed by the presence of deep valleys containing detrimental intermediates

255   with compromised ACE2 function (Fig. 3A). Traversal of evolutionary valleys is thought to be

256   associated with fluctuations in environmental variables, which can change physiological

257   constraints on protein function (*31, 53*). Since ACE2 activity plays major roles in the

258   cardiovascular system, we reasoned that cardiovascular differences between mammalian species

259   may result in differential constraints on ACE2 sequence-function. A central cardiovascular role

260   mediated by ACE2 activity is the regulation of blood pressure *via* enzymatic processing of ANG-

261   (II) to ANG-(1-7) (Fig. 3B). We found extremely low human allele frequencies for missense

262   polymorphisms at these sites (gnomAD), with most at zero (Fig.3C). This is consistent with the

263   role of these sites in mediating ACE2 activity and with the known association of human ACE2

264   polymorphisms with hypertension (*55, 56*). We reasoned that these physiological constraints

265   would be more relaxed in animals with low blood pressures, permitting sequence-function

266   evolution. We conducted a phylogenetic statistical analysis on systolic blood pressure values

267   across mammalian species and observed a significant correlation with body size (r^2=0.26;

268   p=0.001; phylogenetic independent contrast least squares linear regression) (Fig. 3D; Fig. S6).

269   Blood pressures were lowest in rodent and bat species—two lineages immune to SARS-CoV-2

270   (*18, 46*) which also display the highest *ACE2* evolutionary rates relative to other mammals

271   (aBsREL; Fig. 3E; Table S2). This is consistent with bat (*Rhinolophus sinicus*) ACE2 displaying

272   the lowest hydrolysis rates relative to all other species investigated here (except Pangolin, for

273   which blood pressure measurements are not available to our knowledge). Interestingly, rodents

274   are insensitive to the vasodilatory effects of Ang-(1-7) (*39*), suggesting their high ACE2 activity

275   may be unrelated to low blood pressure. Although speculative, relaxed physiological constraints

276   on the vasodilatory role of ACE2 function (due to small body size) may explain why rodents

277   were situated to tolerate an evolutionary trajectory that presumably contained detrimental

278   intermediates as seen in dog and human ACE2 mutants.

279   An alternative explanation is that these evolutionary valleys may not exist in rodents.

280   Specifically, this trajectory may have been opened to mice by permissive mutations (*25, 31*) that

281   compensated for the subsequent introduction of rare variants that are strongly deleterious in dog

282   and human ACE2 backgrounds. To understand if ancestral rodent ACE2 displayed high activity

283   before the introduction of these mutations, we constructed a large jawed vertebrate *ACE2*

284   phylogenetic dataset (Fig. 3F, S5) and reconstructed the ancestral rodent (RodAnc) *ACE2* using

285   likelihood methods (*57*). Reconstructed RodAnc ACE2 had none of the rare sequence variants

286   unique to the RBD binding domain of mice (Fig. 3G). This strongly suggests that these

287   mutations appeared relatively recently in mouse ACE2 evolution. Since the C-terminus of

288   vertebrate *ACE2* was too divergent to align, we constructed the full-length rodent ancestor

289   (RodAnc) ACE2 coding-sequence with either mouse (-M) or human (-H) C-termini, and

290   measured ACE2-specific hydrolysis activity. We found that RodAnc-M displayed high ACE2

291   hydrolysis activity indistinguishable from that of mouse, strongly suggesting that the rodent

292   ancestor had high ACE2 hydrolysis rates (Fig. 3H). Moreover, the RodAnc with the human C-

293   terminus displayed ACE2 activity significantly lower than mouse ACE2, but significantly higher

294   than that of human, suggesting the C-terminus is necessary, but not sufficient to explain the high

295   ACE2 activity seen in the RodAnc. Taken together, these results provide strong evidence that

296   mutations outside the viral binding interface of ACE2 drove increases to ACE2 carboxypeptidase

297   activity during rodent evolution. High ACE2 activity in the rodent ancestor may have therefore

298   compensated for the subsequent introduction of rare variants that are strongly deleterious

299   mutations in dog and human ACE2 backgrounds.

300

301   **Discussion**

302   We have shown that six mutations abolishing SARS-CoV-2 RBD-ACE2 binding also

303   individually disrupt ACE2 hydrolysis activity by up to 50-60%. This provides evidence that like

304   other proteins, the catalytic activity of ACE2 can be mediated by the indirect structural effects of

305   mutations to residues outside the active site (*27, 47, 50*). Although previous work demonstrated

306   that inactivation of the ACE2 active site had no effect on SARS-CoV-1 and -2 binding, these

307   mutations were directly within the active site far from the viral binding interface (*58, 59*). The

308   mutations investigated here are directly within the viral binding interface of SARS-CoV-2, and

309   from species naturally immune to SARS-CoV-2 that still maintain ACE2 activity. The surprising

310   finding that mutations abolishing binding also reduce ACE2 activity is supported by the fact that

311   human ACE2 engineered to increase viral binding (T27Y, L79T, N330Y) also displays reduced

312   ACE2 activity (*24*). Here we discuss the implications of pervasive epistasis in ACE2 for the

313   evolution of ACE2 function, the SARS-CoV-2 host range and therapeutic design.

314   Our work demonstrates the essential importance of the ACE2 viral binding domain on

315   catalytic function. The consequent constraints on the ACE2 amino acid sequence in humans and

316   other primates strongly indicate that protective polymorphisms in the human population are

317   unlikely, in contrast to examples such as CCR5, the co-receptor for HIV (*60*). Our analysis

318   shows that allele frequencies for ACE2 missense polymorphisms at six sites necessary for

319   SARS-CoV-2 binding are lower than the protein-wide average, reflecting the physiological

320   importance of these positions in mediating ACE2 catalytic activity. Although potentially

321   protective ACE2 polymorphisms have been predicted at other sites in the RBD-ACE2 interface

322   (*61, 62*), we show that statistically unlikely combinations of mutations are required to disrupt

323   SARS-CoV-2 binding, each of which have severe effects on ACE2 activity. Given the

324   importance of ACE2 enzymatic processing of Ang II to Ang-(1-7) in protection against

325   pathogenic features of multiple cardiovascular and kidney diseases (*37, 41, 42*), it is possible that

326   that the extant mutational combinations observed in nature (e.g. human, bat, dog, mouse,

327   pangolin) may all represent alternative sequence 'solutions' (*34, 52*) each uniquely required for

328   an animal's physiology, thereby representing fitness 'peaks' in the sequence landscape (*28, 31,*

329   *34, 53, 63-65*). Species variation in ACE2 hydrolysis rates we observed may reflect variation in

330   renin activity levels, which are a major factor in determining circulating Ang-II levels in humans

331   (*39, 66, 67*). It is therefore likely that a preponderance of selective processes has shaped ACE2

332   evolution, imposed by a combination of epistatic and physiological constraints on ACE2

333   sequence-function diversification, as discussed below.

334   Our results strongly imply that the host range of SARS-CoV-2 may be broader than both

335   empirical and computational studies have suggested (*5-11, 15*). In fact, our results suggest that

336   the SARS-CoV-2 host range among non-primate mammals may be essentially unpredictable.

337   This warrants extreme caution for public health efforts aiming to limit the risks of anthroponosis

338   and zoonosis between humans and other species, which can form novel viral reservoirs where

339   new mutations can evolve (*13, 14*). This has already occurred on mink farms, where infection by

340   humans led to the evolution of novel viral strains, which have since reinfected human

341   populations (*11*). The demonstrated infection of minks and other non-felid carnivores by SARS-

342   CoV-2 despite predictions to the contrary (*6, 7, 9-11, 15*) challenges the utility of predicting

343   infection risk from sequence homology to the human ACE2 viral binding interface (*15-17*). Our

344   results reconcile this apparent conundrum by demonstrating that the wide host range of SARS-

345   CoV-2 appears to originate from the structural effects of interacting amino acids within the

346   ACE2 viral binding interface. As in other proteins, this intramolecular epistasis makes the

347   sequence-phenotype relationship unpredictable (*32-36*); intramolecular epistasis has generated

348   sequence diversity driving ACE2 functional diversification (*34, 35*) but through similar

349   conformations all bound by SARS-CoV-2 (*25, 27, 29*). Only in a few lineages has the epistasis

350   of these sites been permitted to open trajectories to new conformations not recognized by SARS-

351   CoV-2 (*46*). Importantly, these new conformations confer ACE2 activity equivalent (bat *vs* dog)

352   or greater (mouse *vs* human) than species susceptible to SARS-CoV-2 infection. Why have

353   these mutational combinations not evolved in other mammalian lineages? The ability to traverse

354   evolutionary valleys is related to effective population sizes (*31*), which are negatively correlated

355   with body size (*70*), suggesting that the ACE2 mutational combinations investigated here may

356   have been generated by the effects of population size on mutational rates. However, given the

357   physiological importance of ACE2 activity discussed above, the existence of functionally

358   detrimental intermediates may also explain why these conformations did not evolve more

359   frequently in evolutionary history (*28, 31, 34, 53, 63, 64*). Our evidence suggests a confluence of

360   ancient permissive mutations and permissive cardiovascular constraints likely opened the

361   available pool of sequence variation, allowing functional *and* conformational diversification (*25,*

362   *28, 29, 31*). The highly effective binding mechanism of SARS-CoV-2 (*21-23, 71*) may therefore

363   synergize with the epistasis of the ACE2 domain to grant the virus a wide-host range—one that

364   is very difficult to predict.

