

Critically-ill COVID-19 susceptibility gene *CCR3* shows natural selection in sub-Saharan Africans

Zewen Sun^{1†}, Lin Pan^{1,2†}, Aowen Tian^{3,4}, Peng Chen^{1,3,4*}

¹Department of Genetics, College of Basic Medical Sciences, Jilin University, Changchun, Jilin, China

²The First Hospital of Jilin University, Changchun, Jilin 130021, China.

³Department of Pathology, College of Basic Medical Sciences, Jilin University, Changchun, Jilin, China

⁴Key Laboratory of Pathobiology, Ministry of Education, Jilin University, Changchun, Jilin, China

† These authors contributed equally in this work.

* Corresponding authors: Peng Chen, PhD.

Email: pchen@jlu.edu.cn

This file includes:

Main Text

Table 1

Figures 1 to 2

1

1

1 **Abstract**

2 The prevalence of COVID-19 critical illness varies across ethnicities, with recent studies suggesting that
3 genetic factors may contribute to this variation. The aim of this study was to investigate natural selection
4 signals of genes associated with critically-ill COVID-19 in sub-Saharan Africans. Severe COVID-19 SNPs
5 were obtained from the HGI website. Selection signals were assessed in 661 sub-Sahara Africans from
6 1000 Genomes Project using integrated haplotype score (iHS), cross-population extended haplotype
7 homozygosity (xpEHH), and fixation index (Fst). Allele frequency trajectory analysis of ancient DNA
8 samples were used to validate the existing of selection in sub-Sahara Africans. We also used Mendelian
9 randomization to decipher the correlation between natural selection and critically-ill COVID-19. We
10 identified that *CCR3* exhibited significant natural selection signals in sub-Sahara Africans. Within the
11 *CCR3* gene, rs17217831-A showed both high iHS (Standardized iHS = 2) and high XP-EHH (Standardized
12 XP-EHH = 2.5) in sub-Sahara Africans. Allele frequency trajectory of *CCR3* rs17217831-A revealed
13 natural selection occurring in the recent 1,500 years. Natural selection resulted in increased *CCR3*
14 expression in sub-Sahara Africans. Mendelian Randomization provided evidence that increased blood
15 *CCR3* expression and eosinophil counts lowered the risk of critically ill COVID-19. Our findings suggest
16 that sub-Saharan Africans are less vulnerable to critically ill COVID-19 due to natural selection and
17 identify *CCR3* as a potential novel therapeutic target.

18

1 **Main Text**

2 **Introduction**

3 The pandemic of COVID-19 has caused unprecedented and devastating influence around the world¹.
4 Most patients are asymptomatic or present with mild disease, while others proceed to severe hypoxic
5 pneumonia or other diseases, which could ultimately lead to death².

6 African ancestry is considered to be associated with the outcome of COVID-19, but the conclusion has
7 not been made. The risk of hospitalization or in-hospital mortality of people with African ancestry is usually
8 confounded by socioeconomic status³. After carefully controlling of these factors, African people manifested
9 equal⁴ or even lower risk of critical illness^{5,6}. At the same time, the vaccination rate and death rate in Africa
10 are both the lowest.

11 Studies have shown that several genetic variants associated with severe COVID-19 were apparently
12 unique in African population. Huffman et al. found that *OAS1* rs10774671 G, which protects against
13 hospitalization for COVID-19, had a higher allele frequency (58%) in Africans than in European population
14 (32%)⁷. It has also been suggested that severe COVID-19 genes were expressed differentially in monocytes
15 between Europeans and Africans⁸. And an average of 69% of COVID severity-associated genes across
16 conditions in monocytes showed increased expression with greater European ancestry. Additionally, higher
17 proportion of European ancestry has been linked to a more robust type I IFN (Interferon) response linked to
18 lower viral titers at later stages, which have significant clinical implications for SARS-CoV-2 infection. All
19 these evidence gives birth to the hypothesis that African people may be genetically invulnerable to critically-
20 ill COVID-19.

21 Pathogenic viruses are among the strongest sources of selection pressure in human evolution⁹. Previous
22 selection analysis shows several genes related to the susceptibility of SARS-CoV-2 infection, including
23 *ACE2*, *TMPRSS2*, *DPP4* and *LY6E* are under natural selection in Africans¹⁰. However, whether natural
24 selection involved in the allele frequency difference or expression regulation of critically-ill COVID-19 genes
25 remains unknown. In this study, we investigate the recent natural selection of the risk alleles associated with
26 critically-ill COVID-19 in sub-Saharan African aiming to explain the population-specific risk of critically-

- 1 ill COVID-19. Our results could inform the development of novel therapeutic targets for critically-ill
- 2 COVID-19.
- 3

1 **Materials and Methods**

2 **SNPs associated with Critically-ill COVID-19**

3 We obtained COVID-19 genome-wide association results from HGI website
4 (<http://www.covid19hg.org/results/>). The round 7 summary statistics without 23andme samples was used in
5 this study¹¹. At the significance level of $p < 1 \times 10^{-5}$, 8,305 SNPs were associated with critically-ill COVID-19.
6 The details of GWAS meta-analysis of COVID-19 can be found in HGI website.

7 **Population differentiation, Natural selection analysis**

8 Natural selection signals in the critical-ill SNPs were tested in the genotype data from the 1000 Genomes
9 Project. In total, we selected 5 superpopulations, including Europeans (N=503), sub-Saharan Africans (N=661),
10 East Asians (N=504), Ad Mixed American (N=347) and South Asian (N=489). The integrated haplotype
11 Score (iHS) is a measure of the amount of extended haplotype homozygosity (EHH) at a given SNP and was
12 used as the screening statistic for natural selection¹². We tested the selection signals at genome-wide level in
13 each superpopulation. The minor allele frequency threshold was set to 5%. SNPs with standardized iHS
14 |exceeded 99% quantile (2.25) or under 1% (-2.73) of genome-wide standardized iHS was recognized as
15 significant selection signals. In addition, cross population extended haplotype homozygosity (XP-EHH) was
16 used as a secondary signal to reinforce the evidence of natural selection. We selected European as the
17 reference population when calculating XP-EHH, as most studies on eQTL and critically-ill COVID-19 are
18 based on European ancestry. Hapbin¹³ was used to calculate standardized iHS and XP-EHH. WRIGHT's Fst
19 statistics was used to estimate the degree of differentiation at single SNP level within five superpopulations¹⁴
20 using PLINK v1.9. Vcftools¹⁵ was used to perform region Fst within five superpopulations.

21 **SNP allele frequency in ancient Human genomes**

22 The genotype data of present and ancient DNA obtained from Allen Ancient DNA Resource (AADR)
23 1240K dataset ([https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-genotypes-](https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dnadata)
24 [present-day-and-ancient-dnadata](https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dnadata). Accessed 1 January 2022.) was downloaded and treated by EIGENSOFT.
25 PLINK v1.9¹⁶ were used to exclude duplicated samples, high correlated samples (correlation cutoff 0.5)
26 and samples with high missing genotype rate (>5%). After quality control, 4741 samples with rs17217831

1 genotyping were used for allele frequency analysis (Supplementary table5). Archaeological ages were
2 indicated in years before 1950 CE.

3 **Infer ancient allele frequency from the genotype data of 1000 Genome Project**

4 Since the allele frequency calculated from ancient DNA can be inflated because of limited sample size in
5 certain ages, we applied CLUES¹⁷, an importance sampling approach for approximating the full likelihood
6 function for the selection coefficient, to the genotype data of African to infer the allele frequency trajectory.
7 To validate selection at SNP of interest, CLUES relies on Markov Chain Monte Carlo (MCMC) samples of
8 the gene tree at the SNP. Genotype data is processed as follow steps: (1) We converted the vcf format
9 genotype data into Relate's required haps/sample input file types. SNPs that are not biallelic were removed.
10 Relate v1.1¹⁸ was used to get a tree with genome-wide genealogies. (2) We also estimated the population
11 size, the branch length and the average mutation rate of the African genome. Given the estimated
12 coalescence rate, the MCMC method was used to sample branch lengths. CLUES was used to perform
13 inference and visualize results (plot trajectories). Time bins in the frequency trajectory were set in 4 different
14 epochs: 0-1500, 1500-3000 and 3000-4500 years.

