

1 **Study of the protective *OAS1* rs10774671-G allele against severe**
2 **COVID-19 in Moroccans suggests a North African origin for**
3 **Neanderthals**

4
5 **Fatima Zohra El Youssfi¹, Abbas Ermilo Haroun^{2,3}, Chaimae Nebhani¹, Jihane**
6 **Belayachi^{2,4}, Omar Askander⁵, Elmostafa El Fahime⁶, Hakima Fares⁷, Khalid Ennibi⁸,**
7 **Redouane Abouqal^{2,3}, Rachid Razine^{2,3}, Ahmed Bouhouche^{1,9*}**

8
9
10 1 Laboratory of Human Genetics, Medical School and Pharmacy, University Mohammed V in Rabat,
11 Morocco

12 2 Laboratory of Biostatistics, Clinical and Epidemiological Research, Department of Public Health,
13 Medical School and Pharmacy, University Mohammed V in Rabat, Morocco

14 3 Laboratory of Community Health, Department of Public Health, Medical School and Pharmacy,
15 University Mohammed V in Rabat, Morocco

16 4 Acute Medical Unit, Ibn Sina University Hospital, Rabat, Morocco

17 5 Faculty of Medical Science, Mohammed VI Polytechnic University, Benguerir, Morocco.

18 6 Molecular Biology and Functional Genomics Platform, National Center for Scientific and Technical
19 Research, Rabat, Morocco.

20 7 Intensive Care Department, Cheikh Zaid International University Hospital, Rabat, Morocco

21 8 Virology, Infectious and Tropical Diseases Center, Hôpital Militaire d'Instruction Mohammed V.
22 Rabat.

23 9 Genomic Center of the Cheikh Zaid Foundation, Abulcasis International University of Health
24 Sciences, Rabat, Morocco

25

26

27

28 * Correspondence author

29 E-mail: a.bouhouche@um5r.ac.ma (AB)

30 **NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ORCID: <https://orcid.org/0000-0001-5279-8649>

31 **Abstract**

32 **Background:** The clinical presentation of COVID-19 has shown high variability
33 between individuals, which is partly due to genetic factors. The *OAS1/2/3* cluster
34 was found to be strongly associated with COVID-19 severity. We aimed to
35 examine this locus for the occurrence of the critical variant, rs10774671, and its
36 respective haplotype blocks within the Moroccan population.

37 **Methods:** The frequency of SNPs at the cluster of *OAS* immunity genes was
38 assessed from an in-house database in 157 unrelated individuals of Moroccan
39 origin. The *OAS1* exon 6 was sequenced by Sanger's method in 71
40 asymptomatic/mild and 74 moderate/severe individuals positive for SARS-CoV-
41 2. Genotypic, allelic, and haplotype frequencies of three SNPs were compared
42 between the two groups. Finally, males in our COVID-19 series were genotyped
43 for the Berber-specific marker E-M81.

44 **Results:** The prevalence of the *OAS1* rs10774671-G allele in present-day
45 Moroccans was 40.4%, close to that of Europeans. However, it was found equally
46 on both the Neanderthal GGG haplotype and the African GAC haplotype with a
47 frequency of 20% each. These two haplotypes, and hence the rs10774671-G
48 allele, were significantly associated with the protection against severe COVID-19
49 ($p = 0.034$, $p = 0.041$, and $p = 0.008$ respectively). Surprisingly, among Berber
50 men, the African haplotype was absent while the prevalence of the Neanderthal
51 haplotype was close to that of Europeans.

52 **Conclusion:** The protective rs10774671-G allele of *OAS1* was found only in the

53 Neanderthal haplotype in Berbers, the indigenous people of North Africa,
54 suggesting that this region may have served as the stepping-stone for the passage
55 of the hominids to the other continents.

56

57

58 **Keywords:** SARS-CoV-2, COVID-19 severity, 2',5'-oligoadenylate synthetase,
59 single nucleotide polymorphism, Neanderthals, North Africa

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75 **1. Introduction**

76 COVID-19, caused by the coronavirus SARS-CoV-2, appeared in November
77 2019 in Wuhan, Hubei province (China), before spreading around the world and
78 causing a pandemic [1, 2]. The consequences of SARS-CoV-2 infection vary
79 greatly from person to person and from population to population. While most
80 infected individuals are asymptomatic or minimally symptomatic, some develop
81 severe or even critical illness. Older people and those with underlying medical
82 condition, such as cardiovascular disease, diabetes, chronic respiratory disease, or
83 cancer, are at greater risk of developing a severe form [3]. However, these risk
84 factors alone do not explain the variability in severity observed. These differences
85 could partly result from individual genetic susceptibility. The scientific
86 community has been mobilized to respond to the health crisis to identify the
87 genetic determinants of COVID-19 that protect people from or predispose them
88 to severe manifestations of the disease. A genetic study conducted on patients of
89 European ancestry identified a risk locus for severe COVID-19 in the 2'-5'-
90 oligoadenylate synthetase 1/2/3 (OAS1/2/3) gene cluster [4]. Subsequently, it was
91 reported that elevated levels of OAS1 in the plasma were associated with reduced
92 COVID-19 susceptibility and severity [5].

93 This genetic region of chromosome 12q24.13 carries a haplotype of
94 approximately 75 kb of Neanderthal origin [6], where the discovery of the causal
95 variant was possible through studies conducted on a multiancestry cohort
96 including patients from both African and European ancestry. It consists of

97 rs10774671 with G>A transition occurring at a splice acceptor site, where the
98 protective ancestral G allele produces a longer OAS1 transcript and a more active
99 protein [7, 8]. Moreover, *OAS1* SNPs have been reported to be associated with
100 susceptibility to West Nile virus infections [9] and to be involved in altered
101 cellular function leading to several autoimmune and infectious diseases,
102 suggesting that the *OAS* genes play a critical role in the innate immune response
103 to viruses [10-14]. For instance, the OAS1 enzyme catalyzes the synthesis of short
104 polyadenylates, which activate ribonuclease L contributing to the degradation of
105 viral RNA and inhibition of viral replication [15, 16]. It produces five transcripts
106 including p42, p44, p46, p48, and p52, and its expression is stimulated by the
107 interferon system [17, 18].

108 In this study, we assessed the prevalence of several variants at the *OAS* gene
109 cluster that comprise the Neanderthal haplotype by analyzing exome data of 157
110 unrelated Moroccan individuals and observed that the *OAS1* rs10774671-G allele
111 was found on both Neanderthal and African haplotypes. We then investigated the
112 rs10774671 variant in 146 Moroccan volunteers positive for SARS-CoV-2
113 infection and showed that it was significantly associated with COVID-19 severity.
114 Interestingly, in Berbers of both paternal and maternal lineages, the African
115 haplotype containing the G allele was absent while the frequency of Neanderthal-
116 derived haplotypes is close to that of Europeans.

