Manuscript bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission. Click here to download Manuscript Ahouty et al. Revised Article.doc

1 **Candidate genes-based investigation of susceptibility to Human African**
2 **Trypanosomiasis in Côte d'Ivoire.**

3 Bernardin Ahouty[1], Mathurin Koffi[2*], Hamidou Ilboudo[3], Gustave Simo[4], Enock Matovu[5],
4 Julius Mulindwa[5], Christiane Hertz-Fowler[6], Bruno Bucheton[7], Issa Sidibé[3], Vincent
5 Jamonneau[7-8], Annette MacLeod[9], Harry Noyes[6], Simon-Pierre N'Guetta[1], for the
6 TrypanoGEN Research Group as members of The H3Africa Consortium.

7 [1]Felix Houphouet Boigny University (UFHB), Cocody, Abidjan, Côte d'Ivoire; [2]Jean Lorougnon Guédé
8 University (UJLoG), Daloa, Côte d'Ivoire;[3]Centre International de Recherche-Développement sur l'Elevage en
9 zone Subhumide (CIRDES), Bobo-Dioulasso, Burkina Faso; [4]University of Dchang, Cameroon, [5]Makerere
10 University, Kampala, Uganda,[6]Centre for Genomic Research, University of Liverpool, UK; [7]Institut de
11 Recherche pour le Développement (IRD), IRD-CIRAD 177, Montpellier, France; [8]Institut Pierre Richet, Bouaké,
12 Côte d'Ivoire, [9]Wellcome Center for Molecular Parasitology, University Place, Glasgow, UK.

13 **Corresponding Author*
14 Email: bernadinahouty@yahoo.fr (BA)
15 *Email: m9koffi@yahoo.fr (MK)
16 Email: hamidou_ilboudo@hotmail.com (HI)
17 Email:gsimoca@yahoo.fr (GS)
18 Email: matovue04@yahoo.com (EM)
19 Email : mujuls@gmail.com (JM)
20 Email :c.Hertz-Fowler@liverpool.ac.uk (CH)
21 Email: bruno.bucheton@ird.fr (BB)
22 Email: isidibe@hotmail.com (IS)
23 Email :vincent.jamonneau@ird.fr (VJ)
24 Email:annette.macleod@glasgow.ac.uk (AM)
25 Email: harry@liverpool.ac.uk (HN)
26 Email: nguettaewatty@yahoo.fr (SPN)

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

27 **Abstract**

28 Human African Trypanosomiasis (HAT) or sleeping sickness is a Neglected Tropical Disease.
29 Long regarded as an invariably fatal disease, there is increasing evidence that infection by *T.*
30 *b. gambiense* can result in a wide range of clinical outcomes, including latent infections,
31 which are long lasting infections with no parasites detectable by microscopy. The
32 determinants of this clinical diversity are not well understood but could be due in part to
33 parasite or host genetic diversity in multiple genes, or their interactions. A candidate gene
34 association study was conducted in Côte d'Ivoire using a case-control design which included
35 a total of 233 subjects (100 active HAT cases, 100 controls and 33 latent infections). All three
36 possible pairwise comparisons between the three phenotypes were tested using 96 SNPs in16
37 candidate genes (*IL1, IL4, IL4R, IL6, IL8, IL10, IL12, IL12R, TNFA, INFG, MIF, APOL1,*
38 *HPR, CFH, HLA-A* and *HLA-G*). Data from 77 SNPs passed quality control. There were
39 suggestive associations at three loci in *IL6* and *TNFA* in the comparison between active cases
40 and controls, one SNP in each of *APOL1*, *MIF* and *IL6* in the comparison between latent
41 infections and active cases and seven SNP in *IL4, HLA-G* and *TNFA* between latent infections
42 and controls. No associations remained significant after Bonferroni correction, but the
43 Benjamini Hochberg false discovery rate test indicated that there were strong probabilities
44 that at least some of the associations were genuine.
45 The excess of associations with latent infections despite the small number of samples
46 available suggests that these subjects form a distinct genetic cluster different from active HAT
47 cases and controls, although no clustering by phenotype was observed by principle
48 component analysis. This underlines the complexity of the interactions existing between host
49 genetic polymorphisms and parasite diversity.

50 **Keywords:** Single nucleotide polymorphism; Candidate genes; Association study; Human
51 African trypanosomiasis; Sleeping sickness, Cote d'Ivoire

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

52 **Author summary**

53 Since it was first identified, human African trypanosomiasis (HAT) or sleeping sickness has
54 been described as invariably fatal. Recent data however suggest that infection by *T. b.*
55 *gambiense* can result in a wide range of clinical outcomes in its human host including long
56 lasting infections, that can be detected by the presence of antibodies, but in which parasites
57 cannot be seen by microscopy; these cases are known as latent infections. While the factors
58 determining this varied response have not been clearly characterized, the effectors of the
59 immune responses have been partially implicated as key players. We collected samples from
60 people with active HAT, latent infections and controls in endemic foci in the Côte d'Ivoire.
61 We tested the role of single nucleotide polymorphisms (SNPs) in 16 genes on
62 susceptibility/resistance to HAT by means of a candidate gene association study. There was
63 some evidence that variants of the genes for *IL4, IL6, APOL1, HLAG, MIF* and *TNFA*
64 modified the risk of developing HAT. These proteins regulate the inflammatory response to
65 many infections or are directly involved in killing the parasites. In this study, the results were
66 statistically weak and would be inconclusive on their own, however other studies have also
67 found associations in these genes, increasing the chance that the variants that we have
68 identified play a genuine role in the response to trypanosome infection in Côte D'Ivoire.

