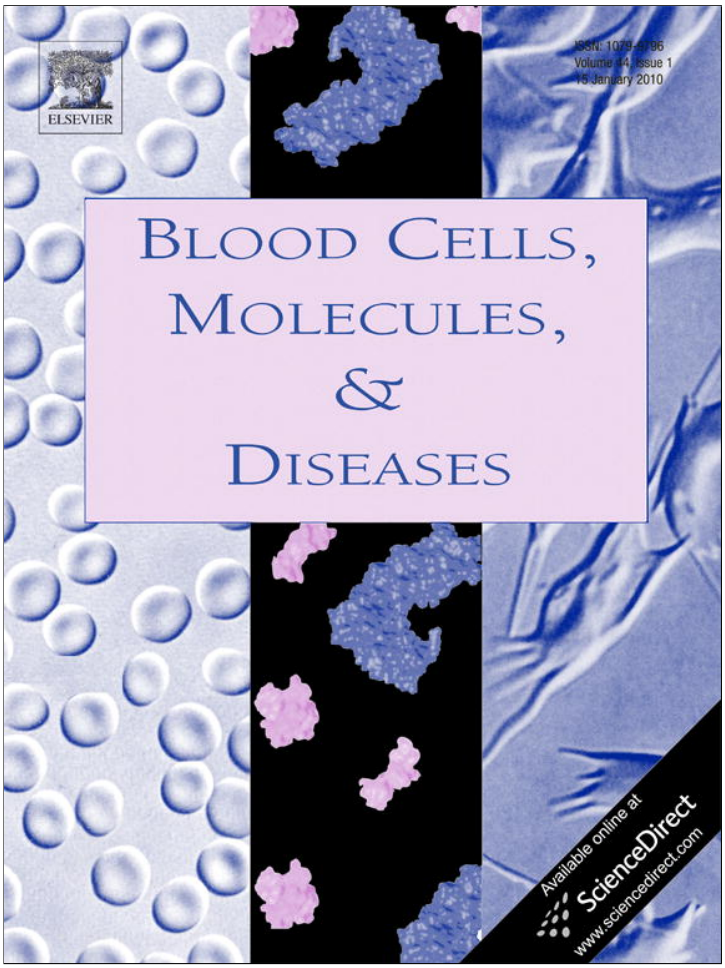

Blood Cells, Molecules, and Diseases 44 (2010) 62–68

Contents lists available at ScienceDirect

Blood Cells, Molecules, and Diseases

journal homepage: www.elsevier.com/locate/ybcmd

Analysis of malaria associated genetic traits in Cabo Verde, a melting pot of European and sub Saharan settlers

Joana Alves [a,b], Patrícia Machado [a], João Silva [a], Nilza Gonçalves [a], Letícia Ribeiro [c], Paula Faustino [d], Virgílio Estólio do Rosário [a], Licínio Manco [e], Leonor Gusmão [f], António Amorim [f,g], Ana Paula Arez [a,*]

[a] Centre for Malaria and Tropical Diseases, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira, 100, 1349-008 Lisbon, Portugal
[b] Ministry of Health, Palácio do Governo, CP 47, Cabo Verde
[c] Hematology Department, Centro Hospitalar de Coimbra, Portugal
[d] Genetics Department, Instituto Nacional de Saúde Dr Ricardo Jorge, Lisbon, Portugal
[e] Centre of Research in Anthropology in Health/Department of Anthropology, Universidade de Coimbra, Portugal
[f] Institute of Molecular Pathology and Immunology of University of Porto (IPATIMUP), Oporto, Portugal
[g] Faculty of Sciences, Universidade do Porto, Portugal

ARTICLE INFO

Article history:
Submitted 31 July 2009
Available online 17 October 2009

(Communicated by Sir D. Weatherall, F.R.S., 17 September 2009)

Keywords:
Hemoglobin S
Glucose-6-Phophate-dehydrogenase
Pyruvate Kinase
Cabo Verde
Malaria

ABSTRACT

Malaria has occurred in the Cabo Verde archipelago with epidemic characteristics since its colonization. Nowadays, it occurs in Santiago Island alone and though prophylaxis is not recommended by the World Health Organization, studies have highlight the prospect of malaria becoming a serious public health problem as a result of the presence of antimalarial drug resistance associated with mutations in the parasite populations and underscore the need for tighter surveillance.

Despite the presumptive weak immune status of the population, severe symptoms of malaria are not observed and many people present a subclinical course of the disease. No data on the prevalence of sickle-cell trait and red cell glucose-6-phosphate dehydrogenase deficiency (two classical genetic factors associated with resistance to severe malaria) were available for the Cabo Verde archipelago and, therefore, we studied the low morbidity from malaria in relation to the particular genetic characteristics of the human host population. We also included the analysis of the pyruvate kinase deficiency associated gene, reported as putatively associated with resistance to the disease.

Allelic frequencies of the polymorphisms examined are closer to European than to African populations and no malaria selection signatures were found. No association was found between the analyzed human factors and infection but one result is of high interest: a linkage disequilibrium test revealed an association of distant loci in the PKLR gene and adjacent regions, only in non-infected individuals. This could mean a more conserved gene region selected in association to protection against the infection and/or the disease.

Introduction

According to de Meira et al. [1] epidemic malaria is known to have occurred in the Cabo Verde archipelago since the remote past . Malaria should have been introduced in the archipelago during its colonization in the XV century. Records from 1507 report that the old Portuguese sailing ships (*caravelas*) from the spice route were not allowed in Cabo Verde ports because of the fear of getting malaria [2]. In 1952, da Costa Monteiro [3] reported malaria as the most serious public health problem in the archipelago, Santiago being the most affected island.

Cabo Verde is comprised of 10 islands in the Atlantic Ocean, 500 km west of Senegal. Santiago is the largest island, where approximately half of the population resides (capital: Praia). Malaria was almost eradicated between 1954 and 1970 and since 1973 autochthonous cases are only observed in this island [4]. The World Health Organization (WHO) [5] considers there to be a limited risk of malaria between September and November. There is no recommendation for prophylaxis but recent studies highlight the prospect of malaria recurring as a serious public health problem in Cabo Verde and underscores the need for a closer and continuous surveillance. The population is considered to be non-immune or semi-immune and irregular outbreaks occur. An outbreak in 1995–1996 in the St. Catarina district was followed by parasitological and molecular analysis during 1 year [6]. Studies indicated that malaria is maintained as asymptomatic and sub-patent infections and that the majority of the circulating parasite populations harbor chloroquine-resistant mutations [7].