117

118

119 **Material and methods**

120 **Frequency of *OAS* gene variants in the Moroccan population**

121 To assess the frequency in the Moroccan population of the Neanderthal haplotype
122 at the cluster of *OAS* immunity genes in the chromosomal region 12q24.13, we
123 analyzed exome data from 157 unrelated patients suspected of having a rare
124 hereditary disease. These patients were recruited from the Department of
125 Neurology of Specialties Hospital and Cheikh Zaid Hospital of Rabat. Part of the
126 data was taken from the in-house exome database and the other part comes from
127 whole exome sequencing (WES) carried out as part of a collaborative project with
128 3Billion® (Seoul, Korea). Data were collected for variants rs1131454,
129 rs10774671, rs1131476, rs1051042 and rs2660 in *OAS1*, rs1859330, rs1859329
130 and rs2285932 in *OAS3*, and the variant rs1293767 in *OAS2*, variants that
131 comprise the Neanderthal haplotype at the *OAS* locus as described elsewhere [19].

132

133 **COVID-19 Patients recruitment**

134 To highlight whether the variants studied are associated with the severity of
135 COVID-19 in the Moroccan population, it was essential to minimize the effects
136 of underlying medical conditions such as comorbidities by reducing the upper age
137 limit of patients. Thus, we have adopted the selection criteria including males or
138 females of Moroccan origin, aged between 18 and 65 years old, with positive RT-
139 PCR test for SARS-COV-2, who have not received any vaccines. Pregnant
140 patients were excluded from the study. 146 COVID-19 patients were enrolled for

141 1 year between 28 July 2020 to 15 July 2021, a period of the pandemic dominated
142 by the alpha and beta variants. All patients filled out a form concerning socio-
143 demographic data for the establishment of ethnicity and medical history, notifying
144 comorbidities, if applicable. Patients were recruited by three different hospital
145 centers involved in a national COVID-19 research consortium, including the
146 Medical Emergency Department (CHU Ibn Sina, Rabat), the Virology
147 Department (Military Hospital, Rabat), and the Intensive Care Unit (Cheikh Zaid
148 international university hospital, Rabat). Outcomes were defined as asymptomatic
149 (without symptoms), mild (non-hospitalized), moderate (hospitalized but not
150 requiring invasive ventilation), and severe (hospitalized requiring invasive
151 ventilation or death). 146 patients were recruited and grouped into the following
152 severity categories, asymptomatic (N = 21), mild COVID-19 (N = 50), moderate
153 COVID-19 (N = 48), and severe COVID-19 (N = 27). Asymptomatic and mild
154 patients were grouped (N=71) and compared to moderate/severe patients (N=75).
155 Five ml of venous blood was collected in an EDTA tube from each participant
156 and stored at -20°C until DNA extraction. The written informed consent was
157 signed by all patients before their enrollment. This study was approved by the
158 ethics committee of Cheikh Zaid Hospital of Rabat (ID: CEFCZ/AB/PR_RFG).

159

160 **Genetic analysis**

161 Genomic DNA (gDNA) of 146 SARS-CoV-2 positive volunteers was extracted
162 using the MagPurix® Blood DNA Extraction Kit with a Zinexts MagPurix EVO

163 24 CE IVD system (Zinexts Life Science Corp., New Taipei City, Taiwan).
164 Quality and concentrations of DNA samples were analyzed by Nanodrop One^c
165 Spectrophotometer (Thermo Fisher Scientific). PCR amplification, covering the
166 *OAS1* exon 6 and its intron boundaries, was performed using two pairs of specific
167 primers (forward: 5'-GTGGCCAGGCTTCTATACCC-3'; and reverse: 5'-
168 TGGAGTGTGCTGGGTCTATG-3') designed with Primer3 Input (V 0.4.0).
169 PCR amplification was performed using 100 ng of template gDNA, 1 μ L of each
170 primer (10 μ M), 5 μ L of 5X MyTaq reaction buffer, and 0.2 μ l of MyTaq DNA
171 polymerase (Bioline) in a total reaction volume of 25 μ L. Thermal conditions
172 were as follows: 95°C for 3 min, 35 cycles of 95°C for 5 sec, 60°C for 15 sec, and
173 72°C for 10 sec, with a final extension of 72°C for 10 min. The PCR products
174 were visualized by 1.5% agarose gel electrophoresis. Five microliters of PCR
175 amplicons were purified using ExoSap IT-express reagent and 2 microliters of the
176 product were used for cycle sequencing with BigDyeTM Terminator v3.1 Ready
177 Reaction Cycle Sequencing Kit. Electrophoresis was performed on SeqStudio
178 Genetic Analyzer, and Sequence data were analyzed by SeqScape 4 software
179 (Thermo Fisher Scientific).

180 The genetic marker, E-M81, reported as being the specific male lineage of
181 autochthonous Berbers of North Africa, was analyzed in all 80 males in the
182 COVID-19 positive cohort of patients. The marker was amplified by PCR and
183 Sanger sequenced as described previously [20].

185 **Statistical analysis**

186 Statistical analysis was performed using *jamovi* software (Version 1.6). The
187 evaluation of demographic variables and disease characteristics was done using
188 descriptive statistics. A chi-square (χ^2) test was used before the association tests
189 to assess deviation from Hardy-Weinberg equilibrium (HWE) in the control
190 group. The genotypic and allelic frequencies were compared between patients and
191 controls to check the associations of the studied SNPs with COVID-19 severity.
192 The χ^2 test was used to identify the differences in genotypic distribution between
193 study groups under an additive genetic model. For allelic and haplotype models,
194 Odds ratios (ORs) at 95% confidence intervals (CI) were estimated using logistic
195 regression analysis, and *p*-values less than 0.05 was considered statistically
196 significant.

197

198 **Results**

199 **Frequency of Neanderthal alleles at *OAS* gene cluster in Moroccans**

200 The genotype of 9 SNPs at the cluster of *OAS* immunity genes in the chromosomal
201 region 12q24.13 (Fig 1) was extracted from an exome sequencing database of 157
202 unrelated individuals, all of Moroccan origin. The results presented in Table 1
203 showed that, overall, the frequencies of these SNPs in Moroccans are close to the
204 frequency of the minor allele frequency (MAF) of the 1000 genome project.
205 Focusing on *OAS1* variants, the reference G allele of the *OAS1* splice variant
206 rs10774671 represented 40.4% in analyzed Moroccan individuals, close to that of

207 Europeans with 35.2%, but much lower than that of Africans with 64.0% and
208 much higher than that of Asians with only 27.2%. The frequencies of the missense
209 variants rs1131476 and rs1051042, and the 3'UTR rs2660 variant were 21.1%
210 identical, indicating that they are in complete linkage disequilibrium, which was
211 also the case in the other populations of European, Asian and African ancestry
212 (Table 1).