69 **Introduction**

70 Sleeping sickness, or human African Trypanosomiasis (HAT), is caused by *Trypanosoma*
71 *brucei gambiense* and *T.b. rhodesiense* and is transmitted by tsetse flies (*Glossina spp*). *T. b.*
72 *gambiense* is associated with a more chronic disease that can take decades to become patent,
73 and *T.b. rhodesiense* causes an acute disease within months of infection. The chronic form of
74 the disease caused by the *T. b. gambiense* is classically characterized by an early
75 hemolymphatic stage (stage 1) associated with non-specific symptoms such as intermittent
76 fevers and headaches, followed by a meningo-encephalitic stage (stage 2) in which the

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

77 parasite invades the central nervous system and causes neurological disorders and death if left
78 untreated. This chronic form is found in Western and Central Africa while the acute form
79 caused by *T. b. rhodesiense* is endemic to Eastern Africa [1,2]. However, recent observations
80 are increasingly indicating that infection by *T. b. gambiense* can result in a wide range of
81 clinical outcomes in its human host [3,4]. Self-cure processes have been described and
82 reviewed in Checchi *et al.* [3]. Furthermore, some authors argued that HAT is not invariably
83 fatal [5], supporting observations made by Garcia et *al.* [4] who followed individuals who
84 remained seropositive for HAT but without detectable parasites for two years. This clinical
85 diversity is not well understood but could be due to parasite genetic diversity [6,7], human
86 immune gene variability [8] or their interaction [9]. It has been suggested that genetic
87 polymorphisms of the parasite could be associated with asymptomatic and very chronic
88 infections [10]. Nevertheless, genes involved in the host immune response have been
89 implicated in the control of infection or susceptibility to HAT [11,12] and also *T. congolense*
90 infections in experimental models [13,14]. Among the genes implicated in the pathogenesis of
91 the disease, are *IL1,IL4, IL6, IL8, IL10, IL12, IFNG, TNFA*, *CFH* and *MIF* [8,11,15-17].

92 Some studies have found associations between certain polymorphisms in genes encoding
93 cytokines. For example, *IL4* plays a role in susceptibility to *T. brucei* infection [18], while
94 polymorphisms in the *IL6* and *IL10* genes have been associated with a decreased risk of
95 developing HAT [11,15]. On the other hand, polymorphisms in *TNFA* genes have been
96 associated with an increased risk of developing the disease [11,15]. Furthermore, MacLean et
97 *al.* [19] reported in their study that plasma levels of *IFNG* significantly decrease during the
98 late stage of the *T.b. rhodesiense* disease [19]. Human blood contains trypanolytic factors
99 (TLF1 and TLF2) that are lytic to almost all African trypanosomes except *T. b. rhodesiense*
100 and *T. b. gambiense* [20]. Human TLF1 contains two primate-specific proteins,
101 apolipoprotein L1 (APOL1) and haptoglobin-related protein (HPR). *APOL1* expression is

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

102 induced by *T. b. gambiense* infection but expression is not associated with susceptibility to
103 sleeping sickness [21]. Macrophage migration inhibitory factor (MIF) contributes to
104 inflammation-associated pathology in the chronic phase of *T. brucei* infection [16].
105 Although genes directly involved in the immune response, like genes encoding cytokines, are
106 very important candidates, genes implicated in the regulation of immunity also have a critical
107 role. Thus, Courtin et *al.* [12] have shown a genetic association between *HLA-G*
108 polymorphisms and susceptibility to HAT.
109 There is now cumulative evidence that polymorphisms in genes involved in the control of the
110 immune response and genes implicated in the regulation of immunity could play a role in
111 HAT infection outcome [11,12].
112 We report a candidate gene association study of the role of single nucleotide polymorphisms
113 (SNPs) in *IL1*, *IL4, IL4R, IL6, IL8, IL10, IL12, IL12R, TNFA, INFG, MIF, HPR, CFH,*
114 *APOL1*, *HLA-A,* and *HLA-G* genes on susceptibility/resistance to HAT.

115

116 **Results**

117 **Characteristics of the population**

118 A total of 100 cases (or former cases), 100 controls and 33 latent infections were enrolled into
119 the study (Table 1). The mean age (range) of the study population was 38.8 (6–84) years. The
120 sex ratio (male: female) was 0.88 (109/124).

121

122

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

123

124 **Table1: Population characteristics**

125

126

127

	Cases (n=100)	Controls (n=100)	Latent infections (n= 33)	Whole population (n=233)
Age mean	41,8	36,8	36	38,8
Males	53	42	14	109
Females	47	58	19	124
TL+	65	0	18	83
TL-	35	100	15	150

128 129 130 131 132 133 134 135 136 137

138 **Table 1:** Population characteristics. Cases: CATT+ve, microscopy +ve; Controls: CATT-ve
139 and microscopy –ve. Trypanolysis negative (TL-ve) cases were from subjects who were
140 retrospectively sampled for the purpose of this study and who had not been subject to the
141 trypanolysis test when originally identified as HAT cases by CATT and microscopy. Latent
142 infection: subjects who were CATT positive (CATTpl ≥1/4) but parasites could not be
143 detected by microscopy at repeated tests for at least two years.

144

145 **Association study**

146 Three association analyses were run on the three possible pairwise combinations of the
147 controls, latent infections and active HAT cases. After filtering out SNP loci that were not in
148 Hardy Weinberg equilibrium (P <0.001), or had > 10% missing genotypes or had minor allele
149 frequencies of zero, 77 SNP loci remained in the latent infection vs active case comparison
150 and 74 SNP loci remained in the other two comparisons. One individual was filtered out of
151 the case sample and one out of the latent infection sample because of low genotyping rates
152 (<90%).

153 F_{ST} between ethnic groups showed that allele frequency differences between the ethnic groups
154 were small, indicating that ethnicity was unlikely to confound results (median Fst -

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

155 0.00011763, Fst maximum 0.015). Scatter plots of the first two principal components between
156 the different phenotypes (cases, latent infections and controls) and different ethnic groups
157 show that this population is homogeneous and that the samples did not cluster by phenotype
158 or by ethnic groups (Fig 1) and consequently the data were not stratified by ethnicity.

159 **Fig 1.** Multidimensional scaling (ms) plots of the genotype data by ethnicity (upper plot) and
160 phenotype (lower plot). The plots show no evidence of clustering either by ethnicity or
161 phenotype.