In the previous two studies, no complicated malaria cases were found in spite of high parasitaemias. Most of patent parasitaemias

* Corresponding author. Fax: +351 213622458.
E-mail address: aparez@ihmt.unl.pt (A.P. Arez).

1079-9796/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bcmd.2009.09.008

were above the range of 1000–10,000 parasites/μl, usually considered a cut-off level for malaria attacks [8]. However, individuals of all ages presented no more than mild symptoms such as fever, headache, nausea and general malaise. This population seemed not to develop severe symptoms of malaria despite its presumptive weak immune status and many persons exhibit a subclinical course. The low morbidity associated with malaria infections in this island may be related to factors of both parasites and host, which may control the severity of the malaria infection.

In the localized outbreak in St. Catarina district [6], we suggest that the genetically homogeneous circulating parasite population could have been a weakly virulent parasite. However, when different localities were studied [7] and *Plasmodium falciparum* heterogeneity was observed this hypothesis proved untenable. Therefore, no evidence is available regarding the contribution of parasite factors to the low morbidity observed in the island.

The establishment of clinical symptoms could be attenuated due to premunition, already described for other areas of unstable and low-level transmission of malaria [9,10,11]. Also, differences in clinical consequences of infection with *P. falciparum* as consequence of host factors have already been demonstrated [12,13] and the most common and best characterized protective polymorphisms are those involving the erythrocyte-specific proteins and enzymes, such as hemoglobin (Hb) and glucose-6-phosphate dehydrogenase (G6PD) variants.

Questioning if the observed low morbidity in Santiago Island could be a consequence of particular characteristics of the host population and since no data on the frequency of these human genetic polymorphisms are available for the Cabo Verde archipelago we studied the prevalence of HbS allele responsible for the sickle-cell trait (heterozygosity for the HbS mutation in β-globin gene, Hb β globin) and the prevalence of G6PD variants, two classical genetic factors strongly associated to resistance against human severe malaria.

Further, both may have a crucial importance in the control and management of malaria cases. Malaria can be one of the major causes of hospitalization and death in patients with sickle cell anemia and as a result, antimalarial prophylaxis is included in the standard management of patients with the disease. However, with the spread of chloroquine resistance there is an on-going debate on which drugs should now be used [14]. Concerning G6PD deficiency, the epidemic conditions of *P. falciparum* malaria justify the use of primaquine as a gametocidal drug but this drug presents potentially fatal side effects in G6PD-deficient individuals [15].

In sub-Saharan Africa, X-linked G6PD is essentially a tri-allelic polymorphism: G6PDB, the most common allele associated to normal enzymatic activity; G6PDA, which results in approximately 85% of the normal enzymatic activity and the G6PDA$^-$ deficiency allele, which implies only around 12% of normal enzymatic activity with a range of 5–25% in sub-Saharan Africa [16,17]. However, considering the history of Cabo Verde settlement and the reported high European contribution, [18] we also searched for the G6PD Mediterranean (Med) variant, the most common in countries surrounding the Mediterranean Sea [19]. This variant is associated with 3% of normal enzyme activity and usually ranges in frequency from 2% to 20% in Europe [20].

More recently, pyruvate kinase (PK) deficiency was associated with resistance to the disease in rodent models [21] and humans [22,23]. Up to now, elevated frequencies of pyruvate kinase liver and red cells (PKLR)-deficient alleles have not been recorded in areas endemic for malaria, although a systematic analysis has not been done. The information about the frequency of PK deficiency in African populations is clearly limited [24,25]. We, therefore, included its analysis in this study. The PKLR gene (1q21) encodes for either PK-L (in liver) or PK-R (in red cells), according to the use of tissue-specific promoters (leading to structural differences in the protein N-terminal region). The coding region is split into 12 exons, 10 of which are shared by the 2 isoforms, while exons 1 and 2 are specific for the erythrocyte and hepatic isozyme, respectively. About 180 mutations associated with PK-deficiency and 8 polymorphic sites have been reported in the PKLR gene [26].

Materials and methods

Study area and Isolates

Biological material–DNA samples obtained from blood–was already available for this analysis. Samples were collected in localities from different Districts of Santiago Island (Praia—south, St Catarina—west, St Cruz—east and Tarrafal—north) in 1995–1996 [6], 1998–2000 and 2003 [7]. From a total of 1056 available samples, a sub-sample of 257 unrelated individuals was used for the present study (99 individuals from Praia, 23 from St Cruz, 119 from St Catarina and 16 from Tarrafal).

Individual data such as gender, age, and malaria history were available. Further, given that each individual was well characterized for *Plasmodium*-infection (species and genotype) and clinical status (most of them asymptomatic and a few with mild symptoms), two groups were defined: 64 infected (I—presence of infection at least once during the collections period) and 188 non-infected (NI—absence of infection throughout the collection period); infection status was uncertain in 5 individuals.

For the analysis of PK polymorphisms, two additional groups were also analyzed–80 adult healthy Portuguese individuals–PT-C (DNA prepared from finger-prick blood samples collected in 2006 at Health Centre of Coruche, Portugal as described in [27]) and 21 Portuguese individuals with hereditary nonspherocytic hemolytic anemia (HNSHA) caused by PK-deficiency—PT-PKD (DNA prepared from venous blood samples). These PK-deficient individuals were previously diagnosed by PK enzyme assay and molecular genetic analysis [28,29].

The investigation was approved by the Ministry of Health of Cabo Verde and by the Ethical Committee at institutions involved in the study. Each person (or parent) was informed of the nature and aims of the study and told that participation was voluntary.

Detection of hemoglobin S allele (HbS)

The mutation at c.6 of the β globin gene was detected using an adaptation of the technique described by Waterfall and Cobb [30] and the homozygous HbSS status was confirmed by a PCR-RFLP technique (details as Supplementary Material).

