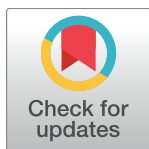RESEARCH ARTICLE

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

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

1 Laboratoire de Génétique, Félix Houphouët Boigny University, Abidjan, Côte d'Ivoire, **2** Unité de Recherche en Génétique et Epidémiology Moléculaire, Jean Lorougnon Guédé University, Daloa, Côte d'Ivoire, **3** Unité Maladies à Vecteurs et Biodiversité, Centre International de Recherche-Développement sur l'Elevage en zone Subhumide, Bobo-Dioulasso, Burkina Faso, **4** Department of Biochemistry, University of Dchang, Dchang, Cameroon, **5** School of Veterinary Medicine, Makerere University, Kampala, Uganda, **6** Centre for Genomic Research, University of Liverpool, Liverpool, United Kingdom, **7** Unité Mixte de Recherche 177 IRD-CIRAD, Institut de Recherche pour le Développement, Montpellier, France, **8** Unité de Recherche Glossines et Trypanosomes, Institut Pierre Richet, Bouaké, Côte d'Ivoire, **9** Wellcome Center for Molecular Parasitology, University of Glasgow, Glasgow, United Kingdom

¶ Membership of The TrypanoGEN Research Group is provided in the Acknowledgments.
* m9koffi@yahoo.fr

Abstract

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

complexity of the interactions existing between host genetic polymorphisms and parasite diversity.

Author summary

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

Introduction

Sleeping sickness, or human African Trypanosomiasis (HAT), is caused by *Trypanosoma brucei gambiense* and *T.b. rhodesiense* and is transmitted by tsetse flies (*Glossina spp*). *T. b. gambiense* is associated with a more chronic disease that can take decades to become patent, and *T. b. rhodesiense* causes an acute disease within months of infection. The chronic form of the disease caused by the *T. b. gambiense* is classically characterized by an early hemolymphatic stage (stage 1) associated with non-specific symptoms such as intermittent fevers and headaches, followed by a meningo-encephalitic stage (stage 2) in which the parasite invades the central nervous system and causes neurological disorders and death if left untreated. This chronic form is found in Western and Central Africa while the acute form caused by *T. b. rhodesiense* is endemic to Eastern Africa [1,2]. However, recent observations are increasingly indicating that infection by *T. b. gambiense* can result in a wide range of clinical outcomes in its human host [3,4]. Self-cure processes have been described and reviewed in Checchi *et al.* [3]. Furthermore, some authors argued that HAT is not invariably fatal [5], supporting observations made by Garcia *et al.* [4] who followed individuals who remained seropositive for HAT but without detectable parasites for two years. This clinical diversity is not well understood but could be due to parasite genetic diversity [6,7], human immune gene variability [8] or their interaction [9]. It has been suggested that genetic polymorphisms of the parasite could be associated with asymptomatic and very chronic infections [10]. Nevertheless, genes involved in the host immune response have been implicated in the control of infection or susceptibility to HAT [11,12] and also *T. congolense* infections in experimental models [13,14]. Among the genes implicated in the pathogenesis of the disease, are interleukins 1, 4, 6, 8, 10 and 12 (*IL1*, *IL4*,

IL6, IL8, IL10 and *IL12*), tumor necrosis factor A (*TNFA*), interferon G (*INFG*), complement factor H (*CFH*), macrophage inhibitory factor (*MIF*), [8,11,15–17].

Some studies have found associations between certain polymorphisms in genes encoding cytokines. For example, *IL4* plays a role in susceptibility to *T. brucei* infection [18], while polymorphisms in the *IL6* and *IL10* genes have been associated with a decreased risk of developing HAT [11,15]. On the other hand, polymorphisms in *TNFA* genes have been associated with an increased risk of developing the disease [11,15]. Furthermore, MacLean *et al.* [19] reported in their study that plasma levels of *IFNG* significantly decrease during the late stage of the *T.b. rhodesiense* disease [19]. Human blood contains trypanolytic factors 1 and 2 (TLF1 and TLF2) that are lytic to almost all African trypanosomes except *T. b. rhodesiense* and *T. b. gambiense* [20]. Human TLF1 contains two primate-specific proteins, apolipoprotein L1 (APOL1) and haptoglobin-related protein (HPR). *APOL1* expression is induced by *T. b. gambiense* infection but expression is not associated with susceptibility to sleeping sickness [21]. MIF contributes to inflammation-associated pathology in the chronic phase of *T. brucei* infection [16].