162 There were suggestive associations at three loci in *IL6* and *TNFA* in the comparison between
163 active cases and controls (Table 2), three loci in *APOL1*, *MIF* and *IL6* in the comparison
164 between latent infections and active cases (Table 3) and five loci in *IL4* (Fig 2), and one each
165 in *HLA-G* and *TNFA* between latent infections and controls (Table 4). After Bonferroni
166 correction, none of these associations remained significant, however Bonferroni correction is
167 very conservative, particularly since there was some linkage between adjacent marker SNP in
168 some genes. The Benjamini-Hochberg false discovery rate (FDR) test, which shows the
169 probability that an observation is a false positive, indicated that at least some of the
170 suggestively positive samples may be true positives. Under the FDR the rs62449495 SNP in
171 *IL6* had an 80% chance of being a true positive and the *TNFA*-308 rs1800629 SNP had a 71%
172 chance of being a true positive (Table 2). The strongest association was with *APOL1*
173 rs73885319, which is also known as the G1 allele, which had a 90% chance of being a true
174 positive in the comparison between latent infections and controls (Table 3). Complete results
175 for all tests that passed quality control are shown in supplementary data tables S1, S2 and S3.
176

177 **Fig 2.** Positions of SNP genotyped within IL4. SNP with suggestive associations (uncorrected
178 p <0.05) are indicated with a star.
179

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

180 **Table 2**: Association tests between cases (n=99) and controls (n=100)

CHR	BP	SNP	Gene	Cons	Min A	Cases	Controls	Maj A	P	Bonf_Corr	FDR_BH	OR	L95	U95	HWE_p
7	22,764,338	rs62449495	IL6	5'	A	0.04	0.13	G	0.0025	0.20	0.20	0.30	0.13	0.68	1.00
6	31,543,031	rs1800629	TNFA	5'	A	0.09	0.19	G	0.0071	0.58	0.29	0.44	0.24	0.80	0.51
7	22,767,137	rs2069830	IL6	Missense	T	0.11	0.06	C	0.0369	1.00	0.81	2.17	1.02	4.61	1.00

181

182 **Table 3**: Association tests between latent infections (n=32) and active HAT cases (n=99)

CHR	BP	SNP	Gene	Cons	Min A	Cases	Latent	Maj A	P	Bonf_Corr	FDR_BH	OR	L95	U95	HWE_p
22	36,661,906	rs73885319	APOL1	Missense	G	0.20	0.41	A	0.0012	0.10	0.10	0.37	0.20	0.68	1.00
22	24,235,455	rs36086171	MIF	5 ‘	G	0.29	0.13	A	0.0062	0.52	0.26	2.90	1.30	6.46	0.05
7	22,764,338	rs62449495	IL6	5’	A	0.04	0.13	G	0.0217	1.00	0.61	0.29	0.10	0.80	1.00

183

184 **Table 4**: Association tests between latent infections (n=32) and controls (n=100)

CHR	BP	SNP	Gene	Cons	Min A	Latent	Controls	Maj A	P	Bonf_Corr	FDR_BH	OR	L95	U95	HWE_p
5	132,009,154	rs2243250	IL4	5’	C	0.42	0.27	T	0.0159	1.00	0.48	2.02	1.13	3.64	0.80
5	132,010,726	rs734244	IL4	Intronic	T	0.25	0.41	C	0.0215	1.00	0.48	0.49	0.26	0.92	0.01
5	132,016,554	rs2243282	IL4	Intronic	A	0.30	0.45	C	0.0347	1.00	0.48	0.53	0.29	0.96	1.00
6	29,799,116	rs1611139	HLAG	3’	T	0.26	0.14	G	0.0354	1.00	0.48	2.13	1.04	4.36	0.39
5	132,009,710	rs2070874	IL4	5’ UTR	T	0.34	0.50	C	0.0372	1.00	0.48	0.53	0.30	0.96	0.69
6	31,543,031	rs1800629	TNFA	5’	A	0.08	0.19	G	0.0396	1.00	0.48	0.37	0.14	0.99	0.51
5	132,014,109	rs2243270	IL4	Intronic	A	0.36	0.23	G	0.0415	1.00	0.48	1.88	1.02	3.45	1.00

185

186 Cons= Consequence from Ensembl Variant predictor, Min A = minor allele, Maj A = major allele, Latent minor allele frequency of Latent cases;
187 Controls minor allele frequency of controls; OR = odds ratio, P = p-value, [L95-U95] = confidence interval of odds ratio, Bonf_Corr= Bonferroni
188 corrected p value, FDR_BH = False Discovery Rate Benjamini-Hochberg (It is probability of falsely rejecting the null hypothesis. The null
189 hypothesis is that allele frequencies in cases and controls are the same), OR = odds ratio, [L95-U95] = confidence interval; HWE P is the Hardy-
190 Weinberg p-value.

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

191 **DISCUSSION**

192 In this study, we investigated the role of single nucleotide polymorphisms on
193 susceptibility/resistance to HAT. We genotyped 96 SNPs within sixteen genes that were
194 investigated to test genetic association with HAT in Côte d'Ivoire. For consistency, all
195 samples were commercially genotyped using two platforms: Genome Transcriptome de
196 Bordeaux, France and LGC Genomics UK.

197 Our results suggest that three SNPs were associated with the development of HAT by
198 controls, another three were associated with the development of active HAT by people with
199 latent infections, and seven SNP were associated with the risk of developing latent infections
200 five of them in *IL4* although none of these results remained significant after Bonferroni
201 correction. Only two of the SNP with suggestive associations were coding and missense (Il6
202 rs2069830 and APOL1 rs73885319) since, with some exceptions, SNP were primarily selected
203 as tags that might be linked to functional SNP rather than for any putative function.

