



HAL
open science

The genetic and evolutionary determinants of COVID-19 susceptibility

Gaspard Kerner, Lluís Quintana-Murci

► **To cite this version:**

Gaspard Kerner, Lluís Quintana-Murci. The genetic and evolutionary determinants of COVID-19 susceptibility. *European Journal of Human Genetics*, 2022, 30 (8), pp.915-921. 10.1038/s41431-022-01141-7 . pasteur-04005972

HAL Id: pasteur-04005972

<https://pasteur.hal.science/pasteur-04005972>

Submitted on 27 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The genetic and evolutionary determinants of COVID-19 susceptibility**

2

3 Gaspard Kerner¹ and Lluís Quintana-Murci,^{1,2}

4

5 ¹Institut Pasteur, Université Paris Cité, CNRS UMR2000, Human Evolutionary Genetics

6 Unit, F-75015 Paris, France

7 ²Collège de France, Chair of Human Genomics and Evolution, F-75005 Paris, France

8

9 Corresponding author: Quintana-Murci, Lluís (quintana@pasteur.fr)

10

11

12 **Abstract**

13 Devastating pandemics, such as that due to COVID-19, can provide strong testimony to our
14 knowledge of the genetic and evolutionary determinants of infectious disease susceptibility
15 and severity. One of the most remarkable aspects of such outbreaks is the stunning
16 interindividual variability observed in the course of infection. In recent decades, enormous
17 progress has been made in the field of the human genetics of infectious diseases, and an
18 increasing number of human genetic factors have been reported to explain, to a great extent,
19 the observed variability for a large number of infectious agents. However, our understanding
20 of the cellular, molecular, and immunological mechanisms underlying such disparities
21 between individuals and ethnic groups, remains very limited. Here, we discuss recent findings
22 relating to human genetic predisposition to infectious disease, from an immunological or
23 population genetic perspective, and show how these and other innovative approaches have
24 been applied to deciphering the genetic basis of human susceptibility to COVID-19 and the
25 severity of this disease. From an evolutionary perspective, we show how past demographic
26 and selection events characterizing the history of our species, including admixture with
27 archaic humans, such as Neanderthals, facilitated modern human adaptation to the threats
28 imposed by ancient pathogens. In the context of emerging infectious diseases, these past
29 episodes of genetic adaptation may contribute to some of the observed population differences
30 in the outcome of SARS-CoV-2 infection and the severity of COVID-19 illness.

31

32 **Introduction**

33 A striking observation in the current COVID-19 pandemic, as observed in several previous
34 outbreaks of other diseases, is that infection with the causal agent (SARS-CoV-2 in this case)
35 does not affect everyone equally. More than five million people have already died from
36 COVID-19, as of February 2022, but death is disproportionately likely in certain individuals,
37 including men and the elderly. After decades of intensive research in human genetics, it is
38 now widely accepted that the large phenotypic variability observed in the course of infection
39 is not random (1-3) and can stem, at least partly, from natural differences in the genetic make-
40 up of humans (4, 5). Our co-evolution with — and co-adaptation to — infectious agents over
41 thousands of years has left molecular signatures in our genomes, which can still be partly
42 disclosed through the use of population and evolutionary genetic approaches. Over the last
43 decade, an increasing number of genes involved in immune functions and host-pathogen
44 interactions have been identified as candidate targets of positive Darwinian selection due to
45 the long-standing pressures imposed by pathogens (4, 6, 7). Remarkably, such pressures may
46 have important consequences for human health today, including greater resistance to
47 emerging diseases, as testified by the iconic case of ancient selection at the *CCR5* locus and
48 present-day resistance to human immunodeficiency virus (8, 9). However, selectively
49 advantageous genetic variants are often unevenly distributed within and between populations,
50 and may, thus, underlie interindividual heterogeneity in the outcome of infection.
51 Furthermore, rare deleterious variants that disrupt or abolish the function of essential immune
52 system genes, can exacerbate disparities in susceptibility to infectious disease.

53 In this review, we will focus on how human and population genetic approaches have
54 facilitated the dissection of the molecular and cellular determinants of critical COVID-19. We
55 will rapidly review how state-of-the-art knowledge on human predisposition to viral disease
56 has led to the discovery of the strongest monogenic predispositions to COVID-19 pneumonia

57 identified to date. We will also review the main loci identified in genome-wide association
58 studies (GWAS) as underlying more complex, but also broader predispositions to COVID-19
59 at the population level. In this context, we will outline previous studies aiming to unravel the
60 grounds for population differences in response to infection, and discuss them in the context of
61 COVID-19. Finally, we will discuss how evolutionary genetics approaches have provided
62 insight into the genetic determinants of COVID-19 susceptibility. Specifically, we will
63 comment on the contribution of past admixture with Neanderthals to present-day
64 susceptibility to viruses, and how particular DNA segments of Neanderthal origin present in
65 the genomes of current non-African populations can alter modern human predisposition to
66 severe COVID-19.

67

68 **Monogenic predisposition to COVID-19**

69 Over the last two years, with the support of unprecedented epidemiological observations of
70 COVID-19 pathophysiology at a worldwide scale, researchers have sought to answer an
71 intriguing question: why would a typically healthy child, teenager or young adult require
72 admission to an intensive care unit following SARS-CoV-2 infection? Since the early 20th
73 century, epidemiological, genetic and molecular studies have garnered compelling evidence
74 that human genetic factors play a fundamental role in determining the outcome of infectious
75 disease (10-12). Definitive proof that inborn errors of immunity (IEI) — rare genetic defects
76 of the immune system — are related to an increase in susceptibility to an individual infectious
77 agent was first provided in 1996, by studies demonstrating that *IFNGR1* deficiency underlies
78 infections with non-virulent mycobacteria (13). Predispositions to many infections have since
79 been shown to be controlled by a narrow set of core genes, specific to each type of infection.
80 For example, since 2015, it has been shown that life-threatening influenza pneumonia is
81 linked to inborn errors of TLR3-, IRF7-, and IRF9-dependent type I interferon (IFN)

82 immunity (14, 15). These groundbreaking discoveries paved the way for the first
83 breakthrough in dissecting the genetic basis of susceptibility to severe COVID-19.

