

1 **TITLE OF PAPER:** Genetic Variant rs1205 is Associated with COVID-19 Outcomes:
2 The Strong Heart Study and Strong Heart Family Study.

3 **AUTHORS:** Best LG,^{1,2} Erdei E,³ Haack K,⁴ Kent JW Jr.,⁴ Malloy K,⁵ Newman DE,⁴
4 O'Leary M,¹ O'Leary R,¹ Sun Q,⁶ Navas-Acien A,⁷ Franceschini N,⁸ and Cole SA,⁴
5

6
7 ¹ Epidemiology Division, Missouri Breaks Industries Research, Inc. Eagle Butte, SD,
8 USA ² Pathology Department, University of North Dakota, Grand Forks, ND, USA
9 ³ Pharmaceutical Sciences, University of New Mexico - Albuquerque, Albuquerque, New
10 Mexico, USA. ⁴ Population Health Program, Texas Biomedical Research Institute, San
11 Antonio, TX, USA. ⁵ Center for American Indian Health Research, Department of
12 Biostatistics and Epidemiology, Hudson College of Public Health, University of
13 Oklahoma Health Sciences Center, Oklahoma City, OK, USA. ⁶ Department of
14 Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA,
15 ⁷ Department of Environmental Health Science, Mailman School of Public Health,
16 Columbia University, NY, New York, USA. ⁸ Department of Epidemiology, University of
17 North Carolina at Chapel Hill, Chapel Hill, NC, USA.
18

19

20 Abstract:

21 Background: Although COVID-19 infection has been associated with a number of
22 clinical and environmental risk factors, host genetic variation has also been associated
23 with the incidence and morbidity of infection. The *CRP* gene codes for a critical
24 component of the innate immune system and *CRP* variants have been reported
25 associated with infectious disease and vaccination outcomes. We investigated possible
26 associations between COVID-19 outcome and a limited number of candidate gene
27 variants including rs1205.

28 Methodology/Principal Findings: The Strong Heart and Strong Heart Family studies
29 have accumulated detailed genetic, cardiovascular risk and event data in geographically
30 dispersed American Indian communities since 1988. Chi-square tests, logistic
31 regression and generalized linear mixed models (implemented in SOLAR) were used in
32 analysis. Genotypic data and 91 COVID-19 adjudicated deaths or hospitalizations from
33 2/1/20 through 3/1/23 were identified among 3,780 participants in two subsets. Among
34 21 candidate variants including genes in the interferon response pathway, *APOE*,
35 *TMPRSS2*, *TLR3*, the HLA complex and the ABO blood group, only rs1205, a 3'
36 untranslated region variant in the *CRP* gene, showed nominally significant association
37 in T-dominant model analyses (odds ratio 1.859, 95%CI 1.001-3.453, p=0.049) after
38 adjustment for age, sex, center, body mass index, and a history of cardiovascular
39 disease. Within the younger subset, association with the rs1205 T-Dom genotype was
40 stronger, both in the same adjusted logistic model and in the SOLAR analysis also
41 adjusting for other genetic relatedness.

42 Conclusion: A T-dominant genotype of rs1205 in the *CRP* gene is associated with
43 COVID-19 death or hospitalization, even after adjustment for relevant clinical factors
44 and potential participant relatedness. Additional study of other populations and genetic
45 variants of this gene are warranted.

46 Author summary:

47 Considerable inter-individual variability in COVID-19 clinical outcome has been noted
48 from the onset of this pandemic. Some risk factors, such as age, diabetes, obesity, prior
49 cardiovascular disease (CVD) are well established. The possibility of inherited, host
50 genetic risk factors has also been examined and a number validated in very large,
51 population based datasets. The present study leveraged on-going clinical surveillance
52 of CVD and available genetic information in an American Indian population to
53 investigate potential host genetic contributors to COVID-19 morbidity and mortality.
54 Although the number of cases ascertained was relatively small, a novel variant in the C-
55 reactive protein gene was found to be associated with COVID-19 outcome. This protein
56 constitutes a critical component of the innate immune system and *CRP* variants have
57 been reported associated with infectious disease and vaccination outcomes. Improved
58 understanding of the pathophysiology of this infection may allow more targeted therapy.

59

60 Introduction

61 The ongoing pandemic of Severe Acute Respiratory Syndrome Coronavirus 2
62 (SARS-CoV-2) infections has taken a devastating toll on American Indian, Alaskan
63 Native (AI/AN) populations, with hospitalization and death rate ratios for COVID-19
64 (compared with White ethnicity) reported as 2.5 fold; and 2.1 fold, respectively.[1]

65 A number of demographic and clinical characteristics, such as male gender,
66 socio-economic deprivation, diabetes, cardiovascular disease and obesity, have been
67 identified as risk factors for COVID-19 associated morbidity and mortality.[2] The
68 hypothesis that this increased burden of COVID-19 morbidity and mortality among
69 ethnic minorities is due to the increased prevalence of these co-morbidities in many
70 populations, has been considered; but the proportion of effect attributable to clinical co-
71 morbidities may, be small.[2] In spite of considerable interest in possible host genetic
72 factors that influence the severity of COVID-19 among all patients, there have been few
73 reliably replicating studies in specific racial/ethnic groups.[3]

74 The expectation that ethnic disparities in COVID-19 outcomes are primarily
75 driven by population specific behavioral or clinical burden of disease may have resulted
76 in less interest in investigating genetic susceptibility among AI/AN communities. This is
77 of importance for the possible identification of therapeutic targets and prevention
78 strategies assisting AI/AN as well as other populations.

79 The Strong Heart Study (SHS) and allied Strong Heart Family Study (SHFS),
80 comprise the largest, ongoing prospective epidemiological cohort study among
81 American Indians in 12 different Tribal communities located in Arizona, North and South
82 Dakotas, and in Oklahoma. These cohorts have the requisite genetic datasets and
83 ongoing surveillance of medical records for cardiovascular disease (CVD) and risk
84 factors, including COVID-19, to also assess genetic contributions to COVID-19 death
85 and morbidity risks. The present results derive from a case/cohort analysis of SNPs
86 reported associated with COVID mortality or morbidity in other populations.

87

88 Materials and Methods

89 *Participants and Cohorts*

90 The SHS/SHFS methodology and design has been described previously.[4,5]
91 Diabetes and hypertensive status were defined as previously;[6] and the dichotomous
92 CVD covariate indicates any history of myocardial infarction, coronary artery disease,
93 congestive heart failure, and atherosclerotic stroke and peripheral vasculature
94 disease.[7] While SHS/SHFS medical record surveillance of participants focuses on
95 ascertainment of outcome events and clinical risk factors related to CVD, in early 2020
96 ascertainment for COVID-19 was begun, with the recognition of a bidirectional
97 relationship between CVD and COVID-19 infection. The physicians conducting medical
98 records reviews identified COVID-19 as either a contributing or definitive factor in death

99 or hospitalization (disease severity). However, medical record review was triggered
100 primarily by CVD related events (for ascertainment criteria, see Lee et al [4]) and a
101 small number of other conditions, but not primarily by infections. Thus, the identified
102 COVID-19 endpoint of hospitalization is probably under-reported, given the surveillance
103 focus on CVD outcomes. However, all deaths associated with COVID-19 during the
104 study period are included in this report.

105 The study period was 2/1/2020 through 3/1/2023, when medical record reviews
106 ascertaining COVID-19 outcomes were available. We included all participants alive on
107 2/1/2020 and identified outcome events (either COVID-19 death or hospitalization)
108 during follow-up. There were no specific ICD-9/10 codes used to identify relevant
109 morbid COVID-19 cases, but physician review of the medical record determined
110 COVID-19 case status based on immune measures and clinical diagnoses, with morbid
111 events for review selected on the basis of standard SHS protocol to ascertain incident
112 CVD.[4]

113 *Rationale for genetic variants selected*

114 A PubMed literature search for candidate variants previously reported to be
115 associated with COVID-19 pathogenesis or clinical outcomes found previously
116 implicated genes in the interferon response pathway,[8,9] *APOE*,[10] *TMPRSS2*,[11]
117 *TLR3*,[9] *ACE*,[12] *FURIN*[13] and the human leukocyte antigen (HLA)[14,15] regions.
118 We also included ABO blood group polymorphisms which have been among the most
119 consistently identified host factors influencing the COVID-19 phenotype.[16-18] The top
120 17 "hits" of of the Covid-19 Host Genetics Initiative Browser, round 7 identified
121 additional variants of interest.[19]

122 Thus our literature search developed a total of 73 plausible candidate variants
123 and resulted in 21 polymorphisms that were available from the genetic resources of the
124 SHS, albeit within two subsets of participants that were genotyped in SHS/SHFS
125 substudies. The candidate SNPs and identifying, published references are found in
126 Table 1.