204 The most striking feature of the results is that most associations were found in the
205 comparisons with latent infections, despite limited power, with only 33 of these samples
206 being available compared with 100 each for the cases and controls. This illustrates the
207 increased power to be gained from using well defined phenotypes. Given the evidence that
208 some people in West Africa can self-cure their *T. b. gambiense* infections [5] it is likely that
209 some of the control subjects are resistant to infection or have recovered from infection, whilst
210 others are naïve and susceptible. These two groups may have very different genetic
211 backgrounds which could confound the association studies. The data also suggest that the
212 latent infections may represent a genetically distinctive group. Although the principle
213 component analysis did not identify any cluster associated with latent infections (Fig 1)
214 further studies of their genetic and immunological profiles are required to test this hypothesis.

215 **Associations with *IL6***

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

216 The minor (A) allele of *IL6* rs62449495 appeared to protect against progression from latent
217 infections to active HAT (Table 3) and against the development of active HAT by controls
218 but these associations did not remain significant after Bonferroni correction (Table 2). The
219 major allele of rs2069830 was also protective against the development of active HAT by
220 controls before but not after Bonferroni correction (Table 2). The rs2069830 SNP causes a
221 proline to serine change, but this is predicted to be benign by both Sift and Polyphen
222 according to the Ensembl Variant Effect predictor. IL6 plays a key role in the acute
223 inflammatory response and in regulation of the production of acute phase proteins such as C-
224 reactive protein [22]. It contributes to the inflammatory response, regulates haematopoiesis,
225 which may contribute to the anaemia associated with HAT, and modifies the permeability of
226 the blood brain barrier, which may contribute to the development of the stage 2 invasion of
227 the CNS [22,23]. The results of our association study are consistent with the involvement of
228 *IL6* variants in the susceptibility to the disease as Courtin *et al* [15] have reported before.
229 However, it should be noted that the polymorphisms of *IL6* in our study (rs62449495 and
230 rs2069830) are different of those found by Courtin *et al*. [15] with $IL6_{4339}$= rs2069849. They
231 showed that in DRC, $IL6_{4339}$ SNP was significantly associated with a decreased risk of
232 developing the disease with a *P*-value = 0.0006 before the Bonferroni correction (0.04 after
233 Bonferroni Correction).

234 **Association with *TNFA***

235 Our results also indicate that subjects carrying the A allele of *TNFA* (rs1800629) A/G had a
236 lower risk of developing active HAT (Table 2) or a latent infection (Table 4), suggesting the
237 possibility of a protective effect. This SNP is also known as the *TNF*-308 SNP and the minor
238 A allele is associated with higher plasma levels of *TNFA* [24]. In a previous study in Côte
239 d'Ivoire, Courtin et *al*. [11] showed that the distribution of the *TNFA*-308G/A polymorphism
240 did not differ significantly between cases and controls in the total population, but found that,

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

241 under a recessive model the AA genotype was associated with risk of HAT in the 39 cases
242 and 57 controls who had been living in the endemic area of Sinfra for less than 10 years
243 (before Bonferroni correction). In contrast, in our study, all participants had lived in the
244 endemic area all their lives and the association was additive rather than recessive, furthermore
245 the MAF in Courtin's studies was 24%, whilst in ours it was 14%, suggesting significant
246 differences in population structure between the two studies. Varying results in *TNFA*
247 associations studies are not uncommon, multiple candidate gene association studies of *TNFA*-
248 308 (rs1800629) and malaria have found inconsistent results in different populations [25,26].
249 Another explanation could be that the *TNFA*-308 has no effect but is in linkage disequilibrium
250 (LD) with another unidentified polymorphism. Varying LD across populations might lead to
251 different findings [27]. Despite the conflicting results from association study, animal models
252 indicate that *TNFA* is likely to be a key mediator in the control of *T. brucei* infections [28]
253 and a direct dose dependent lytic effect of *TNFA* on purified *T.b. gambiense* parasites has
254 been reported suggesting an involvement in parasite growth control [29,30]. Given the
255 experimental evidence for a role for TNFA, it is possible that inconsistent results of
256 association studies are a consequence of different *TNFA* alleles only making a small
257 difference to infection outcome and that therefore larger studies would be needed to detect
258 associations and also to heterogeneity in regulation of *TNFA* between populations.
259

260 **Association with *IL4***

261 Five of the seven SNP that were suggestively associated with the comparison between latent
262 infections and controls were in *IL4* (Table 4, Fig 2). Three of these SNP were adjacent to each
263 other at the 5' end of the gene (Fig 2), although linkage between them was modest (r^2 <0.5).
264 While the individual associations did not remain significant after Bonferroni correction the
265 observation that five of the sixteen SNP tested in *IL4* had suggestive associations increases

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

266 the probability of a genuine association with *IL4* and the risk of developing a latent infection.
267 A study using three *IL4* SNP in DRC did not find any associations, but that was in a
268 comparison between cases and controls [15]. In B10.Q mice deletion in *IL4* lead to increased
269 *T. brucei* parasitaemia levels but longer survival time [18]. The relationship between *IL4*
270 polymorphisms and acute HAT infection requires further study.

271 **Association with *APOL1***

272 The suggestive association with *APOL1* G1 allele rs73885319 and protection against
273 progression from latent infection to active HAT is consistent with the association found in
274 Guinea [31,32], although in those studies this SNP was also associated with increased risk of
275 controls developing a latent infection as well. Moreover, *APOL1* expression is induced by *T.*
276 *b. gambiense* infection but not associated with differential susceptibility to sleeping sickness
277 [21]. *APOL1* rs73885319 is also known as the G1 allele of *APOL1* and is associated with
278 kidney disease in African Americans and the relatively high frequency of this deleterious
279 allele was assumed to be due to selection by HAT [33]. The data presented here is consistent
280 with that hypothesis, but no support was found for the role of the G2 allele of *APOL1*
281 (rs71785313) in HAT although it is implicated in kidney disease.