365   Our study also has relevance for therapeutic design. Human recombinant soluble ACE2

366   (hrsACE2) is in active development as a strategy to neutralize SARS-CoV-2 by binding the viral

367   spike protein (*67, 78*). In human COVID-19 patients, the catalytic activity of hrsACE2 can help

368   reduce angiotensin II levels as well as inflammation associated with COVID-19, likely through

369   elevating Ang-(1-7) levels (*67*). hrsACE2 may therefore have the added benefit of minimizing

370   the injury to multiple organs caused by viral-induced downregulation of ACE2 expression and

371   renin-angiotensin hyperactivation (*40, 73, 79-82*). However, under normal physiological

372   conditions, higher sACE2 plasma activity has been associated with increased pulmonary artery

373   systolic pressure and ventricular systolic dysfunction (*83*). In these instances where patients have

374   pre-existing conditions that can be exacerbated by increasing sACE2 activity it may be beneficial

375   to make use of catalytically inactive hrsACE2 that binds RBD with similar efficiency (*58*). There

376   is likely to be a spectrum of scenarios where variable levels of ACE2 activity is desirable, in

377   which case recombinant non-human ACE2 can serve a key role. Our results also have

378   implications for domesticated animals susceptible to infection, especially minks, for whom the

379   SARS-CoV-2 is particularly transmissible (*10, 11*). In lieu of a vaccine for minks and other non-

380   human mammals important to agricultural production, genetic engineering efforts could feasibly

381   introduce the six residues we observed to abolish viral binding but maintain activity in dog

382   ACE2. This would remove a dangerous reservoir of SARS-CoV-2 wherein the virus can evolve

383   and potentially challenge the efficacy of emerging vaccines.

384

385

References and Notes

1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).

2. K. Xiao *et al.*, Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286-289 (2020).

3. D. K. Chu *et al.*, Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* **395**, 1973-1987 (2020).

4. H. E. Randolph, L. B. Barreiro, Herd Immunity: Understanding COVID-19. *Immunity* **52**, 737-741 (2020).

5. P. J. Halfmann *et al.*, Transmission of SARS-CoV-2 in Domestic Cats. *N Engl J Med*, (2020).

6. Y. I. Kim *et al.*, Infection and Rapid Transmission of SARS-CoV-2 in Ferrets. *Cell Host Microbe* **27**, 704-709 e702 (2020).

7. J. Shi *et al.*, Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* **368**, 1016-1020 (2020).

8. S. F. Sia *et al.*, Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583**, 834-838 (2020).

9. T. H. C. Sit *et al.*, Infection of dogs with SARS-CoV-2. *Nature*, (2020).

10. N. Oreshkova *et al.*, SARS-CoV2 infection in farmed mink, Netherlands, April 2020. *biorxiv*, (2020).

11. B. B. Oude Munnink *et al.*, Jumping back and forth: anthropozoonotic and zoonotic transmission of SARS-CoV-2 on mink farms. *biorxiv*, (2020).

12. V. J. Munster *et al.*, Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature* **585**, 268-272 (2020).

13. K. J. Olival *et al.*, Possibility for reverse zoonotic transmission of SARS-CoV-2 to free-ranging wildlife: A case study of bats. *PLoS Pathog* **16**, e1008758 (2020).

14. J. M. Santini, S. J. L. Edwards, Host range of SARS-CoV-2 and implications for public health. *The Lancet Microbe* **1**, e141-e142 (2020).

15. J. Damas *et al.*, Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc Natl Acad Sci U S A* **117**, 22311-22322 (2020).

16. A. D. Melin, M. C. Janiak, F. Marrone, 3rd, P. S. Arora, J. P. Higham, Comparative ACE2 variation and primate COVID-19 risk. *Commun Biol* **3**, 641 (2020).

17. X. Zhai *et al.*, Comparison of Severe Acute Respiratory Syndrome Coronavirus 2 Spike Protein Binding to ACE2 Receptors from Human, Pets, Farm Animals, and Putative Intermediate Hosts. *J Virol* **94**, (2020).

18. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).

19. W. Tai *et al.*, Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol Immunol* **17**, 613-620 (2020).

20. W. Li *et al.*, Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J* **24**, 1634-1643 (2005).

431   21.   J. Shang *et al.*, Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**,
432   221-224 (2020).
433   22.   A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike
434   Glycoprotein. *Cell* **181**, 281-292 e286 (2020).
435   23.   R. Yan *et al.*, Structural basis for the recognition of SARS-CoV-2 by full-length human
436   ACE2. *Science* **367**, 1444-1448 (2020).
437   24.   K. K. Chan *et al.*, Engineering human ACE2 to optimize binding to the spike protein of
438   SARS coronavirus 2. *Science*, (2020).
439   25.   M. J. Harms, J. W. Thornton, Evolutionary biochemistry: revealing the historical and
440   physical causes of protein properties. *Nat Rev Genet* **14**, 559-571 (2013).
441   26.   D. L. Stern, The genetic causes of convergent evolution. *Nat Rev Genet* **14**, 751-764
442   (2013).
443   27.   J. F. Storz, Causes of molecular convergence and parallelism in protein evolution. *Nat*
444   *Rev Genet* **17**, 239-250 (2016).
445   28.   M. A. DePristo, D. M. Weinreich, D. L. Hartl, Missense meanderings in sequence space:
446   a biophysical view of protein evolution. *Nat Rev Genet* **6**, 678-687 (2005).
447   29.   D. N. Ivankov, A. V. Finkelstein, F. A. Kondrashov, A structural perspective of
448   compensatory evolution. *Curr Opin Struct Biol* **26**, 104-112 (2014).
449   30.   N. Tokuriki, D. S. Tawfik, Stability effects of mutations and protein evolvability. *Curr*
450   *Opin Struct Biol* **19**, 596-604 (2009).
451   31.   C. Pal, B. Papp, Evolution of complex adaptations in molecular systems. *Nat Ecol Evol* **1**,
452   1084-1092 (2017).
453   32.   A. C. Palmer *et al.*, Delayed commitment to evolutionary fate in antibiotic resistance
454   fitness landscapes. *Nat Commun* **6**, 7385 (2015).
455   33.   P. Shah, D. M. McCandlish, J. B. Plotkin, Contingency and entrenchment in protein
456   evolution under purifying selection. *Proc Natl Acad Sci U S A* **112**, E3226-3235 (2015).
457   34.   T. N. Starr, L. K. Picton, J. W. Thornton, Alternative evolutionary histories in the
458   sequence space of an ancient protein. *Nature* **549**, 409-413 (2017).
459   35.   N. C. Wu, L. Dai, C. A. Olson, J. O. Lloyd-Smith, R. Sun, Adaptation in protein fitness
460   landscapes is facilitated by indirect paths. *Elife* **5**, (2016).
461   36.   Z. R. Sailer, M. J. Harms, Molecular ensembles make evolution unpredictable. *Proc Natl*
462   *Acad Sci U S A* **114**, 11938-11943 (2017).
463   37.   N. E. Clarke, A. J. Turner, Angiotensin-converting enzyme 2: the first decade. *Int J*
464   *Hypertens* **2012**, 307315 (2012).
465   38.   A. M. South, D. I. Diz, M. C. Chappell, COVID-19, ACE2, and the cardiovascular
466   consequences. *Am J Physiol Heart Circ Physiol* **318**, H1084-H1090 (2020).
467   39.   R. A. S. Santos *et al.*, The ACE2/Angiotensin-(1-7)/MAS Axis of the Renin-Angiotensin
468   System: Focus on Angiotensin-(1-7). *Physiol Rev* **98**, 505-553 (2018).
469   40.   M. A. Crackower *et al.*, Angiotensin-converting enzyme 2 is an essential regulator of
470   heart function. *Nature* **417**, 822-828 (2002).
471   41.   V. B. Patel *et al.*, ACE2 Deficiency Worsens Epicardial Adipose Tissue Inflammation
472   and Cardiac Dysfunction in Response to Diet-Induced Obesity. *Diabetes* **65**, 85-95
473   (2016).
474   42.   C. Zhang *et al.*, Angiotensin-converting enzyme 2 attenuates atherosclerotic lesions by
475   targeting vascular cells. *Proc Natl Acad Sci U S A* **107**, 15886-15891 (2010).