127

128 **Table 1. Variants genotyped, (associated genes) and allele frequency (AF), when**
129 **the genotype is available.**

| Variant (gene) | RISK ALLELE | SHFS AF of risk allele | SHS AF of risk allele | Literature citation |
|--|-------------|------------------------|-----------------------|---------------------|
| rs16944971 (<i>FURIN</i>) | C | 0.01 | NA* | [13,20] |
| rs7412 (<i>APOE</i>) | C | 0.02 | NA | [3,21] |
| rs1205 (<i>CRP</i>), alternate allele =T | C | 0.49 | 0.46 | [22-25] |
| rs3091244 (<i>CRP</i>) | G | 0.31 | NA | [26-28] |

| | | | | |
|--------------------------------|-----------|------|--------|---------------|
| rs3093068 (CRP) | G | NA | 0.01 | [29,30] |
| rs1800947 (CRP) | C | NA | 0.02 | [23,25,29,31] |
| rs201253322 (IRF7) | C | NA | <0.001 | [3,9] |
| rs116302758 (DPP4) | T | NA | 0.02 | [11] |
| rs56179129 (DPP4) | C | NA | 0.002 | [11] |
| rs12329760 (TMPRSS2) | C | NA | 0.20 | [11,13,32] |
| rs150892504 (ERAP2) | C | NA | 0.001 | [33,34] |
| rs1800795 (IL-6) | G | NA | 0.08 | [23,35,36] |
| rs1799752 (ACE) | Insertion | NA | 0.26 | [3,13] |
| rs8176719 (ABO, O allele) | Insertion | NA | 0.76 | [16,37] |
| rs8176746 (ABO, B allele) | G | 0.02 | 0.02 | [38,39] |
| rs10735079 (OAS3) | A | NA | 0.13 | [40,41] |
| rs1405655 (NR1H2) | G | NA | 0.45 | [42] |
| rs1886814 (FOXP4-AS1) | A | NA | 0.41 | [40] |
| rs2109069 (DPP9) | G | NA | 0.22 | [41,42] |
| rs9380142 (HLA-G) | A | NA | 0.32 | [40,41] |
| rs111837807 (CCHCR1) | A | NA | 0.20 | [42,43] |
| rs2071351 (HLA-DPA1) | A | NA | 0.08 | [40] |
| rs529565 (ABO, intron variant) | A | NA | 0.23 | [40] |
| rs10774671 (OAS1) | A | NA | 0.13 | [40] |
| rs61667602 (LINC02210-CRHR1) | A | NA | 0.06 | [42] |

130 * NA, Not available in this cohort

131

132 Although genotypes were derived from extensive microarray sources, no attempt
133 was made to screen or identify variants of interest on the basis of genome-wide analysis
134 of SHS microarray data.

135 The only exception to the qualification of previously identified association was the
136 choice of *CRP* polymorphisms as explained below. This was felt to be justified due to
137 the central role of *CRP* in the innate immune system and dramatic elevation during
138 COVID-19 and other viral infections.[43]

139 C-reactive protein (*CRP*) levels are typically thought of as a biomarker for
140 inflammatory response and not as a pathogenetic factor, although inherited *CRP*
141 variants have been associated with, and thus possibly a modifying factor in some
142 infections[22,23] response to vaccination,[26] as well as other conditions such as
143 cancer[44,45] and pre-eclampsia.[46,47] For this reason, we selected 4 available
144 variants related to *CRP* expression for our analysis.

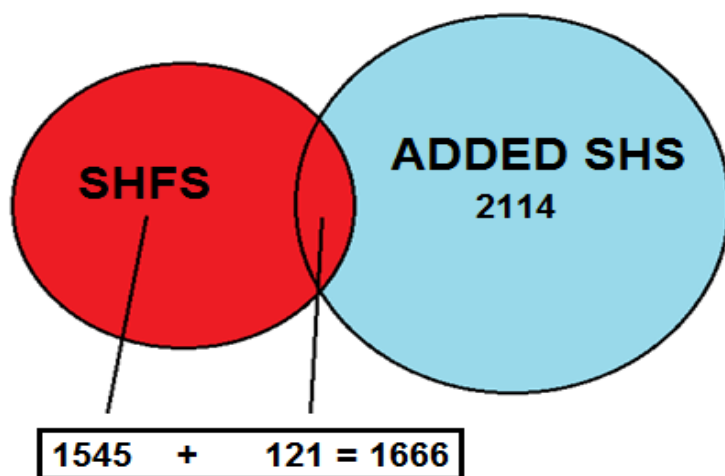
145 *Laboratory methods and considerations*

146 Clinical, anthropometric and laboratory measures are reported as obtained from
147 the most recent SHS exam prior to the study period. Laboratory methodology for
148 determination of serum creatinine,[48] HgbA1c[49] and high sensitivity C-Reactive
149 Protein (hsCRP)[50] have been reported previously

150 The SHFS [5] participants were genotyped with the Illumina Human Cardio-
151 Metabo BeadChip microarray (Illumina, San Diego, CA), incorporating approximately
152 200,000 single nucleotide polymorphisms (SNPs) in loci previously identified as
153 significantly associated with metabolic and CVD traits.[51] These genotypes were
154 generated exclusively from SHFS participants without diabetes mellitus during exams
155 between 1997-2003. Further details related to quality filters and data preparation have
156 been published.[50]

157 The SHS participants were genotyped using the Illumina Infinium Multi-Ethnic
158 Global-8 vs1. Quality control included filtering variants with call rates <95% and
159 duplicated QC samples. The *CRP* rs1205 variant was genotyped from both the SHFS
160 (N=1,666) and SHS (N=2,114) participants, providing a total of (N=3,780) rs1205
161 genotypes. This is presented graphically in Figure 1. Except for rs1205 and rs8176746,
162 genotypic data from other variants were available from either the SHFS or the SHS
163 participants (Table 1).

164 Figure 1. Venn diagram illustrating relationship between analysis cohorts.



165
166 Although two different genotyping platforms were used for the SHS and SHFS
167 participants, the genotyping for both was conducted in the same laboratory at Texas
168 Biomedical Research Institute in San Antonio, Texas. All 121 participants genotyped for
169 rs1205 and the ABO variant rs8176746 on both platforms and meeting the criteria for
170 inclusion in the study had concordant genotyping.

171
172 *Statistical analysis methods*

173 Descriptive statistics presented counts and percentages for discrete variables
174 and means with standard deviations for continuous variables. Between-group

175 comparison of continuous variables used T-tests and assumed independent samples.
176 Discrete variables utilized Pearson chi-square comparisons between groups.

177 Logistic regression models were developed for both univariate and multivariate
178 analysis of COVID-19 associated death or hospitalization during the observation period
179 (2/1/20 through 3/29/22). Only a single COVID-event was counted for each participant.
180 To select covariates included in the final multivariate model, BMI and any history of
181 CVD were chosen on the basis of apparent independent association ($p < 0.05$) when all
182 available covariates (sex, body mass index (BMI), hypertension, diabetes mellitus (DM),
183 impaired fasting glucose (IFG) or normal fasting glucose (NFG), history of CVD event,
184 (tobacco (current, ever, never), serum creatinine, and hemoglobin A1c (HgbA1c)) were
185 considered jointly. Age, sex and center were deemed important, standard adjustments
186 and therefore kept in models. The covariate related to diabetes and dysglycemia was
187 not independently associated with COVID-19 when included with BMI (as seen in the
188 first section of Table 5), so was not retained in the final logistic model. The serum
189 hsCRP was not included as a covariate, given we were testing *CRP* genetic variants
190 with known effects on serum levels. Significance was accepted at a p-value of 0.05,
191 since the genetic variants were chosen on the basis of a priori evidence of potential
192 association with COVID-19 as well as recognizing the hypothesis driven nature of the
193 presented analysis.

