UNIVERSITY OF SÃO PAULO
Ribeirão Preto Medical School
Basic and Applied Immunology Post-graduation Program (IBA)

Paulin Sonon

Plasmodium falciparum and *Schistosoma haematobium* co-infection: role of non-classical HLA class I genes (*HLA-G, -E* and *-F*) in susceptibility to malaria

RIBEIRÃO PRETO
2018

Paulin Sonon

Plasmodium falciparum and *Schistosoma haematobium* co-infection: role of non-classical HLA class I genes (*HLA-G, -E* and *-F*) in susceptibility to malaria

Doctorate thesis presented at Ribeirão Preto Medical School/University of São Paulo to obtain the PhD grade in Sciences.

Area of Concentration: Basic and Applied Immunology

Advisor: Professor Eduardo Antônio Donadi, M.D, PhD.

Foreign collaborators: Dr. David Courtin, PhD, HDR and Professor Kabirou A. Moutairou, PhD.

RIBEIRÃO PRETO

2018

Publication Cataloging
Ribeirão Preto Medical School
University of São Paulo

Paulin Sonon

Co-infecção *Plasmodium falciparum* e *Schistosoma haematobium*: papel dos genes HLA não-clássicos de classe I (*HLA-G, -E* e *-F*) na suscetibilidade à malária

***Plasmodium falciparum* and *Schistosoma haematobium* co-infection: role of non-classical HLA class I genes (*HLA-G, -E* and *-F*) in susceptibility to malaria.**

Ribeirão Preto, 2018.

Número de páginas p.200: il.; 30 cm

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP para obtenção do título de Doutor em Ciências

Área de concentração: Imunologia Básica e Aplicada

Orientador: Profº. Drº. Eduardo Antônio Donadi.
Colaboradores estrangeiros: Drº. David Courtin, e Drº Kabirou A. Moutairou

1. Genes *HLA Ib;* 2. *P. falciparum*; 3. *S. haematobium;* 4. sequenciamento de nova geração; 5 Africano.

SONON, P. ***Plasmodium falciparum*** **and** ***Schistosoma haematobium*** **co-infection: role of non-classical HLA class I genes (*HLA-G, -E* and *-F*) in susceptibility to malaria.** Doctorate thesis presented at Ribeirão Preto Medical School/University of São Paulo to obtain the PhD grade in Sciences.

Approved date: _____/_____/_____

Jury

Prof. Dr._____Institution_____
(President)
Judgment: _____Signature:_____

Prof. Dr._____Institution_____
(Member)
Judgment: _____Signature:_____

Prof. Dr._____Institution_____
(Member)
Judgment: _____Signature:_____

Prof. Dr._____Institution_____
(Member)
Judgment: _____Signature:_____

DEDICATIONS

To:

God thrice Holy…

My Father; **Samuël Sonon…**

My mother; **Koukpétodji Faton…**

My Brothers and Sisters, especially **Janvier Sonon…**

My Wife; **Kèdétin Célestine Kouton…**

My Son; **Fènou Espoir Milagre Sonon…**

My daughter; **Mahounan Fleur Gracias Sonon…**

ACKNOWLEDGEMENTS

To My Advisor Professor Dr. Eduardo Antônio Donadi. *I remember of this day when you informed me about this doctorate opportunity. I was very happy. You gave me the opportunity to achieve a great dream of my life. Thank you for trusting and receiving me in your Laboratory. During my doctorate, here in Brazil, you were not only my Advisor but also my Adoptive Father. Sincerely thank you Professor.*

To Professor Moutairou Kabirou. *Your support to my career (formation) has begun since my Undergraduation and Master degree in ''Immunology and Cell Biology program'' that you headed. And this love for me has continued up to the end of my PhD thesis and never will stop. Thank you, Professor.*

To Dr. David Courtin. *You also is a precious stone of this building that I am. Your presence in my life had still been very important and always had been welcome. I remember of my Master training with you: your simplicity, your helps, your desir to do to myself a great person. You only did not show this (this lovely personality towards myself) during my PhD thesis but you also confirmed this. Thank you very much, Dr. David.*

To Dr. Audrey Sabbagh and Dr. Jacqueline Milet. *For your technical supports, even at distance. This doctorate project would be incomplete without you. Your corroborations had been very useful for my work. Thanks.*

To Dr. Erick C. Castelli and Dr. Celso T. Mendes-Júnior. *For you, I miss words to express really what I fill and receive from you. I thank you specially for your disponobilities. Very thank again to you to have training me for the Next Generation Sequencing Dataset analysis. To Dr. Erick C. Castelli, thank you for your help, even at distance, in this new area. I learned much from you.*

To Dr. Ibrahim Sadissou and his Wife Sally. *You always have been my Angel. I owe you this thesis. Your advice, your helps in various fields gave me the force to accomplish this mission. Thanks.*

Dr. Juliana Doblas Massaro. *It was you who enigmated me, and gave me the basis and the desire to learn bioinformatics. Thank you for your help.*

Laboratory of Molecular Biology. *Specially Dr. Ibrahim A. Sadissou, Dr. Juliana Doblas Massaro, M.Sc Bruna Cristina Bertol, our technicians (Flavia T. de A. Vieira and Sandra S. R. dos santos). Your advices, technical supports, and specifically, your friendship have motivated me to continue, to believe and to never give up. Thank you very much.*

To our pacients and their parents. *I thank the school children and their parents for their free consent and the teachers of the Sô-Ava district for their great collaboration.*

To our collaborators

- **Benin: Professor Moutairou A. Kabirou**, Université d'Abomey-Calavi (UAC/Bénin), **Dr. Ibikounlé Moudachirou**, Université d'Abomey-Calavi (UAC/Bénin) and **Professor Dr. Massougbodji Achille**, Université d'Abomey-Calavi (UAC/Bénin).
- **France: Dr. David Courtin**, IRD-MERIT-UMR216, Paris, France, Université Paris Descartes, 4, Paris, France, **Dr. Audrey Sabbagh**, IRD-MERIT-UMR216, Paris, France, Université Paris Descartes, 4, Paris, France, **Dr. Philippe Moreau**, Hôpital Saint-Louis, Paris, France and **Dr André Garcia,** IRD-MERIT-UMR216, Paris, France, Université Paris Descartes, 4, Paris, France
- **Brazil: Dr. Eduardo A. Donadi**, Universidade de São Paulo (USP), Ribeirão Preto - SP, Brazil, **Dr. Erick C. Castelli**, UNESP-Botucatu and **Dr. Celso T. Mendes-Júnior**, Universidade de São Paulo (USP), Ribeirão Preto - SP, Brazil.

Who worked together, technically for the completion of this project. Thanks.

To Castelli's lab at UNESP-Botucatu. *Specially Dr. Erick C. Castelli, Dra. Luciana C. Veiga-Castelli, Dra. Iane de Oliveira Pires Porto, M.Sc Andreia da Silva Souza, M.Sc Thálitta Hetamaro Ayala Lima, MsC. Michelle Almeida da Paz and Neilton Paulo Bezerra. For the Next Generation Sequencing (NGS) training you gave me at the beginning and during my doctorate project.*

To our financial supports: Institut de Recherche pour le Développement (IRD)/France (JEAI PALUCO), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Brazil (Grant numbers PVE406 594/2013-9, 304 931/2014-1 and 466036/2013-5) and Brazil-France Research Cooperation Program USP/COFECUB (Grant numbers Uc Me 169-17). *For your financial supports that were very useful for the achievement of this work. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code CAPES/PROEX0055042/1768/2013.*

To the Basic and Applied Immunology program. *For the knowledge you gave me that will serve me for all my life. Thank you. I can not forget Ana Cristine Silva Ferreira, our Secretary, for her mother kiss. I thank also all my friends from Applied Immunology program.*

To the Sisters of Vita e Pax. Sister Maria Lúcia dos Santos, Sister Marina Domingues dos Santos and coleagues. *For the hospitality and spiritual supports.*

To all my friends from Santa Tereza D'Avila Parish. *Priest Josirley, Bernardo, Hilton, Giulliano and all my friends. For your spiritual supports.*

Oração de um jovem estudante

Ó Espírito Santo, Amor que procede do Pai e do Filho, Sabedoria divina e Fonte de todo saber! Iluminai-me nesta fase de minha vida, em que ainda sou um estudante, quando tenho a responsabilidade de preparar-me intelectualmente para colocar ao servisse de todo o que tenho o privilégio de estar aprendendo agora. Permita-me Senhor, vencer as dificuldades próprias desta etapa da minha vida, sem brincar com minhas responsabilidades, mas também sem perder o dom de me descontrair e me alegrar, como é próprio do jovem cristão. Ilumina também os meus colegas para que amanhã estejamos todos a serviço da humanidade. Amém.

Pe. Luiz Carlos do Nascimento

Prayer of a young student

Ó Holy Spirit, Love that proceeds from the Father and the Son, divine Wisdom and Source of all knowledge! Enlighten me at this stage of my life, when I am still a student, when I have the responsibility to prepare myself intellectually to put myself in the service of all that I have the privilege of learning now. Allow me, Lord, to overcome the difficulties of this stage of my life, without playing with my responsibilities, but also without losing the gift of relaxing and rejoicing, as is the case of a young Christian. Enlightens also my colleagues so that, tomorrow we will be all in the service of humanity. Amen.

Pe. Luiz Carlos do Nascimento

Resumo

SONON, P. **Co-infecção *Plasmodium falciparum* e *Schistosoma haematobium*: papel dos genes HLA não-clássicos de classe I (*HLA-G, -E* e *-F*) na suscetibilidade à malária.** 2018. 200f. Tese (Doutorado direito) Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2018.

A malária causada pelo *Plasmodium falciparum* (*P. falciparum*) e a bilharzíase urogenital causada pelo *Schistosoma haematobium* (*S. haematobium*) constituem duas doenças infecciosas tropicais alarmantes, sendo ambas endêmicas no Benin. Considerando que a malária (Th1) e a esquistossomose (Th2) apresentem perfis de citocinas divergentes, na presença de co-infecção, o *S. haematobium* poderia modular as respostas especificamente dirigidas contra o *P. falciparum*. Uma vez que, os genes que codificam as moléculas não-clássicas de histocompatibilidade (HLA-G/-E/-F) possuam propriedades imunomoduladoras, pouca atenção tem sido dedicada ao estudo desses genes em populações da África Subsaariana, que são as de maior diversidade genética. O objetivo deste estudo foi avaliar a variabilidade dos genes *HLA-G/-E/-F* na população Toffin do Benin, e identificar fatores genéticos de suscetibilidade/resistência à malária causada pelo *P. falciparum* e/ou bilharziose urogenital, usando uma coorte de crianças (de 4 a 8 anos de idade) não aparentadas. Foram avaliados os segmentos codificantes e reguladores, englobando aproximadamente 5.1 kb do *HLA-G*, 7.7 kb do *HLA-E* e 6.2 kb do *HLA-F*, usando sequenciamento da nova geração. As frequências alélicas e haplotípicas do *HLA-G/-E/-F*, a diversidade genética, a diversidade nucleotídica, o equilíbrio de Hardy-Weinberg (HW) e de Tajima's D foram realizados utilizando o software ARLEQUIN v3.5.2. O desequilíbrio de ligação (LD) entre sítios variáveis com frequência alélica mínima (MAF) acima de 1% foi avaliado calculando o coeficiente de correlação r^2 e os gráficos de LD foram visualizados usando o software Haploview 4.2. O estudo de associação foi implementada nos softwares PLINK v1.90b4.6 e R v3.4.2, usando modelos de regressão linear ou logística múltipla. 96, 37 e 68 sítios variáveis foram detectados ao longo de *HLA-G/-E- F*, respectivamente. Foram identificados 16, 19 e 15 haplótipos da promotora, 19, 15 e 29 haplótipos da codificadora, 12, 7 e 2 haplótipos da região 3' não traduzida (3'UTR), respectivamente, e ainda, 33 , 31 e 32 haplótipos estendidos, respectivamente. Todos os haplótipos promotores/codificadores/3'UTR respeitaram os padrões já descritos na população mundial. O *HLA-E* foi o mais conservado, exibindo principalmente duas proteinas (E*01:01 e E*01:03), seguido do *HLA-F*, três proteínas (F*01:01, F*01:02 e F*01:03) e *HLA-G*, quatro proteínas: três normais (G*01:01, G*01:03 e

G*01:04) e uma truncada (G*01:05N). Embora os alelos do *HLA-G-E/-F* observados na população Toffin tenham sido os mais frequentemente observados em vários países do mundo, as frequências dos haplótipos da região codificadora foram semelhantes às descritas para outras populações africanas (Guiné-Conakry e Burkina-Faso), quando comparadas com os países não-Africanos (Brasileiros), indicando que as variações ao longo desses genes estavam presentes nos Africanos antes da dispersão humana. Foram analisados 105 polimorfismos (MAF> 5%, valor *P* de HW> 0.05 e qualidade de genótipos> 97%) e 56 haplótipos com frequência mínima de 5%. Consideramos significativos, apenas os resultados exibindo valores de *P* <0.01 para polimorfismos e valores de *P* <0.01 antes da correção e valores de *P* <0.05 após correção de Bonferroni, para haplótipos. Encontramos as seguintes associações: **i)** o alelo inserção de 14 pares de bases do *HLA-G* (14 pb Ins) (em modelo dominante) foi associado ao risco de ocorrência de infecção por *P. falciparum* (infecções totais como infecções sintomáticas) e ao número de episódios de infecção (número elevado de episódios de infecções totais como de infecções sintomáticas), **ii)** polimorfismos *HLA-G* (-1155 A e +755 A, em completo LD) (em modelo recessivo), e **iii)** haplótipo *E*.01.03.05-compatível em sinergia com o alelo -1988 C do *HLA-E* (em modelo aditivo) foram associados à proteção contra malária (em níveis de infecção como de densidade parasitária), e **iv)** o haplótipo *E*.Promo.2 foi associado à protecção contra a co-infecção do *P. falciparum* em crianças infectadas por *S. haematobium*. Excetuando o 14 pb Ins, estudos funcionais adicionais são necessários para revelar o papel desses marcadores na expressão dos genes *HLA-G, -E* e *-F*, para entender melhor o mecanismo de associação com doenças.

Palavras-chave: genes *HLA Ib, P. falciparum, S. haematobium*, sequenciamento de nova geração, Africano.

Abstract

SONON, P. ***Plasmodium falciparum* and *Schistosoma haematobium* co-infection: role of non-classical HLA class I genes (*HLA-G, -E* and *-F*) in susceptibility to malaria.** 2018. 200p. Thesis (Doctorate) Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2018.

Plasmodium falciparum (*P. falciparum*) malaria and the urogenital bilharziasis caused by *Schistosoma haematobium* (*S. haematobium*) infection constitute the two alarming tropical infectious diseases; and both are endemic in Benin. Considering that malaria (Th1) and schistosomiasis (Th2) have divergent cytokine profiles, the presence of co-infection could modulate the responses specifically directed against *P. falciparum*. We hypothesize that the non-classical genes and molecules (HLA-G, -E and -F) of major histocompatibility complex (MHC) could be involved in this immunomodulation. Although these molecules have well known immunomodulatory properties, little attention has been devoted to the study of these non-classical class I *HLA* genes in sub-Saharan African populations. This study aimed to evaluate the diversity of *HLA-G, -E* and *-F* gene variable sites in the Beninese Toffin population, and to identify susceptibility/resistance genetic factors associated with *P. falciparum* malaria and/or urogenital bilharziasis in a Beninese cohort of unrelated school-aged children (4 to 8 years old). We evaluated the complete gene variability (5.1 kb for *HLA-G*, 7.7 kb for *HLA-E* and 6.2 kb for *HLA-F*) in the Beninese Toffin population using massive parallel sequencing. *HLA-G, -E* and *-F* allele and haplotype frequencies, gene diversity, average nucleotide diversity, Tajima's D and Hardy–Weinberg (HW) equilibrium were performed using ARLEQUIN v3.5.2 software. The linkage disequilibrium (LD) pattern among variable sites with a minimum allele frequency (MAF) above 1% was evaluated computing the correlation coefficient r^2 and the LD plots were visualized using Haploview 4.2 software. The genetic association study was implemented on PLINK v1.90b4.6 and R v3.4.2 softwares using multiple logistic or linear regression models. Overall, 96, 37 and 68 variable sites were detected along the entire *HLA-G, -E* and *-F*, respectively, arranged into region-specific haplotypes; i.e., promoter haplotypes (16, 19, and 15, respectively), coding haplotypes (19, 15, and 29, respectively), 3' untranslated region (3' UTR) haplotypes (12, 7 and 2, respectively) and extended haplotypes (33, 31 and 32, respectively). All promoter/coding/3'UTR haplotypes followed the patterns already described in worldwide populations. Among the three genes, *HLA-E* was the most conserved, exhibiting mainly two full-length encoded-molecules (E*01:01 and E*01:03), followed by *HLA-F* (three full-length

proteins; F*01:01, F*01:02 and F*01:03) and *HLA-G* (four proteins; three full-length (G*01:01, G*01:03 and G*01:04) and one truncated (G*01:05N)). Although *HLA-G/E/F* alleles in the Toffin population were the most frequently observed worldwide, the frequencies of the coding haplotypes were closely similar to those described for other West African populations (Guinea-Conakry and Burkina-Faso) than for non-African ones (Brazilian population), indicating that variable sites along these genes were present in Africa before human dispersion. A total of 105 polymorphisms and 56 haplotypes were analyzed in the genetic association study to malaria and schistosomiasis susceptibility. Only results exhibiting respectively *P*-values < 0.01 and *P*-values < 0.05 (after Bonferroni correction) for polymorphisms and for haplotypes were considered as significant. The following associations were observed: **i)** *HLA-G* 14 base pairs (14bp) insertion was associated (under dominant model) with the risk of the occurrence of *P. falciparum* infection (all infections and symptomatic infections) and with high number of infection episodes (all infections and symptomatic infections)**, ii)** *HLA-G -1155 A* and *+755 A* alleles (under recessive model) and **iii)** *E*.01.03.05-compatible haplotype in synergy with *HLA-E -1988 C* allele (under additive model) were associated with protection against *P. falciparum* (at infection as well as parasitic levels), and finally **iv)** the *E*.Promo.2 haplotype was associated with protection against malaria-schistosomiasis co-infection. The functional role of the genetic markers (alleles or haplotypes) associated with malaria and schistosomiasis suceptibility/resistance need to be investigated to better understand the mechanism that may explain these associations.

Keywords: *HLA Ib genes; P. falciparum; S. haematobium;* massive parallel sequencing; African

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

HLA-G/E/F	Human leukocyte antigen *(G, E, F)*
UTR	Untranslated region
MHC	Human major histocompatibility complex
DCs	Dendritic cells
NK	Natural killer cells
ILT	Immunoglobulin-like transcript receptor
DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
qPCR	Quantitative real-time PCR
BAM	Binary alignment map
IGV	Integrative genomics viewer
VCF	Variant call format
GATK	Genome analysis toolkit
LD	Linkage disequilibrium
MAF	Minor allele frequency
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
P. f	*Plasmodium falciparum*
Kb	Kilobases (10^3 bases)
mRNA	Messenger RNA (ribonucleic acid)
SINE	**S**hort interspersed nuclear element
S. haematobium	*Schistosoma haematobium*
WHA	World Health Assembly
HIV	Human immunodeficiency virus
AIDS	Acquired immunodeficiency syndrome
EIR	Entomological inoculation rate
HbS	Hemoglobin S genes
WHO	World health organization
LLIN	Long-lasting insecticidal net
IRS	Indoor residual spraying
GPIRM	Global Plan for Insecticide Resistance Management in Malaria Vectors
RDT	Rapid diagnostic test
TBS	Thick and thin blood smears/films
ACT	Artemisinin-based combination therapy
IL-4	Interleukins
Th	T helper
MOH	Ministry of Health
HMIS	Health Management Information System
ARG1	Arginase 1 transcription factor
MRC	Mannose receptor C-type transcription factor
CHI3L3	Chitinase-3-like protein 3 transcription factor
NOS2	Nitric oxide synthase 2

CD	Cluster of differentiation. It is a costimulatory protein found on immune cells and is required for their activation
mDC	Myeloid DC cells
HLA-DR	Human leukocyte antigen class II
STAT6	Signal transducer and activator of transcription 6 gene
CTLA4	Cytotoxic T-lymphocyte–associated antigen 4 gene
NKT	Natural killer T cells
Ig	Immunoglobulin
SPECT	Spororozoite microneme protein essential for cell traversal also called perforin-like protein or PLP
IFN-γ	Interferon-gamma
CSP	Circumsporozoite protein
ADCC	Antibody-dependent cell-mediated cytotoxicity
RBC	Red blood cell
TGF-β	Transforming growth factor beta
AMA-1	Apical membrane antigen 1
MSP-1	Against merozoite surface protein 1
HAT	African trypanosomiasis
URR	Upstream regulatory region
CDS	Coding sequence
GLURP	Glutamate Rich Protein
MSP2	Merozoite Surface Protein-2
NGS	Next Generation Sequencing
TNF-α	Tumor necrosis factor
sHLA-G	Soluble HLA-G
MV	Monthly visit
CU	Emergency consult
µL	Micro liter (10^{-3} Liter)
dNTP	Deoxynucleotide tri-phosphate
$MgCl_2$	Magnesium Chloride
SNV	Single Nucleotide Variation
ISRE	Interferon-stimulated response element
HRE	Heat shock element
LCR	Locus control region

TABLE OF CONTENTS

1. INTRODUCTION

1.1. HUMAN SCHISTOSOMIASIS

1.1.1. Definition

Schistosomiasis, or Bilharzia, is a parasitic disease caused by *trematode helminths* (parasitic worms) that was described for the first time by Theodor Bilharz, in 1852 (Bennett et al. 2015). Helminths are multicellular eukaryotic invertebrates with tube-like or flattened bodies exhibiting bilateral symmetry. The major groups of parasitic helminths include nemathelminths (roundworm) and plathelminths (flatworms). The latter subdivided into cestodes (tapeworms) and trematodes (flukes) (Castro 1996; Cooper et al. 2017). Schistosomiasis is caused specifically by one of the six different species of trematode flatworms of the genus *Schistosoma* **(Figure 1).** Chronologically, the *Schistosoma (S.) haematobium* (Bilharz, 1852) was the first detected, followed by *S. guineensis* (Weinland, 1858), *S. japonicum* (Katsurada, 1904), *S. mansoni* (Sambon, 1907), *S. intercalatum* (Fisher, 1934), and *S. mekongi* (Voge, Bruckner & Bruce, 1978) (Sanogo et al. 2018). In addition, other *Schistosoma* species can be considered such as *S. malayensis* (1988) found and described in Malaysia (Greer et al. 1988) and *S. mattheei*, a zoo-parasite that infects baboon and bovid in Zambia, but can also parasite humans (Weyher et al. 2010). It is very important to highlight the emergence of natural schistosome hybrids [*S. haematobium–S. guineensis* (Webster et al. 2005), *S. haematobium–S. intercalatum* (Southgate et al. 1976; Webster et al. 2003), and *S. haematobium* and the cattle schistosome *S. bovis* (Boissier et al. 2016; Kincaid-Smith et al. 2017)], which may also infect human. Hybrids (*S. haematobium-S. bovis*) may also produce severe clinical signs in humans, rendering more difficult the control and eradicate the disease due to the presence of a potential additional animal reservoir and additional pathogeneic genes.

Most of the registered cases were often due to infection by either *S. haematobium*, *S. japonicum* or *S. mansoni* (Colley et al. 2014; Sanogo et al. 2018). Indeed, we can distinguish urinary Bilharzia (*S. haematobium*) from intestinal (*S. mansoni* and *S. japonicum*) or hepatic bilharzia (*S. mansoni*), depending on the causative agent (Colley et al. 2014; Sanogo et al. 2018).

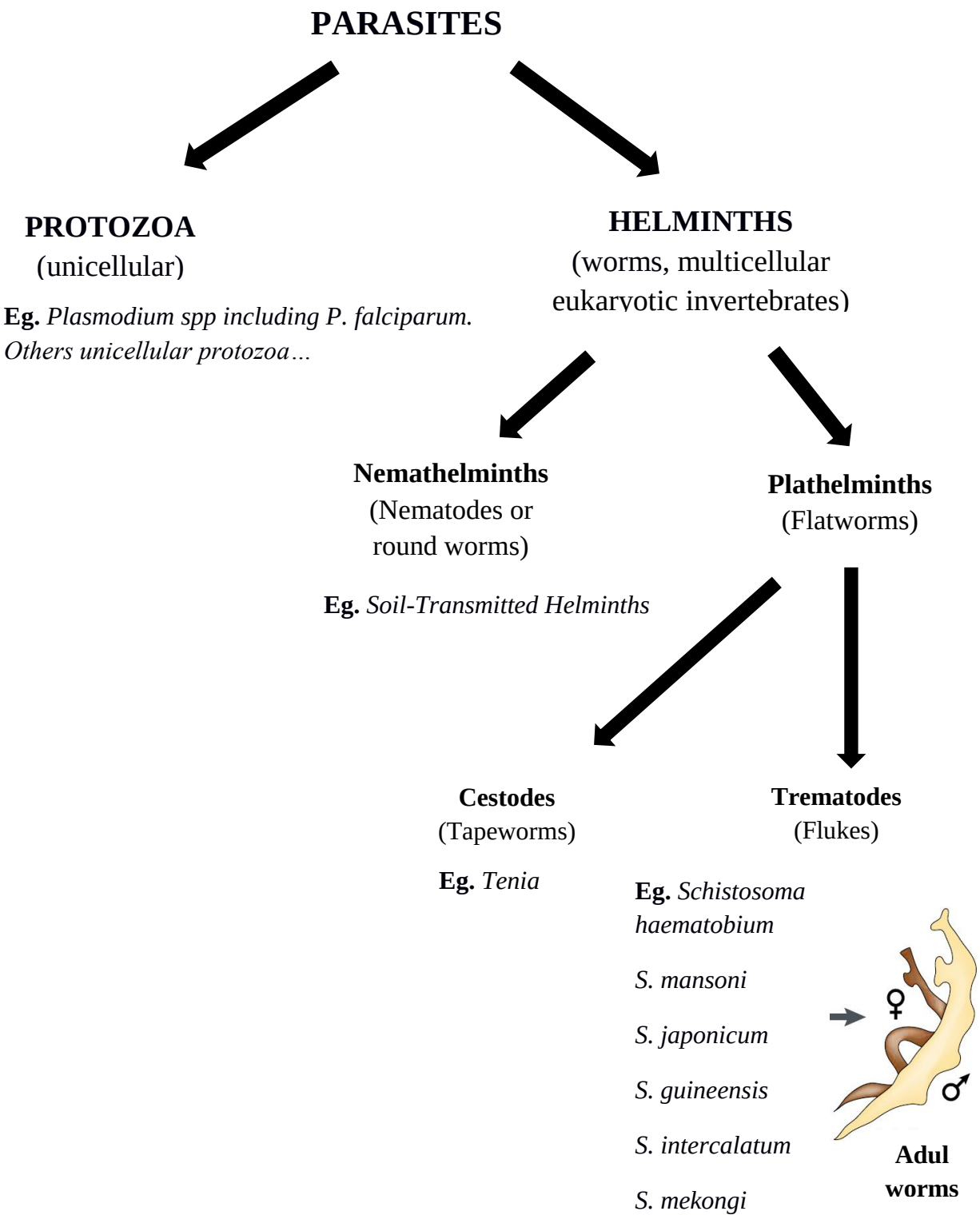


Figure 1: *Simplified taxonomy of Helminths (Example of Schistosoma spp).*

1.1.2. Localization

The human-pathogenic schistosome species are worldwide distributed; however, the African continent harbors almost all the existing schistosome species and even the mixed cases of *S. haematobium* and *S. mansoni infections* **(Figure 2)** (Colley et al. 2014). The sub-

Saharan African countries carry the heaviest burden of parasitic disorders reaching 90% of all cases occurring in the world (Colley et al. 2014; Lai et al. 2015). *S. mansoni* is also found in Latin America, specifically in Brazil (Scholte et al. 2014), while in the People's Republic of China, the Philippines and three small Indonesian *foci* harbor *S. japonicum* (Utzinger et al. 2010). *S. guineensis* and *S. intercalatum* are present along the Congo River and in lower Guinea (in Africa) (Tchuem Tchuenté et al. 2003), whereas the *S. haematobium* is highly endemic in the White Nile River basin, Sudan (Africa) (Ismail et al. 2014). *S. mekongi* is encountered around the border between Cambodia and Lao People's Democratic Republic (Muth et al. 2010).

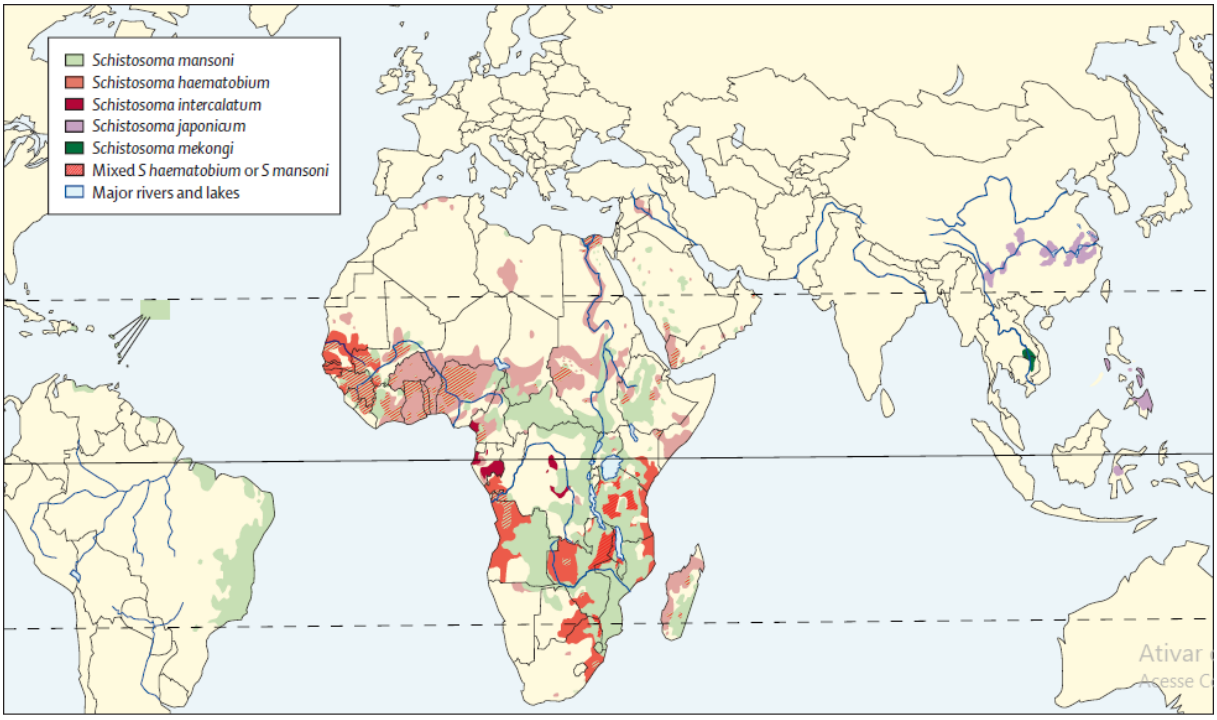


Figure 2: *Global distribution of human schistosomiasis* (Colley et al. 2014). *Adapted from Gryseels et al.* (Gryseels et al. 2006b).

1.1.3. Transmission

Schistosomiasis is often transmitted by body-contact to infected freshwater (contaminated either by urine or by feces), where the intermediate snail host is present (*Biomphalaria, Bulinus* or *Oncomelania spp* snails) (Gryseels and Strickland 2012; Colley et al. 2014; Sanogo et al. 2018). *S mansoni, S haematobium and S japonicum* use freshwater *Biomphalaria*, *Bulinus*, and *Oncomelania spp* snails, respectively, as their intermediate hosts (Gryseels and Strickland 2012; Colley et al. 2014; Sanogo et al. 2018).

The eggs that come from urine (in the case of *S haematobium*) hatch and release miracidia, which swim in fresh water, look after and penetrate in their intermediate snail host,

where they develop into sporocysts. After two-four weeks, the sporocysts become cercariae (Gryseels and Strickland 2012), which can leave the snail under the stimulation of light and swim until they find their definitive mammalian hosts. At least hundreds of the fork-tailed cercariae leave the snail, by using mechanical activity and proteolytic enzymes (Gordon et al. 2012), and enter the skin of the definitive mammalian hosts, when they lose their tails. Cercariae develop into schistosomules that migrate via the blood or lymphatic vessels (Gryseels and Strickland 2012) to the portal circulation, where they mature into adult worms. Depending on the schistosome species, the adult worms can stay in the portal and mesenteric vessels (adult *S. japonicum, S. mekongi* and *S. intercalatum* worms) or in the vesical plexus (adult *S. haematobium* worms). To mate, the male adult schistosome embrace the female worm into its gynaecophoric canal **(Figure 3)** (Gordon et al. 2012). Four to six weeks separate the period of infection (i.e., penetration of the cercaria in the skin) to the mating. Indeed, after this period, the female adult worm starts producing eggs, except for *S. haematobium* worms that take about 60–63 days before oviposition (He et al. 2005). About half number of the eggs produced are excreted in feces or urine, while the remaining are trapped into the tissues causing disease (Colley et al. 2014).

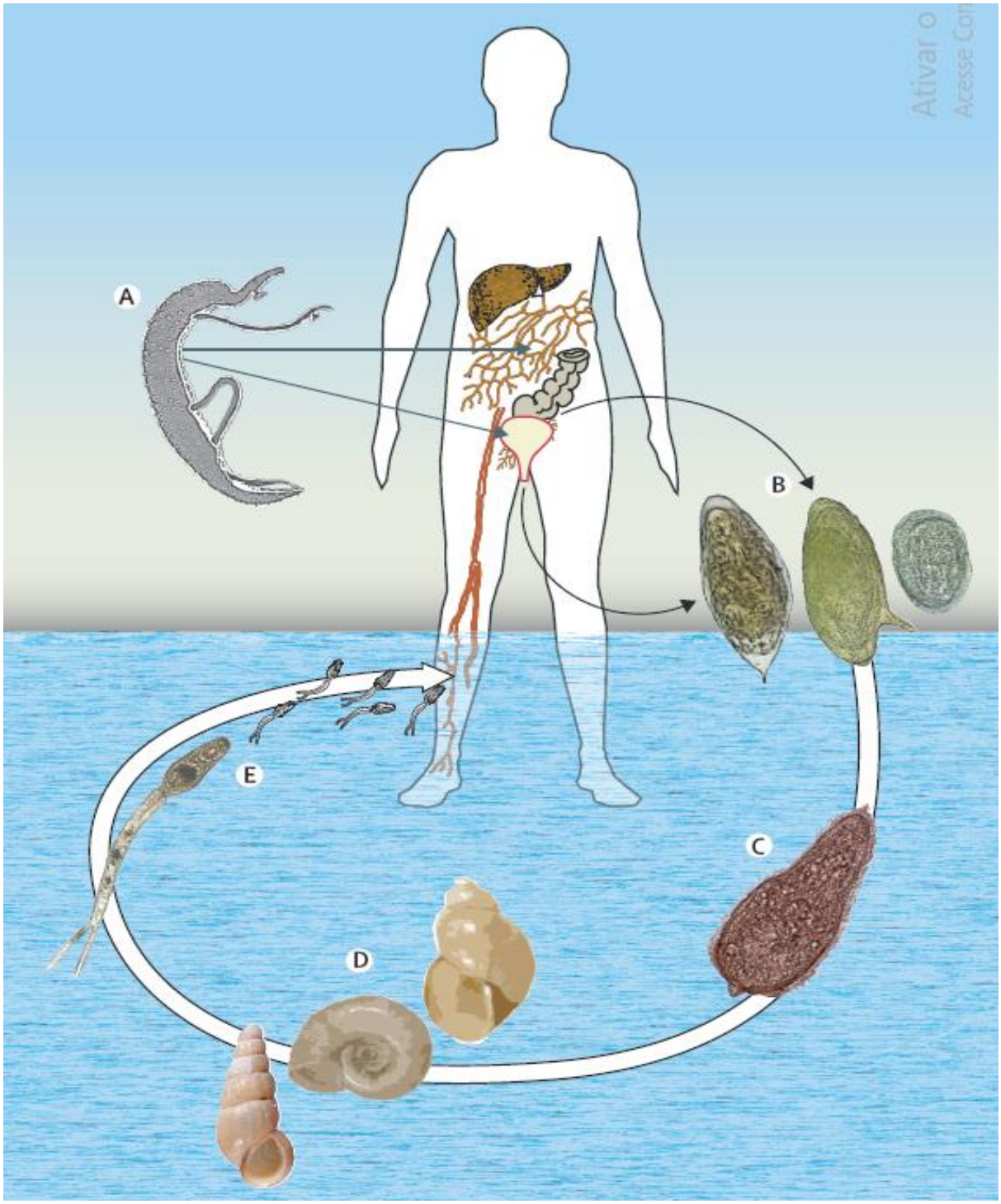


Figure 3: *Lifecycles of Schistosoma mansoni, Schistosoma haematobium, and Schistosoma japonicum* (Colley et al. 2014). ***(A)*** *Paired adult worms (larger male enfolding slender female).* ***(B)*** *Eggs (left to right, S haematobium, S mansoni, S japonicum).* ***(C)*** *Ciliated miracidium.* ***(D)*** *Intermediate host snails (left to right, Oncomelania, Biomphalaria, Bulinus).* ***(E)*** *Cercariae. S mansoni and S haematobium use freshwater Biomphalaria and Bulinus snail species, respectively, whereas S japonicum uses amphibious freshwater Oncomelania spp snails as its intermediate host.*

1.1.4. Clinical manifestations and pathogenesis

The clinical symptoms of human schistosomiasis (*S. haematobium*, *S. mansoni*, *S. intercalatum,* and *S. mekongi*) appear during the different stages of the evolution of the disease, which can range from acute to chronic infection.

Acute infection. School-age children, women, fishermen and farmers living in endemic areas are often infected. Individuals that exert activities such as bathing and swimming in fresh water, scuba diving, water skiing, and rafting, also have at higher risks to be infected (Clerinx and Van Gompel 2011). The onset of infection (mainly characterized by skin penetration by cercariae) is marked by itchy rash also named as "swimmer's itch". It is a localized dermatitis that can result in a pruritic papular or urticarial rash at the site of larval entry (Visser et al. 1995; Kourilova et al. 2004). This hypersensitivity reaction is observed only in individuals who reside in endemic region after repeated exposures. During the evolution of the infection, between three to eight weeks after infection, appears the Katayama fever (an acute schistosomiasis syndrome). It is a systemic hypersensitivity reaction to schistosome antigens due to the presence of circulating immune complexes (Ross et al. 2007). This syndrome is frequently developed in travelers (nonimmune hosts) (Visser et al. 1995; Bottieau et al. 2006) and is manifested by several symptoms, including fever, urticaria and angioedema, chills, myalgia, arthralgia, dry cough, diarrhea, abdominal pain, or headache (Leshem et al. 2008). While eosinophilia is an important marker of acute infection (Alonso et al. 2006; Lambertucci 2010), apparition of neurologic symptoms at this stage is rare (Jauréguiberry et al. 2007).

Chronic infection. Chronic schistosomiasis is restricted to individuals in endemic areas with ongoing exposure. As an exception, chronic manifestations may also be developed by individuals with brief exposure, such as travelers (Whitty et al. 2000; Blanchard 2004; Alonso et al. 2006; Clerinx and Van Gompel 2011); however, the disease does not present systemic symptoms, since they depend on various factors: the number of eggs trapped in tissues, their anatomic distribution, the duration and intensity of infection, and the host immune response (Stephenson 1993; Gryseels et al. 2006b; Colley et al. 2014). Because of organ/tissue tropism, we can distinguish: **(i)** intestinal schistosomiasis due to *S. mansoni*, *S. japonicum*, *S. intercalatum*, *S. mekongi,* and, occasionally, *S. haematobium* (Colley et al. 2014), **(ii)** hepatosplenic schistosomiasis due to *S. mansoni*, *S. japonicum*, *S. intercalatum*, *S. mekongi,* and, occasionally, *S. haematobium* (Colley et al. 2014) and **(iii)** urogenital schistosomiasis, restricted to *S. haematobium* infection (Randrianasolo et al. 2015; Freer et al. 2017), but can be also occasionally caused by *S. mansoni* infection (Feldmeier et al. 1998; Freer et al. 2017).

In the case of human schistosomiasis caused by *S. haematobium* **(studied species in this work)**, distinct disease manifestations may appear depending on the stage of the chronic infection. At the beginning (or early chronic stage), patients present eggs in urine followed by microscopic or macroscopic hematuria and/or pyuria (Devidas et al. 1988; Feldmeier et al. 1999; Da Silva et al. 2006). In most of the uncomplicated cases, blood appears at the end of bladder voiding but it can be also seen during the entire duration of bladder voiding in severe cases (Gryseels et al. 2006b; Colley et al. 2014). Another possible complication in men is hemospermia (Feldmeier et al. 1999). Eggs can produce granulomatous inflammation, ulcerations, and development of pseudopolyps in the vesical and ureteral walls (Devidas et al. 1988; Da Silva et al. 2006). In latter chronic stage, patients can present dysuria and increased urinary frequency. From this moment, the bladder is very predisposed to tumors and becomes fibrosed and may calcify**.** Bladder neck obstruction, hydroureter, and hydronephrosis can be diagnosed at this advanced stage. Finally, the patients can develop renal failure. Epididymis, testicles, spermatic cord, or prostate can also be affected (Devidas et al. 1988; Gryseels et al. 2006b; Colley et al. 2014; Freer et al. 2017). In women, these manifestations may include tubal occlusion, anovulation, hormone perturbation, hypertrophic and ulcerative lesions of the vulva, vagina, and cervix and can lead to the infertility (Chen et al. 2012; Freer et al. 2017).

1.1.5. Factors associated with susceptibility/resistance to Schistosoma haematobium infection

Susceptibility to schistosomiasis may include the influence of environmental (Dessein et al. 1992; Kalinda et al. 2017; Thomas and Short 2017), genetic (Dessein et al. 1992; Isnard and Chevillard 2008b; Mangano and Modiano 2014), and immune factors (details in section 1.4). Other factors such as age, gender, the use of tobacco (Dessein et al. 1992; Remoue et al. 2001; Isnard and Chevillard 2008b; Everts et al. 2010; Zaghloul 2012; Mbanefo et al. 2014) and even the pretreatment intensity (Mbanefo et al. 2014) may also influence the susceptibility to *S. haematobium* infection.

1.1.5.1. Environmental factors

In endemic area, the natural evolution of the schistosomiasis infection depends on the age at the primary exposure and the intensity of the ongoing exposure (Gryseels et al. 2006a; Isnard and Chevillard 2008b), which depends on the magnitude of disease transmission in a particular area (Dessein et al. 1992; Kalinda et al. 2017; Thomas and Short 2017). Several lines of evidence support the idea that the tropical and sub-tropical climates promote the

species richness that can explain not only the high disease burden (Rosenzweig 1992; Schipper et al. 2008; Dunn et al. 2010; CDC 2017), but also intermediate host-parasite interaction (Daszak et al. 2000). Indeed, any climate change due to the anthropological action (Crowley 2000; Matthews et al. 2004; Rosenzweig et al. 2008; Solomon et al. 2009) can lead to intermediate host (eg: snails) spread and consequently to high disease transmission (Gomes et al. 2012; Andrade and Souza-santos 2013; Ajakaye et al. 2016; Kalinda et al. 2017; Thomas and Short 2017). For instance, to perpetuate the snail life cycle, a relatively low water velocity and rainfall with moderate to warm temperatures are required (Ntonifor and Ajayi 2007; Slater and Michael 2012). Human actions, such as the redirection of available water for irrigation systems and the building of dams, may contribute to the implementation of this optimal condition, facilitating an increased snail populations and, consequently, disease transmission (Steinmann et al. 2006; Sokolow et al. 2017; Thomas and Short 2017). Another study also reported that any change in temperature can influence either the schistosoma snail growth, fecundity or survival (Kalinda et al. 2017). This optimal temperature may explain the high *S. haematobium* endemicity in African sub-Saharan regions, like Nigeria (Ajakaye et al. 2016) and Benin (Ibikounlé et al. 2009, 2013) and the high endemicity of *S. mansoni* in Latin-American regions, including Brazil (Gomes et al. 2012; Andrade and Souza-santos 2013).

1.1.5.2. Host genetics

Several studies have reported the influence of genetic factors on helminthiasis susceptibility (including schistosomiasis) (Isnard and Chevillard 2008b; Fumagalli et al. 2010; Isnard et al. 2011; Mangano and Modiano 2014). In 1991 and 1997, two independent groups reported that the major gene region associated with susceptibility to the schistosomiasis was located at 5q31-33 on chromosome 5 (Abel et al. 1991; Marquet et al. 1996). Within this region, are located the genes encoding interleukins 4 (IL-4), IL-5, IL-9, and IL-13, all implicated on the anti-helminthic Th2 response (Sutherland et al. 1988). An additional polymorphic site located at the *IL13* gene promoter region (rs7719175) controlled the level of *S. haematobium* infection (Isnard et al. 2011).

Few studies have addressed the influence of histocompatibility genes on the susceptibility to human schistosome infection. Indeed, no association was observed regarding the *TNF-*α and *TNF-*β (Dessein et al. 1999) genes that are included in the HLA region (6p21); however, a meta-analysis study revealed several HLA class I and class II associations with schistosomiasis (Huy et al. 2011). On the other hand, HLA non-classical class I genes have not been studied yet. Nevertheless, studies in other parasitic diseases found *HLA-G* 3' UTR

polymorphisms exhibiting distinct patterns of associations with both Human African Trypanosomiasis (HAT) and American Trypanosomiasis (Chagas disease) (Courtin et al. 2013; Dias et al. 2015; Sabbagh et al. 2018).

1.1.6. Control, prevention strategies and treatment of urinary bilharzia

On the basis of the life cycle of *S. haematobium*, almost all of the preventive strategies specifically advise to avoid the contact with fresh water in endemic regions (Schwartz and Mawhorter 2013); however, this strategy is far from achieving its goal since fresh water is the only natural water source used by school-age children, women, fishermen and farmers in developing countries to do laundry, crockery, fishing, bathing, agriculture as well as recreation (discussed in (Mbanefo et al. 2014). Conscient on this situation, the World Health Assembly (WHA) (Geneva, 14-22 May 2001) recommendated the mass administration of anthelmintic drugs, periodically (eg, yearly) for the control of schistosomiasis and soil-transmitted helminthiasis. This strategy targeted spefically the high-risk groups [eg, preschool and school-aged children (aged 1–15 years)], since these age-range groups are more susceptible to schistosomiasis and/or to soil-transmitted helminthiasis in endemic regions (WHO 2001, 2006, 2011).