213 Fig 1: OAS gene cluster at chromosome 12q24.13 and structural organization of OAS1 giving
214 the location of the splice variant rs10774671

215

216 We then explored the haplotypes formed by the three SNPs rs10774671,
217 rs1131476, and rs1051042, which allowed for identification of the predominantly
218 protective Neanderthal-derived haplotype in this genetic region. The haplotypes
219 consisting of these three SNPs were GGG, GAC, and AAC, and the graphical
220 representation of their distribution in the present-day Moroccans in comparison
221 with different ancestries are shown in Fig 2. The *OAS1* rs10774671-G allele in
222 Moroccans was included in both Neanderthal GGG and African GAC haplotypes
223 with frequencies of 21% and 19.4% respectively. In comparison to other
224 ancestries, the GGG haplotype accounts for 34.5% of Europeans, 27.1% of
225 Asians, and only 2.7% of Africans. The GAC haplotype is however predominant
226 in Africans with a frequency of 61.3% and is almost absent in Europeans and
227 Asians with 0.7% and 0.001% respectively. Finally, the risk haplotype AAC was
228 the prevalent haplotype in the Asian population with 72.4% (Table 1).

Table 1: SNP Frequencies of the OAS locus in 157 individuals of Moroccan origin in comparison with those of the 1000 genomes project

| | OAS1 | | | | | OAS3 | | | OAS2 |
|----------------------------|--------------------------------------|--|---------------------------------------|---------------------------------------|-----------------------------|------------------------------------|-----------------------------------|-------------------------------------|--------------------------------------|
| | rs1131454 c.484G>A p.Gly162Ser | rs10774671 c.1039-1G>A Splice site | rs1131476 c.1054G>A p.Ala352Pro | rs1051042 c.1082G>C p.Arg361Met | rs2660 c.*84G>A 3'UTR | rs1859330 c.53G>A p.Arg18Lys | rs1859329 c.117C>T p.Ala39= | rs2285932 c.1314T>C p.Ile438= | rs1293767 c.448+41C>G Intronic |
| Minor allele frequency | 0.470 | 0.390 | 0.210 | 0.210 | 0.210 | 0.340 | 0.210 | 0.160 | 0.150 |
| African Ancestry | 0.851 | 0.640 | 0.027 | 0.027 | 0.027 | 0.446 | 0.011 | 0.010 | 0.011 |
| Asian Ancestry | 0.414 | 0.272 | 0.271 | 0.271 | 0.271 | 0.268 | 0.268 | 0.156 | 0.130 |
| European Ancestry | 0.425 | 0.352 | 0.345 | 0.345 | 0.345 | 0.370 | 0.365 | 0.328 | 0.325 |
| Moroccan population | 0.554 | 0.404 | 0.210 | 0.210 | 0.210 | 0.283 | 0.213 | 0.197 | 0.194 |

229 Fig 2. Prevalence of the three possible haplotypes with the three *OAS1* SNPs studied in 157
230 individuals of Moroccan origin in comparison with the data of 1000 Genome Project of
231 Africans, Asians and Europeans.

232

233 **Association study in Moroccan COVID-19 patients**

234 The demographic and clinical data of the two studied groups are shown in Table

235 2. The mean age and standard deviation of patients asymptomatic/mild were 44.1

236 ± 13.5 years and significantly lower than that of moderate/severe with 51.4 ± 10.5

237 years ($p < 0.001$). There was no significant difference regarding the sex ratio

238 between the two groups (OR = 1.04, $p < 0.897$), with a female-to-male ratio of

239 approximately 1:1. As expected, the patients in the asymptomatic/mild group

240 displayed a significantly lower frequency of comorbidity compared to the

241 moderate/severe group (OR = 2.85, $p < 0.004$), with 22.5% and 56%, respectively.

242 Likewise, the mortality rate was 0% among the asymptomatic/mild patients and

243 8% among moderate/severe patients, which was significantly different (OR =

244 0.07, $p < 0.015$).

245 The distribution of genotypes, alleles, and haplotypes of the three *OAS1* SNPs

246 according to severity were presented in Table 3. Among the 292 alleles analyzed,

247 the ancestral G allele represented 40% for the rs10774671 variant and 21.23% for

248 both the rs1131476 and rs1051042 variants. The comparison between the two

249 groups studied, asymptomatic/mild versus moderate/severe, shows a significant

250 difference under both the additive ($p = 0.014$) and allelic models (OR 0.528 [95%

251 CI 0.329-0.848], $p = 0.008$). However, no statistically significant differences in

Table 2: Demographic criteria of COVID-19 studied groups

| | Disease severity | | | | OR (CI 95%) | p value |
|-------------------|-------------------------------|------|-----------------------------|------|-------------------|--------------|
| | Asymptomatic/Mild (n = 71) | | Moderate/Severe (n = 75) | | | |
| | No | % | No | % | | |
| Age (y) Mean ± SD | 44.1 + 13.5 | | 51.4 + 10.5 | | | < .001 |
| Sex | | | | | | |
| Male | 39 | 54.9 | 41 | 54.7 | 1.04 (0.544-2.01) | 0.897 |
| Female | 32 | 45.1 | 34 | 45.3 | | |
| Comorbidity | 16 | 22.5 | 42 | 56.0 | 2.85 (1.39-5.85) | 0.004 |
| Mortality | 00 | 0.0 | 06 | 8.0 | 0.07 (0.004-1.35) | 0.015 |

Table 3. Comparison of genotypic and allelic distribution of the OAS1 rs10774671 and rs11314476 SNPs between COVID-19 asymptomatic/mild and Moderate/Severe groups.

| SNP | Genetic Model | Asymptomatic/Mild (N= 71) | | Moderate /Severe (N = 75) | | <i>p</i> | OR (95% CI) | |
|------------|----------------|------------------------------|-----|------------------------------|-----|----------|--------------|------------------------------|
| | | <i>N</i> | % | <i>N</i> | % | | | |
| rs10774671 | Additive model | GG | 18 | 25.3 | 06 | 08.0 | 0.014 | |
| | | GA | 32 | 45.1 | 37 | 49.3 | | |
| | | AA | 21 | 29.6 | 32 | 42.7 | | |
| | Allelic model | G | 68 | 47.9 | 49 | 32.7 | 0.008 | 0.528 (0.329 – 0.848) |
| | | A | 74 | 52.1 | 101 | 67.3 | Ref | |
| rs11314476 | Additive model | GG | 05 | 07.0 | 02 | 02.7 | 0.248 | |
| | | GA | 26 | 36.6 | 22 | 29.3 | | |
| | | AA | 40 | 56.4 | 51 | 68.0 | | |
| | Allelic model | G | 36 | 25.4 | 26 | 17.3 | 0.096 | 0.617 (0.350 – 1.090) |
| | | A | 106 | 74.6 | 124 | 82.7 | Ref | |

252 the distribution of the rs1131476 polymorphism were observed between the two
253 study groups under both the additive ($p = 0.248$), and allelic models (OR 0.617
254 [95% CI 0.350-1.090], $p = 0.096$), suggesting that this missense mutation was not
255 associated with COVID-19 severity in Moroccans.

256 The comparison of the occurrence of these polymorphisms in their respective
257 haplotype block between the two groups studied is shown in Table 4. The three
258 selected SNPs formed only three haplotypes GGG, GAC and AAC. The two
259 haplotypes GGG, and GAC were common in Asymptomatic/Mild versus
260 Moderate/Severe patients and associated compared to the risk AAC haplotype
261 with OR 0.529 [95% CI 0.285 – 0.973], $p = 0.034$ and OR 0.527 [95% CI 0.028
262 – 0.509], $p = 0.041$ respectively. These two haplotypes thus confer protection
263 against severe COVID compared to the risk haplotype AAC in patients of
264 Moroccan origin.

