

1 **Interleukin-6 receptor genetic variation and tocilizumab treatment response to COVID-19**

2

3 Ammar Ali Almarzooq<sup>1</sup>

4

5 <sup>1</sup>Galore Consultancy, Flat 12, Building 144, Road 336, Block 314, Manama, Bahrain.

6

7 Corresponding author:

8 Ammar Ali Almarzooq,

9 Flat 12, Building 144, Road 336, Block 314, Manama, Bahrain.

10 tel: +973-32171842

11 email: [ammar@galoreconsultancy.com](mailto:ammar@galoreconsultancy.com)

12 **ABSTRACT**

13 Interleukin-6 receptor (IL6R) stimulates the inflammatory pathways as part of the acute-phase  
14 response to infection. Tocilizumab is a monoclonal antibody that inhibits both membrane-bound and  
15 soluble IL6R and is used to treat inflammatory conditions, including COVID-19. Despite the  
16 disproportionate incidence of COVID-19 among underserved, racial, and ethnic minority populations,  
17 the efficacy of tocilizumab in hospitalized COVID-19 patients from these populations is unclear. In  
18 this work, three genetic markers for the *IL6R* gene were analyzed across diverse ethnic backgrounds  
19 to identify population differences in response to tocilizumab treatment. Genetic structure analyses  
20 showed that African populations were significantly different from other described populations. In  
21 addition, mapped frequencies of these alleles showed that Sub-Saharan African populations were 3.4x  
22 more likely to show an impaired response to tocilizumab than East Asian populations, and 1.8x more  
23 likely than European ancestry populations. Existing *IL6R* genotype results may identify populations at  
24 increased therapeutic failure risk. As results from current clinical trials on the efficacy of tocilizumab  
25 treatment for extreme COVID-19 infections are conflicting, more studies are needed across diverse  
26 patient backgrounds to better understand the genetic factors necessary to predict treatment efficacy.  
27 This work demonstrates how pharmacogenomics studies can elucidate genetic variation on treatment  
28 efficacy on COVID-19.

29

30 **KEYWORDS:** COVID-19, tocilizumab, *IL6R*, pharmacogenetics, population genetics

## 31 INTRODUCTION

32 Over 150 million cases and 3.1 million deaths have occurred world-wide as a result of the coronavirus  
33 disease 2019 (COVID-19) pandemic [1]. Treatment options are still limited, with only glucocorticoids  
34 known to improve survival rates among severely ill patients [2] by alleviating the excessive host  
35 inflammatory response [3].

36 Interleukin-6 receptor (IL6R) is released in response to infection and stimulates the  
37 inflammatory pathways as part of the acute-phase response. Tocilizumab, a monoclonal antibody that  
38 inhibits membrane-bound and soluble interleukin-6 receptors, is used to treat inflammatory conditions  
39 such as rheumatoid arthritis and cytokine release syndrome following cell therapy. Clinical use has  
40 been described in COVID-19 [4-6], and following successful trials, tocilizumab has now become a  
41 standard National Health Service NHS treatment for specific COVID-19 patients [7]. In addition, The  
42 Infectious Diseases Society of America (IDSA) has backed the use of tocilizumab for COVID-19  
43 treatment [8]. However, results from clinical trials using this treatment are mixed, with specific trials  
44 finding no benefit of the treatment among patients with severe disease [9-12]. In one case, tocilizumab  
45 treatment made no significant difference in clinical status or mortality rate even after 28 days among  
46 hospitalized patients with severe COVID-19 pneumonia [10]. However, another independent study  
47 demonstrated that survival improved ~70% for patients treated with tocilizumab (or sarilumab),  
48 relative to standard care alone [3]. This is consistent with another U.K. based study showing a slight,  
49 but significant, mortality benefit for critically ill COVID-19 patients treated with tocilizumab [13].  
50 Comparisons across these studies is misleading, however, due to varying sample sizes, demographics,  
51 and methodology.

52 Genetic factors are one of the major contributors to individual and ethnic differences in drug  
53 therapeutic efficacy and toxicity [14-15]. Therefore, medication dosing might need to be altered  
54 based on a patient's genetic information [16-17]. There are several gene variants that can alter  
55 metabolism and processing of tocilizumab, potentially increasing the risk of therapeutic failure [18-  
56 19]. Patients with *IL6R* rs4329505 CC and CT genotypes, may have decreased response to  
57 tocilizumab when compared to patients with TT genotype [18]. Also, rs12083537 AA genotype is  
58 associated with decreased response to tocilizumab and higher risk for asthma compared to AG

59 genotype [18]. Finally, rs11265618 CC genotype is associated with an increased response to  
60 tocilizumab when compared to CT and TT genotypes [19].