Other control strategies are available and already adopted and used by various countries with successful results. For instance, the snail control implemented by various countries including the African ones has provided a high rate of schistosomiasis control (from 80 to 100%) (Sokolow et al. 2018). However, in some countries (specially from West Africa, including Benin), the snail control strategy is still disappointing (Sokolow et al. 2018), indicating that, depending on the country and its geographical localization, this control maybe more difficult than in those countries. Thus, as precaution, Sokolow et al. suggested that it would be important to combine the preventive chemotherapy (to firstly interrupt the infection in human) with the snail control (Sokolow et al. 2018). In this context, the eradication of the human schistosomiasis is is important to:

- educate and train the medical care personnel, premigrant workers, and travelers on the risk of imported schistosomiasis,
- implement studies to well characterize and identify the molecular identity/profile of the invasive snail and its compatible *Schistosoma* species,
- inform travelers of endemic countries and implement strategies to treat all travelers upon their returning from endemic area,

- establish maps detailing the areas, exhibiting snail hosts susceptible to imported *Schistosoma spp*., and areas in which native snails are susceptible to *Schistosoma* hybrids (Sanogo et al. 2018).

The public health education campaigns (including sanitation education) and the govern intervention are encouraged to increase the access of potable water for all populations (Abou-El-Naga 2018). Besides all these recommendations, we can also highlight the numerous ongoing efforts/works in the development of anti-schistosome vaccines (Merrifield et al. 2016; Molehin et al. 2016; Siddiqui and Siddiqui 2017). For instance, the Sm-p80-Based Vaccine is the unique ongoing vaccine based on the large subunit of the *S. mansoni* calcium-activated neutral protease calpain that showed its effectiveness against intestinal/hepatic and urinary schistosomiasis (Siddiqui and Siddiqui 2017).

The diagnosis of urinary bilharzia or *S. haematobium* infection is based on the search of eggs in urine (Doenhoff et al. 2004; Schwartz and Mawhorter 2013). Praziquantel is the recommended treatment used because it is inexpensive (cost-available) and effective against the adult forms of all *Schistosoma* species (Liu et al. 2017), and artemisinin derivatives can also be used against the juvenile forms of *Schistosoma* (Pérez del Villar et al. 2012). The recommended posology of treatment is the following: Praziquantel 20 mg/kg/dose *per os bid* for 1 day in the case of *S. hematobium or S. mansoni* infection (Schwartz and Mawhorter 2013).

1.2. HUMAN *PLASMODIUM* MALARIA

1.2.1. Definition

Human to human malaria transmission is caused by infection with five protozoan species belonging to the genus *Plasmodia* (*Plasmodium falciparum, P. vivax, P. malariae and P. ovale* (sub-species *P. ovale curtisi* e *P. ovale wallikeri* (Sutherland et al. 2010)) transmitted by female *Anopheles* mosquitoes. But, humans can also be infected by *P. knowelesi, a* monkey malaria parasite, endemic in the South-Eastern Asia forested regions (WHO 2015).

Historically, malaria or ‘’bad air’’ (*mal’aria* in Italian) has existed since a long time (a century with real selection pressure exercised for 10,000 years). The ‘’bad air’’ symptoms had already been described in the *Nei Ching* book (the Canon of Medicine, ancient Chinese medical book, written 2700 BC) and some of them, particularly the fever, had been attributed to the bites of certain insects (*The Compendium of Susruta,* a Sanskrit medical treatise). Much later, in 1880, the French Nobel-prized Charles Louis Alphonse Laveran reported for the first time the presence of parasites in the blood of patient suffering from malaria. Chronologically,

the *P. vivax (1890)* and *P. malariae (1890)* were the first described parasites, followed by *P. falciparum* (1897), *P. ovale* (1931) and *P. knowlesi* (1931). In 1886, the Italian scientist Camillo Golgi reported two clinical forms of malaria: one exhibiting fever every other day (tertian periodicity) and another presenting fever every third day (quartan periodicity). In 1897, William Welch named the *P. falciparum* parasite as the responsible for tertian malaria. Between 1897 and 1898, Ronald Ross (British) described mosquitoes as malaria transmission vector. And finally, between 1898 and 1899, an Italian investigator team established that the human *P. falciparum* can be transmitted from infected to uninfected individuals (Cox 2010).

1.2.2. Localization

Of the five *Plasmodium* parasites (*P. falciparum, P. vivax, P. malariae, P. ovale, e P. knowelesi*) that cause malaria, *P. falciparum* and *P. vivax* are those that pose the greatest threat. *P. falciparum* is the most prevalent in the African continent and is responsible for most malaria-related deaths worldwide, whereas *P. vivax* is the dominant malaria parasite in most countries outside of sub-Saharan Africa (WHO 2017).

Generally, *P. falciparum* is the most preponderant in tropical Africa, while *P. vivax* and *P. falciparum* co-exist in South America with *P. vivax* more represented than *P. falciparum.* Both *P. falciparum* and *P. vivax* are prevalent in Southeastern Asia (Durnez et al. 2018) and Western Pacific, while *P. malariae* is present in all malarious areas but with generally low prevalence. Co-infection by both *P. falciparum* and *P. malariae* is sometimes registered in African tropical regions, where *P. ovale* is prevalent. Finally, *P. knowlesi* infection is restricted to certain forested areas of Southeast Asia (Guerra et al. 2008, 2010; Hay et al. 2009; Autino et al. 2012).

1.2.3. Transmission

It has long been established that malaria is transmitted through the bites of female *Anopheles* mosquitoes**.** Of the more than 400 different *Anopheles* mosquito species existing, 30 are described to be the major malaria vectors, usually biting between dusk and dawn. *Anopheles* mosquitoes lay their eggs in water, which hatch into larvae, eventually emerging as adult mosquitoes. The female mosquitoes seek a blood meal to nurture their eggs. The intensity of transmission depends on the places where the mosquito lifespan is longer (where the parasite has time to complete its development inside the mosquito), preferring humans rather than other animals. This is the main reason of the alarming world's malaria cases in Africa (90%) (WHO, 2018).

The life cycle of *Plasmodium spp*. **(Figure 4)** begins when an infected female *Anopheles* mosquito bites its host. The sporozoites are inoculated during the bite, penetrate into the mammalian skin, reach the liver and invade hepatocytes. Approximately thousand merozoites are released and reenter the bloodstream, invading erythrocytes **(end of the pre-erythrocytic cycle).**

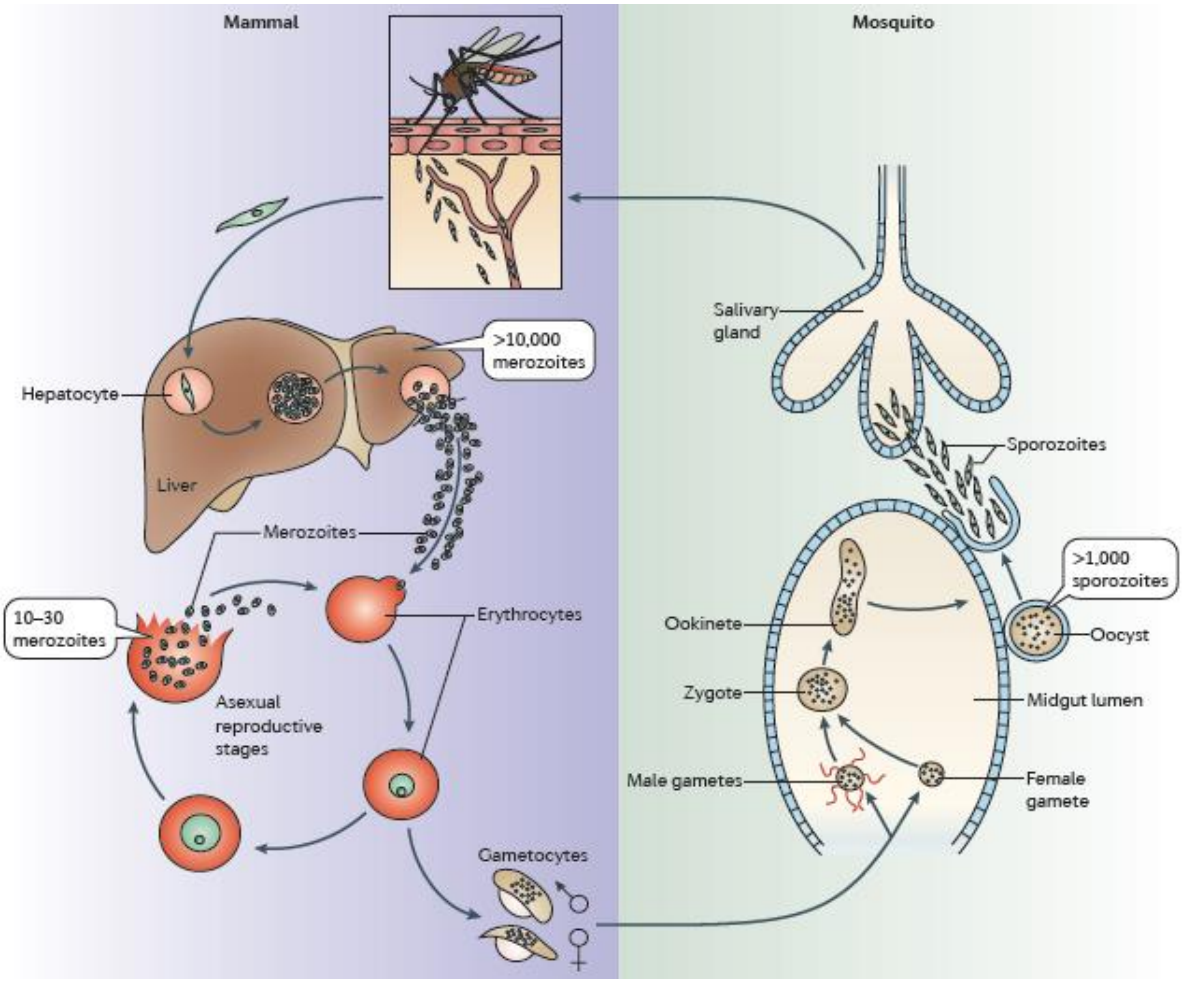


Figure 4: *The Plasmodium spp. life cycle. The life cycle of Plasmodium species that infect mammals* (Ménard et al. 2013).

The erythrocyte cycle begins exactly when a merozoite invades the erythrocyte, where it undergoes 3-5 nuclear divisions and generates ~8–32 new merozoites. This cycle typically lasts ~24 hours and ~48–72 hours in species that infect rodents and humans, respectively. Some intra-erythrocytic parasites transform into male or female gametocytes, which are taken up by a mosquito during a new bite. Gametocytes egress from erythrocytes, activate into gametes and fuse in the mosquito midgut lumen. The motile zygote, called an ookinete, crosses the gut epithelium to transform into an oocyst, in which thousands of sporozoites develop. Sporozoites are released into the mosquito body cavity and later pass through salivary gland cells to enter the salivary ducts (Ménard et al. 2013).

1.2.4. Clinical manifestations and pathogenesis

In a non-immune individual, the first malaria symptoms (fever, headache, and chills) usually appear 10–15 days after the infective mosquito bite. These symptoms are generally related to uncomplicated malaria. If untreated within 24 hours, *P. falciparum* malaria can progress to severe illness and even death. The *P. falciparum* malaria victims are often children under 5 years of age, pregnant women and patients with HIV/AIDS, as well as non-immune migrants, mobile populations and travelers. Children with severe malaria frequently develop one or more of the following symptoms: severe anaemia, respiratory distress in relation to metabolic acidosis, multiple convulsions, hypoglycaemia, acidosis, hyperlactataemia, hyperparasitaemia or cerebral malaria. The frequent and more pronounced symptoms generally encountered in adults with severe malaria include jaundice, acute kidney injury as well as hypoglycaemia, acidosis and hyperlactataemia. In malaria endemic areas, asymptomatic infections may also occur (Health and Malaria 2014; WHO 2015, 2017).

1.2.5. Factors associated with susceptibility/resistance to P. falciparum infection

The susceptibility/resistance to *P. falciparum* malaria depends on various factors, such as those related to the parasite (different species, sporogonic cycle length, drug susceptibility), to the vector (density, larvae breeding sites, temperature, receptivity, feeding pattern, longevity, insecticide susceptibility), to the human host (host immunity and genetics) and to the environment (physical, biological and socio-economic) (Autino et al. 2012; WHO 2017). Among these, we can highlight the environment, host immunity (detailed in section 1.5) and host genetics.

1.2.5.1. Environmental factors

The climate exerts great influence on the geographical distribution of most vector-borne diseases. And this is especially true for malaria transmission, a disease closely dependent on the environment conditions or temperatures favorable for the development of *Anopheles* vectors (Guthmann et al. 2002). These climatic conditions such as rainfall patterns, temperature and humidity may affect the number and survival of mosquitoes. For example, in some endemic areas where the transmission is seasonal, the peak of transmission is often situated during and/or just after the rainy season (Guthmann et al. 2002; Autino et al. 2012). This variation or availability of water and temperature can also affect the duration of the development cycles and ultimately the vector density. Thus, a temperature below 18 °C prevents the existence of the reproductive cycle of *Anopheles*, but temperature variations

beyond this threshold may relatively facilitate the development of *Anopheles* and, therefore, the time required for vectors to reach an epidemiologically dangerous stage. It is also known that there is an average temperature (20-25 °C) and an average humidity of 60% (Martens et al. 1995), which provide an ideal development time, between 9 and 12 days, from larva to the adult *Anopheles*. Ecologic features of vectors are also important for the *Anopheles* subspecies (Mutuku et al. 2006; Djenontin et al. 2010; Djènontin 2015), particularly for *Anopheles gambiae* and *An. funestus*, the main malaria vectors in Southern Benin, where the entomological inoculation rate (EIR), i.e., the number of infective bites per individuals per year, was estimated approximately at 1.38 ± 0.94 (infective bites per human per 100 nights) for *An. gambiae* and 0.67 ± 0.60 for *An. funestus* (Djenontin et al. 2010). Particularly, in Sô-Ava (the site of present study), the *Anopheles* species (*A. coluzzii* and *A. melas*) still predominate; however, only *A. coluzzii* could transmit *P. falciparum* (Djènontin 2015). The features of the domicile represent an essential link to the understanding of malaria risk. Indeed, in endemic areas, there may have significant differences in intensity of transmission from one house to another or from one village to another, that are even separated by short distances (Guthmann et al. 2002; Y?? et al. 2007). Sogoba et al. also showed that the variation of the transmission in time and in space results from the variation of the human-vector contact on a very fine scale (Sogoba et al. 2007).

1.2.5.2. Host genetics

Host genetic factors also represent important contributors to malaria susceptibility, since in the same endemic area, patients with same level of immunity and sharing the same conditions may present distinct responses against malaria; i.e., in the presence of the same amount of parasitemia, one individual may be susceptible to severe malaria and another may develop a milder form. This observation is common among different sympatric ethnic groups, living in the same epidemiological context but differing in their genetic background and susceptibility to malaria infection and disease (Terrenato et al. 1988; Modiano et al. 1996, 1998; Dolo et al. 2005). The best examples are the Fulani, who exhibit greater resistance to malaria infection and mild and severe malaria attacks when compared to other sympatric ethnic groups living in the same villages (Modiano et al. 1996, 1998) (Dolo et al. 2005) (Bereczky et al. 2006). The genetic component of malaria-related phenotypes was also observed by comparing monozygotic twin pairs with dizygotic pairs as well as siblings (Hernandez-Valladares et al. 2014). Two studies have evaluated the genetic influence on susceptibility/resistance to malaria, but only small part was attributed to the hemoglobin S

(*HbS*) and α-thalassemia genes, suggesting that other genes may be involved on the resistance to *P. falciparum* malaria (Mackinnon et al. 2005; Phimpraphil et al. 2008). Besides the heritability of malaria phenotypes, many studies have identified genes or genomic regions associated with susceptibility to *P. falciparum* malaria (Driss et al. 2011; Hernandez-Valladares et al. 2014).

1.2.6. Control, prevention strategies and treatment for human P. falciparum malaria

The World Health Organization (WHO) proposed the following strategies to control, prevent and eradicate malaria from endemic areas and avoid malaria re-establishment in areas where malaria was eliminated or areas that are ongoing malaria control **(Figure 5)** (WHO 2017).

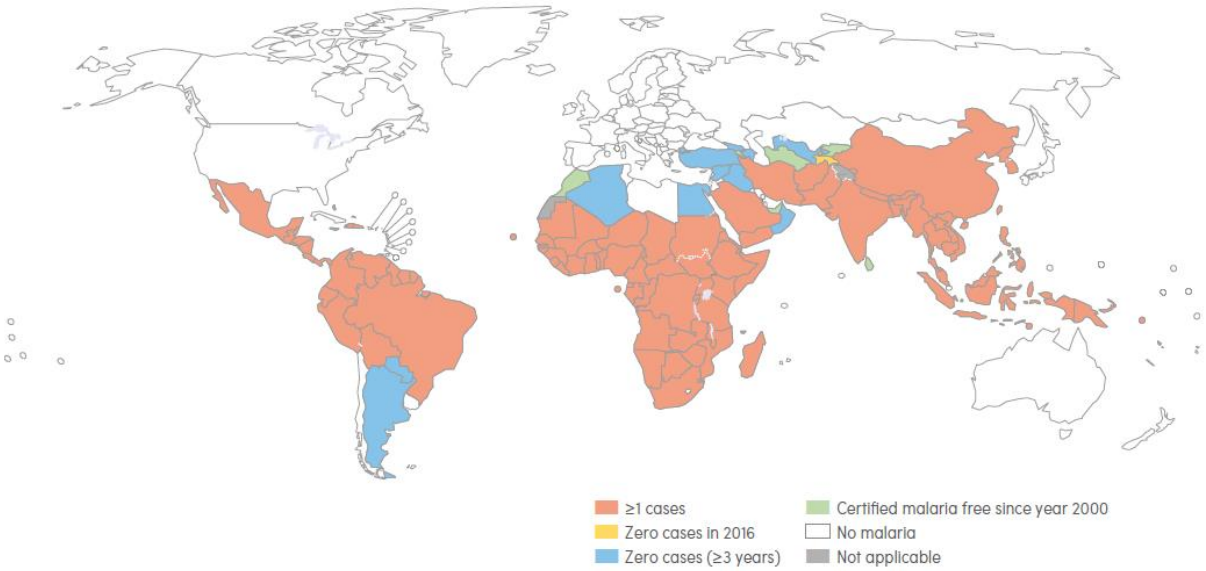


Figure 5: *Countries and territories with indigenous malaria cases in 2000 and their status by 2016. Countries with zero indigenous cases over at least the past 3 consecutive years are eligible to request certification of malaria free status from WHO. All countries in the WHO European Region reported zero indigenous cases in 2016. Kyrgyzstan and Sri Lanka were certified malaria free in 2016. Source: WHO database* (WHO 2017).

1.2.6.1. Prevention

The strategy proposed for prevention is ‘’vector control’’ that can take two forms: **(i)** insecticide-treated mosquito nets and **(ii)** indoor residual spraying. Indeed, WHO recommends the use of free of charge long-lasting insecticidal nets (LLINs) to cover all people at risk of malaria development. In addition, it is important that each country, even sub-locality,

implement communication strategies to inform the population about the importance of the use and adequate maintenance of LLINs. Considering that mosquito resistance to the all 4 forms of insecticides (including pyrethroids, the most recommended and used) has been reported, WHO advises the rotational use of these insecticides for indoor residual spraying (IRS). This spray has been tested, and its efficacy has proved to reduce malaria transmission. Since 2012, WHO have encouraged and supported the "*Global Plan for Insecticide Resistance Management in Malaria Vectors (GPIRM)*", aiming to detect insecticide resistance.

1.2.6.2. Antimalarial drugs

The use of chemoprophylaxis is recommended for the travelers. This measure may suppress not only the blood stage of malaria infections but also will prevent malaria, particularly for the most susceptible group, such as the pregnant women and children under 5 years living in moderate-to-high transmission areas. In these cases, WHO recommends intermittent preventive treatment with sulfadoxine-pyrimethamine at each scheduled antenatal visit after the first trimester, and 3 doses of intermittent preventive treatment with sulfadoxine-pyrimethamine for infants, delivered alongside routine vaccination. Another alternative prophylaxis was introduced in 2012, which consists on the administration of monthly courses of amodiaquine plus sulfadoxine-pyrimethamine to all children under 5 years of age during the high transmission season. But, as drug resistance is not excluded, WHO supports all effort, in any country, in routine monitoring antimalarial drug resistance (WHO 2017).

1.2.6.3. Diagnosis and treatment

To confirm a suspect case of malaria, i.e., children with temperature ≥37% or history of fever for 24h in malaria endemic areas or the presence of palmar pallor or haemoglobin concentration <8g/dL in areas where malaria transmission is seasonal, either light microscopy or rapid diagnostic test (RDT) are the best parasite-based diagnostic tests, recommended before administering treatment. The light microscopy is used to observe the pre-prepared thick and thin blood smears/films (TBS). The thick blood smear detects the infected red cells by visualizing the presence of merozoites (during erythrocytic stage) inside the red cells, while the thin blood smear allows the characterization of the infecting malaria parasite species (*P. falciparum, P. vivax, P. ovale or P. malariae…*). Finally, the RDT is used to detect the

parasite-specific antigen or enzyme and can also tell us the nature of malaria parasite species (WHO 2017).

The best and most used treatment available today against uncomplicate *P. falciparum* malaria is artemisinin-based combination therapy (ACT). Only individuals positive for one or both diagnostic tests are eligible for the treatment. Children and adults (except pregnant women in their first trimester) can benefit of 3-day treatments of the five artemisinin-based combination therapies: artemether+lumefantrine, artesunate+amodiaquine, artesunate-mefloquine, dihydroartemisinin+piperaquine and artesunate+sulfadoxine-pyrimethamine (WHO 2017). In addition, the quinine can also be used to treat either uncomplicated malaria (using tablet formula) or severe malaria (using injectable formula).

1.2.6.4. Surveillance, elimination and vaccines against malaria

The surveillance constitutes the major disease tracking. This means that each country is responsible to collect information about the course or evolution of the disease in its endemic areas and be capable to send suitable data to WHO on the level of endemicity or on the disease distribution. And then, WHO can optimize strategies and respond to novel outbreaks. Surveillance is also centered on detecting, treatment and reporting to a national malaria registry of every infection. All patients diagnosed must be systematically treated by avoiding onward transmission of the disease in the community. By this way, each country can contribute to malaria elimination (WHO 2017).

One approved vaccine, the RTS, S/AS01 (RTS,S)[1] –also known as Mosquirix– had its phase III already completed in 2014 (GlaxoSmithKline 2015; Rts 2016), providing partial protection against malaria in young children. Between 2015 and 2016, this vaccine had already received approbation from different malaria control programs, including the own WHO groups, to implement pilot vaccination projects in various African countries (Ghana, Kenya and Malawi). Likely, this vaccination program using the RTS,S vaccine for the first time will begin in 2018 (WHO 2017).

1

RTS,S is a scientific name and represents the vaccine composition. The 'R' means central repeat region of P. falciparum circumsporozoite protein (CSP); 'T', the T-cell epitopes of the CSP; and 'S', the hepatitis B surface antigen (HBsAg). These are combined in a single fusion protein ('RTS') and co-expressed in yeast cells with free HBsAg. The 'RTS' fusion protein and free 'S' protein spontaneously assemble in 'RTS,S' particles. RTS,S also contains the AS01 adjuvant system, and in scientific papers is usually referred to as 'RTS,S/AS01' (Rts 2016).

1.3. SCHISTOSOMIASIS AND MALARIA: PUBLIC HEALTH PROBLEMS

In tropical regions, malaria and schistosomiasis are at the top of the major public health problems (Adegnika and Kremsner 2012).

1.3.1. Schistosomiasis as public health problem

Schistosomiasis today is present in 78 countries worldwide and remains the most prevalent problem in tropic and subtopic countries. WHO estimated at 700 million, the total individuals at high risk of developing schistosomiasis worldwide, and 240 million of them were confirmed infected, the majority from sub-Saharan African countries (Colley et al. 2014; Collaborators 2017; Mitra and Mawson 2017). The disease, known as the disease of poverty, is responsible for 3.3 million disability-adjusted life years (Bergquist and Whittaker 2012; Hotez et al. 2014). The high-risk groups (e.g, children at school-age) present various sequelae. Although the main clinical symptom is hematuria that affects about 70 million of the affected individuals, anemia, growth retardation, irritability and cognitive impairment and poor performance in school are also common (Chitsulo et al. 2000; Ekpo et al. 2012; Mitra and Mawson 2017). Many of these children, during their adulthood, can present liver fibrosis, urinary tract obstruction and bladder cancer because of chronic infection contracted during their childhood. Besides these innumerous sequelae, schistosomiasis is also the result for many social and economic burden due to social stigma, discrimination, loss of social status, malnutrition and a loss of work-days (Boko et al. 2016; Mitra and Mawson 2017).

In Benin, *S. haematobium* and its major snail vector (genus *Bulinus*) are detected across the territory of Benin (Ibikounlé et al. 2009, 2013, 2014). In 2009, *S. haematobium* presence reached 100% in Borgou region (Northern Benin) (Ibikounlé et al. 2009). In 2016, although decreased, these estimates are still preoccupant, increasing from South (4.6%) to North (38.4%) Benin (Boko et al. 2016).

1.3.2. Malaria as public health problem

The latest worldwide WHO report announced 216 million cases of malaria with *445,000* deaths in 2016, and the highest burden is confined to 15 sub-Saharan countries. Children under 5 years are the main victims and more than 70% of deaths occur within this group, in which one child dies every two minutes. Malaria represents an obstacle to the development of African communities, and despite all the efforts of diverse organizations involved on the control of malaria, new infections and deaths continue to be registered every

day (Garner and Gülmezoglu 2003; Aponte et al. 2009; Pluess et al. 2010; Roca-Feltrer et al. 2010; Kayentao et al. 2013; WHO 2017).

According to 2011 health statistics from Benin's Ministry of Health (MOH), malaria is the cause of morbidity and mortality among children under the age of five and pregnant women. The Beninese Health Management Information System (HMIS) reported that, in 2016, 11,2 million Beninese (100% of the population) were at risk of malaria with 1,324,576 cases confirmed and 1,646 deaths. Moreover, in every 1000 children under 5 years, 115 deaths were registered (WHO 2017).

1.4. *S. HAEMATOBIUM* AND IMMUNE SYSTEM

The advent of the mouse model of direct oviposition (into the bladder lamina propria) (Fu et al. 2012) or into the posterior vaginal walls (Richardson et al. 2014) has opened the way to study and understand the role of immune cells in *S. haematobium* infection. The major features that characterize the innate immune response to *S. haematobium* is the formation of granuloma. It is a structure where the eggs are surrounded by various innate cells (infiltrated or resident), such as epithelioid macrophages, admixed eosinophils, neutrophils as well as T $CD4^+$ cells (Colley et al. 2014; Odegaard and Hsieh 2014; Richardson et al. 2014). The granuloma formation is initiated and maintained mainly by M2-macrophages (expressing *ARG1, MRC* and *CHI3L3* transcription factors*),* which mount a robust type 2 response (Th2) (Herbert et al. 2004; Wilson et al. 2007; Meurs et al. 2014; Odegaard and Hsieh 2014). However, the Th1 markers (M1-macrophage expressing NOS2 and CD40) were also found in the granuloma (Ray et al. 2012), indicating that both macrophage types can act on the context of urogenital oviposition in *S. haematobium* infection.

Although eosinophils mainly act in the pre-egg stage against cercariae and in adult worms, in a IgE-dependent manner (Verwaerde et al. 1987), these cells are also found in the granuloma observed in the bladder and in the vaginal wall of the oviposition mouse model (Fu et al. 2012; Richardson et al. 2014). As well as in the mouse model, eosinophils were also detected in schistosome-infected humans (Sadun et al. 1970; Von Lichtenberg et al. 1971; Richardson et al. 2014). They are the important producer of IL-4 in many immunological contexts (Heredia et al. 2013; Molofsky et al. 2013) and also of IL-13 in hepatosplenic schistosomiasis, promoting the M2-macrophage activation within the granuloma (Reiman et al. 2006). This observation links eosinophil infiltration to the hepatic fibrosis manifestations during *S. haematobium* infection (Reiman et al. 2006). Its functional contribution to fibrosis was also confirmed by another study, in which the DdblGATA or TgPHIL mice (both lack

eosinophils but had preserved IL-5 production) do not exhibit clinical manifestations of the *S. mansoni* infection (Swartz et al. 2006). Despite the crucial role of eosinophils in immunological process during *S. mansoni* infection, its functional contribution in urogenital schistosomiasis remains unclear, even though they infiltrate into the granuloma in the mouse model (Richardson et al. 2014). The presence of eosinophils has also been associated with the pathophysiology manifestations (Rubbery papules) in women with genital *S. haematobium* infection (Randrianasolo et al. 2015). Besides eosinophils, few studies report the participation of other innate cells, such as dendritic cells (DC), natural killer (NK) cells, and basophils/mast cells on the immune response against *S. haematobium* (Odegaard and Hsieh 2014)**.**

Myeloid DC (mDC) cells from patient with *S. haematobium* and *S. mansomi* showed impaired functional profile and become less responsive to the parasite, possibly due to the downregulation of major maturation markers, including HLA-DR, CD80, CD83 and CD86 (Everts et al. 2010; Lopes et al. 2018). In addition, mDC profile was age-dependent; i.e., younger children (aged 5–9 years) exhibit higher level of mDC with a hyporesponsiveness profile, while older children (aged 10–13 and 14–45 years) exhibit higher levels of schistosome specific IgG, accompanied by lower proportion of mDC (Everts et al. 2010).

NK cells may be important for the clearance of schistosome eggs, since NK activation and recruitment within the granuloma is correlated with high IL-12 expression, reduced volume of granuloma and, consequently, reduced hepatic fibrosis cell markers (known as hepatic stellate cells) (Remick et al. 1988; Hashimoto et al. 1990; Asseman et al. 1996; Hou et al. 2012). No recent study has addressed the participation of NK cells on human *S. haematobium* infection. Although, many existing evidences have appointed basophils as IL-4 producer in the *S. mansoni* infection context (Schramm et al. 2007; Odegaard and Hsieh 2014), the role of these cells in *S. haematobium* infection has not been defined (Odegaard and Hsieh 2014).

Various types of T cell response can occur during schistosome infection: T helper (Th)-2, Th1/Th17, and T regulatory (Treg) responses. The Th2 response contributes to granuloma formation/or its maintenance, and is important to control any exacerbation of the immune response (innate or adaptive) that can lead to tissue damage (Herbert et al. 2004; Kouriba et al. 2005; He et al. 2008; Isnard et al. 2011). Each Th2-associated cytokine has specific role during the infection. IL-4 is important for granuloma formation, IL-5 recruits and maintains eosinophils, and IL-13 is the principal inducer of fibrosis (Odegaard and Hsieh 2014).

The Th1/Th17 response, in contrast, was lower whereas the Th2 response was higher in infected mouse model (Fu et al. 2012)**.** Similar Th1/Th17 and Th2 cytokine profiles were observed in humans infected with *S. mansoni* or *S. haematobium* (Burke et al. 2009)**.** This observation was also confirmed by other genetic association studies in humans, for whom the polymorphisms that affect the Th2 pathway (variants at the *IL13* and *STAT6* genes) or enhance the Th1 response (variants in *CTLA4*) were associated to human susceptibility to *S. haematobium* (Isnard and Chevillard 2008a; Maizels 2009). Finally, higher IL-10 level as well as both T or B regulatory responses were observed in humans infected with *S. haematobium* (Burke et al. 2009; Nausch et al. 2011; Lepesheva et al. 2015) and IL-10 is required to regulate both Th1 or Th2 response (Lundy and Boros 2002; Reis et al. 2007; Teixeira-Carvalho et al. 2008).

Whereas NKT cells had been shown (with controversial results) to stimulate Th1/or Th2 response during *S. mansoni* infection in mice (Mallevaey et al. 2006, 2007), no studies addressed the effect of these cells in urogenital schistosomiasis. However, B cells-expressing IgE level was associated to resistance to human *S. haematobium* (Hagan et al. 1991). In endemic regions where individuals were chronically exposed, the later chronic stage of the disease is associated to specific-B cell responses. Indeed, in individuals chronically infected, it was observed high levels of specific antibodies (IgE or IgG) directed against *S. haematobium* parasite or adult worm (Iskander et al. 1981; Grogan et al. 1997; Naus et al. 1998; Burke et al. 2009) and these antibodies were associated with reduced disease-susceptibility in age-dependent manner (Mutapi et al. 2008).

1.5. HUMAN *P. FALCIPARUM* MALARIA AND IMMUNE SYSTEM

Upon infection, the sporozoites, before reaching the hepatocytes, will face the skin barrier resistance mediated by various skin anti-sporozoites antibodies (Rénia and Goh 2016). This skin barrier response can reduce up to 50% the number of sporozoites that can further infect hepatocytes (Guilbride et al. 2012). However, some sporozoites may escape this innate response by expressing specific immune evading proteins (SPECT1 and SPECT2) (Zheng et al. 2014; Gomes et al. 2016).

Inside the hepatocytes, sporozoites initiate replication. This stage is characterized by the recruitment of NK, NKT and γδT cells. They will produce IFN-γ, type-I interferons and other mediators to control the parasite replication in the liver. Specific T $CD8^+$ cells will kill the infected hepatocytes by cytotoxicity (Liehl et al. 2014; Miller et al. 2014; Belachew 2018). The antibodies are also important at this stage to neutralize not only the sporozoites but

also the circumsporozoite protein (CSP), inhibiting the invasion of new hepatocytes. These antibodies also play various other roles, such as activation of complement, opsonization and antibody-dependent cell-mediated cytotoxicity (ADCC) (Belachew 2018).

If any parasite escapes the immunologic control at this hepatic or pre-erythrocytic cycle, it can reach the bloodstream and initiates erythrocytic cycle. Here, the targets are free merozoites and the intraerythrocytic schizonts. This stage is mainly characterized by the pro-inflammatory signatures and a strong participation of humoral and cellular responses (Franklin et al. 2009; Sharma et al. 2011; Belachew 2018). IFN-γ at this stage is important activator of macrophages and consequently mediates the resistance to infection, whereas the antibodies are important to induce ADCC, opsonization, inhibition of red blood cell (RBC) invasion, neutralization of parasite toxins, activation of complement and macrophage for gametocyte killing (Walther et al. 2006; Sponaas et al. 2009; Antonelli et al. 2014; Belachew 2018). However, this anti-*P. falciparum* response, when persistent may lead to tissue damage. To prevent this situation, the T regulatory response pathway, characterized specially by high IL-10 or TGF-β level, is naturally activated to call back the activated immune system into it baseline level (Perkins et al. 2000; Ouma et al. 2008). Nevertheless, sometimes, this also can allow parasite to escape immune control (Belachew 2018).

In the context of a co-infection with both *P. falciparum* and *S. haematobium*, what could be the specific immune response (Th1 or Th2)? Since the control of human urogenital schistosomiasis requires mainly the Th2 response, and at later stage can switch to a T regulatory response in chronic infection.

1.6. *P. FALCIPARUM* AND *S. HAEMATOBIUM* COINFECTION AND IMMUNE SYSTEM

As mentioned at section 1.4, infection with *S. haematobium* early stimulates the Th2 response marked by a strong production of IgE and IgG4 and a later regulatory response with the production of IL-10 and other immunodulatory-mediated markers that inhibit both Th1 and Th2 responses. The Th1 response (detailed at section 1.5), mediated mainly by IgG1, IgG3, T $CD8^+$ and T $CD4^+$, is the adapted one to control the *P. falciparum* infection (Courtin et al. 2009). Therefore, in context of *S. haematobium* and *P. falciparum* co-infection, the immune system faces a challenge because each parasite requires a different appropriate response. In this context, the specific Th2 and the late regulatory response promoted by schistosomes could modulate the host immune response specific against *P. falciparum* antigens and make the host susceptible to malaria.

Many experimental and epidemiological studies have focused on the dual malaria/helminthiasis infection, particularly on malaria/schistosomiasis co-infection. Some studies have suggested an augmented risk for malaria in *S. haematobium* co-infected individuals (Sangweme et al. 2010; Courtin et al. 2011), whereas others reported a protective effect of *S. haematobium* infection against *P. falciparum* malaria. Indeed, children aged 4-8 years, co-infected by *S. haematobium*, exhibited decreased *P. falciparum* parasitemia and decreased number of malaria episodes, and exhibited a delayed time of first clinical malaria infection (Briand et al. 2005; Lyke et al. 2005). This protection was also observed in the group of children with light *S. haematobium* infection (1–9 eggs/10 mL of urine) compared to moderate (10-49 eggs/10 mL of urine) and heavy (≥ 50 10-49 eggs/10 mL of urine) infection (Briand et al. 2005; Lemaitre et al. 2014; Doumbo et al. 2014; Degarege et al. 2016). In addition, any *S. haematobium* co-infected children were prevented from developing anemia (Doumbo et al. 2014; Degarege et al. 2016). This protection might be explained by the high level of Th2 cytokines as observed in the single infection by *S. haematobium* (Lyke et al. 2006), which is sufficient to reduce T regulatory cells (Lyke et al. 2012a), IL-10 production (Lyke et al. 2006) and, consequently, providing protection against *P. falciparum* infection (Lyke et al. 2006, 2012a). In addition, high levels of IgG1 and IgG3 against *S. haematobium* infected children may protect against malaria, a finding that may be due to cross-reactivity between antigens derived both agents (Mutapi et al. 2007; Diallo et al. 2010). These results were further confirmed by Lyke et al., who observed that T $CD4^+$ and B cells with specific memory responses against *P. falciparum* malaria antigens (apical membrane antigen 1 (AMA-1) and against merozoite surface protein 1 (MSP-1)) were observed in individuals co-infected with *S. haematobium* (Lyke et al. 2012b, 2018).

1.7. MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) OR HUMAN LEUCOCYTE ANTIGEN (HLA)

The human major histocompatibility complex (MHC) comprehends more than 200 genes, and among them are the histocompatibility classical class I (*HLA-A, -B and -C*) and non-classical class I (*HLA-E, -F and -G*) genes (Geraghty et al. 1992; Janeway et al. 2001). Differently from the HLA classical class I genes, *HLA-G*, *HLA-E* and *HLA-F* show limited genetic and protein polymorphisms. *HLA-E and HLA-F* are the most conserved ones, exhibiting 27 and 30 alleles, respectively, accompanied by *HLA-G* with 60 different alleles (IPD-IMGT/HLA v.3.32.0, 2018-04-16) (Robinson et al. 2015). *HLA-G, -E,* and *-F* show a restricted expression pattern and are important regulators of the immune system cells

(Carosella et al. 2015; Persson et al. 2017; Sabbagh et al. 2018). Since HLA-G, HLA-E, HLA-F and HLA-C molecules are simultaneously expressed at the fetal-maternal interface, these molecules have been associated with immune tolerance towards the semi-allogeneic fetus (Kovats et al. 1990; Ishitani et al. 2003, 2006).

The non-classical class I genes possess a similar genetic architecture as the classical ones: exon 1 encoding the signal peptide (SP); exons 2–4 encoding the extracellular (EC) domains α1, α2, and α3, respectively; exon 5 encoding the transmembrane (TM) domain; exons 6–7 encoding the intracellular (IC) cytoplasmic tail. The exon 8 or 3' untranslated region (UTR) is untranslated for *HLA-E, -F* and *-G*, whereas exon7 is absent in *HLA-G* and *HLA-F*. Alternative splicing results in at least seven different *HLA-G* mRNA isoforms, with HLA-G1-4 expressed as membrane-bound proteins and HLA-G5-7 expressed as soluble proteins (sHLA-G) (Castelli et al. 2014a; Persson et al. 2017). The IPD-IMGT/HLA coding segment for HLA Ib genes (*HLA-E, -F and -G*) is formed by **i)** a part of promoter (starting by -300), **ii)** all coding region and **iii)** part of 3' UTR for *HLA-E* and *HLA-F* (Robinson et al. 2015).

1.7.1. Overview on the Non-Classical Class I (HLA)-G

HLA-G has a well-recognized immunomodulatory role, inhibiting the function of several immune cells including monocytes/macrophages, dendritic cells (DCs), neutrophils, NK cells, B cells and T cells, by interacting with Ig-like transcript receptor 2 (ILT-2) and/or ILT-4 (Sabbagh et al. 2018). Besides pregnancy, the immunomodulatory role of HLA-G has been reported in several disorders of distinct etiology (Donadi et al. 2011; Carosella et al. 2015; Persson et al. 2017; Sabbagh et al. 2018). In parasite disorders, high levels of soluble HLA-G (sHLA-G) have been observed in patients presenting toxoplasmosis and in the active phase of echinococcosis (Robert-Gangneux et al. 2011; Han et al. 2014; Mariconti et al. 2016). High levels of sHLA-G are also associated with increased susceptibility to malaria (Sadissou et al. 2014; D'Almeida et al. 2016, 2017) and to human African trypanosomiasis (HAT) (Gineau et al. 2016). In contrast, decreased membrane-bound HLA-G has been reported in placenta of patients infected with *P. falciparum* (Sartelet et al. 2005) and in heart and colon of patients presenting Chagas disease (Dias et al. 2015).

Regarding the *HLA-G* gene, several variable sites have been described at regulatory (5' URR and 3' UTR) and coding regions (Castelli et al. 2011, 2014a, b). Divergence exists regarding the exact position of the first ATG (exon 1). Castelli et al. provided a more comprehensive description of the *HLA-G* genetic structure, taking into account the differences

in nomenclature reported by the NCBI/Ensembl and the IPD-IMGT/HLA databases **(Figure 6).** IPD-IMGT/HLA considers the first mRNA segment that is translated as exon 1, while this segment is considered as exon 2 at NCBI/Ensembl (Castelli et al. 2014a). The two probable starting positions of the first translated ATG, as proposed by NCBI/Ensembl (Castelli et al. 2014a), belong to the same frame and might not initiate translation because these sequences do not match the Kozak consensus, i.e., the preferred translation initiation sequence (Nakagawa et al. 2008). In contrast, the third starting position of the first translated ATG, proposed by IMGT/HLA (Castelli et al. 2014a), can initiate the translation and also well define the beginning of the coding sequence (CDS) segment. Moreover, it corresponds to the preferred translation initiation sequence (Nakagawa et al. 2008). Considering the reference transcripts (NCBI/NM_002127.5 and Ensembl/ENST00000428701) as reported by Castelli et al. (Castelli et al. 2014a), the *HLA-G* 5' UTR is formed by a large segment, including intron 1 **(Figure 6)**. This intron is spliced out leading to one promoter segment, which is combined to two adjacent exons (that together, is perhaps considered as exon1 at IMGT/HLA). Indeed, this *HLA-G* 5' UTR segment harbors three potential starting positions: two in 5' UTR and the third one coincides to the first nucleotide A (+1) of the first ATG codon, and can be also considered as the beginning of CDS. Basing on the human genome hg19 transcript (NC_000006.12), used as reference, all nucleotides upstream the +1 have been annotated as negative number and the nucleotides in the CDS segment as positive numbers. Therefore, as defined by the IPD-IMGT/HLA, the *HLA-G* coding region comprises positions from -300 to +2838 (Castelli et al. 2014a).

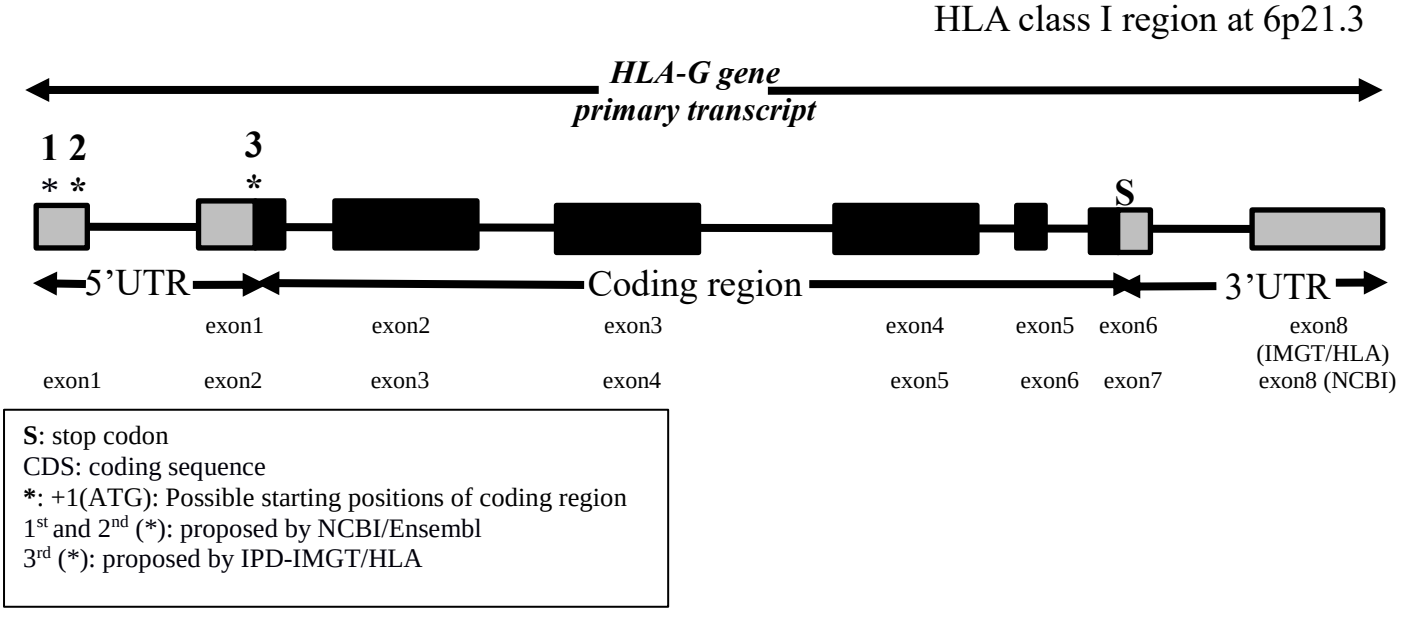


Figure 6. *HLA-G gene structure (Adapted from* (Castelli et al. 2014a)).