265
266 Tables 4. The comparison of the haplotype distribution consisting of rs10774671, rs1131462 and
267 rs1051042 between the asymptomatic/mild group versus moderate/severe group

| Haplotype | Asymptomatic/Mild (N= 71) | | Moderate /Severe (N = 75) | | <i>p</i> | OR (95% CI) |
|-----------|------------------------------|------|------------------------------|------|--------------|------------------------------|
| | <i>N</i> | % | <i>N</i> | % | | |
| GGG | 36 | 25.4 | 26 | 17.3 | 0.034 | 0.529 (0.285 – 0.973) |
| GAC | 32 | 22.5 | 23 | 15.3 | 0.041 | 0.527 (0.028 – 0.509) |
| AAC | 74 | 52.1 | 101 | 67.4 | Ref | |

268

269

270

271 **Distribution of *OAS1* haplotypes in Berbers**

272 Among the 80 males in our cohort of SARS-CoV-2-positive individuals, 41 were
273 positive and 39 were negative for the E-M81 marker. The distribution of the three
274 haplotypes studied GGG, GAC and AAC in these two groups, as well as that of
275 29 males with Parkinson's disease carrying the *LRRK2* G2019S mutation whose
276 Berber origin has been established for both paternal and maternal lineages [20]
277 were presented in Fig 3. The results showed that the prevalence of the African
278 GAC haplotype in males with the E-M81 marker represented only 8.5% and was
279 lower than that of males negative for this marker with 25.7%. Furthermore, this
280 GAC haplotype was completely absent in males whose Berber origin concerned
281 both the maternal and paternal lineages. The prevalence of both GGG and AAC
282 haplotypes in these pure Berber males is close to that of European ancestry.

283
284 Fig 3. Frequencies of the three haplotypes consisting of the three *OAS1* SNPs studied in 39
285 males negative and 41 males positive for E-M81 marker, all positive for SARS-CoV-2, and 29
286 males *LRRK2* G2019S carriers whose Berber origin has been established for both paternal and
287 maternal lineages.

288

289 **Discussion**

290 The clinical presentation of COVID-19 following SARS-CoV-2 infection has
291 shown high interindividual variability, ranging from asymptomatic to death,
292 which is partly due to genetic factors. The chr12q24.13 region containing the *OAS*
293 gene cluster was recently reported to be strongly associated with COVID-19

294 severity and mortality [4], and the rs10774671 splice variant was confirmed to be
295 the functionally critical variant of this chromosomal region that influences the
296 outcomes of SARS-CoV-2 infection [7, 8]. Indeed, the *OAS1* rs10774671-G allele
297 has been shown to increase levels of the p46 isoform and thus the circulating
298 levels of OAS1 protein that were strongly associated with reduced risks of severe
299 COVID-19 [5]. This protective G allele represents about 35% of individuals of
300 European ancestry and was found on a haplotype inherited from the Neanderthal,
301 and it represents 64% of the African population independently of this
302 Neanderthal-derived haplotype [6].

303 To our knowledge, no data currently exist on the prevalence of *OAS* gene cluster
304 variants in the North African population, which by its geographical position in the
305 African continent could give relevant information on how this G allele at
306 rs10774671 was inherited from the ancestral population common to both modern
307 humans and Neanderthals. Our study thus presents a first insight into this
308 chromosomal region in the Moroccan population, which is culturally, historically,
309 and anthropologically similar to the other neighboring countries of the North
310 African subcontinent. We showed that the frequencies of *OAS* cluster SNPs in
311 Moroccans were close to the MAF of the 1000 genome project, where the *OAS1*
312 rs10774671-G allele represented 40.4%. However, analyzing the haplotypes
313 consisting of three *OAS1* variants that captured the Neanderthal haplotype showed
314 that the rs10774671-G allele was found equally on both the Neanderthal GGG
315 haplotype and the African GAC haplotype. This distribution seems different from

316 the European population where the G allele was mostly found at 35% in the
317 Neanderthal haplotype, and from the African population where it was mostly
318 found at 64% in a specific African haplotype, as reported before [6-8].

319 The study of these 3 selected SNPs in a series of 146 SARS-CoV-2 positive
320 individuals of Moroccan origin less than 65 years of age showed that only the
321 splice variant rs10774671 presents a significant association with the severity of
322 COVID-19. Both the AA genotype and the A allele were predictive of severe
323 COVID-19 disease, supporting literature findings that this is the only variant of
324 this locus influencing COVID-19 outcomes [7, 8]. However, even though the
325 prevalence of the protective allele *OAS1* rs10774671-G in the Moroccan
326 population was close to that of Europeans, the severity and mortality caused by
327 SARS-CoV-2 infection are much higher in individuals of European ancestry, even
328 leading to a reduction in life expectancy [21]. This suggests that other genetic
329 factors as well as general health factors and conditions, all interacting in complex
330 ways, could be differentially involved between the two populations and remain to
331 be discovered.

332 Furthermore, both the Neanderthal GGG and African GAC haplotypes,
333 containing the rs10774671-G allele were shown to be protective against severe
334 COVID-19. Overall, in the 146 individuals positive for SARS-CoV-2, the
335 frequency of the G allele and of the two GGG and GAC haplotypes were
336 consistent with those obtained in the control cohort of our WES database, thus
337 confirming the equal prevalence of around 20% for each of the two Neanderthal

338 and African haplotypes in the present-day Moroccans. The presence of the
339 African haplotype could certainly be due to a recent admixture with sub-Saharan
340 African populations, while the chronology of the presence of the Neanderthal
341 haplotype in this population would be difficult to determine precisely.

342 The Moroccan population is known to be a heterogeneous population having
343 received influences from the south, north, and east, and whose ancient origin was
344 essentially indigenous Berber. To provide elements of answers to the hypothesis
345 of the origin of the Neanderthal Haplotype in North Africa, the distribution of the
346 three haplotypes was compared in our series of SARS-CoV-2 positive individuals
347 between males with and without the marker E-M81, known to be the specific male
348 lineage of autochthonous Berbers of North Africa. The frequencies of the African
349 GAC haplotype decreased to 8.5% in favor of the Neanderthal GGG haplotype
350 which increased to 29.3% in males with the E-M81 marker compared to males
351 without this marker at a frequency of 23% and 25.7% respectively. GAC
352 haplotypes in males with the E-M81 marker are likely derived from maternal
353 lineage. Surprisingly, the African GAC haplotype was completely absent in
354 Berber males from both paternal and maternal lineages, and whose frequencies of
355 the two remaining haplotypes GGG and AAC were similar to those of European
356 ancestry.