61 The goal of this study was to use pharmacogenetic and modeling based approaches to better  
62 understand, and potentially predict, genetic factors related to treatment efficacy of tocilizumab. Such  
63 results may elucidate the underlying genetic factors and resolve the conflicting reports from COVID-  
64 19 clinical trials treating with tocilizumab. We show that pharmacogenomics can enhance our  
65 understanding of COVID-19 treatment failure and support advancement of the drug development  
66 pipelines.

67

## 68 **RESULTS**

### 69 *Genetic structure of IL6R across populations*

70 Three variants across 26 global populations from the 1000 Genomes Project Phase III (1kG-p3)  
71 dataset were selected (TAB. S1). SNPs were tested for Hardy–Weinberg equilibrium (HWE), and  
72 populations included in this study met HWE standards ( $p > 0.05$ ) across the three SNPs, except for the  
73 Esan in Nigeria population (ESN) ( $p = 0.039$ ) (TAB. S1). rs4329505 was in pair-wise linkage  
74 disequilibrium (LD) with rs11265618 ( $r^2 \geq 0.90$ ) in European (EUR) and East Asian (EAS)  
75 populations.

76 The two leading principal components from the three variants, showed well-defined  
77 separation between African (AFR) populations and other global populations (FIG. 1a). This was  
78 further validated using pairwise  $F_{ST}$  analyses (TAB. S2). The lowest level of differentiation was  
79 observed between EUR and South Asian (SAS) populations, followed by East Asian and American  
80 (AME) populations. The greatest affinity with AFR was observed with EUR and South Asian  
81 populations (FIG. 1b).

82

### 83 *Pharmacogenetic analyses by biogeographic grouping system*

84 Across the nine biogeographical groups, 80% of subjects were of EUR, Sub-Saharan African (SSA),  
85 South Central Asian, and East Asian origin (20% each), with Latino (LAT) and African  
86 Americans/Afro-Caribbeans (AAC) comprising the remaining 20% (14% and 6%, respectively)

87 (TAB. 1). SNPs were tested for Hardy–Weinberg equilibrium (HWE), and all the studied  
88 biogeographical groups fulfilled HWE ( $p > 0.05$ ) (TAB. 1).

89 Distinct differences were found among these populations, with direct impacts to tocilizumab  
90 clinical outcomes (FIG. 2). rs12083537 allele frequency was significantly higher for SSA (0.337),  
91 followed by AAC (0.295), EUR (0.193), LAT (0.166) and South Central Asians (0.158) indicating a  
92 decreased metabolism and clearance of tocilizumab when compared to East Asians (0.110) (TAB. 1).  
93 rs11265618 allele frequency was significantly higher in SSA (0.347), followed by AAC (0.291),  
94 South Central Asians (0.194), EUR (0.186) and LAT (0.128) indicating a decreased metabolism and  
95 clearance of tocilizumab when compared to East Asians (0.104) (TAB. 1). Suggesting that SSA  
96 populations are 3.4x more likely to show impaired response to tocilizumab than East Asian  
97 populations, and 1.8x more likely than EUR populations (FIG. 3 and TAB. 1). The rs4329505 marker  
98 was excluded from the analysis due to contradictory genotype-phenotype studies toward tocilizumab  
99 (FIG. 2) (18,19).

100

## 101 **DISCUSSION**

102 Despite the disproportionate incidence of COVID-19 infection among racial and ethnic minority  
103 populations [20-28], data regarding the safety and efficacy of repurposed treatments for patients  
104 within these populations are lacking [29]. In addition, conflicting results of the use of these treatment  
105 in clinical trials warrants further investigation. *IL6R* genotyping may help identify populations at an  
106 increased risk of experiencing ADRs or therapeutic failure as a result of tocilizumab treatment.  
107 However, the role of race and ethnicity on the clinical course of COVID-19 is complex due to the  
108 potential interplay of genetics, social determinants, socioeconomic factors, and historical and  
109 structural inequities[23, 25, 28, 30].