The 60 official recognized *HLA-G* alleles (IPD-IMGT/HLA v.3.32.0, 2018-04-16) are randomly reorganized into 22 specific allele groups that encode 3 null (G*01:05N, G*01:13N and G*01:21N) and 19 full-length protein variants. Up to date, approximately 14 variabe sites are frequently observed at the 3' UTR segment in worldwide populations, which are rearranged into several 3' UTR haplotypes, with UTR-01 to -07 the most frequent (minimum frequency of 4%) (Castelli et al. 2010, 2011, 2014a, 2017; Lucena-Silva et al. 2012; Santos et al. 2013; Sabbagh et al. 2014; de Albuquerque et al. 2016; Nilsson et al. 2016). Many studies conducted in several worldwide populations have evaluated *HLA-G* variability at the 5' upstream regulatory region (promoter), coding and/or 3' UTR segments (Tan et al. 2005; Hviid et al. 2006; Castelli et al. 2010, 2011; Consiglio et al. 2011; Lucena-Silva et al. 2012, 2013; Sizzano et al. 2012; Veit et al. 2012, 2014; Carlini et al. 2013; Martelli-Palomino et al. 2013; Santos et al. 2013; Courtin et al. 2013; Garcia et al. 2013; Sabbagh et al. 2014; Catamo et al. 2014, 2015; Gineau et al. 2015; Porto et al. 2015; Garziera et al. 2015; Nilsson et al. 2016; Zambra et al. 2016; de Albuquerque et al. 2016). But, with the exception of a study reporting *HLA-G* variability based on the *1000 Genomes phase I* data (Castelli et al. 2014a), which is characterized by low depth of coverage, the complete *HLA-G* gene variability has not yet been characterized in African population samples.

1.7.2. Overview on Non-Classical Class I (HLA)-E

HLA-E was first detected in different lymphoid and malignant cells (Koller et al. 1988) and, later, its expression was also reported in placenta and extra-villous membrane at first trimester of gestation and at term (Wei and Orr 1990). HLA-E has also been associated with maternal-fetal tolerance and with pregnancy-related disorders (Persson et al. 2017). HLA-E binds to the leader self-peptides derived from other MHC class I molecules (e.g., the aa35–44 peptide from HLA-A2) (Nattermann et al. 2005; Persson et al. 2017) and binds to the inhibitory CD94-NKG2A receptor expressed by NK cells (Nattermann et al. 2005). The recognition of HLA-E-self-peptide complexes by NK cell is a checkpoint for the immune surveillance to avoid NK cytotoxicity (Borrego et al. 1998; Lee et al. 1998; Braud et al. 1999). HLA-E is also a ligand for T CD8 cell receptor (Pietra et al. 2009, 2010), and may present non-self-antigens derived from cytomegalovirus, human immunodeficiency virus, Epstein-Barr virus, influenza and hepatitis C viruses (Sullivan et al. 2008; Pietra et al. 2009, 2010), inhibiting NK cell cytotoxicity, propitiating the proliferation of viral cells. In addition, HLA-E may activate the adaptive immune response through the cytotoxic T cell activation

(Kochan et al. 2013; Marchesi et al. 2013), and also plays a pivotal role in chronic infections and in the context of cellular stress (Jørgensen et al. 2012).

Regarding the *HLA-E* gene, considering as nucleotide +1, the Adenine of the first translated ATG, the *HLA-E* 5′ upstream segment (distal promoter and proximal promoter - 5′ UTR), encompasses nucleotides from -2143 to -1. The coding segment is covered by the IPD-IMGT/HLA database from nucleotide -300 to +3522. The 3′ UTR/exon 8 encompasses nucleotides from +3218 to +4674 (Ramalho et al. 2017).

Currently, 27 *HLA-E* alleles encoding 8 full-length proteins and one null allele (*E**01:08N) have been officially reported (IPD-IMGT/HLA v.3.32.0, 2018-04-16), but only two proteins (E*01:01 and E*01:03) are frequently detected worldwide (Arnaiz-Villena et al. 2007; Liu et al. 2012; Veiga-Castelli et al. 2012; Carvalho dos Santos et al. 2013; Felício et al. 2014; Ramalho et al. 2017). The E*01:01 protein differs from E*01:03 by the presence of an Arginine, instead of a Glycine, encoded by codon 107 at exon 3 and both usually occur at similar frequencies. This may suggest that balancing selection has maintained both alleles (Tamouza et al. 2007; Veiga-Castelli et al. 2012; Ramalho et al. 2017), but the mechanisms underlying this phenomenon are unclear. Many studies have explored *HLA-E* exonic variability in different population samples, including Brazilians (Veiga-Castelli et al. 2012, 2016; Carvalho dos Santos et al. 2013; Felício et al. 2014), Africans (Guinea–Conakry and Burkina Faso) (Castelli et al. 2015), and worldwide population samples from the *1000 genome phase 1* panel (Felício et al. 2014). However, only few studies have evaluated the complete *HLA-E locus*, including the regulatory segments, in Asian (Olieslagers et al. 2017) and Brazilian (Ramalho et al. 2017) samples, and in cell lines from Asian, African American, and European samples (Pyo et al. 2006).

1.7.3. Overview on Non-Classical Class I (HLA)-F

The *HLA-F locus* has been unraveled in the recent years as an immunomodulatory gene (Goodridge et al. 2010). Similar to HLA-G, its expression in the tumor leads to tumor cell escape from the NK cell cytotoxicity and consequently to poor prognosis of the disease (Harada et al. 2015; Xu et al. 2015). HLA-F may play an important role on the maintenance of immunotolerance in feto-maternal interface, since it is expressed at the extravillous trophoblasts, during the first and second trimesters (Ishitani et al. 2003; Shobu et al. 2006), but it is absent at the third one (Hackmon et al. 2017). In addition, HLA-F may be expressed as an open conformer molecule, and physically interacts with inhibitory (KIR3DL2) and activating receptors (KIR2DS4, KIR3DS1) of NK cells and with ILT-2 (Goodridge et al.

2013; Garcia-Beltran et al. 2016). It can also present peptides to T cells and regulate immunity through interactions with NK cell receptors (Dulberger et al. 2017).

Considering the transcript variant NM_018950.2 in Human genome hg19 or hg38, *HLA-F* gene structure is the closest to that accepted by the IPD-IMGT/HLA database (Robinson et al. 2015). Starting from the Adenine of the first translated ATG as nucleotide +1, *HLA-F* gene is fragmented into upstream promoter (segment between -556 and -124), 5' UTR (segment between -124 and -1), coding region (segment between +1 and +2944), 3' UTR (segment between +2945 and +3063) and 3' downstream (all nucleotides downstream the position +3063) (Lima et al. 2016).

Being less polymorphic than *HLA-G*, the 30 alleles officially recognized for *HLA-F* encode five full-length proteins (F*01:01, F*01:02, F*01:03, F*01:04 and F*01:05) (IPD-IMGT/HLA v.3.32.0, 2018-04-16), but only F*01:01 (82.45%) and F*01:03 (16.03%) are frequently observed in worldwide populations (Lima et al. 2016). Some studies have genotyped only the well-described *HLA-F* polymorphisms in Korean and Chinese populations (Moscoso et al. 2007; Kim et al. 2011; Zhang et al. 2012; Pan et al. 2013); however, few reports are available regarding the nucleotide variability of *HLA-F* gene segments or of the entire gene region in Brazilians (Manvailer et al. 2014; Lima et al. 2016) and Asian, African American, and Caucasian samples (Pyo et al. 2006).

1.8. *HLA-IB* POLYMORPHISMS AND *S. HAEMATOBIUM* INFECTION

There are no studies addressing the influence of *HLA-G, -E* and *-F* polymorphisms in urogenital schistosomiasis. Nevertheless, there are studies in other parasitic diseases such as Human African Trypanosomiasis (HAT) and American Trypanosomiasis (or Chagas disease), where the *HLA-G* 3' UTR polymorphisms exhibited distinct patterns of associations with both HAT and Chagas disease (Courtin et al. 2013; Dias et al. 2015; Sabbagh et al. 2018).

1.9. *HLA-IB* POLYMORPHISMS AND *P. FALCIPARUM* MALARIA

The influence of *HLA-G* polymorphisms in *P. falciparum* malaria has been shown in two important family-based association studies. A first family-based association study performed on individuals from Niakhar, Senegal, reported that the +3187G allele was associated with higher malaria transmission to children and lower level of parasite density during asymptomatic *P. falciparum* infection. The *HLA-G* 3' UTR haplotype known as UTR-1 was associated with a decreased level of parasite density during asymptomatic infection under a dominant model, whereas the *HLA-G* UTR-3 haplotype was associated with an

increased level of parasite density during the follow-up and increased intensity of asymptomatic infection under a recessive model (Garcia et al. 2013). A second family-based association study also conducted in the same Senegalese cohort has tested the association of *HLA-G* 3'UTR variants with acquired anti-malarial humoral immunity. The +3010G and +3142C alleles were over transmitted to children with increased total immunoglobulin G (IgG) and IgG1 antibodies levels against Glutamate Rich Protein (GLURP) of *P. falciparum*, and the +3196G allele had a preferential transmission to children with a lower IgG3 response against Merozoite Surface Protein2 (MSP2). The *HLA-G* UTR-*2* haplotype was associated with a decreased IgG3 response against MSP2, suggesting a role of *HLA-G* on the regulation of immune humoral response during *P. falciparum* infection (Sabbagh et al. 2013).

Until now, no studies have addressed the role of *HLA-E* and *-F* polymorphisms in susceptibility to *P. falciparum* malaria.

2. PROBLEMATIC

2.1. PART I: POPULATION GENETIC STUDY

Sub-Saharan African populations are characterized by high levels of intra-population nucleotide and haplotype diversity and inter-population genetic differentiation; however, little attention has been devoted to the study of immunomodulatory HLA genes in these populations, which represent melting pots for several ethnic groups. In Benin, about 50 sociolinguistic groups are spread over specific geographical areas. Although Fon (39.2%), Adja (15.2%), Yoruba (14.5%) and Bariba (9.2%) are the major ethnic groups (http://www.axl.cefan.ulaval.ca/afrique/benin.htm, accessed May 18, 2018), many other groups deserve attention, including the Toffin population, which has its own culture, language and customs. Studies investigating the nucleotide diversity of non-classical HLA class I genes in the Beninese population have primarily focused on some regulatory segments of the *HLA-G* gene, including the promoter region (Gineau et al. 2015) and 3' UTR (Sabbagh et al. 2014) in the Tori ethnic group, living in the Tori-Bossito area in Southern Benin.

Before performing studies regarding the malaria/schistosomiasis co-infection, this section aimed to evaluate *HLA-G, -E* and *-F* entire gene variability by massive parallel sequencing (Next Generation Sequencing; NGS) in Beninese Toffin population. The Toffin, etymologically ''people of water'' is grouped in more than 10 villages, representing small autonomous communities, fugitives from the fights headed by the French soldiers during the pre-colonial period, and have lived simple lives as fishermen for centuries (Principaud 1995).

2.2. PART II: GENETIC ASSOCIATION STUDY

In Benin malaria is seasonal and people are frequently exposed to the risk of uncomplicated or severe malaria, depending on the epidemiological factors. Considering that: i) Toffins live in lake areas and are highly exposed to worms and malaria, ii) immune responses are specific to each parasite and even to species, in the presence of both parasites (*S. haematobium* and *P. falciparum*), a reciprocal influence on the immune system may occur, as reported by some previous studies. In these conditions, what could be the influence of *S. haematobium* on the host immune response against *P. falciparum* and vice versa in co-infected individuals? We hypothesize that the non-classical MHC genes and molecules (especially *HLA-G*) could be associated with the *S haematobium/P. falciparum* co-infection, and to investigate this association we took advantage of the longitudinal PALUCO (*Paludisme et Co-infection*) study to understand the influence of *S. haematobium* worm carriage on *P. falciparum* malaria infection in the Beninese Toffin children.

3. OBJECTIVES

3.1. MAIN OBJECTIVE

This study aimed **i)** to evaluate the diversity of the HLA non-classical class I genes in the Beninese Toffin population, and **ii)** to identify susceptibility or resistance factors within HLA non-classical class I genes associated with *P. falciparum* malaria and/or urogenital bilharziasis in a cohort of children of the Toffin population.

3.2. SPECIFIC OBJECTIVES

Part I:

To identify the HLA non-classical class I *HLA-G, -E* and *-F* gene variability (polymorphisms and haplotypes) at promoter, coding and 3' UTR segments in the Toffin population.

Part II:

1) To perform an genetic association study between *HLA-G, -E* and *-F* genes and the susceptibility or resistance to *P. falciparum* malaria and/or to *S. haematobium* infection;
2) To identify polymorphisms or haplotypes that may modify the risk of malaria infection in children co-infected by *S. haematobium.*

4. MATERIALS AND METHODS

The present project "***Plasmodium falciparum* and *Schistosoma haematobium* co-infection: role of non-classical HLA class I genes (*HLA-G, -E and -F)* in susceptibility to malaria**" represents the genetic component of a major project entitled PALUCO (*Paludisme et Co-infection* - Malaria and Co-infection) developed in Benin, West Africa.

4.1. THE PALUCO PROJECT

The project started on April 1rst, 2012 at the Faculty of Health Sciences (Cotonou, Benin) and was funded by IRD (€ 50,000, for 36 months), and was approved by the Ethics Committee (No.12/03/2012/CEIFSS/UAC) of the "*Faculté des Sciences de la Santé* (FSS)" of Cotonou (economical capital of Benin). It was conducted at the Sô-Ava region, Southern of Benin, 12 km at North of Cotonou (**Figure 7)**, endemic for both *P. falciparum* and *S. haematobium* infections, and encompassed researchers from Benin, France and Brazil.

Sô-Ava is located between 6° 24' and 6° 38' latitude North and between 2° 27 'and 2° 30' longitude East. With an area of 218 km², it has 118,547 inhabitants, of whom 45,493 inhabitants or 38.38% are active workers, primarily laboring as farmers, fishers and hunters, exerting their professional activities in and around Lake Nokoué. The population density of Sô-Ava is 567 inhabitants per km² (Institut National de la Statistique et de l'Analyse Economique 2015). Sô-Ava area was selected because two studies carried out by Ibikounlé et al. (Ibikounlé et al. 2009, 2013) attested of the endemicity of urinary schistosomiasis in the locality. The study took place in three public primary schools: Sô-Tchanhoué (ST), Sô-Tchanhoué-quartier (STQ) and Sô-Zounko (SZ), during two malaria transmission seasons (May to November 2012 and May to November 2013) (**Figure 7**).

Source: Topographic map of Benin at 1/200000

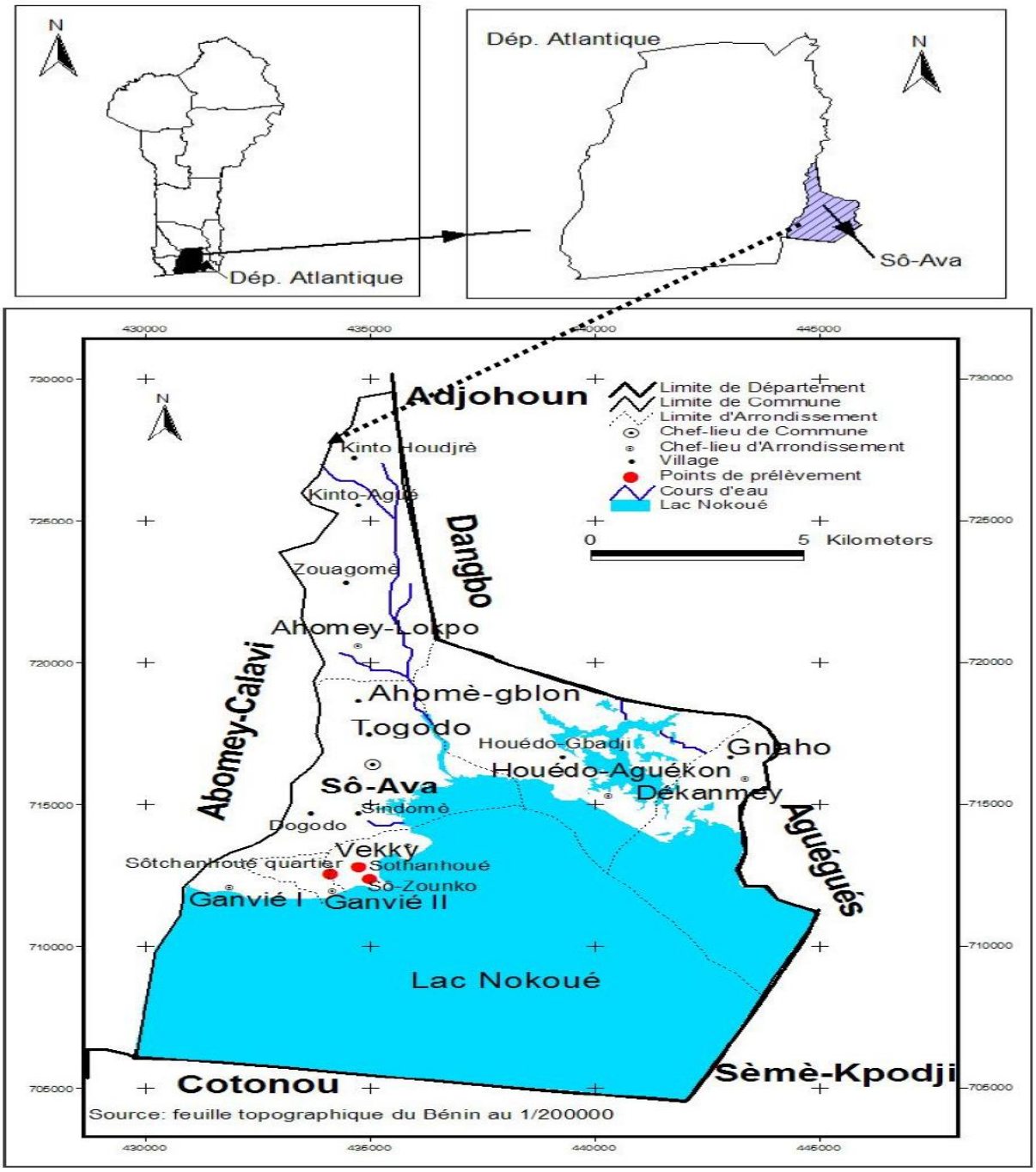
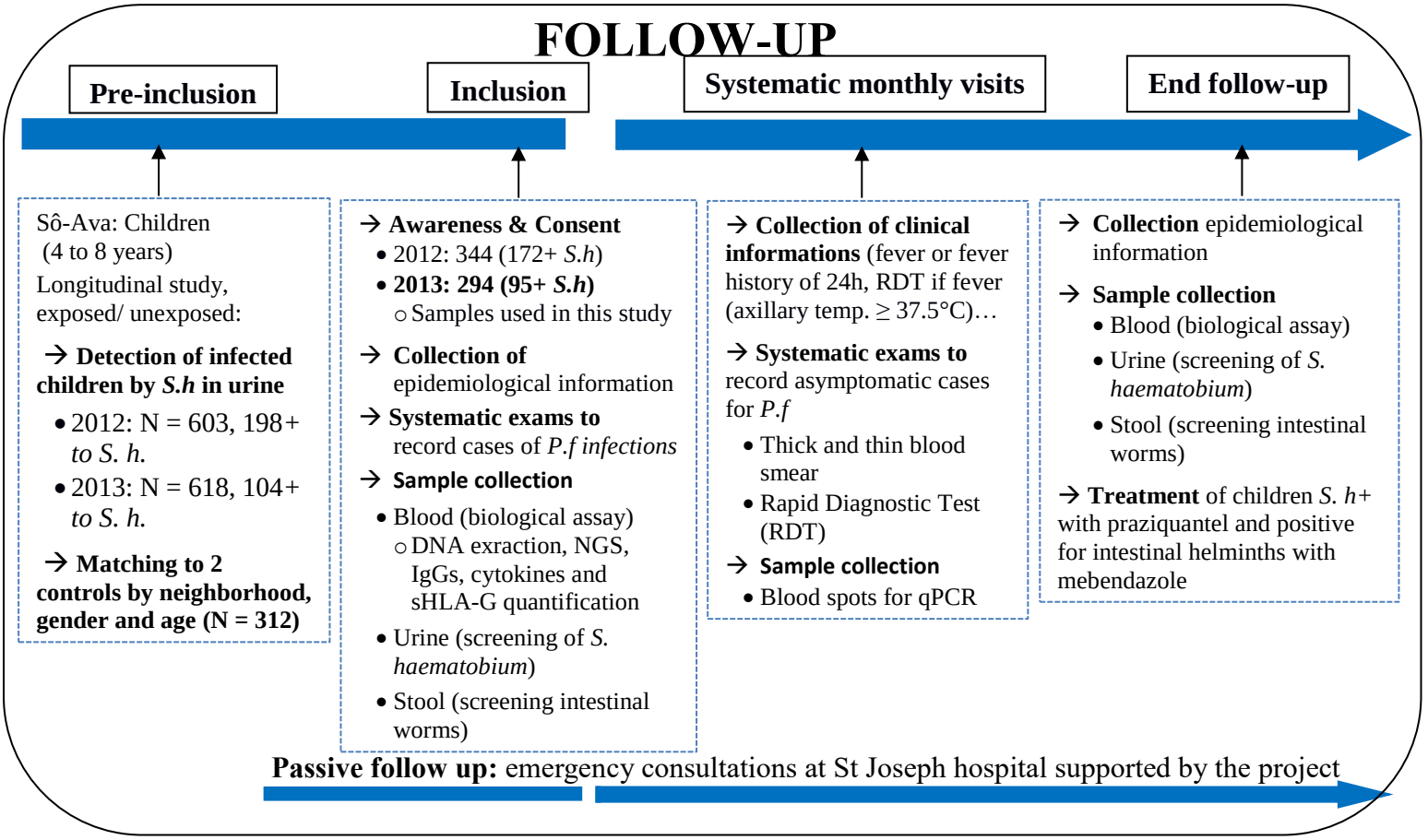


Figure 7: *Location of the surveyed schools. Red colors indicate the sites of study.*

The objective of the "PALUCO" project is to study the impact of infection with schistosomes on susceptibility to malaria in children aged 4 to 8 years, evaluating the relationship between frequency and level of infection with *S. haematobium* and **i)** malaria attacks (period or time of first infection, parasite density and the number of malaria attacks), **ii)** levels of antibodies, especially immunoglobulin-G (IgG) specifically targeted against the most promising *P. falciparum* vaccine antigens (AMA-1, MSP1, MSP2, MSP3, GLURP and

CSP), **iii)** the functionality of anti-*P falciparum* IgG, **iv)** levels of proteins (cytokines/soluble HLA-G) that modulate humoral responses, **v)** non-classical HLA class I gene polymorphisms in the individual response to malaria. The objectives **i** to **iv** were carried out in Benin and objective **(v)** was carried out in Brazil at the Ribeirão Preto Medical School, University of São Paulo (FMRP-USP), and it is the subject of the present thesis.

Six hundred and thirty-eight (638) children (4 to 8 years) were followed during two malaria transmission seasons. The pre-inclusion and inclusion were performed as shown in **Figure 8**.



S.h*+:** positive for *Schistosoma haematobium* after urine screening, **NGS:** Next Generation Sequencing**, IgGs:** anti-*P. falciparum* total IgG, IgG1, IgG3 and anti *Anopheles* saliva IgGs, **Cytokines:** *Il-10*, *TNF-α*, *IFN-γ* and *IL-6*, **sHLA-G:** soluble HLA-G, **RDT:** Rapid Diagnostic Test, ***P.f: *Plasmodium falciparum*, **qPCR:** quantitative real-time PCR.

***Figure 8*: PALUCO project: Description** of the follow-up

During the first malaria transmission season (May to November 2012), after urine screening in volunteers in June 2012, 603 children were pre-selected, of whom 198 were infected with *S. haematobium*. Of whom, 344 were really included (July-August 2012) after signing the informed consent form by their parents. Of these 344 included individuals, 172

were infected by *S. haematobium*. During the second malaria transmission season (May to November 2013), 618 children were pre-included, of whom 104 were infected with *S. haematobium*. Of the 618 children, 294 were really included after an awareness session. In these 294 included, 95 were infected with *S. haematobium*. The exposed children were matched in terms of sex, age and residence village with the non-exposed ones at each malaria transmission season.

At inclusion, a medical examination of the children was performed and stool and blood samples were collected for biological tests. The children were followed-up from July to November 2012, and parents were instructed to consult the medical centre for any health problem of their children. The project freely provided the systematic care for patients in collaboration with the St Joseph Medical Centre, structure set up by the Ministry of Health of Benin. Uncomplicated cases of malaria were treated with quinine 10 mg/kg in tablets *tid* for 7 days and/or injections depending on the clinical state of children. A malaria attack is defined as the combination of axillary temperature ≥ 37.5°C or fever history the day before the visit, and a positive rapid diagnostic test (RDT). Between inclusion and the end of the follow-up, children were actively monitored by monthly visits, during which physicians took medical history with primary emphasis on history of fever within 24 hours, made RDT in case of T ≥ 37.5 °C and thick and thin blood smears. At the end of the season (November 2012), stool, urine and blood samples were also collected for all participants for intestinal worms, urinary schistosomiasis and *P. falciparum* malaria detection. Participants who were positive for schistosomiasis were treated with a single dose of oral praziquantel (40 mg/kg), while those positive for soil-transmitted helminths were treated with albendazole 400 mg at a single oral dose.

The same follow-up and biological assays were performed during the second malaria transmission season (May to November 2013).

4.2. CRITERIA FOR INCLUSION AND EXCLUSION OF CHILDREN IN THE PALUCO PROJECT

- **Inclusion Criteria**

✓ Aged between 4 and 8 years;

✓ Being diagnosed positive or not to *S. haematobium*,

✓ Being present on the day of inclusion;

✓ Being permanent resident (even during holidays) in the study area;

- ✓ Agree to sign the informed consent form.

- **The exclusion criteria**

- ✓ Presence of a chronic disease that may affect the immune response;
- ✓ Presence of a symptomatic infection with *S. haematobium hematuria*, abdominal pain or acute renal failure
- ✓ And, finally, reside more than 10 km from the health center.

4.3. THE LABORATORY TESTS PERFORMED DURING THE FOLLOW-UP

- At inclusion, blood samples collected were used for the obtention of DNA and plasma.
- Thick blood smears were used to detect *P. falciparum* infect*ion and paras*ite density, and urine samples were screened for *Schistosoma haematobium* infection and parasite density detection.
  - The density of *Schistosoma haematobium* was reported as eggs/10mL of urine and was detected as recommended by WHO (2002). After homogenization, 1 mL of urine was deposited in a Petri dish. The eggs were directly counted using microscopy. This operation was repeated 3 times and the mean number of eggs obtained was multiplied by 10 and reported as density of *Schistosoma haematobium* parasites.
- Once a month during the transmission seasons, one thick blood smear (TBS) and three blood spots were systematically performed to detect asymptomatic infection and parasite density by quantitative real-time PCR (qPCR). A diagnosis of *S. haematobium* was performed at the end of follow-up.

These biological exams were organized into four data sets:

- **The immunological data:** anti-*P. falciparum* total IgG, IgG1, IgG3, IgG, cytokines (Il-10, TNF-α, IFN-γ and IL-6) and soluble HLA-G (sHLA-G).
- **The parasitological data:** be infected or not by *P. falciparum* or *S. haematobium*, parasite density, time of first infection, co-infection.
- **Clinical data:** be symptomatic (positive for *P. falciparum* with fever) or asymptomatic (positive for *P. falciparum* without fever), number of malaria attacks.
- **Other variables:** age, sex, neighborhood, hygiene (presence or not of domicile latrines), father and mother profession (farmer, hunter or fisherman…).

Indeed, clinical, immunological, parasitological, hematological and demographical data were recorded during the follow-up and are available for the genetic association study.

Although the PALUCO project had many objectives, this study primarily focused only on the genetic component, which was carried out using 294 samples from the second malaria transmission season (May to November 2013).

4.4. STUDIED SAMPLE

Two hundred and eighty-eight (patients and controls) children of the Toffin ethnic group, aged 4 to 8 years (mean age: 6.7 years ±0.80), were included in the present study, of whom 187 (64.9%) were boys and 101 (35.1%) girls. At the end of the follow-up, we recorded per child, an average number of five monthly visits (MV, average = 5) and two emergency consults (EC). Then, we included in the genetic association analysis, all unrelated children (exclusion of close relatives up to the second degree), who were registered at the end of the follow-up: **a)** at least an average number of three MV (eligible for *P. falciparum* malaria association analysis) and **b)** all unrelated children included at inclusion were eligible for *S. haematobium* association analysis. Finally, 281, 283 and 284 children were successfully sequenced for *HLA-G*, *HLA-E* and *HLA-F*, respectively, as shown in Table 1.

Table 1. Characterization of studied population

Projects	**Follow-up (from inclusion to end)**	**Study population**	**N**
PALUCO Project			
Follow-up n°1	May to November 2012	4 to 8 years	344
Follow-up n°2	May to November 2013	4 to 8 years	294
Human genetics approach	**Description**	Study population	N
Follow-up 2	May to November 2013	4 to 8 years, 6.7 mean age (±0.80)	288
Boys	64.9%	_	187
Girls	35.1%	_	101
Average number of monthly visits (MV) during the follow-up, all children	_	_	5
Average number of emergency consults (EC) during the follow-up, all children	_	_	2
Criteria of inclusion in malaria analysis	Average number of MV ≥ 3 and be unrelated	_	146
Criteria of exclusion from malaria analysis	Average number of MV < 3		8
Eligible for malaria analysis	Average number of MV ≥ 3 and be unrelated	_	146
Criteria of inclusion in schistosomiasis analysis	All unrelated child included at inclusion	_	154
Eligible for schistosomiasis analysis	All unrelated child included at inclusion	_	154
Sequenced	_	_	288
HLA-G	Successful sequenced	_	281
HLA-E	Successful sequenced	_	283
HLA-F	Successful sequenced	_	284

4.5. ETHIC CONSIDERATIONS

PALUCO was approved by Ethics Committee of the ‘*’Faculté des Sciences de la Santé* (FSS)’’ of Cotonou, Benin and registered at No.12/03/2012/CEIFSS/UAC and the genetic component of the PALUCO project was approved by the “*Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto-Universidade de São Paulo*” (HCRP e FMRP-USP) and registered at N° CEP: FMRP-N°710/2016/CEP/MGV. The informed consent was obtained from parents of all participants. Before starting the study at FMRP-USP, we created a biorepository at the laboratory of Pr. Dr. Eduardo Antônio Donadi, where samples were stored under high confidence. All ethic considerations are presented at the Annex section (I to V).

4.6. DNA EXTRACTION

DNA samples were extracted using a commercial kit, according to manufacturer’s instructions (Thermo Fisher Scientific, Waltham, MA). The extracted genomic DNA was stored at -20°C until use.

4.7. DNA AMPLIFICATION AND VISUALIZATION

All DNA samples were diluted at 50 ng/µL; the concentration considered to be optimal for polymerase chain reaction (PCR) reaction. The complete *HLA-E, -F* and *-G* gene segments, including nucleotides upstream the first translated ATG, all exons, all introns, the complete 3’ UTR and 3’ downstream segments were amplified from 288 DNA samples in the same final reaction volume of 50 µL. The *HLA-E* and *-G* genes were amplified using Takara High Fidelity R050A polymerase (PrimeStar GXL DNA Polymerase, Takara, Mountain View, CA), with 10 µL of 5X PrimeSTAR GXL Buffer, 4 µL of dNTP Mixture (2.5 mM each), 1.5 µL of primer at 10 pmol each (forward and reverse), 1 µL of PrimeSTAR GXL DNA Polymerase, 28 µL of Sterilized DNA-free water (27 µL for *HLA-E*) and 4 µL of DNA (50 ng/µL) (5 µL for *HLA-E).* DNA amplification was performed using the following cycling conditions: 30 cycles (35 cycles for *HLA-E*) at 98°C for 10 seconds (denaturation), 60°C for 15 seconds (hybridization or annealing), 68°C for 6 minutes (8 minutes for *HLA-E)* (polymerization) and 4°C overnight. The primers used were: HEPR.F1- 5' GCTTCGCAGTGAATGTGGCA 3' (Forward) and HEUTR.R1- 5' GGACTCCCTGGGCTTTCTCACCG 3' (Reverse) for *HLA-E* and HGPR.F1- 5' ACACTCATAATTCATTCATTCAGC 3' (Forward) and HGUT.R1- 5' TCTTCTGATAACACAGGAACTTC 3' (Reverse) for *HLA-G.*

We amplified *HLA-F* using the Long PCR Enzyme Mix polymerase (Fermentas Int. Inc., Burlington, Ontario). The PCR mixt contained 4 µL of 10X Long PCR buffer with $MgCl_2$ (15 mM), 0.6 µL of $MgCl_2$ (25 mM), 2.5 µL of dNTP Mixture (5 mM each, Fermentas), 1.5 µL of primer at 10 pmol each (forward and reverse), 0.3 µL of Long PCR Enzyme Mix polymerase, 38.1 µL of sterilized DNA-free water and 1.5 µL of DNA (50 ng/µL). The following amplification conditions were used: 1 cycle at 94°C for 5 minutes followed by 15 cycles of 30 seconds at 94°C, 45 seconds at 63°C and 7 minutes at 68°C, 25 cycles of 30 seconds at 94°C, 45 seconds at 63.5°C and 7 minutes at 68°C, 1 cycle of 4 minutes at 68°C and hold at 4°C overnight. The *HLA-F* primers used were: HFPR.F1- 5' GGAGAGAACACTCAGGTGGC 3' (Forward) and HFUTR.R1- 5' CCACTAAACACCCAGCCCAT 3' (Reverse).

The quality of PCR products was assessed on 1% gel agarose (Promega Corporation, Madison, WI) stained with GelRed (Biotium TM, Hayward, CA).

4.8. SAMPLES PREPARATION FOR NEXT GENERATION SEQUENCING (NGS)

Firstly, we proceeded to quantification of amplified products. Using the Qubit® dsDNA HS Assay kit (Thermo Fisher Scientific), we assessed the exact concentration of each amplicon of each gene. This step allowed us to normalize and pooled together the three HLA *class I* genes amplified, before sequencing. Five µL of each pool were purified, using ExoSap (GE Healthcare Life Sciences, Pittsburgh, PA). After a second Qubit quantification of the pool, we diluted each one at 0.2 ng/µL, the input concentration required for NGS sequencing.

4.9. CANDIDATE GENE SEQUENCING

After library preparation with Nextera®XT DNA Sample Preparation Kit (Illumina, San Diego, CA), quality checking was performed using the Bio-Analyser 2100 (Agilent Technologies, Santa Clara, CA) to evaluate the length of the amplicons (from ~250 bp to 1000 bp) and the quantitative real-time PCR (qPCR) was performed, for library quantification, using the KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, MA). Library was sequenced using the Illumina V2 500 cycles flow cell (Illumina) in a paired-end mode (2x250bp).

4.10. READS MAPPING PROCESSING

Prior to mapping, the reads (short DNA segments produced by NGS) were filtered, a step that is quite important for mapping (*hla-mapper software,* (Castelli et al. 2015; Lima et al. 2016), available at www.castelli-lab.net. This software performs a series of filters allowing a suitable mapping procedure when several HLA genes are sequenced together. By using *hla-mapper 2.0* and database version 002.1 with default parameters, we filtered our reads to keep only sequences larger than 80 nucleotides, and presenting at least 80% of bases with a minimum quality of 25 (phred ≥ 25). If a read is removed by any of the applied filter, the entire pair was removed. After the removal of small or low-quality sequences, PCR duplicates (same forward and reverse sequence) were removed. Then, *hla-mapper* addressed each sequence to its specific *HLA* gene (*HLA-E, -F* and *-G*). Finally, the BAM files were generated with all sequences mapped to the reference genome version hg19 (Castelli et al. 2015; Lima et al. 2016) and visualized by Integrative Genomics Viewer (IGV) v.2.3.81 (Robinson et al. 2011; Thorvaldsdóttir et al. 2013).

4.11. VARIANT AND GENOTYPE CALLING

The strategies used here are the same already described elsewhere (Castelli et al. 2015; Lima et al. 2016). By using the Genome Analysis Toolkit (GATK, version 3.7) Haplotype-Caller algorithm, we inferred genotypes using the GVCF mode that generates a specific *HLA.g.vcf* file per individual. After, a VCF (variant call format) file was generated, concatenating all samples together with the GATK GenotypeGVCF algorithm in an unique VCF file per gene (McKenna et al. 2010; DePristo et al. 2011; Van der Auwera et al. 2013). The mixed VCF file generated by GATK was processed by vcfx (available at www.castelli-lab.net), which applied a series of rules to get a reliable genotype calling and assure that only high quality homozygous and heterozygous genotypes are passed forward to the phasing and imputation procedures to be applied afterwards, as detailed by Castelli et al. (Castelli et al. 2015). By using vcfx v1.0, function checkpl (with the minimum genotype likelihood set to 99.9%), which guarantees that only high quality genotypes continue to a further imputation step, we inferred genotypes using GATK and chromosome 6 sequence (hg19) as draft.

4.12. PHASING AND HAPLOTYPE INFERRING

The phasing was performed using two approaches. The first strategy used the GATK routine ReadBackedPhasing, setting the *phaseQualityThresh* parameter at 2000, to assure that only alleles from variable sites that are very close to be present in a same read should be

phased. Considering that variable sites may be quite distant from each other in a same sample and that ReadBackedPhasing performs phase inference only on Single Nucleotide Variation (SNVs) sites, the remaining variable sites (indels and multi-allelic loci) were inferred by using PHASE algorithm (Stephens et al. 2001; Stephens and Donnelly 2003). Prior to phasing, all singletons; i.e., variable sites detected in only one individual and in a heterozygous state, were removed. For the second strategy, we used the *phase-using-known-from-readbackedphasing.pl* perl *Version 2.5* (Castelli et al. 2015; Lima et al. 2016), which used data from a VCF file generated by the GATK ReadBackedPhasing algorithm to create a fragmented *.known* file, that is applied to the PHASE algorithm to fill in the blanks. This *.known* file is usually "fragmented", because GATK ReadBackedPhasing may phase some groups of variable sites, but does not inform the association (or the phase) between these groups. Then, *phase-using-known-from-readbackedphasing.pl* perl runs the PHASE algorithm, considering each of these fragments, compares the PHASE results of multiple runs. The following PHASE algorithm was used: *perl 'phase-using-known-from-readbackedphasing.pl' -t 6 -v 'ReadBackedPhasing_singleton.VCF' -b 'phase.2.1.1.linux/PHASE' -x 'vcfx' -p '1000 1 100'*. Finally, we accepted the samples which present in all runs the same pair of haplotypes when compared to those with the known phases detected by GATK (with no exceptions). The *phase_to_vcf.pl* perl was used to convert the most probable haplotype pair of each sample obtained with the PHASE algorithm into a phased VCF file. We manually included the singletons in the phased VCF file following the criteria established by Castelli et al. (Castelli et al. 2017) .

4.13. HAPLOTYPE DEFINITION

The phased VCF file containing the singletons (manually included) was then converted into a complete *HLA* FASTA (HLA sequences and CDS sequences) file using application *vcfx* function fasta (www.castelli-lab.net/apps/vcfx) and hg19 reference sequence as a draft. By using a IPD-IMGT/HLA database (Robinson et al. 2015) version 3.31.0 in a local BLAST server, we defined for each haplotype the closest known HLA coding alleles. Basing on the previous described nomenclature, we defined the promoter, 3' UTR and extended haplotypes for *HLA-E*, *HLA-F* and *HLA-G* (Tan et al. 2005; Castelli et al. 2010, 2011, 2014a, 2015, 2017; Lucena-Silva et al. 2012; Santos et al. 2013; Sabbagh et al. 2014; Gineau et al. 2015; de Albuquerque et al. 2016; Lima et al. 2016; Nilsson et al. 2016; Ramalho et al. 2017), since no official nomenclature has been assigned yet to these HLA gene regions.

4.14. PHENOTYPE DEFINITION

4.14.1. Phenotypes definition for S. haematobium infection

The following phenotypes were used in this analysis: **i)***''Infected and non-infected''* and **ii)** *''Parasite density''*. We considered infected all individuals who, at the inclusion, exhibited one or more *S. haematobium* eggs/10 ml of urine, while the *''non-infected''* exhibited no eggs in their urine after microscopical examination. After, we separated children with moderate infection (10 to 49 eggs/10 mL) from those with heavy infection (≥ 50 eggs/10 mL). We did not have children with less than 10 eggs/10 mL to be classified as a light infection group (Briand et al. 2005; Lemaitre et al. 2014; Doumbo et al. 2014; Degarege et al. 2016). In addition to these two phenotypes, the major co-variables included gender, age and water-related activity (Briand et al. 2005; Lyke et al. 2005; Satayathum et al. 2006; Mbanefo et al. 2014), as shown in Table 2.

Table 2. Variables included in *S. haematobium* genetic association analysis

Genes/Phenotypes	**Description**	**N**
Genes		
Included in schistosomiasis analysis	All unrelated child included at inclusion	**154**
HLA-E	Clinical and genetic available data	151
HLA-F	Clinical and genetic available data	152
HLA-G	Clinical and genetic available data	149
Phenotypes		
Infected and non-infected	_	**154**
Infected	31.82%	49
Non-infected	68.18%	105
Parasite density and intensity of infection (only in infected group)	**29.87%**	**46**
NA	**Infected but with unavailable DP**	**3**
Light infection (1 to 9 eggs/10 mL)	0%	0
Moderate infection (10 to 49 eggs/10 mL)	16.23%	25
High infection (≥ 50 eggs/10 mL)	13.64%	21
Co-variables		
Age	4 to 8 years	154
Male	66.88%	103
Female	33.12%	51
Water-related activity	**Fishing, bathing, laundry, dishing, agriculture, cross-water on foot**	100
Yes	Who declared exercise one or more water-related activity above	80
No	Who declared contrary	17

4.14.2. Phenotype definition for P. falciparum infection

We included the following phenotypes: [*'Presence of P. falciparum infection'', ''Number of malaria infections'', ''Presence of symptomatic malaria infection'', ''Number of*

symptomatic malaria infections'', ''Presence of asymptomatic infection'', ''Number of asymptomatic malaria infections''], *''P. falciparum parasitemia (PD), ''Resistant/non-resistant, ''Co-infected'' and ''Sub-microscopic infection''* (Table 3).

Within the group of *''Infected''*, we separated the *''Symptomatic''* from *''Asymptomatic''*. A ‘*'Symptomatic''* was registered at emergency consult that was defined as any *P. falciparum* infected children with fever (axillary temperature ≥ 37.5°C) or history of fever within the last 24 hours. An *''Asymptomatic''* phenotype was registered at every MV and defined as any *P. falciparum* infection without fever. The thick and thin blood smears (TBS) were performed to confirm these infections. In each group *(Infected, Symptomatic or Asymptomatic),* we focused on both the prevalence and the number of occurrences of such phenotype. For example, the *''Number of malaria infections'', ''Number of symptomatic episodes''* and *''Number of asymptomatic infections''* were defined based on the number of occurrences of these phenotypes for each child during the follow-up.

The *''Number of symptomatic episodes''* and *''Number of asymptomatic infections''* were defined using the information recorded during emergency consultations and monthly visits, respectively. As mentioned above (at section 4.1), parents were encourage to take their child to the healthcare centre each time they noted any health problem of their children. Medical consultations at the heathcare centre were defined as an "emergency consultations" while systematically medical visits that occurred monthly were define as "monthly visits". Child was considered *''Symptomatic''* if the physician registered positive RDT and fever (axillary temperature ≥ 37.5°C), otherwise, *''Asymptomatic''* if the child was positive for RDT and TBS, but did not exhibit fever during monthly visits. Considering the date between each emergency consultation (≥ 10 days: period sufficient for each child develops and achieves the first malaria episode), we defined the *''Number of symptomatic episodes''* for each child as they attended an emergency consultation once, twice, thrice, and n times. For the *''Number of malaria infection'',* we considered the number of symptomatic or asymptomatic infection recorded for each child during the follow-up. Indeed, per child, we recorded one to eight ‘’*symptomatic episodes''*, one to five ‘’*Asymptomatic infection''* and one to ten *''malaria infection''* (Table 3). The ‘’*Asymptomatic phenotype''* was not given in the Table 3 but it was used to define the ‘’Resistant group’’ (detailed in next paragraph).

Other phenotypes we included in this analysis were *''P. falciparum parasitemia (PD), ''Resistant/non-resistant'', ''Co-infected'' and ''Sub-microscopic infection''*. The parasite density represents the mean level of *P. falciparum* density (PD) performed on log transformed PD (log PD + 1, to allow for 0 count) taking into account all the PD measurements performed

during the follow-up, including negative TBS. This average PD reflects the level of infection children suffered during the season of transmission taking into account the fact that children can be uninfected during the follow-up (Milet et al. 2010; Garcia et al. 2013). This phenotype is to date the more classical phenotype used in genetic epidemiology studies of malaria infection (Abel et al. 1992; Garcia et al. 1998; Rihet et al. 1998).

We defined the *''Resistant/non-resistant''* by considering the *ratio: symptomatic infection/asymptomatic infection* registered per child during the follow-up. If the *ratio <1*; that is, the child registered more asymptomatic infection than symptomatic, he was considered *''Resistant''*. Otherwise, if *ratio* >1, the child was considered to be *''non-resistant''*. *''Co-infected''* children were defined basing on their status positive for *S. heamatobium* eggs in urine (*S. h*$^{+}$) their status (symptomatic or asymptomatic) for malaria at inclusion. Finally, we used only the asymptomatic group to define the submicroscopic phenotype. Indeed, all child, at the end of the follow-up, who was negative for thick blood smear (TBS^{-}) and positive for *P. falciparum* quantitative real-time polymerase-chain reaction (qPCR^{+}) was considered *''sub-microscopically infected''*. This latter group was considered as *''more resistant''* (Table 3).