476 43. K. B. Pedersen, S. Sriramula, K. H. Chhabra, H. Xia, E. Lazartigues, Species-specific
477   inhibitor sensitivity of angiotensin-converting enzyme 2 (ACE2) and its implication for
478   ACE2 activity assays. *Am J Physiol Regul Integr Comp Physiol* **301**, R1293-1299 (2011).

479 44. F. E. Hauser, B. S. Chang, Insights into visual pigment adaptation and diversity from
480   model ecological and evolutionary systems. *Curr Opin Genet Dev* **47**, 110-120 (2017).

481 45. S. R. Tipnis *et al.*, A human homolog of angiotensin-converting enzyme. Cloning and
482   functional expression as a captopril-insensitive carboxypeptidase. *J Biol Chem* **275**,
483   33238-33243 (2000).

484 46. Y. Li *et al.*, SARS-CoV-2 and Three Related Coronaviruses Utilize Multiple ACE2
485   Orthologs and Are Potently Blocked by an Improved ACE2-Ig. *J Virol* **94**, (2020).

486 47. J. Echave, S. J. Spielman, C. O. Wilke, Causes of evolutionary rate variation among
487   protein sites. *Nat Rev Genet* **17**, 109-121 (2016).

488 48. S. L. Pond, S. D. Frost, S. V. Muse, HyPhy: hypothesis testing using phylogenies.
489   *Bioinformatics* **21**, 676-679 (2005).

490 49. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**,
491   1586-1591 (2007).

492 50. K. S. Siddiqui, R. Cavicchioli, Cold-adapted enzymes. *Annu Rev Biochem* **75**, 403-433
493   (2006).

494 51. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines,
495   Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819 (2017).

496 52. G. M. Castiglione, B. S. Chang, Functional trade-offs and environmental variation shaped
497   ancient trajectories in the evolution of dim-light vision. *Elife* **7**, (2018).

498 53. B. Steinberg, M. Ostermeier, Environmental changes bridge evolutionary valleys. *Sci Adv*
499   **2**, e1500921 (2016).

500 54. D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, Darwinian evolution can
501   follow only very few mutational paths to fitter proteins. *Science* **312**, 111-114 (2006).

502 55. N. Lu *et al.*, ACE2 gene polymorphism and essential hypertension: an updated meta-
503   analysis involving 11,051 subjects. *Molecular Biology Reports* **39**, 6581-6589 (2012).

504 56. Y. Luo *et al.*, Association of ACE2 genetic polymorphisms withhypertension-related
505   target organ damages in south Xinjiang. *Hypertension Research* **42**, 681-689 (2019).

506 57. B. S. W. Chang, K. Jönsson, M. A. Kazmi, M. J. Donoghue, T. P. Sakmar, Recreating a
507   Functional Ancestral Archosaur Visual Pigment. *Molecular Biology and Evolution* **19**,
508   1483-1489 (2002).

509 58. P. Liu, X. Xie, L. Gao, J. Jin, Designed variants of ACE2-Fc that decouple anti-SARS-
510   CoV-2 activities from unwanted cardiovascular effects. *Int J Biol Macromol* **165**, 1626-
511   1633 (2020).

512 59. M. J. Moore *et al.*, Retroviruses pseudotyped with the severe acute respiratory syndrome
513   coronavirus spike protein efficiently infect cells expressing angiotensin-converting
514   enzyme 2. *J Virol* **78**, 10628-10635 (2004).

515 60. G. Hütter *et al.*, Long-Term Control of HIV by CCR5 Delta32/Delta32 Stem-Cell
516   Transplantation. *New England Journal of Medicine* **360**, 692-698 (2009).

517 61. C. A. Devaux, J.-M. Rolain, D. Raoult, ACE2 receptor polymorphism: Susceptibility to
518   SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome.
519   *Journal of Microbiology, Immunology and Infection* **53**, 425-435 (2020).

520 62. E. W. Stawiski *et al.*, Human ACE2 receptor polymorphisms predict SARS-CoV-2
521   susceptibility. *bioRxiv*, 2020.2004.2007.024752 (2020).

63. J. A. de Visser, J. Krug, Empirical fitness landscapes and the predictability of evolution. *Nat Rev Genet* **15**, 480-490 (2014).

64. F. J. Poelwijk, D. J. Kiviet, D. M. Weinreich, S. J. Tans, Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383-386 (2007).

65. T. N. Starr, J. W. Thornton, Exploring protein sequence-function landscapes. *Nat Biotechnol* **35**, 125-126 (2017).

66. M. Gheblawi *et al.*, Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System: Celebrating the 20th Anniversary of the Discovery of ACE2. *Circ Res* **126**, 1456-1474 (2020).

67. A. Zoufaly *et al.*, Human recombinant soluble ACE2 in severe COVID-19. *The Lancet Respiratory Medicine* **8**, 1154-1158 (2020).

68. J. D. Jensen *et al.*, The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern and Hahn 2018. *Evolution* **73**, 111-114 (2019).

69. M. Nei, Y. Suzuki, M. Nozawa, The neutral theory of molecular evolution in the genomic era. *Annu Rev Genomics Hum Genet* **11**, 265-289 (2010).

70. R. Lanfear, H. Kokko, A. Eyre-Walker, Population size and the rate of evolution. *Trends Ecol Evol* **29**, 33-41 (2014).

71. E. S. Brielle, D. Schneidman-Duhovny, M. Linial, The SARS-CoV-2 Exerts a Distinctive Strategy for Interacting with the ACE2 Human Receptor. *Viruses* **12**, (2020).

72. A. Banerjee, K. Kulcsar, V. Misra, M. Frieman, K. Mossman, Bats and Coronaviruses. *Viruses* **11**, (2019).

73. W. Li *et al.*, Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676-679 (2005).

74. H. K. Frank, D. Enard, S. D. Boyd, Exceptional diversity and selection pressure on SARS-CoV and SARS-CoV-2 host receptor in bats compared to other mammals. *bioRxiv*, 2020.2004.2020.051656 (2020).

75. A. Latinne *et al.*, Origin and cross-species transmission of bat coronaviruses in China. *Nat Commun* **11**, 4235 (2020).

76. M. Canals, J. Iriarte-Diaz, B. Grossi, in *Biomechanics in Applications*. (2011), chap. Chapter 13.

77. G. Neuweiler, *The Biology of Bats*. (Oxford University Press, 2000), pp. 310.

78. V. Monteil *et al.*, Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* **181**, 905-913 e907 (2020).

79. J. Guo, Z. Huang, L. Lin, J. Lv, Coronavirus Disease 2019 (COVID-19) and Cardiovascular Disease: A Viewpoint on the Potential Influence of Angiotensin-Converting Enzyme Inhibitors/Angiotensin Receptor Blockers on Onset and Severity of Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *J Am Heart Assoc* **9**, e016219 (2020).

80. R. M. Inciardi *et al.*, Cardiac Involvement in a Patient With Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol*, (2020).

81. B. Siripanthong *et al.*, Recognizing COVID-19-related myocarditis: The possible pathophysiology and proposed guideline for diagnosis and management. *Heart Rhythm*, (2020).

82. D. W. Wong *et al.*, Loss of Angiotensin-Converting Enzyme-2 (Ace2) Accelerates Diabetic Kidney Injury. *The American Journal of Pathology* **171**, 438-451 (2007).

83. S. Epelman *et al.*, Soluble angiotensin-converting enzyme 2 in human heart failure: relation with myocardial function and clinical outcomes. *J Card Fail* **15**, 565-571 (2009).

84. W. H. Li *et al.*, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).

85. K. K. Chan *et al.*, Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **369**, 1261-1265 (2020).

86. S. K. Wong, W. Li, M. J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem* **279**, 3197-3201 (2004).

87. J. K. Millet, G. R. Whittaker, Murine Leukemia Virus (MLV)-based Coronavirus Spike-pseudotyped Particle Production and Infection. *Bio-protocol* **6**, e2035 (2016).

88. A. Loytynoja, N. Goldman, webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* **11**, 579 (2010).

89. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321 (2010).