84 A pioneering study led by the international COVID Human Genetic Effort consortium
85 (www.covidhge.com), based on exome or genome sequencing data from 659 critically ill
86 COVID-19 patients of various ancestries, revealed a significant enrichment in biochemically
87 deleterious variants of the three core genes for influenza susceptibility (*TLR3*, *IRF7* and *IRF9*)
88 and 10 other closely related viral genes (odds ratio [OR] = 8.28; $p = 0.01$) in these patients,
89 relative to 534 subjects with asymptomatic or benign infections (16). The genetic
90 homogeneity underlying susceptibility to influenza and COVID-19 was not particularly
91 surprising, given that both these conditions are respiratory infections caused by RNA viruses
92 and transmitted by droplets and small airborne particles.

93 Independently of influenza, the essential role of type I IFNs in COVID-19 pathogenesis
94 was supported by the discovery that critical COVID-19 pneumonia in 16 unrelated men
95 carrying deleterious X-linked *TLR7* variants was caused by low type I IFN production of
96 plasmacytoid dendritic cells (pDCs) in response to SARS-CoV-2 (17). Yet, other studies have
97 brought supporting evidence for the implication of type I IFN-independent genes as
98 monogenic etiologies of COVID-19. For example, genetic variants in the *CFTR* gene,
99 including those causing cystic fibrosis, have been found to be overrepresented among
100 critically ill COVID-19 patients (18-20), consistent with previous associations of the gene
101 with susceptibility to respiratory tract infections (21). Likewise, a genetic polymorphism in
102 the androgen receptor (*AR*) gene that correlates with low serum testosterone level has been
103 associated with severe COVID-19 in men (22). However, immunological and clinical data
104 supporting the implication of these type I IFN-independent genes in COVID-19 pathogenesis
105 remain circumstantial.

106 Further evidence for the pervasive role of type I IFNs in modulating COVID-19 severity
107 was provided by multiple independent observations showing that many individuals with
108 critical disease carry autoantibodies neutralizing type I IFNs (23-26), a phenocopy of the
109 inborn errors. This observation was in line with previous studies describing autoimmune
110 phenocopies of IEI for type II IFN, IL-6, IL-17A/F, and granulocyte-macrophage colony-
111 stimulating factor, underlying mycobacterial disease, staphylococcal disease, mucocutaneous
112 candidiasis, and nocardiosis, respectively. Importantly, in all cases in which data were
113 available, the autoantibodies were found to have been present before SARS-CoV-2 infection,
114 supporting the notion that they are the cause rather than a consequence of the disease (16).
115 The contribution of these studies to improving our understanding of the genetic architecture
116 of COVID-19 severity has been such that it is now estimated that ~20% of patients with
117 critical COVID-19 over 80 years of age, and ~20% of patients of all ages who died from the
118 disease, carried autoantibodies neutralizing type I IFNs. Interestingly, a subsequent study
119 showed that the prevalence of type I IFN autoantibodies in the general population was higher
120 in men than in women (27). This, together with the findings in the *TLR7* (17) and *AR* (22)
121 genes, provides cues as to the genetic and biological mechanisms underlying the observed sex
122 bias among critical COVID-19 cases (28). Furthermore, it was noted that the prevalence of
123 type I IFN autoantibodies increases significantly with age (27), consistent with the early-
124 established higher risk of death from COVID-19 in the elderly population (28). Studies based
125 on clinical genetic approaches have, thus, shown that IEI and their autoimmune phenocopies
126 contribute to the pathogenesis of about 15-20% of patients with critical COVID-19
127 pneumonia, representing a major burden in individuals over 80 years old.

128

129 **Complex predisposition to COVID-19**

130 By contrast to the strong allelic effects of IEI as monogenic determinants of COVID-19
131 illness (ORs typically >5), common variants (frequency >5%) are expected to underlie more
132 subtle, complex patterns of susceptibility (ORs <1.5), usually orchestrated by several genes.
133 Other international efforts were established early in the pandemic, to determine the
134 contribution of common variants to either susceptibility to infection with SARS-CoV-2 or to
135 COVID-19 severity. Specifically, the COVID-19 Host Genetics Initiative (COVID-19 HGI),
136 GenOMICC (Genetics Of Mortality In Critical Care), the 23andMe COVID-19 Team and the
137 Severe Covid-19 GWAS Group conducted GWAS with an unprecedented number of cases
138 and controls in the field of infectious disease genomics research (29-32). The largest study to
139 date, a meta-analysis including 125,584 cases and over 2.5 million controls across 60 studies
140 from 25 countries conducted by COVID-19 HGI, has identified 23 loci significantly
141 associated with disease severity or susceptibility to infection (33). These loci include a region
142 on chromosome 3 (3p21.31) and the ABO locus, these genomic regions being the most
143 frequently replicated to date (six and five times, respectively). The most promising evidence
144 concerning the greater susceptibility to SARS-CoV-2 infection associated with the 3p21.31
145 locus points to involvement of the *SLC6A20* gene, encoding a sodium transporter interacting
146 with ACE2, the well-known receptor for SARS-CoV-2 on the cell surface (34). Consistent
147 with this, the most recent analyses performed by the COVID-19 HGI have identified a
148 polymorphism (rs190509934) close to the *ACE2* gene that is known to lower *ACE2*
149 expression (35) and is associated with a lower risk of infection. Interestingly, an earlier study
150 showed a significant association between *ACE2* allelic variation and COVID-19 severity (36).
151 The ABO locus has also been found to be more strongly associated with susceptibility to
152 infection rather than COVID-19 severity (29), with blood groups O and A protecting against
153 and increasing the risk of infection, respectively.

154 The remaining significant hits in GWAS have been less frequently replicated, but the
155 associations of *IFNAR2* and *TYK2* with COVID-19 severity (29, 30) merit discussion, given
156 the key immunoregulatory functions of the proteins they encode. A potential role of *IFNAR2*,
157 encoding the second chain of the type I IFN receptor, in modulating disease severity is
158 consistent with the known role of type I IFNs in protection against COVID-19 severity. A
159 similar, but more controversial finding is the association of *TYK2* variants with critical
160 COVID-19, with the rs34536443 variant reportedly causal for the underlying phenotype (29).
161 This suggested causal role of rs34536443 is supported by its functional role in protection
162 against multiple autoimmune disorders (37), and the established link between complete *TYK2*
163 deficiency and susceptibility to severe recurrent infections (38). However, the rs34536443-CC
164 genotype has been shown to selectively impair cellular responses to IL-23, but not those to
165 IFN- α or IL-10, consistent with rs34536443 being a common monogenic etiology of
166 tuberculosis, but not of viral infectious diseases (39, 40). Finally, albeit the variant has not
167 been found by GWAS, a common *TLR3* missense variant, impairing perhaps type I IFN
168 immunity, has been reported as a marker of COVID-19 severity (41).