110 In this work, several genetic markers were analyzed across diverse ethnic backgrounds to identify  
111 population level differences in drug responses may be associated with tocilizumab treatment. MDS  
112 analysis showed that the African populations significantly deviated from other populations. MDS  
113 results were further validated by pairwise  $F_{ST}$  values (FIG. 1). Mapped frequencies of these alleles  
114 showed that Sub-Saharan African populations were 3.4x more likely to show impaired *IL6R* than East

115 Asian populations, and 1.8x more likely than European ancestry populations (FIG. 3 and TAB. 1).  
116 Therefore, it is plausible that a higher proportion of Asian and European ancestry populations could  
117 have a “normal” tocilizumab response, and therefore may be less susceptible to complications from  
118 tocilizumab-based treatment. Our analysis excluded rs4329505 marker not only due to conflicting  
119 research on genotype-phenotype for tocilizumab treatment (FIG. 2) [19, 20], but also because the pair-  
120 wise LD with rs11265618 ( $r^2 \geq 0.90$ ) in Europeans and East Asian populations. This is consistent  
121 with the recently updated PharmGKB Clinical Annotations change downgrading Level of Evidence  
122 from three to four [31]. However, polymorphisms in other genes may also be associated with  
123 tocilizumab response. Patients with the *CD69 rs11052877* AA genotype, for example, can have an  
124 increased response to tocilizumab compared to patients with the AG and GG genotypes [32]. Also,  
125 patients with the *FCGR3A rs396991* AA genotype may have an increased response to tocilizumab as  
126 compared with patients with the AC or CC genotypes [33]. Finally, *GALNT18 rs4910008* CC  
127 individuals may also have an increased response to tocilizumab compared with patients with the CT  
128 and TT genotypes [32].

129 Results from this study show that pharmacogenomics studies can enhance our understanding  
130 of adverse reactions to COVID-19 treatments and support advancement of drug development  
131 pipelines.

132

## 133 **METHODS**

### 134 *Population structure analyses*

135 Population structure was examined using Wright’s Fixation Index ( $F_{ST}$ ) for pairwise distances  
136 between populations based on three SNPs across *IL6R* to identify the ancestral relatedness across 26  
137 global populations from the 1000 Genomes Project Phase III (1kG-p3) dataset [34]. Ethnically  
138 defined populations by the 1kG-p3 dataset consisted of 2504 individuals from 26 populations within  
139 five defined ancestral groups (TAB. S1):

- 140 1. African (AFR): Americans of African Ancestry in SW, USA (ASW), Esan in Nigeria  
141 (ESN), Gambian in Western Divisions in the Gambia (GWD), Mende in Sierra Leone

142 (MSL), Luhya in Webuye, Kenya (LWK), Yoruba in Ibadan, Nigeria (YRI) and African  
143 Caribbeans in Barbados (ACB).

144 2. European (EUR): Utah Residents (CEPH) with Northern and Western European  
145 Ancestry (CEU), Finnish in Finland (FIN), British in England and Scotland (GBR),  
146 Iberian Population in Spain (IBS) and Toscani in Italia (TSI).

147 3. East Asian (EAS): Han Chinese in Beijing, China (CHB), Chinese Dai in Xishuangbanna,  
148 China (CDX), Southern Han Chinese (CHS), Japanese in Tokyo, Japan (JPT) and Kinh in  
149 Ho Chi Minh City, Vietnam (KHV).

150 4. South Asian (SAS): Bengali from Bangladesh (BEB), Indian Telugu from the UK (ITU),  
151 Punjabi from Lahore, Pakistan (PJL), Tamil from the UK (STU) and Gujarati Indian from  
152 Houston, Texas (GIH).

153 5. American (AMR): Mexican Ancestry from Los Angeles USA (MXL), Puerto Ricans  
154 from Puerto Rico (PUR), Colombians from Medellin, Colombia (CLM) and Peruvians  
155 from Lima, Peru (PEL).

156 Genotype and allele frequencies were tested for deviations from Hardy-Weinberg equilibrium using a  
157 chi-square ( $\chi^2$ ) test ( $p > 0.05$ ). These were implemented with the HWChisq function in  
158 HardyWeinberg R package [35]. Fixation indices were estimated with the calcFst function in the  
159 polysat R package [36]. Divergences were visualized by multidimensional scaling analysis (MDS)  
160 using the metaMDS function in the vegan R package [37] and with a hierarchical clustering based on  
161  $F_{ST}$  values using the hclust function in base R.

162

### 163 *Pharmacogenetic analyses*

164 Frequencies of the three pharmacogenetics biomarkers were assessed cumulatively for nine  
165 biogeographical groups [38], that consisted of 2,504 individuals from 26 global populations [34].