90. V. Lefort, J.-E. Longueville, O. Gascuel, SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution* **34**, 2422-2424 (2017).

91. S. B. Hedges, J. Marin, M. Suleski, M. Paymer, S. Kumar, Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* **32**, 835-845 (2015).

92. B. Murrell *et al.*, FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* **30**, 1196-1205 (2013).

93. J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B. Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* **44**, W232-W235 (2016).

94. J. P. Bielawski, Z. Yang, A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol* **59**, 121-132 (2004).

95. C. R. White, R. S. Seymour, The role of gravity in the evolution of mammalian blood pressure. *Evolution* **68**, 901-908 (2014).

96. R. Tacutu *et al.*, Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Res* **46**, D1083-D1090 (2018).

97. W. P. Maddison, D. R. Maddison. (http://www.mesquiteproject.org, 2018).

98. T. Garland, A. R. Ives, Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. *American Naturalist* **155**, 346-364 (2000).

99. M. D. Smith *et al.*, Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection. *Molecular Biology and Evolution* **32**, 1342-1353 (2015).

Acknowledgements

612   **Competing Interests:** The authors declare they have no competing interests**.**

613   **Data and materials Availability:** All data is available in the manuscript or in the supplementary
614   materials

615   **List of Supplementary Materials**

616   **Materials and Methods**

617   **Table S1-S5**

618   **Fig S1-S6**

619   **References 86-99**

620

621   Supplementary Materials for:

622   The ancient cardioprotective mechanisms of ACE2 bestow SARS-CoV-2 with a wide host
623   range

624   G.M. Castiglione et al.

Materials and Methods

ACE2 and SARS-CoV-2 constructs

Wild-type hACE2 with a 1D4 (C9) C-terminal tag (TETSQVAPA) was a gift from Hyeryun Choe (Addgene plasmid # 1786 ; http://n2t.net/addgene:1786 ; RRID:Addgene_1786) (*84*). Site-directed mutagenesis primers were designed to induce single amino acid substitutions via PCR (QuickChange II, Agilent). Mouse ACE2 was cloned from cDNA synthesized from an RNA extraction of ileum tissue from a C57/BL6 mouse sacrificed in compliance with all regulations of the Johns Hopkins University Institutional Animal Care and Use committee. ACE2 coding sequences from dog (XM_014111329.2), pangolin (*Manis javanica*; XM_017650263.1), bat (*Rhinolophus sinicus*; XM_019746337.1) and the rodent ancestor were synthesized as gblocks (Integrated DNA technologies). All animal ACE2 inserts were cloned into a pEGFP-N1 vector, with eGFP C-terminal tag. For interspecies comparisons, hACE2 was also cloned into pEGFP-N1. The receptor binding domain of the SARS-CoV-2 S protein was obtained as a pcDNA3-SARS-CoV-2-S-RBD-Fc plasmid, as a gift from Erik Procko (Addgene plasmid # 141183 ; http://n2t.net/addgene:141183 ; RRID:Addgene_141183) (*85*). pcDNA3.1-SARS-Spike was a gift from Fang Li (Addgene plasmid # 145031 ; http://n2t.net/addgene:145031 ; RRID:Addgene_145031) (*21*). We used this plasmid as a template to clone the RBD of SARS-CoV-1 Spike protein into pcDNA3 with a Fc C-terminal tag. This RBD consisted of residues 318 to 510, as previously described (numbering as per AAP13441.1) (*86*).

Flow cytometry and immunoprecipitation

HEK293T cells were transfected with ACE2-1D4, or SARS-CoV-2 S protein RBD-Fc constructs using TransIT-X2 (Mirus). 24hrs after transfection the media of RBD-transfected cells was replaced with OptiPRO SFM (ThermoFisher). 72hrs after transfection, the media from RBD-transfected cells was concentrated using an Amicon Ultra-15 centrifugal filter unit with a 3000 kDa molecular weight cutoff. In parallel, ACE2-transfected cells were harvested and washed with PBS. For flow cytometry, 5.0×10^5 cells were resuspended in 1mL incubation buffer [Dulbecco's PBS containing 0.02% EDTA (Sigma-Aldrich), 50µg/mL DNase I (Worthington) and 5mM $MgCl_2$] and incubated with 20 ug/mL of RBD-Fc for 30 minutes at room temperature as previously described (*19*). Cells were then washed with buffer (5% FBS, 0.1% Sodium Azide in PBS) and incubated with human ACE2 Alexa Fluro 647-conjugated antibody ($1µg/10^6$ cell, #FAB9332R, R&D Systems), human IgG Fc PE-conjugated antibody ($10µg/10^6$ cell, #FAB110P, R&D Systems) or human IgG Fc APC-conjugated antibody ($10µg/10^6$ cell, #FAB110A, R&D Systems) for 30 minutes at room temperature. LSR II (BD bioscience) was used to collect the data and Flowjo (Flowjo, LLC.) was used for analysis, conducted as previously described (*19*). For immunoprecipitation, ACE2-transfected cells were lysed in a PBS buffer containing 1% CHASPO (Sigma-Aldrich), and incubated with Dynabeads Protein G (ThermoFisher) and 2µg of RBD-Fc concentrate. Dynabeads were washed with PBS (0.5% CHAPSO), and elutions were immunoblotted using antibodies against 1D4 (Abcam; ab5417) and Human IgG Fc (Abcam; ab97225).

Pseudovirus Assay

665 SARS-CoV-2 spike pseudotyped murine leukemia virus (MLV) were generated based on a
666 published protocol (*87*). Briefly, HEK293T cells were transfected with three plasmids, which
667 express MLV Gag and Pol, firefly luciferase reporter and SARS-CoV-2 S protein. 48 h after
668 transfection, culture medium containing pseudotyped particles were centrifuged and then filtered
669 through a sterile 0.45 µm pore-sized filter to remove cell debris. For transduction, HEK293T
670 cells were first transfected with ACE2 orthologs using lipofectamine 3000. 24 hours later,
671 200µL of SARS-CoV-2 Spike pseudotyped MLV were added to each well and incubated for 2
672 days. The transduction efficiency was quantified by measuring the activity of luciferase using
673 luciferase assay system (Promega) and GloMax 20/20 luminometer (Promega).

674 *ACE2 hydrolysis assay*

675 HEK293T cells were transfected with ACE2 constructs using TransIT-X2 (Mirus). 72hrs later
676 cells were washed in PBS and lysed in ACE2 reaction buffer pH 6.5 (1M NaCl, 0.5% Triton X-
677 100, 0.5mM $ZnCl_2$, 75mM Tris-HCl). Cell lysates were diluted in reaction buffer to 0.5µg
678 protein, incubated ± the ACE2-specific inhibitor 10µM DX600 (Cayman Chemical) for 20
679 minutes at room temperature, followed by addition of 100µM Mca-YVADAPK(Dnp) (R&D
680 Systems) and incubation at 37°C. Fluorescence emission at 405 nm was measured at 10, 60, and
681 120 minutes using a microplate reader (BMG Labtech) after excitation at 320 nm. Hydrolysis
682 rates were quantified as fluorescence units per minute, using the slope of fluorescence
683 development between 10 and 120 minutes, as previously described (*43*). An ANOVA general
684 linear model with fluorescence as the response variable, mutation as the factor, and assay time as
685 the covariate was fit to fluorescence data generated in the ACE2 hydrolysis assay. Statistical
686 differences in ACE2 hydrolysis rates were determined using a cross factor between the mutation
687 factor and time covariate.