357 These results provide evidence for the presence of the Neanderthal protective
358 haplotype among Berbers who are the ancient indigenous inhabitants of North
359 Africa over the last 10,000 years. However, the question that remains to be asked

360 is whether this haplotype was imported to North Africa from the “back to Africa”
361 migration or exported to Europe and Asia by the earlier “out of Africa” one.
362 Several recent pieces of scientific evidence lean towards the second scenario.
363 First, it has been reported that present-day North Africans share a majority of their
364 ancestry with Europeans and West Asians but not with sub-Saharan Africans [22-
365 24]. Along the same lines and in a slightly earlier period, the genetic study of
366 human fossils from the Taforalt region dating back 15,000 years excluded gene
367 flow from Southern Europe into northern Africa, as well as a genetic affinity with
368 early Holocene Near Easterners, suggesting a late Pleistocene connection between
369 North Africa and the Near East [25]. However, due to the absence in this region
370 from ancient genomic data at a similar time, the authors could not define the
371 epicenter or the direction of this connection.

372 The study of the G2019S mutation of the *LRRK2* gene, a founder-effect mutation
373 found in the haplotype 1 identified worldwide but with varying prevalence, was
374 previously determined to originate from a Near-Eastern founder at least 4000
375 years ago and reintroduced by recurrent gene flow to European and North-African
376 populations [26]. Using uniparental markers, we showed that the mutation arose
377 in a Berber founder and its epicenter was indeed in North Africa, where the
378 prevalence was the highest with 40%, and the gene flow by migration took place
379 in the opposite direction [20].

380 In the same way, the site of Jebel Irhoud located in Morocco has provided valuable
381 information missing in the puzzle of the history of hominids and in particular their

382 first exit from Africa to other continents. Irhoud hominids were originally
383 considered as North African *H. sapiens* interbred with Neanderthals, an African
384 form of Neanderthals, or a North African archaic population [27]. Recently,
385 Hublin et al. [28] thermoluminescently dated the Irhoud fossils to 315,000 years
386 ago, making Jebel Irhoud the oldest African hominid site with features of
387 anatomically modern humans and more primitive cranial morphology.
388 Phylogenetic modeling of cranial morphology showed phenotypic affinities of
389 Irhoud fossils with both Neanderthals and early *H. sapiens* suggesting that they
390 have been introgressed into Neanderthals at the end of the Middle Pleistocene
391 contributing to the evolution of the Neanderthals [29]. Finally, we performed a
392 whole genome sequencing of only three present-day Moroccans and identified
393 over 200,000 SNV absent from 1000G and gnomAD databases, suggesting that
394 the Moroccan population would have more genetic variability and therefore
395 would be older than the European population. Principal components analysis
396 showed that Moroccan genomes form a distinct population and placed them
397 between European and African 1000G populations [30].
398 Sánchez-Quinto et al. [31] analyzed 780,000 SNPs in individuals from different
399 locations in North Africa including Morocco and showed that North African
400 populations have a significant excess of derived alleles shared with Neanderthals,
401 whereas Sub-Saharan populations are not affected by this admixture event. The
402 authors also showed that the Neanderthal's genetic signal was higher in
403 populations with a local, pre-Neolithic North African ancestry, confirming that

404 this detected ancient admixture is not due to recent Near Eastern or European
405 migrations. Furthermore, the first discovered Neanderthal fossil was at the Jbel
406 Irhoud site [32], where Jean-Jacques Hublin et al. [28] dated the fossils found at
407 300,000 years old, that are accompanied by a Mousterian-type lithic industry, a
408 lithic culture that has long been associated only with Neanderthals in Eurasia.
409 Lithic data (Mousterians) have been reported long before this, and the oldest lithic
410 remains discovered in Morocco belong to the Acheulean and are dated to around
411 1.3 million years [33, 34]. These data therefore constitute genomic, fossil and
412 lithic evidence of the existence of Neanderthals in Morocco and their local
413 contributions to anatomically modern humans, reinforcing the idea of a local
414 introgression of the *OAS1* rs10774671-G allele.

415 Admittedly, Africa is the cradle of humanity, but the evolution of the hominid
416 ancestors of *Homo Sapiens* did not occur in a linear way within this continent, but
417 rather, through a focus dictated by severe geographical and climatic conditions.
418 Therefore, the evolution of hominids in North Africa, isolated to the south by a
419 largely arid and desert region, would have been different from that of sub-Saharan
420 Africa, and would probably be at the origin of the initial dispersal out of Africa.
421 Their conquest of the other continents would have been possible via the Strait of
422 Gibraltar during the Ice Age, and later along the Mediterranean via the Fertile
423 Crescent. The answers to this scenario are most likely to be found in the whole
424 genome sequencing of a large series of present-day individuals and also in
425 hominid fossils from North Africa.

426 **Conclusion**

427 The rs10774671-G allele of *OAS1* confers protection against severe COVID-19
428 in the Moroccan population. Surprisingly, this protective allele was found in the
429 Neanderthal haplotype in Berbers, the indigenous people of North Africa
430 suggesting that this subcontinent could be the stepping-stone for the passage of
431 the hominids to other continents.

432

433 **Supporting information**

434 **S1 Table. Demographic, COVID-19 phenotype, genotype and haplotype of 3**
435 **OAS1 SNP of 146 individuals positive for SARS-COV-2. (xlsx)**

436 **S1 Table. Genotype of Y-M81 marker in males positive for SARS-COV-2.**
437 **(xlsx)**

438

439 **Acknowledgments**

440 We thank all the participants in this study, and 3billion® (Seoul, Korea) for
441 providing part of the whole exome sequencing free of charge.

442

443

444

445

446

447

448 **Author Contributions**

449 **Conceptualization:** Ahmed Bouhouche.

450 **Formal analysis:** Rachid Razine.

451 **Data Curation:** Omar Askander, Elmostafa El Fahime, Ahmed Bouhouche.

452 **Funding acquisition:** Ahmed Bouhouche.

453 **Investigation:** Chaimae Nebhani, Fatima-Zahra El Youssfi, Ahmed Bouhouche, Redouane

454 Abouqal, Khalid Ennibi, Hakima Fares, Abbas Ermilo Haroun, Jihane Belayachi.

455 **Methodology:** Ahmed Bouhouche.

456 **Project administration:** Ahmed Bouhouche.

457 **Supervision:** Ahmed Bouhouche.

458 **Validation:** Ahmed Bouhouche.

459 **Writing – review & editing:** Ahmed Bouhouche.

460

461 **Data availability statement**

462 All relevant data are within the manuscript and its Supporting Information files.

463

464 **References**

465 1. Harapan H, Itoh N, Yufika A, Winardi W, Keam S, Te H, et al. Coronavirus disease 2019
466 (COVID-19): A literature review. *J Infect Pub Health*. 2020; 13:667-673.

467 <https://doi.org/10.1016/j.jiph.2020.03.019>

468 2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients
469 with pneumonia in China, 2019. *New Engl J Med*. 2020; 382:727–733.

470 <https://doi.org/10.1056/nejmoa2001017>

471 3. Ranjan J, Ravindra A, Mishra B. Gender and genetic factors impacting COVID-19
472 severity. *J Family Med Prim Care*. 2021; 10:3956-3963.