15 **Association analysis of critically-ill COVID-19 SNPs in Africans**

16 The UK Biobank is consisted of data from over 500,000 individuals aged 37–73 years (99.5% were between
17 40 and 69 years of age) who were recruited between 2006 and 2010 from across the UK¹⁹. 3,460 participants
18 reported their ethnic background as African (AFR). We recoded the critically-ill COVID-19 in UK biobank
19 African population. Critically-ill COVID-19 was determined by COVID-related death, Intensive care unit
20 (ICU) admission or ventilation machine use 7 days before and up to 30 days after a patient's first positive test
21 for SARS-CoV-2 and Diagnoses of ICD10. The association between critically-ill COVID-19 and natural
22 selected genes was calculated using PLINK. Age, BMI, sex, the first 5 genetic principal components were
23 used as covariates.

24 **Colocalization, mendelian Randomization Analysis and instrumental variable selection**

25 We further estimated the causal association between *CCR3* blood expression level, eosinophil counts and
26 critically-ill COVID-19. We first performed colocalization test using eQTL data from Genotype-Tissue
27 Expression (GTEx, <https://gtexportal.org/>) database²⁰ and HGI critically-ill COVID-19 GWAS result with R

1 package ‘coloc 2.0’²¹. We selected 575 significant eQTL ($p < 1e-5$) in the colocalization test. To further
2 discover the causal role of *CCR3* and critically-ill COVID-19, two-sample Mendelian randomization analysis
3 was conducted using R package ‘MR-PRESSO’²² which was used to correct for the pleiotropy effect. The
4 weighted median estimator and MR-egger were used as robust tests against MR-PRESSO using the R
5 package ‘TwoSampleMR 0.5.6’²³. Another causal relationship between eosinophil counts and critically-ill
6 COVID-19 was tested using previous reported eosinophil count GWAS summary²⁴ and HGI critically-ill
7 COVID-19 GWAS result. After clumping, 1,193 SNPs were selected as independent instrumental variables.
8 Two-sample Mendelian randomization analysis was performed using same methods. Mendelian
9 randomization power was calculated mRnd²⁵.

10

1 **Results**

2 **SNPs at *CCR3* showed recent natural selection evidence**

3 We investigated the natural selection in Sub-Sahara African population at 8,305 critically-ill COVID-19
4 SNPs, which were reported by recent Host Genetics Initiative (HGI) GWAS, as of date April 2022 (version
5 7). We evaluated the genome-wide integrated haplotype score (iHS) in Sub-Sahara Africans.

6 The most significant selection signal was rs13086063, located in upstream of *CCR3*, with its standardized
7 iHS = 6.31 exceeding 99.97% quantile of whole genome standardized iHS (Fig.1 A, Table 1). The natural
8 selection signal on *CCR3* was also validated by XP-EHH. The top standardized XP-EHH of *CCR3* SNP
9 rs3091311 was 2.57, which is greater than the 99% quantile of whole genome standardized XP-EHH, 2.04
10 (Fig.1 B, Table 1). We also identified a *CCR3* SNP rs17217831 showed natural selection validated by two
11 index (standardized iHS = 2.49 & standardized XP-EHH = 2.00).

12 *CCR3* showed population differentiation as evidenced both by single SNP F_{st} (Fig 1.C, Supplementary
13 table 1) and regional F_{st} (weighted F_{st} = 0.18 for *CCR3* upstream chr3: 45900001- 46200000 and weighted
14 F_{st} = 0.10 for *CCR3* chr3: 46200001-46500000) within five superpopulations in 1000 genome data.

15 **Allele frequency trajectories across ancient time**

16 By looking through the frequency trajectory from ancient time to present, we found the evidence of
17 positive selection that induced the descent of the critically-ill COVID-19 risk allele *CCR3* rs17217831-A in
18 Sub-Saharan Africans. Using genotypes of ancient DNA²⁶, we calculated the allele frequency of *CCR3*
19 rs17217831-A in three archaeological ages starting from 1950 CE (present, 0-1500, 1500+ years).

20 We detected a declining trend of *CCR3* rs17217831-A frequency in Africans. The frequency showed a
21 trend of descent from 23% at 1500+ years to 6% at present (Figure 1.D & Supplementary table 2). However,
22 the trajectory was sparse due to the limited number of samples in certain ages. We further inferred the
23 frequency trajectory from the genotype data at present based on likelihood¹⁷. The inferred frequency
24 trajectory coincided well with the frequencies calculated from ancient DNA. We observed an inflating trend
25 from 3000 years ago, but a sharp decline in roughly 1500 years, indicating recent unfavored selection on the
26 critically-ill allele (Figure 1.E).

27 **Natural selection on *CCR3* eQTLs in African**

1 We identified four eQTLs (rs11926063, rs13062523, rs35984988 and rs41537946) showed suggestive-
2 level ($iHS > 1$ and $XP-EHH > 1$) natural selection signals. Importantly, these eQTLs exhibited negative
3 correlations ($BETAs < -0.1$) with *CCR3* expression in African (Figure 1F) showing the expression of *CCR3*
4 may increase following natural selection in African.

5 **Association between SNPs under natural selection and critically-ill COVID-19 in African ancestry**

6 We next investigated the association between critically-ill COVID-19 and *CCR3* rs17217831 in cohorts
7 of African ancestry. As of May 2021, 504 participants of African ancestry in the UK Biobank were tested
8 positive for SARS-CoV-2 infection, of which 36 presented with critical illness. *CCR3* rs17217831 was not
9 associated with critical illness ($p = 0.381$). The meta-analysis in HGI African GWAS results (version 7) also
10 did not show significant association ($p = 0.56$). These insignificant associations could be attributed to the low
11 frequency of the risk allele in modern African population (risk allele frequency *CCR3* rs17217831 =0.05).

12 **Causal relationship between *CCR3* expression and critically-ill COVID-19**

13 The instrumental variables for *CCR3* expression were obtained from GTEx database. We obtained 575
14 *CCR3* eQTLs in blood samples of European ancestry and 5 remained for further analysis after LD-clumping
15 ($P < 5 \times 10^{-7}$, Supplementary Table 3). Using these eQTLs, we found the genetic susceptibility of *CCR3*
16 expression and critically-ill COVID-19 were colocalized. The probability of the causal variants of these
17 phenotypes sharing the same locus was 1 (Fig.2 A). A significant horizontal pleiotropic effect of these eQTL
18 was estimated (MR-PRESSO global test $p < 0.001$), which also indicated the shared genetic basis between
19 *CCR3* expression level and critical illness. We further conducted a Mendelian Randomization analysis which
20 corrected for this pleiotropy effect. The result revealed that *CCR3* expression level (the exposure) was
21 causally associated with decreased risk of critical illness (OR = 0.89 [0.87, 0.92], $p = 0.015$, Fig.2 B). The
22 robust test with weighted median model obtained the consistent result (Supplementary Table 4). Our results
23 indicated *CCR3* as another important therapeutic target for critically-ill COVID-19.

24

1 Discussion

2 Coronavirus disease (COVID-19) presents a wide range of clinical manifestations, from asymptomatic
3 to severe clinical course. Recent studies have shown that ethnicity is relevant to severe COVID-19
4 outcomes^{27,28}. Genes have undergone natural selection in sub-Saharan Africa. For example, studies have
5 shown that ADH gene regions, which includes the ADH1B*48His allele, has experienced recent positive
6 selection in multiple agriculturalist populations from sub-Saharan Africa over the past 2,000 years or so²⁹. In
7 this study, we confirmed that the protective alleles of critically-ill COVID-19, at least in part, were more
8 prevalent in African people due to natural selection. We identified significant selection signal of *CCR3*
9 rs17217831-A in Sub-Saharan African. We further revealed that *CCR3* expression level was causally
10 associated with decreased risk of critical illness.