688 *Phylogenetic Comparative Methods*

689 ACE and ACE2 coding sequences from 107 species representing all major jawed vertebrate
690 lineages were obtained from GenBank (Table S4), and aligned using PRANK followed by
691 manual adjustment (*88*). This alignment was used to estimate a gene tree using PhyML 3.1 (*89*)
692 (Fig. S5), with GTR selected using automatic model selection based on AIC values (*90*), and
693 aLRT SH-like branch support. This ML tree was rooted using *ACE*, and recapitulated all major
694 phylogenetic relationships (Fig. S5 (*91*)). Ancestral sequences and posterior probability
695 distributions were inferred using the best fitting models in the codeml package of PAML 4.9
696 (Table S5) (*49*). For estimation of dN/dS values using random sites models in the codeml
697 package (*49*) and FUBAR (*92*), ACE sequences were pruned from the dataset, non-mammalian
698 *ACE2* was also pruned, and additional mammalian *ACE2* sequences added to increase sampling.
699 This alignment represented 89 species (Table S1), and was used to infer a ML gene tree using
700 IQ-Tree (Fig. S2), with the substitution model auto selected, followed by ultrafast bootstrap
701 analysis and SH-aLRT branch tests (*93*). PAML random sites models were used to investigate
702 evidence of positive selection (Table S2). To test for dN/dS differences amongst branches in the
703 phylogeny, clade model D (CmD) (*94*) was used to analyze an *ACE2* alignment containing only
704 those sites of interest in the ACE2 viral binding interface (Table S3). M3 with three site classes

705   was used as the null model for CmD. All random sites and clade model PAML model pairs were
706   statistically evaluated for significance by likelihood ratio tests (LRT) with a $\chi 2$ distribution.
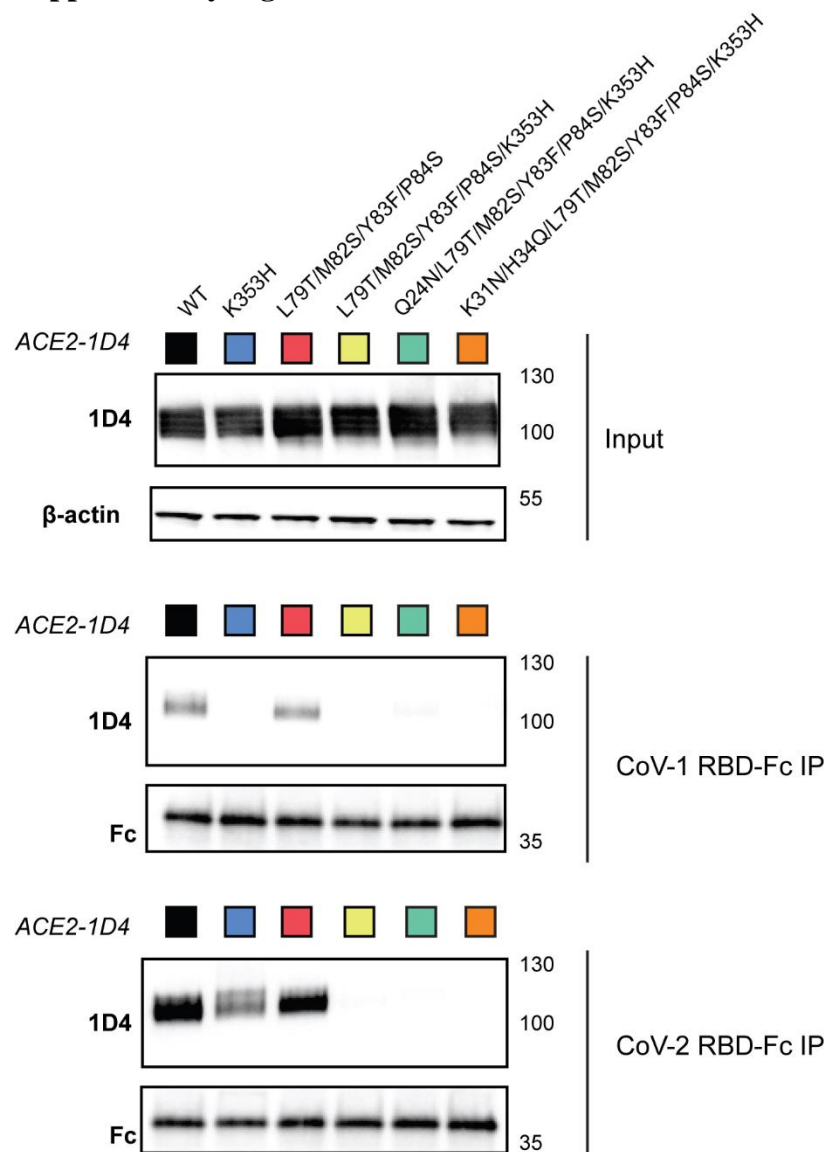
707   We conducted a phylogenetic comparative analysis on systolic blood pressure dataset by
708   combining bat data with mammalian data compiled from a previous study (*77, 95*) . We pruned
709   this dataset to include only species with fossil calibrated divergence times (*91*) (Fig. S6). Body
710   size values (grams) were obtained from the Ageing Genomic Resources AnAge database (*96*).
711   We conducted a phylogenetically independent correlation analysis (least squares linear
712   regression; Fig. S6) by using the PDTREE program of the PDAP module of MESQUITE (*97,*
713   *98*) to calculate phylogenetically independent contrasts of $\log_{10}$ body mass (grams) systolic blood
714   pressures (mmHg) as the dependent variable. Before independent contrasts were calculated,
715   branch lengths reflecting divergence times were ln transformed to meet the assumptions of
716   independent contrast analysis (*97, 98*). To produce dN/dS estimates for branches in the
717   untransformed phylogeny, we pruned the mammalian *ACE2* alignment to match the species
718   represented in the blood pressure dataset, and subjected the phylogeny and the alignment to
719   analysis by aBSREL(*99*).

720

721   **Supplementary Figure S1.**

722
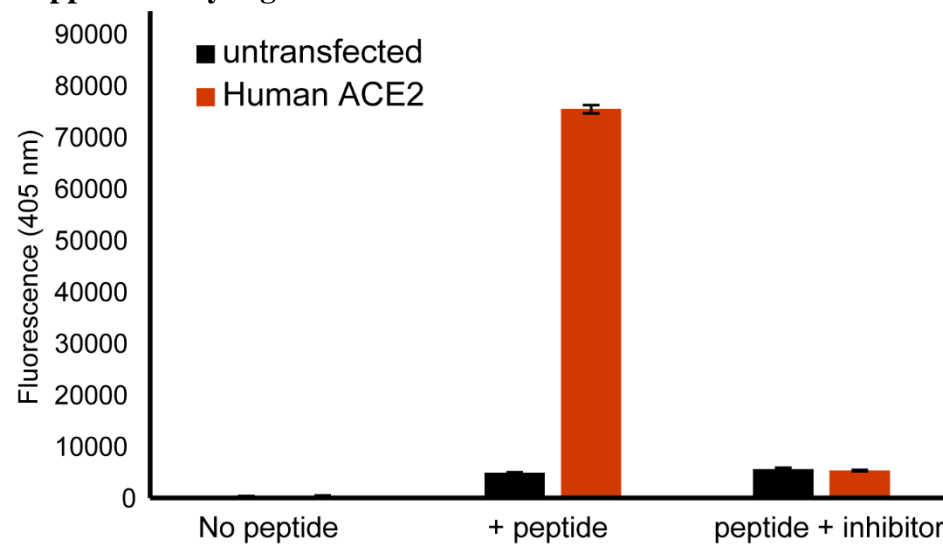
723   **Recombinant ACE2 hydrolysis activity in solubilized transfected HEK293T cells.** ACE2
724   hydrolysis activity was measured using a fluorogenic peptide substrate (Mca-YVADAPK(Dnp)-
725   OH) incubated with lysates of HEK293T cells for 2 hours. Cells were either untransfected or
726   were transfected with a human ACE2 construct. Incubation of lysate-peptide mixture with a
727   ACE2-specific inhibitor (DX600) reduced fluorescence attributable to ACE2 hydrolysis activity.

728

729 **Supplementary Figure S2.**

730 Maximum likelihood phylogeny used in PAML analyses. aLRT-SH like branch support values

731 (PhyML) are shown

732

733

734

735   **Supplementary Figure S3.**
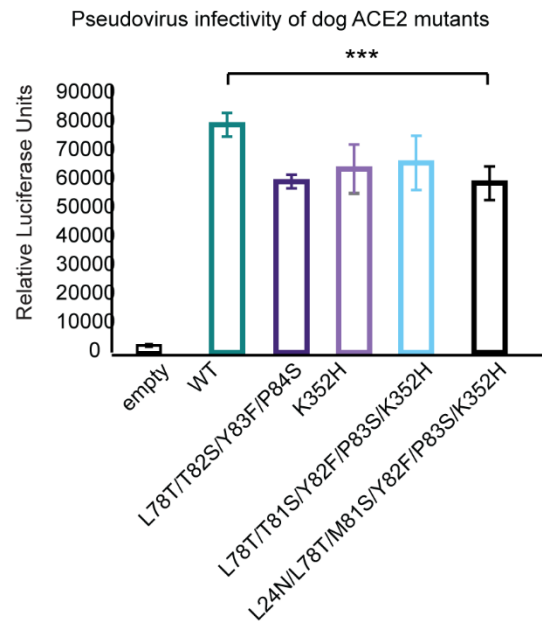

736

737   **Targeted mutations to human ACE2 disrupt binding to the RBD of SARS-CoV-1 and**
738   **SARS-CoV-2.** Western blots of immunopreciptations and cell lysates of HEK293T cells co-
739   transfected with an Fc-tagged SARS-CoV-2 S protein RBD, and 1D4-tagged (C9) human ACE2
740   construct.