473 https://doi.org/10.4103/jfmpe.jfmpe_769_21

- 474 4. Pairo-Castineira E, Clohisey, S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic
475 mechanisms of critical illness in COVID-19. *Nature*. 2020; 591:92-98.
476 <https://doi.org/10.1038/s41586-020-03065-y>
- 477 5. Zhou S, Hill CS, Sarkar S, Tse LV, Woodburn BMD, Schinazi RF, et al. β -d-N 4-
478 hydroxycytidine inhibits SARS-CoV-2 through lethal mutagenesis but is also mutagenic
479 to mammalian cells. *J Infect Dis*. 2021; 224:415-419.
480 <https://doi.org/10.1093/infdis/jiab247>
- 481 6. Zeberg H, Pääbo S. A genomic region associated with protection against severe COVID-
482 19 is inherited from Neandertals. *Proc Natl AcaD Sci USA*. 2021; 118:e2026309118.
483 <https://doi.org/10.1073/pnas.2026309118>
- 484 7. Banday AR, Stanifer ML, Florez-Vargas O, Onabajo OO, Papenberg BW, Zahoor M et al.
485 Genetic regulation of OAS1 nonsense-mediated decay underlies association with COVID-
486 19 hospitalization in patients of European and African ancestries. *Nat genet*. 2022;
487 54:1103-1116. <https://doi.org/10.1038/s41588-022-01113-z>
- 488 8. Huffman JE, Butler-Laporte G, Khan A, Pairo-Castineira E, Drivas TG, Peloso GM, et al.
489 Multi-ancestry fine mapping implicates OAS1 splicing in risk of severe COVID-19. *Nat*
490 *genet*. 2022; 54:125-127. <https://doi.org/10.1038/s41588-021-00996-8>
- 491 9. Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM, Johnson B, et al. Genetic variation
492 in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS pathogens*.
493 2009. 5:e1000321. <https://doi.org/10.1371/journal.ppat.1000321>
- 494 10. O'Brien M, Lonergan R, Costelloe L, O'Rourke K, Fletcher JM, Kinsella K, et al. OAS1:
495 a multiple sclerosis susceptibility gene that influences disease severity. *Neurology*. 2010;
496 75:411-418. <https://doi.org/10.1212/wnl.0b013e3181ebdd2b>
- 497 11. Mandal S, Abebe F, Chaudhary J. 2'-5' oligoadenylate synthetase 1 polymorphism is
498 associated with prostate cancer. *Cancer*. 2011; 117:5509-5518.
499 <https://doi.org/10.1002/cncr.26219>
- 500 12. Sams AJ, Dumaine A, Nédélec Y, Yotova V, Alfieri C, Tanner JE, et al. Adaptively
501 introgressed Neandertal haplotype at the OAS locus functionally impacts innate immune
502 responses in humans. *Genome Biol*. 2016; 17:246. <https://doi.org/10.1186/s13059-016-1098-6>
- 503
504 13. Liu X, Xing H, Gao W, Yu D, Zhao Y, Shi X, et al. A functional variant in the OAS1 gene
505 is associated with Sjögren's syndrome complicated with HBV infection. *Sci Rep*. 2017;
506 7:17571. <https://doi.org/10.1038/s41598-017-17931-9>

- 507 **14.** Wu S, Wang Y, Chen G, Zhang M, Wang M, He JQ. 2'-5'-Oligoadenylate synthetase 1
508 polymorphisms are associated with tuberculosis: a case-control study. *BMC Pulm Med.*
509 2018; 18:180. <https://doi.org/10.1186/s12890-018-0746-x>
- 510 **15.** Choi UY, Kang JS, Hwang YS, Kim YJ. Oligoadenylate synthase-like (OASL) proteins:
511 dual functions and associations with diseases. *Exp Mol Med.* 2015; 47:e144.
512 <https://doi.org/10.1038/emm.2014.110>
- 513 **16.** Wickenhagen A, Sugrue E, Lytras S, Kuchi S, Noerenberg M, Turnbull ML, et al. A
514 prenylated dsRNA sensor protects against severe COVID-19. *Science.* 2021;
515 374:eabj3624. <https://doi.org/10.1126/science.abj3624>
- 516 **17.** Kjær KH, Pahus J, Hansen MF, Poulsen JB, Christensen EI, Justesen J, et al. Mitochondrial
517 localization of the OAS1 p46 isoform associated with a common single nucleotide
518 polymorphism. *BMC cell biol.* 2014; 15:33. <https://doi.org/10.1186/1471-2121-15-33>
- 519 **18.** Zhou Q, Chen V, Shannon CP, Wei XS, Xiang X, Wang X, et al. Interferon- α 2b Treatment
520 for COVID-19. *Front Immunol.* 2020; 11:1061. <https://doi.org/10.3389/fimmu.2020.01061>
- 521 **19.** Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from
522 Neanderthals. *Nature.* 2020; 587:610–612. <https://doi.org/10.1038/s41586-020-2818-3>
- 523 **20.** Ben El Haj R, Salmi A, Regragui W, Moussa A, Bouslam N, Tibar H, et al. Evidence for
524 prehistoric origins of the G2019S mutation in the North African Berber population. *PLoS*
525 *One.* 2017; 12:e0181335. <https://doi.org/10.1371/journal.pone.0181335>
- 526 **21.** Schöley J, Aburto JM, Kashnitsky I, Kniffka MS., Zhang L, Jaadla H, et al. Life expectancy
527 changes since COVID-19. *Nat Hum Behav.* 2022; 6(12):1649-1659.
528 <https://doi.org/10.1038/s41562-022-01450-3>
- 529 **22.** Botigué LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, et al. Gene flow
530 from North Africa contributes to differential human genetic diversity in southern Europe.
531 *Proc Natl Acad Sci USA.* 2013; 110:11791-11796. <https://doi.org/10.1073/pnas.1306223110>
- 532 **23.** Schuenemann VJ, Peltzer A, Welte B, van Pelt WP, Molak M, Wang CC, et al. Ancient
533 Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-
534 Roman periods. *Nat Commun.* 2017; 8:15694. <https://doi.org/10.1038/ncomms15694>
- 535 **24.** Pakstis AJ, Gurkan C, Dogan M, Balkaya HE, Dogan S, Neophytou PI, et al. Genetic
536 relationships of European, Mediterranean, and SW Asian populations using a panel of 55
537 AISNPs. *Eur J Hum Genet.* 2019; 27:1885-1893. <https://doi.org/10.1038/s41431-019-0466-6>
- 538 **25.** van de Loosdrecht M, Bouzouggar A, Humphrey L, Posth C, Barton N, Aximu-Petri A, et
539 al. Pleistocene North African genomes link near Eastern and sub-Saharan African human
540 populations. *Science.* 2018; 360, 548–552. <https://doi.org/10.1126/science.aar8380>