166 These nine groups were defined by global autosomal genetic structure and based on data from large-  
167 scale sequencing initiatives, and are used to illustrate the broad diversity of global allele frequencies  
168 in this study. Genotype and allele frequencies were tested for deviations from Hardy-Weinberg  
169 equilibrium using a  $\chi^2$  test ( $p > 0.05$ ). These were implemented with the HWChisq function in

170 HardyWeinberg R package [35]. LD analysis was performed using the LDlink tool to generate  $r^2$   
171 values [39]. This biogeographic grouping system meets a key need in pharmacogenetics research by  
172 enabling consistent communication of the scale of variability in global allele frequencies and is now  
173 used by PharmGKB and CPIC [38].

174 1. American (AME): The American genetic ancestry group includes populations from both  
175 North and South America with ancestors predating European colonization, including  
176 American Indian, Alaska Native, First Nations, Inuit, and Métis in Canada, and  
177 Indigenous peoples of Central and South America.

178 2. Central/South Asian (SAS): The Central and South Asian genetic ancestry group  
179 includes populations from Pakistan, Sri Lanka, Bangladesh, India, and ranges from  
180 Afghanistan to the western border of China.

181 3. East Asian (EAS): The East Asian genetic ancestry group includes populations from  
182 Japan, Korea, and China, and stretches from mainland Southeast Asia through the islands  
183 of Southeast Asia. In addition, it includes portions of central Asia and Russia east of the  
184 Ural Mountains.

185 4. European (EUR): The European genetic ancestry group includes populations of primarily  
186 European descent, including European Americans. We define the European region as  
187 extending west from the Ural Mountains and south to the Turkish and Bulgarian border.

188 5. Near Eastern (NEA): The Near Eastern genetic ancestry group encompasses populations  
189 from northern Africa, the Middle East, and the Caucasus. It includes Turkey and African  
190 nations north of the Saharan Desert.

191 6. Oceanian (OCE): The Oceanian genetic ancestry group includes pre-colonial populations  
192 of the Pacific Islands, including Hawaii, Australia, and Papua New Guinea.

193 7. Sub-Saharan African (SSA): The Sub-Saharan African genetic ancestry group includes  
194 individuals from all regions in Sub-Saharan Africa, including Madagascar.

195 8. African American/Afro-Caribbean (AAC): Individuals in the African American/Afro-  
196 Caribbean genetic ancestry group reflect the extensive admixture between African,  
197 European, and Indigenous ancestries and, as such, display a unique genetic profile



198 compared to individuals from each of those lineages alone. Examples within this cluster  
199 include the Coriell Institute’s African Caribbean in Barbados (ACB) population, the  
200 African Americans from the Southwest US (ASW) population, and individuals from  
201 Jamaica and the US Virgin Islands.

202 9. Latino (LAT): The Latino genetic ancestry group is not defined by an exclusive  
203 geographic region, but includes individuals of Mestizo descent, individuals from Latin  
204 America, and self-identified Latino individuals in the United States. Like the African  
205 American/Afro-Caribbean group, the admixture in this population creates a unique  
206 genetic pattern compared to any of the discrete geographic regions, with individuals  
207 reflecting mixed native and indigenous American, European, and African ancestry.

208 The total frequency of the three SNPs within these nine geographically-defined groups were mapped  
209 for global impact visualization of allele frequency on tocilizumab response (TAB. 1, FIG. 2). Inferred  
210 frequencies for contradictory genotype-phenotype were excluded from our biogeographical analyses.