Table 3. Variables included in *P. falciparum* genetic association analysis

Genes/Phenotypes	Description	N
Genes		
Included in malaria analysis	Average number of MV ≥ 3 in unrelated children	**146**
HLA-E	Clinical and genetic available data	143
HLA-F	Clinical and genetic available data	145
HLA-G	Clinical and genetic available data	141
Phenotypes		
Infected and non-infected		**146**
Male	67.12%	98
Female	32.88%	48
Infected (RDT+, TBS+ with fever or no)	**91.78%**	**134**
Non-infected (RDT-, TBS- with fever or no)	**8.22%**	**12**
PD	**Average PD median =32**	115
PD<median	39.04%	57
PD>median	39.73%	58
Number of malaria infection per child during follow-up	**Ranging from 1 to 10**	134
1 or 2 infections	33.56%	49
3 or 4 infections	30.82%	45
> 4 infections	27.40%	40
Symptomatics (Fever and TBS+ or RDT+)	**87.67%**	128
Number of symptomatic malaria per child during follow-up	**Ranging from 1 to 8**	128
1 or 2 episodes	39.73%	58
3 or 4 episodes	30.14%	44
> 4 episodes	17.81%	28
Resistant and non-resistant	**yes, if ratio sympto/asympto <1; no, if >1**	96
Yes	4.11%	6
No	61.64%	90
Co_infection	***S. h*+ with *P.f* infection (all infections) at inclusion**	154
Yes	20.13%	31
No	77.92%	120
Not available (NA)	Known *S. h* status but unknown *P. f* status	3
Submicroscopic infection (in asymptomatic group)	**TBS- and qPCR+**	**69**
Yes	40.41%	59
No	6.85%	10

4.15. STATISTICAL ANALYSIS

HLA-G, -E and -F allele and haplotype frequencies, gene diversity, average nucleotide diversity and Tajima's D were calculated using ARLEQUIN v3.5.2 software (Excoffier et al. 2005). A Tajima's D *p-value* was computed by comparing the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data. It represents the probability of obtaining a simulated Tajima's D larger than the observed one. The Hardy–Weinberg equilibrium was

tested by the exact test of Guo and Thompson (Guo et al. 1992), using the ARLEQUIN v3.5.2 software (Excoffier et al. 2005). The linkage disequilibrium (LD) pattern among polymorphic sites was evaluated computing the correlation coefficient r2 (Devlin and Risch 1995). LD plots were visualized using Haploview 4.2 software (Barrett et al. 2005), considering only variable sites with a minimum allele frequency (MAF) above 1%.

The association study (SNPs *versus* phenotypes, and haplotypes *versus* phenotypes) was performed using PLINK v1.90b4.6 (Purcell et al. 2007) and R software v3.4.2 (R Core Team 2017), respectively. The associations between polymorphisms (all polymorphisms detected within the evaluated *HLA-G, -E* and *-F* segments and with MAF ≥5%) and between each specific phenotype were performed with PLINK using a mixed logistic model (for binary phenotypes) and linear model (for quantitative traits); using as covariables: sex and age for *P. falciparum* and sex, age and water-related activity, for *S. haematobium* (Table 2 and 3). The same approach was used to perform the association between the haplotypes (all haplotypes including the extended ones detected within the evaluated *HLA-G, -E* and *-F* segments and presenting a minimum frequency of 5%) and phenotypes using R packages.

However, the following traits (‘*’Number of malaria infection’’ and ‘’Number of symptomatic malaria episodes’’)* were firstly submitted to negative binomial regression (Poisson regression) adjusted for ‘’sex’’ and ‘’age’’ (using R packages) and the residuals (quantitative values) were then used as quantitative traits.

The final linear model fitness was checked using the **i)** residual fitness, **ii)** normality of standardized residuals, **iii)** scale-location of standardized residuals and **iv)** leverage of standardized residuals. The choice of the final logistic model was made on the basis of the lowest Akaike information criterion (AIC) (Akaike 1974).

5. RESULTS

RESULTS: PART I

PART I: POPULATION GENETIC STUDY

HLA-G, -E and *-F* regulatory and coding region variability and haplotypes in the Beninese Toffin population sample

5.1. OVERVIEW OF THE *HLA-G, -E* AND *-F* VARIABLE SITES

Based on recorded information (during the follow-up), a total of 154 unrelated children (exclusion of close relatives up to the second degree) were enrolled in the present study, exhibiting DNA samples successfully-sequenced for *HLA-G* (*n* = 149), *HLA-E* (*n* = 150) and *HLA-F* (*n* = 152) (Table 4).

Table 4. Variables included in the population genetics study.

Variables	Description	Study population	N
Related and unrelated	_	4 to 8 years, 6.7 mean age (±0.80)	288
Number of family	154 different families were counted based on survey information recorded	_	154
Related children	46.53%	_	134
Unrelated children	53.47% (exclusion of close relatives up to the second degree accepting just one individual per family)	_	154
HLA-E		_	**283**
Related	47%	_	133
Unrelated	53%	_	150
HLA-F	_	_	**284**
Related	46.48%	_	132
Unrelated	53.52%	_	152
HLA-G	_	_	**281**
Related	46.98%	_	132
Unrelated	53.02%	_	149

The evaluation of *HLA-G* (from position -1377 to +3275), *HLA-E* (from -2143 to +4420) and *HLA-F* (from -1709 to +3537) *loci*, revealed 96, 37 and 68 variable sites in the Beninese Toffin sample, respectively (Table 5). A complete list of variable sites, SNPid, and MAF is provided in Tables S1 to S3 for *HLA-G*, *HLA-E* and *HLA-F*, respectively (Appendix I to III).

Regarding the *HLA-G* gene, the classically-defined promoter region, encompassing positions -1 to -1377, exhibited 29 variable sites, among which 27 (93.1%) were polymorphic (MAF ≥ 1%). Fifty-eight variable sites were identified in the IPD-IMGT/HLA region (-300 to +2838), among which 56 occurred with a MAF ≥ 1%. In the 3'UTR segment (from +2960 to +3275), 16 variable sites were identified, among them, 12 (75%) were polymorphic (Table 5).

For the promoter region of *HLA-E,* encompassing positions -1 to -2143, we observed 19 variable sites, of which 15 (78.95%) were polymorphic, and 15 of these variable sites were described in the IPD-IMGT/HLA region (-300 to +3522) and 10 were polymorphic. In the 3' UTR segment, from +3468 to +4420, we identified seven variable sites, all polymorphic except one (Table 5).

For the *HLA-F* gene, the region upstream the first transcribed ATG encompassed three segments: **i)** the proximal (from -556 to -1), **ii)** the distal promoter (nucleotides upstream position -556), and **iii)** the segment encompassing positions -547 (hg 19 position 29690694) to -1013 (hg 19 position 29690228), which exhibited extended short tandem repeats (STR) with low sequence complexity. The genotype quality checking using the *vcfx checkpl* function exhibited high levels of missing alleles, and it was excluded from the analysis. According to this strategy, we detected 16 and 8 variable sites at the distal and proximal promoter, respectively (Table 5), and 20 out of the total of 24 (83.3%) variable sites presented a MAF >1%. All the 24 variable sites were previously described in the IPD-IMGT/HLA region (-300 to +3250), and 39 were polymorphic presenting MAFs >1%. We identified eight variable sites in the region downstream the stop codon, extending from positions +3061 to +3537, and six presented a MAF ≥ 1% and one variable site [+3061 (0.66%)] in the 3' UTR segment (Table 5).

Table 5. *HLA-G, -E and -F* variable sites, detected along all gene segments, and observed in the Beninese Toffin population.

Variable sites detected in evaluated segment highlighting the segment defined by IPD-IMGT/HLA v.3.31.0 database											
		5' upstream [a]									
Genes	In entire evaluated segment	Distal promoter [b]	Proximal promoter	Promoter polymorphic sites (MAF >1%)	Total exons	Total introns	**Region evaluated by IMGT [c]**	**Polymorphic sites in region tracked by IMGT [c]**	Variable sites at 3'UTR	Polymorphic sites at 3'UTR	Region downstream the codon stop [d]
HLA-G	96 [e] **(**-1377 to +3275)	–	29 (-1 to -1377)	27	14	40	**58**	**56**	16 (+2960 to +3275)	12	16
HLA-E	37 [e] **(**-2143 to +4420)	–	19 (-1 to -2143)	15	6	5	**15**	**10**	7 (+3468 to +4420)	6	7
HLA-F	68 [e] **(**-1709 to +3537)	16 (-1013 to -1709)	8 (-1 to -556)	20	10	26	**44**	**39**	1 (+2945 to +3063)	0	8

[a] All nucleotides upstream the +1 were considered to be the promoter region.

[b] As the distal promoter was not included in the *HLA-F* structure previously described (Lima et al., 2016), we considered all nucleotides upstream the nucleotide -556 as distal *HLA-F* promoter (Table S3, appendix III).

[c] IMGT coding region: *HLA-G* (-300 to 2838), *HLA-E* (-300 to +3522) and *HLA-F* (-300 to +3250), considering the Adenine of the first translated ATG as +1 (IPD-IMGT/HLA v.3.31.0 database).

[d] ''Region downstream the codon stop'' encompasses positions (+2798 to +3275 for *HLA-G*), (+3468 to +4420 for *HLA-E*) and (+3061 to +3537 for *HLA-F*) (Table S1, S2 and S3, appendix I to IV).

[e] Many of these variable sites were also included in *1000 Genome phase 3* project (Table S1, S2 and S3, appendix I to IV)**.**

MAF: minor allele frequency, **UTR:** untranslated region.

5.2. *HLA-G* HAPLOTYPES

5.2.1. HLA-G 5' upstream regulatory region haplotypes

We observed 29 promoter variable sites, arranged into 16 promoter haplotypes, which can be grouped into four lineages (PROMO-G010101, PROMO-G010102, PROMO-G0103 and PROMO-G0104), according to previous studies (Tan et al. 2005; Castelli et al. 2011, 2014a, 2017; Gineau et al. 2015; Nilsson et al. 2016) (Table 6). Accordingly, **i)** the PROMO-G010101 lineage clustered PROMO-G010101a to PROMO-G010101k, and the PROMO-G010101a was the most frequent (8.7%), **ii)** the PROMO-G010102 lineage grouped PROMO-G010102a and PROMO-G010102f, of which the PROMO-G010102a was the most frequent (33.2%), **iii)** the PROMO-G0103 lineage grouped PROMO-G0103a to PROMO-G0103h, and the most frequent was PROMO-G0103e (5.4%) and **iv)** the PROMO-G0104 lineage grouped the PROMO-G0104a to PROMO-G0104c, of which the PROMO-G0104a was the most frequent (25.8%). The PROMO-G010101k, PROMO-G010102f and PROMO-G0103h were classified as new haplotypes, and were observed as single copies in the Toffin population (Table 6).

Table 6. *HLA-G* promoter haplotypes observed in the Beninese Toffin population

	***HLA-G* Promoter Haplotypes [a]**																														
5' UTR Haplotypes[b]	**-1377**	**-1305**	**-1179**	**-1155**	**-1140**	**-1138**	**-1121**	**-964**	**-762**	**-725**	**-716**	**-689**	**-666**	**-646**	**-633**	**-546**	**-541**	**-539**	**-509**	**-486**	**-483**	**-477**	**-400**	**-391**	**-369**	**-256**	**-201**	**-90**	**-56**	**Frequency (2*n* = 298)**
PROMO-*G*010101a	T	G	A	G	A	A	C	G	C	C	T	A	G	A	G	A	GA	A	C	A	A	C	G	G	C	TC	G	G	C	0.0872
PROMO-*G*010101b	T	G	A	G	A	A	C	G	C	G	T	A	G	A	G	A	GA	A	C	A	A	C	G	G	C	TC	G	G	C	0.0537
PROMO-*G*010101d	T	G	A	G	A	A	C	G	C	C	T	A	G	A	G	A	GA	A	C	A	G	C	G	G	C	TC	G	G	C	0.0369
PROMO-*G*010101f	T	G	A	G	A	A	C	G	C	C	T	A	G	A	G	A	G	A	C	A	A	C	G	G	C	TC	G	G	C	0.0772
PROMO-*G*010101h	G	G	A	G	A	A	C	G	C	C	T	A	G	A	G	A	GA	A	C	A	A	C	G	G	C	TC	G	G	C	0.0302
PROMO-*G*010101k	**T**	**G**	**A**	**G**	**A**	**A**	**C**	**G**	**C**	**G**	**T**	**A**	**G**	**A**	**G**	**A**	**GA**	**A**	**C**	**A**	**A**	**C**	**G**	**G**	**C**	**TC**	**G**	**A**	**C**	**0.0034**
PROMO-*G*010102a	T	A	G	G	T	A	C	A	T	C	G	G	T	A	A	A	GA	A	C	C	A	G	G	G	A	TC	A	G	C	0.3322
PROMO-*G*010102f	**T**	**A**	**G**	**G**	**T**	**A**	**C**	**G**	**T**	**C**	**G**	**G**	**T**	**A**	**A**	**A**	**GA**	**A**	**C**	**C**	**A**	**G**	**G**	**G**	**A**	**TC**	**A**	**G**	**C**	**0.0034**
PROMO-*G*0103a	T	G	G	G	A	G	C	G	C	T	T	A	G	A	G	AG	GA	A	G	A	A	G	A	A	A	TC	G	G	T	0.0134
PROMO-*G*0103c	T	G	G	G	A	G	C	G	C	T	T	A	G	G	G	A	GA	A	G	A	A	G	G	G	A	TC	G	G	T	0.0302
PROMO-*G*0103e	T	G	G	G	A	G	C	G	C	T	T	A	G	G	G	A	GA	A	G	A	A	G	A	A	A	TC	G	G	T	0.0537
PROMO-*G*0103*g*	T	G	G	G	A	G	C	G	C	T	T	A	G	A	G	AG	GA	A	C	A	A	G	A	A	A	TC	G	G	T	0.0034
PROMO-*G*0103h	**T**	**G**	**G**	**G**	**A**	**G**	**C**	**A**	**C**	**T**	**T**	**A**	**G**	**G**	**G**	**A**	**GA**	**A**	**G**	**A**	**A**	**G**	**A**	**A**	**A**	**TC**	**G**	**G**	**T**	**0.0034**
PROMO-*G*0104a	T	A	G	A	A	A	C	A	T	C	G	G	T	A	A	A	GA	A	C	C	A	G	G	G	A	TC	A	G	C	0.2584
PROMO-*G*0104b	T	A	G	A	A	A	C	A	T	C	G	G	T	A	A	A	GA	G	C	C	A	G	G	G	A	TC	A	G	C	0.0034
PROMO-*G*0104c	T	A	G	A	A	A	C	A	T	C	G	G	T	A	A	A	GA	A	C	C	A	G	G	G	A	T	A	G	C	0.0101
																									Nucleotide diversity					0.0060 ± 0.0031
																									Gene diversity					0.8027 ± 0.0145
																									Tajima's *D*					2.5522, *P* = 0.0087

[a] Positions were inferred based on the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as +1. The human genome hg19 was used as draft: the first and second alternative alleles are marked in light grey and dark gray respectively.

[b] Haplotypes were named according to previous studies (Castelli et al., 2017; Castelli et al., 2011; Castelli et al., 2014b; Nilsson et al., 2016; Gineau et al., 2015, Tan et al., 2005). New haplotypes are marked in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.2.2. HLA-G IPD-IMGT/HLA region haplotypes

Nineteen coding region haplotypes (*HLA-G* alleles) were inferred in our population (Table 7), of which 13 have already been described at the IPD-IMGT/HLA v3.31.0 database, six have already been identified in Brazilians (Castelli et al. 2017) and four exhibited variable sites that have not been described in the IPD-IMGT/HLA database. Among the new alleles, the *G**01:01:02:01$^{(new1452C)}$ (1.01%) and *G**01:04:01$^{(new1523C)}$ (1.01%) corresponded to new variable sites, whereas the *G**01:01:01:01$^{(new192T)}$ (0.34%) and *G**01:01:01:05$^{(new-90A)}$ (0.34%) carried singletons. The most frequent haplotypes were *G**01:01:02:01 (19.8%), *G**01:04:04 (19.8%), *G**01:05N (11.4%) and *G**01:03:01:02 (10.4%) (Table 7).

Table 7. *HLA-G* haplotypes observed the Beninese Toffin population sample, considering the whole segment described by IPD-IMGT/HLA from -300 to +2838

Coding haplotypes [a]	**Notes**	**Frequency (2*n* = 298)** [b]
*G**01:01:02:01		0.1980
*G**01:04:04		0.1980
*G**01:05N		0.1141
*G**01:03:01:02		0.1040
*G**01:01:01:01		0.0772
*G**01:01:01:04		0.0772
*G**01:04:01		0.0537
*G**01:01:01:05		0.0369
***G**01:01:09 (nodel at +615)**		**0.0302**
*G**01:01:01:08		0.0235
*G**01:01:15-compatible		0.0168
***G**01:01:01:05 (+99*G*, +1147C, +2412A)**		**0.0134**
*G**01:01:02:02		0.0134
*G**01:01:02:01(new1452C)	A, new haplotype	0.0101
*G**01:04:01 (new1523C)	A, new haplotype	0.0101
*G**01:04:05-compatible		0.0101
*G**01:01:01:06		0.0067
*G**01:01:01:01 (new192T)	B, new haplotype	0.0034
*G**01:01:01:05 (new-90A)	B, new haplotype	0.0034
	Nucleotide diversity	0.0058 ± 0.0029
	Gene diversity	0.8821 ± 0.0081
	Tajima's *D*	2.8836, *P* = 0.0034

[a] Haplotypes were named according to the closest official haplotype at the IPD-IMGT/HLA v3.31.0 database highlighting the differences observed in some positions. The word **"new"** indicates the variable sites not recognized by IPD-IMGT/HLA v3.31.0 database. The term **"compatible"** was used when the allele did not have a complete sequence defined in IMGT, but its sequence is compatible with that which is defined in IMGT.

[b] Haplotypes were classified in descending order of their frequency.

Notes:

(A) indicates the variable sites described in 1000 genome phase 3 data (Table S1, appendix I).

(B) indicates a singleton, i.e, variable site detected in a single heterozygous individual from the population sample. Alleles shared with the Brazilian population (Castelli et al., 2017) are shown in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed. The complete sequences of coding *HLA-G* region are available upon request.

5.2.3. HLA-G 3' UTR haplotypes

We detected 13 variable sites at 3' UTR (from +2960 to +3275) that were arranged into 12 haplotypes, and named according to our previous reports (Castelli et al. 2010, 2011, 2014a, 2017; Lucena-Silva et al. 2012; Santos et al. 2013; Sabbagh et al. 2014; de Albuquerque et al. 2016; Nilsson et al. 2016) (Table 8). The *HLA-G*-3UTR-01 (11.74%), *HLA-G*-3UTR-02 (33.56%) and *HLA-G*-3UTR-03 (26.51%) were the most frequent, while *HLA-G*-3UTR-47 and -48, new haplotypes, were found at frequency < 1%.

Table 8. *HLA-G 3'* UTR haplotypes observed in the Beninese Toffin population.

***HLA-G* 3' UTR haplotypes[a]**														
Haplotypes[b]	**14bp (+2960)**	**+3001**	**+3003**	**+3010**	**+3032**	**+3035**	**+3038**	**+3142**	**+3187**	**+3196**	**+3204**	**+3227**	**+3275**	**Frequency (2*n* = 298)**
HLA-G-3UTR-01	G	C	T	G	G	C	C	C	G	C	G	G	C	0.1174
HLA-G-3UTR-02	GATTTGTTCATGCCT	C	T	C	G	C	C	G	A	G	G	G	C	0.3356
HLA-G-3UTR-03	G	C	T	C	G	C	C	G	A	C	G	G	C	0.2651
HLA-G-3UTR-04	G	C	C	G	G	C	C	C	A	C	G	G	C	0.0772
HLA-G-3UTR-05	GATTTGTTCATGCCT	C	T	C	G	T	C	G	A	C	G	G	C	0.0772
HLA-G-3UTR-06	G	C	T	G	G	C	C	C	A	C	G	G	C	0.0772
HLA-G-3UTR-17	GATTTGTTCATGCCT	T	T	C	G	T	C	G	A	C	G	G	C	0.0201
HLA-G-3UTR-18	G	C	T	G	G	C	C	C	A	C	G	A	C	0.0101
HLA-G-3UTR-20	G	C	T	G	C	C	C	C	A	C	G	G	C	0.0034
HLA-G-3UTR-46	G	C	T	C	G	C	T	G	A	C	G	G	C	0.0067
***HLA-G*-3UTR-47**	**G**	**C**	**T**	**G**	**G**	**C**	**C**	**C**	**A**	**C**	**C**	**A**	**C**	**0.0034**
***HLA-G*-3UTR-48**	**GATTTGTTCATGCCT**	**C**	**T**	**C**	**G**	**T**	**C**	**G**	**A**	**C**	**G**	**G**	**T**	**0.0067**
									Nucleotide diversity					0.0268 ± 0.0137
									Gene diversity					0.7875 ± 0.0134
									Tajima's *D*					0.0032, *P* = 0.4292

[a] The positions are those accepted by the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as +1. The Human genome hg19 was used as draft: the alternative alleles are marked in grey.

[b] Haplotypes were named according to previous studies (Nilsson et al., 2016; Sabbagh et al., 2014; Gineau et al., 2015; Castelli et al., 2010; Castelli et al., 2011; Lucena-Silva et al., 2012; de Albuquerque et al., 2016; Santos et al., 2013; Castelli et al., 2014b, Castelli et al., 2017). New haplotype is marked in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.2.4. HLA-G extended haplotypes

Considering positions -1377 to +3275, we observed 33 extended haplotypes, shown in Table 9. Twenty (60.61%) out of 33 occurred with a MAF ≥ 1%, and 25 have already been identified in the Brazilian population (Castelli et al. 2017). Due to the high level of LD detected across the whole gene **(Figure 9)**, strong associations were observed between haplotypes of different gene segments, for instance: **i)** all *G**01:01:01:01 alleles were in LD with PROMO-G010101a and *HLA-G*-3UTR-01, **ii)** all *G**01:01:01:04 alleles were associated with PROMO-G010101f and *HLA-G*-3UTR-06 group (*HLA-G*-3UTR-18, -20 and -47), **iii)** *G**01:01:02:01 was associated with PROMO-G010102a or similar (PROMO-G010102f) and *HLA-G*-3UTR-02, and **iv)** the *G**01:04:01 or *G**01:04:04 was associated with PROMO-G0104a and the *HLA-G*-3UTR-03 group (*HLA-G*-3UTR-46).

Table 9. *HLA-G* extended region (from -1377 to +3275) haplotypes observed in the Beninese Toffin population.

Promoter haplotypes [a]	**Coding haplotypes** [b]	**3' UTR haplotypes** [c]	**Notes**	**Frequency ($2n$ = 298)** [d]
PROMO-*G*010102a	*G**01:01:02:01	*HLA-G*-3UTR-02		0.1946
PROMO-*G*0104a	*G**01:04:04	*HLA-G*-3UTR-03		0.1846
PROMO-*G*010102a	*G**01:05N	*HLA-G*-3UTR-02		0.1141
PROMO-*G*010101a	*G**01:01:01:01	*HLA-G*-3UTR-01		0.0772
PROMO-*G*010101f	*G**01:01:01:04	*HLA-G*-3UTR-06		0.0604
PROMO-*G*0104a	*G**01:04:01	*HLA-G*-3UTR-03		0.0537
PROMO-*G*010101b	*G**01:01:01:05	*HLA-G*-3UTR-04		0.0336
PROMO-*G*010101h	*G**01:01:09 (nodel at +615)	*HLA-G*-3UTR-04		0.0302
PROMO-*G*0103e	*G**01:03:01:02	*HLA-G*-3UTR-05		0.0302
PROMO-*G*0103c	*G**01:03:01:02	*HLA-G*-3UTR-05		0.0302
PROMO-*G*010101d	*G**01:01:01:08	*HLA-G*-3UTR-01		0.0235
PROMO-*G*010101b	*G**01:01:15-compatible	*HLA-G*-3UTR-06		0.0168
PROMO-*G*0103e	*G**01:03:01:02	*HLA-G*-3UTR-17		0.0168
PROMO-*G*010101d	*G**01:01:01:05 (+99*G*, +1147C, +2412A)	*HLA-G*-3UTR-01		0.0134
PROMO-*G*010102a	*G**01:01:02:02	*HLA-G*-3UTR-02		0.0134
PROMO-*G*0103a	*G**01:03:01:02	*HLA-G*-3UTR-05		0.0134
PROMO-*G*010101f	*G**01:01:01:04	*HLA-G*-3UTR-18		0.0101
PROMO-*G*010102a	*G**01:01:02:01(new1452C)	*HLA-G*-3UTR-02	New haplotype	0.0101
PROMO-*G*0104a	*G**01:04:01 (new1523C)	*HLA-G*-3UTR-03	New haplotype	0.0101
PROMO-*G*0104c	*G**01:04:05-compatible	*HLA-G*-3UTR-03		0.0101
PROMO-*G*010101a	*G**01:01:01:06	*HLA-G*-3UTR-04		0.0067
PROMO-*G*0103e	*G**01:03:01:02	*HLA-G*-3UTR-48	New haplotype	0.0067
PROMO-*G*0104a	*G**01:04:04	*HLA-G*-3UTR-46		0.0067
PROMO-*G*010101a	*G**01:01:01:01 (new192T)	*HLA-G*-3UTR-01	New haplotype	0.0034
PROMO-*G*010101f	*G**01:01:01:04	*HLA-G*-3UTR-47	New haplotype	0.0034
PROMO-*G*010101f	*G**01:01:01:04	*HLA-G*-3UTR-20		0.0034
PROMO-*G*010101k	*G**01:01:01:05 (new-90A)	*HLA-G*-3UTR-04	New haplotype	0.0034
PROMO-*G*010101b	*G**01:01:01:05	*HLA-G*-3UTR-03		0.0034
PROMO-*G*010102f	*G**01:01:02:01	*HLA-G*-3UTR-02	New haplotype	0.0034
PROMO-*G*0103g	*G**01:03:01:02	*HLA-G*-3UTR-05		0.0034
PROMO-*G*0103h	*G**01:03:01:02	*HLA-G*-3UTR-17	New haplotype	0.0034
PROMO-*G*0104b	*G**01:04:04	*HLA-G*-3UTR-03		0.0034
PROMO-*G*0104a	*G**01:04:04	*HLA-G*-3UTR-04		0.0034
			Nucleotide diversity	0.0074 ± 0.0036
			Gene diversity	0.8994 ± 0.0086
			Tajima's *D*	2.5973, P = 0.0061

[a] The list of promoter haplotypes are depicted at Table 6. [b] The list of coding haplotypes are depicted at Table 7.

[c] The list of 3' UTR haplotypes are depicted at Table 8. [d] Haplotypes were classified in descending order of their frequency. A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

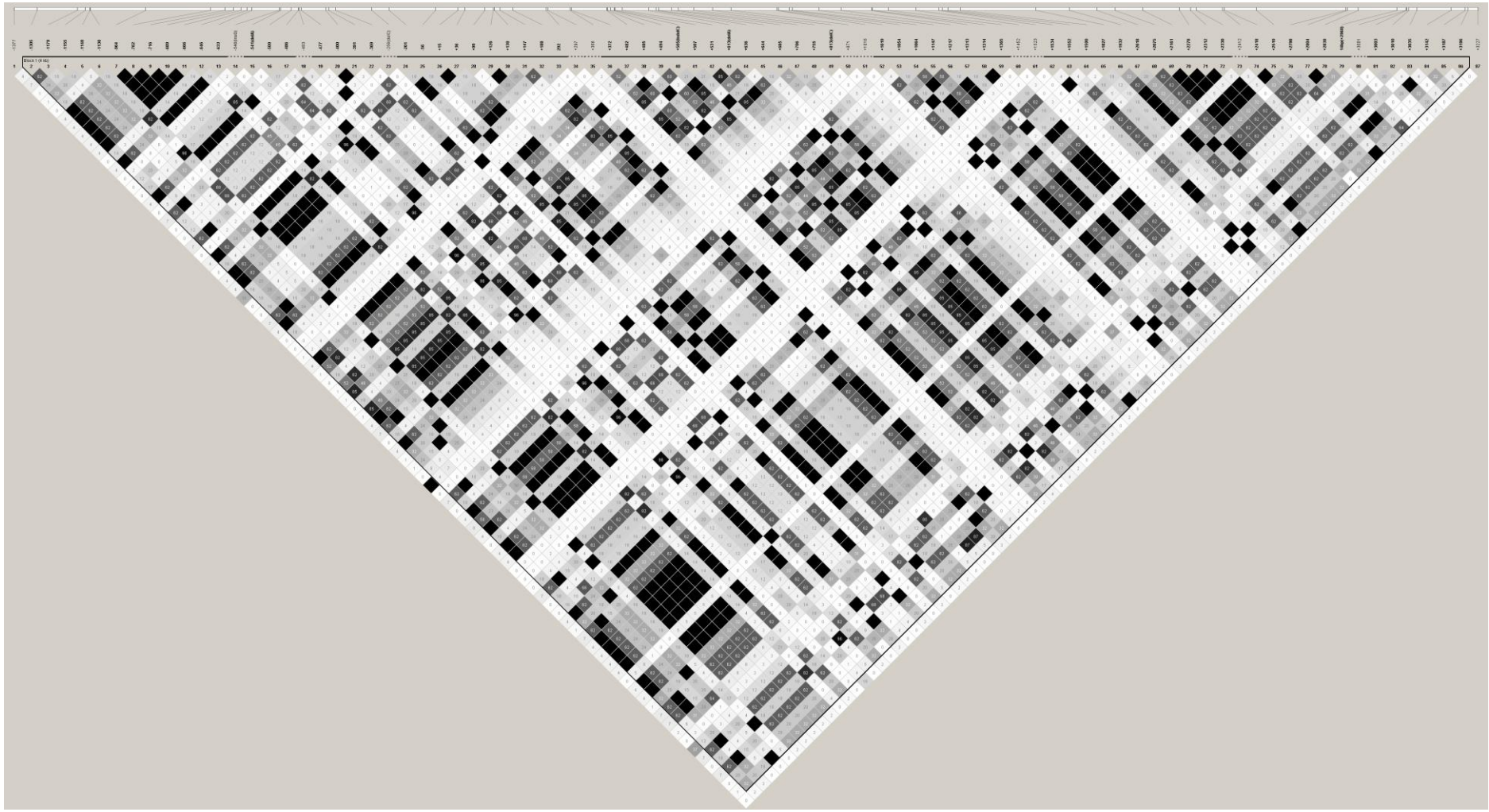


Figure 9. *LD patterns for the entire HLA-G region encompassing positions -1377 to +3275 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (x100) for the marker pairs are listed in the corresponding boxes. The LD plot was generated using SNPs with a minimum allele frequency (MAF) of 1%.*

5.3. *HLA-E* HAPLOTYPES

5.3.1. HLA-E 5' upstream regulatory region haplotypes

The region encompassing positions -1 to -2143 exhibited 19 variable sites, arranged into 19 promoter haplotypes, named as *E*-Promo-01 to *E*-Promo-39, according to our previous study (Ramalho et al. 2017). Sixteen out of 19 variable sites exhibited minimum frequencies of 1%, while others, including *E*-Promo-14 (0.67%), *E*-Promo-23 (0.67) and *E*-Promo-38 (0.33%), were less frequent, and the *E*-Promo-38 (0.33%) and *E*-Promo-39 (1%) represented new haplotypes (Table 10).

Table 10. *HLA-E 5′* upstream regulatory region haplotypes observed in the Beninese Toffin population

***HLA-E* Promoter Haplotypes [a]**																				
Haplotypes [b]	**-2143**	**-2142**	**-2123**	**-2106**	**-2069**	**-2015**	**-1988**	**-1981**	**-1796**	**-1423**	**-1389**	**-1262**	**-1167**	**-1159**	**-1158**	**-1079**	**-113**	**-104**	**-26**	**Frequency ($2n$ = 300)**
E-Promo-1	T	G	G	G	G	C	T	G	A	G	A	GT	A	A	T	G	T	A	G	0.3767
E-Promo-7	C	G	G	G	G	G	T	G	A	G	A	GT	A	G	T	G	T	A	G	0.1300
E-Promo-2	T	G	G	G	G	C	T	G	A	G	G	GT	A	A	T	G	T	A	G	0.1200
E-Promo-20	C	A	G	G	G	G	C	G	A	G	A	GT	A	G	T	G	T	G	G	0.0700
E-Promo-9	T	G	A	G	G	C	T	G	A	G	A	GT	A	A	T	G	T	A	T	0.0500
E-Promo-15	T	G	G	G	G	C	T	G	A	G	G	GT	A	G	T	G	C	A	G	0.0567
E-Promo-6	T	G	G	G	C	C	T	G	A	G	A	GT	A	A	T	G	T	A	G	0.0400
E-Promo-8	C	G	G	G	G	G	C	G	A	G	A	GT	A	G	T	G	T	G	G	0.0300
E-Promo-5	T	G	G	A	G	C	T	G	A	G	A	GT	A	A	T	G	T	A	G	0.0233
E-Promo-10	T	G	G	G	G	G	C	G	A	G	A	GT	G	A	T	G	T	A	G	0.0167
E-Promo-18	T	G	G	G	G	C	T	G	A	A	A	GT	A	A	T	G	T	G	G	0.0167
E-Promo-11	T	G	G	G	G	G	C	G	A	G	A	GT	A	A	T	T	T	A	G	0.0133
E-Promo-17	T	G	G	G	G	C	T	G	G	G	G	GT	A	A	T	G	T	A	G	0.0100
E-Promo-21	T	G	G	G	G	C	T	A	A	G	G	GT	A	A	T	G	T	A	G	0.0100
E-Promo-31	C	G	G	G	G	G	T	G	A	G	A	GT	A	A	T	G	T	A	G	0.0100
***E*-Promo-39**	**T**	**G**	**G**	**G**	**G**	**C**	**T**	**G**	**A**	**G**	**A**	**GT**	**A**	**G**	**C**	**G**	**T**	**A**	**G**	**0.0100**
E-Promo-14	T	G	G	A	G	C	T	G	A	G	A	G	A	A	T	G	T	A	G	0.0067
E-Promo-23	T	G	G	G	G	C	T	G	A	G	A	GT	A	G	T	G	T	A	G	0.0067
***E*-Promo-38**	**C**	**G**	**G**	**G**	**G**	**G**	**T**	**G**	**A**	**G**	**A**	**GT**	**A**	**G**	**T**	**G**	**T**	**G**	**G**	**0.0033**
																			Nucleotide diversity	0.0013 ± 0.0008
																			Gene diversity	0.8146 ± 0.0175
																			Tajima's *D*	-0.1764, P = 0.5000

[a] Positions were inferred based on the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as +1. The Human genome hg19 was used as draft: the alternative alleles are marked in gray. Haplotypes were listed in descending order of their frequency.

[b] Haplotypes were named according to a previous study (Ramalho et al., 2017). New haplotypes are marked in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.3.2. HLA-E IPD-IMGT/HLA region haplotypes

Regarding the coding region, 15 haplotypes (or *HLA-E* alleles) were inferred (Table 11), of which: **i)** 6 have been recognized by IPD-IMGT/HLA v.3.31.0 database, **ii)** the *E**01:03:02:01$^{(new2269C)}$ and *E**01:03:01:01$^{(new1322A)}$ have already been observed in African and Brazilian samples (Castelli et al. 2015; Ramalho et al. 2017), **iii)** the *E**01:01:01:01$^{(new3468C)}$ was also observed in the Brazilian population (Ramalho et al. 2017), and **iv)** the six remaining haplotypes were considered new alleles (summed frequency of 4.66%), exhibiting one or more variable sites that have not been described in the IPD-IMGT/HLA v.3.31.0 database. Of these, the *E**01:01:01:01$^{(new758C)}$ and *E**01:01:01:01$^{(new3042C)}$ exhibited variable sites at exons 3 and 7, respectively (Table 11 and Table S2, appendix II). These variations (+758 C and +3042 C) lead to a G>C synonymous substitution at codon 107 (*E**01:01:01:01$^{(codon107:AGG->AGC)}$) and to an A>C substitution at codon 336 (*E**01:01:01:01$^{(codon336:AGC->CGC)}$), respectively (Table 11). The first three most frequent haplotypes identified here were *E**01:01:01:01 (40.0%), *E**01:03:02:01 (24.7%) and *E**01:03:05$^{-compatible}$ (11.7%).

Table 11. *HLA-E* haplotypes observed in the Beninese Toffin population, considering the segment reported by IPD-IMGT/HLA from -300 to +3522.

				***HLA-E* coding Haplotypes** [a]																
***HLA*-E coding Haplotypes** [b]	**Notes**	**Coding haplotypes (without introns)**	**Encoded proteins** [c]	**-113**	**-104**	**-26**	**+424**	**+458**	**+756**	**+758**	**+1322**	**+1625**	**+1994**	**+2269**	**+2626**	**+2944**	**+3042**	**+3468**	**Frequency ($2n$ = 300)** [d]
*E**01:01:01:01		*E**01:01:01	E*01:01	T	A	G	C	G	A	G	G	G	C	T	C	C	A	A	0.4000
*E**01:03:02:01		*E**01:03:02	E*01:03	T	A	G	T	G	G	G	G	G	C	T	C	C	A	A	0.2467
*E**01:03:05-compatible		*E**01:03:05	E*01:03	T	G	G	C	G	G	G	G	C	C	T	C	C	A	A	0.1167
*E**01:05-compatible		*E**01:05	E*01:05	C	A	G	C	A	G	G	G	G	C	T	C	C	A	A	0.0567
*E**01:01:01:06-compatible		*E**01:01:01	E*01:01	T	A	T	C	G	A	G	G	G	C	T	C	C	A	C	0.0500
***E**01:01:01:01(new3468C)**	**A,B**	***E**01:01:01**	**E*01:01**	**T**	**A**	**G**	**C**	**G**	**A**	**G**	**G**	**G**	**C**	**T**	**C**	**C**	**A**	**C**	**0.0400**
*E**01:01:01:01(new2626T)	A, new haplotype	*E**01:01:01	E*01:01	T	A	G	C	G	A	G	G	G	C	T	T	C	A	A	0.0233
*E**01:03:01:01		*E**01:03:01	E*01:03	T	A	G	C	G	G	G	G	G	C	T	C	C	A	A	0.0233
***E**01:03:02:01(new2269C)**	**A,B**	***E**01:03:02**	**E*01:03**	**T**	**A**	**G**	**T**	**G**	**G**	**G**	**G**	**G**	**C**	**C**	**C**	**C**	**A**	**A**	**0.0133**
*E**01:01:01:01(new3042C)	A,D, new haplotype	*E**01:01:01:01 (codon336:AGC->CGC)	E*01:01	T	A	G	C	G	A	G	G	G	C	T	C	C	C	A	0.0067
*E**01:01:01:01(new1994T, new3468C)	A,B, new haplotype	*E**01:01:01	E*01:01	T	A	G	C	G	A	G	G	G	T	T	C	C	A	C	0.0067
***E**01:03:01:01(new1322A)**	**A,B**	***E**01:03:01**	**E*01:03**	**T**	**A**	**G**	**C**	**G**	**G**	**G**	**A**	**G**	**C**	**T**	**C**	**C**	**A**	**A**	**0.0067**
*E**01:01:01:01(new758C)	C, E, new haplotype	E*01:01:01:01 (codon107:AGG->AGC)	E*01:01	T	A	G	C	G	A	C	G	G	C	T	C	C	A	A	0.0033
*E**01:01:01:01(new2944T)	A,C new haplotype	*E**01:01:01	E*01:01	T	A	G	C	G	A	G	G	G	C	T	C	T	A	A	0.0033
*E**01:03:01:01(new-104G, +1625C, new3468C)	A,B, new haplotype	*E**01:03:01	E*01:03	T	G	G	C	G	G	G	G	C	C	T	C	C	A	C	0.0033
																		Nucleotide diversity		0.0005 ± 0.0003
																		Gene diversity		0.7593 ± 0.0172
																		Tajima's *D*		-0.4737, *P* = 0.6294

[a] Positions were inferred based on the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as +1. The Human genome hg19 was used as draft: the alternative alleles are marked in gray.

[b] Haplotypes were named according to the closest official haplotypes in IPD-IMGT/HLA v3.31.0 database highlighting the differences observed in some positions. The word **''new''** indicates the variable sites not recognized by IPD-IMGT/HLA v3.31.0 database. The term **''compatible''** was used when the allele did not have a complete sequence defined in IMGT, but its sequence is compatible with that which is defined in IMGT.

[c] Represent the full-length protein encoded by each allele (haplotype). Three different full-length proteins (E*01:01, F*01:03 and F*01:05) were detected in our population.

[d] Haplotypes were classified in descending order of their frequency.

Notes:

(A) indicates the variable sites described in 1000 genome phase 3 database except +1994T (Table S2, appendix II).

(B) indicates the variable sites described in other populations (Castelli et al., 2015; Ramalho et al., 2017) except +1994T.

(C) indicates a singleton, i.e., variable site detected in a single heterozygous individual from the population sample.

(D) Synonymous mutation on exon 7 (Table S2, appendix II) leading to a substitution of A nucleotide into C nucleotide in codon 12.

(E) Synonymous mutation on exon 3 (Table S2, appendix II) leading to a substitution of G nucleotide into C nucleotide in codon 17. Alleles shared with the Brazil population sample (Ramalho et al., 2017) and other African population sample (Castelli et., 2015) are shown in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.3.3. HLA-E 3' UTR haplotypes

Seven haplotypes spanning from +3468 to +4420 were identified, which were named as *HLA-E*-3UTR-01 to *HLA-E*-3UTR-14, according to our previous study (Castelli et al. 2015; Ramalho et al. 2017), and the most frequent were *HLA-E*-3UTR-01 (58.7%), *HLA-E*-3UTR-02 (12.0%), *HLA-E*-3UTR-03 (10.0%), and *HLA-E*-3UTR-04 (13.0%) (Table 12).

Table 12. *HLA-E* 3' UTR haplotypes observed in the Beninese Toffin population

***HLA-E* 3' UTR haplotypes [a]**								
Haplotypes [b]	**+3468**	**+3634**	**+3777**	**+3778**	**+3824**	**+4297**	**+4420**	**Frequency (2*n* = 300)**
HLA-E-3UTR-1	A	G	A	A	T	G	C	0.5867
HLA-E-3UTR-2	A	G	G	A	T	G	C	0.1200
HLA-E-3UTR-3	C	G	A	A	T	G	C	0.1000
HLA-E-3UTR-4	A	A	A	A	T	A	C	0.1300
HLA-E-3UTR-5	A	G	A	G	T	G	C	0.0367
HLA-E-3UTR-13	A	G	A	A	T	G	T	0.0233
HLA-E-3UTR-14	A	G	A	A	TG	G	C	0.0033
						Nucleotide diversity		0.0010 ± 0.0008
						Gene diversity		0.6147 ± 0.0276
						Tajima's *D*		0.0152, *P* = 0.4317

[a] Positions were inferred based on the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as +1. The Human genome hg19 was used as draft: the alternative alleles are marked in grey.

[b] Haplotypes were named according to previous studies (Castelli et al., 2015; Ramalho et al., 2017).

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.3.4. HLA-E extended haplotypes

Thirty-one *HLA-E* extended haplotypes were detected in the region encompassing positions -143 to +4420 (Table 13). Sixteen of them occurred at a frequency above 1% and 23 have already been described in the Brazilian population (Ramalho et al. 2017). The remaining 8 new haplotypes, together, presented a summed frequency of 5.66%. Two patterns of haplotype combinations were observed: **i)** the *E*-Promo-01 and similar haplotypes (*E*-Promo-05, -06, -14, -31 and -39) were associated with the *E**01:01:01 coding allele (or derivatives) and to the *HLA-E*-3UTR-01 haplotype, and **ii)** the *E*-Promo-02 and similar haplotypes (*E*-Promo-17) were associated to the *E**01:03 coding allele (or derivatives) and to *HLA-E*-3UTR-

02. Although these associations were common, others were also observed due to the weak LD among the different *HLA-E* segments **(Figure 10),** findings that were also observed in our previous studies (Felício et al. 2014; Castelli et al. 2017; Ramalho et al. 2017).

Table 13. *HLA-E* extended region (from -2143 to +4420) haplotypes observed in the Beninese Toffin population.

5' Upstream haplotypes[a]	Coding haplotypes [b]	3' UTR haplotypes [c]	Notes	Frequency ($2n = 300$) [d]
E-Promo-1	*E**01:01:01:01	*HLA-E*-3UTR-1		0.3233
E-Promo-7	*E**01:03:02:01	*HLA-E*-3UTR-4		0.1300
E-Promo-2	*E**01:03:02:01	*HLA-E*-3UTR-2		0.0967
E-Promo-20	*E**01:03:05-compatible	*HLA-E*-3UTR-1		0.0700
E-Promo-15	*E**01:05-compatible	*HLA-E*-3UTR-1		0.0567
E-Promo-9	*E**01:01:01:06-compatible	*HLA-E*-3UTR-3		0.0500
E-Promo-6	*E**01:01:01:01	*HLA-E*-3UTR-1		0.0400
E-Promo-1	*E**01:01:01:01(new3468C)	*HLA-E*-3UTR-3		0.0333
E-Promo-5	*E**01:01:01:01(new2626T)	*HLA-E*-3UTR-1	New haplotype	0.0233
E-Promo-8	*E**01:03:05-compatible	*HLA-E*-3UTR-13		0.0233
E-Promo-10	*E**01:01:01:01	*HLA-E*-3UTR-5		0.0167
E-Promo-18	*E**01:03:05-compatible	*HLA-E*-3UTR-1		0.0167
E-Promo-11	*E**01:03:01:01	*HLA-E*-3UTR-5		0.0133
E-Promo-2	*E**01:03:02:01(new2269C)	*HLA-E*-3UTR-2		0.0133
E-Promo-39	*E**01:01:01:01	*HLA-E*-3UTR-1		0.0100
E-Promo-17	*E**01:03:02:01	*HLA-E*-3UTR-2		0.0100
E-Promo-14	*E**01:01:01:01	*HLA-E*-3UTR-1		0.0067
E-Promo-1	*E**01:01:01:01(new3042C)	*HLA-E*-3UTR-1	New haplotype	0.0067
E-Promo-1	*E**01:01:01:01(new1994T, new3468C)	*HLA-E*-3UTR-3	New haplotype	0.0067
E-Promo-31	*E**01:01:01:01(new3468C)	*HLA-E*-3UTR-3	New haplotype	0.0067
E-Promo-21	*E**01:03:01:01	*HLA-E*-3UTR-1		0.0067
E-Promo-23	*E**01:03:01:01(new1322A)	*HLA-E*-3UTR-1		0.0067
E-Promo-2	*E**01:03:02:01	*HLA-E*-3UTR-1		0.0067
E-Promo-31	*E**01:01:01:01	*HLA-E*-3UTR-1		0.0033
E-Promo-1	*E**01:01:01:01(new758C)	*HLA-E*-3UTR-1	New haplotype	0.0033
E-Promo-1	*E**01:01:01:01(new2944T)	*HLA-E*-3UTR-1	New haplotype	0.0033
E-Promo-21	*E**01:03:01:01	*HLA-E*-3UTR-5		0.0033
E-Promo-8	*E**01:03:01:01(new-104G, +1625C, new3468C)	*HLA-E*-3UTR-3	New haplotype	0.0033
E-Promo-2	*E**01:03:02:01	*HLA-E*-3UTR-5		0.0033
E-Promo-38	*E**01:03:05-compatible	*HLA-E*-3UTR-14	New haplotype	0.0033
E-Promo-8	*E**01:03:05-compatible	*HLA-E*-3UTR-1		0.0033
			Nucleotide diversity	0.0008 ± 0.0004
			Gene diversity	0.8561 ± 0.0148
			Tajima's *D*	-0.3062, P = 0.5518

[a] The list of promoter haplotypes are depicted at Table 10.