194 The SHFS cohort enrolled large families, and therefore we accounted for the
195 relatedness using random effects of pedigree relationships (established with
196 participants at recruitment) in generalized linear mixed models implemented in
197 SOLAR.[52,53] For discrete phenotypes such as affection status, SOLAR employs a
198 classic liability model.[54] Briefly: the unmeasured, presumed multifactorial, basis of
199 liability to disease is modeled as a standard normal distribution (mean=0, SD=1), with
200 affected individuals represented by an upper tail of the distribution that has the same
201 density as the incidence of affection status in the general population. For this study we
202 used the mean weekly incidence of COVID-19 in the US population during the study
203 period (217.72 cases/100,000 population, 0.218%)[55] Heritability of liability is
204 estimated from genetic correlations between relatives[56] and the effects of measured
205 factors (age, sex, etc., as well as measured genotypes) are estimated as deviations of
206 the liability threshold from the population incidence.

207 *Ethics Statement*

208 All participants in this study have provided written, informed consent to allow
209 genetic and other research related to CVD and its risk factors. Consent was obtained at
210 least once in all cases since Phase I of SHS in 1988, and in most instances has been
211 renewed during each the 6 subsequent phases until the present. Approval has been
212 obtained from the following institutional review boards (IRB) of record: Arizona Area
213 Indian Health Service IRB, MedStar Health Research Institute IRB, Great Plains Area
214 Indian Health Service IRB, Oglala Sioux Tribe Research Review Board, University of
215 Oklahoma Health Sciences IRB, and Oklahoma Area Indian Health Service IRB. In

216 addition to these formal, IRB approvals, all of the participants' Tribal governments have
 217 approved the conduct of the SHS/SHFS studies for these purposes. The close
 218 association between CVD (especially stroke) and COVID-19 infection, often in a
 219 bidirectional manner, justified IRB approval for this substudy.

220 Results

221 Table 1 shows the genotyped variants selected, their associated genes and allele
 222 frequencies in the SHFS or SHS cohorts. The characteristics of both the SHS and
 223 SHFS cohorts are summarized in Table 2.

224 **Table 2. Demographic and clinical measures stratified by enrolled cohort.**

| | SHFS | SHS | Total | P value |
|--|-------------|-------------|-------------|--------------------|
| Participants, N (%) | 1666 (44.1) | 2114 (55.9) | 3780 (100) | |
| Male Sex, N (%) | 660 (39.6) | 916 (43) | 1576 (42) | .021* |
| Mean Age (SD), years on 2/1/2020 | 52.0 (14.8) | 83.9 (8.4) | 69.8 (19.7) | <.001** |
| Mean Age (SD), years at last clinical exam | 41.7 (14.6) | 62.9 (8.7) | 53.6 (15.7) | <.001 |
| Diabetes status, N, (%) | | | | <.001 |
| Diabetes Mellitus (DM) | 214 (13) | 1001 (49) | 1215 (33) | |
| Impaired fasting glucose tolerance (IFG) | 347 (21) | 500 (24) | 847 (23) | |
| Normal fasting glucose tolerance (NFG) | 1080 (65) | 563 (27) | 1643 (44) | |
| Hypertension, N (%) | 470 (28) | 1110 (53) | 1580 (42) | <.001 |
| Cardiovascular disease*** N, (%) | 169 (10) | 870 (41) | 1039 (28) | <.001 |
| Mean BMI (SD), kg/m ² | 32.3 (7.7) | 30.3 (6.5) | 31.2 (7.1) | <.001 |
| Current smoker N (%) | 644 (39) | 618 (30) | 1310 (35) | 0.657 ^Δ |

| | | | | |
|--|------------------------------|------------------------------|-------------------------------|---------------------|
| Ever smoker, N (%) | 391 (24) | 749 (37) | 1140 (31) | <.001 ^{ΔΔ} |
| Never smoker, N (%) | 619 (37) | 618 (29) | 1237 (34) | |
| Mean serum creatinine (SD), mg/dL | 0.87 (0.36) | 1.14 (1.22) | 1 (0.90) | <.001 |
| Mean % HgB A1c (SD) | 6.36 (1.6) | 7.10 (2.22) | 6.90 (2.10) | <.001 |
| Mean hsCRP (SD) ^{****} | 4.12 (3.58) | 4.35 (3.31) | 4.23 (3.45) | .074 |
| COVID-19 any case, N (%) | 48 (2.9) | 43 (2.0) | 91 (2.4) | 0.09 |
| COVID-19 death, N (%) | 21 (1.3) | 26 (1.2) | 47 (1.2) | 0.93 |
| Participants with either fatal or non-fatal COVID | | | | |
| | SHFS | SHS | Total | P value |
| Male sex, N (%) | 18 (37) | 15 (35) | 33 (36) | 0.796 |
| Mean Age (SD), years on 2/1/2020 | 61.9 (14.7) | 78.1 (5.0) | 69.6 (13.8) | <0.001 |
| DM/IFG/NFG, N (%) | 19/5/23 (40.4,10.6,48.9) | 17/13/13 (39.5,30.2,30.2) | 36/18/36 (40.0, 20.0,40.0) | 0.043 |
| HTN, N (%) | 22 (45.8) | 21 (48.8) | 43 (47.3) | 0.774 |
| CVD, N (%) | 13 (27.1) | 25 (58.1) | 38 (41.8) | 0.003 |
| Mean BMI (SD) | 36.07 (9.10) | 31.04 (6.64) | 33.7 (8.38) | 0.001 |
| Smoking Current/ever/never (%) | 15/15/18 (31.3,31.3,37.5) | 14/11/18 (32.6,25.6,41.9) | 29/26/36 (31.9,28.6,39.6) | 0.828 |
| | | | | |
| Serum Creatinine (SD), mg/dL | 0.85 (0.20) | 1.12 (1.86) | 0.98 (1.27) | 0.362 |
| Mean HgbA1c (SD), % | 7.44 (2.14) | 6.65 (1.74) | 7.00 (1.95) | 0.191 |

| | | | | |
|---------------------------|-------------|-------------|-------------|-------|
| Mean hsCRP (SD), mg/L**** | 5.00 (3.78) | 4.48 (3.40) | 4.76 (3.60) | 0.715 |
|---------------------------|-------------|-------------|-------------|-------|

225 Values are mean (SD) or N (%). * Pearson chi square test for count data. **T-
 226 test for mean (SD) data *** Any history of coronary heart disease, stroke, congestive
 227 heart failure, or other cardiovascular disease ^Δ smoking "current" vs smoking "never" ^{ΔΔ}
 228 smoking "ever" vs smoking "never" **** hsCRP values >15mg/L excluded

229

230 Unadjusted results limited to the primary variant of interest from the SHFS, the
 231 SHS cohort or combined SHFS/SHS cohorts, using additive and dominant models
 232 based on risk or non-risk alleles are shown in Table 3.

233

234 **Table 3. Primary results of interest, rs1205: univariant associations with either**
 235 **fatal or non-fatal COVID-19.**

| Cohort | SNP | Risk allele | Model | Chi-square p value* | OR** | 95% CI | P value |
|-----------------|---------------|-------------|--------------|---------------------|--------------|----------------------|--------------|
| SHS | rs1205 | C | Add | 0.564 | 0.922 | 0.603 - 1.411 | 0.709 |
| SHS | rs1205 | C | C-DOM | 0.815 | 1.084 | 0.553 - 2.124 | 0.815 |
| SHS | rs1205 | T | T-DOM | 0.368 | 1.452 | 0.642 - 3.285 | 0.370 |
| SHFS | rs1205 | C | Add | 0.039 | 0.599 | 0.396 - 0.906 | 0.015 |
| SHFS | rs1205 | C | C-DOM | 0.099 | 0.609 | 0.336 - 1.104 | 0.102 |
| SHFS | rs1205 | T | T-DOM | 0.016 | 2.976 | 1.171 - 7.565 | 0.022 |
| SHS/SHFS | rs1205 | C | Add | 0.066 | 0.743 | 0.553 - 0.998 | 0.049 |
| SHS/SHFS | rs1205 | C | C-DOM | 0.332 | 0.804 | 0.516 - 1.251 | 0.333 |
| SHS/SHFS | rs1205 | T | T-DOM | 0.020 | 2.038 | 1.105 - 3.758 | 0.023 |

236

237 * Pearson chi-square, asymptotic significance, two-sided

238 ** OR: univariate logistic regression odds ratio

239

240 Results for the remaining variants are found in Supplemental Table S1. For
 241 variants with low frequency, logistic models were often unstable and not reported.
 242 Analyses in the SHFS cohort showed nominally significant associations for the rs1205 T
 243 allele dominant (T-Dom) genotype in the combined SHFS/SHS cohort and both additive
 244 and T-Dom models in the SHFS cohort).