282 **Association with *MIF***

283 Although, our data show that subjects carrying the G allele of *MIF* (rs36086171) G/A had a
284 risk of developing active HAT (table 3), we did not find a significant difference after
285 correction (BONF = 0.52). Macrophage inhibitory factor (MIF) is a ubiquitously expressed
286 protein that has proinflammatory, hormonal and enzymatic activities [34]. It is implicated in
287 many inflammatory diseases [35]. It functions by recruiting myeloid cells to the site of
288 inflammation [36], by inducing their differentiation towards M1 cells secreting TNF [37] and
289 by suppressing p53-dependent apoptosis of inflammatory cells. While there are no human

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

290 studies directly linking *MIF* to HAT, murine studies show that *MIF* plays a role as a mediator
291 of the inflammation which is a key feature in trypanosomiasis-associated pathology [16,38].
292

293 **Association with *HLA-G***

294 *HLA-G* molecule plays an important role on immune response regulation and has been
295 associated with the risk of human immunodeficiency virus (HIV) infection [39], human
296 papilloma virus [40] and herpes simplex virus type 1 [41]. The results of our association study
297 are consistent with the involvement of HLA-G genetic variants in the susceptibility to the
298 disease [12]. The rs1611139 T allele of the *HLA-G* gene showed a suggestive association with
299 increased risk of controls a latent infection. This result could be due to differences in selection
300 pressures possibly driven by variability in the immuno-pathology of the diseases. It is known
301 that cytokines such as IL-10 can induce HLA-G expression by affecting mRNA transcripts
302 and protein synthesis by human monocytes and trophoblasts, thus having a significant impact
303 on parasitic infections [42].

304 **Conclusion**

305 The results discussed in this paper should be used with caution as no loci remained
306 significantly associated with HAT after the Bonferroni correction for multiple testing,
307 although some had high probabilities of being associated using a false discovery rate test.
308 Some of the SNP loci identified here were also significant in other studies. Multiple
309 independent observations of marginally significant effects suggest that these effects may be
310 genuine but that the effect size is not large enough to be detected by the numbers available to
311 be tested.

312 Our data support the findings from Guinea about the role of the *APOL1* G1 allele and also
313 suggest that the polymorphisms of *IL4, IL6, HLA-G* and *TNFA* present interesting candidates
314 for the investigation of the genetic susceptibility / resistance to HAT.

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

315 The large number of suggestive associations with latent infections despite the small number
316 of samples indicate the people with latent infections form a genetically distinctive group that
317 merit further investigation.
318

319 **Material and Methods**

320 **Population study, definition of phenotypes and study design**

321 The study took place in western-central Côte d'Ivoire in the main HAT foci of Bonon,
322 Bouafle, Zoukougbeu, Oume, and Sinfra. Samples were collected in three stages, (i)
323 previously archived samples; existing collections of previously archived samples were
324 centralized to CIRDES (Centre International de Recherche-Développement sur l'Elevage en
325 zone Subhumide) Burkina-Faso. (ii) retrospectively collected samples; study sites were
326 revisited to consent and resample previously diagnosed and treated patients and (iii)
327 prospectively collected samples including new HAT patients. Three phenotypes were
328 considered: (i) cases, defined as individuals in whom the presence of trypanosomes was
329 confirmed by microscopy; (ii) controls, defined as individuals living in the endemic area who
330 are card agglutination test for trypanosomiasis (CATT) negative and trypanolysis test (TL)
331 negative, and without evidence of previous HAT infection and (iii) latent infections, defined
332 as subjects who were CATT positive (CATTpl $\geq$1/4) but in whom parasites could not be
333 detected by microscopy at repeated tests for at least two years.

334 A total of 233 individuals with 100 former HAT cases, 100 controls and 33 latent infections
335 were included in the study. Samples were frozen directly in the field at -20°C and kept at that
336 temperature until use. For each individual, an aliquot of plasma was used to perform the
337 immune trypanolysis test that detects Litat 1.3 and Litat 1.5 variable surface antigens specific
338 for *T.b. gambiense* [43]. All control individuals (n=100) included in this study were negative
339 on the trypanolysis test (TL-ve) (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

340 This study was one of five studies of populations of HAT endemic areas in Cameroon, Côte
341 d'Ivoire, Guinea, Malawi and Uganda by the TrypanoGEN consortium [44]. The studies were
342 designed to have 80% power to detect odds ratios (OR) >2 for loci with disease allele
343 frequencies of 0.15 – 0.65 and 100 cases and 100 controls with the 96 SNPs genotyped.
344 Power calculations were undertaken using the genetics analysis package gap in R [45].

345 **Selection of SNPs and genotyping**

346 Genomic DNA was obtained from peripheral blood samples. Extraction was performed using
347 the Qiagen DNA extraction kit (QIAamp DNA Blood Midi Kit) according to the
348 manufacturer's instructions and quantified by Nanodrop assay. DNA was stored at -20 °C
349 until analysis.

350 Ninety-six SNPs were genotyped in *IL1*, *IL4*, *IL4R*, *IL6*, *IL8*, *IL10*, *IL12*, *IL12R*, *TNFA*,
351 *INFG*, *MIF*, *HPR*, *CFH*, *APOL1*, *HLA-A*, and *HLA-G*. The SNPs were selected by two
352 strategies: 1) SNP in *IL4*, *IL6*, *IL8*, *HLAG* and *IFNG* were designed as markers for linkage
353 disequilibrium (LD) scans of each gene [46], 2) SNPs in other genes were selected based on
354 reports in the literature that they were associated with trypanosomiasis or other infectious
355 diseases. The LD scans were designed using 1000 Genomes Project data [47] ,merged with
356 low fold coverage (8-10x) whole genome shotgun data generated from 230 residents living in
357 regions (DRC, Guinea Conakry, Côte D'Ivoire and Uganda) where trypanosomiasis is
358 endemic (TrypanoGEN consortium, European Nucleotide Archive study
359 EGAS00001002482). Loci with minor allele frequency < 5% in the reference data were
360 excluded and an r^2 of 0.5 was used to select SNP in linkage disequilibrium.