Detection of glucose-6-phosphate dehydrogenase polymorphisms

Mutations in the G6PD gene were detected by a PCR-RFLP method as described in Tishkoff et al. [20] (details as Supplementary Material). The possible nine genotypes were grouped as follows: hemizygous males G6PDB and G6PDA, homozygous females G6PDBB and G6PDAA and heterozygous females G6PDBA (variants with a putative normal enzyme activity) as g6pd$^+$; hemizygous males G6PDA$^-$ and homozygous females G6PDA$^-$A$^-$ (putative deficient variants) as g6pd$^-$ and heterozygous females G6PDBA$^-$ and G6PDAA$^-$ (variants with a putative intermediate enzyme activity) as g6pd$^\pm$ [31].

Detection of pyruvate kinase polymorphisms

Analysis of PKLR gene was done by two approaches: (1) typing of polymorphic loci and searching for relevant mutations associated to PK-deficiency previously described and (2) search for new microsatellite regions—short tandem repeats (STRs) in the gene and adjacent regions. In total, a PKLR gene spanning region of 95 kb was analyzed.

Analysis of binary polymorphisms

Two mutations were investigated: 269T>A (90Ile>Asn) at exon 3, the mutation identified in mice as associated to malaria protection [21], and already described in PK-deficient individuals [32] (technical details as Supplementary Material) and 1456C>T (486Arg>Trp) at exon 11, the most common mutation responsible for PK deficiency in humans from Portugal and some Sub-Saharan regions [33,34,35]. Also, two polymorphisms were analyzed: the single nucleotide polymorphism (SNP) 1705A/C at exon 12 [35,36,] and the T10/19 repeat at intron 10 [37], common polymorphic sites in São Tomé e Príncipe [24].

Analysis of STRs

After searching for STRs in the PKLR gene (accession nr AY316591) and downstream/upstream adjacent regions (accession nr AL713999), 4 loci were chosen for analysis: 2 inside (intron 3-IVS3 and intron 11-IVS11) and 2 downstream the gene (25 kb—locus PKA and 65 kb —PKV). IVS11 was the only one already described as polymorphic [38] (see Supplementary Material for amplification conditions and analysis of PCR products).

Statistical analysis

Pearson χ^2 test was used for comparison of populations from different districts and malaria I and NI groups. Additionally, PK polymorphisms were also compared with the two Portuguese groups, PT-C and PT-PKD. Allelic frequencies and selection signatures were investigated (genetic diversity, Hardy–Weinberg equilibrium deviation and linkage disequilibrium) with Arlequin 3.11. for Windows [39]. For all tests, a significance level of 0.05 was considered.

Results

Hemoglobin polymorphisms

The β globin genotype was successfully defined for a total of 217 individuals (84%). From these, 92% were HbAA, 7% HbAS and 1% HbSS. HbS allele was only found in Praia (11% of HbAS) and St Catarina (4% of HbAS and 3% of HbSS) districts with a very low frequency (0.05).

I and NI individuals distributed similarly among HbAA and HbAS genotypes [21% I and 79% NI in the HbAA group (unknown infection status of 3 individuals) and 23% I and 77% NI in the HbAS group (unknown infection status of 1 individual)]. All 3 HbSS individuals were I.

Glucose-6-phosphate dehydrogenase polymorphisms

G6PD genotype was measured in a total of 176 (68%) individuals, 77 males and 99 females. Seventy-four percent of males presented G6PDB genotype, 25% G6PDA and 1% G6PDA$^-$; 61% of females presented G6PDBB genotype, 29% G6PDBA, 6% G6PDAA and 4% G6PDAA$^-$ (no genotypes G6PDBA$^-$ and G6PDA$^-$A$^-$ were found).

In the total population, allelic frequencies were $f(B) = 0.95$, $f(A) = 0.04$ and $f(A^-) = 0.008$, respectively but A$^-$ allele was only found in Praia and Tarrafal districts, being much more frequent in the latter—0.019 and 0.115, respectively ($P = 0.007$), which reflected the presence of 3 G6PDAA$^-$ genotypes. G6PDMed variant was not found.

Ninety-seven percent of individuals were G6PD$^+$, 2% were G6PD$^\pm$ and 1% were G6PD$^-$. Normal condition seems to be equally prevalent in both genders (99% G6PD$^+$ in males and 96% in females); 1% and 0% of G6PD$^-$ in males and females, respectively and 4% of G6PD$^\pm$ in females.

Among A$^-$ carriers, all except one G6PDAA$^-$ female were NI. I and NI individuals distributed similarly among G6PD$^+$ and G6PD$^\pm$ groups [33% I and 64% NI in the G6PD$^+$ group (unknown infection status of 5 individuals) and 25% I and 75% NI in the G6PD$^\pm$ group]. The only G6PD$^-$individual was NI.

Pyruvate kinase polymorphisms

Binary polymorphisms

The 269T>A (exon 3) and 1456C>T (exon 11) mutations were screened with success in 253 (98%) and 255 (98%) individuals respectively and mutated alleles were not found.

Polymorphisms 1705A/C (exon 12) and T10/19 (intron 10) were accomplished in a total of 200 individuals (78%). Regarding 1705A/C, 19.5% were of AA genotype, 33% CC and 47.5% AC. Allelic frequencies were $f(A) = 0.43$ and $f(C) = 0.57$. Regarding (T)10/19, 27% were of 10/10 genotype, 20.5% 19/19 and 52.5% 10/19. Allelic frequencies were determined to be $fT(10) = 0.53$ and $fT(19) = 0.47$. The analysis of possible haplotypes revealed that 1705C/(T)10 exhibited a frequency of 0.52 and 1705A/(T)19 a frequency of 0.42. The other two, 1705A/(T)10 and 1705C/(T)19, presented very low frequencies (0.01 and 0.05, respectively).