Although genes directly involved in the immune response, like genes encoding cytokines, are very important candidates, genes implicated in the regulation of immunity also have a critical role. Thus, Courtin *et al.* [12] have shown a genetic association between human leukocyte antigen G (*HLA-G*) polymorphisms and susceptibility to HAT.

There is now cumulative evidence that polymorphisms in genes involved in the control of the immune response and genes implicated in the regulation of immunity could play a role in HAT infection outcome [11,12].

We report a candidate gene association study of the role of single nucleotide polymorphisms (SNP) in *IL1, IL4, IL6, IL8, IL10, IL12*, interleukin 4 receptor (*IL4R*), interleukin 12 receptor (*IL12R*), *TNF, INFG, MIF*, haptoglobin-related protein (*HPR*), *CFH, APOL1, HLA-A* and *HLA-G* genes on susceptibility/resistance to HAT.

Results

Characteristics of the population

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

Association study

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

F_{ST} between ethnic groups showed that allele frequency differences between the ethnic groups were small, indicating that ethnicity was unlikely to confound results (median Fst -0.00011763, Fst maximum 0.015). Scatter plots of the first two principal components between the different phenotypes (cases, latent infections and controls) and different ethnic groups show that this population is homogeneous and that the samples did not cluster by phenotype or by ethnic groups (Fig 1) and consequently the data were not stratified by ethnicity.

There were suggestive associations at three loci in *IL6* and *TNFA* in the comparison between active cases and controls (Table 2), three loci in *APOL1, MIF* and *IL6* in the

Table 1. Population characteristics.

	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+ve	65	0	18	83
TL-ve	35	100	15	150

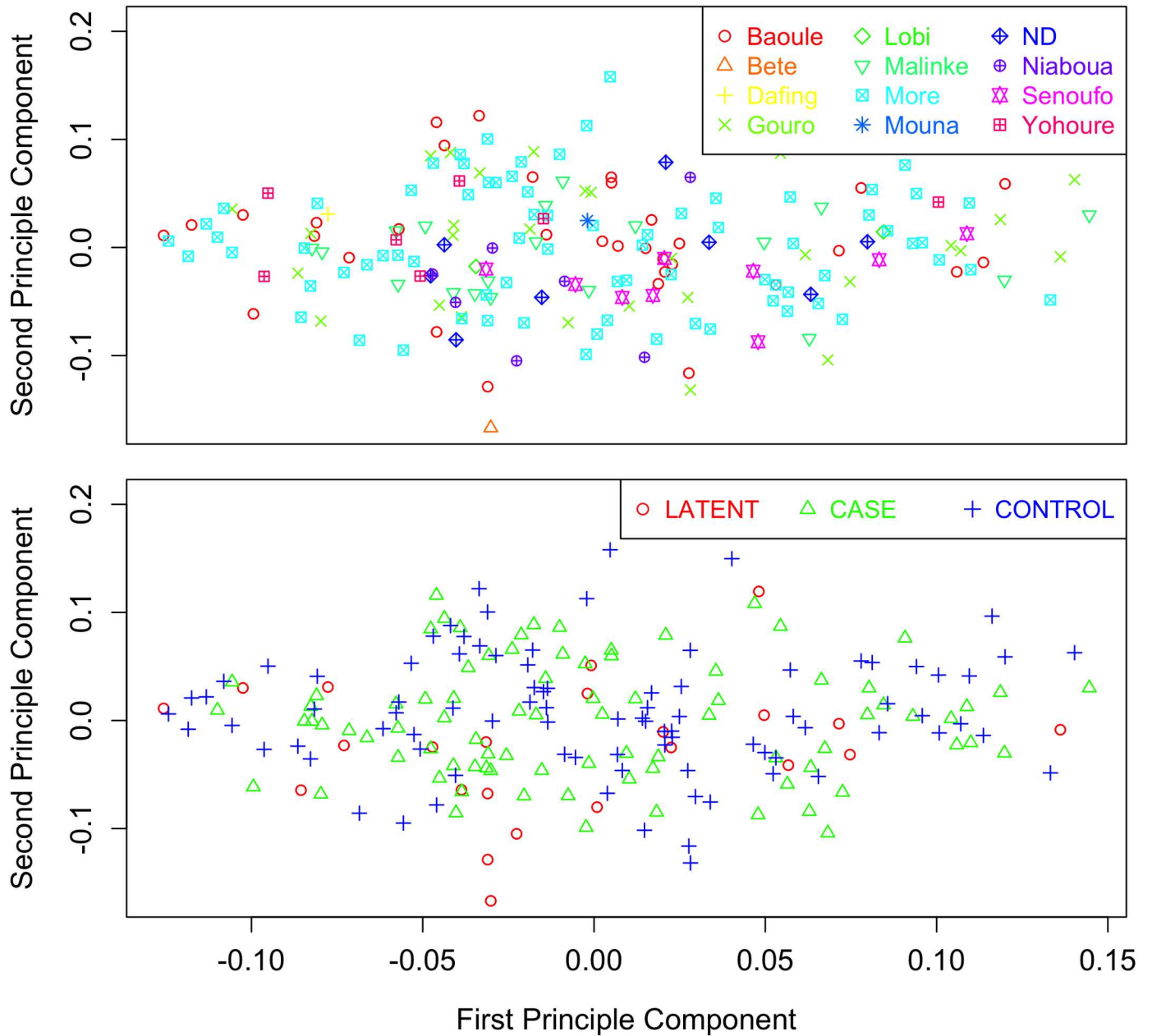https://doi.org/10.1371/journal.pntd.0005992.t001