245 Table 4 shows the distribution of COVID-19 cases among recruitment centers,
 246 highlighting the relatively larger number of cases identified in the Dakota center
 247 compared to other centers, and provides the allelic prevalence of rs1205 by center.

248 Among the combined SHFS/SHS cohorts, the prevalence of the rs1205 T allele was
 249 0.525 (95% CI 0.51 - 0.54).

250

251 **Table 4. Combined SHFS/SHS participant and SNP results from adjusted analysis**
 252 **by recruitment center.**

| Center* | AZ | DK | OK | p value | analysis |
|---|--------------------|--------------------|--------------------|---------|----------|
| N | 488 | 1705 | 1587 | | |
| Age mean (SD) | 70.65 (19.48) | 68.39 (20.29) | 71.13 (19.05) | <0.001 | 1 |
| Male sex N (%) | 184 (38) | 734 (43) | 658 (41) | 0.104 | 2 |
| BMI mean (SD) | 34.78 (9.88) | 30.25 (6.36) | 31.15 (6.55) | <0.001 | 1 |
| ANY_CVD_YN N(%) | 119 (24) | 547 (32) | 373 (23) | <0.001 | 2 |
| SHFS cohort N(%) | 165 (34) | 698 (44) | 803 (47) | <0.001 | 2 |
| Fatal/non-fatal COVID-19 N (% of total) | 11 (12.1) | 62 (68.1) | 18 (19.8) | <0.001 | 2 |
| Risk of outcome % (95% CI) | 2.25% (0.94-3.57) | 3.64% (2.75-4.52) | 1.13% (0.61-1.66) | | 3 |
| rs1205,C allele count (95% CI) | 0.48 (0.45 - 0.51) | 0.58 (0.56 - 0.59) | 0.48 (0.47 - 0.50) | <0.001 | 4 |
| rs1205 T-Dom (95% CI) | 0.72 (0.68 - 0.76) | 0.81 (0.79 - 0.83) | 0.73 (0.71 - 0.75) | <0.001 | 2 |
| rs1205 T-Dom p value | 0.254 | 0.333 | 0.147 | | 5 |

253 * AZ = Arizona, DK = Dakotas, OK = Oklahoma centers

254 1= Independent samples, T-test

255 2= between center Pearson chi-square p value, 3X2 table

256 3= within center binomial confidence interval on proportion

257 4= between center Pearson chi-square p value, 3X3 table

258 5= within center COVID fatal/non-fatal outcome via logistic regression adjusted for age, sex, BMI,
 259 ANY_CVD

260

261

262

263 Multivariate logistic regression model findings for the primary SNPs of interest
 264 are reported in Table 5. Results of the rs1205 T-Dom model is attenuated after
 265 adjustment for covariates, but retains nominally significant association (OR 1.859, 95%
 266 CI 1.001-3.453, p=0.049) with COVID-19 in in the combined SHS/SHFS cohorts. Within

267 the younger SHFS cohort, the same model showed an odds ratio of 2.857, 95%CI
 268 1.108-7.362, p=0.030.
 269

270 **Table 5. Results of multivariate adjusted models evaluating the composite**
 271 **outcome of either fatal or non-fatal COVID-19 within the combined SHFS and SHS**
 272 **cohorts.**

| SHFS and SHS COMBINED | | | | | | |
|-------------------------------|--------------|--------------------|--------------|--------------|--------------|--------------|
| | | | | 95% CI | | |
| Multivariate model including: | N | OR | Lower | Upper | P value | |
| Age | 3780 | 0.999 | 0.987 | 1.012 | 0.923 | |
| Sex | 3780 | 0.817 | 0.527 | 1.267 | 0.367 | |
| AZ center | 3780 | Indicator variable | | | | |
| DK center | 3780 | 2.132 | 1.064 | 4.272 | 0.033 | |
| OK center | 3780 | 0.649 | 0.296 | 1.423 | 0.280 | |
| BMI | 3780 | 1.050 | 1.023 | 1.079 | <0.001 | |
| ANY CVD (Y/N) | 3780 | 1.871 | 1.153 | 3.036 | 0.011 | |
| SHFS COHORT | | | | | | |
| | | | | 95% CI | | |
| SNP* | model | N | OR | Lower | Upper | P value |
| rs1205 | C-Add | 1647 | 0.587 | 0.381 | 0.905 | 0.016 |
| rs1205 | T-DOM | 1647 | 2.857 | 1.108 | 7.362 | 0.030 |
| SHS COHORT | | | | | | |
| rs1205 | C-Add | 2042 | 1.057 | 0.683 | 1.634 | 0.804 |
| rs1205 | T-DOM | 2042 | 1.243 | 0.541 | 2.860 | 0.608 |
| SHFS and SHS COMBINED | | | | | | |
| rs1205 | C-Add | 3689 | 0.799 | 0.592 | 1.078 | 0.142 |
| rs1205 | T-DOM | 3689 | 1.859 | 1.001 | 3.453 | 0.049 |

273 * Genetic variant included in model adjusted for age, sex, center, BMI, and ANY CVD.

274
 275 In the SHFS analysis, using a generalized linear threshold model in SOLAR to
 276 adjust for the random effect of family relatedness, and adjusting for fixed effects of sex,
 277 age, BMI, CVD, and center, either fatal or non-fatal COVID-19 was significantly
 278 associated with the fixed effect of the rs1205 T allele-dominant genotype (p=0.0003).
 279 The regression beta for this association was negative (= -0.50 standard deviations),
 280 indicating a tendency to increase the affected upper tail of the liability distribution (thus
 281 confirming rs1205-T Dom genotype as a risk allele).

282
 283 While those with rs1205 T-Dom genotypes trended toward lower mean hsCRP
 284 levels at baseline than others (p=0.087 excluding levels over 15 mg/L to avoid spurious
 285 elevations due to intercurrent infections), mean hsCRP levels were not significantly
 286 different between COVID-19 cases or controls. (Supplemental Table S2).

287

288 Discussion

289 We present previously unreported evidence of association between a common
290 human variant (rs1205) and the clinical impact of the SARS-CoV-2 virus causing
291 COVID-19. While our study has a modest number of cases, these results derive from a
292 large cohort of American Indian individuals who are being followed longitudinally using
293 standardized, physician review of medical records[4] The rs1205 variant has been
294 recognized as functionally affecting serum levels of CRP,[25,28,29] in several
295 populations and, in the present SHFS population as well.[50] CRP is an important
296 component of the innate immune system, thus supporting a role of the gene and this
297 variant's potential effects on the pathophysiology of COVID-19. A genome-wide
298 association meta-analysis has reported an intron variant (rs67579710) associated with
299 COVID-19 hospitalization among 24,741 cases and 2,835,201 controls.[42] This variant
300 is 4.5Mb from rs1205 [57] and within the thrombospondin3 gene, thus it may play a
301 role in the thrombosis associated with COVID-19, rather than inflammatory pathways.

302 There are ample theoretical reasons that heritable genetic variation could affect
303 SARS-CoV-2 infection, beginning with the viral requirement for surface receptors to gain
304 entry to the cytoplasm[58] and running through multiple immune response pathways
305 generating the "cytokine storm" which has been prominently noted in the pathogenesis
306 of COVID-19.[59] A number of genetic variants have been associated with morbidity
307 and mortality, including variants in the angiotensin converting enzyme 1 (*ACE1*)
308 gene,[60] the *TMPRSS2* gene,[61] the *IL-6* gene[62] and others extensively reviewed by
309 Ishak et al.[63] The COVID-19 Host Genetics Initiative, a massive ascertainment of
310 nearly 50,000 cases from 19 countries has identified 13 genome-wide loci related to
311 either initial infection, or morbidity.[64]

312 The rs1205 variant is a C/T single nucleotide polymorphism in the 3' untranslated
313 region of the C-reactive protein gene (*CRP*), with a prevalence of the C allele ranging
314 from ~ 0.67 in European populations, to ~ 0.40 in Asian populations,[65] and 0.51 to
315 0.54 in the present study. The 4 *CRP* variants were chosen for their role in the innate
316 immune system and availability in the SHS dataset, but no literature was found that
317 examined *CRP* variants related to COVID-19 severity, perhaps because CRP has
318 traditionally been viewed as a biomarker and not as a potential causal factor. There are,
319 however, reports of *CRP* variants associated with the clinical outcomes of other
320 infections[22,23] and the response to vaccination.[26] Additionally, reduced expression
321 of serum CRP is consistently reported associated with the rs1205 T allele, with the beta
322 coefficient for each additional T allele ranging between -0.17 and -0.27;[25,28,29] and
323 estimated at -0.23 in the SHS Dakota center.[50] As noted in Table S2 the mean serum
324 CRP level is lower in the present study, among those with the rs1205 T-Dom genotype
325 compared with those homozygous for the C allele, although the difference is not
326 significant.