11 A study identified that a chromosome 3 haplotype which was inherited from Neanderthals, increases the
12 risk of severe COVID-19³⁰. Close to the locus where this haplotype is identified, locates two chemokine
13 receptor genes, *CCR2* which has been associated with severe COVID-19³¹, and *CCR3*. *CCR2* is a
14 monocyte/macrophage specific chemokine receptor, while *CCR3* is expressed mainly in eosinophils and
15 basophils, but also a subset of Th2 lymphocytes, mast cells and myeloid-derived suppressor cells (MDSC)
16 (Figure 2.B and C)³²⁻³⁴. Recent studies have confirmed that these *CCR3*+ cells involved in a variety of
17 diseases including asthma³⁵, tumors³⁶ and parasitic infection³⁷. Over the past few years, extensive research
18 has confirmed Chemokine signaling through *CCR3* is a key regulatory pathway for eosinophil associated
19 with allergic inflammation and asthma^{38,39}. Furthermore, investigations have indicated that the activation of
20 *CCR3* on Th-2 cells and basophils can stimulate the release of inflammatory factors such as IL-4, thereby
21 facilitating the activation of eosinophils^{40,41}. Notably, in the context of chronic *Staphylococcus aureus*
22 infection, MDSCs exhibit eosinophilic granules and express variable levels of *CCR3*³⁴. These specialized
23 cells, known as Eo-MDSCs, accumulate at the infection site, playing a substantial role in exacerbating the
24 infection. These findings collectively underscore the indispensable role of eosinophils, orchestrated through
25 the *CCR3* surface receptor, in the pathogenesis of various diseases. Of note, the *CCR3* rs17217831-A under
26 selection in Africans was associated with lower eosinophil counts⁴² which was also causally associated with
27 higher risk of critically-ill COVID-19 in our Mendelian randomization analysis. As a result, our study

1 indicated that *CCR3* rs17217831-A may increase the risk of critically-ill COVID-19 either by reducing blood
2 *CCR3* expression level or eosinophil counts. In support for this notion, recent *in vivo* studies showed that
3 *CCR3* blockage or knockout could inhibit the recruitment and migration of eosinophils to the blood, airway
4 and lung airspace^{38,43}. Multiple studies of the immune profile of COVID-19 patients have also demonstrate
5 that decreased eosinophil count is associated with a worse disease outcome⁴⁴⁻⁴⁶.

6 Nevertheless, the role of eosinophils in the progression of COVID-19 remains controversial and the
7 underlying reasons deserve further investigations. Lucas et. al. reported that increased eosinophil counts was
8 associated with severe COVID-19 in 253 US residents⁴⁷. To be noted, the participants in this study achieved
9 a relatively high percentage (29.2%) of black people as compared with those in other reports, while ethnicity
10 was not adjusted in their analysis. On the other hand, excessive eosinophils in lung could also lead to a
11 specific type of pneumonia⁴⁸. However, this may represent a distinct etiological pathway as from the
12 pneumonia induced by SARS-cov-2 infection. From a physiology point of view, eosinophils are both pro-
13 inflammatory and anti-viral. It is plausible to believe that eosinophils help reducing the viral load in the early
14 stages of COVID-19, while cause respiratory failure in the late stages by a synergetic action with the cytokine
15 storm. As a result, observed association between eosinophil level and disease severity may vary by the time
16 when samples were taken. In addition, most of the studies investigated the eosinophil counts in peripheral
17 blood, while the status of blood eosinophils could be different from their lung-resident counterparts. Using
18 blood eosinophil as a proxy of the severity of lung disease may lead to a confounded conclusion. Furthermore,
19 the limited sample size and the difference of ethnic composition among cohorts could also contribute to the
20 controversial results⁴⁹.

21 *IFNAR2* showed a suggestive selection signal in African population with a missense variant rs2229207
22 standardized iHS = 1.85, which exceeded 94.3% quantile of African genome-wide standardized iHS. *IFNAR2*
23 encodes the type I interferon receptor (IFN- α/β)⁵⁰. It was associated with critically-ill COVID-19⁵¹ and
24 confirmed by a Mendelian randomization analysis aimed for drug repurposing⁵² which validate the rationale
25 of using natural selection to detect drug target.

26 Our study was limited in the genomic analysis of 1000 Genomes Project data. No biological experiments
27 were undertaken to ascertain the involvement of *CCR3* or eosinophils in severe COVID-19 disease. Future

1 studies should validate the therapeutic efficacy of *CCR3* and eosinophils in the context of critically-ill
2 COVID-19.

3 Taken together, the *CCR3* risk allele of critical ill COVID-19 was under recent natural selection in
4 Africans, which may make Africans less vulnerable to the critical illness. A possible mechanism of this
5 protective effect could originate from a higher eosinophil count. However, the mechanism and efficacy of
6 regulating eosinophil in the treatment of critically-ill COVID-19 shall be further investigated.

7 **Conclusion**

8 From a population genetics point-of-view, our study indicated *CCR3* as a potential therapeutic target for
9 critically-ill COVID-19. The indicated therapeutic mechanism builds a bridge between the natural selection
10 of the critical illness allele and the development of novel critical illness treatments.

11

1 **Acknowledgments**

2 This study used the UK Biobank resources under the application number 53562. We thank the David Reich
3 Lab for providing Ancient DNA genotype data. We would also acknowledge the participants and
4 investigators of the FinnGen study.

5 **Author Contributions**

6 P.C conceptualized and supervised the study. P.C collected the data. Z.S and A.T analyzed the data. P.C,
7 Z.S and L.P interpreted the results. Z.S and P.C wrote the original manuscript. P.C revised the manuscript.
8 All authors approved the submission.

9 **Competing Interest Statement**

10 The authors declared no competing interests.

11 **Funding**

12 Not applicable.

13 **Data Availability Statement**

14 The individual-level genetic data are available on 1000 genome project website
15 (<https://www.internationalgenome.org/>) and David Reich Lab website (<https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data>). IHS, XPEHH, Fst
16 values will be available through application after publication of this study.
17

18

1 References

- 2 1 Hamilton, M. A., Hamilton, D., Soneye, O., Ayeyemi, O. & Jaradat, R. An analysis of the
3 impact of policies and political affiliation on racial disparities in COVID-19 infections and
4 deaths in the USA. *International Journal of Data Science and Analytics* **13**, 63-76,
5 doi:10.1007/s41060-021-00284-y (2022).
- 6 2 Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the
7 Coronavirus Disease 2019 (COVID-19) Outbreak in China. *JAMA* **323**, 1239,
8 doi:10.1001/jama.2020.2648 (2020).
- 9 3 Mackey, K. *et al.* Racial and Ethnic Disparities in COVID-19-Related Infections,
10 Hospitalizations, and Deaths : A Systematic Review. *Ann Intern Med* **174**, 362-373,
11 doi:10.7326/M20-6306 (2021).
- 12 4 Price-Haywood, E. G., Burton, J., Fort, D. & Seoane, L. Hospitalization and Mortality
13 among Black Patients and White Patients with Covid-19. *N Engl J Med* **382**, 2534-2543,
14 doi:10.1056/NEJMsa2011686 (2020).
- 15 5 McPadden, J. *et al.* Clinical characteristics and outcomes for 7,995 patients with SARS-
16 CoV-2 infection. *PLoS One* **16**, e0243291, doi:10.1371/journal.pone.0243291 (2021).
- 17 6 Ogedegbe, G. *et al.* Assessment of Racial/Ethnic Disparities in Hospitalization and
18 Mortality in Patients With COVID-19 in New York City. *JAMA Netw Open* **3**, e2026881,
19 doi:10.1001/jamanetworkopen.2020.26881 (2020).
- 20 7 Huffman, J. E. *et al.* Multi-ancestry fine mapping implicates OAS1 splicing in risk of
21 severe COVID-19. *Nat Genet* **54**, 125-127, doi:10.1038/s41588-021-00996-8 (2022).
- 22 8 Randolph, H. E. *et al.* Genetic ancestry effects on the response to viral infection are
23 pervasive but cell type specific. *Science* **374**, 1127-1133, doi:10.1126/science.abg0928
24 (2021).
- 25 9 Enard, D. & Petrov, D. A. Evidence that RNA Viruses Drove Adaptive Introgression
26 between Neanderthals and Modern Humans. *Cell* **175**, 360-371 e313,
27 doi:10.1016/j.cell.2018.08.034 (2018).
- 28 10 Zhang, C. *et al.* Impact of natural selection on global patterns of genetic variation and
29 association with clinical phenotypes at genes involved in SARS-CoV-2 infection. *Proc Natl
30 Acad Sci U S A* **119**, e2123000119, doi:10.1073/pnas.2123000119 (2022).
- 31 11 The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host
32 genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic.
33 *European Journal of Human Genetics* **28**, 715-718, doi:10.1038/s41431-020-0636-6
34 (2020).
- 35 12 Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A Map of Recent Positive Selection
36 in the Human Genome. *PLoS Biology* **4**, e72, doi:10.1371/journal.pbio.0040072 (2006).
- 37 13 Maclean, C. A., Chue Hong, N. P. & Prendergast, J. G. hapbin: An Efficient Program for
38 Performing Haplotype-Based Scans for Positive Selection in Large Genomic Datasets.
39 *Mol Biol Evol* **32**, 3027-3029, doi:10.1093/molbev/msv172 (2015).
- 40 14 Wright, S. Evolution in Mendelian Populations. *Genetics* **16**, 97-159,
41 doi:10.1093/genetics/16.2.97 (1931).
- 42 15 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158,
43 doi:10.1093/bioinformatics/btr330 (2011).
- 44 16 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based
45 linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