F_{ST} values were calculated for all pairs of districts and only St Catarina and St Cruz revealed significant differences ($P = 0.045 \pm 0.02$). Concerning both 1705A/C and (T)10/19 allelic frequencies, while St Catarina follows the general trend [$f(A) = 0.41$ and $f(C) = 0.59$; $f(T)10 = 0.54$ and $f(T)19 = 0.46$], in St Cruz values are inverted [$f(A) = 0.57$ and $f(C) = 0.43$; $f(T)10 = 0.39$ and $f(T)19 = 0.61$]. Haplotype frequencies were also different in St Cruz—on the opposite to the general population, 1705A/(T)19 was the predominant haplotype with a frequency of 0.57, followed by 1705C/(T)10 with 0.39; 1705C/(T)19 was present with a frequency of 0.05 and 1705A/(T)10 was absent.

In total population, no significant differences were found between I and NI. However, when districts were compared separately, certain differences were found in St Catarina as regards locus (T)10/19—the group of I individuals showed a significantly higher heterozygosity than expected ($P = 0.009$) and allelic frequencies were inverted comparing to the general trend [$fT(10) = 0.48$ and $fT(19) = 0.52$] in the NI. Regarding haplotypes, in the NI group, both 1705C/(T)10 and 1705A/(T)19 showed similar frequencies (0.47 and 0.45, respectively) and 1705C/(T)19 showed higher frequency than in the other groups (0.07).

STRs

The 4 STR loci in the PKLR gene and downstream adjacent region–IVS3 (intron 3), IVS11 (intron 11), PKA (25 kb downstream) and PKV (65 kb downstream) (Fig. 1)–were screened in 252 individuals (98%).

All STRs were confirmed to be polymorphic with variable number of repeats—the number of (ATT) repeats in the IVS11 locus varied between 7 and 18, the number of (AAAT) repeats in the PKA locus varied between 6 and 21 and the number of (TTTA) repeats in the PKV locus varied between 8 and 13. The IVS3 locus is the most polymorphic with 8 repeat regions and it is interrupted. The consensus sequence determined, allele classification, etc. are presented as Supplementary Material. The number of repeats in this locus varied between 27 and 43.2, which were nomenclature as alleles 1 to 26.

In the overall population of Cabo Verde, IVS3 locus presented the greatest diversity indices with the larger allele number and expected heterozygosity (Table 1). Observed heterozygosity was according to Hardy–Weinberg expected frequencies for all loci, except for IVS3, which it is significantly below the expected ($P = 0.000$). All pairs of loci revealed a marked Linkage Disequilibrium (LD) ($P = 0.000$), i.e. a significant LD for a ≈75 kbp spanning region (IVS3 was not considered as it was not in Hardy–Weinberg equilibrium).

When districts were compared and F_{ST} values calculated, significant values were obtained for all pairs including St Cruz (vs. Praia—0.012, $P = 0.018$; vs. St Catarina—0.015, $P = 0.009$; vs. Tarrafal—0.012, $P = 0.045$). All the other three revealed no differences between each

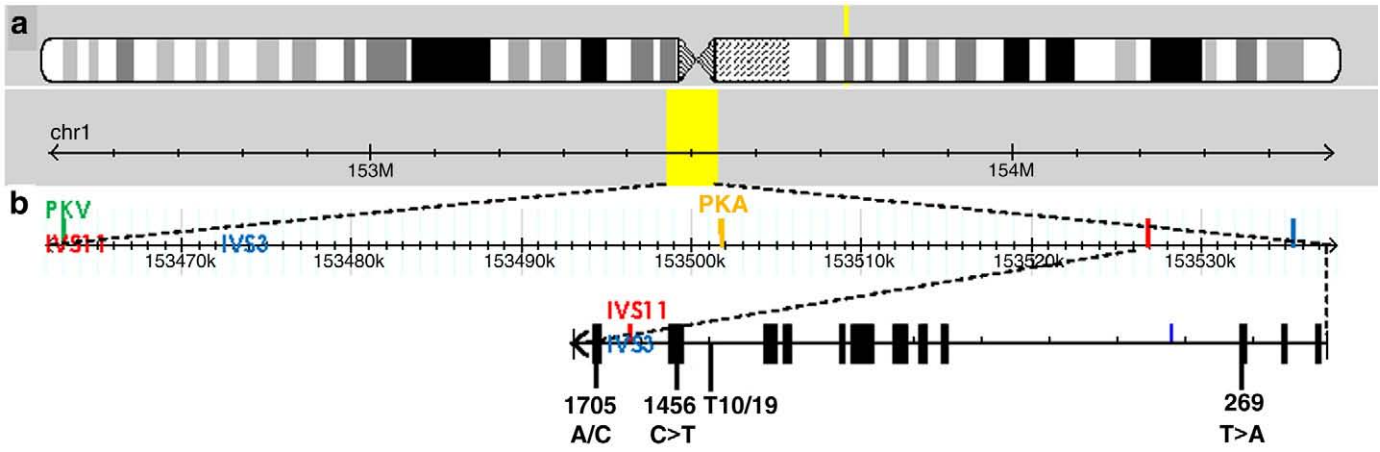

Fig. 1. The 95-kbp fragment analyzed, including PKLR gene and flanking regions. (a) Localization in chromosome 1q21; (b) localization of all mutations and polymorphisms (269T>A, 1456C>T, 1705A/C, (T)10/19, PKV, PKA, IVS11 and IVS3) genotyped in the present study.

other. No conspicuous differences seemed to exist regarding allelic frequencies or inferred haplotypes except that it is the only district when IVS3 observed heterozygosity was according to Hardy–Weinberg expected frequencies.

Regarding the studied Portuguese groups—PT-C and PT-PKD, IVS3 locus also presented the greatest diversity indices with the larger allele number and expected heterozygosity (Table 2). Observed heterozygosity was according to Hardy–Weinberg expected frequencies for all loci in the PT-C but not in PT-PKD. In this one, both IVS3 and IVS11 were significantly below the expected ($P=0.000$ and $P=0.002$, respectively). Again excluding IVS3 from the analysis, PT-C only showed LD for the closer loci (PKV/PKA and PKA/IVS11), while PT-PKD just had LD for PKV/IVS11 but since this latter, as IVS3, was not in Hardy–Weinberg equilibrium, we may say that no LD was observed between loci in this group.