169 GWAS have identified other genetic variants associated with either critical illness or
170 susceptibility to infection with SARS-CoV-2, including variants in the *OAS-RNase L* and
171 *DPP4-DPP9* clusters, which we will discuss later, *MUC5B* and *FOXP4*, previously
172 associated with lung-related phenotypes, and three human leukocyte antigens (HLA) binding
173 to epitope peptides prompting pathogen recognition by the immune system (29, 30). HLA-G
174 rs9380142 and rs143334143, in the vicinity of HLA-C, were found to be associated with
175 COVID-19 severity in both the HGI and GenOMICC studies, whereas HLA-DPB1 rs2071351
176 reached genome-wide significance in the susceptibility analysis of the HGI. Many of these
177 GWAS hits have already been identified in previous studies of lung-related phenotypes,
178 autoimmune or inflammatory diseases (37, 42-44), but further replications and analyses of the

179 underlying cellular mechanisms are required to improve our understanding of the clinical
180 value of these findings in the context of COVID-19.

181

182 **Genetic ancestry and differences in COVID-19 susceptibility**

183 The human genetic factors reported to be associated with COVID-19 in GWAS, despite the
184 uneven representation of different ancestries in these studies, are not equally distributed
185 between populations across the globe. We can therefore wonder whether, as with age, sex or
186 socioeconomic status (45, 46), genetic ancestry, which reflects past differences in population
187 demographic history and adaptation, can also help to explain COVID-19-related health
188 disparities between populations of individuals from different ethnic backgrounds. This is
189 probably not the case for individuals presenting with IEI, who are, by definition,
190 immunocompromised, a notion extending beyond the differences imposed by geographic
191 barriers. For COVID-19, this is consistent with the high risk (OR >8) associated with the
192 carriage of monogenic lesions in type I IFN-related genes (16), and the observation that the
193 two strongest genetic associations reported by the trans-ancestry analysis of the 23andMe
194 COVID-19 Team, at the ABO and 3p21.31 loci, did not explain differences in risk between
195 populations (32). However, the same study also showed that non-European ancestry was a
196 significant risk factor for hospitalization, after accounting for sociodemographic factors and
197 pre-existing health conditions, supporting the notion that complex genetic architectures, as
198 opposed to strong allelic effects, may account for population-level differences in the outcome
199 of SARS-CoV-2 infection. In support of this hypothesis, efforts to quantify differences
200 between ancestries in rates of SARS-CoV-2 infection and COVID-19 clinical manifestations
201 in England showed that ethnic minorities had higher risks of testing positive for SARS-CoV-2
202 and of adverse COVID-19 outcomes, after accounting for differences in sociodemographic,
203 clinical, and household characteristics (47).

204 More generally, these observations are supported by previous studies exploring the extent
205 and nature of population differences in immune responses (48, 49). For example, differential
206 transcriptional responses to viral and bacterial stimuli have been described between
207 individuals of African and European descent. More recently, a single-cell RNA-sequencing
208 study of influenza-infected immune cells from individuals of European and African ancestry
209 reported ancestry-dependent gene signatures under the control of human genetic factors (here,
210 *cis*-expression quantitative trait loci, eQTLs) differing between ancestries (50). Interestingly,
211 the overlap in the immune system genes underlying susceptibility to both influenza and
212 COVID-19 suggests that a similar situation may apply to COVID-19, although single-cell
213 studies investigating population differences in immune responses to SARS-CoV-2 infection
214 are still lacking.

215 The underlying causes of such differences in immunity post-infection between ethnic
216 groups can also be investigated by adopting a different strategy, based on population and
217 evolutionary genetics approaches. In this context, one recent study sought to quantify the
218 impact of past coronavirus-like epidemics across the globe, by screening for signatures of
219 selection at loci for 420 coronavirus-interacting human proteins (CoV-VIPs) in different
220 populations (51). Surprisingly, none of the 26 populations examined worldwide presented a
221 signature of adaptation at the CoV-VIP loci, with the exception of populations of East Asian
222 ancestry, consistent with the geographic origin of several modern coronavirus epidemics. In
223 short, the authors found a cluster of 42 CoV-VIP loci displaying consistent patterns of
224 adaptive evolution, dating back to 20,000 years ago, in East Asian populations, reflecting
225 long-term selection pressures exerted by coronavirus-like viruses on the ancestors of modern-
226 day East Asians. Possible geographically restricted positive selection at immune loci has also
227 been detected in a Japanese population, in which the DPB1*04:01 HLA allele was found to
228 have undergone a strong recent increase in frequency (52). In line with a putative protective

229 effect of this variant against hepatitis B virus (HBV) infection (53), the authors speculated
230 that this increase in frequency in the Japanese population resulted from past pathogen-driven
231 selection. Collectively, these studies support a role for local human adaptation in response to
232 past and present infectious agents in increasing immune response disparities between
233 populations around the world.

234

235 **Ancient admixture with Neanderthals and RNA viruses**

236 The observed differences in the genetic make-up of populations from different ancestries
237 result not only from their past demographic or adaptive history, but also from differences in
238 their past history of admixture (or hybridization) with other types of humans that are now
239 extinct (**Figure 1**). Anatomically modern humans interbred with these archaic hominins, such
240 as Neanderthals and Denisovans, on multiple occasions and in several locations (54). As a
241 result, all non-African groups share ~2% Neanderthal ancestry in their genomes, whereas
242 some south-east Asian and Oceanian populations have accumulated up to 5-6% of combined
243 (Neanderthal and Denisovan) archaic ancestry (55). There is evidence to suggest that
244 purifying selection has been the dominant selective force acting against the introgression of
245 archaic DNA material, leading to a steady decrease in haplotypes of an archaic nature in the
246 genomes of modern humans over time (56).