- 541 **26.** Lesage S, Patin E, Condroyer C, Leutenegger AL, Lohmann E, Giladi N, et al. Parkinson's
542 disease-related LRRK2 G2019S mutation results from independent mutational events in
543 humans. *Hum Mol Genet.* 2010; 19:1998-2004. <https://doi.org/10.1093/hmg/ddq081>
- 544 **27.** Bruner E, Pearson O. Neurocranial evolution in modern humans: the case of Jebel Irhoud
545 1. *Anthropol Sci.* 2013; 121:31-41. <https://doi.org/10.1537/ase.120927>
- 546 **28.** Hublin JJ, Ben-Ncer A, Bailey SE, Freidline SE, Neubauer , Skinner MM, et al. New fossils
547 from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature.* 2017;
548 546:289–292. <https://doi.org/10.1038/nature22336>
- 549 **29.** Mounier A, Mirazón LM. Deciphering African late middle Pleistocene hominin diversity
550 and the origin of our species. *Nat Commun.* 2019; 10:3406.
551 <https://doi.org/10.1038/s41467-019-11213-w>
- 552 **30.** Crooks L, Cooper-Knock J, Heath PR, Bouhouche A, Elfahime M, Azzouz M, et al.
553 Identification of single nucleotide variants in the Moroccan population by whole-genome
554 sequencing. *BMC Genet.* 2020; 21:111. <https://doi.org/10.1186/s12863-020-00917-4>
- 555 **31.** Sánchez-Quinto F, Botigué LR, Civit S, Arenas C, Avila-Arcos MC, Bustamante CD, et
556 al. North African populations carry the signature of admixture with Neandertals. *PLoS*
557 *One.* 2012; 7:e47765. <https://doi.10.1371/journal.pone.0047765>
- 558 **32.** Arambourg C. Découverte au Maroc d'un crâne de Néanderthalien. *Bulletin de la Société*
559 *Préhistorique Française.* 1962; 59-7-8:513-514.
- 560 **33.** Mohib A, Raynal JP, Gallotti R, Daujeard C, El Graoui M, Fernandes P, et al. Forty Years
561 of Research at Casablanca (Morocco): New Insights in the Early/Middle Pleistocene
562 Archaeology and Geology. *Hespéris-Tamuda, Eléments d'Archéologie préhistorique*
563 *marocaine.* 2019; LIV (3):25-56. hal-03067997
- 564 **34.** Gallotti R, Muttoni G, Lefèvre D, Degeai JP, Geraads D, Zerboni A, et al. First high
565 resolution chronostratigraphy for the early North African Acheulean at Casablanca
566 (Morocco). *Sci Rep.* 2021; 11:15340. <https://doi.10.1038/s41598-021-94695-3>