211

212 **REFERENCES**

- 213 1. World Health Organization. Coronavirus disease (COVID-19) pandemic.  
214 <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. (2021).
- 215 2. The WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group.  
216 Association between administration of systemic corticosteroids and mortality among critically  
217 ill patients with COVID-19: a meta-analysis. *JAMA*. **324**, 1330-1341 (2020).
- 218 3. The REMAP-CAP Investigators. Interleukin-6 Receptor Antagonists in Critically Ill Patients  
219 with Covid-19. *N. Engl. J. Med.* 10.1056/NEJmoa2100433 (2021).
- 220 4. Xu, X. *et al.* Effective treatment of severe COVID-19 patients with tocilizumab. *Proc. Natl.*  
221 *Acad. Sci. USA*. **117**, 10970-10975 (2020).
- 222 5. Gremese, E. *et al.* Sarilumab use in severe SARS-CoV-2 pneumonia. *EClinical Medicine*. **27**,  
223 100553-100553 (2020).
- 224 6. Somers, E. C. *et al.* Tocilizumab for treatment of mechanically ventilated patients with  
225 COVID-19. *Clin. Infect. Dis.* 10.1093/cid/ciaa954 (2020).
- 226 7. NHS. Interim Clinical Commissioning Policy: Tocilizumab for hospitalised patients with  
227 COVID-19 pneumonia (adults). 1-5 (2021).
- 228 8. Bhimraj, A. *et al.* Infectious Disease Society of America Guidelines on the Treatment and  
229 Management of Patients with COVID-19. *Infectious Disease Society of America*. Version  
230 4.2.0 (2021).
- 231 9. Stone, J. H. *et al.* Efficacy of Tocilizumab in Patients Hospitalized with Covid-19. *N. Engl. J.*  
232 *Med.* **383**, 2333-2344 (2020).
- 233 10. Rosas, I. O. *et al.* Tocilizumab in Hospitalized Patients with Severe Covid-19 Pneumonia. *N.*  
234 *Engl. J. Med.* 10.1056/NEJMoa2028700 (2021).
- 235 11. Salvarani, C. *et al.* Effect of Tocilizumab vs Standard Care on Clinical Worsening in Patients  
236 Hospitalized With COVID-19 Pneumonia: A Randomized Clinical Trial. *JAMA Intern. Med.*  
237 181, 24-31 (2021).

- 238 12. Hermine, O. et al. Effect of Tocilizumab vs Usual Care in Adults Hospitalized With COVID-  
239 19 and Moderate or Severe Pneumonia: A Randomized Clinical Trial. *JAMA Intern. Med.*  
240 181, 32-40 (2021).
- 241 13. Horby, P. W. *et al.* Tocilizumab in patients admitted to hospital with COVID-19  
242 (RECOVERY): preliminary results of a randomised, controlled, open-label, platform trial.  
243 *MedRxIV*. 10.1101/2021.02.11.21249258 (2021).
- 244 14. Gray, J. A. The shift to personalised and population medicine. *Lancet*. **382**, 200-201 (2013).
- 245 15. Peterson, J. F. *et al.* Building evidence and measuring clinical outcomes for genomic  
246 medicine. *Lancet*. **394**, 604-610 (2019).
- 247 16. Burchard, E. G. *et al.* The importance of race and ethnic background in biomedical research  
248 and clinical practice. *N. Engl. J. Med.* **348**, 1170-1175 (2003).
- 249 17. Gemmati, D. & Tisato, V. Genetic Hypothesis and Pharmacogenetics Side of Renin-  
250 Angiotensin-System in COVID-19. *Genes (Basel)*. **11**, 1044 (2020).
- 251 18. Enevold, C. *et al.* Interleukin-6-receptor polymorphisms rs12083537, rs2228145, and  
252 rs4329505 as predictors of response to tocilizumab in rheumatoid arthritis. *Pharmacogenet.*  
253 *Genomics*. **24**, 401-5 (2014).
- 254 19. Maldonado-Montoro, M., Cañadas-Garre, M., González-Utrilla, A., & Ángel Calleja-  
255 Hernández, M. Influence of *IL6R* gene polymorphisms in the effectiveness to treatment with  
256 tocilizumab in rheumatoid arthritis. *Pharmacogenomics J.* **18**, 167-172 (2018).
- 257 20. Pan, D. *et al.* The impact of ethnicity on clinical outcomes in COVID-19: a systematic  
258 review. *EClinical Medicine*. **23**, 100404-100404 (2020).
- 259 21. Centers for Disease Control and Prevention. Cases, data and surveillance - COVID-19  
260 hospitalization and death by race/ethnicity. [https://www.cdc.gov/coronavirus/2019-  
261 ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html).  
262 Accessed November 30, 2020.
- 263 22. Williamson, E. J. *et al.* Factors associated with COVID-19-related death using  
264 OpenSAFELY. *Nature*. **584**, 430-436 (2020).