247 However, in some cases, hybridization events appear to have facilitated the acquisition of
248 advantageous traits, a phenomenon known as ‘adaptive introgression’ (57). Neanderthals and
249 Denisovans inhabited Eurasia for at least 300,000 years before modern humans arrived, and
250 are thought to have become genetically adapted to their local climates, nutritional resources
251 and pathogens over this period. Unsurprisingly, there is increasing evidence to suggest that
252 archaic introgression has facilitated the acquisition, by modern humans, of beneficial variants
253 of immunity-related genes, attesting to the long-term adaptation of the archaic species to

254 pathogens outside of Africa. Interestingly, an early, influential work showed an enrichment in
255 Neanderthal ancestry among innate immunity genes, as a whole, in Europeans (58). Similarly,
256 high levels of Neanderthal or Denisovan ancestry have been detected in the genomes of
257 modern humans at the loci for antiviral OAS genes, the *TLR1-6-10* gene cluster or the
258 inflammation-related *TNFAIP3* gene in several non-African populations around the world
259 (59-61).

260 Remarkably, two studies found an enrichment in Neanderthal ancestry for genetic variants
261 associated with gene expression variation, eQTLs, in monocytes and macrophages in
262 Europeans. This enrichment was particularly evident for genetic variants associated with
263 antiviral responses (48, 49), suggesting that Neanderthal introgression, in particular,
264 facilitated the genetic adaptation of early Eurasians to viral challenges. Consistent with this
265 theory, human genes encoding proteins that interact with viruses were also found to display
266 significantly high levels of Neanderthal ancestry, particularly those encoding proteins
267 interacting with RNA viruses, such as influenza, hepatitis C virus or coronaviruses (62).

268 Another study explored the impact of Neanderthal ancestry on the human regulatory genetic
269 landscape, including promoters, enhancers and miRNA-mediated regulation (63). A massive
270 colocalization of Neanderthal variants with active enhancers in adipose-related tissues and
271 various types of primary T cells was observed. Collectively, these studies shed light on the
272 extent to which archaic introgression has contributed to modern human adaptation to new
273 environments, by modulating human immunity to newly encountered pathogens, including
274 RNA viruses in particular.

275

276 **Neanderthal heritage and the current COVID-19 pandemic**

277 Recent studies driven by the COVID-19 pandemic have provided additional support for the
278 links between Neanderthal introgression and human immunity to viruses. Two GWAS hits

279 from the aforementioned human genetic determinants of COVID-19 susceptibility and
280 severity overlap with genomic regions inherited from Neanderthals. One of these regions
281 includes genetic variants at the chr12q24.13 locus, which are mostly absent from Africans but
282 were present in Neanderthals. These variants define a ~75 kb haplotype in individuals of
283 European ancestry that has been found to be associated with a 22% lower risk of
284 hospitalization for COVID-19 (29, 30, 64). The locus concerned covers the OAS-RNase L
285 cluster, which encodes enzymes essential for antiviral immunity (65, 66). However, the
286 cellular mechanism underlying this improvement in COVID-19 outcomes is not yet fully
287 understood. In this context, one recent study sought to delineate the causal variant associated
288 with COVID-19 protection from more than a hundred linked variants within the same
289 associated haplotype, by focusing on the rs10774671 variant, a candidate *OAS1* splice
290 acceptor variant (67). The authors tested associations in different ancestry groups with
291 markedly different levels of linkage disequilibrium (LD). By focusing on individuals of
292 African ancestry, in whom LD levels are lowest and rs10774671 segregates independently of
293 the other variants, the authors were able to identify a causal connection between the
294 rs10774671-G allele and COVID-19 illness. These findings highlight the role of the *OAS1*
295 isoform p46, encoded by the splice site variant at this locus, in effective protection against
296 hospitalization for COVID-19, at least in individuals of European or African descent
297 (OR = 0.92-0.94, $p = 5.8 \times 10^{-10}$ -0.03).

298 In stark contrast, the other COVID-19-related genomic region of Neanderthal origin, a
299 locus on chromosome 3, 3p21.31, has been associated with greater susceptibility to the
300 development of severe forms of COVID-19. It spans a 50-kb haplotype containing six genes
301 (*SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6* and *XCRI*). The contribution of one of these
302 genes, *SLC6A20*, to susceptibility to infection has already been discussed. GWAS data
303 suggest that the risk of hospitalization for COVID-19 is 60% higher in carriers of the

304 Neanderthal haplotype, which is at least three times more frequent in individuals of South
305 Asian descent (> 50%) than in individuals with European ancestry (16%) (68). This, together
306 with unaccounted sociodemographic factors, may partly explain the higher risk of infection or
307 hospitalization for COVID-19 in minorities of South Asian ancestry living in the UK (47).
308 Two independent studies aiming to determine the cellular basis of the increase in the risk of
309 severe disease associated with this locus have suggested that this outcome may result from a
310 decrease in CXCR6 levels (69, 70).

311 Finally, a genetic variant in the promoter region of the *DPP4* gene, encoding a receptor for
312 another coronavirus, MERS-CoV, but not for SARS-CoV-2, was also inherited from
313 Neanderthals and has been shown to double the risk of developing critical COVID-19. The
314 Neanderthal variant did not reach genome-wide significance in GWAS studies, but variants of
315 *DPP9*, a homolog of *DPP4*, are, on the contrary, significantly associated with severe COVID-
316 19 in GWAS, supporting a potential role for *DPP4* in COVID-19 pathogenesis (30).
317 Together, these observations lend further support to the notion that Neanderthal introgression
318 has had profound consequences for the adaptation of our species to viral challenges, and that
319 such past adaptation events can affect the present-day health status of individuals infected
320 with SARS-CoV-2.