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

https://doi.org/10.1371/journal.pntd.0005992.g001

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

https://doi.org/10.1371/journal.pntd.0005992.t002

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

Discussion

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

Our results suggest that three SNPs were associated with the development of HAT by controls, another three were associated with the development of active HAT by people with latent infections, and seven SNP were associated with the risk of developing latent infections five of them in *IL4* although none of these results remained significant after Bonferroni correction. Only two of the SNP with suggestive associations were coding and missense (IL6 rs2069830 and APOL1 rs73885319) since, with some exceptions, SNP were primarily selected as tags that might be linked to functional SNP rather than for any putative function.

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

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

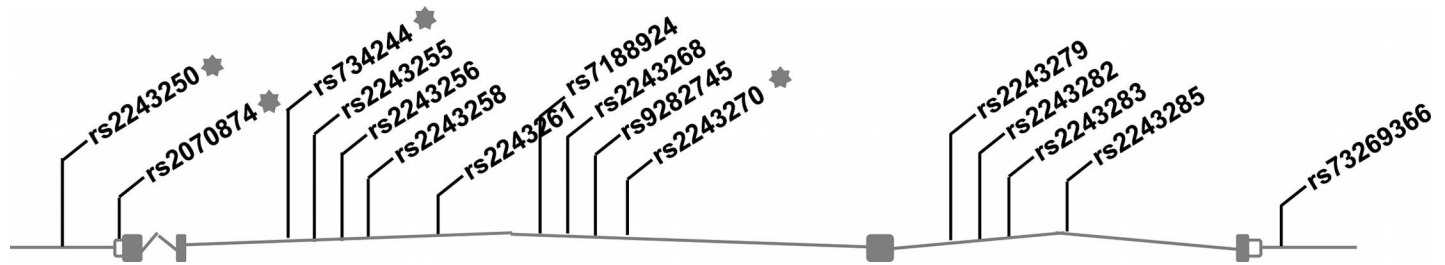| CHR	BP	SNP	Gene	Cons	Min A	Cases	Latent	Maj A	P	Bonf_Corr	FDR_BH	OR	L95	U95	HWE_p
22	36,661,906	rs73885319	APOL1	Missense	G	0.20	0.41	A	0.0012	0.10	0.10	0.37	0.20	0.68	1.00
22	24,235,455	rs36086171	MIF	5 '	G	0.29	0.13	A	0.0062	0.52	0.26	2.90	1.30	6.46	0.05
7	22,764,338	rs62449495	IL6	5'	A	0.04	0.13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

https://doi.org/10.1371/journal.pntd.0005992.t003

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

cluster associated with latent infections (Fig 1) further studies of their genetic and immunological profiles are required to test this hypothesis.