[b] The list of coding haplotypes are depicted at Table 11.

[c] The list of 3' UTR haplotypes are depicted at Table 12.

[d] Haplotypes were classified in descending order of their frequency.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

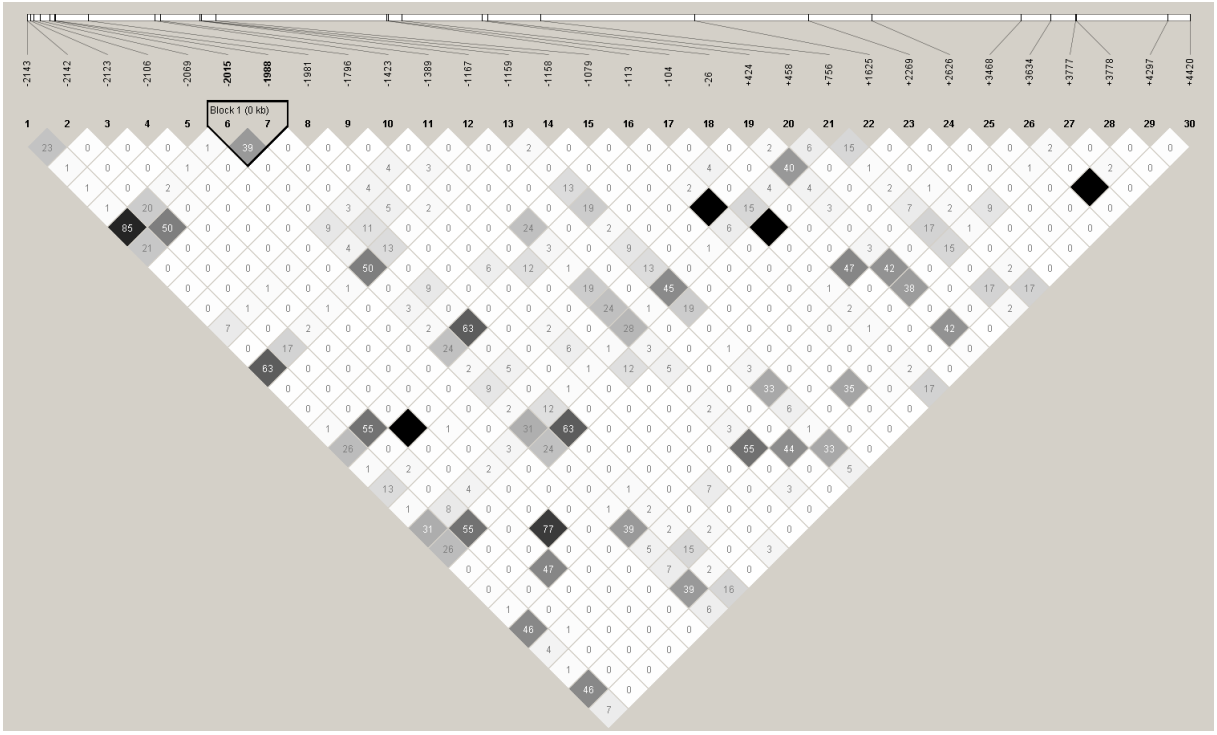


Figure 10. *LD patterns for the entire HLA-E region encompassing the positions -2143 to +4420 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (x100) for the marker pairs are listed in the corresponding boxes. The LD plot was generated using SNPs with a minimum allele frequency (MAF) of 1%.*

5.4. *HLA-F* HAPLOTYPES

5.4.1. HLA-F 5' upstream regulatory region haplotypes

The region upstream the first transcribed ATG was evaluated in two segments, a proximal promoter (from -556 to -1) and a distal promoter (nucleotides upstream position -556) (detailed at section 5.1). The distal promoter variable sites were arranged into eight different haplotypes (Table 14). Considering that this region has never been described nor considered in IPD-IMGT/HLA, we named its haplotypes as *F**Distal-A to *F**Distal-J (Table

14), following the same pattern reported by Lima et al. to name the proximal haplotypes (Lima et al. 2016). The *F**Distal-A to *F**Distal-F haplotypes were observed with frequencies greater than 1%, with *F**Distal-A being the most frequent (45.1%) one. In the same Table 14, we identified and presented seven F-proximal promoter haplotypes, which were named as *F**upstream-A to *F**upstream-H according to the same study reported by Lima et al. (Lima et al. 2016). The most frequent proximal haplotypes were *F**upstream-A (42.4%), *F**upstream-C (15.43%), *F**Upstream-D (16.1%), and *F**upstream-B (15.1%), while the *F**upstream-F was the least frequent one (0.3%).

Table 14. *HLA-F* 5' distal (A) and proximal (B) promoter haplotypes observed in the Beninese Toffin population.

(A) *HLA-F* Distal Promoter Haplotypes

Haplotypes	-1709	-1651	-1610	-1558	-1524	-1499	-1475	-1416	-1268	-1239	-1215	-1207	-1185	-1116	-1114	-1013	Frequency ($2n$ = 304)
*F**Distal-A	T	T	T	A	G	T	G	A	C	G	T	C	C	G	T	C	0.4507
*F**Distal-B	G	G	T	AG	T	T	G	A	T	T	C	C	C	A	T	G	0.2303
*F**Distal-C	T	T	T	A	G	T	G	A	C	G	T	C	T	G	T	C	0.1118
*F**Distal-D	T	T	C	A	G	T	G	A	C	G	T	C	C	G	T	C	0.1645
*F**Distal-E	T	T	T	A	G	T	A	A	C	G	T	C	C	G	T	C	0.0197
*F**Distal-F	T	T	T	A	G	G	G	A	C	G	T	C	C	G	T	C	0.0132
*F**Distal-I	T	T	C	A	G	T	G	A	C	G	T	T	C	G	C	C	0.0033
*F**Distal-J	G	G	T	AG	T	T	G	G	T	T	C	C	C	A	T	G	0.0066
																Nucleotide diversity	0.0055 ± 0.0031
																Gene diversity	0.7060 ± 0.0170
																Tajima's *D*	1.1128, *P* = 0.1136

(B) *HLA-F* Proximal Promoter Haplotypes

Haplotypes	-547	-526	-500	-403	-222	-151	-144	-101	Frequency ($2n$ = 304)
*F**Upstream-A	C	T	G	G	G	T	A	C	0.4243
*F**Upstream-B	G	T	G	G	A	A	A	C	0.1513
*F**Upstream-C	C	T	G	G	G	T	T	G	0.1743
*F**Upstream-D	C	T	G	G	G	A	A	C	0.1612
*F**Upstream-F	C	T	A	A	G	T	T	G	0.0033
*F**Upstream-G	C	A	G	G	G	A	A	C	0.0263
*F**Upstream-H	C	T	A	G	G	T	T	G	0.0592
							Nucleotide diversity		0.0042 ± 0.0027
							Gene diversity		0.7389 ± 0.0164
							Tajima's *D*		0.9808, *P* = 0.1481

(A, B) Positions were inferred based on the IPD-IMGT/HLA database, considering the Adenine of the first translated ATG as 1. The Human genome hg19 was used as draft: the alternative alleles are marked in gray.

(A) As the distal promoter was not included in this *HLA-F* structure previously described (Lima et al., 2016), we considered all nucleotides upstream the nucleotide -556 as distal *HLA-F* promoter (Table S3, appendix III). Haplotypes were listed in alphabetic order.

(B) Haplotypes were named according to a previous study (Lima et al., 2016).

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

5.4.2. HLA-F IPD-IMGT/HLA region haplotypes

We inferred 29 haplotypes in this IPD-IMGT/HLA coding region (Table 15), of which, five (*F**01:01:01:09, *F**01:01:01:08, *F**01:01:01:11, *F**01:03:01:01 and *F**01:01:01:01) were officially recognized by the IPD-IMGT/HLA v.3.31.0 database. Other 19 haplotypes exhibited variable sites that have already been described in IPD-IMGT/HLA v.3.31.0 database or shared with the Brazilian population sample (Lima et al. 2016), and the five remaining haplotypes can be considered as new ones (Table 15), since they carried variants that were not described in IPD-IMGT/HLA v.3.31.0 database. The coding haplotypes encode three different HLA-F proteins: F*01:01 (77.6%), F*01:02 (3.0%) and F*01:03 (19.4%) (Table 15). Noteworthy, the F*01:02 protein (produced by *F**01:02$^{(1193C,\ 1943G)}$ and *F**01:02$^{(1193C,\ 1943G,\ new1946C)}$) was absent in a previously studied Brazilian population sample (Lima et al. 2016).

Table 15. *HLA-F* haplotypes observed in the Beninese Toffin population sample, considering the whole segment reported by IPD-IMGT/HLA from -300 to +3250.

***HLA-F* coding haplotypes (with introns) [a]**	**Notes**	***HLA-F* coding haplotypes (without introns)**	**Encoded proteins [b]**	**Frequency ($2n$ = 304) [c]**
***F**01:01:01:05(1943G, [TG]$_{12}$)**		***F**01:01:01**	**F*01:01**	**0.1513**
***F**01:03:01:01(1383G)**		***F**01:03:01**	**F*01:03**	**0.1513**
*F**01:01:01:09		*F**01:01:01	F*01:01	0.1250
*F**01:01:02:05(1943G, [TG]$_{11}$)		*F**01:01:02	F*01:01	0.1020
***F**01:01:01:01([TG]$_{12}$)**		***F**01:01:01**	**F*01:01**	**0.0954**
***F**01:01:02:03(1943G, 2208C, [TG]$_{11}$)**		***F**01:01:02**	**F*01:01**	**0.0691**
***F**01:01:02:03(1943G, [TG]$_{13}$)**		***F**01:01:02**	**F*01:01**	**0.0526**
*F**01:01:01:08		*F**01:01:01	F*01:01	0.0395
*F**01:01:01:11		*F**01:01:01	F*01:01	0.0329
*F**01:01:01:08(-222G, 2698G, [TG]$_{12}$)		*F**01:01:01	F*01:01	0.0263
*F**01:01:01:09(new205A)	A, B, D, new haplotype	*F**01:01:01(codon4:TCC->TCA)	F*01:01	0.0230
*F**01:02(1193C, 1943G)		*F**01:02	F*01:02	0.0164
*F**01:03:01:01		*F**01:03:01	F*01:03	0.0164
*F**01:02(1193C, 1943G, new1946C)	A, new haplotype	*F**01:02	F*01:02	0.0132
***F**01:03:01:01(new1378T, 1383G, [TG]$_{11}$)**	**A, B**	***F**01:03:01**	**F*01:03**	**0.0132**
*F**01:01:01:01		*F**01:01:01	F*01:01	0.0099
*F**01:01:01:05(new1497C, 1943G, [TG]$_{12}$)	A, B, new haplotype	*F**01:01:01	F*01:01	0.0099
*F**01:01:01:09(new2123T)	A, new haplotype	*F**01:01:01	F*01:01	0.0099
*F**01:01:01:09(3189T)		*F**01:01:01	F*01:01	0.0066
*F**01:01:02:03(1943G, [TG]$_{14}$)		*F**01:01:02	F*01:01	0.0066
*F**01:01:01:01(new145A)	A, C, new haplotype	*F**01:01:01	F*01:01	0.0033
*F**01:01:01:01([TG]$_{11}$)		*F**01:01:01	F*01:01	0.0033
*F**01:01:01:09([TG]$_{12}$)		*F**01:01:01	F*01:01	0.0033
*F**01:01:02:03(1943G, [TG]$_{12}$)		*F**01:01:02	F*01:01	0.0033
*F**01:01:02:06(1225nolargedel, 1230A)		*F**01:01:02	F*01:01	0.0033
*F**01:03:01:01(1383G, [TG]$_{11}$)		*F**01:03:01	F*01:03	0.0033
*F**01:03:01:01([TG]$_{10}$)		*F**01:03:01	F*01:03	0.0033
***F**01:03:01:01([TG]$_{11}$)**		***F**01:03:01**	**F*01:03**	**0.0033**
***F**01:03:01:01([TG]$_{13}$)**		***F**01:03:01**	**F*01:03**	**0.0033**
			Nucleotide diversity	0.0054 ± 0.0027
			Gene diversity	0.9093 ± 0.0061
			Tajima's *D*	1.7086, P = 0.0416

[a] Haplotypes were named according to the closest official haplotypes in IPD-IMGT/HLA v3.31.0 database highlighting the differences observed in some positions. [TG]$_{10}$ to [TG]$_{14}$ are microsatellites (TG repeats). This *HLA-F* short tandem repeats (STR) occurs at position 3097 (IPD-IMGT/HLA) (Table S3, appendix III). The total number of dinucleotide TG repeats is indicated when different from the known *HLA-F* allele. The word **"new"** indicates the variable sites not recognized by IPD-IMGT/HLA v3.31.0 database.

[b] Represent the full-length protein encoded by each allele (haplotype). Three different proteins or *HLA-F* transcripts (F*01:01, F*01:02 and F*01:03) were detected in our population.

[c] Haplotypes were classified in descending order of their frequency.

Notes:

(A) indicates the variable sites described in 1000 genome phase 3 data base (Table S3, appendix III).

(B) indicates the variable sites described in other populations (Lima et al., 2016).

(C) indicates a singleton, i.e., variable site detected in a single heterozygous individual from the population sample.

(D) Synonymous mutation on exon 2 (Table S3, appendix III) leading to a substitution of C nucleotide into A nucleotide in codon 4.

Alleles shared with the Brazilian population are shown in bold.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

The complete sequences of coding *HLA-F* region are available upon request.

5.4.3. HLA-F 3' UTR haplotypes

The *HLA-F* 3'UTR segment encompasses the positions +2945 to +3063 when considering the transcript variant NM_018950.2 (Lima et al. 2016). In the present study, we detected only one 3' UTR variable site at position +3061 (Table S3), which led to two 3' UTR haplotypes named as *F**3UTR-A (+3061A, 99.34%) and *F**3UTR-B (+3061T, 0.66%) (Table 16).

5.4.4. HLA-F extended haplotypes

Considering the entire gene from positions -1709 to +3250, 32 different extended haplotypes were inferred (Table 16). Fifteen of them presented polymorphic frequencies ranging from 1.3% to 15.1%, whereas 17 had frequencies below 1% (Table 16). Overall, we observed the following associations: **i)** *F**01:01:01:01 or derivatives were associated with *F**Distal-C, *F**upstream-B and *F**3UTR-A, **ii)** *F**01:01:01:09 or *F**01:02 or derivatives were associated to *F**Distal-A, *F**upstream-A and *F**3UTR-A, and **iii)** *F**01:03:01:01 or derivatives were associated with *F**Distal-A, *F**upstream-A or to *F**Distal-D or *F**upstream-A and *F**3UTR-A. The LD pattern between *HLA-F* coding and regulatory regions shown in **Figure 11**, already described elsewhere (Lima et al. 2016), may explain these associations.

Table 16. *HLA-F* extended region (from -1709 to +3250) haplotypes observed in the Beninese Toffin population.

5' Distal Promoter [a]	5' Proximal Promoter [a]	Coding haplotypes (with introns) [b]	3' UTR haplotypes [c]	Frequency ($2n$ = 304) [d]
*F**Distal-A	*F**Upstream-D	*F**01:01:01:05(1943G, $[TG]_{12}$)	*F**3UTR-A	0.1513
*F**Distal-D	*F**upstream-A	*F**01:03:01:01(1383G)	*F**3UTR-A	0.1480
*F**Distal-A	*F**upstream-A	*F**01:01:01:09	*F**3UTR-A	0.1250
*F**Distal-B	*F**upstream-C	*F**01:01:02:05(1943G, $[TG]_{11}$)	*F**3UTR-A	0.1020
*F**Distal-C	*F**upstream-B	*F**01:01:01:01($[TG]_{12}$)	*F**3UTR-A	0.0954
*F**Distal-B	*F**upstream-C	*F**01:01:02:03(1943G, 2208C, $[TG]_{12}$)	*F**3UTR-A	0.0691
*F**Distal-B	*F**upstream-H	*F**01:01:02:03(1943G, $[TG]_{13}$)	*F**3UTR-A	0.0526
*F**Distal-A	*F**upstream-B	*F**01:01:01:08	*F**3UTR-A	0.0395
*F**Distal-A	*F**upstream-G	*F**01:01:01:08(-222G, 2698G, $[TG]_{12}$)	*F**3UTR-A	0.0263
*F**Distal-A	*F**upstream-A	*F**01:01:01:09(new205A)	*F**3UTR-A	0.0230
*F**Distal-E	*F**upstream-A	*F**01:01:01:11	*F**3UTR-A	0.0197
*F**Distal-A	*F**upstream-A	*F**01:02(1193C, 1943G)	*F**3UTR-A	0.0164
*F**Distal-A	*F**upstream-A	*F**01:01:01:11	*F**3UTR-A	0.0132
*F**Distal-A	*F**upstream-A	*F**01:02(1193C, 1943G, new1946C)	*F**3UTR-A	0.0132
*F**Distal-D	*F**upstream-A	*F**01:03:01:01(new1378T, 1383G, $[TG]_{11}$)	*F**3UTR-A	0.0132
*F**Distal-C	*F**upstream-B	*F**01:01:01:01	*F**3UTR-A	0.0099
*F**Distal-A	*F**Upstream-D	*F**01:01:01:05(new1497C, 1943G, $[TG]_{12}$)	*F**3UTR-A	0.0099
*F**Distal-A	*F**upstream-A	*F**01:01:01:09(new2123T)	*F**3UTR-A	0.0099
*F**Distal-*F*	*F**upstream-A	*F**01:03:01:01	*F**3UTR-A	0.0099
*F**Distal-A	*F**upstream-A	*F**01:01:01:09(3189T)	*F**3UTR-B	0.0066
*F**Distal-J	*F**upstream-H	*F**01:01:02:03(1943G, $[TG]_{14}$)	*F**3UTR-A	0.0066
*F**Distal-A	*F**upstream-A	*F**01:03:01:01	*F**3UTR-A	0.0066
*F**Distal-C	*F**upstream-B	*F**01:01:01:01(new145A)	*F**3UTR-A	0.0033
*F**Distal-C	*F**upstream-B	*F**01:01:01:01($[TG]_{11}$)	*F**3UTR-A	0.0033
*F**Distal-A	*F**upstream-A	*F**01:01:01:09($[TG]_{12}$)	*F**3UTR-A	0.0033
*F**Distal-B	*F**upstream-F	*F**01:01:02:03(1943G, $[TG]_{12}$)	*F**3UTR-A	0.0033
*F**Distal-B	*F**upstream-C	*F**01:01:02:06(1225nolargedel, 1230A)	*F**3UTR-A	0.0033
*F**Distal-I	*F**upstream-A	*F**01:03:01:01(1383G)	*F**3UTR-A	0.0033
*F**Distal-D	*F**upstream-A	*F**01:03:01:01(1383G, $[TG]_{11}$)	*F**3UTR-A	0.0033
*F**Distal-F	*F**upstream-A	*F**01:03:01:01($[TG]_{10}$)	*F**3UTR-A	0.0033
*F**Distal-A	*F**upstream-A	*F**01:03:01:01($[TG]_{11}$)	*F**3UTR-A	0.0033
*F**Distal-A	*F**upstream-A	*F**01:03:01:01($[TG]_{13}$)	*F**3UTR-A	0.0033
			Nucleotide diversity	0.0048 ± 0.0023
			Gene diversity	0.9109 ± 0.0061
			Tajima's *D*	1.4647, P = 0.0618

[a] The list of promoter haplotypes are depicted at Table 14. [b] The list of coding haplotypes are depicted at Table 15.

[c] As the 3'UTR haplotypes are represented by an unique variable site at position +3061 (Table S3, appendix III), we contented to cite it in the text (at sections 5.4.3). The extended sequences are available upon request. [d] Haplotypes were classified in descending order of their frequency.

A Tajima's *D p-value* was computed by the comparison of the estimated statistic to a distribution of estimates computed for 50,000 random samples of the same sample size and level of polymorphism as the observed data and represents the probability of obtaining a simulated Tajima's *D* larger than the observed.

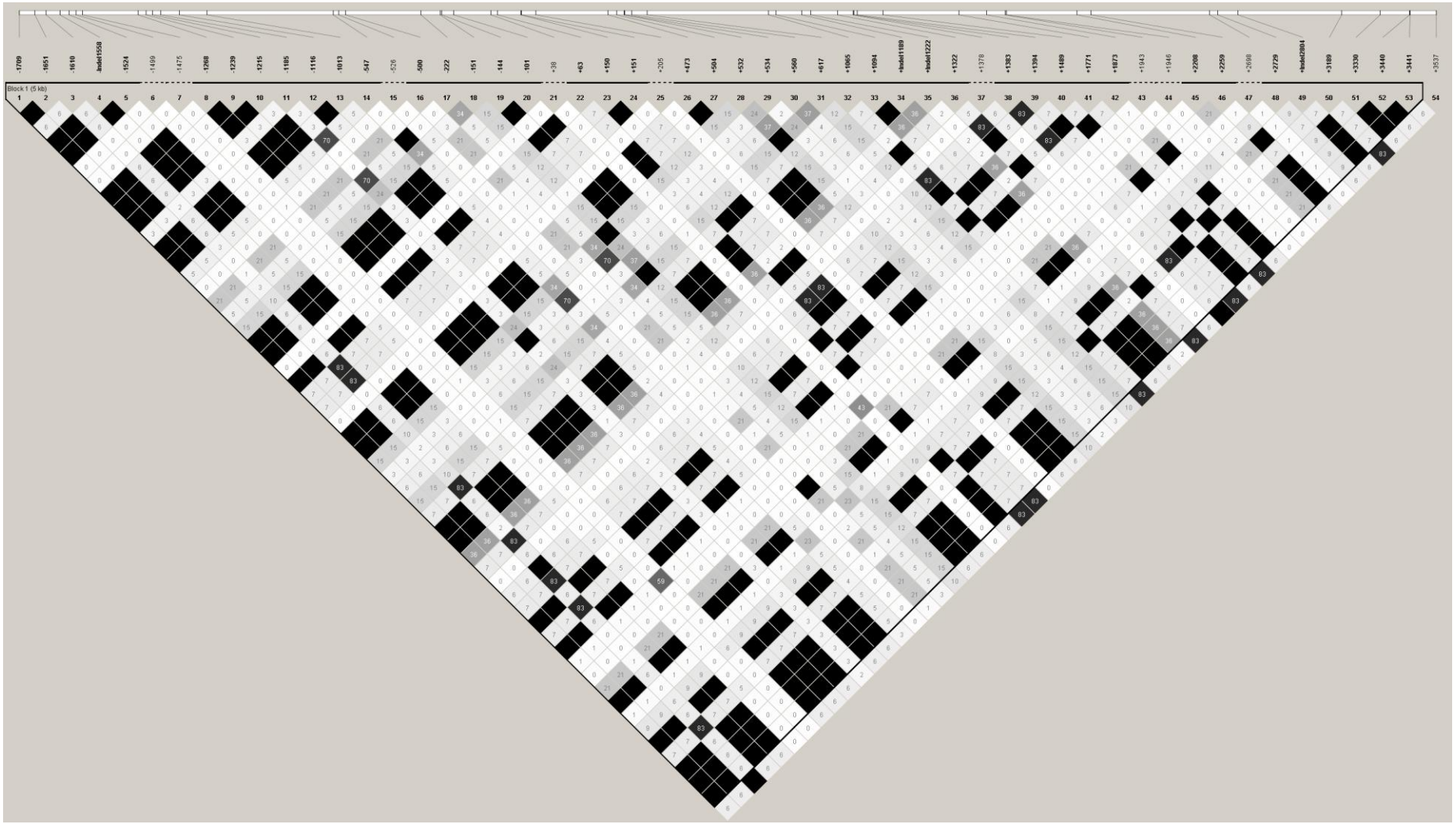


Figure 11. *LD patterns for the entire HLA-F region encompassing positions -1709 to +3537 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (x100) for the marker pairs are listed in the corresponding boxes. The LD plot was generated using SNPs with a minimum allele frequency (MAF) of 1%.*

RESULTS: PART II

PART II: GENETIC ASSOCIATION STUDY

Association between *HLA-G, -E and -F* polymorphisms (SNPs and Indel), and haplotypes with human *P. falciparum* malaria and urogenital schistosomiasis in Beninese Toffin young children.

5.5. *HLA-G, -E* AND *-F* POLYMORPHISMS (SNPs and Indels), HAPLOTYPES AND PHENOTYPES INCLUDED IN ASSOCIATION STUDY

Before starting the association studies, some quality control needed to be performed on the unphased genotype data. All variable sites observed at *HLA-E, -F* and *-G* were filtrated using the PLINK software, and only polymorphisms that: **i)** exhibited MAF≥ 0.05, **ii)** fit the Hardy-Weinberg equilibrium (α set to 0.05), and **iii)** presented a minimum genotype rate ≥ 90%, were analyzed (Table 17A).

Firstly, the variable sites (without phenotypes) were filtrered against genotypes and Hardy-Weinberg expections, using the quality control filters mentioned above. Out of the 196 variable sites used as input, none was removed due to missing genotypes, and 24 variants were removed due to non-adherence to the Hardy-Weinberg exact test. The overall genotyping rate was higher than 97%, and 172 variable sites were passed forward after these filters.

Secondly, the 172 variable sites were again filtered, using the MAF filter (MAF≥ 0.05) and adjusted for both, *P. falciparum* malaria and urogenital bilharzia phenotypes; i.e., (*''Infected/non-infected'', ''Symptomatic infection/non-symptomatic infection'', ''Resistant/non-resistant, ''Co-infected/non-co-infected'', ''Sub-microscopic infection/non-submicroscopic infection'', ''parasite density (PD)'', ''number of P. falciparum infection''* and *''number of P. falciparum symptomatic infection'')* for *P. falciparum and (*''*Infected/non-infected''* and *"Intensity of infection")* for *S. haematobium*. Overall, 67 variable sites were removed due to MAF < 0.05. Finally, 105 variable sites were passed forwad after the filter and were included in the association analysis for both infectious diseases (Table 17A).

The association tests were performed using haplotypes shown in Table 17B, including the 5' URR, coding region, 3' UTR, and extended haplotypes of each gene. The coding haplotypes were defined as ''protein-encoded allele groups''. For instance, ''*G**01:01-encoded allele group'' (Table 17B) encompasses all coding alleles (official alleles described at IPD-IMGT/HLA database, as well as derivatives ones) that generate the ''G*01:01 protein". The same strategy was used to define the *HLA-E* and *-F* coding haplotypes (Table 17B).

Table 17. *HLA-G, -E* and *-F* polymorphisms (A) and haplotypes (B) included in the association studies tests, performed before (unfiltered) and after filtering using the PLINK software.

(A) Minor allele frequency (MAF) of *HLA-G, -E* and *-F* polymorphisms included in association tests [a]

Genes	Evaluated segments	Number of polymorphisms (unfiltrated)	Filtrated (MAF≥ 0.05; p-HW ≥ 0.05; Min. genotype rate > 0.97)	Number of haplotypes with frequency ≥ 0.05
HLA-G	Entire gene	91	66	19
HLA-E	Entire gene	40	18	19
HLA-F	Entire gene	65	21	18
Total		**196**	**105**	**56**

(B) List *of HLA-G, -E* and *-F* haplotypes exhibiting minimum frequency of 0.05 included in the association tests

Genes	5' URR [b]	IPD-IMGT/HLA region [c]	3' UTR [d]	Extended haplotypes [e]
HLA-G	PROMO-*G*0104a, *G*010101a, *G*010101b, *G*010102a, *G*010101f	*G**01:01, *G**01:03, *G**01:04, *G**01:05N-(encoded-allele groups)	*HLAG*-3UTR-01, -2, -3, -4, -5, -6	extGH1, -2, -5, 10
HLA-E	*E*-Promo-1, -2, -6, -7, -15, -20	*E**01:01, E*01:03, *E**01:05-(encoded-allele group)	*HLA-E*-3UTR-1, -2, -3, **-5**	extHE1, -3, -4, -6, -7, -8
HLA-F	*F**Distal (A, -B, -C, -D), *F**upstream (A, -B, -C, -D, -H)	*F**01:01, F*01:02, *F**01:03-(encoded-allele group)	_	extHF1, -15, -19, -24, 35, 38

[a] **MAF:** minor allele frequency, **p-HW:** Hardy-Weinberg p-value.

[b] Upstream regulatory region (URR) haplotypes shown in Tables 6, 10 and 14.

[c] **IPD-IMGT/HLA** region haplotypes shown in Tables 7, 11 and 15.

[d] 3' Untranslated region (UTR) haplotypes shown in Tables 8 and 12.

[e] **ext:** Extended haplotypes shown in Tables 9, 13 and 16. The list of frequent extended haplotypes (frequency > 5%) and specific haplotypes they carried shown in Table S4, appendix IV.

The quality of the final linear model used to analyze the associations between haplotypes and phenotypes, using the R-Packages, are shown in Figures **(12)** *''Number of malaria infection'',* **(13)** *''Number of symptomatic malaria episodes''* and **(14)** *''P. falciparum parasite density (PD)''*. The results obtained by the evaluation of the residual fitness, normality of standardized residuals, scale-location of standardized residuals and leverage of standardized residuals indicated that the model used was adequate for the proposed analyses. Noteworthy, these variables fit the normality expectations and the

standardized residual values for leverage fluctuated around ‘’zero’’, indicating that atypical values (low or high outliers) could not influence the significance of the results. Altogether, these analyses confirmed the best fitness of the model.

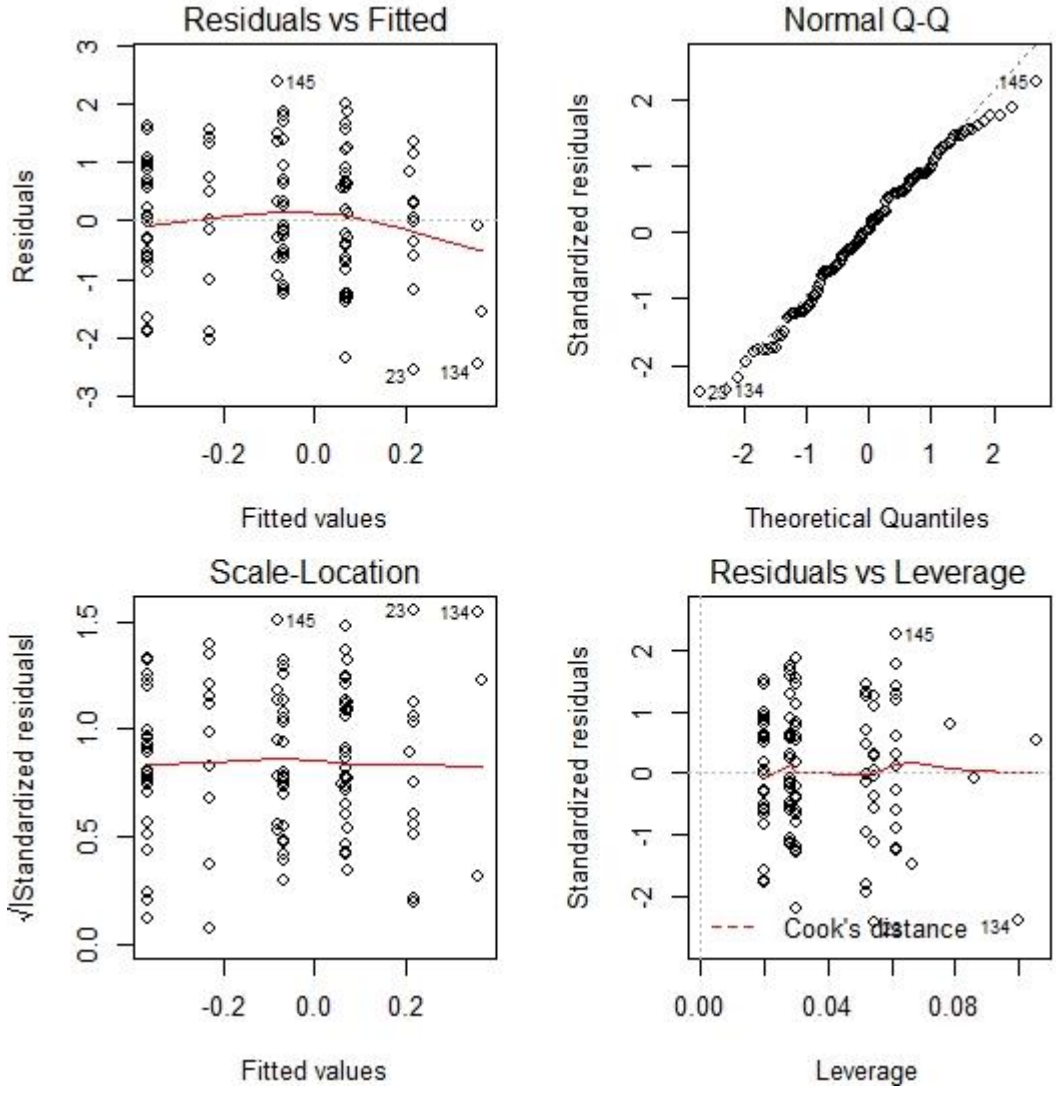


Figure 12. *Quality of the final linear model, analyzed by the R-packages, regarding the ‘’Number of malaria infection’’ as an explicative variable.*

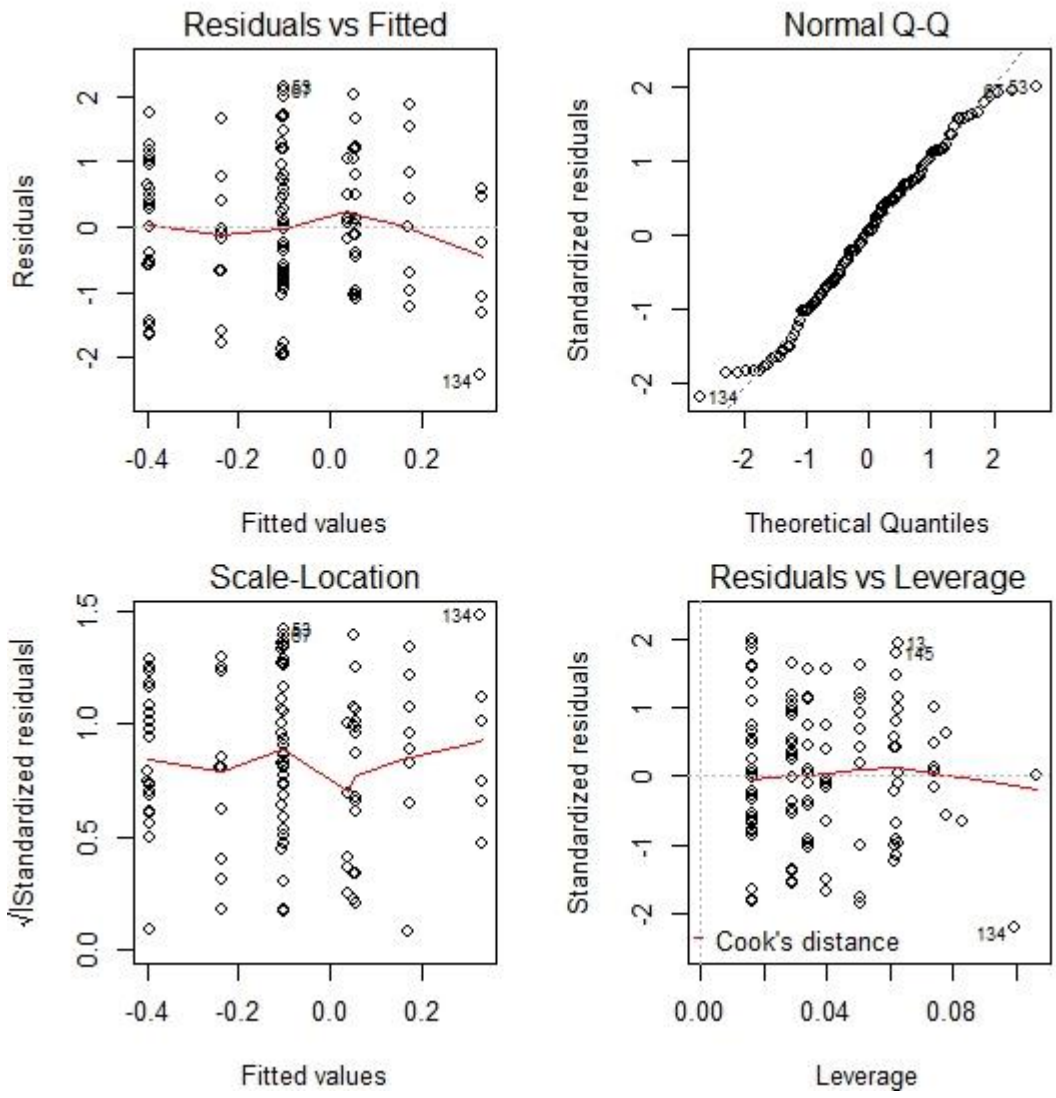


Figure 13. *Quality of the final linear model, analyzed by the R-packages, regarding the ''Number of symptomatic malaria episodes'' as an explicative variable.*

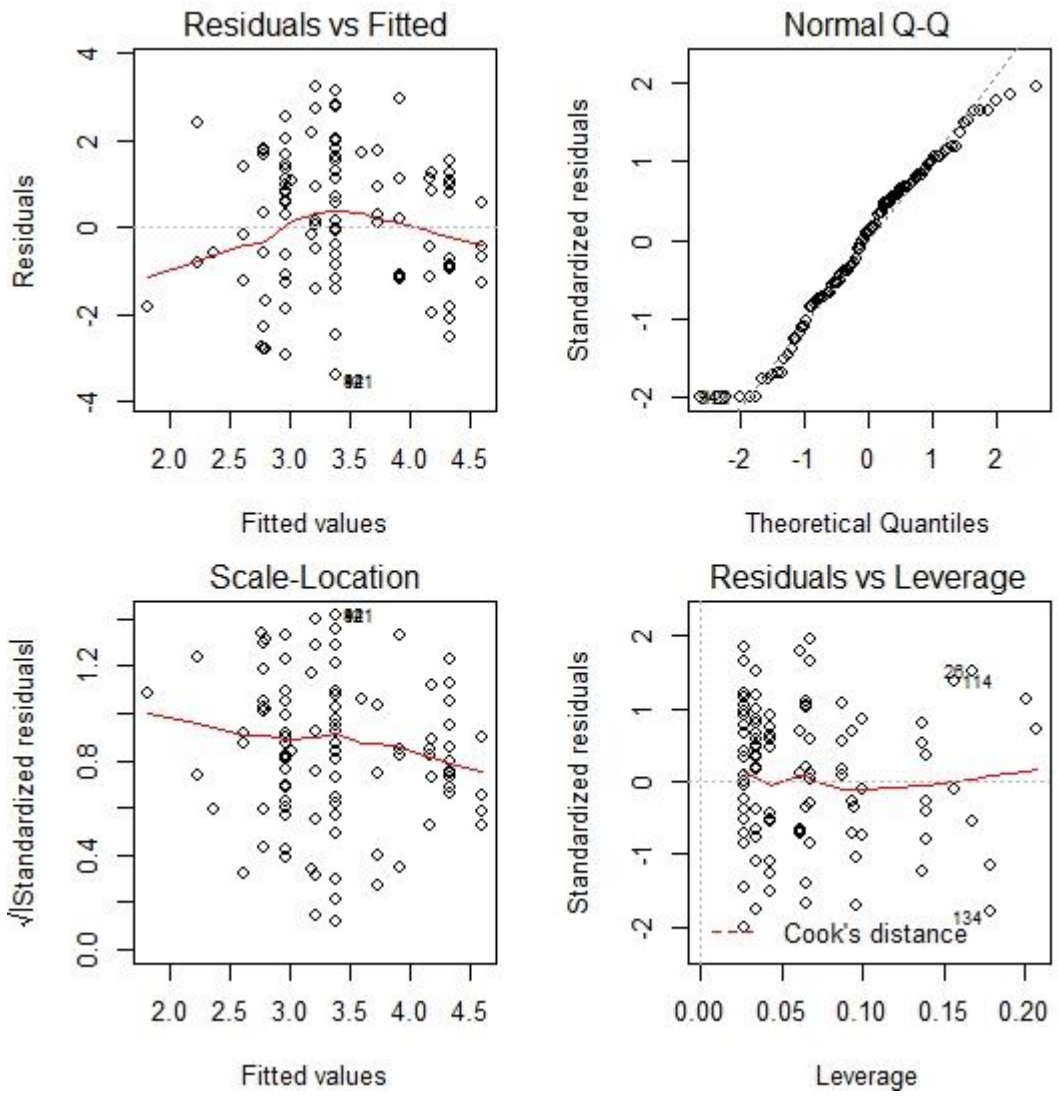


Figure 14. *Quality of the final linear model, analyzed by the R-packages, regarding the ''P. falciparum parasite density (DP)'' as an explicative variable.*

Only the results exhibiting *P*-values < 0.01 were considered as significant in the analysis of genetic association between polymorphisms and our phenotypes of interest to avoid too much false positive results and to assure the reliability of ours results (Benjamin et al. 2018). Although significant results at $P < 0.05$ threshold are shown in Tables, we will only discuss the results that fit the *P*-value 0.01 threshold.

For haplotypic results, only the results exhibiting *P*-values < 0.01 before correction and *P*-values < 0.05 after Bonferroni correction were considered significant.

5.6. *HLA-G, -E* AND *-F* POLYMORPHISMS ASSOCIATED TO SUSCEPTIBILITY OR PROTECTION TO *P. FALCIPARUM* MALARIA IN BENINESE TOFFIN YOUNG CHILDREN

In this section, we describe the significant associations found between *HLA-G, -E, -F* genes polymorphisms and susceptibility or protection against *P. falciparum* malaria.

5.6.1. HLA-G, -E and -F polymorphisms associated to susceptibility to P. falciparum malaria in Beninese Toffin young children

After the multivariate analysis (using either a logistic or linear model), run on different *P. falciparum* malaria clinical phenotypes (Table 18 to 22) using the 105 filtered polymorphisms observed across the whole *HLA-G, -E and -F* genes (Table 17A), we identified the *HLA-G* 14bp insertion allele at position +2960 as associated with susceptibility to malaria. This significant association can be observed in Tables 18 to 22, in which the *HLA-G* 14bp Ins was associated under a dominant model with an increased risk for *P. falciparum* infection, both for the occurrence of infection (***P* = 0.0241, OR = 4.060, CI95% 1.202 to 13.720,** Table 18, for all infections, and ***P* = 0.0022, OR = 5.220, CI95% 1.809 to 15.060,** Table 19, for symptomatic infections), and for the number of infection episodes (***P* = 0.0024, Beta = 0.611, CI95% 0.224 to 0.997,** Table 20, for all infections, and ***P* = 0.0002, Beta = 0.742, CI95% 0.364 to 1.120,** Table 21, for symptomatic infections), and finally, also associated to non-resistant phenotype (i.e., children that registered more symptomatic *P. falciparum* malaria infections than asymptomatic, during the follow-up; ***P* = 0.0127, OR = 0.087, CI95% 0.013 to 0.594,** Table 22).

Table 18. Polymorphisms associated with susceptibility to *P. falciparum* infection when infected individuals were compared to non-infected, using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/ HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P
6	*HLA-G*	Indel 14bp	rs3711 94629	297985 81	14bp (+2960)	3' UTR	Ins	**DOM**	141	4.060	0.621	1.202	13.720	2.256	0.0241
6	*HLA-E*	T>C	rs1059 510	304577 32	+424	Exon2	T	DOM	142	4.757	0.795	1.001	22.610	1.961	0.0499

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. DOM; for dominant model. The model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age.

Table 19. Polymorphisms associated to susceptibility to *P. falciparum* symptomatic infection when symptomatic individuals were compared to non-symptomatic, using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r^2 *
6	*HLA-G*	Indel 14bp	rs371194629	29798581	14bp (+2960)	3' UTR	Ins	ADD	141	2.569	0.424	1.120	5.894	2.227	0.0259	
								DOM	141	5.220	0.541	1.809	15.060	3.057	**0.0022**	
6	*HLA-E*	T>C	rs1059510	30457732	+424	Exon2	T	DOM	142	3.574	0.603	1.097	11.640	2.113	0.0346	
6	*HLA-E*	G>A	rs17195376	30460942	+3634	3' UTR	A	DOM	142	8.095	1.063	1.009	64.96	1.968	0.0491	1
6	*HLA-E*	G>A	rs9283	30461605	+4297	3' UTR	A	DOM	142	8.095	1.063	1.009	64.96	1.968	0.0491	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, DOM; for dominant model. The model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age. ***** +3634 G>A and +4297 G>A are in complete LD (r^2=1)

Table 20. Polymorphisms associated with an increased number of *P. falciparum* infections using multiple linear regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	Beta *	SE	L95	U95	Stat	P
6	*HLA-G*	Indel 14bp	rs371194629	29798581	14bp (+2960)	3' UTR	Ins	**DOM**	141	0.611	0.197	0.224	0.997	3.095	**0.0024**

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. DOM; for dominant model. The model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age. ***** Beta represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of *P. falciparum* infection).

Table 21. Polymorphisms associated with an increased number of symptomatic *P. falciparum* infections using multiple linear regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	Beta *	SE	L95	U95	Stat	P	LD r^2 **
6	*HLA-G*	A>T	rs1736934	29794482	-1140	Promo	T	ADD	141	0.299	0.145	0.016	0.583	2.069	0.0404	0.644 and 1
								DOM	141	0.407	0.183	0.048	0.766	2.224	0.0278	
6	*HLA-G*	C>T	rs1130356	29796327	+706	Exon3	T	ADD	141	0.299	0.145	0.016	0.583	2.069	0.0404	0.644 and 1
								DOM	141	0.407	0.183	0.048	0.766	2.224	0.0278	
6	*HLA-G*	C>T	rs1049033	29797639	+2018	Exon5	T	ADD	141	0.299	0.145	0.016	0.583	2.069	0.0404	0.644 and 1
								DOM	141	0.407	0.183	0.048	0.766	2.224	0.0278	
6	*HLA-G*	G>A	rs915670	29798419	+2798	Intron6	A	ADD	141	0.299	0.145	0.016	0.583	2.069	0.0404	0.644 and 1
								DOM	141	0.407	0.183	0.048	0.766	2.224	0.0278	
6	*HLA-G*	Indel 14bp	rs371194629	29798581	+2960	3' UTR	Ins	ADD	141	0.348	0.133	0.087	0.609	2.616	0.0099	
								DOM	141	0.742	0.193	0.364	1.120	3.847	**0.0002**	
6	*HLA-G*	C>G	rs1610696	29798803	+3196	3' UTR	G	ADD	141	0.299	0.145	0.016	0.583	2.069	0.0404	0.644 and 1
								DOM	141	0.407	0.183	0.048	0.766	2.224	0.0278	

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, DOM; for dominant model. The model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age. ***** Beta represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of symptomatic *P. falciparum* infection). ****** r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms; $r^2 = 0.644$ for LD between *14bpIndel* and other SNPs, $r^2 = 1$ for LD between SNPs excepted *14bp Indel*.