327 Although the prevalent view of COVID-19 pathophysiology implicates a
328 hyperactive immune response (often characterized by increased CRP serum levels as
329 well as coordinated production of other cytokines), it also seems plausible that a
330 relative, baseline insufficiency of an innate immune factor, particularly in the early phase
331 of exposure, could also result in increased morbidity.

332 *CRP* genotypes correlating with increased baseline CRP levels have been
333 positively associated with risk of pre-eclampsia[46,47,66] and conversely, as in the
334 current study, *CRP* genotypes (including rs1205 T-Dom) correlated with lower CRP
335 levels have been associated with increased infectious burden.[22-24] These divergent
336 effects invite speculation that *CRP* polymorphisms may have experienced balanced
337 selection during evolution.

338 Our multivariate logistic regression results are in accord with prior studies that
339 showed the contribution of commonly reported clinical factors, such as age,[67] BMI[64]
340 and pre-existing CVD,[68,69] which negatively impact COVID-19 outcomes. Although
341 the mean BMI of the SHFS is somewhat greater than the SHS cohort, it should be noted
342 that the SHFS cohort has a low prevalence of DM (13%) and is younger than the SHFS,
343 due to selection for that substudy of those without diabetes at baseline. Perhaps this
344 allows a clearer genetic (vs an acquired clinical or environmental) association to
345 become apparent.

346 The small number of cases when stratified by Center suggests cautious
347 interpretation, but is presented since the number of cases from the Dakota Center is
348 considerably higher than from the comparable Oklahoma center. Although the
349 possibility of population stratification confounding our results is present, we believe the
350 fact that all 3 centers show an effect of the T-Dom genotype in the same direction,
351 including the Arizona Center, which is likely to be most distant in their genetic
352 background, suggests genetic background effects are not likely. The statistical
353 significance of the center covariate could be due to various differences in regional,
354 environmental factors, such as household density, weather, and accessibility of medical
355 care.

356 The lack of significant results in the other SNPs is not surprising, given the
357 number of these variants with low allele frequencies and the small number of COVID-19
358 cases. Further surveillance of the SHS cohort will likely identify new cases and improve
359 the power for other variants (eg rs111837807 and rs2109069).

360 The limitations of this analysis are chiefly related to the relatively small number of
361 cases obtained to this point. The morbid case ascertainment is subject to undercounting
362 since the choice of medical records for review is based primarily on CVD outcomes.
363 This does not affect ascertainment of mortal cases, since all deaths are reviewed and
364 recorded as soon as possible. The possibility of population stratification is noted,
365 although seems unlikely for reasons noted above. P values have not been adjusted for
366 multiple testing.

367 The study's strengths include a robust physician assignment of outcomes using
368 medical records and comprehensive analysis of genetic variabilities as well as
369 adjustment for relatedness between individuals.

370 It is clear that race/ethnicity is a valid and critical risk marker for other underlying
371 conditions affecting the complexity of COVID-19 disease, such as structural racism,
372 discrimination and socioeconomic status,[70] lack of health care access,[71] and
373 exposure to infectious agents related to high risk and service industry occupations.[72]
374 Although one can never be confident that all confounding factors have been adjusted
375 properly, Williamson et al[2] showed that ethnicity among a very large cohort in England
376 was associated with COVID-19 outcomes, even after adjusting for socioeconomic
377 factors. Thus, even though we are committed to improving our understanding of the
378 complex social and behavioral factors influencing COVID-19 disease landscape among
379 Tribal communities we also believe that investigating the potential influence of common
380 genetic variants among the high risk American Indian/Alaska Native populations may
381 provide opportunities for therapeutic or preventive options. Although severely impacted
382 by COVID-19,[73,74] Tribal communities are, at times, excluded from scientific studies
383 due to their smaller proportion in the U.S. population and various socio-political factors.

384 In conclusion, a statistically significant association was found between the rs1205
385 T-Dom genotype and risk of COVID-19 death or hospitalization among American Indian
386 participants, employing chi-square tests, logistic regression models adjusting for age,
387 sex, center, BMI and prior history of CVD, and SOLAR software analysis also adjusting
388 for relatedness within the SHFS cohort. The direction of effect suggests a lower level of
389 CRP during early phases of infection may increase the risk of subsequent
390 complications, a novel finding we will continue to investigate further as the pandemic
391 continues and our samples size tragically continues to increase.

392 Supporting Information:

393 Supplemental Table 1. Additional univariant associations with either fatal or non-
394 fatal COVID-19.

395 Supplemental Table 2. Baseline hsCRP measures* in relation to rs1205 T-Dom
396 genotype and COVID-19 case/control status.

397

398 Acknowledgements:

399 We thank the study participants, Indian Health Service facilities, and participating Tribal
400 communities for their extraordinary cooperation and involvement, which has been
401 critical to the success of the Strong Heart Study since 1988.

402 References:

- 403 1. Risk for COVID-19 Infection, Hospitalization, and Death By Race/Ethnicity,
404 [https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-
406 discovery/hospitalization-death-by-race-ethnicity.html](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-
405 discovery/hospitalization-death-by-race-ethnicity.html) (accessed 1/7/2022)
- 407 2. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-
408 19-related death using OpenSAFELY. *Nature*. 2020 Aug;584(7821):430-436. doi:
409 10.1038/s41586-020-2521-4. Epub 2020 Jul 8. PMID: 32640463; PMCID:
410 PMC7611074.
- 411 3. Ferreira de Araújo JL, Menezes D, Saraiva-Duarte JM, et al. Systematic review
412 of host genetic association with Covid-19 prognosis and susceptibility: What have
413 we learned in 2020? *Rev Med Virol*. 2021 Aug 2:e2283. doi: 10.1002/rmv.2283.
414 Epub ahead of print. PMID: 34338380.
- 415 4. Lee ET, Welty TK, Fabsitz R, Cowan LD, Le NA, Oopik AJ, et al. The Strong
416 Heart Study. A study of cardiovascular disease in American Indians: design and
417 methods. *Am J Epidemiol*. 1990 Dec;132(6):1141-55. PMID: 2260546.
- 418 5. North KE, Howard BV, Welty TK, Best LG, Lee ET, Yeh JL, et al. Genetic and
419 environmental contributions to cardiovascular disease risk in American Indians:
420 the strong heart family study. *Am J Epidemiol*. 2003 Feb 15;157(4):303-14.
421 PMID: 12578801.
- 422 6. Xu J, Lee ET, Peterson LE, Devereux RB, Rhoades ER, Umans JG, et al.
423 Differences in risk factors for coronary heart disease among diabetic and
424 nondiabetic individuals from a population with high rates of diabetes: the Strong
425 Heart Study. *J Clin Endocrinol Metab*. 2012 Oct;97(10):3766-74. doi:
426 10.1210/jc.2012-2110. Epub 2012 Jul 16. PMID: 22802089; PMCID:
427 PMC3674295.
- 428 7. Lu W, Resnick HE, Jablonski KA, Jones KL, Jain AK, Howard WJ, et al. Non-
429 HDL cholesterol as a predictor of cardiovascular disease in type 2 diabetes: the
430 strong heart study. *Diabetes Care*. 2003 Jan;26(1):16-23. doi:
431 10.2337/diacare.26.1.16. PMID: 12502653.
- 432 8. Gómez J, Albaiceta GM, Cuesta-Llavona E, García-Clemente M, López-Larrea
433 C, Amado-Rodríguez L, et al. The Interferon-induced transmembrane protein 3
434 gene (IFITM3) rs12252 C variant is associated with COVID-19. *Cytokine*. 2021
435 Jan;137:155354. doi: 10.1016/j.cyto.2020.155354. Epub 2020 Oct 23. PMID:
436 33113474.
- 437 9. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, et al. Inborn errors of
438 type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020
439 Oct 23;370(6515):eabd4570. doi: 10.1126/science.abd4570. Epub 2020 Sep 24.
440 PMID: 32972995; PMCID: PMC7857407.
- 441 10. Kuo CL, Pilling LC, Atkins JL, Masoli JAH, Delgado J, Kuchel GA, et al. APOE e4
442 Genotype Predicts Severe COVID-19 in the UK Biobank Community Cohort. *J
443 Gerontol A Biol Sci Med Sci*. 2020 Oct 15;75(11):2231-2232. doi:
444 10.1093/gerona/glaa131. PMID: 32451547; PMCID: PMC7314139.
- 445 11. Latini A, Agolini E, Novelli A, Borgiani P, Giannini R, Gravina P, et al. COVID-19
446 and Genetic Variants of Protein Involved in the SARS-CoV-2 Entry into the Host
447 Cells. *Genes (Basel)*. 2020 Aug 27;11(9):1010. doi: 10.3390/genes11091010.
PMID: 32867305; PMCID: PMC7565048.