Associations with IL6

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

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

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

Association with *TNFA*

Our results also indicate that subjects carrying the A allele of *TNFA* (rs1800629) A/G had a lower risk of developing active HAT (Table 2) or a latent infection (Table 4), suggesting the possibility of a protective effect. This SNP is also known as the *TNF*-308 SNP and the minor A allele is associated with higher plasma levels of *TNFA* [24]. In a previous study in Côte d'Ivoire, Courtin et *al.* [11] showed that the distribution of the *TNFA*-308G/A polymorphism did not differ significantly between cases and controls in the total population, but found that, under a recessive model the AA genotype was associated with risk of HAT in the 39 cases and 57 controls who had been living in the endemic area of Sinfra for less than 10 years (before Bonferroni correction). In contrast, in our study, all participants had lived in the endemic area all their lives and the association was additive rather than recessive, furthermore the MAF in Courtin's studies was 24%, whilst in ours it was 14%, suggesting significant differences in population structure between the two studies. Varying results in *TNFA* associations studies are not uncommon, multiple candidate gene association studies of *TNFA*-308 (rs1800629) and malaria have found inconsistent results in different populations [25,26]. Another explanation could be that the *TNFA*-308 has no effect but is in linkage disequilibrium (LD) with another unidentified polymorphism. Varying LD across populations might lead to different findings [27]. Despite the conflicting results from association study, animal models indicate that *TNFA* is likely to be a key mediator in the control of *T. brucei* infections [28] and a direct dose dependent lytic effect of *TNFA* on purified *T.b. gambiense* parasites has been reported suggesting an involvement in parasite growth control [29,30]. Given the experimental evidence for a role for *TNFA*, it is possible that inconsistent results of association studies are a consequence of different *TNFA* alleles only making a small difference to infection outcome and that therefore larger studies would be needed to detect associations and also to heterogeneity in regulation of *TNFA* between populations.

Association with *IL4*

Five of the seven SNP that were suggestively associated with the comparison between latent infections and controls were in *IL4* (Table 4, Fig 2). Three of these SNP were adjacent to each other at the 5' end of the gene (Fig 2), although linkage between them was modest (r^2 <0.5). While the individual associations did not remain significant after Bonferroni correction the observation that five of the sixteen SNP tested in *IL4* had suggestive associations increases the probability of a genuine association with *IL4* and the risk of developing a latent infection. A study using three *IL4* SNP in DRC did not find any associations, but that was in a comparison between cases and controls [15]. In B10.Q mice deletion in *IL4* lead to increased *T. brucei* parasitaemia levels but longer survival time [18]. The relationship between *IL4* polymorphisms and acute HAT infection requires further study.

Association with *APOL1*

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

Association with *MIF*

Although, our data show that subjects carrying the G allele of *MIF* (rs36086171) G/A had a risk of developing active HAT ([Table 3](#)), we did not find a significant difference after correction (BONF = 0.52). MIF is a ubiquitously expressed protein that has proinflammatory, hormonal and enzymatic activities [34]. It is implicated in many inflammatory diseases [35]. It functions by recruiting myeloid cells to the site of inflammation [36], by inducing their differentiation towards M1 cells secreting TNF [37] and by suppressing p53-dependent apoptosis of inflammatory cells. While there are no human studies directly linking *MIF* to HAT, murine studies show that *MIF* plays a role as a mediator of the inflammation which is a key feature in trypanosomiasis-associated pathology [16,38].

Association with *HLA-G*

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

Conclusion

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

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

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

Materials and methods

Population study, definition of phenotypes and study design

The study took place in western-central Côte d'Ivoire in the main HAT foci of Bonon, Bouafle, Zoukougbeu, Oume, and Sinfra. Samples were collected in three stages: (i) previously archived samples; existing collections of previously archived samples were centralized to Centre International de Recherche-Développement sur l'Elevage en zone Subhumide (CIRDES) in Burkina-Faso, (ii) retrospectively collected samples for which study sites were revisited for consent and resample previously diagnosed and treated patients and (iii) prospectively collected samples including new HAT cases. Three phenotypes were considered: (i) cases, defined as individuals in whom the presence of trypanosomes was confirmed by microscopy and trypanolysis

test positive (TL+ve) but trypanolysis negative (TL-ve) for some cases who were retrospectively sampled for the purpose of this study and who had not been subject to the trypanolysis test (TL) when originally identified as HAT cases by microscopy; (ii) controls, defined as individuals living in the endemic area who are card agglutination test for trypanosomiasis (CATT) negative (CATT-ve) and trypanolysis test negative (TL-ve), and without evidence of previous HAT infection and (iii) latent infections, defined as subjects who were CATT positive (CATT +ve) with plasma dilution higher or equal to ¼ (CATTpl ≥1/4) but in whom parasites could not be detected by microscopy at repeated tests for at least two years.