Table 22. Polymorphisms associated with non-resistant phenotype when resistant individuals were compared to non-resistant[a], using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r^2 *
6	*HLA-G*	A>T	rs1736934	29794482	-1140	Promo	T	ADD	93	0.096	1.117	0.011	0.859	-2.096	0.0361	0.644 and 1
								DOM	93	0.084	1.146	0.009	0.793	-2.163	0.0306	
6	*HLA-G*	C>T	rs1130356	29796327	+706	Exon3	T	ADD	93	0.096	1.117	0.011	0.859	-2.096	0.0361	0.644 and 1
								DOM	93	0.084	1.146	0.009	0.793	-2.163	0.0306	
6	*HLA-G*	C>T	rs1049033	29797639	+2018	Exon5	T	ADD	93	0.096	1.117	0.011	0.859	-2.096	0.0361	0.644 and 1
								DOM	93	0.084	1.146	0.009	0.793	-2.163	0.0306	
6	*HLA-G*	G>A	rs915670	29798419	+2798	Intron6	A	ADD	93	0.096	1.117	0.011	0.859	-2.096	0.0361	0.644 and 1
								DOM	93	0.084	1.146	0.009	0.793	-2.163	0.0306	
6	*HLA-G*	Indel 14bp	rs371194629	29798581	+2960	3' UTR	Ins	ADD	93	0.101	0.943	0.016	0.640	-2.433	0.0150	
								DOM	93	0.087	0.982	1.013	0.594	-2.491	**0.0127**	
6	*HLA-G*	C>G	rs1610696	29798803	+3196	3' UTR	G	ADD	93	0.096	1.117	0.011	0.859	-2.096	0.0361	0.644 and 1
								DOM	93	0.084	1.146	0.009	0.793	-2.163	0.0306	

[a] **non-resistants** were children that registered more *P. falciparum* symptomatic infections than asymptomatic during the follow-up

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, DOM; for dominant model. The model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age. * r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms; r^2 = 0.644 for LD between *14bpIndel* and other SNPs, r^2 = 1 for LD between SNPs excepted *14bp Indel*.

5.6.2. HLA-G, -E and -F polymorphisms associated to protection against P. falciparum malaria in Beninese Toffin young children

Three polymorphisms showed significant associations with protection against *P. falciparum* malaria at the 0.01 significance threshold (Tables 23 to 26). Two located in *HLA-G* (i.e., *-1155 G>A* and *+755 C>A*, in complete LD with each other (r^2 = 1)) conferred protection against most *P. falciparum* clinical phenotypes evaluated in this series (Tables 23 to 26). Indeed, the individuals carrying two A alleles (*A/A*) showed a lower prevalence of: **i)** "*P. falciparum* infection" (***P* = 0.0069, OR = 0.096, CI95% 0.018 to 0.525,** Table 23), **ii)** "*P. falciparum* symptomatic infection" (***P* = 0.0069, OR = 0.104**, **CI95% 0.020 to 0.536,** Table 24). In addition, the A/A genotype showed suggestive associations with "low *P. falciparum* parasitemia" (***P* = 0.0360, Beta = -1.486**, **CI95% -2.858 to -0.115,** Table 25) and "reduced number of *P. falciparum* symptomatic infections" (***P* = 0.0217, Beta = -0.353**, **CI95% -0.651 to -0.055,** Table 26) at the 0.05 threshold. Finally, one polymorphism located at the *HLA-E* promoter *(-1988 T>C)* was also significantly associated with protection against "*P. falciparum* infection", under the additive model (***P* = 0.0061, OR = 0.204**, **CI95% 0.066 to 0.636,** Table 23).

Moreover, three *HLA-F* polymorphisms *(-151 T>A, +532 T>C and +617 C>G,* in complete LD with each other, showed suggestive association with "low *P. falciparum* parasitemia" (***P* = 0.0155, Beta = -0.802**, **CI95% -1.440 to -0.163,** Table 25), under the dominant model, at the 0.05 significance level.

Table 23. Polymorphisms associated with protection against *P. falciparum* infection when infected individuals were compared to non-infected, using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/ HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r^2 *
6	*HLA-G*	G>A	rs3823321	29794467	-1155	Promo	A	**REC**	141	0.096	0.866	0.018	0.525	-2.703	**0.0069**	1
6	*HLA-G*	C>A	rs12722477	29796376	+755	Exon3	A	**REC**	141	0.096	0.866	0.018	0.525	-2.703	**0.0069**	1
6	*HLA-E*	T>C	rs17875360	30455321	-1988	Promo	C	**ADD**	142	0.204	0.579	0.066	0.636	-2.741	**0.0061**	0.637
								DOM	142	0.187	0.662	0.051	0.685	-2.532	0.0113	
6	*HLA-E*	A>G	rs61356961	30457205	-104	Promo	G	DOM	142	0.273	0.637	0.078	0.950	-2.040	0.0413	1
6	*HLA-E*	G>C	rs11548296	30458933	1625	Exon4	C	DOM	142	0.273	0.637	0.078	0.950	-2.040	0.0413	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference.

IMGT/HLA: position at IPD-IMGT/HLA database.

Test: test performed by PLINK. ADD; for additive model, DOM; for dominant model and REC; for recessive model. Model was chosen basing on the more significant p-value.

NMISS: number of tested samples, **OR:** odds ratio, SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age.

***** r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms; $r^2 = 1$ for LD between *HLA-G* SNPs or for LD between *HLA-E -104 G* and *+1625 C*; $r^2 = 0.637$ for LD between *-1988 C* and (*-104 G* and *+1625 C*) polymorphisms within *HLA-E* gene.

Table 24. Polymorphisms associated with protection against *P. falciparum* symptomatic infection when symptomatic individuals were compared to non-symptomatic, using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r^2 *
6	*HLA-G*	G>A	rs3823321	29794467	-1155	Promo	A	**REC**	141	0.104	0.839	0.020	0.536	-2.704	**0.0069**	1
6	*HLA-G*	C>A	rs12722477	29796376	+755	Exon3	A	**REC**	141	0.104	0.839	0.020	0.536	-2.704	**0.0069**	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, REC; for recessive model. Model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, **OR:** odds ratio, SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age. * *-1155 G>A* and *+755 C>A* are in complete LD (r2=1).

Table 25. Polymorphisms associated with low *P. falciparum* parasitemia (PD) using multiple linear regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	Beta *	SE	L95	U95	Stat	P	LD r^2 **
6	*HLA-G*	G>A	rs3823321	29794467	-1155	Promo	A	REC	110	-1.486	0.700	-2.858	-0.115	-2.124	0.0360	1
6	*HLA-G*	C>A	rs12722477	29796376	+755	Exon3	A	REC	110	-1.486	0.700	-2.858	-0.115	-2.124	0.0360	1
6	*HLA-F*	T>A	rs1362125	29691090	-151	Proximal promo	A	DOM	113	-0.802	0.326	-1.440	-0.163	-2.460	0.0155	1
6	*HLA-F*	T>C	rs2076179	29691772	532	Exon2	C	DOM	113	-0.802	0.326	-1.440	-0.163	-2.460	0.0155	1
6	*HLA-F*	C>G	rs2072895	29691857	617	Exon2	G	DOM	113	-0.802	0.326	-1.440	-0.163	-2.460	0.0155	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. REC; for recessive model, DOM; for dominant model. Model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age.

* Beta represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (*P. falciparum* parasitemia). ** r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms; $r^2 = 1$ for LD between SNPs within both genes (*HLA-G* or *HLA-F*).

Table 26. Polymorphisms associated with low number of symptomatic *P. falciparum* infection episodes using multiple linear regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	Beta *	SE	L95	U95	Stat	P	LD r2 **
6	*HLA-G*	G>A	rs3823321	29794467	-1155	Promo	A	ADD	141	-0.353	0.152	-0.651	-0.055	-2.322	0.0217	1
								DOM	141	-0.367	0.180	-0.721	-0.014	-2.036	0.0436	
6	*HLA-G*	C>A	rs12722477	29796376	+755	Exon3	A	ADD	141	-0.353	0.152	-0.651	-0.055	-2.322	0.0217	1
								DOM	141	-0.367	0.180	-0.721	-0.014	-2.036	0.0436	

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, DOM; for dominant model. Model was chosen basing on the more significant p-value. **NMISS:** number of tested samples, SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age.

* Beta represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of symptomatic *P. falciparum* infection). ** r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms; $r^2 = 1$ for LD between SNPs.

5.7. *HLA-G, -E AND -F* HAPLOTYPES ASSOCIATED WITH SUSCEPTIBILITY OR PROTECTION TO *P. FALCIPARUM* MALARIA IN BENINESE TOFFIN YOUNG CHILDREN

The ***HLA.E*.3UTR.3** and ***E*.01.03.encoded-allele group** haplotypes were associated with protection against *P. falciparum* malaria (Tables 27 to 33). After Bonferroni correction, only the ***E*.01.03.encoded-allele group** continued conferring protection against ''*P. falciparum* symptomatic infection'' (***Pc* = 0.0364**, **OR = 0.033**, **CI95% 0.003 to 0.240,** Table 28) and marginally against higher ''*P. falciparum* parasitemia'' (***Pc* = 0.0876, regression coefficient = -0.948**, **CI95% -1.695 to -0.200,** Table 29). In addition, without reaching significance at $P < 0.05$, the ***E*.01.03.encoded-allele group** was associated to **i)** protection against *P. falciparum* infection (Table 27), **ii)** to a reduced number of *P. falciparum* infection and *P. falciparum* symptomatic infection episodes (Table 30 and 31).

Table 27. Haplotypes associated with susceptibility or protection against *P. falciparum* infection when infected individuals were compared to non-infected, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**
E.Promo.15-presence	6.958E+06	1.605E+03	1.481E-43	NA	0.9922
***E*.01.03.encoded allele group-presence**	0.123	1.198	0.005	0.858	0.0798
F.Upstream.D-presence	0.921	0.720	0.225	4.074	0.9092
F.Upstream.H-presence	1.406	1.121	0.216	27.740	0.7611
F.01.01.encoded allele group-presence	9.569	1.507	0.336	276.179	0.1341
F.01.02.encoded allele group-presence	0.118	1.055	0.013	1.067	0.0428

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*****Covariates included in logistic multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 146 children.

Table 28. Haplotypes associated with susceptibility or protection against *P. falciparum* infection when symptomatic individuals were compared to non-symptomatic, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**	**Pc**
age (5 years)	2.68E-07	4.18E+03	NA	6.20E+114	0.9971	
age (6 years)	1.91E-07	4.18E+03	NA	4.13E+96	0.9970	
age (7 years)	2.94E-08	4.18E+03	NA	1.12E+259	0.9967	
age (8 years)	1.00E-08	4.18E+03	NA	3.60E+249	0.9965	
E.Promo.1-presence	0.394	0.786	0.076	1.753	0.2357	
E.Promo.7-presence	40.473	1.341	4.184	1.09E+03	0.0058	
***E*.01.03.encoded allele group-presence**	0.033	1.135	0.003	0.240	**0.0028**	**0.0364**
HLA.E.3UTR.2-presence	4.938	0.974	0.853	44.786	0.1011	
HLA.E.3UTR.3-presence	0.066	1.022	0.007	0.428	0.0080	0.1040
F.Upstream.D-presence	0.530	0.802	0.109	2.695	0.4291	
F.upstream.H-presence	2.11E-08	1.69E+03	1.23E-168	8.00E+30	0.9917	
F.01.01.encoded allele group-presence	4.427	1.668	0.125	1.64E+02	0.3725	
G.01.03.encoded allele group-presence	1.25E+07	1.69E+03	3.02E-40	NA	0.9923	

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value. **Pc:** corrected p-value. 13 tests performed (α = 0.0038). Therefore, we considered significative, only association with Bonfferroni corrected **Pc** < 0.05.

*****Covariates included in logistic multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 128 children.

Table 29. Haplotypes associated with high or low *P. falciparum* parasitemia (PD), using linear multivariate regression

Covariates *	**Estimation**	**SE**	**L95**	**U95**	**P**	**Pc**
***E*.01.03.encoded allele group-presence**	-0.948	0.382	-1.695	-0.200	**0.0146**	**0.0876**
F.Upstream.D-presence	-0.417	0.368	-1.138	0.305	0.2606	
F.upstream.H-presence	-1.394	0.751	-2.866	0.079	0.0665	
F.01.02.encoded allele group-presence	-1.144	0.684	-2.485	0.197	0.0975	
extHF35-presence	-0.592	0.455	-1.485	0.300	0.1964	
G.01.03.encoded allele group-presence	1.229	0.646	-0.038	2.495	0.0600	

SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value. **Pc:** corrected p-value. 6 tests performed (α = 0.0083). Therefore, we considered significative, only association with Bonfferroni corrected **Pc** < 0.05. *Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (exhibited here) was available for 115 children.

Estimation represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (*P. falciparum* parasite density) due to the presence of the independent ones.

Table 30. Haplotypes associated with high or low number of *P. falciparum* infection episodes using linear multivariate regression

Covariates *	Estimation	SE	L95	U95	P
E.01.01.encoded allele group-presence	-0.435	0.236	-0.897	0.026	0.0668
***E*.01.03.encoded allele group-presence**	-0.298	0.213	-0.716	0.120	0.1644
HLAG.3UTR.04-presence	0.138	0.257	-0.366	0.642	0.5925
HLAG.3UTR.06-presence	0.286	0.267	-0.238	0.810	0.2863

SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*****Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (exhibited here) was available for 146 children.

Estimation represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of *P. falciparum* infection episodes) due to the presence of the independent ones.

Table 31. Haplotypes associated with high or low number of *P. falciparum* symptomatic infection episodes using linear multivariate regression

Covariates *	Estimation	SE	L95	U95	P
***E*.01.03.encoded allele group-presence**	-0.159	0.201	-0.552	0.234	0.4290
F.Distal.D-presence	-0.293	0.203	-0.691	0.105	0.1510
HLAG.3UTR.04-presence	-0.004	0.255	-0.504	0.495	0.9860
HLAG.3UTR.06-presence	0.275	0.267	-0.248	0.798	0.3050

SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*****Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (exhibited here) was available for 146 children.

Estimation represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of *P. falciparum* symptomatic infection episodes) due to the presence of the independent ones.

Since ***E*.01.03.encoded allele group** conferred protection against various *P. falciparum* malaria phenotypes (Tables 27 to 31), we decided to identify the specific haplotype associated with protection. Considering the coding *HLA-E* alleles exhibiting minimum frequency of 5%, the **E*01:03** protein can be encoded by *E**01:03:02:01 and *E**01:03:05-compatible alleles

(Table 11). Only the ***E*.01.03.05-compatible** was associated to *E*-Promo-20 in extended haplotype (7%, Table 13). In addition, the *E*-Promo-20 (Table 10) carries the *-1988 C allele*; associated with protection against malaria (Table 23). Thus, the ***E*.01.03.05-compatible** could be the target. Once targeted the specific haplotype, we performed the same association tests that had been performed with ***E*.01.03.encoded allele group.** The results obtained (Table 32 to 35) confirmed the protective effect of this haplotype against *P. falciparum* malaria, considered at the levels of: **i)** infection (***Pc* = 0.0312, OR = 0.067, CI95% 0.008 to 0.375,** Table 32), **ii)** *P. falciparum* symptomatic infection (Table 33, not significant at $P < 0.05$), and **iii)** number of malaria episodes (Table 34 and 35, not significant at $P < 0.05$).

Table 32. Influence of *E*.01.03.05.compatible on *P. falciparum* infection when infected individuals were compared to non-infected, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**	**Pc**
E.Promo.15-presence	3.70E+06	2.47E+03	1.52E-79	NA	0.9951	
***E*.01.03.05.compatible-presence**	0.067	0.940	0.008	0.375	**0.0039**	**0.0312**
HLA.E.3UTR.5-presence	0.026	1.204	0.002	0.262	0.0024	
extHE6-presence	0.136	0.927	0.017	0.719	0.0313	
F.Upstream.D-presence	0.256	0.859	0.044	1.382	0.1124	
F.01.01.encoded allele group-presence	6.859	2.239	0.084	5.39E+02	0.3897	
F.01.02.encoded allele group-presence	0.061	1.140	0.006	0.623	0.0139	
extHF35-presence	0.101	1.186	0.008	1.092	0.0536	

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value. **Pc:** corrected p-value. 8 tests performed ($\alpha = 0.0063$). Therefore, we considered significative, only association with Bonfferroni corrected **Pc** < 0.05.

*****Covariates included in logistic multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 146 children.

Table 33. Influence of *E*.01.03.05.compatible on *P. falciparum* infection when symptomatic individuals were compared to non-symptomatic, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**
Age (5 years)	7.10E-07	4.56E+03	NA	9.08E+126	0.9975
Age (6 years)	7.72E-07	4.56E+03	NA	9.58E+119	0.9975
Age (7 years)	1.03E-07	4.56E+03	NA	8.68E+276	0.9972
Age (8 years)	1.13E-07	4.56E+03	NA	8.02E+254	0.9972
E.Promo.1-presence	0.312	0.752	0.061	1.230	0.1211
E.Promo.7-presence	4.936	1.230	0.600	1.12E+02	0.1944
***E*.01.03.05.compatible-presence**	0.371	0.667	0.096	1.364	0.1375
HLA.E.3UTR.5-presence	0.055	1.088	0.005	0.409	0.0078
F.Upstream.D-presence	0.704	0.749	0.165	3.274	0.6397
F.upstream.H-presence	4.21E-08	1.63E+03	NA	1.22E+40	0.9917
G.01.03.encoded allele group-presence	1.43E+07	1.63E+03	1.88E-37	NA	0.9919
HLAG.3UTR.04-presence	0.532	0.776	0.119	2.663	0.4153

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*Covariates included in logistic multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 128 children.

Table 34. Influence of *E*.01.03.05.compatible on number of *P. falciparum* infection episodes using linear multivariate regression

Covariates *	**Estimation**	**SE**	**L95**	**U95**	**P**
E.01.01.encoded_allele_group-presence	-0.377	0.225	-0.818	0.063	0.0957
***E*.01.03.05.compatible-presence**	-0.126	0.245	-0.607	0.355	0.6079
extHF15-presence	-0.055	0.219	-0.485	0.376	0.8039
PROMO.G0104a-presence	-0.387	0.215	-0.809	0.035	0.0743
PROMO.G010101f-presence	0.358	0.376	-0.379	1.094	0.3428
HLAG.3UTR.04-presence	0.085	0.260	-0.424	0.594	0.7444
HLAG.3UTR.06-presence	0.165	0.385	-0.590	0.920	0.6685

SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (exhibited here) was available for 146 children.

Estimation represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of *P. falciparum* infection episodes) due to the presence of the independent ones.

Table 35. Influence of *E*.01.03.05.compatible on number of *P. falciparum* symptomatic infection episodes using linear multivariate regression

Covariates *	**Estimation**	**SE**	**L95**	**U95**	**P**
E.01.01.encoded_allele_group-presence	-0.441	0.229	-0.890	0.007	0.0558
***E*.01.03.05.compatible-presence**	-0.335	0.225	-0.777	0.106	0.1393
F.Upstream.D-presence	0.034	0.202	-0.362	0.429	0.8674
G.01.03.encoded allele group-presence	0.187	0.238	-0.280	0.654	0.4341
HLAG.3UTR.04-presence	0.211	0.256	-0.290	0.712	0.4107
HLAG.3UTR.06-presence	0.437	0.276	-0.103	0.978	0.1149

SE: standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value.

*Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (exhibited here) was available for 146 children.

Estimation represents regression coefficients which can be positive or negative. Regression coefficients measure the increase (positive value) or decrease (negative value) of the dependent variable (number of *P. falciparum* symptomatic infection episodes) due to the presence of the independent ones.

5.8. *HLA-G, -E* AND *-F* POLYMORPHISMS AND HAPLOTYPES ASSOCIATED WITH SUSCEPTIBILITY OR PROTECTION TO *S. HAEMATOBIUM* INFECTION IN BENINESE TOFFIN YOUNG CHILDREN

At the $P < 0.01$ significance level, no polymorphism was associated with susceptibility or protection against *S. haematobium* disease phenotypes evaluated here (Tables 36 and 37). However, at the $P < 0.05$ significance level, the three *HLA-F* polymorphisms *(-151 T>A, +532 T>C and +617 C>G)* in complete LD with each other, that were previously shown to protect against "high *P. falciparum* parasitemia", under dominant model, (Table 25), also showed protection against *S. haematobium* infection, under recessive model **(*P* = 0.0428, OR = 0.115, CI95% 0.014 to 0.932,** Table 36**)**.

At haplotypic level, the ***E*.Promo.7** was marginally associated with the moderate *S. haematobium* infection (10–49 eggs/10 mL of urine) after Bonferroni correction **(*Pc* = 0.0525, OR = 0.070, CI95% 0.008 to 0.415,** Table 38**)**.

Table 36. Polymorphisms associated with protection against *S. haematobium* infection when infected individuals were compared to non-infected, using multiple logistic regression, adjusted for sex, age and water-related activity.

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r2 *
6	*HLA-F*	T>A	rs1362125	29691090	-151	Proximal promo	A	REC	98	0.115	1.066	0.014	0.932	-2.026	0.0428	1
6	*HLA-F*	T>C	rs2076179	29691772	+532	Exon2	C	REC	98	0.115	1.066	0.014	0.932	-2.026	0.0428	1
6	*HLA-F*	C>G	rs2072895	29691857	+617	Exon2	G	REC	98	0.115	1.066	0.014	0.932	-2.026	0.0428	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, REC; for recessive model, **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex, age and water-related activity (Fishing, bathing, laundry, dishing, agriculture, cross-water on foot). ***** r^2=1 shows the complete Linkage Disequilibrium (LD) between polymorphisms.

Table 37. Polymorphisms associated with "heavy *S. haematobium* infection" (>49 eggs/10 mL urine) using multiple logistic regression, adjusted for sex, age and water-related activity

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r2 *
6	*HLA-E*	G>A	rs17195376	30460942	+3634	3' UTR	A	ADD	33	6.458	0.930	1.043	40.000	2.005	0.0450	1
								DOM	33	6.458	0.930	1.043	40.000	2.005	0.0450	
6	*HLA-E*	G>A	rs9283	30461605	+4297	3' UTR	A	ADD	33	6.458	0.930	1.043	40.000	2.005	0.0450	1
								DOM	33	6.458	0.930	1.043	40.000	2.005	0.0450	
6	*HLA-F*	T>A	rs1362125	29691090	-151	Proximal promo	A	DOM	33	6.273	0.855	1.175	33.490	2.149	0.0317	1
6	*HLA-F*	T>C	rs2076179	29691772	+532	Exon2	C	DOM	33	6.273	0.855	1.175	33.490	2.149	0.0317	1
6	*HLA-F*	C>G	rs17875381	29691800	+560	Exon2	G	DOM	33	15.250	1.210	1.423	163.400	2.251	0.0244	
6	*HLA-F*	C>G	rs2072895	29691857	+617	Exon2	G	DOM	33	6.273	0.855	1.175	33.490	2.149	0.0317	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, DOM; for dominant model, **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex, age and water-related activity (Fishing, bathing, laundry, dishing, agriculture, cross-water on foot). ***** r^2 shows the strength of Linkage Disequilibrium (LD) between polymorphisms within each gene. Moderate *S. haematobium* infection = (10–49 eggs/10 mL of urine). Heavy *S. haematobium* infection = (>49 eggs/10 mL of urine).

Table 38. Haplotypes that could influence the intensity (moderate or heavy) infection of *S. haematobium*, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**	**Pc**
age (5 years)	1.71E+13	3.39E+03	6.49E-218	NA	0.9928	
age (6 years)	1.50E+07	2.40E+03	2.68E-205	NA	0.9945	
age (7 years)	4.95E+06	2.40E+03	8.09E-206	NA	0.9949	
age (8 years)	4.35E+07	2.40E+03	1.32E-204	NA	0.9942	
***E*.Promo.7-presence**	0.070	0.995	0.008	0.415	**0.0075**	**0.0525**
0.00525E.Promo.15-presence	0.044	1.407	0.002	0.524	0.0261	
F.Distal.B-presence	7.958	0.975	1.415	72.096	0.0335	

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value. **Pc:** corrected p-value. 8 tests performed (α = 0.0071). Therefore, we considered significative, only association with Bonfferroni corrected **Pc** < 0.05.

*****Covariates included in linear multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 46 children aged 4-8 years.

Moderate *S. haematobium* infection = (10–49 eggs/10 mL of urine)

Heavy *S. haematobium* infection = (>49 eggs/10 mL of urine)

5.9. HLA-G, -E AND *-F* POLYMORPHISMS AND HAPLOTYPES ASSOCIATED TO SUSCEPTIBILITY OR PROTECTION TO *P. FALCIPARUM* MALARIA IN YOUNG CHILDREN CO-INFECTED BY *S. HAEMATOBIUM*

From the 105 polymorphisms tested, *HLA-F (-151 T>A, +532 T>C* and *+617 C>G)* polymorphisms (in complete LD with each other), under additive model, were associated with the protection against *P. falciparum-S. haematobium* co-infection. However, this result dit not reach significance at the $P < 0.01$ (Table 39).

For the haplotypes, the *E*.Promo.2 protected against *P. falciparum- S. haematobium* co-infection, even after Bonferroni correction **(*Pc* = 0.0365, OR = 0.041, CI95% 0.004 to 0.264,** Table 40).

When we attempted to see whether the presence of *S. haematobium* could influence the susceptibility to *P. falciparum* infection, we did not observe significant differences between the *S. haematobium* infected group and *P. falciparum* infected group (Tables 41 A and B).

Table 39. Polymorphisms associated with the protection against *P. falciparum-S. haematobium* co-infection when co-infected individuals were compared to non-coinfected, using multiple logistic regression, adjusted for sex and age

Chr.	Genes	Alleles	idSNP	Hg19	IMGT/ HLA	Gene segments	Associated allele	Test	NMISS	OR	SE	L95	U95	Stat	P	LD r2 *
6	*HLA-F*	T>A	rs1362125	29691090	-151	Proximal promo	A	**ADD**	149	0.470	0.337	0.243	0.910	-2.240	0.0251	1
6	*HLA-F*	T>C	rs2076179	29691772	+532	Exon2	C	**ADD**	149	0.470	0.337	0.243	0.910	-2.240	0.0251	1
6	*HLA-F*	C>G	rs2072895	29691857	+617	Exon2	G	**ADD**	149	0.470	0.337	0.243	0.910	-2.240	0.0251	1

Chr.: chromosome, **Hg19:** polymorphism position at human genome reference, **IMGT/HLA:** position at IPD-IMGT/HLA database, **Test:** test performed by PLINK. ADD; for additive model, **NMISS:** number of tested samples, **OR:** odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value adjusted for sex and age ***** $r^2=1$ shows the complete Linkage Disequilibrium (LD) between polymorphisms.

Table 40. Haplotypes associated with susceptibility/resistance to *S. haematobium-P. falciparum* co-infection, when co-infected individuals were compared to non-coinfected, using logistic multivariate regression

Covariates *	**OR**	**SE**	**L95**	**U95**	**P**	**Pc**
E.Promo.2-presence	0.041	1.050	0.004	0.264	**0.0023**	**0.0365**
E.01.01.encoded-allele group-presence	0.083	1.023	0.010	0.546	0.0149	
HLA.E.3UTR.3-presence	0.197	0.922	0.027	1.042	0.0777	
extHE6-presence	2.862	0.936	0.482	20.009	0.2613	
F.Distal.A-presence	0.102	0.827	0.018	0.480	**0.0058**	0.0928
F.Distal.D-presence	0.076	1.263	0.005	0.821	0.0410	
F.upstream.A-presence	0.468	0.982	0.061	3.017	0.4392	
F.upstream.H-presence	1.323	1.282	0.115	20.296	0.8270	
F.01.01.encoded-allele group-presence	0.197	1.457	0.011	5.545	0.2652	
F.01.03.encoded-allele group-presence	7.569	1.331	0.597	124.209	0.1284	
extHF1-presence	8.623	1.008	1.343	74.670	0.0325	
G.01.03.encoded-allele group-presence	1.536	1.150	0.131	13.885	0.7091	
HLAG.3UTR.01-presence	7.178	0.972	1.010	49.747	0.0426	
HLAG.3UTR.04-presence	10.415	0.887	1.878	64.982	**0.0082**	0.1314
HLAG.3UTR.06-presence	1.351	0.877	0.229	7.541	0.7317	
extHG5-presence	0.419	1.197	0.038	4.567	0.4673	

OR: odds ratio, **SE:** standard error, **L95:** low confidence interval (CI 95%), **U95:** high CI 95%, **P:** asymptomatic p-value. **Pc:** corrected p-value. 16 tests performed (α = 0.0031). Therefore, we considered significative, only association with Bonfferroni corrected **Pc** < 0.05.

*Covariates included in logistic multivariate regression were sex, age and 56 haplotypes (promoter, coding, 3'UTR and extended *HLA-G, -E, -F* haplotypes). Information concerning the covariates in final model (presented here) was available for 151 children aged 4-8 years.

Table 41. Association between *S. haematobium* infection and susceptibility to *P. falciparum* infection.

(A) Influence of the presence of *S. haematobium* on susceptibility to *P. falciparum* malaria				
Level	*P. f* uninfected	*P. f* infected	P	Test
n	70	81		
S. h uninfected (%)	52 (74.3)	50 (61.7)	0.118	Exact
S. h infected (%)	18 (25.7)	31 (38.3)		

(B) Association between intensity of *S. haematobium* infection and susceptibility to *P. falciparum* infection				
Level	*P. f* uninfected	*P. f* infected	P	Test
n	70	81		
S. h heavy infection (>49 eggs/10 mL urine) (%)	10 (14.5)	11 (13.9)	0.114	Exact
S. h moderate infection (10–49 eggs/10 mL urine) (%)	7 (10.1)	18 (22.8)		
S. h uninfected (0 egg) (%)	52 (75.4)	50 (63.3)		

S. h = Schistosoma haematobium

P. f = Plasmodium falciparum

6. DISCUSSION

DISCUSSION: PART I

Non-synonymous substitutions at HLA class Ib gene variable sites observed at coding regions may lead to protein modifications, which may modify molecule binding to leukocyte receptors. Variable sites at regulatory regions (enhancers, promoters and 3' UTRs), may modify the binding sites to transcriptional (promoter region) and post-transcriptional (3' UTR by miRNAs) factors. In this study, we evaluated the nucleotide diversity across the major HLA class Ib genes, considered either as gene segments or as whole gene haplotypes. This approach may provide a better understanding of how gene variable sites may influence protein structure and the magnitude of protein production.

The study of HLA class Ib genes in African populations may help to understand the evolution of these immune checkpoint genes, emphasizing the study of the Toffin population that have moved from the Godomé Coast to escape from the war at Aja-Tado and kept their villages in the Northern shore of Nokoué lake (Principaud 1995). Since Toffins have lived simple lives as fishermen in the Beninese Sô-Ava region, and since they have built their stilt house in Nokoué lake, they are highly exposed to schistosomiasis (Ibikounlé et al. 2009, 2013, 2018) and also to malaria as in most African countries. The discussion of this component of the present study will be performed taking into account each region of the HLA class Ib gene as well as the gene as a whole.

6.1. OVERVIEW OF THE VARIABILITY AT THE ENTIRE *HLA-G, -E* AND *-F* GENES

Most of the variable sites observed for HLA class Ib genes were considered to be polymorphic, since 92.7% (*HLA-G*), 73.0% (*HLA-E*) and 86.8% (*HLA-F*) exhibited frequencies higher than 1% (Table 42). Most of these variable sites have already been reported by: **i)** the *1000 Genomes phase 3* project (Auton et al. 2015); **ii)** the IPD-IMGT/HLA v3.31.0 database (Robinson et al. 2015) or **iii)** specific population studies (Castelli et al. 2015, 2017; Lima et al. 2016; Ramalho et al. 2017). Noteworthy, three singletons at the *HLA-G locus* at positions -90, +192 and +3204, four variable sites at the *HLA-E locus* at positions -1262, +758, +1994 and +3824 and six variable sites at the *HLA-F locus*, at positions -1558, -1207, -1114, +2804, +3061 and +3096 have never been previously reported.

All variable sites for *HLA-G* and *HLA-E* did fit the Hardy-Weinberg expectations (Table 42). In contrast, the 26 variable sites at *HLA-F* did not fit Hardy-Weinberg expectations, primarily occurring at the *F**distal promoter (positions upstream -1013) and in some introns (Table S3, appendix III) that usually presented low read coverage (Castelli et al. 2015, 2017; Lima et al. 2016), possibly due to the type of kit used for DNA library preparation (Nextera

Kit transposase) (Wang et al. 2011; Shiina et al. 2012). To overcome genotype errors during genotype inference, the *hla-mapper* and *vcfx* software (available at www.castelli-lab.net) were used (Castelli et al. 2015, 2017; Lima et al. 2016), and only 0.31% of missing alleles introduced by *vcfx* needed to be imputed by the PHASE program (Stephens et al. 2001; Stephens and Donnelly 2003). Most of the variable sites (24 out of 26) had been already described at *1000 Genomes phase 3* project (Table S3, appendix III), IPD-IMGT/HLA v3.31.0 database or at a specific population (Lima et al. 2016).

Considering the 58 variable sites identified for *HLA-G* IPD-IMGT/HLA region (-300 to +2838), 56 occurred with a MAF ≥ 1% (Table 42), and 53 (95%) of these polymorphisms have already been described in the IPD-IMGT/HLA v3.31.0 database. The three remaining variable sites at positions -256, +1452 and +1523 were observed in intron 3 or in regulatory regions, and two of them (at positions [+1452 (0.12%) and +1523 (0.02%)] have also been described in the *1000 Genomes phase 3* panel (Auton et al. 2015), and the variation site observed at -256 (0.2%) position has been previously reported (Castelli et al. 2017).

Ten polymorphic sites were observed for *HLA-E*, and five (50%) have been described in the IPD-IMGT/HLA region (-300 to +3522). The variable sites observed at positions -113, -104, +2269, +2626 and +3468 (Table 42) have also been described in the *1000 Genomes phase 3* panel, occurring in intronic or in regulatory regions (Table S2, appendix II).

Thirty-four (87.2%) of the polymorphic sites identified in *HLA-F* IPD-IMGT/HLA region were also described in IPD-IMGT/HLA v3.31.0 database (Table 42 and Table S3 in appendix III). The remaining five, at positions +205, +1378, +1946, +2804 and +3097 were not described in the IPD-IMGT/HLA database; however, they have also been described in the *1000 Genomes phase 3* panel (Table S3, appendix III).

Table 42. *HLA-G, -E and -F* variable sites detected by NGS in entire evaluated segment in Beninese Toffin population.

Variable sites detected in evaluated segment highlighting the segment tracked by IPD-IMGT/HLA v.3.31.0 database										
						New variable sites [d]				
Genes	In entire evaluated segment	% of polymorphic sites in entire genes	Variable sites already described at public database or by other studies [a]	**Polymorphic sites in region tracked by IMGT [b]**	**% polymorphic sites already described at IMGT/v.3.31.0 [c]**	In promoter	Transcribed region	In 3' and downstream UTR	Hardy-Weinberg p-Value ($\alpha = 0.01$)	Average reads per each 26 sites (*HLA-F*)
HLA-G	96 [e] (-1377 to +3275)	89 (92.7%)	93 (97%)	**56**	**95.00 (53)**	-90 [f]	+192 [f]	+3204 [f]	≥ 0.01	_
HLA-E	37 [e] (-2143 to +4420)	27 (73.0%)	33 (89.2%)	**10**	**50.00 (5)**	-1262	+758, +1994	+3824	≥ 0.01	_
HLA-F	68 [e] (-1709 to +3537)	59 (86.8%)	62 (91.2%)	**39**	**87.18 (34)**	-1558, -1207, -1114	+2804	+3061, +3096	26 variable sites p-HW< 0.01	1228

[a] The majority of these variable sites have already been reported either by the *1000 Genomes phase 3* project (Auton et al. 2015) or by the IPD-IMGT/HLA v3.31.0 database (Robinson et al. 2015) or by specific population studies (Castelli et al. 2015, 2017; Lima et al. 2016; Ramalho et al. 2017).

[b] IMGT coding region: *HLA-G* (-300 to 2838), *HLA-E* (-300 to +3522) and *HLA-F* (-300 to +3250), considering the Adenine of the first translated ATG as +1 (IPD-IMGT/HLA v.3.31.0 database).

[c] In brackets the number of variable sites with MAF >1% identified here that had also been included in IPD-IMGT/HLA v.3.31.0 database.

[d] Variable sites that were never described at any public database nor by specific studies.

[e] Many of these variable sites were also included in *1000 Genome phase 3* project (Table S1, S2 and S3, appendix I to III).

[f] Indicates a singleton, i.e, variable site detected in a single heterozygous individual from the population sample.

6.2. *HLA-G, -E* AND *-F* UPSTREAM REGULATORY REGION (URR) HAPLOTYPES

The promoter region of the *HLA-G* gene encompassed the positions -1 to -1377 (Table 6); i. e., the same region previously evaluated for the Beninese Tori sample (Gineau et al. 2015). Consistent with the Tori results, the promoter haplotype lineage PROMO-G010102a (33.2%) was the most frequent in Toffins, followed by PROMO-G0104a (25.8%), PROMO-G010101a (8.7%) and PROMO-G0103a (1.3%) (Table 6). Noteworthy, the PROMO-G0104a was frequent in Toffins but not so frequent in African and non-African populations (Tan et al. 2005; Castelli et al. 2011, 2014a, b, 2017; Gineau et al. 2015; Nilsson et al. 2016). As in the Tori ethnic group (Gineau et al. 2015), the *HLA-G* promoter region appeared to be highly conserved in Toffins, with a similar level of nucleotide diversity (0.0060 ± 0.0031) [Table 6 and ref. (Gineau et al. 2015)] within this gene region. This finding is consistent with the idea that a balancing selection is acting in this region [Tajima's *D* value (2.5522, *P* = 0.0087) (Table 6)], as also observed in many other populations (Ober et al. 2003; Tan et al. 2005; Castelli et al. 2011, 2017; Santos et al. 2013; Gineau et al. 2015).

The promoter region of *HLA-E* evaluated for Toffins was the same evaluated in a Brazilian sample (Ramalho et al. 2017), which was primarily composed by individuals of European ancestry (approximately 80%), a typical sample from the Southeastern Brazilian region. Contrary to the Brazilian population, many of the variable sites were polymorphic (with MAF ≥1%) and the number of promoter haplotypes exhibiting at least 1% of frequency was higher in Toffins when compared to Brazilians (Table 10). Several reasons may account for these differences: **i)** the difference in nucleotide variability within promoter region between both populations [0.0013 ± 0.0008 (Toffins) *versus* 0.0010 ± 0.0006 (Brazilians, according to Ramalho et al. 2017), and **ii)** the negative Tajima's *D* value (-0.1764, *P* = 0.5000, Table 10), reflecting a possible assumption of non-described new alleles or population expansion (demography), and **iii)** some variants detected in Toffins with higher frequency were less frequent in Brazilians, for instance the -104G exhibited a frequency of 2.38% in Brazilians and 12% in Toffins of the present study and in other African samples from Guinea-Conakry and Burkina Faso (11.39%) (Castelli et al. 2015), **iv)** while in Brazilians the variants with lower frequency were between -440 to -1 (except position −104) (Veiga-Castelli et al. 2016; Ramalho et al. 2017), the variable sites we identified in the same region were more frequent in Toffins (Table S2, appendix II). Despite of these differences in frequencies of variable sites, the pattern of variability along the entire *HLA-E* gene between Brazilians and Toffins was quite similar. Likewise in Brazilians, the promoter region variability (Table 10) is

greater than the coding (Table 11) and 3' UTR segments (Table 12), and the 3' UTR segment is more variable than the coding region. Interestingly, despite this similar pattern of nucleotide variability between the *HLA-E* segments, the gene region (coding and regulatory segments) variability (shown as nucleotide variability) we observed was always higher than that observed for Brazilians (Ramalho et al. 2017). This variability may also explain the observation of the two new haplotypes in Toffins (*E*-Promo-38 and *E*-Promo-39) (Table 10). Similarly to the Brazilian population, the frequent promoter variants identified for Toffins were also concentrated upstream position −1389 (Veiga-Castelli et al. 2016; Ramalho et al. 2017). In addition, the most frequent haplotype detected for Toffins was also *E*-Promo-01 (37.67%). Although promoter region variability occurs among distinct populations, additional functional studies are needed to further evaluate the influence of promoter variability on *HLA-E* function and expression.

As mentioned at the Results section, two regions within the *HLA-F* promoter were evaluated: the distal (from -1709 to -1013) and proximal (from -547 to -101) ones (Table 14), as previously described (Lima et al. 2016). The F-distal promoter had already been evaluated in Asian, African American, and Caucasian population samples (Pyo et al. 2006). Since we observed a strong LD between *F**Distal and *F**proximal promoters **(Figure 11)**, we also observed the same haplotype frequency distribution pattern for the distal and proximal segments (i.e., four most frequent haplotypes with minimum frequency of 11%) (Table 14). The *F*-proximal promoter is more conserved than the other *F*-distal promoter and of the coding region, as shown by nucleotide diversity within these regions (Tables 14 and 15). This same pattern of *F*-promoter conservation may be characteristic of this gene, as it was also reported by other authors (Auton et al. 2015; Lima et al. 2016), and may reflect the probable influence of the *F*-proximal promoter on the *HLA-F* gene expression (Pyo et al. 2006, Lima et al. 2016).

6.3. *HLA-G, E AND -F* IPD-IMGT/HLA REGION HAPLOTYPES

Considering the IPD-IMGT/HLA v3.31.0 database, the *HLA-G* coding region evaluated here (positions -300 to +2838) generated 19 different haplotypes (Table 7), coding only four different (full length or truncated) proteins [G*01:01 (51.02%), G*01:03 (10.40 %), G*01:04 (27.19%) and G*01:05N (11.41%)]. Taken into account the well-recognized role of HLA-G as an immune checkpoint molecule, which is expressed in the placenta counteracting the mother immune system (Kovats et al. 1990; Rouas-Freiss et al. 1997; Allan et al. 1999; Fournel et al. 2000; Lila et al. 2001; Petroff et al. 2002; Ishitani et al. 2003, 2006; Morandi et al. 2010; Huang et al. 2010; Baudhuin et al. 2013; Naji et al. 2014; Sampangi et al. 2015; Grange et al. 2015; Nazari and Farjadian 2016; Bian et al. 2016; Eguchi et al. 2016; Kï¿½stlin et al. 2017), a low protein variability would be expected, reinforcing the hypothesis of a conserved coding region, due to the action of balancing selection in this region, as shown by the high Tajima's *D* (2.8, $P = 0.0034$). In addition, among the four most frequent haplotypes identified: *G**01:01:02:01 (19.80%), *G**01:04:04 (19.80), *G**01:05N (11.41%) and *G**01:03:01:02 (10.40%) (Table 7), two of them (*G**01:04:04 and *G**01:05N) exhibited frequencies relatively higher than those observed worldwide (Castelli et al. 2014a). Although these two alleles are observed at low frequency in Occidental regions of Europe and Asia, they exhibited higher frequencies in regions on natural pressure (endemic for parasite infection), including Africa (Carlini et al. 2013; Castelli et al. 2014a) and the Brazilian Amazon (Mendes-Junior et al. 2013), where the *HLA-G* coding region was shown to depart from neutrality (Aldrich et al. 2002; Mendes-Junior et al. 2013).

We identified 15 different alleles for *HLA-E* that encoded the two most frequent and worldwide distributed E*01:01 and E*01:03 proteins (Arnaiz-Villena et al. 2007; Liu et al. 2012; Veiga-Castelli et al. 2012; Carvalho dos Santos et al. 2013; Felício et al. 2014; Ramalho et al. 2017). Although the immunomodulatory properties of HLA-E have not been so extensively studied compared to HLA-G, the conservation of the HLA-E structure is expected (Sullivan et al. 2008; Pietra et al. 2010). In Toffins, the frequencies of *HLA-E**01:01 (53.66%) and *E**01:03 (40.67%) are closely similar to several worldwide populations (Arnaiz-Villena et al. 2007; Liu et al. 2012; Veiga-Castelli et al. 2012; Carvalho dos Santos et al. 2013; Felício et al. 2014; Ramalho et al. 2017). The sum of the frequency of these alleles in worldwide population samples may reach 94-95% of *HLA-E* encoded-molecules, a finding that is similar to that observed for Toffins (Table 11). The gene conservation may justify the few (6 out of 15) new alleles identified even in African population, as observed in the present study, frequently displaying intronic or synonymous mutations (Auton et al. 2015; Castelli et

al. 2015; Ramalho et al. 2017). The *HLA-E* coding region is the most conserved amongst the HLA Ib genes, exhibiting much lower nucleotide diversity (0.0005 ± 0.0003) than *HLA-F* (0.0054 ± 0.0027) and *HLA-G* (0.0058 ± 0.0029) (Tables 7, 11 and 15). Considering that nucleotide diversity observed at the coding region (Table 11) is lower than those observed at the promoter (Table 10) and 3' UTR in Toffins (Table 12) as well as in Brazilians (Ramalho et al. 2017), we may hypothesize that the regulation of *HLA-E* expression was more relevant than the structure of HLA-E.