- 1 17 Stern, A. J., Wilton, P. R. & Nielsen, R. An approximate full-likelihood method for
2 inferring selection and allele frequency trajectories from DNA sequence data. *PLoS*
3 *Genet* **15**, e1008384, doi:10.1371/journal.pgen.1008384 (2019).
- 4 18 Speidel, L., Forest, M., Shi, S. & Myers, S. R. A method for genome-wide genealogy
5 estimation for thousands of samples. *Nat Genet* **51**, 1321-1329, doi:10.1038/s41588-
6 019-0484-x (2019).
- 7 19 Rusk, N. The UK Biobank. *Nature Methods* **15**, 1001-1001, doi:10.1038/s41592-018-
8 0245-2 (2018).
- 9 20 Consortium, G. T. The GTEx Consortium atlas of genetic regulatory effects across human
10 tissues. *Science* **369**, 1318-1330, doi:10.1126/science.aaz1776 (2020).
- 11 21 Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic
12 Association Studies Using Summary Statistics. *PLoS Genetics* **10**, e1004383,
13 doi:10.1371/journal.pgen.1004383 (2014).
- 14 22 Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Publisher Correction: Detection of
15 widespread horizontal pleiotropy in causal relationships inferred from Mendelian
16 randomization between complex traits and diseases. *Nat Genet* **50**, 1196,
17 doi:10.1038/s41588-018-0164-2 (2018).
- 18 23 Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between
19 imprecisely measured traits using GWAS summary data. *PLoS Genet* **13**, e1007081,
20 doi:10.1371/journal.pgen.1007081 (2017).
- 21 24 Vuckovic, D. *et al.* The Polygenic and Monogenic Basis of Blood Traits and Diseases. *Cell*
22 **182**, 1214-1231 e1211, doi:10.1016/j.cell.2020.08.008 (2020).
- 23 25 Brion, M. J., Shakhbazov, K. & Visscher, P. M. Calculating statistical power in Mendelian
24 randomization studies. *Int J Epidemiol* **42**, 1497-1501, doi:10.1093/ije/dyt179 (2013).
- 25 26
- 26 *David Reich Lab. Allen Ancient DNA Resource (AADR): Downloadable genotypes*
27 *of present-day and ancient DNA data, version 42.4.*
28 [https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-](https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dnadata)
29 [genotypes-present-day-and-ancient-dnadata](https://reich.hms.harvard.edu/allenancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dnadata). Accessed 1 January 2022.
- 30 27 Velavan, T. P. *et al.* Host genetic factors determining COVID-19 susceptibility and
31 severity. *EBioMedicine* **72**, 103629, doi:10.1016/j.ebiom.2021.103629 (2021).
- 32 28 Niedzwiedz, C. L. *et al.* Ethnic and socioeconomic differences in SARS-CoV-2 infection:
33 prospective cohort study using UK Biobank. *BMC Med* **18**, 160, doi:10.1186/s12916-020-
34 01640-8 (2020).
- 35 29 McQuillan, M. A. *et al.* Signatures of Convergent Evolution and Natural Selection at the
36 Alcohol Dehydrogenase Gene Region are Correlated with Agriculture in Ethnically
37 Diverse Africans. *Mol Biol Evol* **39**, doi:10.1093/molbev/msac183 (2022).
- 38 30 Zeberg, H. & Pääbo, S. The major genetic risk factor for severe COVID-19 is inherited
39 from Neanderthals. *Nature* **587**, 610-612, doi:10.1038/s41586-020-2818-3 (2020).
- 40 31 Pairo-Castineira, E. *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* **591**,
41 92-98, doi:10.1038/s41586-020-03065-y (2021).
- 42 32 Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**,
43 1260419, doi:10.1126/science.1260419 (2015).
- 44 33 Willems, L. I. & Ijzerman, A. P. Small molecule antagonists for chemokine CCR3
45 receptors. *Med Res Rev* **30**, 778-817, doi:10.1002/med.20181 (2010).

- 1 34 Goldmann, O., Beineke, A. & Medina, E. Identification of a Novel Subset of Myeloid-Derived Suppressor Cells During Chronic Staphylococcal Infection That Resembles
2 Immature Eosinophils. *J Infect Dis* **216**, 1444-1451, doi:10.1093/infdis/jix494 (2017).
3
- 4 35 Bertrand, C. P. & Ponath, P. D. CCR3 blockade as a new therapy for asthma. *Expert Opin Investig Drugs* **9**, 43-52, doi:10.1517/13543784.9.1.43 (2000).
5
- 6 36 Korbecki, J. *et al.* CC Chemokines in a Tumor: A Review of Pro-Cancer and Anti-Cancer Properties of the Ligands of Receptors CCR1, CCR2, CCR3, and CCR4. *Int J Mol Sci* **21**,
7 doi:10.3390/ijms21218412 (2020).
8
- 9 37 Turner, J. D. *et al.* Interleukin-4 activated macrophages mediate immunity to filarial
10 helminth infection by sustaining CCR3-dependent eosinophilia. *PLoS Pathog* **14**,
11 e1006949, doi:10.1371/journal.ppat.1006949 (2018).
- 12 38 Grozdanovic, M. *et al.* Novel peptide nanoparticle-biased antagonist of CCR3 blocks
13 eosinophil recruitment and airway hyperresponsiveness. *J Allergy Clin Immunol* **143**,
14 669-680 e612, doi:10.1016/j.jaci.2018.05.003 (2019).
- 15 39 Shah, S. A. *et al.* Platelets Independently Recruit into Asthmatic Lungs and Models of
16 Allergic Inflammation via CCR3. *Am J Respir Cell Mol Biol* **64**, 557-568,
17 doi:10.1165/rcmb.2020-0425OC (2021).
- 18 40 Cheng, L. E. *et al.* IgE-activated basophils regulate eosinophil tissue entry by modulating
19 endothelial function. *J Exp Med* **212**, 513-524, doi:10.1084/jem.20141671 (2015).
- 20 41 Foster, P. S. *et al.* Modeling T(H) 2 responses and airway inflammation to understand
21 fundamental mechanisms regulating the pathogenesis of asthma. *Immunol Rev* **278**, 20-
22 40, doi:10.1111/imr.12549 (2017).
- 23 42 Chen, M. H. *et al.* Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667
24 Individuals from 5 Global Populations. *Cell* **182**, 1198-1213 e1114,
25 doi:10.1016/j.cell.2020.06.045 (2020).
- 26 43 Wen, T., Besse, J. A., Mingler, M. K., Fulkerson, P. C. & Rothenberg, M. E. Eosinophil
27 adoptive transfer system to directly evaluate pulmonary eosinophil trafficking in vivo.
28 *Proc Natl Acad Sci U S A* **110**, 6067-6072, doi:10.1073/pnas.1220572110 (2013).
- 29 44 Zein, J. G. *et al.* Eosinophilia Is Associated with Improved COVID-19 Outcomes in Inhaled
30 Corticosteroid-Treated Patients. *J Allergy Clin Immunol Pract* **10**, 742-750 e714,
31 doi:10.1016/j.jaip.2021.12.034 (2022).
- 32 45 Ferastroaru, D. *et al.* Eosinophilia in Asthma Patients Is Protective Against Severe
33 COVID-19 Illness. *J Allergy Clin Immunol Pract* **9**, 1152-1162 e1153,
34 doi:10.1016/j.jaip.2020.12.045 (2021).
- 35 46 Qin, C. *et al.* Dysregulation of Immune Response in Patients With Coronavirus 2019
36 (COVID-19) in Wuhan, China. *Clin Infect Dis* **71**, 762-768, doi:10.1093/cid/ciaa248
37 (2020).
- 38 47 Lucas, C. *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19.
39 *Nature* **584**, 463-469, doi:10.1038/s41586-020-2588-y (2020).
- 40 48 De Giacomo, F., Vassallo, R., Yi, E. S. & Ryu, J. H. Acute Eosinophilic Pneumonia. Causes,
41 Diagnosis, and Management. *Am J Respir Crit Care Med* **197**, 728-736,
42 doi:10.1164/rccm.201710-1967CI (2018).
- 43 49 Glickman, J. W., Pavel, A. B., Guttman - Yassky, E. & Miller, R. L. The role of circulating
44 eosinophils on COVID - 19 mortality varies by race/ethnicity. *Allergy* **76**, 925-927,
45 doi:10.1111/all.14708 (2021).