A total of 233 individuals with 100 HAT cases, 100 controls and 33 latent infections were included in the study. Samples were frozen directly in the field at -20˚C and kept at that temperature until use. For each individual, an aliquot of plasma was used to perform the immune TL test that detects Litat 1.3 and Litat 1.5 variable surface antigens specific for *T.b. gambiense* [43].

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

Selection of SNPs and genotyping

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

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

DNA was genotyped by two commercial service providers: INRA- Site de Pierroton, Plateforme Genome Transcriptome de Bordeaux, France and ii- LGC genomics Hoddesden UK. At INRA multiplex design (two sets of 40 SNPs) was performed using Assay Design Suite v2.0 (Agena Biosciences). SNP genotyping was achieved with the iPLEX Gold genotyping kit (Agena Biosciences) for the Mass Array iPLEX genotyping assay, following the manufacturer's instructions. Products were detected on a Mass Array mass spectrophotometer and data were acquired in real time with Mass Array RT software (Agena Biosciences). SNP clustering and validation was carried out with Typer 4.0 software (Agena Biosciences). At LGC Genomics SNP were genotyped using the PCR based KASP assay[48].

Statistical analysis

All statistical analyses were performed using the Plink 1.9 and R v 3.2.1 software. Individuals were excluded who had > 10% missing SNP data. SNP loci were excluded that > 10% missing genotypes or if the control samples were not in Hardy-Weinberg equilibrium (p<0.001). Case-control association analysis using SNP alleles/genotypes was undertaken using Fisher's

exact test. The difference between the ethnic groups was estimated using the F_{ST}, with values potentially varying from zero (no population differentiation) to one (complete differentiation) [47]. Ethnicity was not used as a risk factor for HAT because we observed no significant differences in F_{ST} between ethnic groups (F_{ST} maximum 0.015) or clustering by linguistic group by multidimensional scaling in Plink. Some studies have used ethnicity and age as risk factors for HAT but found no significant association [11,15]. 77 SNP remained in at least one comparison after filtering. *P* values were adjusted for multiple testing using the Bonferroni corrections and Benjamini Hocheberg false discovery rate test as implemented in Plink.

Ethics statement

The population of the study was informed. All adult subjects provided written informed consent. For children, a parent or guardian of any child participant provided written informed consent on their behalf. The protocol of the study was approved by the traditional authorities (chief and village committee) and by National ethics committee hosted by the Public Health Ministry of Côte d'Ivoire with the number: N˚38/MSLS/CNERm-dkn 5th May 2014. This study is part of the TrypanoGen project which aims to a better understanding of genetic determinism of human susceptibility to HAT and the TrypanoGen-CI samples were archived in the TrypanoGen Biobank hosted by CIRDES in Bobo-Dioulasso, Burkina Faso [44].

Supporting information

S1 Table. Quality control test between HAT cases and controls.
(XLSX)

S2 Table. Quality control test between latent infections and HAT cases.
(XLSX)

S3 Table. Quality control test between latent infections and controls.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Mathurin Koffi.

Data curation: Christiane Hertz-Fowler, Harry Noyes.

Formal analysis: Julius Mulindwa, Harry Noyes.

Funding acquisition: Mathurin Koffi, Gustave Simo, Enock Matovu, Christiane Hertz-Fowler, Bruno Bucheton, Issa Sidibé, Vincent Jamonneau, Annette MacLeod.

Methodology: Hamidou Ilboudo, Julius Mulindwa, Christiane Hertz-Fowler, Bruno Bucheton, Harry Noyes, Simon-Pierre N'Guetta.

Supervision: Mathurin Koffi, Gustave Simo, Enock Matovu, Vincent Jamonneau, Annette MacLeod, Harry Noyes, Simon-Pierre N'Guetta.

Writing – original draft: Bernardin Ahouty, Harry Noyes.

Writing – review & editing: Bernardin Ahouty, Mathurin Koffi, Hamidou Ilboudo, Harry Noyes, Simon-Pierre N'Guetta.