- 448 12. Ghafouri-Fard S, Noroozi R, Vafae R, Branicki W, Pośpiech E, Pyrc K, et al.
449 Effects of host genetic variations on response to, susceptibility and severity of
450 respiratory infections. *Biomed Pharmacother*. 2020 Aug;128:110296. doi:
451 10.1016/j.biopha.2020.110296. Epub 2020 May 29. PMID: 32480226; PMCID:
452 PMC7258806.
- 453 13. Jafarpour R, Pashangzadeh S, Dowran R. Host factors: Implications in
454 immunopathogenesis of COVID-19. *Pathol Res Pract*. 2021 Dec;228:153647.
455 doi: 10.1016/j.prp.2021.153647. Epub 2021 Oct 12. PMID: 34749207; PMCID:
456 PMC8505027.
- 457 14. Saulle I, Vicentini C, Clerici M, Biasin M. Antigen presentation in SARS-CoV-2
458 infection: the role of class I HLA and ERAP polymorphisms. *Hum Immunol*. 2021
459 Aug;82(8):551-560. doi: 10.1016/j.humimm.2021.05.003. Epub 2021 May 10.
460 PMID: 34116863; PMCID: PMC8108382.
- 461 15. Littera R, Campagna M, Deidda S, Angioni G, Cipri S, Melis M, et al. Human
462 Leukocyte Antigen Complex and Other Immunogenetic and Clinical Factors
463 Influence Susceptibility or Protection to SARS-CoV-2 Infection and Severity of
464 the Disease Course. The Sardinian Experience. *Front Immunol*. 2020 Dec
465 4;11:605688. doi: 10.3389/fimmu.2020.605688. PMID: 33343579; PMCID:
466 PMC7746644.
- 467 16. Gómez J, Albaiceta GM, García-Clemente M, García-Gala JM, Coto E. DNA
468 genotyping of the ABO gene showed a significant association of the A-group
469 (A1/A2 variants) with severe COVID-19. *Eur J Intern Med*. 2021 Jun;88:129-132.
470 doi: 10.1016/j.ejim.2021.02.016. Epub 2021 Feb 25. PMID: 33750629; PMCID:
471 PMC7906510.
- 472 17. Pereira E, Felipe S, de Freitas R, Araújo V, Soares P, Ribeiro J, et al. ABO blood
473 group and link to COVID-19: A comprehensive review of the reported
474 associations and their possible underlying mechanisms. *Microb Pathog*. 2022
475 Jun 25;169:105658. doi: 10.1016/j.micpath.2022.105658. Epub ahead of print.
476 PMID: 35764188; PMCID: PMC9233352.
- 477 18. Tamayo-Velasco Á, Peñarrubia-Ponce MJ, Álvarez FJ, de la Fuente I, Pérez-
478 González S, Andaluz-Ojeda D. ABO Blood System and COVID-19 Susceptibility:
479 Anti-A and Anti-B Antibodies Are the Key Points. *Front Med (Lausanne)*. 2022
480 Apr 25;9:882477. doi: 10.3389/fmed.2022.882477. PMID: 35547235; PMCID:
481 PMC9081929.
- 482 19. <https://www.covid19hg.org>, accessed 3/1/23.
- 483 20. de Araújo JLF, Menezes D, de Aguiar RS, de Souza RP. *IFITM3*, *FURIN*, *ACE1*,
484 and *TNF-α* Genetic Association With COVID-19 Outcomes: Systematic Review
485 and Meta-Analysis. *Front Genet*. 2022 Apr 1;13:775246. doi:
486 10.3389/fgene.2022.775246. PMID: 35432458; PMCID: PMC9010674.
- 487 21. Gkouskou K, Vasilogiannakopoulou T, Andreacos E, Davanos N, Gazouli M,
488 Sanoudou D, et al. COVID-19 enters the expanding network of apolipoprotein E4-
489 related pathologies. *Redox Biol*. 2021 May;41:101938. doi:
490 10.1016/j.redox.2021.101938. Epub 2021 Mar 10. PMID: 33730676; PMCID:
491 PMC7943392.
- 492 22. Liu C, Jin P, Luo Y, Xu J, Kong C, Chen J, et al. Association of Single-Nucleotide
493 Polymorphisms of C-Reactive Protein Gene with Susceptibility to Infantile Sepsis

- 494 in Southern China. *Med Sci Monit.* 2018 Jan 30;24:590-595. doi:
495 10.12659/msm.908602. PMID: 29379005; PMCID: PMC5800323.
- 496 23. Chou SC, Ko HW, Lin YC. CRP/IL-6/IL-10 Single-Nucleotide Polymorphisms
497 Correlate with the Susceptibility and Severity of Community-Acquired
498 Pneumonia. *Genet Test Mol Biomarkers.* 2016 Dec;20(12):732-740. doi:
499 10.1089/gtmb.2016.0156. Epub 2016 Oct 5. Erratum in: *Genet Test Mol*
500 *Biomarkers.* 2019 Jun;23(6):433. PMID: 27705004.
- 501 24. Ponasenko AV, Kutikhin AG, Khutornaya MV, Rutkovskaya NV, Kondyukova NV,
502 Odarenko YN, et al. Inherited Variation in Cytokine, Acute Phase Response, and
503 Calcium Metabolism Genes Affects Susceptibility to Infective Endocarditis.
504 *Mediators Inflamm.* 2017;2017:7962546. doi: 10.1155/2017/7962546. Epub 2017
505 Jun 4. PMID: 28659664; PMCID: PMC5474236.
- 506 25. Kocarnik JM, Pendergrass SA, Carty CL, Pankow JS, Schumacher FR, Cheng I,
507 et al. Multiancestral analysis of inflammation-related genetic variants and C-
508 reactive protein in the population architecture using genomics and epidemiology
509 study. *Circ Cardiovasc Genet.* 2014 Apr; 7(2):178–88. <https://doi.org/10.1161/CIRCGENETICS.113.000173> Epub 2014 Mar 12. PMID: 24622110;
510 PubMed Central PMCID: PMC4104750.
- 511 26. Ovsyannikova IG, White SJ, Larrabee BR, Grill DE, Jacobson RM, Poland GA.
512 Leptin and leptin-related gene polymorphisms, obesity, and influenza A/H1N1
513 vaccine-induced immune responses in older individuals. *Vaccine.* 2014 Feb
514 7;32(7):881-7. doi: 10.1016/j.vaccine.2013.12.009. Epub 2013 Dec 18. PMID:
515 24360890; PMCID: PMC3922536.
- 516 27. Mõlkänen T, Rostila A, Ruotsalainen E, Alanne M, Perola M, Järvinen A. Genetic
517 polymorphism of the C-reactive protein (CRP) gene and a deep infection focus
518 determine maximal serum CRP level in *Staphylococcus aureus* bacteremia. *Eur*
519 *J Clin Microbiol Infect Dis.* 2010 Sep;29(9):1131-7. doi: 10.1007/s10096-010-
520 0978-z. Epub 2010 Jun 16. PMID: 20552244.
- 521 28. Ridker PM, Pare G, Parker A, Zee RY, Danik JS, Buring JE, et al. Loci related to
522 metabolic-syndrome pathways including LEPR, HNF1A, IL6R, and GCKR
523 associate with plasma C-reactive protein: the Women's Genome Health Study.
524 *Am J Hum Genet.* 2008 May; 82(5):1185–92. [https://doi.org/10.](https://doi.org/10.1016/j.ajhg.2008.03.015)
525 [1016/j.ajhg.2008.03.015](https://doi.org/10.1016/j.ajhg.2008.03.015) Epub 2008 Apr 24. PMID: 18439548; PubMed Central
526 PMCID: PMC2427311.
- 527 29. Wang S, Xu H, Zhou N, Zhao W, Wu D, Shen B. Combined Effects of Single
528 Nucleotide Polymorphisms (SNPs) within C-reactive Protein (CRP) and
529 Environmental Parameters on Risk and Prognosis for Diabetic Foot
530 Osteomyelitis Patients. *Exp Clin Endocrinol Diabetes.* 2020 Aug;128(8):528-539.
531 doi: 10.1055/a-0754-5622. Epub 2020 Jun 15. PMID: 32542640.
- 532 30. Myburgh PH, Nienaber-Rousseau C, Kruger IM, Towers GW. Education,
533 Smoking and *CRP* Genetics in Relation to C-Reactive Protein Concentrations in
534 Black South Africans. *Int J Environ Res Public Health.* 2020 Sep 11;17(18):6646.
535 doi: 10.3390/ijerph17186646. PMID: 32933066; PMCID: PMC7558133.
- 536 31. Schick UM, Auer PL, Bis JC, Lin H, Wei P, Pankratz N, et al.; Cohorts for Heart
537 and Aging Research in Genomic Epidemiology; National Heart, Lung, and Blood
538 Institute GO Exome Sequencing Project. Association of exome sequences with
539