- 1 50 Wu, L., Zhu, J., Liu, D., Sun, Y. & Wu, C. An integrative multiomics analysis identifies
2 putative causal genes for COVID-19 severity. *Genetics in Medicine* **23**, 2076-2086,
3 doi:10.1038/s41436-021-01243-5 (2021).
- 4 51 Akter, S., Roy, A., Tonmoy, M. & Islam, M. in silico Deleterious single nucleotide
5 polymorphisms (SNPs) of human IFNAR2 gene facilitate COVID-19 severity in patients: a
6 comprehensive approach. *Journal of biomolecular structure & dynamics*, 1-17,
7 doi:10.1080/07391102.2021.1957714 (2021).
- 8 52 Gaziano, L. *et al.* Actionable druggable genome-wide Mendelian randomization
9 identifies repurposing opportunities for COVID-19. *Nat Med* **27**, 668-676,
10 doi:10.1038/s41591-021-01310-z (2021).

11

12

1 **Table**

2 **Table 1. Natural selection signatures for *CCR3* SNPs**

GENE	SNP	CHR:POS	Ancestry	Derived	AFR standardized XPEHH	AFR standardized iHS	EUR standardized iHS	Fst	AFR Freq
<i>CCR3</i>	rs13086063	3:46200860	T	C	0.81	6.31	0.41	0.10	0.22
<i>CCR3</i>	rs3091311	3:46306476	T	C	2.57	1.52	0.11	0.06	0.38
<i>CCR3</i>	rs17217831	3:46305441	C	A	2.49	2.00	0.55	0.07	0.06

3

4 Ancestry, ancestry allele; Derived, derived allele; standardized XPEHH, standardized XP-EHH value

5 calculated in African population; AFR standardized iHS, iHS calculated in African population; EUR

6 standardized iHS, iHS calculated in European population; Fst, measure of population differentiation in

7 1000G five superpopulation (European, East Asian, South Asian, American and African); AFR Freq, derived

8 allele frequency in African population.

9

1 **Figure 1. Natural selection signals of critically-ill COVID-19 SNPs.**

2 (A&B). Manhattan plot of critically-ill COVID-19 SNPs standardized iHS and standardized XP-EHH

3 calculated between African and European. The x-axis is the chromosomal position and y-axis is the

4 significance of nature selection. Red lines indicate genome-wide significance level (standardized iHS |&

5 standardized XP-EHH > 99% quantile statics). Red points indicate the significant variants. (C). Locus plot

6 for plotting Fst of variants in CCR. Linkage disequilibrium information is from 1000 Genomes African.

7 (D). Frequency trajectory of CCR3 rs17217831 from ancient to present. (E). Frequency trajectory of CCR3

8 rs17217831 inferred from 1000 Genomes African genotype. The x-axis is generation before present.