As reported for Brazilians, we did not observe within *HLA-F* some frequent polymorphisms described in IPD-IMGT/HLA v.3.31.0 database, such as the G deletion at position +1943. The 29 coding haplotypes (Table 15) encoded only three different proteins: F*01:01, F*01:02 and F*01:03 (Table 15). This gene conservation can be observed in the IPD-IMGT/HLA v.3.31.0 database and was also confirmed by other *HLA-F* studies (Pyo et al. 2006; Lima et al. 2016) and by *1000 Genomes phase 3* project (Auton et al. 2015). The protein F*01:01 was observed in 77.63% of the individuals, accompanied by F*01:03 (19.41%) and F*01:02 (2.96%) (Table 15). These frequencies are closely similar to those reported in several studies, which mainly detected the two proteins (F*01:01 and F*01:03), for which F*01:01 is usually detected at higher frequency, and F*01:02, when detected, was always at low frequency (Pyo et al. 2006; Moscoso et al. 2007; Zhang et al. 2012; Pan et al. 2013; Manvailer et al. 2014; Lima et al. 2016). The presence of dinucleotide repeats at position +3097, as detected here (Table S3 in appendix III and Table 15) and in other studies (Auton et al. 2015; Lima et al. 2016), confirmed the existence of microsatellites and consequently the existence of new haplotypes harboring these microsatellites. However, a proper nomenclature must be assigned by the WHO Nomenclature Committee for Factors of the HLA System.

6.4. *HLA-G* AND *-E* 3’ UTR HAPLOTYPES

The *HLA-G* 3’ UTR evaluated here revealed well-documented haplotypes (Tan et al. 2005; Hviid et al. 2006; Castelli et al. 2010, 2011, 2014a, 2017; Consiglio et al. 2011; Lucena-Silva et al. 2012, 2013; Sizzano et al. 2012; Veit et al. 2012, 2014; Carlini et al. 2013; Martelli-Palomino et al. 2013; Santos et al. 2013; Courtin et al. 2013; Garcia et al. 2013; Sabbagh et al. 2014; Catamo et al. 2014, 2015; Garziera et al. 2015; Gineau et al. 2015; Porto et al. 2015; Nilsson et al. 2016; Zambra et al. 2016; de Albuquerque et al. 2016) (Table 8). The six most frequent haplotypes detected in this series (*HLA-G*-3UTR-01 to -06) were also the most worldwide represented. However, *HLA-G*-3UTR-17 and -18 were detected at 2.01%

and 1.01%, respectively (Table 8), but were absent in the Beninese Tori-Bossito (Tori ethnic group) sample (Sabbagh et al. 2014). The first possible explanation for this observation may be related to the Sanger sequencing procedure that was used by Sabbagh et al, possibly detecting only the most frequent or well-documented polymorphisms, which would be reflected in the lower variability in the Tori ethnic group (nucleotide diversity = 0.007) (Sabbagh et al. 2014) than in the Toffin population (nucleotide diversity = 0.027; Table 8). The second possible explanation is that the 3' UTR segment may be on balancing selection in the Tori ethnic group (Tajima's D = 1.84, P = 0.011) (Sabbagh et al. 2014), but not in the Toffin population (Tajima's D = 0.0032, P = 0.4292). This could justify the conservation of the 3' UTR segment in the Tori population, guaranteeing the presence or conservation of existing polymorphisms, perhaps, protecting Tori people against parasite infection pressure in this region (Djenontin et al. 2010; Le Port et al. 2012), where some malaria vectors are also resistant to insecticide (Djenontin et al. 2010).

Of the seven *HLA-E* 3' UTR haplotypes (Table 12), the first five and most frequent ones (*HLA-E*-3UTR-1 to -5) had already been detected in African samples from Guinea-Conakry and Burkina Faso (Castelli et al. 2015), whereas the last two ones (*HLA-E*-3UTR-13 and -14) were shared with the Brazilian population sample (Ramalho et al. 2017). The *HLA-E* 3' UTR segment is also very conserved (Felício et al. 2014; Castelli et al. 2015; Ramalho et al. 2017), showing the lowest nucleotide diversity when compared to the promoter and 3' UTR of other HLA Ib genes (Tables 6, 8, 12, and 14). This conservation at 3' UTR segment may reflect the low variability of HLA-E expression by post-transcriptional mechanisms, modulated by microRNA binding (Bartel 2004, 2009) and mRNA secondary/tertiary structure (Chen et al. 2006). Noteworthy, the presence of the highly polymorphic variable site at position +3777 (Table 12, Table S2), coinciding with the poly Adenine tail of a known Alu element [SINE (short interspersed nuclear element)] insertion, is not targeted by any microRNA (Ramalho et al. 2017).

6.5. *HLA-G, -E AND -F* EXTENDED HAPLOTYPES

We observed a strong LD between the *HLA-G* 5' UTR/coding/3' UTR segments (Table 9), and even the new haplotypes identified at the promoter or coding region of the gene do support this relationship. This strong association is relatively conserved and may be an intrinsic signature of the gene, since the same association was also observed (Castelli et al. 2010, 2011, 2014a, 2017; Donadi et al. 2011; Lucena-Silva et al. 2012; Santos et al. 2013; Nilsson et al. 2016). The balancing selection acting on the entire gene segment, as shown by

the significant Tajima's *D* (Table 9), reinforces this idea of conservation. In addition, no recombination hotspot exists within the *HLA-G* segment, at least from the promoter to a polymorphism 20Kb downstream the gene, which may prevent the *HLA-G* extended segment from high diversity (Kulski et al. 2001; Santos et al. 2013). Interestingly, Toffins did not present the *G**01:01:03/UTR-07 combination (Table 7, Table 8), which is frequently observed in Europe and Asia, but absent in African countries (Santos et al. 2013; Castelli et al. 2014a), suggesting that the haplotypes observed here and around the world may be quite old.

In agreement to previous studies (Felício et al. 2014; Castelli et al. 2015; Ramalho et al. 2017), *HLA-E* exhibited very low association between the promoter/coding/3' UTR haplotypes (Table 13). This may be explained by the weak LD between these gene segments (Felício et al. 2014; Castelli et al. 2015; Ramalho et al. 2017) **(Figure 10)**. However, some associations detected here followed the same pattern already described (Ramalho et al. 2017) (Table 13). Moreover, the number of haplotypes or new haplotypes we found may result from the high recombination rate in this extended segment, probably due to the presence of various *Alu* elements observed along the *HLA-E* gene (Kent et al. 2002; Deininger 2011; Ramalho et al. 2017).

As only one variable site was detected at *HLA-F* 3' UTR at position +3061 (Table S3), only two 3' UTR haplotypes were detected. The variable sites they carried, A (99.34%) for *HLA-F**3UTR-A and T (0.66%) for *HLA-F**3UTR-B, were not detected at IPD-IMGT/HLA-A nor at *1000 Genomes phase 3* databases. However, they fit the Hardy-Weinberg expectations (Table S3). The *HLA-F**3UTR-B was detected in two unrelated individuals and it was also associated to the same extended haplotype (Table 16). Likewise, the two alleles (*F**01:02$^{(1193C,\ 1943G)}$ and *F**01:02$^{(1193C,\ 1943G,\ new1946C)}$) detected here and absent in Brazilians were also associated to the same extended haplotype (Table 16).

6.6. BENIN TOFFIN ETHNIC GROUP AS A RICH REPOSITORY FOR GENETIC VARIATION

Some *HLA-G, -E and -F* coding alleles (Table 7, 11 and 15), which were shared with the Brazilian population (Lima et al. 2016; Castelli et al. 2017; Ramalho et al. 2017), were more frequent in Toffins than in Brazilians. In the present study, both the most frequent [*E**01:03:05$^{-compatible}$ (11.67%), *E**01:05$^{-compatible}$ (5.67%) and *E**01:01:01:01$^{(new3468C)}$ (4%)] and the less frequent [*E**01:01:01:01$^{(new1994T,\ new3468C)}$ (0.67%) and *E**01:03:01:01$^{(new1322A)}$ (0.67%)] alleles in the Toffin population (Table 11) were more frequent than in Brazilians (Ramalho et al. 2017), but exhibited highly similar frequencies when compared to other

African populations from Guinea-Conakry and Burkina Faso (Castelli et al. 2015). This observation is also true for *HLA-F* and *HLA-G* coding alleles that were shared with Brazilians (Table 7 and 15). Moreover, the difference in nucleotide diversity along the coding (Table 7, 11 and 15) and extended (Table 9, 13 and 16) segments for *HLA-G, -E and -F* disclosed differences between African and Brazilian populations. Overall, these findings corroborate the idea that African countries (including Benin) could be considered as a rich repository for genetic variation and that these alleles might have arisen in Africa before human dispersion.

6.7. PARTIAL CONCLUSION: PART I

This is the first study evaluating the complete structure of HLA Ib genes in a secluded African population, using massive parallel sequencing. Although known and new *HLA-G, -E* and *-F* coding alleles were described, these gene regions remained relatively conserved, since all the coding alleles converged to encode few molecules for each gene. *HLA-G* and *HLA-E* were more conserved at the coding region (Table 7 and 11) than at promoter (Tables 6 and 10) and at 3' UTR (Tables 8 and 12). *HLA-F* was more conserved at the regulatory [proximal promoter (Table 14) and 3' UTR (Table 16)] than at the coding (Table 15) and distal promoter (Table 14) regions; however, the coding region is still conserved. These observations support the immunomodulatory role of these genes, rather than antigen presentation, since an eventual high variability at the coding regions would lead to modifications at the ligand-binding sites on target cells, compromising their function not only in pregnancy and transplantation but also in diseases.

DISCUSSION: PART II

6.8. *HLA-G, -E AND -F* GENETIC ASSOCIATION STUDIES AND SUSCEPTIBILITY OR RESISTANCE TO *P. FALCIPARUM MALARIA*

Specific genetic epidemiology approaches can be used to identify genetic factors influencing host susceptibility/resistance to infectious diseases. Two common approaches are case-control studies and family-based association studies that have revealed many association/linkage between genes mainly coding proteins of immune system or involved in the physiology of host cells and malaria susceptibility/resistance (Garcia et al. 2010; Milet et al. 2010; Hernandez-Valladares et al. 2014; Marquet 2017). Many candidate genes have been located at 5q31-q33 and 6p21-p23 chromosomal regions and have been associated with the control of parasitemia in the context of both, mild and asymptomatic malaria. The region 5q31-q33 encompasses genes controlling the Th1/Th2 responses and the region 6p21-p23 is the MHC region (including HLA I, II and *TNF* gene in HLA-III) (Garcia et al. 1998, 2010; Milet et al. 2010; Driss et al. 2011; Hernandez-Valladares et al. 2014). The genes closer to *TNF* region showed significant association with uncomplicated malaria phenotype (Flori et al. 2003, 2005). Although the studies aforementioned revealed MHC *loci* (6p21-p23) as probable genetic factors that could influence the susceptibility/resistance to uncomplicated malaria, it still unknown the role of HLA Class Ib *loci* in the susceptibility/resistance to uncomplicated malaria.

6.9. *HLA-G* 14 BASE PAIR (14BP) INSERTION (UNDER DOMINANT MODEL) ASSOCIATED TO SUSCEPTIBILITY TO *P. FALCIPARUM* MALARIA IN BENINESE TOFFIN YOUNG CHILDREN

To our knowledge, the role of *HLA-E* and *-F* in malaria susceptibility or resistance has never been previously investigated, and only few human *HLA-G* studies have been reported in the context of malaria. Most of the studies evaluating the genetic association with *P. falciparum* malaria clinical phenotypes focused on the *HLA-G* 3' UTR segment (Sabbagh et al. 2013, 2018; Garcia et al. 2013). Thus, except the study conducted by Gineau et al. (2016) that evaluated the association of variable sites in the entire *HLA-G* (including 5' UR, coding and 3' UT regions) with human African trypanosomiasis (HAT) in Guineans (Gineau et al. 2016), no studies have addressed the influence of genetic variants of the entire *HLA-G* gene on *P. falciparum* malaria clinical outcomes. In our cohort, the *HLA-G* 14bp Indel was the unique polymorphism, among 105 polymorphisms analyzed in the entire *HLA-G, -E* and *-F* genes, associated with the susceptibility to *P. falciparum* infection and clinical phenotypes (Tables 18 to 22, $P < 0.01$; where it associated with the risk of *P. falciparum* infection,

symptomatic infections, the number of infection and the number of symptomatic infections episodes).

The 14 bp Indel, belonging to *HLA-G* 3' UTR, has been extensively studied, and numerous genetic associations were reported between this variant and disease mordidity or susceptibility (Wang et al. 2013; Zhang et al. 2014; Fan et al. 2014, 2017; Ge et al. 2014; Catamo et al. 2015; Kim et al. 2015; Lee et al. 2015; Li et al. 2015; Pabalan et al. 2015; Coelho et al. 2016; Shi et al. 2017; de Almeida et al. 2018). Functional studies reported an influence of 14 bp polymorphism on the post-transcriptional regulation of the gene (Tan et al. 2008; Veit and Chies 2009). The 14 bp Ins allele has been associated to mRNA stability and to differencial expression of mRNA and mRNA isoforms, in which the presence of 14bp (insertion) was associated to a significantly lower mRNA isoform expression compared to the absence of 14bp (deletion) (Hiby et al. 1999; O'Brien et al. 2001; Hviid et al. 2003; Rousseau et al. 2003). The presence of the 14 bp (insertion) is associated with a deletion of 92 bp from the mature *HLA-G* mRNA transcript (Hiby et al. 1999; Hviid et al. 2003), leading to a smaller and more stable mRNA (Rousseau et al. 2003), which is the preferred target for miRNA compared to the unstable one. However, at protein level, the 14 bp Ins/Ins, Ins/Del and Del/Del genotypes leaded to lower, and higher sHLA-G expression, respectively (Chen et al. 2008; Martelli-Palomino et al. 2013). Considering that sHLA-G downregulates the immune system by interacting with immune cells through immunoglobulin-like transcript (ILT)-2 and -4 receptors (Donadi et al. 2011; Carosella et al. 2015; Sabbagh et al. 2018), sHLA-G desactivates all immune cells implicated on the host defense against parasites and specifically against *P. falciparum* (Sabbagh et al. 2018).

The observation that the 14 bp Ins (in dominant model) was associated with susceptibility to *P. falciparum* contrasts with previous evidences regarding the relationship between 14 bp Ins allele and the sHLA-G expression, as detailed above. As in other parasitic adult disorders, increased production of sHLA-G has been associated with susceptibility instead of protection (Sabbagh et al. 2018). Although the previous functional studies regarding the influence of 14 bp polymorphism on *HLA-G* expression profile were implemented with the most sensible and appropriate methodologies, the mechanism involved on this regulation may be more complex than expected, and other epigenetic mechanisms may be involved (Moreau et al. 2003, 2009a; Mouillot et al. 2005; Polakova et al. 2007; Holling et al. 2009). In addition, the observed associations between 14 bp polymorphism and disease susceptibility were evaluated at both allelic and haplotypic level and the findings obtained with haplotypic analyses seem more conclusive than those observed with allelic analyses (Courtin et al. 2013; Garcia et al. 2013;

Dias et al. 2015; de Almeida et al. 2018; Sabbagh et al. 2018). In other words, various polymorphisms (at the *HLA-G* 3' UT, coding or 5' UR regions, or even at other genes, closer to *HLA-G,* like *HLA-E* and *–F*) in LD with 14 bp polymorphism could also influence disease association (de Almeida et al. 2018). Moreover, a previous study conducted on other Beninese ethnic group (the Tori) revealed the existence of a balancing selection acting at the 3' UTR segment of *HLA-G* (Tajima's D = 1.84, P = 0.011) (Sabbagh et al. 2014), reinforcing the importance of the 3' UTR segment on the gene regulation, where any modification or variation in this gene segment may affect the gene expression profile, leading to disease susceptibility/resistance. Finally, it should be stressed that most studies regarding the association of the 14 bp Ins/Del with sHLA-G levels were performed in adult healthy subjects (Chen et al. 2008; Martelli-Palomino et al. 2013), and little is known regarding the influence of this variation site in healthy or ill children, who may exhibit a differential expression of microRNAs at each condition. Further studies should be conducted to evaluate the levels of sHLA-G in healthy and ill children as well as the miRNA profiles in each condition.

6.10. *HLA-G (-1155 A* AND *+755 A)* POLYMORPHISMS (UNDER RECESSIVE MODEL) ASSOCIATED TO PROTECTION AGAINST *P. FALCIPARUM* MALARIA IN BENINESE TOFFIN YOUNG CHILDREN

The presence of the A allele at position +755 (exon3) leads to a substitution of a Cytosine (C) to Adenine (A) at codon 110 and, consequently, to a substitution of Leucine by Isoleucine (IPD-IMGT/HLA database (release 3.33.0, 2018-07-11) (Robinson et al. 2015)). This missense variant is carried by the *G**01:04:04/*G**01:04:01 coding alleles and could affect the regular protein, i.e., the substitution of G*01:04 protein by a truncated one. This may be explain the protective effect confered by this allele to children that carry this allele (Tables 23 to 26). Indeed, a truncated molecule would have less or no affinity for ILT-2 and -4 receptors present on host immune cells, leading these immune cells to escape from the HLA-G immunomodulatory activities (Donadi et al. 2011; Carosella et al. 2015; Sabbagh et al. 2018), and, consequently, conferring protection to the *+755 A/A* genotype carriers.

The region between positions -56 and -1377 (including -1155) has previously reported as segment that harbored important complex transcription modules, such as the enhancer A, the interferon-stimulated response element (ISRE), a heat shock element (HRE), and the locus control region (LCR), among others (Moreau et al. 2009b), suggesting that any nucleotide variation in this segment could lead to alteration in the regulation of gene expression and, consequently, to disease susceptibility or resistance (Castelli et al. 2014b; Cordeiro et al.

2017). However, it remains unknown the mechanism underlying this process (gene expression), since few functional studies have addressed the influence of these polymorphisms (encompassing the positions -1179 to -964, for e.g) on *HLA-G* expression (Castelli et al. 2014b). For instance, the *-1179 G>A* (Agrawal et al. 2015), *-1305 G>A* (Kim et al. 2011), *-964 G>A* (Nicolae et al. 2005; Misra et al. 2013), *-725 C>G* (Ober et al. 2003; Wiśniewski et al. 2010; Misra et al. 2013; Agrawal et al. 2015) and *-486 A>C* (Misra et al. 2013; Agrawal et al. 2015) variable sites were studied regarding their association with diseases. Of them, the *-725 C>G* was the unique that was functionally known to upregulate the HLA-G expression (Ober et al. 2006; Jassem et al. 2012). Moreover, the nucleotide variation within the segment encompassing the positions -1305 to -810, including -1155, which coincides with the LCR module, was recently shown to influence the cervical cancer outcome in patients with cervical intraepithelial neoplasias (Cordeiro et al. 2017). Thus, further studies are needed to unveil the functional activity of the *-1155 G>A* polymorphism.

Considering that the *+755 C>A* and the *-1155 G>A* are in complete LD, these variable sites may: **i**) act in synergy to confer protection to children carrying these polymorphisms, **ii)** be conserved due to balancing selection acting in the promoter and conding *HLA-G* regions, as shown by the significant Tadjima's D values (Tables 6 and 7), to protect individuals from infectious diseases (e.g, *P. falciparum*) in endemic regions, like Benin, West of Africa. The high frequency of the *G**01:04:04 (19.80%) and *G**01:05N (11.41%) alleles observed in our population (Table 7) corroborates this hypothesis. These frequencies were relatively higher than that observed worldwide (Castelli et al. 2014a) and may reflect the non-neutral natural selection acting at the *HLA-G* coding region (Aldrich et al. 2002; Mendes-Junior et al. 2013), due to parasite pressures in endemic regions, including African regions (Carlini et al. 2013; Castelli et al. 2014a) or in Amerindian populations from the Brazilian Amazon (Mendes-Junior et al. 2013).

6.11. E.01.03.05-COMPATIBLE HAPLOTYPE IN SYNERGY WITH *HLA-E -1988 C* ALLELE (UNDER ADDITIVE MODEL) PROTECTED BENINESE TOFFIN YOUNG CHILDREN AGAINST *P. FALCIPARUM* MALARIA

The *E*.01.03 allele group was associated with protection (at $P < 0.01$ and $Pc < 0.05$) against almost all phenotypes we analyzed, i.e., *P. falciparum* infection (all infections and symptomatic infection), parasite density and occurrences of infections (number of all infections and number of symptomatic infections). Since this haplotype was associated with various *P. falciparum* malaria phenotypes (Tables 27 to 31), we decided to identify the

specific haplotype that could be associated with this protection, and the ***E*.01.03.05-compatible** was our target. Interestingly, the results obtained (Table 32 to 35) confirmed again the protective effect of this haplotype against *P. falciparum* malaria, at the level of: **i)** infection ($P < 0.01$, $Pc = 0.0312$**,** Table 32), **ii)** *P. falciparum* symptomatic infection (Table 33, not significant at $P < 0.05$), and **iii)** number of malaria episodes (Table 34 and 35, not significant at $P < 0.05$).

The *E*.01.03.05-compatible haplotype carries the *-104 G* and *+1625 C* alleles (in complete LD between each other), which also protected from *P. falciparum* malaria infections at $P < 0.05$ (Table 23). Considering extended haplotypes, the same coding haplotype was associated with *E*-Promo-20 and *HLA-E*-3UTR-1 (Table 13). The *HLA-E* promoter allele (*-1988 C, rs17875360),* carried by the *E*-Promo-20, also conferred strong protection against '*P. falciparum* infections, under the additive model (Table 23, $P < 0.01$). These SNPs could lead to synergic protective effect, even though no functional studies evaluated the influences of them on *HLA-E* expression pattern.

The *HLA-E*-3UTR-1 haplotype is formed by variable sites ranging from +3468 to +4420 (including *+3777 A* allele). The miR-376a(e) may target two different microRNA-binding sites, the first at positions +3402 and +3417 (upstream the position +3468), and the second at positions +4553 to +4573 (downstream the position +4420) (Nachmani et al. 2014). Therefore, no microRNA binding sites has ben reported in this 3' UTR segment that constitute the *HLA-E*-3UTR-1 haplotype. Moreover, the *+3777 A* allele is the unique alternative allele carried by this haplotype (since the remaining variable sites were present with their reference alleles: +3468A/+3634G/**+3777A**/+3778A/+3824T/+4297G/+4420C) (Table 12). The *+3777 A* allele is also the unique and most frequent variable site (MAF = 88%, Table 12) in our population that coincides with the poly Adenine tail of a known Alu element (Ramalho et al. 2017), which theoretically may not bind to any microRNA. Notwithstanding, no functional post-transcriptional regulation study is availbale to justify this hypothesis.

Although a synergic protection against *P. falciparum* malaria outcome was conferred by the *-1988 C* allele (in promoter), the *-104 G* and *+1625 C* alleles (in E.01.03.05-compatible coding haplotype) and the *+3777 A* allele (in 3' UTR), the lack of functional studies prevent the understanding of the mechanism by which the *HLA-E* regulatory regions could influence gene expression and the disease association. Despite of these comments, the *+3777 A* allele would have an influence on the *HLA-E* post-trancriptional regulation, since another 3' UTR

haplotype (the *HLA-E*-3UTR-3) that carried this variation site, also protected against ‘’*P. falciparum* symptomatic infection’’ ($P < 0.01$, $Pc > 0.05$, Table 28).

6.12. *HLA-G, -E* AND *-F* POLYMORPHISMS AND HAPLOTYPES ASSOCIATED TO SUSCEPTIBILITY OR PROTECTION WITH EITHER *S. HAEMATOBIUM* INFECTION OR *P. FALCIPARUM-S. HAEMATOBIUM* CO-INFECTION IN BENINESE TOFFIN YOUNG CHILDREN

No *HLA-E, -F* or *-G* polymorphisms were associated with susceptibility or to protection against *S. haematobium* disease nor with *P. falciparum-S. haematobium* co-infection, at the $P < 0.01$ significance level (Tables 36, 37 and 39). In addition, no association was detected between infection with *S. haematobium* and the risk to be infected by *P. falciparum* (Tables 41 A and B).

Considering the haplotypes, the *E*.Promo.2 conferred protection against *P. falciparum-S. haematobium* co-infection ($Pc = 0.0365$, OR = 0.041, CI95% 0.004 to 0.264, Table 40). To our knowledge, no studies have addressed the influence of *HLA-E* haplotypes (at regulatory regions or coding region) on *P. falciparum* and *S. haematobium* infections, as well as on any parasite diseases. The *E*.Promo.2 haplotype has a strong association with the *E*.01.03.encoded-allele group at the extended level (Table 13); i. e., all *E*.Promo.2 carriers also carried the *E*.01.03.encoded-allele. We previously observed that the *E*.01.03.encoded-allele group was associated with protection against *P. falciparum* malaria ($P < 0.01$ and $Pc < 0.05$, Tables 27 to 31). Therefore, this means that *E*.Promo.2 children infected with *S. haematobium*, would be protected from *P. falciparum* co-infection, because the child also inherits the *E*.01.03.encoded-allele, suggesting that the *E*.Promo.2 haplotype may influence the *HLA-E* expression at transcriptional level. Similar to other non-classical HLA class Ib molecules, HLA-E also has immunomodulatory properties in pregnancy-related disorders (Persson et al. 2017) and in the context of viral infection (Sullivan et al. 2008; Pietra et al. 2009, 2010). The HLA-E molecule expressed by the virus-infected cells can inhibit NK cell cytotoxicity, propitiating the proliferation of viral cells (Sullivan et al. 2008; Pietra et al. 2009, 2010). In this context, HCV-infected hepatic cells may induce HLA-E expression and NKG2A inhibitory receptors (Nattermann et al. 2005), propitiating the propagation of the viral infection. Considering that *P. falciparum* has a hepatic cycle, it is possible that the agent may alter HLA-E expression. Although the functional role of the *HLA-E*.Promo.2 haplotype has not been evaluated, a decreased *HLA-E* promoter response may produce a reduced number of trancripts and protein, contributing to the protection effect against *P. falciparum*

co-infection in *S. haematobium* infected children. Finally, this protective effect against *P. falciparum* infection corroborates almost all the previous studies regarding the protective effect of *S. haematobium* infection against *P. falciparum* malaria (Briand et al. 2005; Lyke et al. 2005, 2006, 2012a, b, 2018; Mutapi et al. 2007; Diallo et al. 2010; Doumbo et al. 2014; Degarege et al. 2016).

6.13. PARTIAL CONCLUSION: PART II

We evaluated for the first time, in the Beninese Toffin population, the association between the non-classical HLA class I genes (*HLA-G, -E* and *-F*) with two of the most alarming tropical infectious diseases (*P. falciparum* malaria and the urogenital bilharziasis caused by *S. haematobium* infection) observed in African populations. Evaluating the complete segment of each gene (5' and 3' regulatory regions and coding region including all introns), we observed the following associations: **i)** *HLA-G* 14 base pair (14bp) insertion (under dominant model) was associated with the susceptibility to *P. falciparum* malaria**, ii)** *HLA-G -1155 A* and *+755 A* polymorphisms (in complete LD between each other*)* (under recessive model) were associated with protection against *P. falciparum* malaria, **iii)** the *E*.01.03.05-compatible haplotype in synergy with *HLA-E -1988 C* allele (under additive model) was associated with the protection against *P. falciparum* malaria, and finally **iv)** the *E*.Promo.2 haplotype associated with the protection against *P. falciparum-S. haematobium* co-infection. Exception made to the 14 bp Ins that possess functional evaluation, additional studies are needed to identify the functional role of the genetic markers associated with these major pathogenic agents.

7. CONCLUSION

The Beninese Toffin population, like other African populations, can be considered as a genetic repository. The nucleotide diversity along the regulatory and extended segments of *HLA-G, -E* and *-F* disclosed differences between African and Brazilian populations; however, the *HLA-G, -E* and *-F* coding regions remained conserved, even in the Beninese Toffin ethnic group, reinforcing the balancing selection acting in the coding region of HLA Ib genes, imputing to these genes an immunomodulatory role rather than antigen presenting properties.

Although the most remarkable association with susceptibility to *P. falciparum* malaria in Toffins was observed for the *HLA-G* 14bp insertion, under a dominant model, other variable sites in linkage disequilibrium (promoter/coding *HLA-G -1155 A* and *+755 A*, and the *E*.01.03.05-compatible haplotype in synergy with *HLA-E -1988 C* allele) were associated with protection against *P. falciparum* malaria. Only the *E*.Promo.2 haplotype was associated with the protection against *P. falciparum-S. haematobium* co-infection. Since many variable sites were observed along the entire HLA class Ib genes, and since a strict correction was necessary to adjust the *P* values, it was not possible to detect striking differences between the profiles of immunomodulatory genes in the *P. falciparum-S. haematobium* co-infection. Further studies encomapsssing larger series of patients may unveil major similarities or differences.

8. BIBLIOGRAPHY

Bibliography

Abel L, Cot M, Mulder L, et al (1992) Segregation analysis detects a major gene controlling blood infection levels in human malaria. Am J Hum Genet 50:1308–17

Abel L, Demenais F, Prata a, et al (1991) Evidence for the segregation of a major gene in human susceptibility/resistance to infection by Schistosoma mansoni. Am J Hum Genet 48:959–970

Abou-El-Naga IF (2018) Towards elimination of schistosomiasis after 5000 years of endemicity in Egypt. Acta Trop 181:112–121. doi: 10.1016/j.actatropica.2018.02.005

Adegnika AA, Kremsner PG (2012) Epidemiology of malaria and helminth interaction: A review from 2001 to 2011

Agrawal D, Prakash S, Misra MK, et al (2015) Implication of HLA-G 5′ upstream regulatory region polymorphisms in idiopathic recurrent spontaneous abortions. Reprod Biomed Online 30:82–91. doi: 10.1016/j.rbmo.2014.09.015

Ajakaye OG, Olusi TA, Oniya MO (2016) Environmental factors and the risk of urinary schistosomiasis in Ile Oluji/Oke Igbo local government area of Ondo State. Parasite Epidemiol Control 1:98–104. doi: 10.1016/j.parepi.2016.03.006

Akaike H (1974) A new look at statistical model identification. IEEE Trans Automat Contr 19:716–723. doi: 10.1109/TAC.1974.1100705

Aldrich C, Wambebe C, Odama L, et al (2002) Linkage disequilibrium and age estimates of a deletion polymorphism (1597ΔC) in HLA-G suggest non-neutral evolution. Hum Immunol 63:405–412. doi: 10.1016/S0198-8859(02)00377-4

Allan DS, Colonna M, Lanier LL, et al (1999) Tetrameric complexes of human histocompatibility leukocyte antigen (HLA)-G bind to peripheral blood myelomonocytic cells. J Exp Med 189:1149–1156. doi: 10.1084/jem.189.7.1149

Alonso D, Muñoz J, Gascón J, et al (2006) Failure of standard treatment with praziquantel in two returned travellers with Schistosoma haematobium infection. Am J Trop Med Hyg 74:342–344

Andrade R, Souza-santos R (2013) Biological and environmental factors associated with risk of schistosomiasis mansoni transmission in Porto de Galinhas , Pernambuco State , Brazil Fatores biológicos e ambientais associados ao risco de transmissão da esquistossomose mansoni na localidade . 29:357–367. doi: 10.1590/S0102-311X2013000200022

Antonelli LRVR V, Leoratti FMSMS, Costa P a C a C, et al (2014) The CD14 + CD16 + Inflammatory Monocyte Subset Displays Increased Mitochondrial Activity and Effector Function During Acute Plasmodium vivax Malaria. PLoS Pathog 10:e1004393. doi: 10.1371/journal.ppat.1004393

Aponte JJ, Schellenberg D, Egan A, et al (2009) Efficacy and safety of intermittent preventive treatment with sulfadoxine-pyrimethamine for malaria in African infants: a pooled analysis of six randomised, placebo-controlled trials. Lancet 374:1533–1542. doi: 10.1016/S0140-6736(09)61258-7

Arnaiz-Villena A, Vargas-Alarcon G, Serrano-Vela JI, et al (2007) HLA-E polymorphism in Amerindians from Mexico (Mazatecans), Colombia (Wayu) and Chile (Mapuches): Evolution of MHC-E gene. In: Tissue Antigens. pp 132–135

Asseman C, Pancré V, Quatennens B, Auriault C (1996) Schistosoma mansoni-infected mice show augmented hepatic fibrosis and selective inhibition of liver cytokine production after treatment with anti-NK1.1 antibodies. Immunol Lett 54:11–20. doi: 10.1016/S0165-2478(96)02634-X

Autino B, Noris A, Russo R, Castelli F (2012) Epidemiology of malaria in endemic areas. Mediterr J Hematol Infect Dis 4:. doi: 10.4084/MJHID.2012.060

Auton A, Abecasis GR, Altshuler DM, et al (2015) A global reference for human genetic

variation. Nature 526:68–74. doi: 10.1038/nature15393
Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265. doi: 10.1093/bioinformatics/bth457
Bartel DP (2009) MicroRNAs: Target Recognition and Regulatory Functions. Cell 136:215–233
Bartel DP (2004) MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell 116:281–297
Baudhuin J, Migraine J, Faivre V, et al (2013) Exocytosis acts as a modulator of the ILT4-mediated inhibition of neutrophil functions. Proc Natl Acad Sci 110:17957–17962. doi: 10.1073/pnas.1221535110
Belachew EB (2018) Review Article Immune Response and Evasion Mechanisms of Plasmodium falciparum Parasites. 2018:. doi: 10.1155/2018/6529681
Benjamin DJ, Berger JO, Johannesson M, et al (2018) Redefine statistical significance. Nat Hum Behav 2:6–10. doi: 10.1038/s41562-017-0189-z
Bennett JE (John E, Dolin R, Blaser MJ, Douglas RG (Robert G (2015) Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, 8th edition
Bereczky S, Dolo A, Maiga B, et al (2006) Spleen enlargement and genetic diversity of Plasmodium falciparum infection in two ethnic groups with different malaria susceptibility in Mali, West Africa. Trans R Soc Trop Med Hyg 100:248–257. doi: 10.1016/j.trstmh.2005.03.011
Bergquist R, Whittaker M (2012) Control of neglected tropical diseases in asia pacific: Implications for health information priorities. Infect Dis Poverty 1:1–4. doi: 10.1186/2049-9957-1-3
Bian X, Si Y, Zhang M, et al (2016) Down-expression of miR-152 lead to impaired anti-tumor effect of NK via upregulation of HLA-G. Tumor Biol 37:3749–3756. doi: 10.1007/s13277-015-3669-7
Blanchard TJ (2004) Schistosomiasis. Travel Med Infect Dis 2:5–11. doi: 10.1016/j.tmaid.2004.02.011
Boissier J, Grech-Angelini S, Webster BL, et al (2016) Outbreak of urogenital schistosomiasis in Corsica (France): an epidemiological case study. Lancet Infect Dis 16:971–979. doi: 10.1016/S1473-3099(16)00175-4
Boko PM, Ibikounle M, Onzo-Aboki A, et al (2016) Schistosomiasisand soil transmitted helminths distribution in Benin: A baseline prevalence survey in 30 districts. PLoS One 11:1–17. doi: 10.1371/journal.pone.0162798
Borrego F, Ulbrecht M, Weiss EH, et al (1998) Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. J Exp Med 187:813–8. doi: 10.1084/jem.187.5.813
Bottieau E, Clerinx J, de Vega MR, et al (2006) Imported Katayama fever: Clinical and biological features at presentation and during treatment. J Infect 52:339–345. doi: 10.1016/j.jinf.2005.07.022
Braud VM, Allan DS, McMichael AJ (1999) Functions of nonclassical MHC and non-MHC-encoded class I molecules. Curr. Opin. Immunol. 11:100–108
Briand V, Watier L, LE Hesran J-Y, et al (2005) Coinfection with Plasmodium falciparum and schistosoma haematobium: protective effect of schistosomiasis on malaria in senegalese children? Am J Trop Med Hyg 72:702–707
Burke ML, Jones MK, Gobert GN, et al (2009) Immunopathogenesis of human schistosomiasis. Parasite Immunol. 31:163–176
Carlini F, Traore K, Cherouat N, et al (2013) HLA-G UTR haplotype conservation in the Malian population: Association with soluble HLA-G. PLoS One 8:. doi:

10.1371/journal.pone.0082517
Carosella ED, Rouas-Freiss N, Roux DT Le, et al (2015) HLA-G. An Immune Checkpoint Molecule., 1st edn. Elsevier Inc.
Carvalho dos Santos L, Tureck LV, Wowk PF, et al (2013) HLA-E polymorphisms in an Afro-descendant Southern Brazilian population. Hum Immunol 74:199–202. doi: 10.1016/j.humimm.2012.10.020
Castelli EC, Gerasimou P, Paz MA, et al (2017) HLA-G variability and haplotypes detected by massively parallel sequencing procedures in the geographicaly distinct population samples of Brazil and Cyprus. Mol Immunol 83:115–126. doi: 10.1016/j.molimm.2017.01.020
Castelli EC, Mendes-Junior CT, Deghaide NHS, et al (2010) The genetic structure of 3′untranslated region of the HLA-G gene: polymorphisms and haplotypes. Genes Immun 11:134–141. doi: 10.1038/gene.2009.74
Castelli EC, Mendes-Junior CT, Sabbagh A, et al (2015) HLA-E coding and 3' untranslated region variability determined by next-generation sequencing in two West-African population samples. Hum Immunol 76:945–953. doi: 10.1016/j.humimm.2015.06.016
Castelli EC, Mendes-Junior CT, Veiga-Castelli LC, et al (2011) A comprehensive study of polymorphic sites along the HLA-G gene: Implication for gene regulation and evolution. Mol Biol Evol 28:3069–3086. doi: 10.1093/molbev/msr138
Castelli EC, Ramalho J, Porto IOP, et al (2014a) Insights into HLA-G genetics provided by worldwide haplotype diversity. Front Immunol 5:. doi: 10.3389/fimmu.2014.00476
Castelli EC, Veiga-Castelli LC, Yaghi L, et al (2014b) Transcriptional and posttranscriptional regulations of the HLA-G gene. J Immunol Res 2014:. doi: 10.1155/2014/734068
Castro GA (1996) Helminths: Structure, Classification, Growth, and Development
Catamo E, Addobbati C, Segat L, et al (2014) HLA-G gene polymorphisms associated with susceptibility to rheumatoid arthritis disease and its severity in Brazilian patients. Tissue Antigens 84:308–315. doi: 10.1111/tan.12396
Catamo E, Addobbati C, Segat L, et al (2015) Comprehensive analysis of polymorphisms in the HLA-G 5??? upstream regulatory and 3??? untranslated regions in Brazilian patients with systemic lupus erythematosus. Tissue Antigens 85:458–465. doi: 10.1111/tan.12545
CDC (2017) Parasites - Neglected Parasitic Infections (NPIs). Centers Dis Control Prev 1
Chen JM, Férec C, Cooper DN (2006) A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes II: The importance of mRNA secondary structure in assessing the functionality of 3′ UTR variants. Hum. Genet. 120:301–333
Chen W, Flynn EA, Shreefter MJ, Blagg NA (2012) Schistosomiasis: An unusual finding of the cervix. Obstet Gynecol 119:472–475. doi: 10.1097/AOG.0b013e31822da6a4
Chen XY, Yan WH, Lin A, et al (2008) The 14 bp deletion polymorphisms in HLA-G gene play an important role in the expression of soluble HLA-G in plasma. Tissue Antigens 72:335–341. doi: 10.1111/j.1399-0039.2008.01107.x
Chitsulo L, Engels D, Montresor A, Savioli L (2000) The global status of schistosomiasis and its control. Acta Trop 77:41–51. doi: 10.1016/S0001-706X(00)00122-4
Clerinx J, Van Gompel A (2011) Schistosomiasis in travellers and migrants. Travel Med. Infect. Dis. 9:6–24
Coelho AVC, Moura RR, Crovella S, Celsi F (2016) HLA-G genetic variants and hepatocellular carcinoma: a meta-analysis. Genet Mol Res. doi: 10.4238/gmr.15038263
Collaborators G 2016 C of D (2017) Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 390:1151–1210. doi: 10.1016/S0140-6736(17)32152-9
Colley DG, Bustinduy AL, Secor WE, King CH (2014) Human schistosomiasis. Lancet

383:2253–2264. doi: 10.1016/S0140-6736(13)61949-2
Consiglio CR, Veit TD, Monticielo OA, et al (2011) Association of the HLA-G gene +3142C>G polymorphism with systemic lupus erythematosus. Tissue Antigens 77:424–430. doi: 10.1111/j.1399-0039.2011.01635.x
Cooper AJR, Dholakia S, Holland C V, Friend PJ (2017) Review Helminths in organ transplantation. Lancet Infect Dis 17:e166–e176. doi: 10.1016/S1473-3099(16)30533-3
Cordeiro JC, da Silva JS, Roxo VS, da Graça Bicalho M (2017) A pilot study on Hla-G locus control region haplotypes and cervical intraepithelial neoplasias. Hum Immunol 78:281–286. doi: 10.1016/j.humimm.2016.12.004
Courtin D, Djilali-Saïah A, Milet J, et al (2011) Schistosoma haematobium infection affects Plasmodium falciparum-specific IgG responses associated with protection against malaria. Parasite Immunol 33:124–131. doi: 10.1111/j.1365-3024.2010.01267.x
Courtin D, Milet J, Sabbagh A, et al (2013) HLA-G 3' UTR-2 haplotype is associated with Human African trypanosomiasis susceptibility. Infect Genet Evol 17:1–7. doi: 10.1016/j.meegid.2013.03.004
Courtin D, Oesterholt M, Huismans H, et al (2009) The quantity and quality of African children's IgG responses to merozoite surface antigens reflect protection against Plasmodium falciparum malaria. PLoS One 4:e7590. doi: 10.1371/journal.pone.0007590
Cox FE (2010) History of the discovery of the malaria parasites and their vectors. http://www.ncbi.nlm.nih.gov/pubmed/20205846. Parasit Vectors 3:5. doi: 10.1186/1756-3305-3-5
Crowley T (2000) Causes of climate change over the past 1000 years. Science (80-) 289:270–277. doi: 10.1126/science.289.5477.270
D'Almeida TC, Sadissou I, Cottrell G, et al (2016) Evolution of the levels of human leukocyte antigen G (HLA-G) in Beninese infant during the first year of life in a malaria endemic area: Using latent class analysis. Malar J 15:1–10. doi: 10.1186/s12936-016-1131-y
D'Almeida TC, Sadissou I, Milet J, et al (2017) Soluble human leukocyte antigen -G during pregnancy and infancy in Benin: Mother/child resemblance and association with the risk of malaria infection and low birth weight. PLoS One 12:e0171117. doi: 10.1371/journal.pone.0171117
Da Silva IM, Thiengo R, Conceição MJ, et al (2006) Cystoscopy in the diagnosis and follow-up of urinary schistosomiasis in Brazilian soldiers returning from Mozambique, Africa. Rev Inst Med Trop Sao Paulo 48:39–42. doi: /S0036-46652006000100008
Daszak P, Cunningham AA, Hyatt AD (2000) Emerging infectious diseases of wildlife - Threats to biodiversity and human health. Science (80-.). 287:443–449
de Albuquerque RS, Mendes-Junior CT, Lucena-Silva N, et al (2016) Association of HLA-G 3' untranslated region variants with type 1 diabetes mellitus. Hum Immunol 77:358–364. doi: 10.1016/j.humimm.2016.02.001
de Almeida BS, Muniz YCN, Prompt AH, et al (2018) Genetic association between HLA-G 14-bp polymorphism and diseases: a systematic review and meta-analysis. Hum Immunol 0–1. doi: 10.1016/j.humimm.2018.08.003
Degarege A, Degarege D, Veledar E, et al (2016) Plasmodium falciparum Infection Status among Children with Schistosoma in Sub-Saharan Africa: A Systematic Review and Meta-analysis. PLoS Negl Trop Dis 10:1–18. doi: 10.1371/journal.pntd.0005193
Deininger P (2011) Alu elements: know the SINEs. Genome Biol 12:236. doi: gb-2011-12-12-236 [pii]\r10.1186/gb-2011-12-12-236
DePristo MA, Banks E, Poplin RE, et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. doi: 10.1038/ng.806

Dessein a J, Hillaire D, Elwali NE, et al (1999) Severe hepatic fibrosis in Schistosoma mansoni infection is controlled by a major locus that is closely linked to the interferon-gamma receptor gene. Am J Hum Genet 65:709–721. doi: 10.1086/302526

Dessein AJ, Couissinier P, Demeure C, et al (1992) Environmental, genetic and immunological factors in human resistance to schistosoma mansoni. Immunol Invest 21:423–453. doi: 10.3109/08820139209069383

Devidas A, Lamothe F, Develoux M, et al (1988) [Morbidity due to bilharziasis caused by S. haematobium. Relationship between the bladder lesions observed by ultrasonography and the cystoscopic and anatomo-pathologic lesions]. Acta Trop 45:277–287

Devlin B, Risch N (1995) A Comparison of Linkage Disequilibrium Measures for Fine-Scale Mappingitle. Genomics 29:311–322. doi: S0888-7543(85)79003-9 [pii] 10.1006/geno.1995.9003

Diallo TO, Remoue F, Gaayeb L, et al (2010) Schistosomiasis coinfection in children influences acquired immune response against Plasmodium falciparum malaria antigens. PLoS One 5:e12764. doi: 10.1371/journal.pone.0012764

Dias FC, Mendes-Junior CT, Silva MC, et al (2015) Human leucocyte antigen-G (HLA-G) and its murine functional homolog Qa2 in the trypanosoma cruzi infection. Mediators Inflamm 2015:. doi: 10.1155/2015/595829

Djènontin A (2015) Mosquitoes fauna diversity , Plasmodium falciparum infection and insecticide resistance status in malaria vectors in a lagoon area in Southern Benin , West Africa. Int J Multidiscip Curent Res 5:1389–1393. doi: 10.14741/ijmcr/v.5.1.23

Djenontin A, Bio-Bangana S, Moiroux N, et al (2010) Culicidae diversity, malaria transmission and insecticide resistance alleles in malaria vectors in Ouidah-Kpomasse-Tori district from Benin (West Africa): A pre-intervention study. Parasit. Vectors 83

Doenhoff MJ, Chiodini PL, Hamilton J V. (2004) Specific and sensitive diagnosis of schistosome infection: Can it be done with antibodies? Trends Parasitol. 20:35–39

Dolo A, Modiano D, Maiga B, et al (2005) Difference in susceptibility to malaria between two sympatric ethnic groups in Mali. Am J Trop Med Hyg 72:243–248. doi: 72/3/243 [pii]

Donadi EA, Castelli EC, Arnaiz-Villena A, et al (2011) Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. Cell Mol Life Sci 68:369–395. doi: 10.1007/s00018-010-0580-7

Doumbo S, Tran TM, Sangala J, et al (2014) Co-infection of long-term carriers of Plasmodium falciparum with Schistosoma haematobium enhances protection from febrile malaria: a prospective cohort study in Mali. PLoS Negl Trop Dis 8:e3154. doi: 10.1371/journal.pntd.0003154

Driss A, Hibbert JM, Wilson NO, et al (2011) Genetic polymorphisms linked to susceptibility to malaria. Malar J 10:1–10. doi: 10.1186/1475-2875-10-271