- 265 23. Egede, L. E. & Walker, R. J. Structural racism, social risk factors, and Covid-19 - a  
266 dangerous convergence for Black Americans. *N. Engl. J. Med.* **383**, e77-e77 (2020).
- 267 24. Karaca-Mandic, P., Georgiou, A., & Sen, S. Assessment of COVID-19 hospitalizations by  
268 race/ethnicity in 12 states. *JAMA Intern. Med.* **181**, 131-134 (2021).
- 269 25. Evans, M. K. Covid's color line - infectious disease, inequity, and racial justice. *N. Engl. J.*  
270 *Med.* **383**, 408-410 (2020).
- 271 26. Price-Haywood, E. G., Burton, J., Fort, D., & Seoane, L. Hospitalization and mortality among  
272 Black patients and White patients with Covid-19. *N. Engl. J. Med.* **382**, 2534-2543 (2020).
- 273 27. Chowkwanyun, M. & Reed, A. L. Jr. Racial health disparities and Covid-19 - caution and  
274 context. *N. Engl. J. Med.* **383**, 201-203 (2020).
- 275 28. Golestaneh, L. *et al.* The association of race and COVID-19 mortality. *EClinical Medicine*  
276 **25**, 100455-100455 (2020).
- 277 29. Salama, C. *et al.* Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia. *N. Engl. J.*  
278 *Med.* **384**, 20-30 (2021).
- 279 30. Department of Health and Human Services. Social determinants of health. 2021.
- 280 31. Whirl-Carrillo, M. *et al.* Pharmacogenomics Knowledge for Personalized Medicine. *Clin.*  
281 *Pharmacol. Ther.* **92**, 414-417 (2012).
- 282 32. Maldonado-Montoro, M. Cañadas-Garre, M., González-Utrilla, A., Plaza-Plaza, J. C., &  
283 Calleja-Hernández, M. Y. Genetic and clinical biomarkers of tocilizumab response in patients  
284 with rheumatoid arthritis. *Pharmacol. Res.* **111**, 264-271 (2016).
- 285 33. Jiménez Morales, A. *et al.* FCGR2A/FCGR3A gene polymorphisms and clinical variables as  
286 predictors of response to tocilizumab and rituximab in patients with rheumatoid arthritis. *J.*  
287 *Clin. Pharmacol.* **59**, 517-531 (2019).
- 288 34. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation.  
289 *Nature.* **526**, 68-74 (2015).
- 290 35. Graffelman, J. Exploring Diallelic Genetic Markers: The HardyWeinberg Package. *J. Stat.*  
291 *Software.* **64**, 1-23 (2015).

- 292 36. Clark, L. V. & Schreier, A. D. Resolving microsatellite genotype ambiguity in populations of  
293 allopolyploid and diploidized autopolyploid organisms using negative correlations between  
294 allelic variables. *Mol. Ecol. Resour.* 17, 1090-1103 (2017).
- 295 37. Oksanen, J. B. et al. Vegan: Community Ecology Package. [https://cran.r-](https://cran.r-project.org/web/packages/vegan/index.html)  
296 [project.org/web/packages/vegan/index.html](https://cran.r-project.org/web/packages/vegan/index.html). (2020).
- 297 38. Huddart, R. et al. Standardized Biogeographic Grouping System for Annotating Populations  
298 in Pharmacogenetic Research. *Clin. Pharmacol. Ther.* 105, 1256-1262 (2019).
- 299 39. Machiela, M. J. & Chanock, S. J. LD link: a web-based application for exploring population-  
300 specific haplotype structure and linking correlated alleles of possible functional variants.  
301 *Bioinformatics.* 31, 3555-3557 (2015).
- 302

303 **ACKNOWLEDGEMENTS**

304 N/A

305

306 **AUTHORS' CONTRIBUTIONS**

307 AA conceived the research study, performed all the analyses reported and wrote the manuscript.

308

309 **CONSENT FOR PUBLICATION**

310 N/A

311

312 **COMPETING INTERESTS**

313 The author declares that there is no conflict of interest.

314

315 **DATA AVAILABILITY**

316 The datasets generated during the current study are included in the supplementary files.

317

318 **ETHICS DECLARATIONS**

319 N/A

320

321 **FIGURE LEGENDS**

322 **Figure 1:** Population structure analysis. **a.** MDS plot of the three variants across 26 global  
323 populations from the 1000 Genomes Project Phase III (1kG-p3) dataset. **b.** Heat map showing  
324 pairwise  $F_{ST}$  values with hierarchical clustering. Yellow for the lowest and blue for the highest  
325  $F_{ST}$  values. All the abbreviations names and detail information of the reference populations are in  
326 TAB. S1 and S2.

327

328 **Figure 2:** Biogeographical groups and pharmacogenetic analyses for *IL6R* and tocilizumab  
329 associations. Frequencies of three pharmacogenetics (PGx) genotypes and phenotypes assessed  
330 cumulatively against the nine biogeographical groups.