321

322 **Conclusions**

323 The COVID-19 pandemic has had devastating humanitarian consequences. Vaccines were
324 rapidly developed and have prevented the most harmful outcomes of infection with the virus.
325 Nevertheless, the “miracle” brought about by vaccines simply reinforces the importance of
326 prior decades of basic scientific research, without which no amount of investment could have
327 provided us with better solutions so rapidly. In this respect, the field of human genetics of
328 infectious diseases, driven by clinical, population and evolutionary genetic studies, has made

329 spectacular breakthroughs over the last decade. Prompted by duty, and perhaps a fear of
330 COVID-19, the scientific community has rapidly organized unprecedented international
331 efforts, and provided scientific results more successfully and rapidly than ever before. For
332 example, the discovery that the pathogenesis of the disease in 20% of patients with critical
333 COVID-19 pneumonia can be explained by either IELs of type I IFN immunity, or pre-
334 existing autoantibodies neutralizing type I IFNs is an outstanding finding for common
335 infections, for which monogenic lesions have never been shown to underlie more than 1% of
336 cases for other conditions (39). The contributions of GWAS have also been substantial,
337 particularly given the unprecedented number of cases (125,000) and controls (2.5 million)
338 recruited for such studies, the second largest infectious GWAS in terms of case numbers
339 being that performed by the 23andMe consortium on 107,769 cases of chickenpox and 15,982
340 controls (71).

341 Importantly, trans-ancestry analyses in GWAS have revealed population disparities in
342 terms of susceptibility to infection or disease severity. The strongest genetic determinants of
343 COVID-19 reported to date have similar effects on individuals across the globe, but non-
344 European ancestries have been shown to confer a higher risk of developing severe forms of
345 COVID-19. These studies have once again highlighted the importance of including diverse
346 and underrepresented human populations in genomic studies, to delineate variants differing
347 between ancestries that may, under complex genetic architectures, underlie population
348 differences in disease outcome (72). Such ancestry-inclusive efforts are of major importance
349 in the context of drug development, as drug efficiency may depend strongly on the genetic
350 make-up of the population. An eQTL study in the context of influenza infection in
351 populations of African and European ancestry has already supported the notion that, in some
352 conditions, variants differentially represented across ancestries can result in different genetic
353 signatures, possibly attesting to an ancestry-specific activation of biological pathways (50).

354 Finally, the sequencing of archaic hominin genomes has made it possible to unravel some of
355 the essential features of immunity responsible for facilitating the adaptation of early non-
356 Africans to newly encountered pathogenic environments. Studies of the genetic legacy of
357 archaic hominins in the genomes of modern humans extend well beyond questions relating to
358 molecular anthropology, as attested by the surprising finding that the genetic legacy of ancient
359 admixture with Neanderthals around 50,000 years ago still affects the health of humans today,
360 even in the specific context of the COVID-19 pandemic.

361

362 **Competing Interests**

363 None declared

364

365 **Ethical Approval**

366 Approval by an ethics committee was not required as this is a systematic review of published
367 literature.

368

369 **Acknowledgments**

370 We thank Yann Aquino and Sebastian Cuadros Espinoza for helpful discussions. The
371 Laboratory of Human Evolutionary Genetics is supported by the Institut Pasteur, the *Collège*
372 *de France*, the *Centre Nationale de la Recherche Scientifique* (CNRS), the *Agence Nationale*
373 *de la Recherche* (ANR) grants CNSVIRGEN (ANR-19-CE15-0009-02) and COVID-19-
374 POPCELL (ANR-21-CO14-0003-01), the French Government's *Investissement d'Avenir*
375 program, *Laboratoires d'Excellence* "Integrative Biology of Emerging Infectious Diseases"
376 (ANR-10-LABX-62-IBEID) and "*Milieu Intérieur*" (ANR-10-LABX-69-01), the *Fondation*
377 *pour la Recherche Médicale* (Equipe FRM DEQ20180339214), the *Fondation Allianz-Institut*
378 *de France*, and the *Fondation de France* (no. 00106080). G.K. is supported by a Pasteur-
379 Roux-Cantarini fellowship.

380

381 **References**

- 382 1. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al.
383 Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO
384 Clinical Characterisation Protocol: prospective observational cohort study. *Bmj*.
385 2020;369:m1985.
- 386 2. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for
387 mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort
388 study. *Lancet*. 2020;395(10229):1054-62.
- 389 3. Casanova JL, Abel L. Inborn errors of immunity to infection: the rule rather than the
390 exception. *J Exp Med*. 2005;202(2):197-201.
- 391 4. Quintana-Murci L. Human Immunology through the Lens of Evolutionary Genetics.
392 *Cell*. 2019;177(1):184-99.
- 393 5. Casanova JL, Abel L, Quintana-Murci L. Immunology taught by human genetics. *Cold*
394 *Spring Harb Symp Quant Biol*. 2013;78:157-72.
- 395 6. Barreiro LB, Quintana-Murci L. From evolutionary genetics to human immunology:
396 how selection shapes host defence genes. *Nat Rev Genet*. 2010;11(1):17-30.
- 397 7. Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L, Nielsen R. Signatures
398 of environmental genetic adaptation pinpoint pathogens as the main selective pressure
399 through human evolution. *PLoS Genet*. 2011;7(11):e1002355.
- 400 8. Lalani AS, Masters J, Zeng W, Barrett J, Pannu R, Everett H, et al. Use of chemokine
401 receptors by poxviruses. *Science*. 1999;286(5446):1968-71.
- 402 9. Bamshad MJ, Mummidi S, Gonzalez E, Ahuja SS, Dunn DM, Watkins WS, et al. A
403 strong signature of balancing selection in the 5' cis-regulatory region of CCR5. *Proc*
404 *Natl Acad Sci U S A*. 2002;99(16):10539-44.
- 405 10. Dubos RJ. Second Thoughts on the Germ Theory. *Scientific American*. 1955;192:31-5.