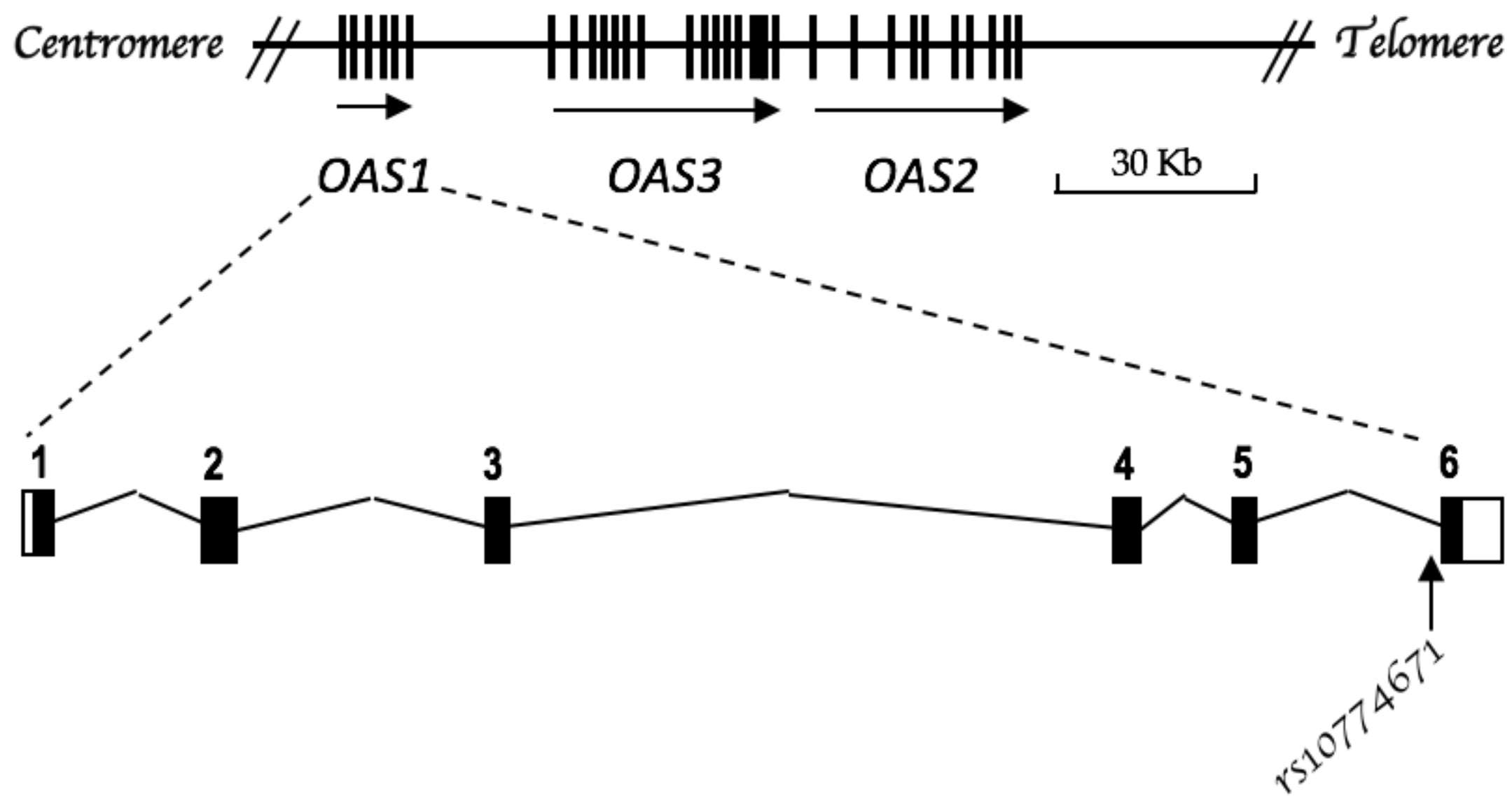


Figure 1

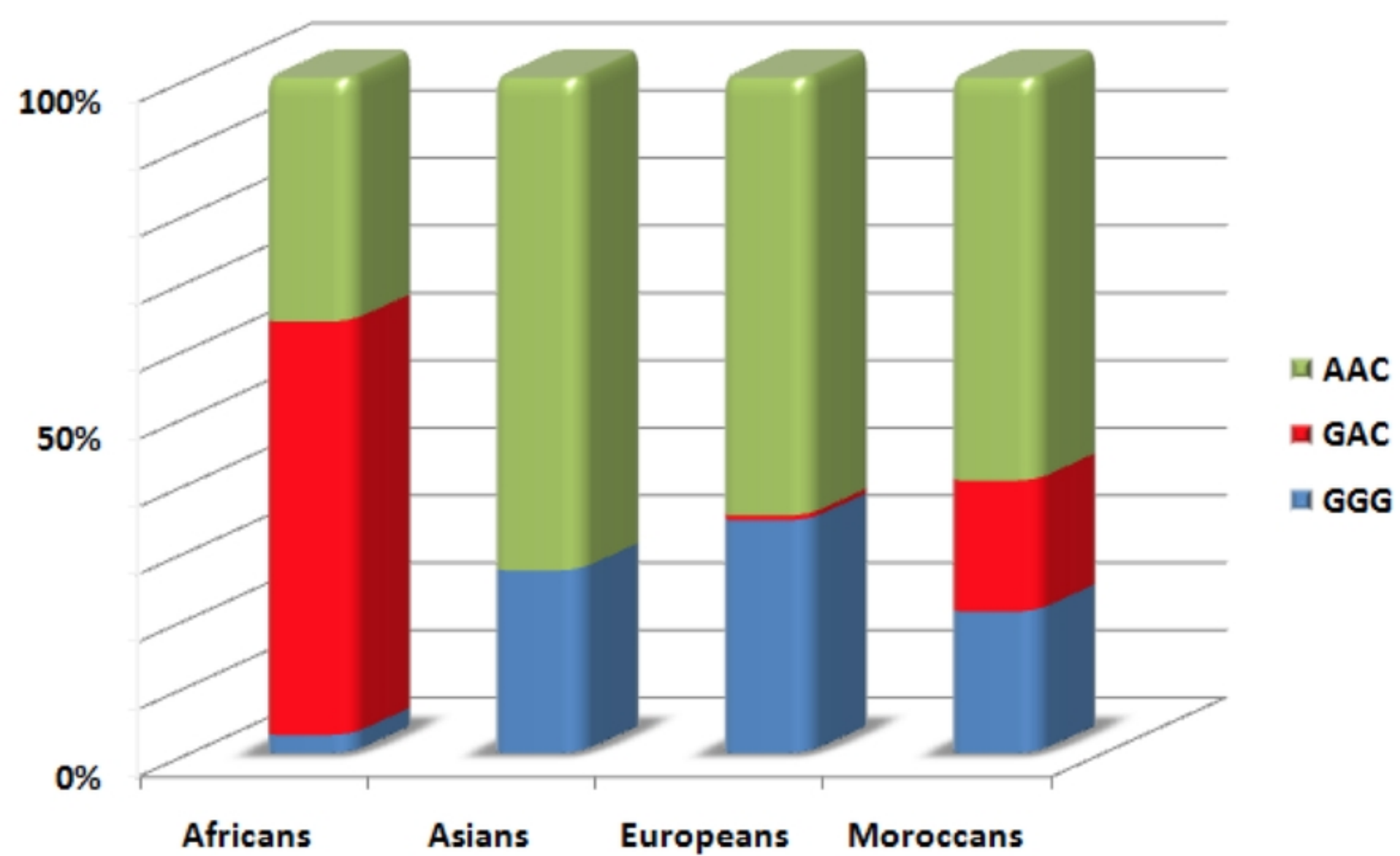


Figure 2

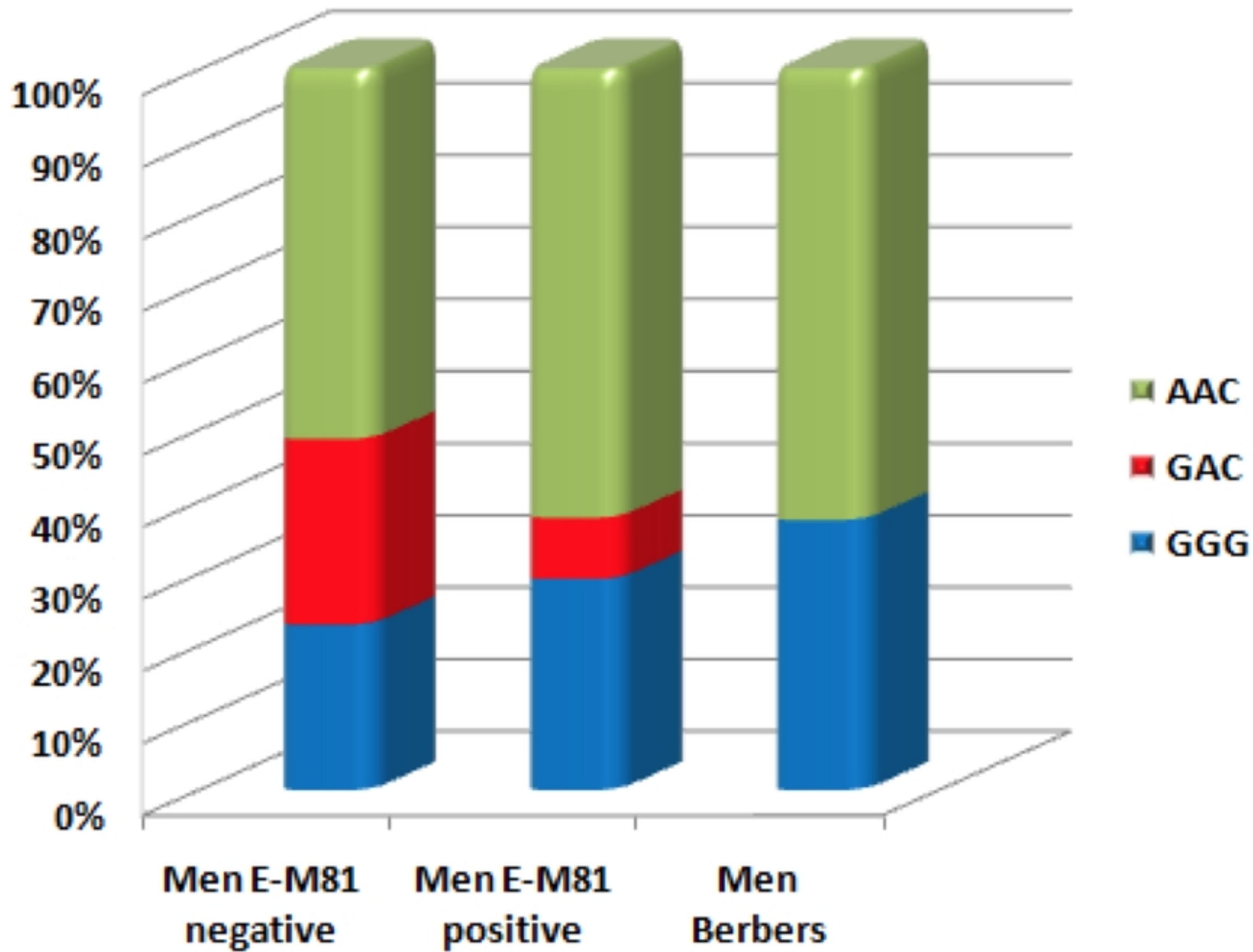


Figure 3