741

Supplementary Figure S4.
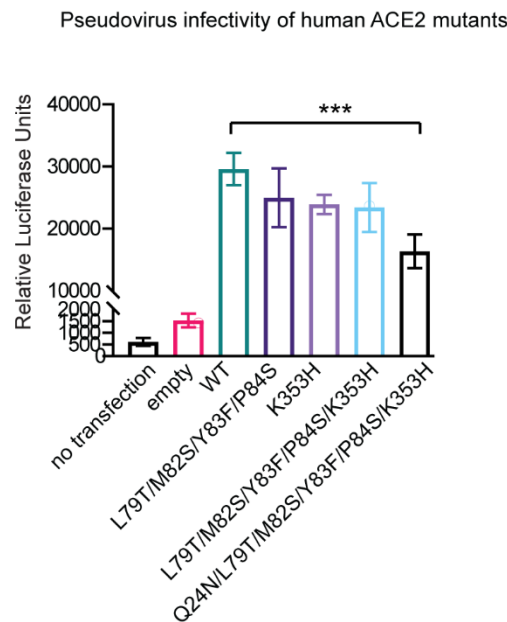
Effect of mouse substitutions on pseudovirus infection of cells expressing human and dog ACE2. Infection of HEK293T cells transfected with either (**A**) human or (**B**) dog ACE2, containing the indicated mutations. Cells were exposed to VSV-G pseudotyped with SARS-CoV-2 S protein, and cellular infection measured as a function of luciferase luminescence. (**C**) ACE2 hydrolysis activity of dog ACE2 with indicated mutations. ACE2 activity was measured as fluorescence units per minute.

751 **Supplementary Figure S5.**

752 Maximum likelihood phylogeny used in ancestral reconstruction of Rodent ACE2. aLRT-SH like

753 branch support values (IQ-Tree) are shown

754
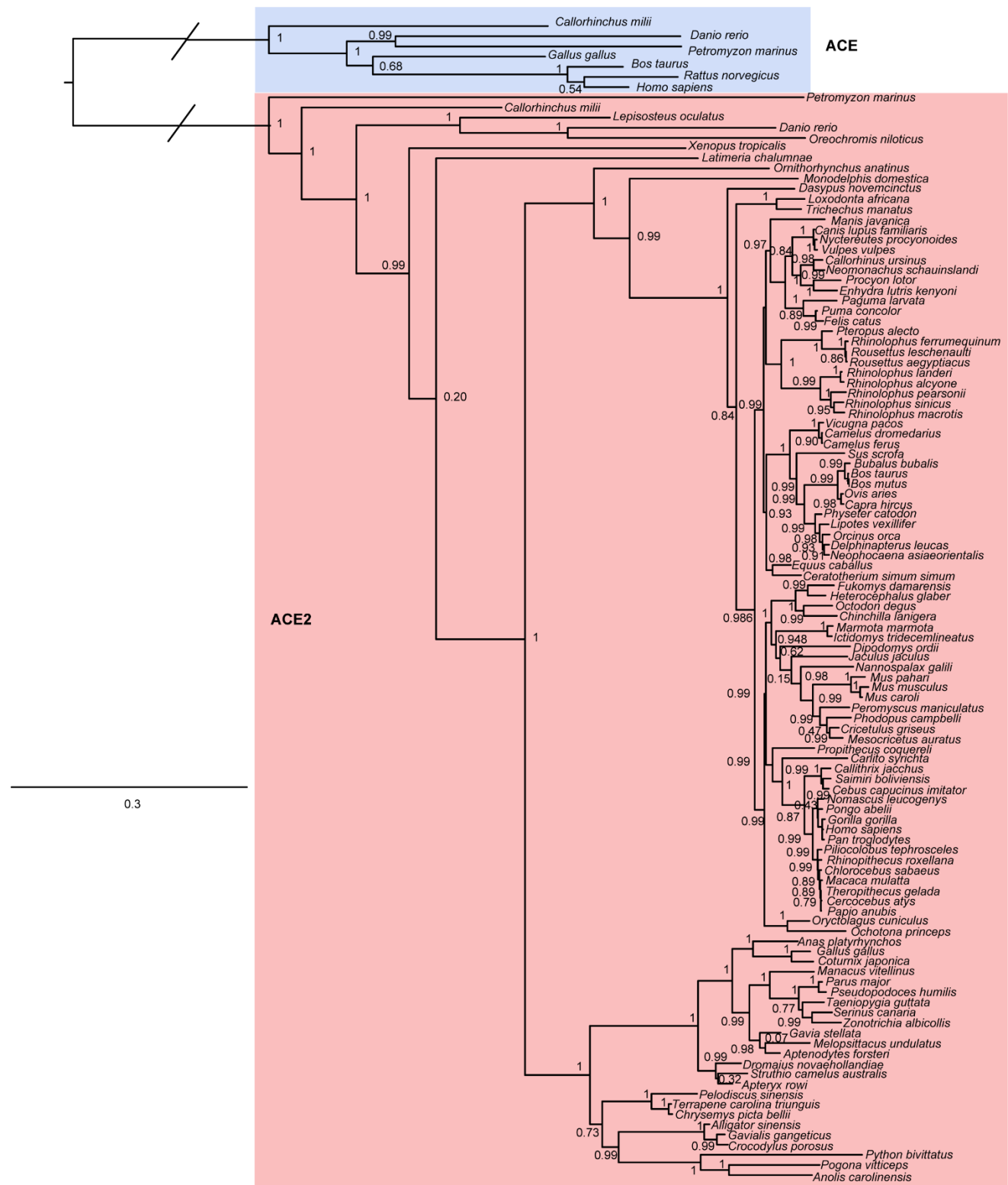
755 **Supplementary Figure S6.**

756 Species phylogeny and least squares linear regression using phylogenetically independent

757 contrasts of systolic blood pressure and body mass

758

759
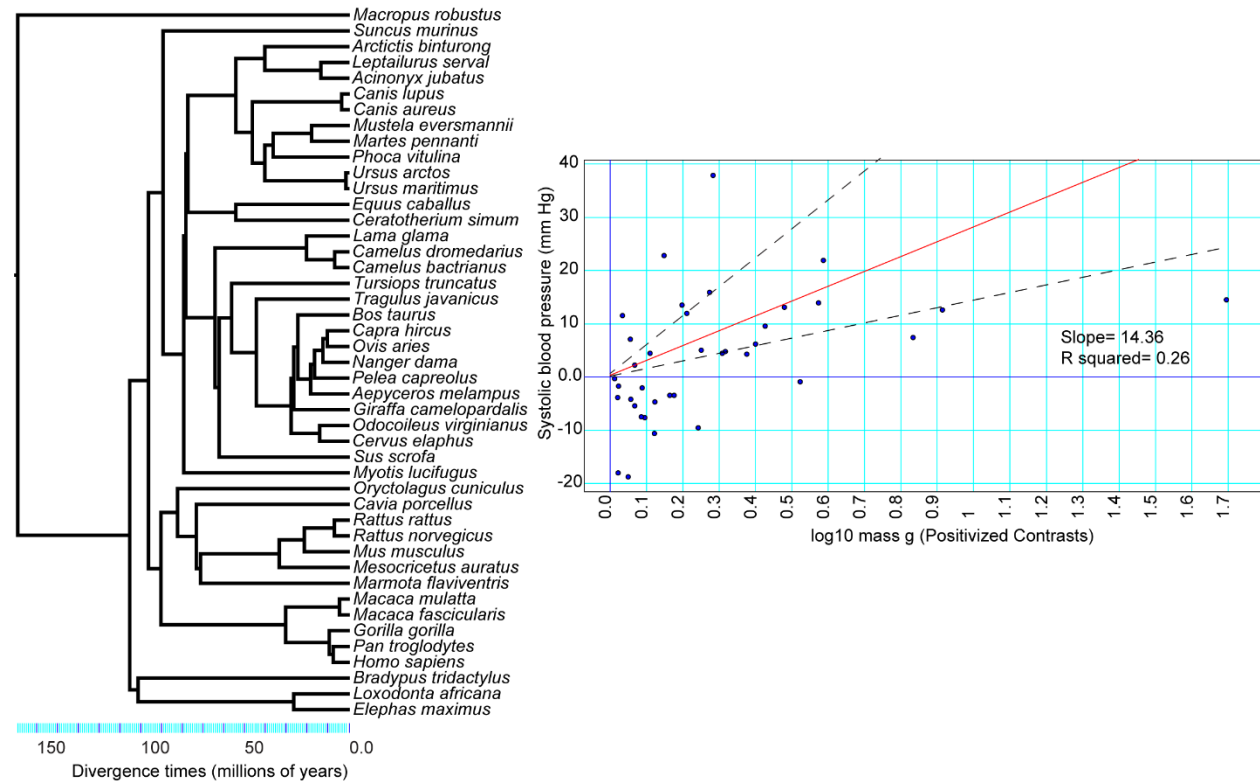
760 **Supplementary Table 1. ACE2 accession numbers used in dN/dS estimates**