- 406 11. Allison AC. Protection afforded by sickle-cell trait against subtertian malareal infection.
407 Br Med J. 1954;1(4857):290-4.
- 408 12. Casanova JL, Abel L. Lethal Infectious Diseases as Inborn Errors of Immunity: Toward
409 a Synthesis of the Germ and Genetic Theories. *Annu Rev Pathol.* 2021;16:23-50.
- 410 13. Jouanguy E, Altare F, Lamhamedi S, Revy P, Emile JF, Newport M, et al. Interferon-
411 gamma-receptor deficiency in an infant with fatal bacille Calmette-Guérin infection. *N*
412 *Engl J Med.* 1996;335(26):1956-61.
- 413 14. Ciancanelli MJ, Huang SX, Luthra P, Garner H, Itan Y, Volpi S, et al. Infectious
414 disease. Life-threatening influenza and impaired interferon amplification in human
415 IRF7 deficiency. *Science.* 2015;348(6233):448-53.
- 416 15. Casanova JL, Abel L. The human genetic determinism of life-threatening infectious
417 diseases: genetic heterogeneity and physiological homogeneity? *Hum Genet.*
418 2020;139(6-7):681-94.
- 419 16. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of
420 type I IFN immunity in patients with life-threatening COVID-19. *Science.*
421 2020;370(6515).
- 422 17. Asano T, Boisson B, Onodi F, Matuozzo D, Moncada-Velez M, Maglorius Renkilaraj
423 MRL, et al. X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with
424 life-threatening COVID-19. *Science immunology.* 2021;6(62):eabl4348.
- 425 18. Baldassarri M, Fava F, Fallerini C, Daga S, Benetti E, Zguro K, et al. Severe COVID-
426 19 in Hospitalized Carriers of Single CFTR Pathogenic Variants. *J Pers Med.*
427 2021;11(6):558.
- 428 19. Corvol H, de Miranda S, Lemonnier L, Kemgang A, Reynaud Gaubert M, Chiron R, et
429 al. First Wave of COVID-19 in French Patients with Cystic Fibrosis. *J Clin Med.*
430 2020;9(11):3624.

- 431 20. Fallerini C, Picchiotti N, Baldassarri M, Zguro K, Daga S, Fava F, et al. Common, low-
432 frequency, rare, and ultra-rare coding variants contribute to COVID-19 severity. *Human*
433 *genetics*. 2022;141(1):147-73.
- 434 21. Miller AC, Comellas AP, Hornick DB, Stoltz DA, Cavanaugh JE, Gerke AK, et al.
435 Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related
436 conditions. *Proceedings of the National Academy of Sciences of the United States of*
437 *America*. 2020;117(3):1621-7.
- 438 22. Baldassarri M, Picchiotti N, Fava F, Fallerini C, Benetti E, Daga S, et al. Shorter
439 androgen receptor polyQ alleles protect against life-threatening COVID-19 disease in
440 European males. *EBioMedicine*. 2021;65:103246-.
- 441 23. Chauvineau-Grenier A, Bastard P, Servajean A, Gervais A, Rosain J, Jouanguy E, et al.
442 Autoantibodies neutralizing type I interferons in 20% of COVID-19 deaths in a French
443 hospital. *Res Sq*. 2021.
- 444 24. Troya J, Bastard P, Planas-Serra L, Ryan P, Ruiz M, de Carranza M, et al. Neutralizing
445 Autoantibodies to Type I IFNs in >10% of Patients with Severe COVID-19 Pneumonia
446 Hospitalized in Madrid, Spain. *J Clin Immunol*. 2021;41(5):914-22.
- 447 25. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann HH, Zhang Y, et al.
448 Autoantibodies against type I IFNs in patients with life-threatening COVID-19.
449 *Science*. 2020;370(6515).
- 450 26. Koning R, Bastard P, Casanova JL, Brouwer MC, van de Beek D. Autoantibodies
451 against type I interferons are associated with multi-organ failure in COVID-19 patients.
452 *Intensive Care Med*. 2021;47(6):704-6.
- 453 27. Bastard P, Gervais A, Le Voyer T, Rosain J, Philippot Q, Manry J, et al. Autoantibodies
454 neutralizing type I IFNs are present in ~4% of uninfected individuals over 70 years old
455 and account for ~20% of COVID-19 deaths. *Sci Immunol*. 2021;6(62).

- 456 28. Ahrenfeldt LJ, Otavova M, Christensen K, Lindahl-Jacobsen R. Sex and age differences
457 in COVID-19 mortality in Europe. *Wien Klin Wochenschr.* 2021;133(7-8):393-8.
- 458 29. Initiative C-HG. Mapping the human genetic architecture of COVID-19. *Nature.*
459 2021;600(7889):472-7.
- 460 30. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al.
461 Genetic mechanisms of critical illness in COVID-19. *Nature.* 2021;591(7848):92-8.
- 462 31. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al.
463 Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J*
464 *Med.* 2020;383(16):1522-34.
- 465 32. Shelton JF, Shastri AJ, Ye C, Weldon CH, Filshtein-Sonmez T, Coker D, et al. Trans-
466 ancestry analysis reveals genetic and nongenetic associations with COVID-19
467 susceptibility and severity. *Nat Genet.* 2021;53(6):801-8.
- 468 33. Initiative C-HG, Ganna A. Mapping the human genetic architecture of COVID-19: an
469 update. *medRxiv.* 2021:2021.11.08.21265944.
- 470 34. Vuille-dit-Bille RN, Camargo SM, Emmenegger L, Sasse T, Kummer E, Jando J, et al.
471 Human intestine luminal ACE2 and amino acid transporter expression increased by
472 ACE-inhibitors. *Amino Acids.* 2015;47(4):693-705.
- 473 35. Horowitz JE, Kosmicki JA, Damask A, Sharma D, Roberts GHL, Justice AE, et al.
474 Genome-wide analysis in 756,646 individuals provides first genetic evidence that ACE2
475 expression influences COVID-19 risk and yields genetic risk scores predictive of severe
476 disease. *medRxiv.* 2021.
- 477 36. Benetti E, Tita R, Spiga O, Ciolfi A, Birolo G, Bruselles A, et al. ACE2 gene variants
478 may underlie interindividual variability and susceptibility to COVID-19 in the Italian
479 population. *Eur J Hum Genet.* 2020;28(11):1602-14.