331

332 **Figure 3:** Global frequency map of variability in of *IL6R* reduced response phenotypes on  
333 tocilizumab response. Since this map indicates the borders of each geographical group based on the  
334 location of genetic ancestors pre-Diaspora and pre-colonization, the two admixed groups (AAC and  
335 LAT) are absent. It should also be noted that, due to the large geographical areas designated for each  
336 group, a single population may not accurately capture the large amount of genetic diversity found in  
337 that defined region.

338

339 **TABLES**

340 **Table 1:** Distribution of *IL6R* alleles and genotypes assessed cumulatively against the

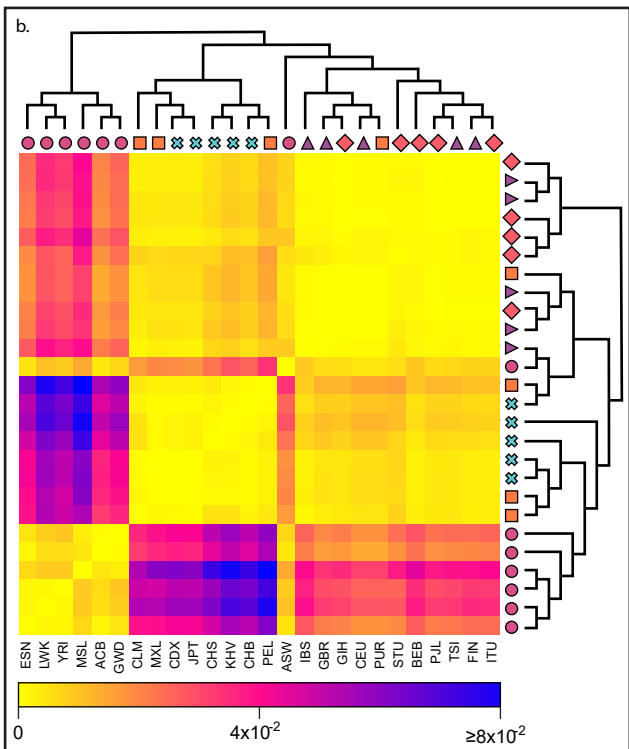
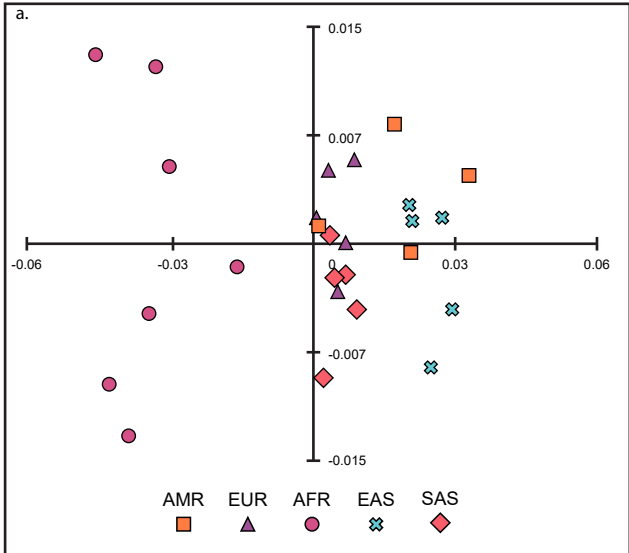
341 biogeographical groups (rs4329505, rs12083537, rs11265618)