When F_{ST} values were calculated for the two Portuguese groups, a significant value was obtained (0.025, $P=0.009$). When those were calculated for all the studied populations pairs, significant values ($P=0.000$) were obtained for all: CV-Total vs. PT-C—0.068 and vs. PT-PKD—0.111; CV-St Cruz vs. PT-C—0.111 and vs. PT-PKD—0.170; CV-I-St Catarina vs. PT-C—0.076 and vs. PT-PKD—0.122; CV-NI-St Catarina vs. PT-C—0.076 and vs. PT-PKD—0.124.

When groups of I and NI were analyzed separately, lower number of alleles was observed in I for all loci (IVS3: I—20, NI—25; IVS11: I—9, NI—11; PKA:I—10, NI—11) except for PKV (5 alleles in both groups) but this may be due to the smaller sample size of the I group (I—128, NI—376 alleles). Calculation of F_{ST} revealed no significant differences between the groups, both presenting the same most frequent alleles for all loci and no specific haplotypes being associated to any of them.

Yet, LD analysis revealed different results. While in the NI, as in overall population, a marked LD was observed between all pair of loci, in the I this effect was not found between the most distant loci—IVS11 and PKV. This could also be related with the smaller sample size of the I group, as it also happened in those districts with smaller sample size when were analyzed separately (St Cruz—46 alleles and Tarrafal—32 alleles). However, when I and NI from St Catarina, which have similar sample sizes (I—112 alleles and NI—118), were compared, the same was observed—a marked LD between all pair of loci in the overall population and NI alone and no linkage between IVS11 and PKV in the I. Besides, I and NI from St Catarina revealed no significant differences between them but IVS3 observed heterozygosity was according to Hardy–Weinberg expected frequencies in the I group.

Discussion

The study of malaria epidemiology is crucial for control, especially in countries like Cabo Verde where mosquito vectors are in close proximity to susceptible host populations and tourists. In Cabo Verde, we are addressing the three biological entities involved in the complex malaria life-cycle doing both parasitological [6,7] and entomological studies (on-going). The present study addresses some human host genetic polymorphisms in association to malaria.

Sickle cell disease affects millions of people worldwide and it is most common among people whose ancestors come from sub-Saharan Africa, India, Saudi Arabia and Mediterranean countries. Frequencies of the heterozygous state for the sickle cell gene (HbAS) range from 2% to 38% in sub-Saharan Africa where HbS allele frequencies frequently exceed 25% [14,16,40]. Sickle-cell trait is the best described host-specific factor shown to confer strong protection against *P. falciparum* in numerous studies over the course of the last 50 years [41,42,43,44].

Deficient G6PD alleles are distributed worldwide with a global prevalence of deficiency of 4.9% and an estimate of nearly 330 million people carrying a deficiency-associated mutation in the G6PD gene

Table 1
Diversity indices for the studied short tandem repeats in the Cabo Verde population.

Loci	Number of alleles	Heterozygosity		
		Observed	Expected	*P*-value
IVS3	26	0.825	0.927	0.000
IVS11	11	0.873	0.850	0.458
PKA	11	0.766	0.804	0.256
PKV	6	0.619	0.640	0.404

Table 2
Diversity indices for the studied short tandem repeats in the Portuguese groups.

Loci	PT-C				PT-PKD			
	Number of alleles	Heterozygosity			Number of alleles	Heterozygosity		
		Obs	Exp	*P*		Obs	Exp	*P*
IVS3	19	0.913	0.906	0.389	11	0.524	0.792	0.000
IVS11	9	0.738	0.682	0.636	5	0.476	0.708	0.002
PKA	8	0.488	0.512	0.162	4	0.143	0.139	1.000
PKV	5	0.588	0.601	0.697	4	0.476	0.580	0.294

PT-C: Portuguese healthy individuals; PT-PKD: Portuguese individuals with hereditary nonspherocytic hemolytic anemia (HNSHA) caused by PK-deficiency; Obs: Observed Heterozygosity; Exp: Expected Heterozygosity; *P*: *P*-value.

[45]. The highest prevalence is reported in Africa, southern Europe, the Middle East, Southeast Asia, and the central and southern Pacific islands; however, because of fairly recent migration, deficient alleles are nowadays quite prevalent in North and South America and in parts of northern Europe [19]. In most areas of high prevalence of G6PD deficiency, several polymorphic alleles are found but tropical regions of Africa are one exception, where the variant G6PD A^- accounts for about 90% of G6PD deficiency with frequencies of 5–25% [16,17]. The coincident worldwide distribution of malaria and mutated G6PD alleles made "The malaria/G6PD hypothesis" generally accepted [46]. Further evidence of protection against severe *P falciparum* malaria comes from both epidemiological studies [47] as well as from *in vitro* work [48,49].

PK deficiency along with G6PD deficiency, are the two most frequent enzyme disorders causing chronic hemolytic anemia worldwide. In families with no consanguinity, PK-deficient individuals are usually compound heterozygotes and prevalence of heterozygous individuals is estimated to be 1–2% in most studies [50]. The highest frequencies of the PK deficiency associated alleles are found in Europe and Asia with a prevalence ranging from 1% to 3.6% [26,33]. As these regions were historically endemic for malaria, it could have been responsible for maintaining this frequency or the ~180 mutations resulting in PK-deficiency are simply the product of random variation or other population genetic phenomena. However, in Africa, although the prevalence of PK deficiency is not known, the perception exists that it is rare, which may reflect a lack of testing [23]. If the marked *in vitro* protective effect of homozygosity for PK deficiency against malaria translates into the field (further supported by the murine model data), the argument that malaria has maintained the polymorphic frequency of the abnormal alleles may be plausible. In addition, the large number of PKLR mutations *per se* also suggests that these have been maintained by a selective force [23].