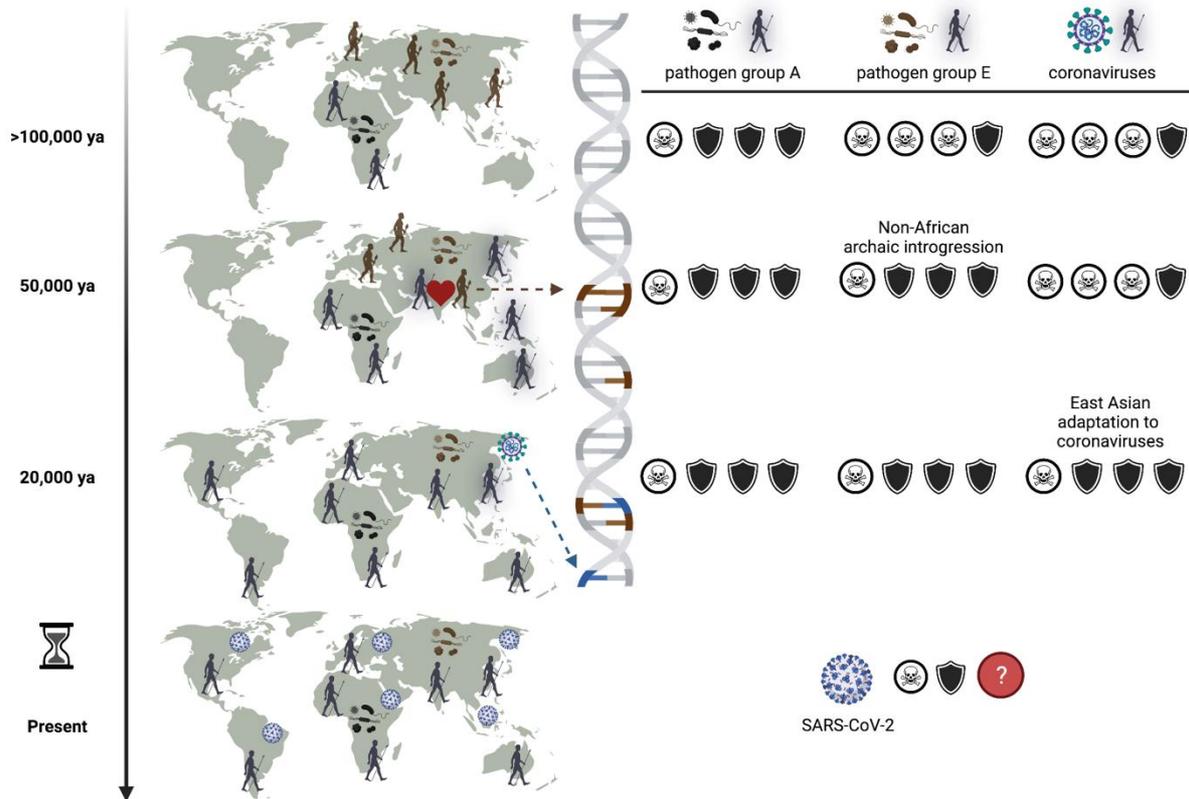
- 480 37. Dendrou CA, Cortes A, Shipman L, Evans HG, Attfield KE, Jostins L, et al. Resolving
481 TYK2 locus genotype-to-phenotype differences in autoimmunity. *Sci Transl Med*.
482 2016;8(363):363ra149.
- 483 38. Kreins AY, Ciancanelli MJ, Okada S, Kong XF, Ramírez-Alejo N, Kilic SS, et al.
484 Human TYK2 deficiency: Mycobacterial and viral infections without hyper-IgE
485 syndrome. *J Exp Med*. 2015;212(10):1641-62.
- 486 39. Kerner G, Ramirez-Alejo N, Seeleuthner Y, Yang R, Ogishi M, Cobat A, et al.
487 Homozygosity for TYK2 P1104A underlies tuberculosis in about 1% of patients in a
488 cohort of European ancestry. *Proc Natl Acad Sci U S A*. 2019;116(21):10430-4.
- 489 40. Boisson-Dupuis S, Ramirez-Alejo N, Li Z, Patin E, Rao G, Kerner G, et al.
490 Tuberculosis and impaired IL-23-dependent IFN- γ immunity in humans homozygous
491 for a common TYK2 missense variant. *Sci Immunol*. 2018;3(30).
- 492 41. Croci S, Venneri MA, Mantovani S, Fallerini C, Benetti E, Picchiotti N, et al. The
493 polymorphism L412F in TLR3 inhibits autophagy and is a marker of severe COVID-19
494 in males. *Autophagy*. 2021:1-11.
- 495 42. Dai J, Lv J, Zhu M, Wang Y, Qin N, Ma H, et al. Identification of risk loci and a
496 polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese
497 populations. *Lancet Respir Med*. 2019;7(10):881-91.
- 498 43. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, et al. Genome-
499 wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat*
500 *Genet*. 2013;45(6):613-20.
- 501 44. Manichaikul A, Wang XQ, Sun L, Dupuis J, Borczuk AC, Nguyen JN, et al. Genome-
502 wide association study of subclinical interstitial lung disease in MESA. *Respir Res*.
503 2017;18(1):97.

- 504 45. de Glanville WA, Thomas LF, Cook EAJ, Bronsvort BMC, Wamae NC, Kariuki S, et
505 al. Household socio-economic position and individual infectious disease risk in rural
506 Kenya. *Sci Rep.* 2019;9(1):2972.
- 507 46. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors
508 associated with COVID-19-related death using OpenSAFELY. *Nature.*
509 2020;584(7821):430-6.
- 510 47. Mathur R, Rentsch CT, Morton CE, Hulme WJ, Schultze A, MacKenna B, et al. Ethnic
511 differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive
512 care unit admission, and death in 17 million adults in England: an observational cohort
513 study using the OpenSAFELY platform. *Lancet.* 2021;397(10286):1711-24.
- 514 48. Nédélec Y, Sanz J, Baharian G, Szpiech ZA, Pacis A, Dumaine A, et al. Genetic
515 Ancestry and Natural Selection Drive Population Differences in Immune Responses to
516 Pathogens. *Cell.* 2016;167(3):657-69.e21.
- 517 49. Quach H, Rotival M, Pothlichet J, Loh YE, Dannemann M, Zidane N, et al. Genetic
518 Adaptation and Neandertal Admixture Shaped the Immune System of Human
519 Populations. *Cell.* 2016;167(3):643-56.e17.
- 520 50. Randolph HE, Fiege JK, Thielen BK, Mickelson CK, Shiratori M, Barroso-Batista J, et
521 al. Genetic ancestry effects on the response to viral infection are pervasive but cell type
522 specific. *Science.* 2021;374(6571):1127-33.
- 523 51. Souilmi Y, Lauterbur ME, Tobler R, Huber CD, Johar AS, Moradi SV, et al. An ancient
524 viral epidemic involving host coronavirus interacting genes more than 20,000 years ago
525 in East Asia. *Curr Biol.* 2021;31(16):3504-14.e9.
- 526 52. Kawashima M, Ohashi J, Nishida N, Tokunaga K. Evolutionary analysis of classical
527 HLA class I and II genes suggests that recent positive selection acted on DPB1*04:01 in
528 Japanese population. *PLoS One.* 2012;7(10):e46806.