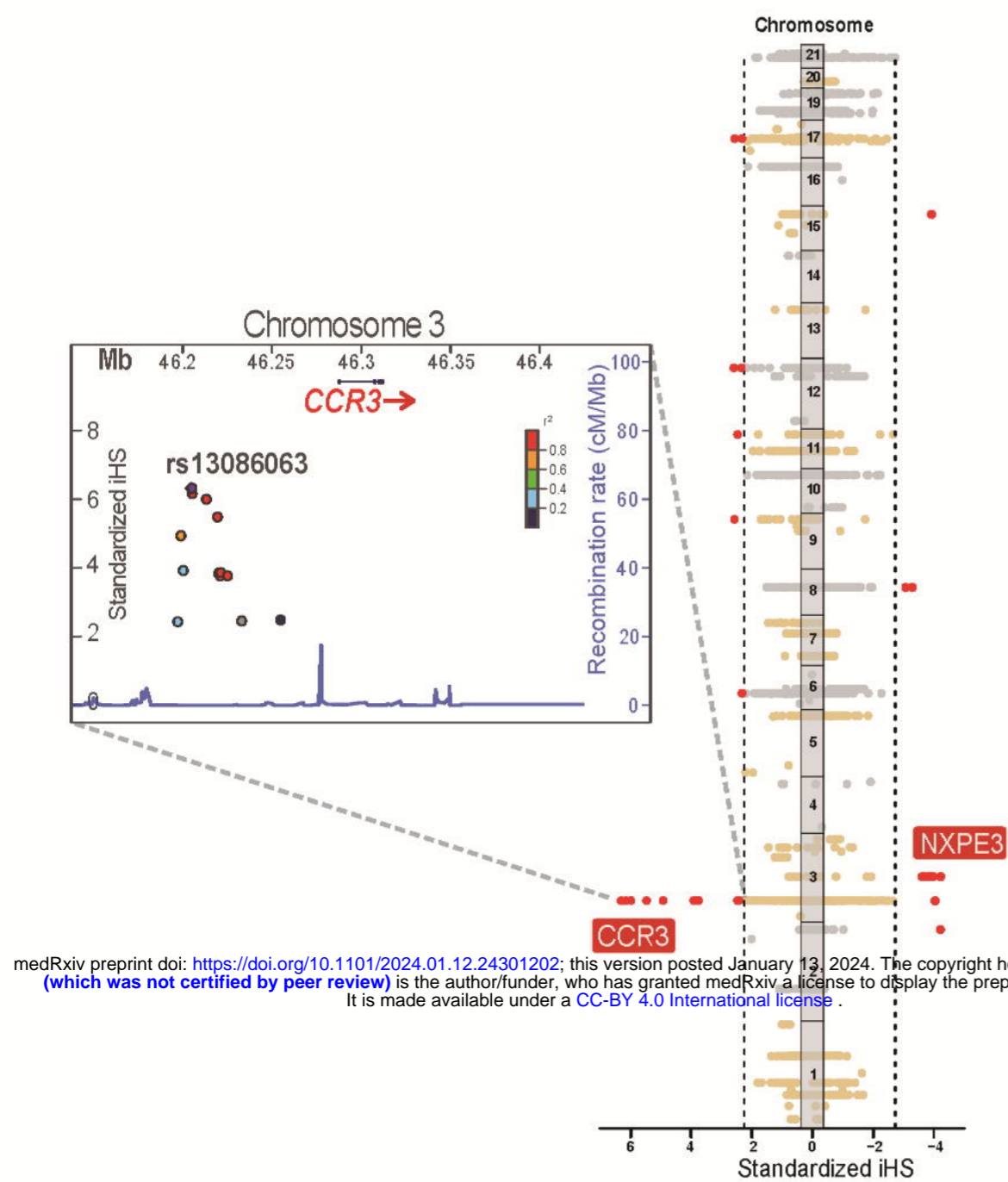
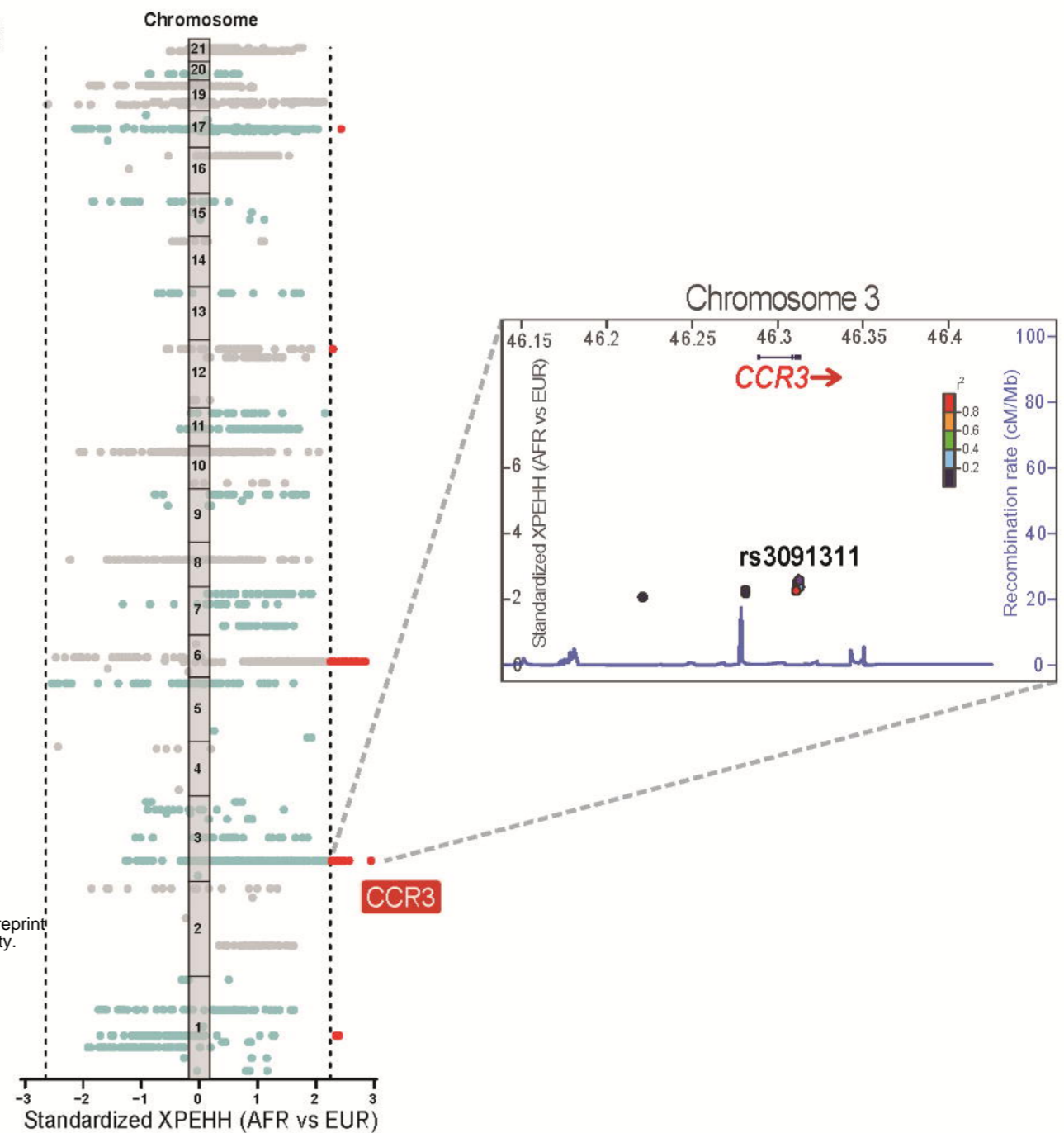
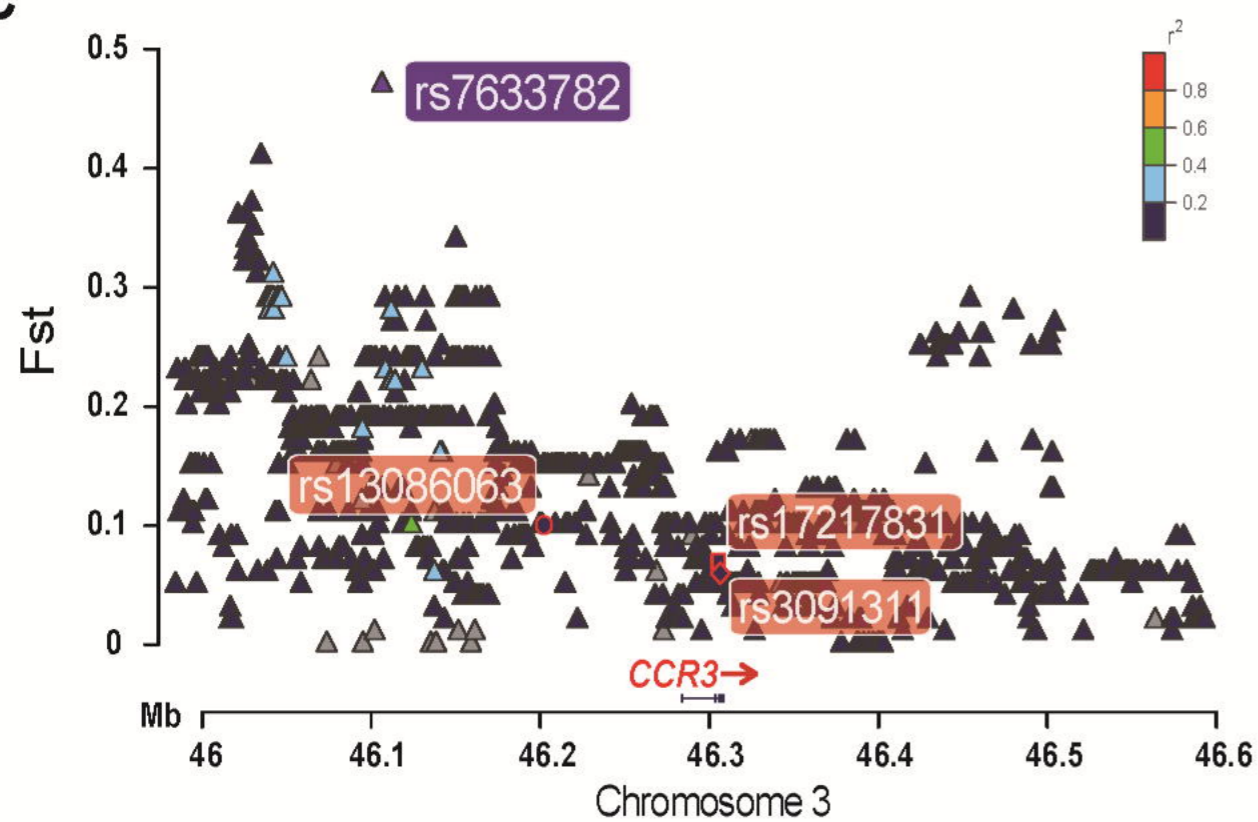
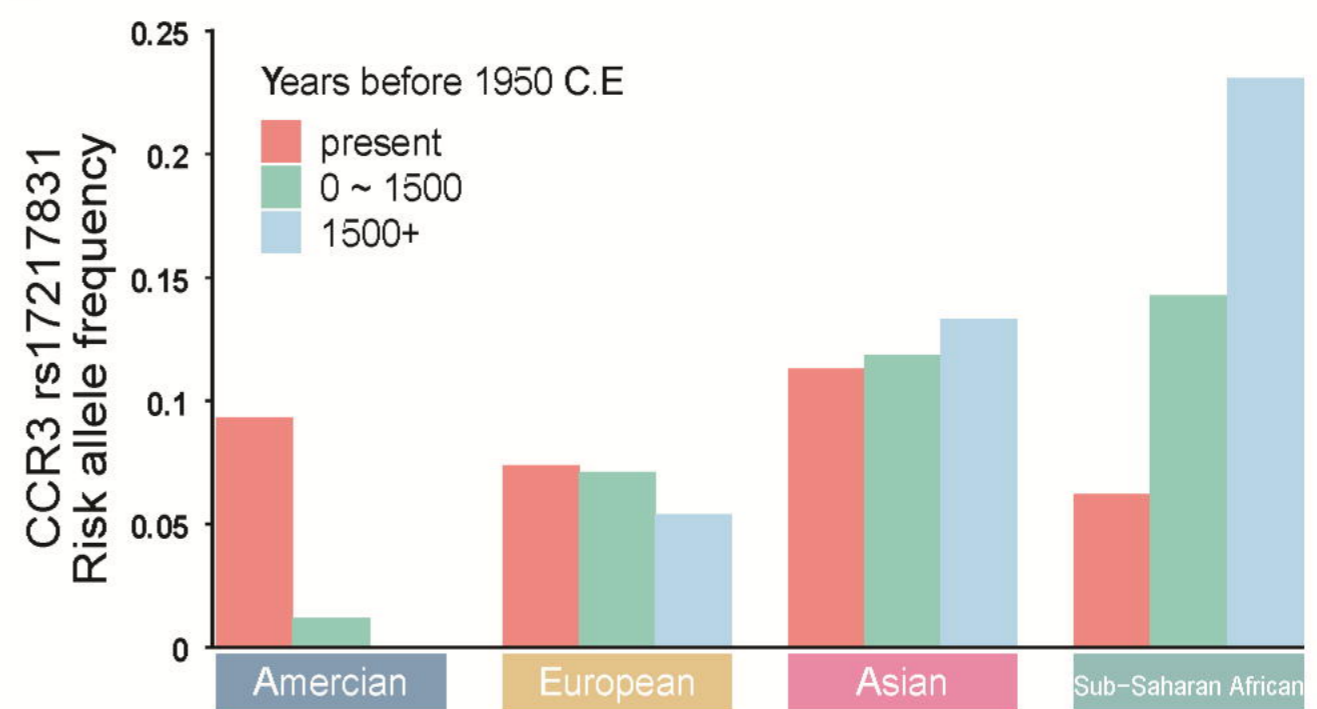