In the present study of the β globin chain of Hb, 6% of HbAS individuals and a frequency 0.05 of HbS allele are low values for a sub-Saharan region. Also G6PD deficiency associated mutations occurs in a very low frequency in this population (0.6%). Concerning PKLR polymorphisms, frequencies of alleles or haplotypes also differ from those described for African populations. Allelic frequencies of polymorphism 1705A/C [f(A) = 43%, f(C) = 57%] are closer to the European populations [f(A) ~ 29%, f(C) ~ 71%] than to Saharawi population from North Africa [f(A) ~ 62%, f(C) ~ 38%] or sub-Saharan populations [f(A) ~ 67%, f(C) ~ 33%] [25]. Allelic frequencies of the repeat (T)10/19 [f(10) = 53%, f(C) = 47%] are closer to the Portuguese population [f(10) ~ 78%, f(19) ~ 22%] than to São Tomé e Príncipe (Gulf of Guinea, West Africa) [f(10) ~ 36%, f(19) ~ 64%] [24]. Allelic frequencies of all these polymorphisms seem always be closer to the European, particularly to the Portuguese populations. The most frequent haplotypes 1705C/(T)10 and 1705A/(T)19, were the only two observed in Portugal and Central Europe [37]. However, the other two, 1705A/(T)10 and 1705C/(T)19 also occurred in low frequencies. As in São Tomé e Príncipe [24], there is a strong but not total association for combinations among these two biallelic systems.

Such low frequencies of traits HbS and G6PDMED are somehow unexpected. It could be due to the already well known importance of Caucasian admixture in the population of Cabo Verde [18] but these traits are quite prevalent in the Mediterranean region, an endemic region for malaria in the past. Further, Santiago Island should have had less contribution from Caucasians as demonstrated before in previous studies with mtDNA [51], Y-chromosome lineages [52] and autosomal STR [53].

Nevertheless, particular settlements with a strong African contribution to the genetic composition of the population seemed to persist as it may be the case of St Cruz. This district located in the east coast of the island showed F_{ST} values significantly different with all other studied populations both considering loci 1705A/C and (T)10/19 or STRs analysis. Moreover, allelic frequencies of the first two loci [f(A) = 57%, f(C) = 43%; fT(10) = 39%, fT(19) = 61%] were closer to the Saharawi population from North Africa and São Tomé e Príncipe (see above). Although we do not have historical reports about St Cruz or its capital Pedra Badejo (former Port of São Tiago), it is commonly said that the escaping slaves (Cabo Verde became an important provisioning station for slaves headed for the Americas) used to hide in this area, from where they could escape to the Island of Maio. This could justify such a stronger African contribution for the genetic background of this population but this should be further analyzed with more balanced sample sizes.

In the present study, no malaria related clinical data were available but regarding the infection status no association seems to occur with either the Hb β globin or the G6PD genotype. Also no haplotype or polymorphism of PKLR gene was associated to infected or non-infected individuals. Nevertheless some striking results related with PKLR analysis deserve a special remark. A linkage disequilibrium test revealed an association of distant loci only in non-infected individuals. This could mean a more conserved gene region in these individuals, which could happen if it would confer any protection against the infection and/or disease. Further, other peculiarities were found in the two groups. Infected individuals from St Catarina showed a significantly higher heterozygosity than expected in the locus T10/19 and on the opposite, it was the only group where IVS3 observed heterozygosity was within Hardy–Weinberg expected frequencies. Non-infected individuals from this district showed inverted allelic frequencies of the locus T10/19 comparing to the general trend and haplotypes 1705C/T10 and 1705A/T19 presented similar frequencies and 1705C/T19 showed higher frequency than in other studied groups. Further studies are needed to assess if these findings have a real biological meaning or are simply sampling artifacts.

Concluding remarks

This was the first study where data on sickle cell trait and G6PD deficiency frequencies were obtained for Cabo Verde human populations.

In this study no association was found between the analyzed human genetic factors and infection status of individuals. Three main reasons may have contributed for this: (1) the role of erythrocyte polymorphisms are usually associated and much easier demonstrated in severe than in mild or asymptomatic cases [54], (2) the cross-sectional sampling makes the infected/non-infected classification a faint case definition for an association study and 3) selective pressure of malaria, even if it had occurred, could never had a strong effect in this area due to its epidemic character.

Nonetheless, the finding of a very low frequency of G6PD deficiency associated alleles (A^- and MED) have important implications for the malaria control strategies defined by the National Program to Fight against Malaria (*Programa Nacional de Luta contra o Paludismo*, PNLP) viewing that it is recommended by WHO [55] that primaquine (potentially lethal in G6PD-deficient individuals) should be added to the drug regimen to block transmission in epidemic conditions such as Cabo Verde.

Regarding the PKLR gene, responsible for PK deficiency, recently reported as conferring protection against malaria in rodent and *in vitro* models, this study has not shown any clear association with malaria infection. Selective advantage afforded individuals protection from severe life-threatening complications of malaria and did not necessarily decrease their susceptibility to infection. Further, pyruvate kinase deficiency is a heterogeneous condition and most of the clinical phenotypes are mild or moderate in severity [26]. This suggests that the reproductive cost of PK deficiency was not limiting and mutations/polymorphisms would be spread in apparently healthy individuals.

Nevertheless, this is, to our knowledge, the first genetic population study about this putative association and results such as the region in linkage identified in the non-infected group deserve further investigation. Also, to further assess the assumption of a

protective effect of PK deficiency, further studies are being performed in other African populations from malaria highly endemic areas with well-defined malaria clinical cases (different severity level), well-characterized *Plasmodium*-infection and Hb β globin and G6PD status (to control for negative epistasis) and immediate enzymatic activity dosage at collection.

Acknowledgments

We are grateful to the population of Santiago Island, Cabo Verde who accepted to collaborate in this study. We thank the Health Delegates and technicians of Health Care Units of St Cruz, Tarrafal, St Catarina (especially Ana Veiga, Aníbal Monteiro, Antonino Monteiro and Edna Semedo) and Praia (especially Ernesto Cabral), Jorge de Pina (National Program against Malaria, Cabo Verde), Encarnação Horta and Marta Remédios (Institute of Hygiene and Tropical Medicine, Portugal) for technical assistance. We are also grateful to Doutor João Pinto for his participation in some sampling periods.