References

1. Simarro PP, Diarra A, Ruiz Postigo JA, Franco JR, Jannin JG (2011) The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000–2009: the way forward. PLoS Negl Trop Dis 5: e1007. https://doi.org/10.1371/journal.pntd.0001007 PMID: 21364972

2. WHO (2013) Control and surveillance of human African trypanosomiasis. WHO Technical Report Series 984 Geneva.

3. Checchi F, Filipe JA, Barrett MP, Chandramohan D (2008) The natural progression of *Gambiense* sleeping sickness: what is the evidence? PLoS Negl Trop Dis 2: e303. https://doi.org/10.1371/journal.pntd.0000303 PMID: 19104656

4. Garcia A, Jamonneau V, Magnus E, Laveissiere C, Lejon V, et al. (2000) Follow-up of Card Agglutination Trypanosomiasis Test (CATT) positive but apparently aparasitaemic individuals in Cote d'Ivoire: evidence for a complex and heterogeneous population. Trop Med Int Health 5: 786–793. PMID: 11123826

5. Jamonneau V, Ilboudo H, Kabore J, Kaba D, Koffi M, et al. (2012) Untreated human infections by Trypanosoma brucei gambiense are not 100% fatal. PLoS Negl Trop Dis 6: e1691. https://doi.org/10.1371/journal.pntd.0001691 PMID: 22720107

6. Kabore J, Macleod A, Jamonneau V, Ilboudo H, Duffy C, et al. (2011) Population genetic structure of Guinea Trypanosoma brucei gambiense isolates according to host factors. Infect Genet Evol 11: 1129–1135. https://doi.org/10.1016/j.meegid.2011.04.011 PMID: 21515408

7. Koffi M, De Meeus T, Bucheton B, Solano P, Camara M, et al. (2009) Population genetics of *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. Proc Natl Acad Sci U S A 106: 209–214. https://doi.org/10.1073/pnas.0811080106 PMID: 19106297

8. Ilboudo H, Bras-Goncalves R, Camara M, Flori L, Camara O, et al. (2014) Unravelling human trypanotolerance: IL8 is associated with infection control whereas IL10 and TNFalpha are associated with subsequent disease development. PLoS Pathog 10: e1004469. https://doi.org/10.1371/journal.ppat.1004469 PMID: 25375156

9. Bucheton B, MacLeod A, Jamonneau V (2011) Human host determinants influencing the outcome of Trypanosoma brucei gambiense infections. Parasite Immunol 33: 438–447. https://doi.org/10.1111/j.1365-3024.2011.01287.x PMID: 21385185

10. Jamonneau V, Ravel S, Garcia A, Koffi M, Truc P, et al. (2004) Characterization of *Trypanosoma brucei s.l.* infecting asymptomatic sleeping-sickness patients in Cote d'Ivoire: a new genetic group? Ann Trop Med Parasitol 98: 329–337. https://doi.org/10.1179/000349804225003406 PMID: 15228714

11. Courtin D, Argiro L, Jamonneau V, N'Dri L, N'Guessan P, et al. (2006) Interest of tumor necrosis factor-alpha -308 G/A and interleukin-10–592 C/A polymorphisms in human African trypanosomiasis. Infect Genet Evol 6: 123–129. https://doi.org/10.1016/j.meegid.2005.03.002 PMID: 15894515

12. Courtin D, Milet J, Sabbagh A, Massaro JD, Castelli EC, et al. (2013) HLA-G 3' UTR-2 haplotype is associated with Human African trypanosomiasis susceptibility. Infect Genet Evol 17: 1–7. https://doi.org/10.1016/j.meegid.2013.03.004 PMID: 23541412

13. Iraqi F, Clapcott SJ, Kumari P, Haley CS, Kemp SJ, et al. (2000) Fine mapping of trypanosomiasis resistance loci in murine advanced intercross lines. Mamm Genome 11: 645–648. PMID: 10920233

14. Kemp SJ, Iraqi F, Darvasi A, Soller M, Teale AJ (1997) Localization of genes controlling resistance to trypanosomiasis in mice. Nat Genet 16: 194–196. https://doi.org/10.1038/ng0697-194 PMID: 9171834

15. Courtin D, Milet J, Jamonneau V, Yeminanga CS, Kumeso VK, et al. (2007) Association between human African trypanosomiasis and the IL6 gene in a Congolese population. Infect Genet Evol 7: 60–68. https://doi.org/10.1016/j.meegid.2006.04.001 PMID: 16720107

16. Stijlemans B, Leng L, Brys L, Sparkes A, Vansintjan L, et al. (2014) MIF contributes to Trypanosoma brucei associated immunopathogenicity development. PLoS Pathog 10: e1004414. https://doi.org/10.1371/journal.ppat.1004414 PMID: 25255103