- 540 plasma C-reactive protein levels in >9000 participants. *Hum Mol Genet.* 2015 Jan
541 15; 24(2):559–71. <https://doi.org/10.1093/hmg/ddu450> Epub 2014 Sep 3. PMID:
542 25187575; PubMed Central PMCID:PMC4334838.
- 543 32. Mohammadpour S, Torshizi Esfahani A, Halaji M, Lak M, Ranjbar R. An updated
544 review of the association of host genetic factors with susceptibility and resistance
545 to COVID-19. *J Cell Physiol.* 2021 Jan;236(1):49-54. doi: 10.1002/jcp.29868.
546 Epub 2020 Jun 15. PMID: 32542735; PMCID: PMC7323230.
- 547 33. Saulle I, Vanetti C, Goglia S, Vicentini C, Tombetti E, Garziano M, et al. New
548 ERAP2/Iso3 Isoform Expression Is Triggered by Different Microbial Stimuli in
549 Human Cells. Could It Play a Role in the Modulation of SARS-CoV-2 Infection?
550 *Cells.* 2020 Aug 24;9(9):1951. doi: 10.3390/cells9091951. PMID: 32847031;
551 PMCID: PMC7563522.
- 552 34. D'Amico S, Tempora P, Lucarini V, Melaiu O, Gaspari S, Algeri M, et al. ERAP1
553 and ERAP2 Enzymes: A Protective Shield for RAS against COVID-19? *Int J Mol*
554 *Sci.* 2021 Feb 8;22(4):1705. doi: 10.3390/ijms22041705. PMID: 33567739;
555 PMCID: PMC7914632.
- 556 35. Karcioğlu Batur L, Hekim N. Correlation between interleukin gene polymorphisms
557 and current prevalence and mortality rates due to novel coronavirus disease
558 2019 (COVID-2019) in 23 countries. *J Med Virol.* 2021 Oct;93(10):5853-5863.
559 doi: 10.1002/jmv.27127. Epub 2021 Jun 10. PMID: 34081354; PMCID:
560 PMC8242628.
- 561 36. Smieszek SP, Przychodzen BP, Polymeropoulos VM, Polymeropoulos CM,
562 Polymeropoulos MH. Assessing the potential correlation of polymorphisms in the
563 IL6R with relative IL6 elevation in severely ill COVID-19 patients'. *Cytokine.* 2021
564 Dec;148:155662. doi: 10.1016/j.cyto.2021.155662. Epub 2021 Jul 29. Erratum in:
565 *Cytokine.* 2022 Jan;149:155752. PMID: 34353696; PMCID: PMC8318728.
- 566 37. Severe Covid-19 GWAS Group; Ellinghaus D, Degenhardt F, Bujanda L, Buti M,
567 Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19
568 with Respiratory Failure. *N Engl J Med.* 2020 Oct 15;383(16):1522-1534. doi:
569 10.1056/NEJMoa2020283. Epub 2020 Jun 17. PMID: 32558485; PMCID:
570 PMC7315890.
- 571 38. Kolin DA, Kulm S, Christos PJ, Elemento O. Clinical, regional, and genetic
572 characteristics of Covid-19 patients from UK Biobank. *PLoS One.* 2020 Nov
573 17;15(11):e0241264. doi: 10.1371/journal.pone.0241264. PMID: 33201886;
574 PMCID: PMC7671499.
- 575 39. Lehrer S, Rheinstein PH. ABO blood groups, COVID-19 infection and mortality.
576 *Blood Cells Mol Dis.* 2021 Jul;89:102571. doi: 10.1016/j.bcmd.2021.102571.
577 Epub 2021 Apr 21. PMID: 33894687; PMCID: PMC8059281.
- 578 40. Velavan TP, Pallerla SR, Rüter J, Augustin Y, Kremsner PG, Krishna S, et al.
579 Host genetic factors determining COVID-19 susceptibility and severity.
580 *EBioMedicine.* 2021 Oct;72:103629. doi: 10.1016/j.ebiom.2021.103629. Epub
581 2021 Oct 13. PMID: 34655949; PMCID: PMC8512556.
- 582 41. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al.
583 Genetic mechanisms of critical illness in COVID-19. *Nature.* 2021
584 Mar;591(7848):92-98. doi: 10.1038/s41586-020-03065-y. Epub 2020 Dec 11.
585 PMID: 33307546.

- 586 42. <https://app.covid19hg.org/variants>, accessed 12/15/22.
- 587 43. Szyda J, Dobosz P, Stojak J, Sypniewski M, Suchocki T, Kotlarz K, et al. Beyond
588 GWAS-Could Genetic Differentiation within the Allograft Rejection Pathway
589 Shape Natural Immunity to COVID-19? *Int J Mol Sci.* 2022 Jun 3;23(11):6272.
590 doi: 10.3390/ijms23116272. PMID: 35682950; PMCID: PMC9181155.
- 591 44. Nimptsch K, Aleksandrova K, Fedirko V, Jenab M, Gunter MJ, Siersema PD, et
592 al. Pre-diagnostic C-reactive protein concentrations, CRP genetic variation and
593 mortality among individuals with colorectal cancer in Western European
594 populations. *BMC Cancer.* 2022 Jun 24;22(1):695. doi: 10.1186/s12885-022-
595 09778-9. PMID: 35739525; PMCID: PMC9229883.
- 596 45. Pan HY, Mi YY, Xu K, Zhang Z, Wu H, Zhang W, et al. Association of C-reactive
597 protein (CRP) rs1205 and rs2808630 variants and risk of cancer. *J Cell Physiol.*
598 2020 Nov;235(11):8571-8584. doi: 10.1002/jcp.29701. Epub 2020 Apr 24. PMID:
599 32329054.
- 600 46. Best LG, Saxena R, Anderson CM, Barnes MR, Hakonarson H, Falcon G, et al.
601 Two variants of the C-reactive protein gene are associated with risk of pre-
602 eclampsia in an American Indian population. *PLoS One.* 2013 Aug
603 5;8(8):e71231. doi: 10.1371/journal.pone.0071231. Erratum in: *PLoS One.* 2013
604 Sep 23;8(9):null. PMID: 23940726; PMCID: PMC3733916.
- 605 47. Best LG, Nadeau M, Davis K, Lamb F, Bercier S, Anderson CM. Genetic
606 variants, immune function, and risk of pre-eclampsia among American Indians.
607 *Am J Reprod Immunol.* 2012 Feb;67(2):152-9. doi: 10.1111/j.1600-
608 0897.2011.01076.x. Epub 2011 Oct 17. PMID: 22004660; PMCID: PMC3253256.
- 609 48. Shara NM, Wang H, Valaitis E, Pehlivanova M, Carter EA, Resnick HE, et al.
610 Comparison of estimated glomerular filtration rates and albuminuria in predicting
611 risk of coronary heart disease in a population with high prevalence of diabetes
612 mellitus and renal disease. *Am J Cardiol.* 2011 Feb 1;107(3):399-405. doi:
613 10.1016/j.amjcard.2010.09.036. PMID: 21257005; PMCID: PMC3035999.
- 614 49. Wang H, Shara NM, Lee ET, Devereux R, Calhoun D, de Simone G, et al.
615 Hemoglobin A1c, fasting glucose, and cardiovascular risk in a population with
616 high prevalence of diabetes: the strong heart study. *Diabetes Care.* 2011
617 Sep;34(9):1952-8. doi: 10.2337/dc11-0329. Epub 2011 Jul 25. PMID: 21788631;
618 PMCID: PMC3161283.
- 619 50. Best LG, Balakrishnan P, Cole SA, Haack K, Kocarnik JM, Pankratz N, et al.
620 Genetic analysis of hsCRP in American Indians: The Strong Heart Family Study.
621 *PLoS One.* 2019 Oct 17;14(10):e0223574. doi: 10.1371/journal.pone.0223574.
622 PMID: 31622379; PMCID: PMC6797125.
- 623 51. Voight BF, Kang HM, Ding J, et al. The metabochip, a custom genotyping array
624 for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS*
625 *Genet.* 2012;8(8):e1002793. doi: 10.1371/journal.pgen.1002793. Epub 2012 Aug
626 2. Erratum in: *PLoS Genet.* 2013 Apr;9(4). doi:10.1371/annotation/0b4e9c8b-
627 35c5-4dbd-b95b-0640250fbc87. PMID: 22876189; PubMed Central PMCID:
628 PMC3410907.
- 629 52. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general
630 pedigrees. *Am J Hum Genet.* 1998. 62:1198-1211.
- 631 53. <https://solar-eclipse-genetics.org>