361 DNA was genotyped by two commercial service providers: INRA- Site de Pierroton,
362 Plateforme Genome Transcriptome de Bordeaux, France and ii- LGC genomics Hoddesden
363 UK. At INRA multiplex design (two sets of 40 SNPs) was performed using Assay Design
364 Suite v2.0 (Agena Biosciences). SNP genotyping was achieved with the iPLEX Gold

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

365 genotyping kit (Agena Biosciences) for the Mass Array iPLEX genotyping assay, following
366 the manufacturer's instructions. Products were detected on a Mass Array mass
367 spectrophotometer and data were acquired in real time with Mass Array RT software (Agena
368 Biosciences). SNP clustering and validation was carried out with Typer 4.0 software (Agena
369 Biosciences). At LGC Genomics SNP were genotyped using the PCR based KASP assay[48].

370 **Statistical analysis**

371 All statistical analyses were performed using the Plink 1.9 and R v 3.2.1 software. Individuals
372 were excluded who had > 10% missing SNP data. SNP loci were excluded that > 10%
373 missing genotypes or if the control samples were not in Hardy-Weinberg equilibrium
374 ($p<0.001$). Case-control association analysis using SNP alleles/genotypes was undertaken
375 using Fisher's exact test. The difference between the ethnic groups was estimated using the
376 F_{ST}, with values potentially varying from zero (no population differentiation) to one
377 (complete differentiation) [47]. Ethnicity was not used as a risk factor for HAT because we
378 observed no significant differences in F_{ST} between ethnic groups (F_{ST} maximum 0.015) or
379 clustering by linguistic group by multidimensional scaling in Plink. Some studies have used
380 ethnicity and age as risk factors for HAT but found no significant association [11,15]. 77 SNP
381 remained in at least one comparison after filtering. P values were adjusted for multiple testing
382 using the Bonferroni corrections and Benjamini Hocheberg false discovery rate test as
383 implemented in Plink.

384 **Ethics Statement**

385 The population of the study was informed. All adult subjects provided written informed
386 consent. For children, a parent or guardian of any child participant provided written informed
387 consent on their behalf. The protocol of the study was approved by the traditional authorities
388 (chief and village committee) and by National ethics committee hosted by the Public Health

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

389 Ministry of Côte d'Ivoire with the number: N°38/MSLS/CNERm-dkn 5th May 2014. This
390 study is part of the TrypanoGen project which aims to a better understanding of genetic
391 determinism of human susceptibility to HAT and the TrypanoGen-CI samples were archived
392 in the TrypanoGen Biobank hosted by CIRDES in Bobo-Dioulasso, Burkina Faso [44].

393

394 **Acknowledgements:** We thank subjects who generously donated their specimens and the
395 field workers from the Ivorian HAT foci for their dedication in collecting and processing
396 these specimens.

397

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

398

399 **References**

400

401 1. Simarro PP, Diarra A, Ruiz Postigo JA, Franco JR, Jannin JG (2011) The human African
402 trypanosomiasis control and surveillance programme of the World Health Organization 2000-
403 2009: the way forward. PLoS Negl Trop Dis 5: e1007.
404 2. WHO (2013) Control and surveillance of human African trypanosomiasis. WHO Technical Report
405 Series 984 Geneva.
406 3. Checchi F, Filipe JA, Barrett MP, Chandramohan D (2008) The natural progression of *Gambiense*
407 sleeping sickness: what is the evidence? PLoS Negl Trop Dis 2: e303.
408 4. Garcia A, Jamonneau V, Magnus E, Laveissiere C, Lejon V, et al. (2000) Follow-up of Card
409 Agglutination Trypanosomiasis Test (CATT) positive but apparently aparasitaemic individuals
410 in Cote d'Ivoire: evidence for a complex and heterogeneous population. Trop Med Int Health
411 5: 786-793.
412 5. Jamonneau V, Ilboudo H, Kabore J, Kaba D, Koffi M, et al. (2012) Untreated human infections by
413 Trypanosoma brucei gambiense are not 100% fatal. PLoS Negl Trop Dis 6: e1691.
414 6. Kabore J, Macleod A, Jamonneau V, Ilboudo H, Duffy C, et al. (2011) Population genetic structure
415 of Guinea Trypanosoma brucei gambiense isolates according to host factors. Infect Genet
416 Evol 11: 1129-1135.
417 7. Koffi M, De Meeus T, Bucheton B, Solano P, Camara M, et al. (2009) Population genetics of
418 *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. Proc Natl
419 Acad Sci U S A 106: 209-214.
420 8. Ilboudo H, Bras-Goncalves R, Camara M, Flori L, Camara O, et al. (2014) Unravelling human
421 trypanotolerance: IL8 is associated with infection control whereas IL10 and TNFalpha are
422 associated with subsequent disease development. PLoS Pathog 10: e1004469.
423 9. Bucheton B, MacLeod A, Jamonneau V (2011) Human host determinants influencing the outcome
424 of Trypanosoma brucei gambiense infections. Parasite Immunol 33: 438-447.
425 10. Jamonneau V, Ravel S, Garcia A, Koffi M, Truc P, et al. (2004) Characterization of *Trypanosoma*
426 *brucei s.l.* infecting asymptomatic sleeping-sickness patients in Cote d'Ivoire: a new genetic
427 group? Ann Trop Med Parasitol 98: 329-337.
428 11. Courtin D, Argiro L, Jamonneau V, N'Dri L, N'Guessan P, et al. (2006) Interest of tumor necrosis
429 factor-alpha -308 G/A and interleukin-10 -592 C/A polymorphisms in human African
430 trypanosomiasis. Infect Genet Evol 6: 123-129.
431 12. Courtin D, Milet J, Sabbagh A, Massaro JD, Castelli EC, et al. (2013) HLA-G 3' UTR-2 haplotype is
432 associated with Human African trypanosomiasis susceptibility. Infect Genet Evol 17: 1-7.
433 13. Iraqi F, Clapcott SJ, Kumari P, Haley CS, Kemp SJ, et al. (2000) Fine mapping of trypanosomiasis
434 resistance loci in murine advanced intercross lines. Mamm Genome 11: 645-648.
435 14. Kemp SJ, Iraqi F, Darvasi A, Soller M, Teale AJ (1997) Localization of genes controlling resistance
436 to trypanosomiasis in mice. Nat Genet 16: 194-196.
437 15. Courtin D, Milet J, Jamonneau V, Yeminanga CS, Kumeso VK, et al. (2007) Association between
438 human African trypanosomiasis and the IL6 gene in a Congolese population. Infect Genet
439 Evol 7: 60-68.
440 16. Stijlemans B, Leng L, Brys L, Sparkes A, Vansintjan L, et al. (2014) MIF contributes to Trypanosoma
441 brucei associated immunopathogenicity development. PLoS Pathog 10: e1004414.
442 17. Tiberti N, Hainard A, Lejon V, Robin X, Ngoyi DM, et al. (2010) Discovery and verification of
443 osteopontin and Beta-2-microglobulin as promising markers for staging human African
444 trypanosomiasis. Mol Cell Proteomics 9: 2783-2795.