This study was supported by "Financiamento Programático do Laboratório Associado CMDT.LA/IHMT", POCI—Programa Operacional Ciência e Inovação 2010 (IPATIMUP) and POCI/SAU-ESP/55110/2004 (FCT/MCTES, Portugal). J. Alves and A.P. Arez were funded by FCT/MCTES Portugal (SFRH/BD/153451/2005 and SFRH/BPD/1624/2000—until 2007, respectively).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bcmd.2009.09.008.

References

[1] M.T.V. de Meira, T.S. Simões, J.F. Pinto Nogueira, Observações sobre sezonismo nas ilhas do Sal, Boa Vista e S. Nicolau (Cabo Verde), An. Inst. Med. Trop. (Lisb) IV (1947) 213-238.
[2] F.J.C. Cambournac, H. Santa Rita Vieira, M.A. Coutinho, et al., Note sur l'éradication du paludisme dans l'île de Santiago (Republique du Cap-Vert), An. Inst. Hig. Med. Trop. (Lisb) 10 (1984) 23–34.
[3] M. da Costa Monteiro, O sezonismo em Cabo Verde, An. Inst. Med. Trop. (Lisb) IX (1952) 461-483.
[4] J. Alves, Programme National de Lutte contre le paludisme. Plan d'Action 1994-98. Ministère de la Santé. République du Cap-Vert, 1994.
[5] WHO, International travel and health. Vaccination requirements and malaria situation. World Health Organization, 2005, http://whqlibdoc.who.int/publications/2005/9241580364_country_list.pdf (accessed 30 Jul 2009).
[6] A.P. Arez, G. Snounou, J. Pinto, et al., A clonal *Plasmodium falciparum* population in an isolated outbreak of malaria in the Republic of Cabo Verde, Parasitology 118 (1999) 347–355.
[7] J. Alves, A.L. Roque, P. Cravo, et al., Epidemiological characterization of *Plasmodium falciparum* in the Republic of Cabo Verde: implications for potential large-scale re-emergence of malaria, Malar. J. 5 (2006) 32.
[8] B. Carme, M.P. Hayette, A. Mbitsi, H. Moudzeo, J.C. Bouquety, *Plasmodium falciparum* index and level of parasitemia: diagnostic and prognostic value in the Congo, Ann. Soc. Belg. Med. Trop. 75 (1995) 33–41.
[9] I.M. Elhassan, L. Hviid, P.H. Jakobsen, et al., High proportion of subclinical *Plasmodium falciparum* infections in an area of seasonal and unstable malaria in Sudan, Am. J. Trop. Med. Hyg. 53 (1995) 78–83.
[10] J.M. Gonzalez, V. Olano, J. Vergara, et al., Unstable, low-level transmission of malaria on the Colombian Pacific Coast, Ann. Trop. Med. Parasitol. 91 (1997) 349–358.
[11] F.M. Leoratti, R.R. Durlacher, M.V. Lacerda, et al., Pattern of humoral immune response to *Plasmodium falciparum* blood stages in individuals presenting different clinical expressions of malaria, Malar. J. 7 (2008) 186.
[12] A.V. Hill, The genomics and genetics of human infectious disease susceptibility, Annu. Rev. Genomics Hum. Genet. 2 (2001) 373–400.
[13] F. Verra, V.D. Mangano, D. Modiano, Genetics of susceptibility to *Plasmodium falciparum*: from classical malaria resistance genes towards genome-wide association studies, Parasite Immunol. 31 (2009) 234–253.
[14] J. Makani, T.N. Williams, K. Marsh, Sickle cell disease in Africa: burden and research priorities, Ann. Trop. Med. Parasitol. 101 (2007) 3–14.
[15] M.J. Bouma, M. Goris, T. Akhtar, et al., Prevalence and clinical presentation of glucose-6-phosphate dehydrogenase deficiency in Pakistani Pathan and Afghan refugee communities in Pakistan; implications for the use of primaquine in regional malaria control programmes, Trans. R. Soc. Trop. Med. Hyg. 89 (1995) 62–64.
[16] A. Enevold, L.S. Vestergaard, J. Lusingu, et al., Rapid screening for glucose-6-phosphate dehydrogenase deficiency and haemoglobin polymorphisms in Africa by a simple high-throughput SSOP-ELISA method, Malar. J. 4 (2005) e61.
[17] M.A. Saunders, M. Slatkin, C. Garner, et al., The extent of linkage disequilibrium caused by selection on G6PD in humans, Genetics 171 (2005) 1219–1229.
[18] H. Spínola, J. Bruges-Armas, D. Middleton, A. Brehm, HLA polymorphisms in Cabo Verde and Guiné-Bissau inferred from sequence-based typing, Hum. Immunol. 66 (2005) 1082–1092.
[19] M.D. Cappellini, G. Fiorelli, Glucose-6-phosphate dehydrogenase deficiency, Lancet 371 (2008) 64–74.
[20] S.A. Tishkoff, R. Varkonyi, N. Cahinhinan, et al., Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance, Science 293 (2001) 455–462.
[21] G. Min-Oo, A. Fortin, M.F. Tam, et al., Pyruvate kinase deficiency in mice protects against malaria, Nat. Genet. 35 (2003) 357–362.
[22] K. Ayi, G. Min-Oo, L. Serghideset al, Pyruvate kinase deficiency and malaria, N. Engl. J. Med. 358 (2008) 1805–1810.
[23] P.M. Durand, T.L. Coetzer, Pyruvate kinase deficiency protects against malaria in humans, Haematologica 93 (2008) 939–940.
[24] L. Manco, A.L. Oliveira, C. Gomes, et al., Population genetics of four PKLR intragenic polymorphisms in Portugal and São Tomé e Princípe (Gulf of Guinea), Hum. Biol. 73 (2001) 467–474.
[25] E. Mateu, A. Perez-Lezaun, R. Martinez-Arias, et al., PKLR-GBA region shows almost complete linkage disequilibrium over 70 kb in a set of worldwide populations, Hum. Genet. 110 (2002) 532–544.
[26] A. Zanella, E. Fermo, P. Bianchi, et al., Pyruvate kinase deficiency: the genotype-phenotype association, Blood Rev. 21 (2007) 217–231.