17. Tiberti N, Hainard A, Lejon V, Robin X, Ngoyi DM, et al. (2010) Discovery and verification of osteopontin and Beta-2-microglobulin as promising markers for staging human African trypanosomiasis. Mol Cell Proteomics 9: 2783–2795. https://doi.org/10.1074/mcp.M110.001008 PMID: 20724469

18. Bakhiet M, Jansson L, Buscher P, Holmdahl R, Kristensson K, et al. (1996) Control of parasitemia and survival during Trypanosoma brucei brucei infection is related to strain-dependent ability to produce IL-4. J Immunol 157: 3518–3526. PMID: 8871651

19. MacLean L, Chisi JE, Odiit M, Gibson WC, Ferris V, et al. (2004) Severity of human african trypanosomiasis in East Africa is associated with geographic location, parasite genotype, and host inflammatory cytokine response profile. Infect Immun 72: 7040–7044. https://doi.org/10.1128/IAI.72.12.7040-7044.2004 PMID: 15557627

20. Pays E, Vanhollebeke B, Vanhamme L, Paturiaux-Hanocq F, Nolan DP, et al. (2006) The trypanolytic factor of human serum. Nat Rev Microbiol 4: 477–486. https://doi.org/10.1038/nrmicro1428 PMID: 16710327

21. Ilboudo H, Berthier D, Camara M, Camara O, Kabore J, et al. (2012) APOL1 expression is induced by Trypanosoma brucei gambiense infection but is not associated with differential susceptibility to sleeping sickness. Infect Genet Evol 12: 1519–1523. https://doi.org/10.1016/j.meegid.2012.05.010 PMID: 22691369

22. Yamada Y, Metoki N, Yoshida H, Satoh K, Ichihara S, et al. (2006) Genetic risk for ischemic and hemorrhagic stroke. Arterioscler Thromb Vasc Biol 26: 1920–1925. https://doi.org/10.1161/01.ATV.0000229694.97827.38 PMID: 16741147

23. Raj DS (2009) Role of interleukin-6 in the anemia of chronic disease. Semin Arthritis Rheum 38: 382–388. https://doi.org/10.1016/j.semarthrit.2008.01.006 PMID: 18336871

24. Perera MK, Herath NP, Pathirana SL, Phone-Kyaw M, Alles HK, et al. (2013) Association of high plasma TNF-alpha levels and TNF-alpha/IL-10 ratios with TNF2 allele in severe P. falciparum malaria patients in Sri Lanka. Pathog Glob Health 107: 21–29. https://doi.org/10.1179/2047773212Y.0000000069 PMID: 23432860

25. Gichohi-Wainaina WN, Melse-Boonstra A, Swinkels DW, Zimmermann MB, Feskens EJ, et al. (2015) Common Variants and Haplotypes in the TF, TNF-alpha, and TMPRSS6 Genes Are Associated with Iron Status in a Female Black South African Population. J Nutr 145: 945–953. https://doi.org/10.3945/jn.114.209148 PMID: 25809685

26. Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, et al. (1999) A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. Nat Genet 22: 145–150. https://doi.org/10.1038/9649 PMID: 10369255

27. Tu IP, Whittemore AS (1999) Power of association and linkage tests when the disease alleles are unobserved. Am J Hum Genet 64: 641–649. https://doi.org/10.1086/302253 PMID: 9973303

28. Magez S, Caljon G (2011) Mouse models for pathogenic African trypanosomes: unravelling the immunology of host-parasite-vector interactions. Parasite Immunol 33: 423–429. https://doi.org/10.1111/j.1365-3024.2011.01293.x PMID: 21480934

29. Daulouede S, Bouteille B, Moynet D, De Baetselier P, Courtois P, et al. (2001) Human macrophage tumor necrosis factor (TNF)-alpha production induced by Trypanosoma brucei gambiense and the role of TNF-alpha in parasite control. J Infect Dis 183: 988–991. https://doi.org/10.1086/319257 PMID: 11237819

30. Magez S, Radwanska M, Beschin A, Sekikawa K, De Baetselier P (1999) Tumor necrosis factor alpha is a key mediator in the regulation of experimental Trypanosoma brucei infections. Infect Immun 67: 3128–3132. PMID: 10338530

31. Cooper A, Ilboudo H, Alibu VP, Ravel S, Enyaru J, et al. (2017) APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis. Elife 6.