scientific name	accession
Ornithorhynchus_anatinus	XM_001515547.4
Monodelphis domestica	XM_007500874.2
Dasypus novemcinctus	XM_004449067.3
Loxodonta africana	XM_023555192.1
Trichechus manatus latirostris	XM 004386324.2
Manis javanica	XM 017650263.1
Paguma larvata	AY881174.1
Felis catus	AB211997.1
Puma concolor	XM 025934632.1
Procyon lotor	AB211998
Enhydra lutris kenyoni	XM 022518370.1
Ailuropoda melanoleuca	XM 002930611.3
Ursus maritimus	XM 008696415.1
Neomonachus schauinslandi	XM 021680805.1
Callorhinus ursinus	XM 025857612.1
Nyctereutes procyonoides	EU024940.1
Vulpes vulpes	XM 025986727.1
Canis lupus familiaris	XM 014111329.2
Canis lupus dingo	XM 025437140.1
Eptesicus fuscus	XM 008154928.2
Hipposideros armiger	XM 019667391.1
Rhinolophus landeri	KR559015.1
Rhinolophus alcyone	KR559016.1
Rhinolophus pearsonii	EF569964.1
Rhinolophus sinicus	GQ262791.1
Rhinolophus macrotis	GQ999932.1
Rhinolophus pusillus	GQ999938.1
Rhinolophus ferrumequinum	FJ598617.1
Rousettus leschenaulti	AB299376.1
Rousettus aegyptiacus	XM 016118926.1
Pteropus vampyrus	XM 011362973.2
Pteropus alecto	XM 006911647.1
Sus scrofa	GQ262781.1
Bubalus bubalis	XM 006041540.2
Bos mutus	XM 005903111.1
Bos taurus	BC105340.1
Capra hircus	XM 005701072.3
Ovis aries	XM 012106267.3
Physeter catodon	XM 024115511.2

Lipotes vexillifer	XM 007466327.1
Orcinus orca	XM 004269657.1
Delphinapterus leucas	XM 022562652.2
Neophocaena asiaeorientalis	XM 024744126.1
Vicugna pacos	XM 006212647.3
Camelus ferus	XM 006194201.2
Camelus dromedarius	XM 010993415.2
Ceratotherium simum simum	XM 004435149.2
Equus caballus	XM 001490191.5
Equus przewalskii	XM 008544773.1
Propithecus coquereli	XM 012638732.1
Carlito syrichta	XM 008064619.1
Callithrix jacchus	XM 017968359.1
Saimiri boliviensis	XM 010336623.1
Cebus capucinus imitator	XM 017512376.1
Nomascus leucogenys	XM 003261084.3
Pongo abelii	NM 001131132.2
Gorilla gorilla gorilla	XM 019019204.1
Homo sapiens	NM 001371415.1
Pan paniscus	XM 008974180.1
Pan troglodytes	XM 016942979.1
Piliocolobus tephrosceles	XM 023199053.2
Rhinopithecus roxellana	XM 010366065.2
Chlorocebus sabaeus	XM 007991113.1
Papio anubis	XM 021933040.1
Macaca mulatta	XM 015126958.2
Macaca nemestrina	XM 011735203.2
Cercocebus atys	XM 012035808.1
Theropithecus gelada	XM 025372062.1
Oryctolagus cuniculus	GQ262787.1
Ochotona princeps	XM 004597492.2
Dipodomys ordii	XM 013032118.1
Heterocephalus glaber	XM 004866100
Fukomys damarensis	XM 010645175.2
Chinchilla lanigera	NM 001282361.1
Octodon degus	XM 023719547.1
Ictidomys tridecemlineatus	XM 005315994.3
Marmota marmota	XM 015488054.1
Jaculus jaculus	XM 004671466.2
Nannospalax galili	XM 008840876.2
Phodopus campbelli	GQ262790.1

Mesocricetus auratus	GQ262794.1
Cricetulus griseus	XM 003503235.4
Peromyscus maniculatus bairdii	XM 006973207.2
Mus caroli	XM 021153479.2
Mus pahari	XM 021188276.2
Mus musculus	NM 001130513.1

761

762

Supplementary Table 2.

Analyses of selection on Mammalian *ACE2* using PAML random sites models.

Model	ΔAIC[1]	*ln*L	Parameters[2]			Null	*p* [df][3]
			ω_0/p	ω_1/q	ω_2/ω_p		
M0	3746.8	-36519.69	0.31	-	-	N/A	-
M1a	569.2	-34929.93	0.08 (67%)	1.00 (33%)	-	M0	**0.000** [1]
M2a	369.6	-34828.09	0.08 (66%)	1.00 (29%)	2.39 (5%)	M1a	**0.000** [2]
M3	110.3	-34697.44	0.04 (57%)	0.48 (31%)	1.51 (12%)	M2	**0.000** [1]
M7	186.8	-34738.71	0.24	0.55	-	N/A	-
M8a	141.3	-34714.97	0.35	1.66	1.00	N/A	-
M8	**0.0***	-34643.31	0.29	0.87	1.84	M7	**0.000** [2]
						M8a	**0.000** [1]

[1]All ΔAIC values are calculated from the lowest AIC model. The best fits are bolded with an asterisk (*).

[2]ω values of each site class are shown are shown for model M0-M3 (ω_0– ω_2) with the proportion of each site class in parentheses. For M7 and M8, the shape parameters, p and q, which describe the beta distribution are listed instead. In addition, the ω value for the positively selected site class (ω_p, with the proportion of sites in parentheses) is shown for M8.

[3]Significant *p*-values (α ≤ 0.05) are bolded. Degrees of freedom are given in square brackets after the *p*-values. Significance was determined through a likelihood-ratio test of null and alternative models, with reference to a χ^2 distribution.

Abbreviations—*ln*L, ln Likelihood; *p*, *p*-value; **N/A**, not applicable.

Supplementary Table 3.

Results of Clade Model D (CmD) analyses of Mammalian *ACE2* under various partitions.

Model & Foreground[1]	ΔAIC[2]	lnL	Parameters[3]			Null	p [df][4]
			ω_0	ω_1	ω_2/ω_d		
M3	6.51	-722.34	0.00 (0%)	0.14 (25%)	1.17 (75%)	N/A	-
CmD_*Bats*	4.31	-720.24	0.00 (0%)	0.14 (25%)	0.97 (75%) Bats: 2.44	M3	**0.04** [1]
CmD_*Rodents*	1.66	-718.92	0.00 (0%)	0.14 (25%)	1.63 (75%) Rodents: 0.52	M3	**0.000** [1]
CmD_*Primates*	5.63	-720.90	0.00 (0%)	0.14 (25%)	1.31 (75%) Primates: 0.39	M3	0.09 [1]
CmD_*RodentsBats*	1.91	-718.04	0.00 (0%)	0.13 (25%)	1.37 (75%) Rodents: 0.52 Bats: 2.44	M3	**0.03** [2]
CmD_*RodentsBatsPrimates*	**0.00***	-716.07	0.00 (0%)	0.13 (25%)	1.77 (75%) Rodents: 0.52 Bats: 2.44 Primates: 0.42	M3	**0.006** [3]

[1]The foreground partition is listed after the underscore for the clade models and consists of either: the clade of 'Bats, 'Rodents', or the clade 'Primates'. In any partitioning scheme, the entire clade was tested, and all non-foreground data are present in the background partition.

[2]All ΔAIC values are calculated from the lowest AIC model. The best fits are bolded with an asterisk (*).

[3]ω_d is the divergent site class, which has a separate value for the foreground and background partitions.

[4]Significant p-values ($\alpha \leq 0.05$) are bolded. Degrees of freedom are given in square brackets after the p-values. Significance was determined through a likelihood-ratio test of null and alternative models, with reference to a χ^2 distribution.