- 632 54. Falconer DS, Mackay FC. *Introduction to Quantitative Genetics*, 4th ed. Harlow:
633 Prentice Hall. pp. 299-309.
- 634 55. US CDC. COVID data tracker: weekly case trends – the United States.
635 <https://covid.cdc.gov/covid-data-tracker/#datatracker-home> . Accessed
636 10/24/2022.
- 637 56. Williams JT, Blangero J. Power of variance component linkage analysis – II.
638 Discrete traits. *Ann Hum Genet*, 2004. 68:620-632.
- 639 57. <http://genome.ucsc.edu/>, accessed 12/15/22
- 640 58. Sturrock BR, Milne KM, Chevassut TJ. The renin-angiotensin system - a
641 therapeutic target in COVID-19? *Clin Med (Lond)*. 2020 Jul;20(4):e72-e75. doi:
642 10.7861/clinmed.2020-0146. Epub 2020 May 15. PMID: 32414711; PMCID:
643 PMC7385758.
- 644 59. Rasmi Y, Hatamkhani S, Naderi R, Shokati A, Nayeb Zadeh V, Hosseinzadeh F,
645 et al. Molecular signaling pathways, pathophysiological features in various
646 organs, and treatment strategies in SARS-CoV2 infection. *Acta Histochem*. 2022
647 Jul;124(5):151908. doi: 10.1016/j.acthis.2022.151908. Epub 2022 May 25. PMID:
648 35662001; PMCID: PMC9130726.
- 649 60. Saad H, Jabotian K, Sakr C, Mahfouz R, Akl IB, Zgheib NK. The Role of
650 Angiotensin Converting Enzyme 1 Insertion/Deletion Genetic Polymorphism in
651 the Risk and Severity of COVID-19 Infection. *Front Med (Lausanne)*. 2021 Dec
652 23;8:798571. doi: 10.3389/fmed.2021.798571. PMID: 35004773; PMCID:
653 PMC8733297.
- 654 61. Rokni M, Heidari Nia M, Sarhadi M, Mirinejad S, Sargazi S, Moudi M, et al.
655 Association of TMPRSS2 Gene Polymorphisms with COVID-19 Severity and
656 Mortality: a Case-Control Study with Computational Analyses. *Appl Biochem
657 Biotechnol*. 2022 Aug;194(8):3507-3526. doi: 10.1007/s12010-022-03885-w.
658 Epub 2022 Apr 7. PMID: 35386063; PMCID: PMC8986508.
- 659 62. Giannitrapani L, Augello G, Mirarchi L, Amodeo S, Veronese N, Sasso BL, et al.
660 Outcome predictors in SARS-CoV-2 disease (COVID-19): The prominent role of
661 IL-6 levels and an IL-6 gene polymorphism in a western Sicilian population. *J
662 Infect*. 2022 Aug;85(2):174-211. doi: 10.1016/j.jinf.2022.04.043. Epub 2022 Apr
663 29. PMID: 35490738; PMCID: PMC9050196.
- 664 63. Ishak A, Mehendale M, AlRawashdeh MM, Sestacovschi C, Sharath M, Pandav
665 K, et al. The association of COVID-19 severity and susceptibility and genetic risk
666 factors: A systematic review of the literature. *Gene*. 2022 Jun 14;836:146674.
667 doi: 10.1016/j.gene.2022.146674. Epub ahead of print. PMID: 35714803;
668 PMCID: PMC9195407.
- 669 64. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of
670 COVID-19. *Nature*. 2021 Dec;600(7889):472-477. doi: 10.1038/s41586-021-
671 03767-x. Epub 2021 Jul 8. PMID: 34237774; PMCID: PMC8674144.
- 672 65. <https://www.ncbi.nlm.nih.gov/snp/rs1205>, 7/24/23.
- 673 66. Wang Y, Wang Q, Guo C, Wang S, Wang X, An L, et al. Association between
674 CRP gene polymorphisms and the risk of preeclampsia in Han Chinese women.
675 *Genet Test Mol Biomarkers*. 2014 Nov;18(11):775-80. doi:
676 10.1089/gtmb.2014.0142. Epub 2014 Oct 14. PMID: 25314633.

- 677 67. Rod JE, Oviedo-Trespalacios O, Cortes-Ramirez J. A brief-review of the risk
678 factors for covid-19 severity. *Rev Saude Publica*. 2020;54:60. doi:
679 10.11606/s1518-8787.2020054002481. Epub 2020 Jun 1. PMID: 32491116;
680 PMCID: PMC7263798.
- 681 68. Dan S, Pant M, Upadhyay SK. The Case Fatality Rate in COVID-19 Patients
682 With Cardiovascular Disease: Global Health Challenge and Paradigm in the
683 Current Pandemic. *Curr Pharmacol Rep*. 2020;6(6):315-324. doi:
684 10.1007/s40495-020-00239-0. Epub 2020 Sep 15. PMID: 32953401; PMCID:
685 PMC7490208.
- 686 69. Wilk-Sledziowska K, Sielatycki PJ, Uscinska N, Bujno E, Rosolowski M,
687 Kakareko K, et al. The Impact of Cardiovascular Risk Factors on the Course of
688 COVID-19. *J Clin Med*. 2022 Apr 18;11(8):2250. doi: 10.3390/jcm11082250.
689 PMID: 35456343; PMCID: PMC9026388.
- 690 70. Andrasik MP, Maunakea AK, Oseso L, Rodriguez-Diaz CE, Wallace S, Walters
691 K, et al. Awakening: The Unveiling of Historically Unaddressed Social Inequities
692 During the COVID-19 Pandemic in the United States. *Infect Dis Clin North Am*.
693 2022 Jun;36(2):295-308. doi: 10.1016/j.idc.2022.01.009. Epub 2022 Feb 1.
694 PMID: 35636901; PMCID: PMC8806123.
- 695 71. Khanijahani A, Iezadi S, Gholipour K, Azami-Aghdash S, Naghibi D. A systematic
696 review of racial/ethnic and socioeconomic disparities in COVID-19. *Int J Equity
697 Health*. 2021 Nov 24;20(1):248. doi: 10.1186/s12939-021-01582-4. PMID:
698 34819081; PMCID: PMC8611382.
- 699 72. Koh D. Occupational risks for COVID-19 infection. *Occup Med (Lond)*. 2020 Mar
700 12;70(1):3-5. doi: 10.1093/occmed/kqaa036. PMID: 32107548; PMCID:
701 PMC7107962.
- 702 73. Kovich H. Rural Matters - Coronavirus and the Navajo Nation. *N Engl J Med*.
703 2020;383(2):105-107. doi:10.1056/NEJMp2012114
- 704 74. Cohen, P. N. (2020). The COVID-19 Epidemic in Rural U.S. Counties. *European
705 Journal of Environment and Public Health*, 4(2),
706 em0050. <https://doi.org/10.29333/ejeph/8331>
707
708