32. Kabore WJ, Ilboudo H, Noyes H, Camara O, Kabore J, et al. (2017) Candidate gene study between human African trypanosomiasis clinical phenotypes in Guinea. PLoS Negl Trop Dis in press.

33. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, et al. (2010) Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science 329: 841–845. https://doi.org/10.1126/science.1193032 PMID: 20647424

34. Barton A, Lamb R, Symmons D, Silman A, Thomson W, et al. (2003) Macrophage migration inhibitory factor (MIF) gene polymorphism is associated with susceptibility to but not severity of inflammatory polyarthritis. Genes Immun 4: 487–491. https://doi.org/10.1038/sj.gene.6364014 PMID: 14551601

35. Rosado Jde D, Rodriguez-Sosa M (2011) Macrophage migration inhibitory factor (MIF): a key player in protozoan infections. Int J Biol Sci 7: 1239–1256. PMID: 22110378

36. Bernhagen J, Krohn R, Lue H, Gregory JL, Zernecke A, et al. (2007) MIF is a noncognate ligand of CXC chemokine receptors in inflammatory and atherogenic cell recruitment. Nat Med 13: 587–596. https://doi.org/10.1038/nm1567 PMID: 17435771

37. Calandra T, Bernhagen J, Metz CN, Spiegel LA, Bacher M, et al. (1995) MIF as a glucocorticoid-induced modulator of cytokine production. Nature 377: 68–71. https://doi.org/10.1038/377068a0 PMID: 7659164

38. Stijlemans B, Beschin A, Magez S, Van Ginderachter JA, De Baetselier P (2015) Iron Homeostasis and Trypanosoma brucei Associated Immunopathogenicity Development: A Battle/Quest for Iron. Biomed Res Int 2015: 819389. https://doi.org/10.1155/2015/819389 PMID: 26090446

39. Lajoie J, Hargrove J, Zijenah LS, Humphrey JH, Ward BJ, et al. (2006) Genetic variants in nonclassical major histocompatibility complex class I human leukocyte antigen (HLA)-E and HLA-G molecules are associated with susceptibility to heterosexual acquisition of HIV-1. J Infect Dis 193: 298–301. https://doi.org/10.1086/498877 PMID: 16362895

40. Ferguson R, Ramanakumar AV, Richardson H, Tellier PP, Coutlee F, et al. (2011) Human leukocyte antigen (HLA)-E and HLA-G polymorphisms in human papillomavirus infection susceptibility and persistence. Hum Immunol 72: 337–341. https://doi.org/10.1016/j.humimm.2011.01.010 PMID: 21256910

41. Lafon M, Prehaud C, Megret F, Lafage M, Mouillot G, et al. (2005) Modulation of HLA-G expression in human neural cells after neurotropic viral infections. J Virol 79: 15226–15237. https://doi.org/10.1128/JVI.79.24.15226-15237.2005 PMID: 16306594

42. Zhao M, Zhang R, Xu X, Liu Y, Zhang H, et al. (2013) IL-10 reduces levels of apoptosis in Toxoplasma gondii-infected trophoblasts. PLoS One 8: e56455. https://doi.org/10.1371/journal.pone.0056455 PMID: 23418570

43. Jamonneau V, Bucheton B, Kabore J, Ilboudo H, Camara O, et al. (2010) Revisiting the immune trypanolysis test to optimise epidemiological surveillance and control of sleeping sickness in West Africa. PLoS Negl Trop Dis 4: e917. https://doi.org/10.1371/journal.pntd.0000917 PMID: 21200417

44. Ilboudo H, Noyes H, Mulindwa J, Kimuda MP, Koffi M, et al. (2017) Introducing the TrypanoGEN biobank: A valuable resource for the elimination of human African trypanosomiasis. PLoS Negl Trop Dis 11: e0005438. https://doi.org/10.1371/journal.pntd.0005438 PMID: 28570558

45. Zhao JH (2007) Gap: genetic analysis package. J Stat Soft 23: 1–18.

46. Gray IC, Campbell DA, Spurr NK (2000) Single nucleotide polymorphisms as tools in human genetics. Hum Mol Genet 9: 2403–2408. PMID: 11005795

47. Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evol 38: 1358–1370.

48. Semagn k, Babu R, Hearn S, Olsen MS (2013) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. Mol Breed 33: 1–14.