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

445 18. Bakhiet M, Jansson L, Buscher P, Holmdahl R, Kristensson K, et al. (1996) Control of parasitemia
446 and survival during Trypanosoma brucei brucei infection is related to strain-dependent ability
447 to produce IL-4. J Immunol 157: 3518-3526.
448 19. MacLean L, Chisi JE, Odiit M, Gibson WC, Ferris V, et al. (2004) Severity of human african
449 trypanosomiasis in East Africa is associated with geographic location, parasite genotype, and
450 host inflammatory cytokine response profile. Infect Immun 72: 7040-7044.
451 20. Pays E, Vanhollebeke B, Vanhamme L, Paturiaux-Hanocq F, Nolan DP, et al. (2006) The
452 trypanolytic factor of human serum. Nat Rev Microbiol 4: 477-486.
453 21. Ilboudo H, Berthier D, Camara M, Camara O, Kabore J, et al. (2012) APOL1 expression is induced
454 by Trypanosoma brucei gambiense infection but is not associated with differential
455 susceptibility to sleeping sickness. Infect Genet Evol 12: 1519-1523.
456 22. Yamada Y, Metoki N, Yoshida H, Satoh K, Ichihara S, et al. (2006) Genetic risk for ischemic and
457 hemorrhagic stroke. Arterioscler Thromb Vasc Biol 26: 1920-1925.
458 23. Raj DS (2009) Role of interleukin-6 in the anemia of chronic disease. Semin Arthritis Rheum 38:
459 382-388.
460 24. Perera MK, Herath NP, Pathirana SL, Phone-Kyaw M, Alles HK, et al. (2013) Association of high
461 plasma TNF-alpha levels and TNF-alpha/IL-10 ratios with TNF2 allele in severe P. falciparum
462 malaria patients in Sri Lanka. Pathog Glob Health 107: 21-29.
463 25. Gichohi-Wainaina WN, Melse-Boonstra A, Swinkels DW, Zimmermann MB, Feskens EJ, et al.
464 (2015) Common Variants and Haplotypes in the TF, TNF-alpha, and TMPRSS6 Genes Are
465 Associated with Iron Status in a Female Black South African Population. J Nutr 145: 945-953.
466 26. Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, et al. (1999) A polymorphism that affects
467 OCT-1 binding to the TNF promoter region is associated with severe malaria. Nat Genet 22:
468 145-150.
469 27. Tu IP, Whittemore AS (1999) Power of association and linkage tests when the disease alleles are
470 unobserved. Am J Hum Genet 64: 641-649.
471 28. Magez S, Caljon G (2011) Mouse models for pathogenic African trypanosomes: unravelling the
472 immunology of host-parasite-vector interactions. Parasite Immunol 33: 423-429.
473 29. Daulouede S, Bouteille B, Moynet D, De Baetselier P, Courtois P, et al. (2001) Human macrophage
474 tumor necrosis factor (TNF)-alpha production induced by Trypanosoma brucei gambiense
475 and the role of TNF-alpha in parasite control. J Infect Dis 183: 988-991.
476 30. Magez S, Radwanska M, Beschin A, Sekikawa K, De Baetselier P (1999) Tumor necrosis factor
477 alpha is a key mediator in the regulation of experimental *Trypanosoma brucei* infections.
478 Infect Immun 67: 3128-3132.
479 31. Cooper A, Ilboudo H, Alibu VP, Ravel S, Enyaru J, et al. (2017) APOL1 renal risk variants have
480 contrasting resistance and susceptibility associations with African trypanosomiasis. Elife 6.
481 32. Kabore WJ, Ilboudo H, Noyes H, Camara O, Kabore J, et al. (2017) Candidate gene study between
482 human African trypanosomiasis clinical phenotypes in Guinea. PLoS Negl Trop Dis in press.
483 33. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, et al. (2010) Association of
484 trypanolytic ApoL1 variants with kidney disease in African Americans. Science 329: 841-845.
485 34. Barton A, Lamb R, Symmons D, Silman A, Thomson W, et al. (2003) Macrophage migration
486 inhibitory factor (MIF) gene polymorphism is associated with susceptibility to but not
487 severity of inflammatory polyarthritis. Genes Immun 4: 487-491.
488 35. Rosado Jde D, Rodriguez-Sosa M (2011) Macrophage migration inhibitory factor (MIF): a key
489 player in protozoan infections. Int J Biol Sci 7: 1239-1256.
490 36. Bernhagen J, Krohn R, Lue H, Gregory JL, Zernecke A, et al. (2007) MIF is a noncognate ligand of
491 CXC chemokine receptors in inflammatory and atherogenic cell recruitment. Nat Med 13:
492 587-596.
493 37. Calandra T, Bernhagen J, Metz CN, Spiegel LA, Bacher M, et al. (1995) MIF as a glucocorticoid-
494 induced modulator of cytokine production. Nature 377: 68-71.