[27] C. Alves, V. Gomes, M.J. Prata, et al., Population data for Y-chromosome haplotypes defined by 17 STRs (AmpFlSTR YFiler) in Portugal, Forensic Sci. Int. 171 (2007) 250–255.
[28] L. Manco, M.L. Ribeiro, H. Almeida, et al., PK-LR gene mutations in pyruvate kinase deficient Portuguese patients, Br. J. Haematol. 105 (1999) 591–595.
[29] L. Manco, M.L. Ribeiro, V. Máximo, et al., A new PKLR gene mutation in the R-type promoter region affects the gene transcription causing pyruvate kinase deficiency, Br. J. Haematol. 110 (2000) 993–997.
[30] C.M. Waterfall, B.D. Cobb, Single tube genotyping of sickle cell anaemia using PCR-based SNP analysis, Nucleic Acids Res. 29 (2001) e119.
[31] J. May, C.G. Meyer, L. Gro terlinden, et al., Red cell glucose-6-phosphate dehydrogenase status and pyruvate kinase activity in a Nigerian population, Trop. Med. Int. Health 5 (2000) 119–123.
[32] W.W. van Solinge, R. van Wijk, R.J. Kraaijenhagen, et al., Novel mutations in the human red cell type pyruvate kinase gene: two promoter mutations in cis, a splice site mutation, a nonsense- and three missense mutations, Blood 90 (1997) 1197.
[33] E. Beutler, T. Gelbart, Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population, Blood 95 (2000) 3585–3588.
[34] L. Manco, A. Abade, Pyruvate kinase deficiency: prevalence of the 1456 C→T mutation in the Portuguese population, Clin. Genet. 60 (2001) 472–473.
[35] C. Lenzner, P. Nurnberg, B.-J. Thiele, et al., Mutations in the pyruvate kinase L gene in patients with hereditary hemolytic anemia, Blood 83 (1994) 2817–2822.
[36] H. Kanno, H. Fujii, A. Hirono, et al., Identical point mutations of the R-type pyruvate kinase (PK) cDNA found in unrelated PK variants associated with hereditary hemolytic anemia, Blood 79 (1992) 1347–1350.
[37] C. Lenzner, P. Nurnberg, G. Jacobasch, B.-J. Thiele, Complete genomic sequence of the human PK-L/R-gene includes four intragenic polymorphisms defining different haplotype backgrounds of normal and mutant Pk-genes, DNA Seq. 8 (1997) 45–53.
[38] C. Lenzner, G. Jacobash, A. Reis, et al., Trinucleotide repeat polymorphism at the PKLR locus, Hum. Mol. Genet. 3 (1994) 523.
[39] L. Excoffier, G. Laval, S. Schneider, Arlequin ver. 3.0: An integrated software package for population genetics data analysis, Evol. Bioinform. Online 1 (2005) 47–50.
[40] B. Modell, M. Darlison, Global epidemiology of haemoglobin disorders and derived service indicators, Bull. WHO 86 (2008) 480–487.
[41] D.A. Hassan, A.P. Arez, H.S. Mohamed, et al., Reduced sequestration of *P. falciparum* infected erythrocytes among sickle cell trait malaria patients: an *in vivo* evidence among a West African tribe in Sudan, Ann. Trop. Med. Parasitol. 102 (2008) 743–748.
[42] M. Aidoo, D.J. Terlouw, M.S. Kolczak, et al., Protective effects of the sickle cell gene against malaria morbidity and mortality, Lancet 359 (2002) 1311–1312.
[43] J.R. Aluoch, Higher resistance to *Plasmodium falciparum* infection in patients with homozygous sickle cell disease in western Kenya, Trop. Med. Int. Health 2 (1997) 568–571.
[44] J. Miranda, N. Gonçalves, I. Picanço, et al., Sickle-cell trait and red cell glucose-6-phosphate dehydrogenase status and malaria morbidity in Angola, Proc. 5th Eur. Congr. Trop. Med. Int. Health (2007) 65–68.
[45] E.T. Nkhoma, C. Poole, V. Vannappagari, et al., The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis, Blood Cells Mol. Dis. 42 (2009) 267–278.
[46] L. Luzzatto, U. Bienzle, The malaria/G6PD hypothesis, Lancet 1 (1979) 1183–1184.
[47] C. Ruwende, S.C. Khoo, R.W. Snow, et al., Natural selection of hemi and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria, Nature 376 (1995) 246–249.
[48] M. Cappadoro, G. Giribaldi, E. O'Brien, et al., Early phagocytosis of Glucose-6-Phosphate Dehydrogenase (G6PD)-deficient erythrocytes parasitized by *Plasmodium falciparum* may explain malaria protection in G6PD deficiency, Blood 92 (1998) 2527–2534.
[49] E.F. Roth Jr., M.C. Calvin, I. Max-Audit, et al., The enzymes of the glycolytic pathway in erythrocytes infected with *Plasmodium falciparum* malaria parasites, Blood 72 (1988) 1922–1925.

[50] I. Max-Audit, Anémie hémolytique due à un déficit en pyruvate kinase du globule rouge. Orphanet. http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=FR&Expert=766.0, 2001 (accessed 31 Jul 2009).
[51] A. Brehm, L. Pereira, H.J. Bandelt, A. Amorim, Mithocondrial portrait of the Cabo Verde archipelago: the Senegambian outpost of Atlantic slave trade, Ann. Hum. Genet. 66 (2002) 49–60.
[52] R. Gonçalves, A. Rosa, A. Freitas, et al., Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers, Hum. Genet. 113 (2003) 467–472.
[53] A.T. Fernandes, R. Velosa, J. Jesus, et al., Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms, Ann. Hum. Genet. 67 (2003) 340.
[54] F. Migot-Nabias, S. Pelleau, L. Watier, et al., Red blood cell polymorphisms in relation to *Plasmodium falciparum* asymptomatic parasite densities and morbidity in Senegal, Microbes Infect. 8 (2006) 2352–2358.
[55] WHO, Malaria epidemics: prediction, preparedness and control. WHO Expert Committee on Malaria. Twentieth report, http://apps.who.int/malaria/docs/ecr20_7.htm, 1998 (accessed 03 Jul 2009).