Abbreviations—*ln*L, ln Likelihood; *p*, *p*-value; **AIC**, Akaike information criterion

799 **Supplementary Table 4. ACE and ACE2 accession numbers used in ancestral**
800 **reconstruction**

scientific name	accession	Gene
Petromyzon marinus	XM 032951728.1	ACE
Callorhinchus milii	XM 007907341.1	
Danio rerio	XM 689244.9	
Gallus gallus	NM 001167732.1	
Bos taurus	NM 001206668.1	
Homo sapiens	NM 000789.4	
Rattus norvegicus	NM 012544.1	
Petromyzon marinus	XM 032979141.1	ACE2
Callorhinchus milii	XM 007891654.1	
Lepisosteus oculatus	XM 006639122.2	
Danio rerio	NM 001007297.1	
Latimeria chalumnae	XM 005997853.2	
Oreochromis niloticus	XM 005466811.4	
Xenopus tropicalis	XM 002938247.4	
Anolis carolinensis	XM 008107249.1	
Python bivittatus	XM 007431880.3	
Pogona vitticeps	XM 020786763.1	
Pelodiscus sinensis	XM 006122829.3	
Chrysemys picta bellii	XM 005287788.2	
Terrapene carolina triunguis	XM 026648166.1	
Alligator sinensis	XM 025210843.1	
Gavialis gangeticus	XM 019525515.1	
Crocodylus porosus	XM 019529281.1	
Struthio camelus australis	XM 009669200.1	
Dromaius novaehollandiae	XM 026120784.1	
Apteryx rowi	XM 026087161.1	
Coturnix japonica	XM 015886577.2	
Anas platyrhynchos	XM 013094461.3	
Gallus gallus	MK560199.1	
Taeniopygia guttata	XM 002194267.4	
Manacus vitellinus	XM 018084005.2	
Parus major	XM 015631329.1	
Serinus canaria	XM 009089674.3	
Aptenodytes forsteri	XM 009276865.1	
Gavia stellata	XM 009817825.1	
Melopsittacus undulatus	XM 005151459.1	
Zonotrichia albicollis	XM 005491775.2	
Pseudopodoces humilis	XM 005516655.1	

Ornithorhynchus_anatinus	XM_001515547.4
Monodelphis domestica	XM_007500874.2
Dasypus novemcinctus	XM_004449067.3
Loxodonta africana	XM_023555192.1
Trichechus manatus latirostris	XM 004386324.2
Manis javanica	XM 017650263.1
Felis catus	AB211997.1
Puma concolor	XM 025934632.1
Procyon lotor	AB211998
Enhydra lutris kenyoni	XM 022518370.1
Ursus maritimus	XM 008696415.1
Neomonachus schauinslandi	XM 021680805.1
Callorhinus ursinus	XM 025857612.1
Nyctereutes procyonoides	EU024940.1
Vulpes vulpes	XM 025986727.1
Canis lupus familiaris	XM 014111329.2
Eptesicus fuscus	XM 008154928.2
Rhinolophus landeri	KR559015.1
Rhinolophus alcyone	KR559016.1
Rhinolophus pearsonii	EF569964.1
Rhinolophus sinicus	GQ262791.1
Rhinolophus macrotis	GQ999932.1
Rhinolophus ferrumequinum	FJ598617.1
Rousettus leschenaulti	AB299376.1
Rousettus aegyptiacus	XM 016118926.1
Pteropus alecto	XM 006911647.1
Sus scrofa	GQ262781.1
Bubalus bubalis	XM 006041540.2
Bos mutus	XM 005903111.1
Bos taurus	BC105340.1
Capra hircus	XM 005701072.3
Ovis aries	XM 012106267.3
Physeter catodon	XM 024115511.2
Lipotes vexillifer	XM 007466327.1
Orcinus orca	XM 004269657.1
Delphinapterus leucas	XM 022562652.2
Neophocaena asiaeorientalis	XM 024744126.1
Vicugna pacos	XM 006212647.3
Camelus ferus	XM 006194201.2
Camelus dromedarius	XM 010993415.2
Ceratotherium simum simum	XM 004435149.2

Equus caballus	XM 001490191.5	
Propithecus coquereli	XM 012638732.1	
Carlito syrichta	XM 008064619.1	
Callithrix jacchus	XM 017968359.1	
Saimiri boliviensis	XM 010336623.1	
Cebus capucinus imitator	XM 017512376.1	
Nomascus leucogenys	XM 003261084.3	
Pongo abelii	NM 001131132.2	
Gorilla gorilla gorilla	XM 019019204.1	
Homo sapiens	NM 001371415.1	
Pan troglodytes	XM 016942979.1	
Piliocolobus tephrosceles	XM 023199053.2	
Rhinopithecus roxellana	XM 010366065.2	
Chlorocebus sabaeus	XM 007991113.1	
Papio anubis	XM 021933040.1	
Macaca mulatta	XM 015126958.2	
Macaca nemestrina	XM 011735203.2	
Cercocebus atys	XM 012035808.1	
Theropithecus gelada	XM 025372062.1	
Oryctolagus cuniculus	GQ262787.1	
Ochotona princeps	XM 004597492.2	
Dipodomys ordii	XM 013032118.1	
Heterocephalus glaber	XM 004866100	
Fukomys damarensis	XM 010645175.2	
Chinchilla lanigera	NM 001282361.1	
Octodon degus	XM 023719547.1	
Ictidomys tridecemlineatus	XM 005315994.3	
Marmota marmota	XM 015488054.1	
Jaculus jaculus	XM 004671466.2	
Nannospalax galili	XM 008840876.2	
Phodopus campbelli	GQ262790.1	
Mesocricetus auratus	GQ262794.1	
Cricetulus griseus	XM 003503235.4	
Peromyscus maniculatus bairdii	XM 006973207.2	
Mus caroli	XM 021153479.2	
Mus pahari	XM 021188276.2	
Mus musculus	NM 001130513.1	

801

802 **Supplementary Table 5.**

803 Results of random sites analyses of vertebrate *ACE2*, with the best fitting mode (M8) used for
804 the ancestral reconstruction of mammalian ACE2.

Model	ΔAIC	*ln*L	Parameters[2]			Null	*p* [df][3]
			ω_0/p	ω_1/q	ω_2/ω_p		
M3	231.6	-63042.66	0.02 (42%)	0.23 (33%)	0.78 (25%)	N/A	-
M7	46.32	-62953.03	0.35	0.84	-	N/A	-
M8a	**53.04**	-62955.39	0.49	1.84	1		**-**
M8	**0.0***	-62927.87	0.41	1.35	1.22	M7	**0.000** [2]
						M8a	**0.000** [1]

813 [1]All ΔAIC values are calculated from the lowest AIC model. The best fits are bolded with an asterisk (*).

814 [2]ω values of each site class are shown are shown for model M0-M3 (ω_0– ω_2) with the proportion of each
815 site class in parentheses. For M7 and M8, the shape parameters, p and q, which describe the beta
816 distribution are listed instead. In addition, the ω value for the positively selected site class (ω_p, with the
817 proportion of sites in parentheses) is shown for M8.

818 [3]Significant *p*-values ($\alpha \leq 0.05$) are bolded. Degrees of freedom are given in square brackets after the *p*-
819 values. Significance was determined through a likelihood-ratio test of null and alternative models, with
820 reference to a χ^2 distribution.

821 Abbreviations—*ln*L, ln Likelihood; *p*, *p*-value; **N/A**, not applicable.
822

References

86. S. K. Wong, W. Li, M. J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem* **279**, 3197-3201 (2004).

87. J. K. Millet, G. R. Whittaker, Murine Leukemia Virus (MLV)-based Coronavirus Spike-pseudotyped Particle Production and Infection. *Bio-protocol* **6**, e2035 (2016).

88. A. Loytynoja, N. Goldman, webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* **11**, 579 (2010).

89. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321 (2010).

90. V. Lefort, J.-E. Longueville, O. Gascuel, SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution* **34**, 2422-2424 (2017).

91. S. B. Hedges, J. Marin, M. Suleski, M. Paymer, S. Kumar, Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* **32**, 835-845 (2015).

92. B. Murrell *et al.*, FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* **30**, 1196-1205 (2013).

93. J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B. Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* **44**, W232-W235 (2016).

94. J. P. Bielawski, Z. Yang, A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol* **59**, 121-132 (2004).

95. C. R. White, R. S. Seymour, The role of gravity in the evolution of mammalian blood pressure. *Evolution* **68**, 901-908 (2014).

96. R. Tacutu *et al.*, Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Res* **46**, D1083-D1090 (2018).

97. W. P. Maddison, D. R. Maddison. (http://www.mesquiteproject.org, 2018).

98. T. Garland, A. R. Ives, Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. *American Naturalist* **155**, 346-364 (2000).

99. M. D. Smith *et al.*, Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection. *Molecular Biology and Evolution* **32**, 1342-1353 (2015).