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

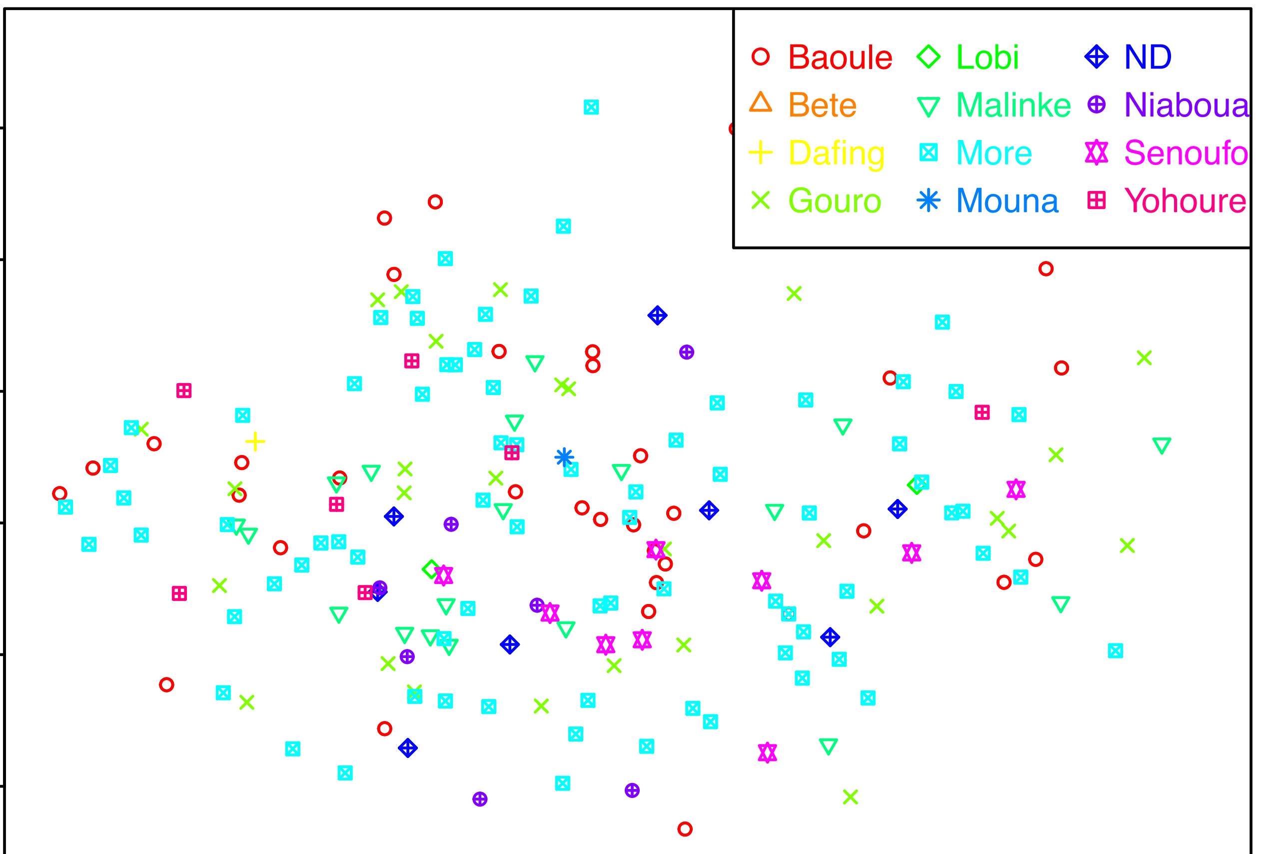
495 38. Stijlemans B, Beschin A, Magez S, Van Ginderachter JA, De Baetselier P (2015) Iron Homeostasis
496 and Trypanosoma brucei Associated Immunopathogenicity Development: A Battle/Quest for
497 Iron. Biomed Res Int 2015: 819389.
498 39. Lajoie J, Hargrove J, Zijenah LS, Humphrey JH, Ward BJ, et al. (2006) Genetic variants in
499 nonclassical major histocompatibility complex class I human leukocyte antigen (HLA)-E and
500 HLA-G molecules are associated with susceptibility to heterosexual acquisition of HIV-1. J
501 Infect Dis 193: 298-301.
502 40. Ferguson R, Ramanakumar AV, Richardson H, Tellier PP, Coutlee F, et al. (2011) Human leukocyte
503 antigen (HLA)-E and HLA-G polymorphisms in human papillomavirus infection susceptibility
504 and persistence. Hum Immunol 72: 337-341.
505 41. Lafon M, Prehaud C, Megret F, Lafage M, Mouillot G, et al. (2005) Modulation of HLA-G
506 expression in human neural cells after neurotropic viral infections. J Virol 79: 15226-15237.
507 42. Zhao M, Zhang R, Xu X, Liu Y, Zhang H, et al. (2013) IL-10 reduces levels of apoptosis in
508 Toxoplasma gondii-infected trophoblasts. PLoS One 8: e56455.
509 43. Jamonneau V, Bucheton B, Kabore J, Ilboudo H, Camara O, et al. (2010) Revisiting the immune
510 trypanolysis test to optimise epidemiological surveillance and control of sleeping sickness in
511 West Africa. PLoS Negl Trop Dis 4: e917.
512 44. Ilboudo H, Noyes H, Mulindwa J, Kimuda MP, Koffi M, et al. (2017) Introducing the TrypanoGEN
513 biobank: A valuable resource for the elimination of human African trypanosomiasis. PLoS
514 Negl Trop Dis 11: e0005438.
515 45. Zhao JH (2007) Gap: genetic analysis package. J Stat Soft 23: 1-18.
516 46. Gray IC, Campbell DA, Spurr NK (2000) Single nucleotide polymorphisms as tools in human
517 genetics. Hum Mol Genet 9: 2403-2408.
518 47. Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure.
519 Evol 38: 1358-1370.
520 48. Semagn k, Babu R, Hearn S, Olsen MS (2013) Single nucleotide polymorphism genotyping using
521 Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop
522 improvement. Mol Breed 33: 1-14.

523

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Click here to download Figure Figure 1
Ahouty et al, candidate genes



bioRxiv preprint doi: https://doi.org/10.1101/186718; this version posted September 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Click here to download Figure Figure 2 Abouty et al. Candidate genes study_IL4.pdf

Figure 2