- 529 53. Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, et al.
530 A genome-wide association study identifies variants in the HLA-DP locus associated
531 with chronic hepatitis B in Asians. *Nat Genet.* 2009;41(5):591-5.
- 532 54. Dannemann M, Racimo F. Something old, something borrowed: admixture and
533 adaptation in human evolution. *Curr Opin Genet Dev.* 2018;53:1-8.
- 534 55. Choin J, Mendoza-Revilla J, Arauna LR, Cuadros-Espinoza S, Cassar O, Larena M, et
535 al. Genomic insights into population history and biological adaptation in Oceania.
536 *Nature.* 2021;592(7855):583-9.
- 537 56. Harris K, Nielsen R. The Genetic Cost of Neanderthal Introgression. *Genetics.*
538 2016;203(2):881-91.
- 539 57. Taylor SA, Larson EL. Insights from genomes into the evolutionary importance and
540 prevalence of hybridization in nature. *Nat Ecol Evol.* 2019;3(2):170-7.
- 541 58. Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova JL, et al. Genomic
542 Signatures of Selective Pressures and Introgression from Archaic Hominins at Human
543 Innate Immunity Genes. *Am J Hum Genet.* 2016;98(1):5-21.
- 544 59. Dannemann M, Andrés AM, Kelso J. Introgression of Neandertal- and Denisovan-like
545 Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors. *Am J*
546 *Hum Genet.* 2016;98(1):22-33.
- 547 60. Sams AJ, Dumaine A, Nédélec Y, Yotova V, Alfieri C, Tanner JE, et al. Adaptively
548 introgressed Neandertal haplotype at the OAS locus functionally impacts innate
549 immune responses in humans. *Genome Biol.* 2016;17(1):246.
- 550 61. Zammit NW, Siggs OM, Gray PE, Horikawa K, Langley DB, Walters SN, et al.
551 Denisovan, modern human and mouse TNFAIP3 alleles tune A20 phosphorylation and
552 immunity. *Nat Immunol.* 2019;20(10):1299-310.

- 553 62. Enard D, Petrov DA. Evidence that RNA Viruses Drove Adaptive Introgression
554 between Neanderthals and Modern Humans. *Cell*. 2018;175(2):360-71.e13.
- 555 63. Silvert M, Quintana-Murci L, Rotival M. Impact and Evolutionary Determinants of
556 Neanderthal Introgression on Transcriptional and Post-Transcriptional Regulation. *Am*
557 *J Hum Genet*. 2019;104(6):1241-50.
- 558 64. Zeberg H, Pääbo S. A genomic region associated with protection against severe
559 COVID-19 is inherited from Neandertals. *Proc Natl Acad Sci U S A*. 2021;118(9).
- 560 65. Choi UY, Kang JS, Hwang YS, Kim YJ. Oligoadenylate synthase-like (OASL)
561 proteins: dual functions and associations with diseases. *Exp Mol Med*. 2015;47(3):e144.
- 562 66. Gresser I, Vignaux F, Belardelli F, Tovey MG, Maunoury MT. Injection of mice with
563 antibody to mouse interferon alpha/beta decreases the level of 2'-5' oligoadenylate
564 synthetase in peritoneal macrophages. *J Virol*. 1985;53(1):221-7.
- 565 67. Huffman JE, Butler-Laporte G, Khan A, Pairo-Castineira E, Drivas TG, Peloso GM, et
566 al. Multi-ancestry fine mapping implicates OAS1 splicing in risk of severe COVID-19.
567 *Nat Genet*. 2022.
- 568 68. Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited
569 from Neanderthals. *Nature*. 2020;587(7835):610-2.
- 570 69. Dai Y, Wang J, Jeong HH, Chen W, Jia P, Zhao Z. Association of CXCR6 with
571 COVID-19 severity: delineating the host genetic factors in transcriptomic regulation.
572 *Hum Genet*. 2021;140(9):1313-28.
- 573 70. Katz DH, Tahir UA, Ngo D, Benson MD, Bick AG, Pampana A, et al. Proteomic
574 Profiling in Biracial Cohorts Implicates DC-SIGN as a Mediator of Genetic Risk in
575 COVID-19. *medRxiv*. 2020.

- 576 71. Tian C, Hromatka BS, Kiefer AK, Eriksson N, Noble SM, Tung JY, et al. Genome-
577 wide association and HLA region fine-mapping studies identify susceptibility loci for
578 multiple common infections. *Nat Commun.* 2017;8(1):599.
- 579 72. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic
580 Studies. *Cell.* 2019;177(4):1080.
- 581



582

583

584

585 **Figure 1. Ancient admixture and present-day immunity to infection.**

586 Graphical representation of the contribution of admixture with archaic humans and exposure
 587 to ancient viruses to differences in the response to infection between present-day populations.
 588 On the left of the figure, genetic material from archaic humans, such as Neanderthals or
 589 Denisovans, is shown to be inherited by non-Africans, a process that began ~50,000 years
 590 ago. When beneficial, this event is known as ‘adaptive introgression’, and is thought to have
 591 facilitated the acquisition of advantageous variants by modern humans, accelerating their
 592 adaptation to Eurasian pathogens (“pathogen group E”), which here are hypothesized to be
 593 different from African pathogens (“pathogen group A). Continuous high-level exposure of the
 594 ancestors of modern East Asians to coronaviruses over the last ~20,000 years has left
 595 signatures of selection at CoV-VIP loci in the genomes of modern East Asians (51), a pattern
 596 that is not observed in other human populations. Together with other genetic or environmental
 597 factors, these historical events underlie some of the disparities observed today in
 598 predisposition to COVID-19 between human populations. Created with BioRender.com.