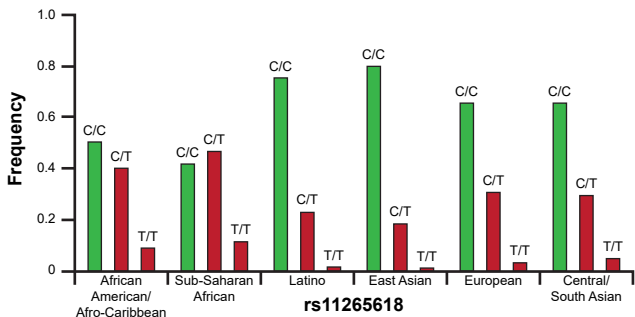
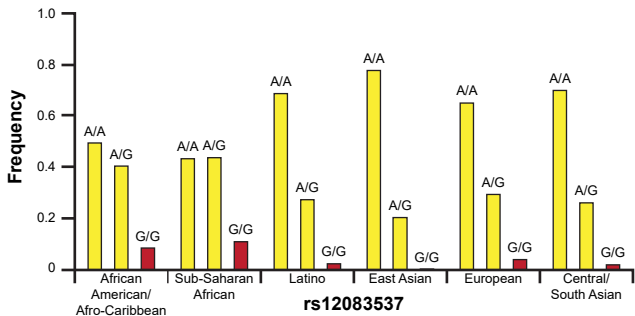
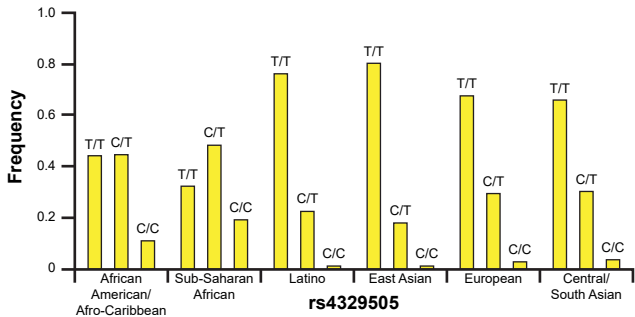
342

		<b>T</b>	<b>C</b>	<b>T/T</b>	<b>C/T</b>	<b>C/C</b>	<b>N</b>	<b>%</b>	<b><math>\chi^2</math></b>	<b>P-value</b>
		<b>(count)</b>	<b>(count)</b>	<b>(count)</b>	<b>(count)</b>	<b>(count)</b>				
rs4329505	African									
	American/ Afro- Caribbean	0.668 (208)	0.333 (106)	0.443 (69)	0.449 (70)	0.109 (18)	157	0.06	0.00	1.00
	Sub-Saharan African	0.565 (571)	0.435 (437)	0.323 (164)	0.484 (243)	0.193 (97)	504	0.20	0.17	0.92
	Latino	0.879 (610)	0.121 (84)	0.767 (266)	0.225 (78)	0.009 (3)	347	0.14	1.11	0.58
	East Asian	0.898 (905)	0.102 (103)	0.808 (407)	0.181 (91)	0.012 (6)	504	0.20	0.13	0.94
	European	0.826 (831)	0.174 (175)	0.678 (341)	0.296 (149)	0.026 (13)	503	0.20	0.48	0.79
	Central/ South Asian	0.812 (794)	0.188 (184)	0.661 (323)	0.303 (148)	0.037 (18)	489	0.20	0.04	0.98
rs12083537	African									
	American/ Afro- Caribbean	0.705 (218)	0.295 (96)	0.501 (76)	0.410 (66)	0.091 (15)	157	0.06	0.01	0.99
	Sub-Saharan African	0.663 (671)	0.337 (337)	0.441 (224)	0.44 (223)	0.115 (57)	504	0.20	0.02	0.99
	Latino	0.834 (579)	0.166 (115)	0.695 (241)	0.280 (97)	0.026 (9)	347	0.14	0.04	0.98
	East Asian	0.890 (897)	0.110 (111)	0.786 (396)	0.208 (105)	0.006 (3)	504	0.20	2.00	0.37
	European	0.807 (812)	0.193 (194)	0.658 (331)	0.298 (150)	0.044 (22)	503	0.20	0.89	0.64
	Central/ South Asian	0.842 (823)	0.158 (155)	0.708 (346)	0.268 (131)	0.025 (12)	489	0.20	0.01	1.00
rs11265618	African									
	American/ Afro- Caribbean	0.709 (220)	0.291 (94)	0.5075 (78)	0.402 (64)	0.091 (15)	157	0.06	0.13	0.94
	Sub-Saharan African	0.653 (658)	0.347 (350)	0.419 (211)	0.467 (236)	0.113 (57)	504	0.20	0.55	0.76
	Latino	0.872 (605)	0.128 (89)	0.758 (263)	0.228 (79)	0.014 (5)	347	0.14	0.12	0.94
	East Asian	0.896 (903)	0.104 (105)	0.804 (405)	0.185 (93)	0.012 (6)	504	0.20	0.06	0.97
	European	0.814 (819)	0.186 (187)	0.660 (332)	0.308 (155)	0.032 (16)	503	0.20	0.17	0.92
	Central/ South Asian	0.806 (788)	0.194 (190)	0.658 (322)	0.294 (144)	0.047 (23)	489	0.20	1.72	0.42

343







■ Better response    
 ■ Contradictory studies    
 ■ Decreased or poorer response