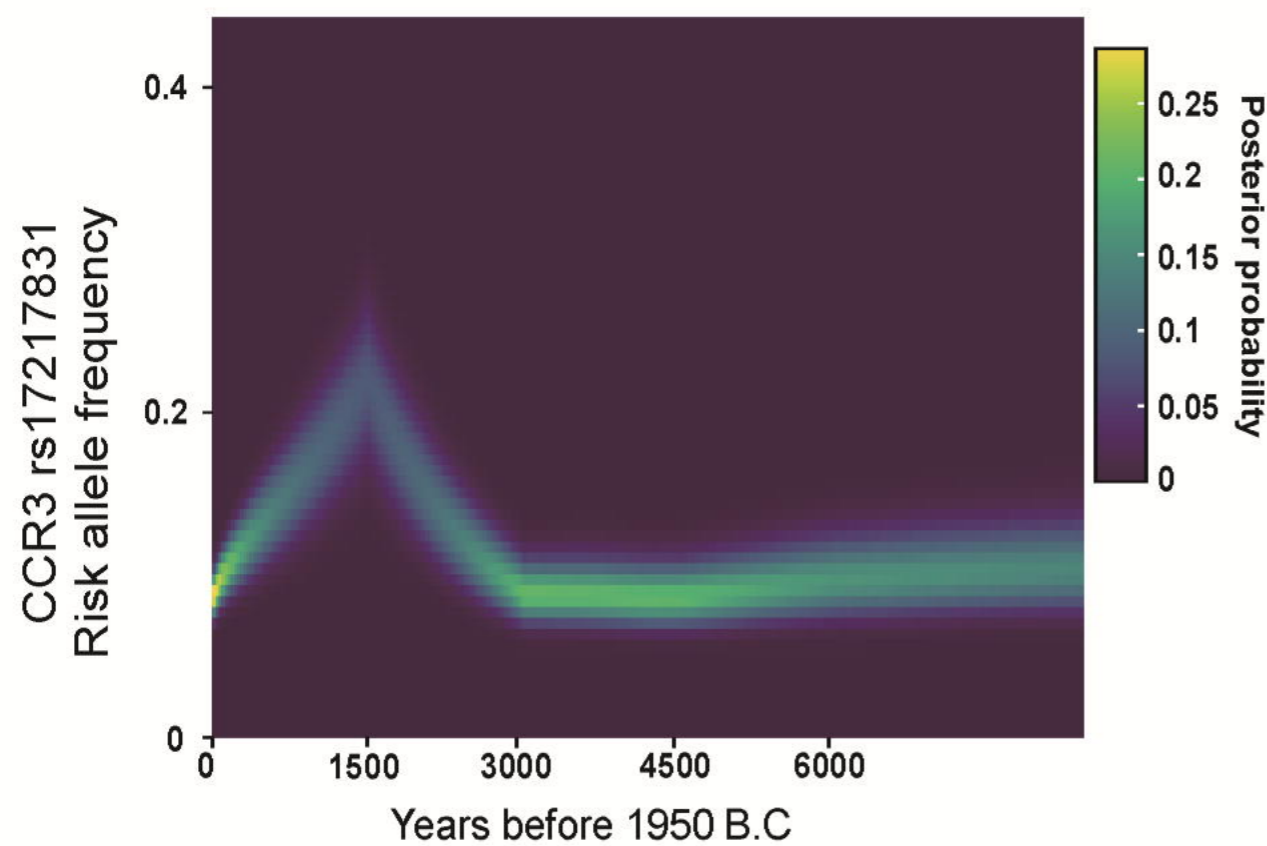
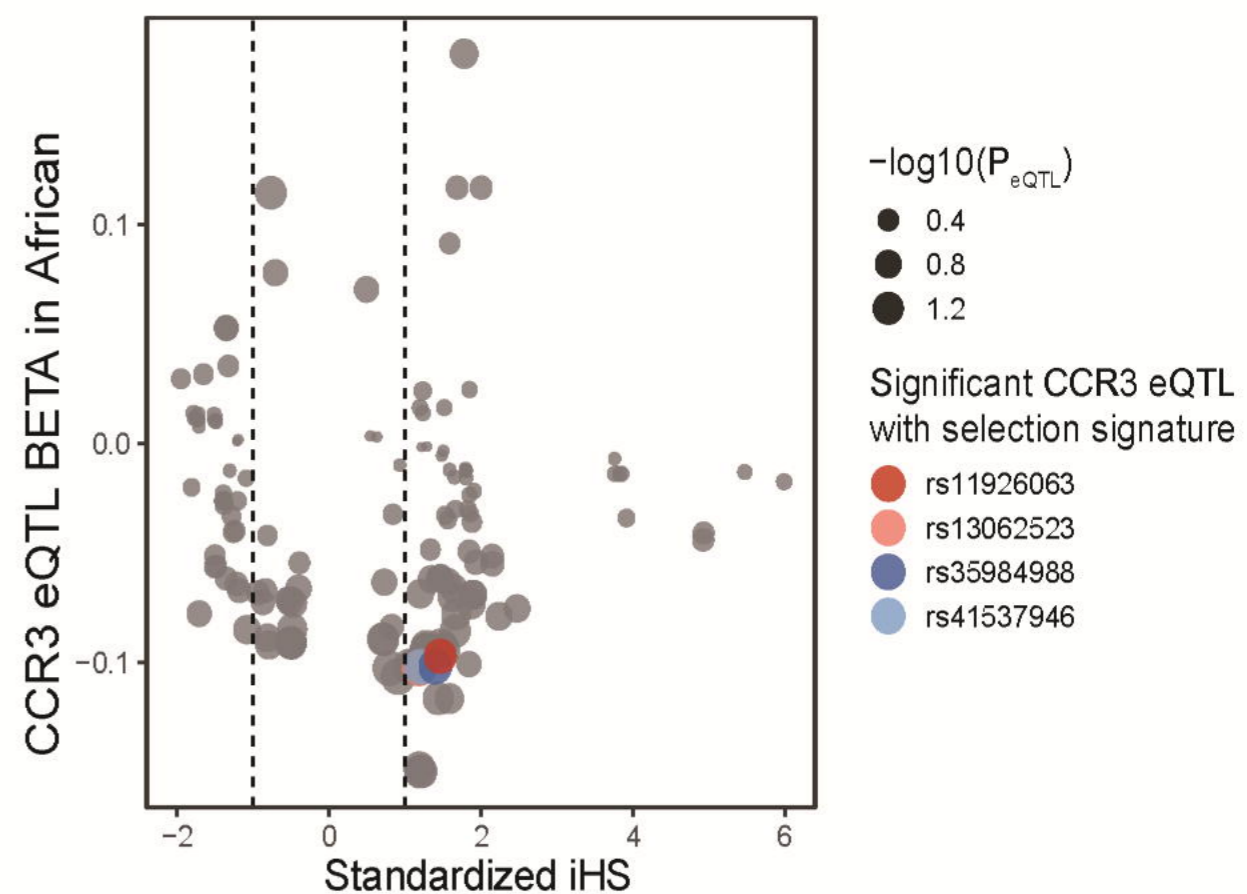
9 Yellow linear shows high probability of such frequency trajectory happened. (F) BETA, effect size of

10 eQTL.

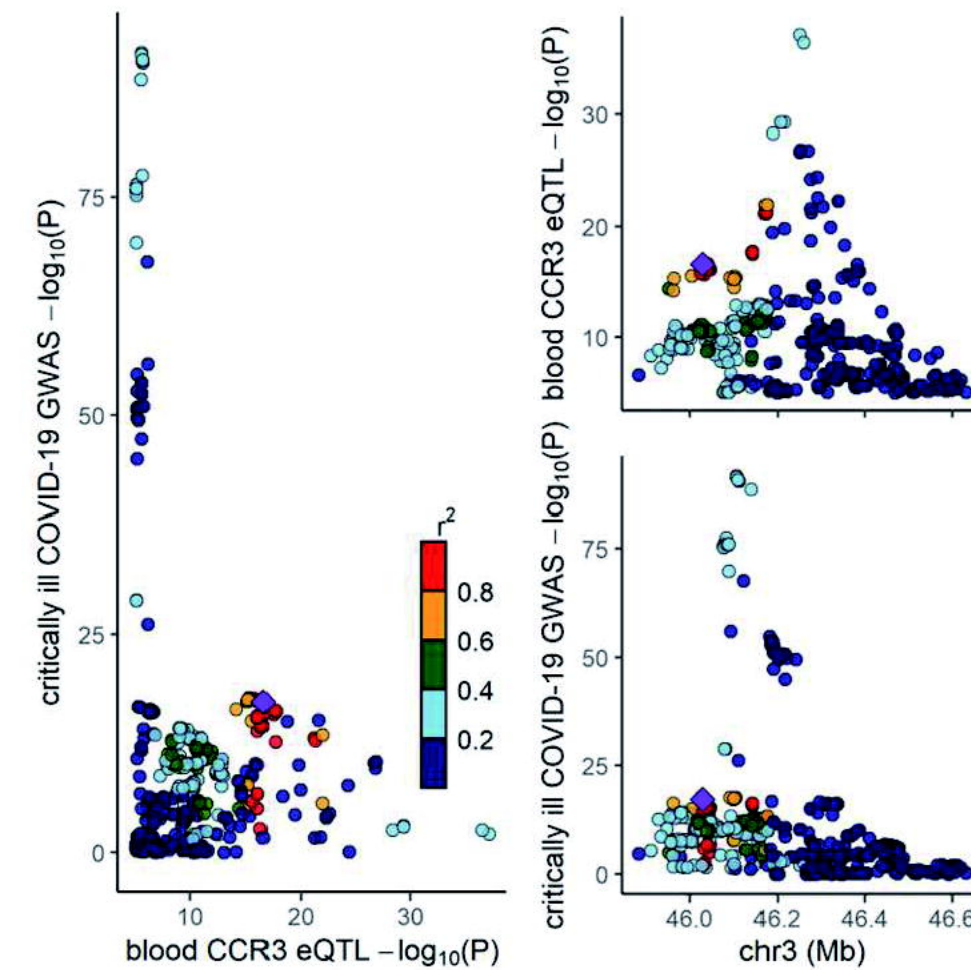
11

1 **Figure 2. Colocalization and mendelian randomization analysis for CCR3 and critically-ill COVID-**
2 **19.**

3 (A) Comparing the significant signals of *CCR3* eQTL and critically-ill COVID-19 in a same locus. r^2
4 indicates the linkage disequilibrium of these variants. (B) Relationship between *CCR3* rs17217831, blood
5 *CCR3* expression, critically-ill COVID-19 and eosinophil counts. Red arrows mean they are positively
6 correlated. Blue blunts mean they are negatively correlated. (C) Expression levels of *CCR3* in different
7 immune cell

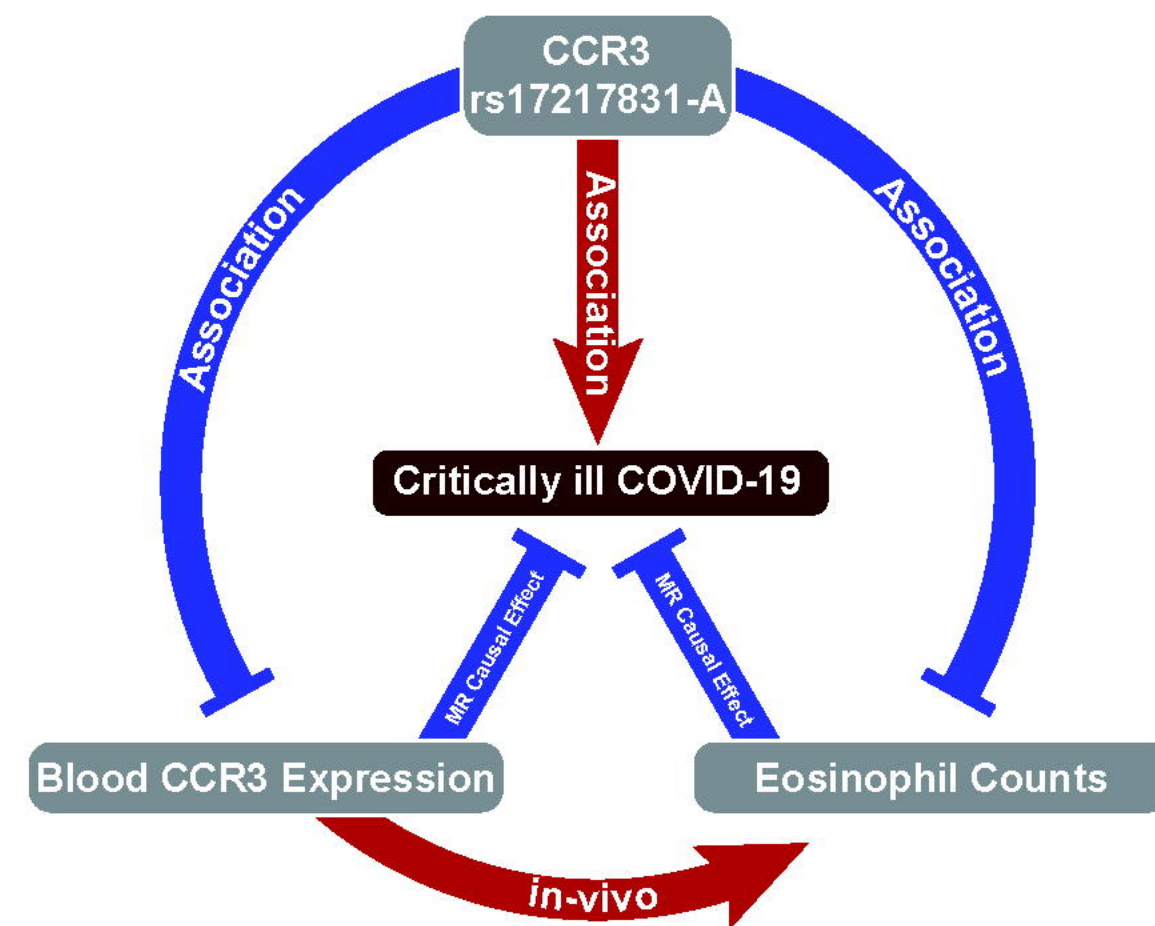
A**B****C****D****E****F**

A



B

Negative Selection Evidence in AFR



C

Immune cell type specificity: Group enriched (basophil, eosinophil)