Dulberger CL, McMurtrey CP, Hölzemer A, et al (2017) Human Leukocyte Antigen F Presents Peptides and Regulates Immunity through Interactions with NK Cell Receptors. Immunity 46:1018–1029.e7. doi: 10.1016/j.immuni.2017.06.002

Dunn RR, Davies TJ, Harris NC, Gavin MC (2010) Global drivers of human pathogen richness and prevalence. Proc R Soc B Biol Sci 277:2587–2595. doi: 10.1098/rspb.2010.0340

Durnez L, Pareyn M, Mean V, et al (2018) Identification and characterization of areas of high and low risk for asymptomatic malaria infections at sub-village level in Ratanakiri, Cambodia. Malar J 17:1–14. doi: 10.1186/s12936-017-2169-1

Eguchi H, Maeda A, Lo PC, et al (2016) HLA-G1, but Not HLA-G3, Suppresses Human Monocyte/Macrophage-mediated Swine Endothelial Cell Lysis. Transplant Proc 48:1285–1287. doi: 10.1016/j.transproceed.2015.10.070

Ekpo UF, Oluwole AS, Abe EM, et al (2012) Schistosomiasis in infants and pre-school-aged children in sub-Saharan Africa: Implication for control. Parasitology 139:835–841

Everts B, Adegnika AA, Kruize YCM, et al (2010) Functional Impairment of Human Myeloid Dendritic Cells during Schistosoma haematobium Infection. PLoS Negl Trop Dis 4:e667. doi: 10.1371/journal.pntd.0000667

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50. doi: 10.1016/j.ympev.2005.04.031

Fan W, Huang Z, Li S, Xiao Z (2017) The HLA-G 14-bp polymorphism and recurrent implantation failure: a meta-analysis. J Assist Reprod Genet. doi: 10.1007/s10815-017-0994-3

Fan W, Li S, Huang Z, Chen Q (2014) Relationship between HLA-G polymorphism and susceptibility to recurrent miscarriage: A meta-analysis of non-family-based studies. J Assist Reprod Genet. doi: 10.1007/s10815-013-0155-2

Feldmeier H, Daccal RC, Martins MJ, et al (1998) Genital Manifestations of Schistosomiasis Mansoni in Women: Important but Neglected. Mem Inst Oswaldo Cruz 93 SUPPL.:127–133. doi: 10.1590/S0074-02761998000700018

Feldmeier H, Leutscher P, Poggensee G, Harms G (1999) Editorial: Male genital schistosomiasis and haemospermia. Trop. Med. Int. Heal. 4:791–793

Felício LP, Porto IOP, Mendes-Junior CT, et al (2014) Worldwide HLA-E nucleotide and haplotype variability reveals a conserved gene for coding and 3′ untranslated regions. Tissue Antigens 83:82–93. doi: 10.1111/tan.12283

Flori L, Delahaye NF, Iraqi FA, et al (2005) TNF as a malaria candidate gene: Polymorphism-screening and family-based association analysis of mild malaria attack and parasitemia in Burkina Faso. Genes Immun 6:472–480. doi: 10.1038/sj.gene.6364231

Flori L, Sawadogo S, Esnault C, et al (2003) Linkage of mild malaria to the major histocompatibility complex in families living in Burkina Faso. Hum. Mol. Genet. 12:375–378

Fournel S, Aguerre-Girr M, Huc X, et al (2000) Cutting Edge: Soluble HLA-G1 Triggers CD95/CD95 Ligand-Mediated Apoptosis in Activated CD8+ Cells by Interacting with CD8. J Immunol 164:6100–6104. doi: 10.4049/jimmunol.164.12.6100

Franklin BS, Parroche P, Ataíde MA, et al (2009) Malaria primes the innate immune response due to interferon-gamma induced enhancement of toll-like receptor expression and function. Proc Natl Acad Sci U S A 106:5789–94. doi: 10.1073/pnas.0809742106

Freer JB, Bourke CD, Durhuus GH, et al (2017) Schistosomiasis in the first 1000 days. Lancet Infect Dis 3099:. doi: 10.1016/S1473-3099(17)30490-5

Fu CL, Odegaard JI, Herbert DR, Hsieh MH (2012) A novel mouse model of schistosoma haematobium egg-induced immunopathology. PLoS Pathog 8:. doi: 10.1371/journal.ppat.1002605

Fumagalli M, Pozzoli U, Cagliani R, et al (2010) The landscape of human genes involved in the immune response to parasitic worms. BMC Evol Biol 10:. doi: 10.1186/1471-2148-10-264

Garcia-Beltran WF, Hölzemer A, Martrus G, et al (2016) Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. Nat Immunol 17:1067–1074. doi: 10.1038/ni.3513

Garcia A, Marquet S, Bucheton B, et al (1998) Linkage analysis of blood Plasmodium falciparum levels: Interest of the 5q31-q33 chromosome region. Am J Trop Med Hyg 58:705–709. doi: 10.4269/ajtmh.1998.58.705

Garcia A, Milet J, Courtin D, et al (2013) Association of HLA-G 3'UTR polymorphisms with response to malaria infection: A first insight. Infect Genet Evol 16:263–269. doi:

10.1016/j.meegid.2013.02.021
Garcia A, Milet J, Watier L, et al (2010) Genome wide linkage and association study of plasmodium falciparum infection and mild malaria attack. Tissue Antigens Conference:495–496. doi: http://dx.doi.org/10.1111/j.1399-0039.2010.01488.x
Garner P, Gülmezoglu AM (2003) Drugs for preventing malaria-related illness in pregnant women and death in the newborn. Cochrane Database Syst Rev CD000169. doi: 10.1002/14651858.CD000169
Garziera M, Bidoli E, Cecchin E, et al (2015) HLA-G 3'UTR polymorphisms impact the prognosis of stage II-III CRC patients in fluoropyrimidine-based treatment. PLoS One 10:. doi: 10.1371/journal.pone.0144000
Ge YZ, Ge Q, Li MH, et al (2014) Association between human leukocyte antigen-G 14-bp insertion/deletion polymorphism and cancer risk: a meta-analysis and systematic review. Hum Immunol. doi: 10.1016/j.humimm.2014.06.004
Geraghty DE, Koller BH, Hansen JA, Orr HT (1992) The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. J Immunol 149:1934–46. doi: 10.4049/jimmunol.171.10.5264
Gineau L, Courtin D, Camara M, et al (2016) Human Leukocyte Antigen-G: A Promising Prognostic Marker of Disease Progression to Improve the Control of Human African Trypanosomiasis. Clin Infect Dis 63:1189–1197. doi: 10.1093/cid/ciw505
Gineau L, Luisi P, Castelli EC, et al (2015) Balancing immunity and tolerance: Genetic footprint of natural selection in the transcriptional regulatory region of HLA-G. Genes Immun 16:57–70. doi: 10.1038/gene.2014.63
GlaxoSmithKline (2015) Fact sheet : The RTS , S malaria vaccine candidate clinical trials. 1–4
Gomes EC de S, Leal-Neto OB, Albuquerque J, et al (2012) Schistosomiasis transmission and environmental change: a spatio-temporal analysis in Porto de Galinhas, Pernambuco - Brazil. Int J Health Geogr 11:1–11. doi: 10.1186/1476-072X-11-51
Gomes PS, Bhardwaj J, Rivera-Correa J, et al (2016) Immune escape strategies of malaria parasites. Front. Microbiol. 7
Goodridge JP, Burian A, Lee N, Geraghty DE (2010) HLA-F Complex without Peptide Binds to MHC Class I Protein in the Open Conformer Form. J Immunol 184:6199–6208. doi: 10.4049/jimmunol.1000078
Goodridge JP, Burian A, Lee N, Geraghty DE (2013) HLA-F and MHC Class I Open Conformers Are Ligands for NK Cell Ig-like Receptors. J Immunol 191:3553–3562. doi: 10.4049/jimmunol.1300081
Gordon CA, Acosta LP, Gray DJ, et al (2012) High Prevalence of Schistosoma japonicum Infection in Carabao from Samar Province, the Philippines: Implications for Transmission and Control. PLoS Negl Trop Dis 6:. doi: 10.1371/journal.pntd.0001778
Grange C, Tapparo M, Tritta S, et al (2015) Role of HLA-G and extracellular vesicles in renal cancer stem cell-induced inhibition of dendritic cell differentiation. BMC Cancer 15:1009. doi: 10.1186/s12885-015-2025-z
Greer GJ, Ow-Yang CK, Yong H-S (1988) Schistosoma malayensis n. sp.: A Schistosoma japonicum-Complex Schistosome from Peninsular Malaysia. J Parasitol 74:471–480. doi: 10.2307/3282058
Grogan JL, Kremsner PG, van Dam GJ, et al (1997) Anti-schistosome IgG4 and IgE at 2 years after chemotherapy: infected versus uninfected individuals. J Infect Dis 176:1344–1350
Gryseels B, Polman K, Clerinx J, Kestens L (2006a) Human Schistosomiasis. Lancet 43:323. doi: 10.1016/S0140-6736(06)69440-3
Gryseels B, Polman K, Clerinx J, Kestens L (2006b) Human schistosomiasis. Lancet 43:323.

doi: 10.1016/S0140-6736(06)69440-3
Gryseels B, Strickland GT (2012) Schistosomiasis. In: Hunter's Tropical Medicine and Emerging Infectious Disease: Ninth Edition. pp 867–883
Guerra CA, Gikandi PW, Tatem AJ, et al (2008) The limits and intensity of Plasmodium falciparum transmission: Implications for malaria control and elimination worldwide. PLoS Med 5:0300–0311. doi: 10.1371/journal.pmed.0050038
Guerra CA, Howes RE, Patil AP, et al (2010) The international limits and population at risk of Plasmodium vivax transmission in 2009. PLoS Negl Trop Dis 4:. doi: 10.1371/journal.pntd.0000774
Guilbride DL, Guilbride PDL, Gawlinski P (2012) Malaria's deadly secret: A skin stage. Trends Parasitol. 28:142–150
Guo SW, Guo SW, Thompson E a, Thompson E a (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372. doi: 10.2307/2532296
Guthmann JP, Llanos-Cuentas a, Palacios a, Hall a J (2002) Environmental factors as determinants of malaria risk. A descriptive study on the northern coast of Peru. Trop Med Int Health 7:518–525. doi: 10.1046/j.1365-3156.2002.00883.x
Hackmon R, Pinnaduwage L, Zhang J, et al (2017) Definitive class I human leukocyte antigen expression in gestational placentation: HLA-F, HLA-E, HLA-C, and HLA-G in extravillous trophoblast invasion on placentation, pregnancy, and parturition. Am J Reprod Immunol 77:. doi: 10.1111/aji.12643
Hagan P, Blumenthal UJ, Dunn D, et al (1991) Human IgE, IgG4 and resistance to reinfection with Schistosomahaematobium. Nature 349:243–245. doi: 10.1038/349243a0
Han M, Jiang Y, Lao K, et al (2014) sHLA-G Involved in the Apoptosis of Decidual Natural Killer Cells Following Toxoplasma gondii Infection. Inflammation 37:1718–1727. doi: 10.1007/s10753-014-9900-7
Harada A, Ishigami S, Kijima Y, et al (2015) Clinical implication of human leukocyte antigen (HLA)-F expression in breast cancer. Pathol Int 65:569–574. doi: 10.1111/pin.12343
Hashimoto A, Pincelli C, Fujioka A, et al (1990) Relationship between NK cells and granulomatous inflammation in mice. J Clin {&} Lab Immunol 33:41–47
Hay SI, Guerra CA, Gething PW, et al (2009) A world malaria map: plasmodium falciparum endemicity in 2007. PLoS Med 6:0286–0302. doi: 10.1371/journal.pmed.1000048
He H, Isnard A, Kouriba B, et al (2008) A STAT6 gene polymorphism is associated with high infection levels in urinary schistosomiasis. Genes Immun 9:195–206. doi: 10.1038/gene.2008.2
He YX, Salafsky B, Ramaswamy K (2005) Comparison of skin invasion among three major species of Schistosoma. Trends Parasitol. 21:201–203
Health I, Malaria RB (2014) Severe Malaria. 19:7–131. doi: 10.1111/tmi.12313
Herbert DR, Hölscher C, Mohrs M, et al (2004) Alternative macrophage activation is essential for survival during schistosomiasis and downmodulates T helper 1 responses and immunopathology. Immunity 20:623–635. doi: 10.1016/S1074-7613(04)00107-4
Heredia JE, Mukundan L, Chen FM, et al (2013) Type 2 innate signals stimulate fibro/adipogenic progenitors to facilitate muscle regeneration. Cell 153:376–388. doi: 10.1016/j.cell.2013.02.053
Hernandez-Valladares M, Rihet P, Iraqi FA (2014) Host susceptibility to malaria in human and mice: compatible approaches to identify potential resistant genes. Physiol Genomics 46:1–16. doi: 10.1152/physiolgenomics.00044.2013
Hiby SE, King A, Sharkey A, Loke YW (1999) Molecular studies of trophoblast HLA-G: Polymorphism, isoforms, imprinting and expression in preimplantation embryo. Tissue Antigens. doi: 10.1034/j.1399-0039.1999.530101.x
Holling TM, Bergevoet MWT, Wierda RJ, et al (2009) Genetic and epigenetic control of the

major histocompatibility complex class Ib Gene HLA-G in trophoblast cell lines. Ann N Y Acad Sci. doi: 10.1111/j.1749-6632.2009.04660.x

Hotez PJ, Alvarado M, Basáñez M-G, et al (2014) The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases. PLoS Negl Trop Dis 8:e2865. doi: 10.1371/journal.pntd.0002865

Hou X, Yu F, Man S, et al (2012) Negative regulation of Schistosoma japonicum egg-induced liver fibrosis by natural killer cells. PLoS Negl Trop Dis 6:. doi: 10.1371/journal.pntd.0001456

Huang J, Burke P, Yang Y, et al (2010) Soluble HLA-G Inhibits Myeloid Dendritic Cell Function in HIV-1 Infection by Interacting with Leukocyte Immunoglobulin-Like Receptor B2. J Virol 84:10784–10791. doi: 10.1128/JVI.01292-10

Huy NT, Hamada M, Kikuchi M, et al (2011) Association of HLA and post-schistosomal hepatic disorder: A systematic review and meta-analysis. Parasitol Int 60:347–356. doi: 10.1016/j.parint.2011.05.008

Hviid TVF, Hylenius S, Rorbye C, Nielsen LG (2003) HLA-G allelic variants are associated with differences in the HLA-G mRNA isoform profile and HLA-G mRNA levels. Immunogenetics 55:63–79. doi: 10.1007/s00251-003-0547-z

Hviid TVF, Rizzo R, Melchiorri L, et al (2006) Polymorphism in the 5′ Upstream Regulatory and 3′ Untranslated Regions of the HLA-G Gene in Relation to Soluble HLA-G and IL-10 Expression. Hum Immunol 67:53–62. doi: 10.1016/j.humimm.2005.12.003

Ibikounlé M, Gbédjissi LG, Ogouyèmi-Hounto A, et al (2014) Schistosomose et géohelminthoses dans le nord-est du Bénin: cas des écoliers des communes de Nikki et de Pèrèrè. Bull la Soc Pathol Exot 107:171–176. doi: 10.1007/s13149-014-0344-y

Ibikounlé M, Mouahid G, Mintsa Nguema R, et al (2013) Snail intermediate host/Schistosoma haematobium relationships from three transmission sites in Benin (West Africa). Parasitol Res 112:227–233. doi: 10.1007/s00436-012-3129-x

Ibikounlé M, Mouahid G, Sakiti NG, et al (2009) Freshwater snail diversity in Benin (West Africa) with a focus on human schistosomiasis. Acta Trop 111:29–34. doi: 10.1016/j.actatropica.2009.02.001

Ibikounlé M, Onzo-Aboki A, Doritchamou J, et al (2018) Results of the first mapping of soil-transmitted helminths in Benin: Evidence of countrywide hookworm predominance. PLoS Negl Trop Dis 12:1–17. doi: 10.1371/journal.pntd.0006241

Institut National de la Statistique et de l'Analyse Economique (2015) Rgph4 : Que Retenir Des Effectifs De Population En 2013 ? République du Bénin 35

Ishitani A, Sageshima N, Hatake K (2006) The involvement of HLA-E and -F in pregnancy. J. Reprod. Immunol. 69:101–113

Ishitani A, Sageshima N, Lee N, et al (2003) Protein Expression and Peptide Binding Suggest Unique and Interacting Functional Roles for HLA-E, F, and G in Maternal-Placental Immune Recognition. J Immunol 171:1376–1384. doi: 10.4049/jimmunol.171.3.1376

Iskander R, Das PK, Aalberse RC (1981) IgG4 antibodies in Egyptian patients with schistosomiasis. Int Arch Allergy Appl Immunol 66:200–207

Ismail HAHA, Hong S-T, Babiker ATEB, et al (2014) Prevalence, risk factors, and clinical manifestations of schistosomiasis among school children in the White Nile River basin, Sudan. Parasit Vectors 7:478. doi: 10.1186/s13071-014-0478-6

Isnard A, Chevillard C (2008a) Recent advances in the characterization of genetic factors involved in human susceptibility to infection by schistosomiasis. Curr Genomics 9:290–300. doi: 10.2174/138920208785133262

Isnard A, Chevillard C (2008b) Recent advances in the characterization of genetic factors involved in human susceptibility to infection by schistosomiasis. Curr Genomics 9:290–300. doi: 10.2174/138920208785133262

Isnard A, Kouriba B, Doumbo O, Chevillard C (2011) Association of rs7719175, located in the IL13 gene promoter, with Schistosoma haematobium infection levels and identification of a susceptibility haplotype. Genes Immun 12:31–39. doi: 10.1038/gene.2010.43

Janeway CA, Travers P, Walport M, Shlomchik M (2001) Immunobiology: The Immune System In Health And Disease. Immuno Biol 5 892. doi: 10.1111/j.1467-2494.1995.tb00120.x

Jassem RM, Shani WS, Loisel DA, et al (2012) HLA-G polymorphisms and soluble HLA-G protein levels in women with recurrent pregnancy loss from Basrah province in Iraq. Hum Immunol 73:811–817. doi: 10.1016/j.humimm.2012.05.009

Jauréguiberry S, Ansart S, Perez L, et al (2007) Acute neuroschistosomiasis: Two cases associated with cerebral vasculitis. Am J Trop Med Hyg 76:964–966. doi: 10.1007/s11908-013-0329-1

Jørgensen PB, Livbjerg AH, Hansen HJ, et al (2012) Epstein-Barr virus Peptide Presented by HLA-E is Predominantly Recognized by CD8bright Cells in multiple Sclerosis Patients. PLoS One 7:. doi: 10.1371/journal.pone.0046120

Kalinda C, Chimbari M, Mukaratirwa S (2017) Implications of changing temperatures on the growth, fecundity and survival of intermediate host snails of schistosomiasis: A systematic review. Int J Environ Res Public Health 14:. doi: 10.3390/ijerph14010080

Kayentao K, Garner P, Maria van Eijk A, et al (2013) Intermittent Preventive Therapy for Malaria During Pregnancy Using 2 vs 3 or More Doses of Sulfadoxine-Pyrimethamine and Risk of Low Birth Weight in Africa. JAMA 309:594. doi: 10.1001/jama.2012.216231

Kent WJ, Sugnet CW, Furey TS, et al (2002) The Human Genome Browser at UCSC. Genome Res 12:996–1006. doi: 10.1101/gr.229102

Kï¿½stlin N, Ostermeir AL, Spring B, et al (2017) HLA-G promotes myeloid-derived suppressor cell accumulation and suppressive activity during human pregnancy through engagement of the receptor ILT4. Eur J Immunol 47:374–384. doi: 10.1002/eji.201646564

Kim SK, Hong MS, Shin MK, et al (2011) Promoter polymorphisms of the HLA-G gene, but not the HLA-E and HLA-F genes, is associated with non-segmental vitiligo patients in the Korean population. Arch Dermatol Res 303:679–684. doi: 10.1007/s00403-011-1160-x

Kim SK, Jeong KH, Kang IJ, et al (2015) Relationship between the HLA-G 14bp insertion/deletion polymorphism and susceptibility to autoimmune disease: A meta-analysis. Genet Mol Res. doi: 10.4238/2015.December.1.35

Kincaid-Smith J, Rey O, Toulza E, et al (2017) Emerging Schistosomiasis in Europe: A Need to Quantify the Risks. Trends Parasitol. 33:600–609

Kochan G, Escors D, Breckpot K, Guerrero-Setas D (2013) Role of non-classical MHC class I molecules in cancer immunosuppression. Oncoimmunology 2

Koller BH, Geraghty DE, Shimizu Y, et al (1988) HLA-E. A novel HLA class I gene expressed in resting T lymphocytes. J Immunol 141:897–904

Kouriba B, Chevillard C, Bream JH, et al (2005) Analysis of the 5q31-q33 Locus Shows an Association between IL13-1055C/T IL-13-591A/G Polymorphisms and Schistosoma haematobium Infections. J Immunol 174:6274–6281. doi: 10.4049/jimmunol.174.10.6274

Kourilova P, Hogg KG, Kolarova L, Mountford AP (2004) Cercarial Dermatitis Caused by Bird Schistosomes Comprises Both Immediate and Late Phase Cutaneous Hypersensitivity Reactions. J Immunol 172:3766–3774. doi: 10.4049/jimmunol.172.6.3766

Kovats S, Main EK, Librach C, et al (1990) A class I antigen, HLA-G, expressed in human trophoblasts. Science 248:220–223. doi: 10.1126/science.2326636

Kulski JK, Martinez P, Longman-Jacobsen N, et al (2001) The association between HLA-A alleles and an Alu dimorphism near HLA-G. J Mol Evol 53:114–123. doi: 10.1007/s002390010199

Lai Y-S, Biedermann P, Ekpo UF, et al (2015) Spatial distribution of schistosomiasis and treatment needs in sub-Saharan Africa: a systematic review and geostatistical analysis. Lancet Infect Dis 15:927–940. doi: 10.1016/S1473-3099(15)00066-3

Lambertucci JR (2010) Acute schistosomiasis mansoni: Revisited and reconsidered. Mem Inst Oswaldo Cruz 105:422–435. doi: 10.1590/S0074-02762010000400012

Le Port A, Cottrell G, Martin-Prevel Y, et al (2012) First malaria infections in a cohort of infants in Benin: Biological, environmental and genetic determinants. Description of the study site, population methods and preliminary results. BMJ Open 2:1–12. doi: 10.1136/bmjopen-2011-000342

Lee N, Goodlett DR, Ishitani A, et al (1998) HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. J Immunol 160:4951–60. doi: 10.1038/nature01413

Lee YH, Bae SC, Song GG (2015) Meta-analysis of associations between functional HLA-G polymorphisms and susceptibility to systemic lupus erythematosus and rheumatoid arthritis. Rheumatol. Int.

Lemaitre M, Watier L, Briand V, et al (2014) Coinfection with Plasmodium falciparum and Schistosoma haematobium: Additional evidence of the protective effect of schistosomiasis on malaria in Senegalese children. Am J Trop Med Hyg 90:329–334. doi: 10.4269/ajtmh.12-0431

Lepesheva GI, Hargrove TY, Rachakonda G, et al (2015) cr ipt ce pt cr ipt Ac ce pt us. 1–31

Leshem E, Maor Y, Meltzer E, et al (2008) Acute Schistosomiasis Outbreak: Clinical Features and Economic Impact. Clin Infect Dis 47:1499–1506. doi: 10.1086/593191

Li T, Huang H, Liao D, et al (2015) Genetic polymorphism in HLA-G 3′UTR 14-bp ins/del and risk of cancer: a meta-analysis of case–control study. Mol Genet Genomics. doi: 10.1007/s00438-014-0985-3

Liehl P, Zuzarte-Luís V, Chan J, et al (2014) Host-cell sensors for Plasmodium activate innate immunity against liver-stage infection. Nat Med 20:47–53. doi: 10.1038/nm.3424

Lila N, Rouas-Freiss N, Dausset J, et al (2001) Soluble HLA-G protein secreted by allo-specific CD4+ T cells suppresses the allo-proliferative response: a CD4+ T cell regulatory mechanism. Proc Natl Acad Sci U S A 98:12150–5. doi: 10.1073/pnas.201407398

Lima THA, Buttura RV, Donadi EA, et al (2016) HLA-F coding and regulatory segments variability determined by massively parallel sequencing procedures in a Brazilian population sample. Hum Immunol 77:841–853. doi: 10.1016/j.humimm.2016.07.231

Liu L-X, Li-Li J, Qiong C, Xiao-Lin F (2017) Recent advances in the synthesis of antischistosomal drugs and agents. Mini-Reviews Med Chem 17:. doi: 10.2174/1389557513666131119204558

Liu XX, Pan FH, Tian W (2012) Characterization of HLA-E polymorphism in four distinct populations in Mainland China. Tissue Antigens 80:26–35. doi: 10.1111/j.1399-0039.2012.01873.x

Lopes DM, de Almeida TVVS, de Souza R da P, et al (2018) Susceptibility of dendritic cells from individuals with schistosomiasis to infection by Leishmania braziliensis. Mol Immunol 93:173–183. doi: 10.1016/j.molimm.2017.11.018

Lucena-Silva N, de Souza VSB, Gomes RG, et al (2013) HLA-G 3’ Untranslated Region Polymorphisms Are Associated with Systemic Lupus Erythematosus in 2 Brazilian

Populations. J Rheumatol 40:1104–1113. doi: 10.3899/jrheum.120814
Lucena-Silva N, Monteiro AR, de Albuquerque RS, et al (2012) Haplotype frequencies based on eight polymorphic sites at the 3' untranslated region of the HLA-G gene in individuals from two different geographical regions of Brazil. Tissue Antigens 79:272–278. doi: 10.1111/j.1399-0039.2012.01842.x
Lundy SK, Boros DL (2002) Fas ligand-expressing B-1a lymphocytes mediate CD4+-T-cell apoptosis during schistosomal infection: Induction by interleukin 4 (IL-4) and IL-10. Infect Immun 70:812–819. doi: 10.1128/IAI.70.2.812-819.2002
Lyke KE, Dabo A, Arama C, et al (2012a) Reduced T regulatory cell response during acute Plasmodium falciparum infection in Malian children co-infected with Schistosoma haematobium. PLoS One 7:e31647. doi: 10.1371/journal.pone.0031647
Lyke KE, Dabo A, Arama C, et al (2018) Long-term Maintenance of CD4 T Cell Memory Responses to Malaria Antigens in Malian Children Coinfected with Schistosoma haematobium. Front Immunol 8:1–11. doi: 10.3389/fimmu.2017.01995
Lyke KE, Dabo A, Sangare L, et al (2006) Effects of concomitant Schistosoma haematobium infection on the serum cytokine levels elicited by acute Plasmodium falciparum malaria infection in malian children. Infect Immun 74:5718–5724. doi: 10.1128/IAI.01822-05
Lyke KE, Dicko A, Dabo A, et al (2005) Association of Schistosoma haematobium infection with protection against acute Plasmodium falciparum malaria in Malian children. Am J Trop Med Hyg 73:1124–1130. doi: 10.4269/ajtmh.2005.73.1124
Lyke KE, Wang A, Dabo A, et al (2012b) Antigen-specific B memory cell responses to Plasmodium falciparum malaria antigens and Schistosoma haematobium antigens in co-infected Malian children. PLoS One 7:e37868. doi: 10.1371/journal.pone.0037868
Mackinnon MJ, Mwangi TW, Snow RW, et al (2005) Heritability of malaria in Africa. PLoS Med 2:1253–1259. doi: 10.1371/journal.pmed.0020340
Maizels RM (2009) Parasite immunomodulation and polymorphisms of the immune system. J. Biol. 8
Mallevaey T, Fontaine J, Breuilh L, et al (2007) Invariant and noninvariant natural killer T cells exert opposite regulatory functions on the immune response during murine schistosomiasis. Infect Immun 75:2171–2180. doi: 10.1128/IAI.01178-06
Mallevaey T, Zanetta JP, Faveeuw C, et al (2006) Activation of Invariant NKT Cells by the Helminth Parasite Schistosoma mansoni. J Immunol 176:2476–2485. doi: 10.4049/jimmunol.176.4.2476
Mangano VD, Modiano D (2014) Host genetics and parasitic infections. Clin Microbiol Infect 20:1265–1275. doi: 10.1111/1469-0691.12793
Manvailer LFS, Wowk PF, Mattar SB, et al (2014) HLA-F polymorphisms in a Euro-Brazilian population from Southern Brazil. Tissue Antigens 84:554–559. doi: 10.1111/tan.12444
Marchesi M, Andersson E, Villabona L, et al (2013) HLA-dependent tumour development: A role for tumour associate macrophages? J Transl Med 11:. doi: 10.1186/1479-5876-11-247
Mariconti M, Meroni V, Badulli C, et al (2016) Correlation of serum sHLA-G levels with cyst stage in patients with cystic echinococcosis: is it an immune evasion strategy? Parasite Immunol 38:414–418. doi: 10.1111/pim.12328
Marquet S (2017) Overview of human genetic susceptibility to malaria: From parasitemia control to severe disease. Infect. Genet. Evol.
Marquet S, Abel L, Hillaire D, et al (1996) Genetic localization of a locus controlling the intensity of infection by Schistosoma mansoni on chromosome 5q31-q33. Nat Genet 14:181–184. doi: 10.1038/ng1096-181
Martelli-Palomino G, Pancotto JA, Muniz YC, et al (2013) Polymorphic Sites at the 3'

Untranslated Region of the HLA-G Gene Are Associated with Differential hla-g Soluble Levels in the Brazilian and French Population. PLoS One 8:. doi: 10.1371/journal.pone.0071742

Martens WJ, Niessen LW, Rotmans J, et al (1995) Potential impact of global climate change on malaria risk. Environ Health Perspect 103:458–464. doi: 10.1289/ehp.95103458

Matthews HD, Weaver AJ, Meissner KJ, et al (2004) Natural and anthropogenic climate change: incorporating historical land cover change, vegetation dynamics and the global carbon cycle. Clim Dyn 22:. doi: 10.1007/s00382-004-0392-2

Mbanefo EC, Huy NT, Wadagni AA, et al (2014) Host Determinants of Reinfection with Schistosomes in Humans: A Systematic Review and Meta-analysis. PLoS Negl Trop Dis 8:. doi: 10.1371/journal.pntd.0003164

McKenna A, Hanna M, Banks E, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. doi: 10.1101/gr.107524.110

Ménard R, Tavares J, Cockburn I, et al (2013) Looking under the skin: the first steps in malarial infection and immunity. Nat Rev Microbiol 11:701–712. doi: 10.1038/nrmicro3111

Mendes-Junior CT, Castelli EC, Meyer D, et al (2013) Genetic diversity of the HLA-G coding region in Amerindian populations from the Brazilian Amazon: A possible role of natural selection. Genes Immun 14:518–526. doi: 10.1038/gene.2013.47

Merrifield M, Hotez PJ, Beaumier CM, et al (2016) Advancing a vaccine to prevent human schistosomiasis. Vaccine 34:2988–2991. doi: 10.1016/j.vaccine.2016.03.079

Meurs L, Mbow M, Boon N, et al (2014) Cytokine Responses to Schistosoma mansoni and Schistosoma haematobium in Relation to Infection in a Co-endemic Focus in Northern Senegal. PLoS Negl Trop Dis 8:. doi: 10.1371/journal.pntd.0003080

Milet J, Nuel G, Watier L, et al (2010) Genome wide linkage study, using a 250K SNP map, of Plasmodium falciparum infection and mild malaria attack in a senegalese population. PLoS One 5:. doi: 10.1371/journal.pone.0011616

Miller JL, Sack BK, Baldwin M, et al (2014) Interferon-Mediated Innate Immune Responses against Malaria Parasite Liver Stages. Cell Rep 7:436–447. doi: 10.1016/j.celrep.2014.03.018

Misra MK, Prakash S, Kapoor R, et al (2013) Association of HLA-G promoter and 14-bp insertion-deletion variants with acute allograft rejection and end-stage renal disease. Tissue Antigens 82:317–326. doi: 10.1111/tan.12210

Mitra AK, Mawson AR (2017) Neglected Tropical Diseases: Epidemiology and Global Burden. Trop Med Infect Dis 2:36. doi: 10.3390/tropicalmed2030036

Modiano D, Chiucchiuini A, Petrarca V, et al (1998) Humoral response to Plasmodium falciparum Pf155/ring-infected erythrocyte surface antigen and Pf332 in three sympatric ethnic groups of Burkina Faso. Am J Trop Med Hyg 58:220–224. doi: 10.4269/ajtmh.1998.58.220

Modiano D, Petrarca V, Sirima BS, et al (1996) Different response to Plasmodium falciparum malaria in west African sympatric ethnic groups. Proc Natl Acad Sci U S A 93:13206–13211. doi: 10.1073/pnas.93.23.13206

Molehin AJ, Rojo JU, Siddiqui SZ, et al (2016) Development of a schistosomiasis vaccine. Expert Rev Vaccines 15:619–627. doi: 10.1586/14760584.2016.1131127

Molofsky AB, Nussbaum JC, Liang H-E, et al (2013) Innate lymphoid type 2 cells sustain visceral adipose tissue eosinophils and alternatively activated macrophages. J Exp Med 210:535–549. doi: 10.1084/jem.20121964

Morandi F, Ferretti E, Bocca P, et al (2010) A novel mechanism of soluble HLA-G mediated immune modulation: Downregulation of T cell chemokine receptor expression and

impairment of chemotaxis. PLoS One 5:. doi: 10.1371/journal.pone.0011763
Moreau P, Flajollet S, Carosella ED (2009a) Non-classical transcriptional regulation of HLA-G: An update. J. Cell. Mol. Med.
Moreau P, Flajollet S, Carosella ED (2009b) Non-classical transcriptional regulation of HLA-G: An update. J. Cell. Mol. Med. 13:2973–2989
Moreau P, Mouillot G, Rousseau P, et al (2003) HLA-G gene repression is reversed by demethylation. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.0337539100
Moscoso J, Serrano-Vela JI, Arnaiz-Villena A (2007) MHC-F polymorphism and evolution. In: Tissue Antigens. pp 136–139
Mouillot G, Marcou C, Rousseau P, et al (2005) HLA-G gene activation in tumor cells involves cis-acting epigenetic changes. Int J Cancer. doi: 10.1002/ijc.20682
Mutapi F, Burchmore R, Mduluza T, et al (2008) Age-Related and Infection Intensity–Related Shifts in Antibody Recognition of Defined Protein Antigens in a Schistosome-Exposed Population. J Infect Dis 198:167–75. doi: 10.1086/589511
Mutapi F, Roussilhon C, Mduluza T, Druilhe P (2007) Anti-malaria humoral responses in children exposed to Plasmodium falciparum and Schistosoma haematobium. Mem Inst Oswaldo Cruz 102:405–409
Muth S, Sayasone S, Odermatt-Biays S, et al (2010) Schistosoma mekongi in Cambodia and Lao People's Democratic Republic. Adv Parasitol 72:179–203. doi: 10.1016/S0065-308X(10)72007-8
Mutuku FM, Bayoh MN, Gimnig JE, et al (2006) Pupal habitat productivity of Anopheles gambiae complex mosquitoes in a rural village in western Kenya. Am J Trop Med Hyg 74:54–61. doi: 74/1/54 [pii]
Nachmani D, Zimmermann A, Oiknine Djian E, et al (2014) MicroRNA Editing Facilitates Immune Elimination of HCMV Infected Cells. PLoS Pathog 10:. doi: 10.1371/journal.ppat.1003963
Naji A, Menier C, Morandi F, et al (2014) Binding of HLA-G to ITIM-Bearing Ig-like Transcript 2 Receptor Suppresses B Cell Responses. J Immunol 192:1536–1546. doi: 10.4049/jimmunol.1300438
Nakagawa S, Niimura Y, Gojobori T, et al (2008) Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. Nucleic Acids Res 36:861–871. doi: 10.1093/nar/gkm1102
Nattermann J, Nischalke HD, Hofmeister V, et al (2005) The HLA-A2 restricted T cell epitope HCV core35-44stabilizes HLA-E expression and inhibits cytolysis mediated by natural killer cells. Am J Pathol 166:443–453. doi: 10.1016/S0002-9440(10)62267-5
Naus CWA, van Dam GJ, Naus CWA, et al (1998) Human IgE, IgG Subclass, and IgM Responses to Worm and Egg Antigens in Schistosomiasis Haematobium: A 12-Month Study of Reinfection in Cameroonian Children. Clin Infect Dis 26:1142–1147. doi: 10.1086/520310
Nausch N, Midzi N, Mduluza T, et al (2011) Regulatory and activated T cells in human Schistosoma haematobium infections. PLoS One 6:. doi: 10.1371/journal.pone.0016860
Nazari N, Farjadian S (2016) Natural Killer Cell Cytotoxicity Against SKOV3 after HLA-G Downregulation by shRNA. Iran J Immunol 13:178–185. doi: IJIv13i3A3
Nicolae D, Cox NJ, Lester LA, et al (2005) Fine Mapping and Positional Candidate Studies Identify HLA-G as an Asthma Susceptibility Gene on Chromosome 6p21. Am J Hum Genet 76:349–357. doi: 10.1086/427763
Nilsson LL, Djurisic S, Andersen AMN, et al (2016) Distribution of HLA-G extended haplotypes and one HLA-E polymorphism in a large-scale study of mother–child dyads with and without severe preeclampsia and eclampsia. HLA 88:172–186. doi: 10.1111/tan.12871

Ntonifor HN, Ajayi JA (2007) Studies on the ecology and distribution of some medically important freshwater snail species in Bauchi State, Nigeria. Int J Biol Chem Sci 1:121–127

O'Brien M, McCarthy T, Jenkins D, et al (2001) Altered HLA-G transcription in pre-eclampsia is associated with allele specific inheritance: Possible role of the HLA-G gene in susceptibility to the disease. Cell Mol Life Sci. doi: 10.1007/PL00000828

Ober C, Aldrich CL, Chervoneva I, et al (2003) Variation in the HLA-G Promoter Region Influences Miscarriage Rates. Am J Hum Genet 72:1425–1435. doi: 10.1086/375501

Ober C, Billstrand C, Kuldanek S, Tan Z (2006) The miscarriage-associated HLA-G -725G allele influences transcription rates in JEG-3 cells. Hum Reprod 21:1743–1748. doi: 10.1093/humrep/del036

Odegaard JI, Hsieh MH (2014) Immune responses to Schistosoma haematobium infection. Parasite Immunol 36:428–438. doi: 10.1111/pim.12084

Olieslagers TI, Voorter CEM, Groeneweg M, et al (2017) New insights in HLA-E polymorphism by refined analysis of the full-length gene. HLA 89:143–149. doi: 10.1111/tan.12965

Ouma C, Davenport GC, Were T, et al (2008) Haplotypes of IL-10 promoter variants are associated with susceptibility to severe malarial anemia and functional changes in IL-10 production. Hum Genet 124:515–524. doi: 10.1007/s00439-008-0578-5

Pabalan N, Jarjanazi H, Sun C, Iversen AC (2015) Meta-analysis of the human leukocyte antigen-G (HLA-G) 14 bp insertion/deletion polymorphism as a risk factor for preeclampsia. Tissue Antigens 86:. doi: 10.1111/tan.12627

Pan FH, Liu XX, Tian W (2013) Characterization of HLA-F polymorphism in four distinct populations in Mainland China. Int J Immunogenet 40:369–376. doi: 10.1111/iji.12047

Pérez del Villar L, Burguillo FJ, López-Abán J, Muro A (2012) Systematic Review and Meta-Analysis of Artemisinin Based Therapies for the Treatment and Prevention of Schistosomiasis. PLoS One 7:. doi: 10.1371/journal.pone.0045867

Perkins DJ, Weinberg JB, Kremsner PG (2000) Reduced Interleukin-12 and Transforming Growth Factor–β1 in Severe Childhood Malaria: Relationship of Cytokine Balance with Disease Severity. J Infect Dis 182:988–992. doi: 10.1086/315762

Persson G, Melsted WN, Nilsson LL, Hviid TVF (2017) HLA class Ib in pregnancy and pregnancy-related disorders. Immunogenetics 69:581–595

Petroff MG, Sedlmayr P, Azzola D, Hunt JS (2002) Decidual macrophages are potentially susceptible to inhibition by class Ia and class Ib HLA molecules. J Reprod Immunol 56:3–17. doi: 10.1016/S0165-0378(02)00024-4

Phimpraphil W, Paul R, Witoonpanich B, et al (2008) Heritability of P. falciparum and P. vivax Malaria in a Karen population in Thailand. PLoS One 3:. doi: 10.1371/journal.pone.0003887

Pietra G, Romagnani C, Manzini C, et al (2010) The emerging role of HLA-E-restricted CD8+ T lymphocytes in the adaptive immune response to pathogens and tumors. J Biomed Biotechnol 2010:907092. doi: 10.1155/2010/907092

Pietra G, Romagnani C, Moretta L, Mingari M (2009) HLA-E and HLA-E-Bound Peptides: Recognition by Subsets of NK and T Cells. Curr Pharm Des 15:3336–3344. doi: 10.2174/138161209789105207

Pluess B, Tanser FC, Lengeler C, Sharp BL (2010) Indoor residual spraying for preventing malaria (Review). In: COCHRANE Collab. http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006657.pub2/pdf

Polakova K, Bandzuchova E, Tirpakova J, et al (2007) Modulation of HLA-G expression. Neoplasma

Porto IOP, Mendes-Junior CT, Felício LP, et al (2015) MicroRNAs targeting the

immunomodulatory HLA-G gene: A new survey searching for microRNAs with potential to regulate HLA-G. Mol Immunol 65:230–241. doi: 10.1016/j.molimm.2015.01.030

Principaud J (1995) La pêche en milieu lagunaire dans le sud-est du Bénin. L'exemple de l'exploitation des acadjas (en danger) sur le lac Nokoué et la basse Sô. Cah d'outre-mer 519–546

Purcell S, Neale B, Todd-Brown K, et al (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 81:559–575. doi: 10.1086/519795

Pyo CW, Williams LM, Moore Y, et al (2006) HLA-E, HLA-F, and HLA-G polymorphism: Genomic sequence defines haplotype structure and variation spanning the nonclassical class I genes. Immunogenetics 58:241–251. doi: 10.1007/s00251-005-0076-z

R Core Team (2017) R Core Team (2017). R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria URL http//wwwR-project.org/ R Foundation for Statistical Computing

Ramalho J, Veiga-Castelli LC, Donadi EA, et al (2017) HLA-E regulatory and coding region variability and haplotypes in a Brazilian population sample. Mol Immunol 91:173–184. doi: 10.1016/j.molimm.2017.09.007

Randrianasolo BS, Jourdan PM, Ravoniarimbinina P, et al (2015) Gynecological manifestations, histopathological findings, and schistosoma-specific polymerase chain reaction results among women with schistosoma haematobium infection: A cross-sectional study in Madagascar. In: Journal of Infectious Diseases. pp 275–284

Ray D, Nelson TA, Fu CL, et al (2012) Transcriptional Profiling of the Bladder in Urogenital Schistosomiasis Reveals Pathways of Inflammatory Fibrosis and Urothelial Compromise. PLoS Negl Trop Dis 6:. doi: 10.1371/journal.pntd.0001912

Reiman RM, Thompson RW, Feng CG, et al (2006) Interleukin-5 (IL-5) augments the progression of liver fibrosis by regulating IL-13 activity. Infect Immun 74:1471–1479. doi: 10.1128/IAI.74.3.1471-1479.2006

Reis EAG, Azevedo TM, McBride AJA, et al (2007) Na??ve donor responses to Schistosoma mansoni soluble egg antigens. Scand J Immunol 66:662–670. doi: 10.1111/j.1365-3083.2007.02024.x

Remick DG, Chensue SW, Hiserodt JC, et al (1988) Flow-cytometric evaluation of lymphocyte subpopulations in synchronously developing Schistosoma mansoni egg and Sephadex bead pulmonary granulomas. Am J Pathol 131:298–307

Remoue F, To Van D, Schacht A-M, et al (2001) Gender-dependent specific immune response during chronic human Schistosomiasis haematobia. Clin Exp Immunol 124:62–68. doi: 10.1046/j.1365-2249.2001.01495.x

Rénia L, Goh YS (2016) Malaria parasite: the great escape. Front Immunol 7:463. doi: 10.3389/fimmu.2016.00463

Richardson ML, Fu C-L, Pennington LF, et al (2014) A New Mouse Model for Female Genital Schistosomiasis. PLoS Negl Trop Dis 8:e2825. doi: 10.1371/journal.pntd.0002825

Rihet P, Traoré Y, Abel L, et al (1998) Malaria in humans: Plasmodium falciparum blood infection levels are linked to chromosome 5q31-q33. Am J Hum Genet 63:498–505. doi: 10.1086/301967

Robert-Gangneux F, Gangneux JP, Vu N, et al (2011) High level of soluble HLA-G in amniotic fluid is correlated with congenital transmission of Toxoplasma gondii. Clin Immunol 138:129–134. doi: 10.1016/j.clim.2010.12.004

Robinson J, Halliwell JA, Hayhurst JD, et al (2015) The IPD and IMGT/HLA database: Allele variant databases. Nucleic Acids Res 43:D423–D431. doi: 10.1093/nar/gku1161

Robinson JT, Thorvaldsdóttir H, Winckler W, et al (2011) Integrative genomics viewer. Nat. Biotechnol. 29:24–26

Roca-Feltrer A, Carneiro I, Smith L, et al (2010) The age patterns of severe malaria syndromes in sub-Saharan Africa across a range of transmission intensities and seasonality settings. Malar J 9:. doi: 10.1186/1475-2875-9-282

Rosenzweig C, Karoly D, Vicarelli M, et al (2008) Attributing physical and biological impacts to anthropogenic climate change. Nature 453:353–357. doi: 10.1038/nature06937

Rosenzweig ML (1992) Species Diversity Gradients: We Know More and Less Than We Thought. J Mammal 73:715–730. doi: 10.2307/1382191

Ross AG, Vickers D, Olds GR, et al (2007) Katayama syndrome. Lancet Infect. Dis. 7:218–224

Rouas-Freiss N, Goncalves RM-B, Menier C, et al (1997) Direct evidence to support the role of HLA-G in protecting the fetus from maternal uterine natural killer cytolysis. Proc Natl Acad Sci 94:11520–11525. doi: 10.1073/pnas.94.21.11520

Rousseau P, Le Discorde M, Mouillot G, et al (2003) The 14 bp Deletion-Insertion Polymorphism in the 3′ UT Region of the HLA-G Gene Influences HLA-G mRNA Stability. In: Human Immunology. pp 1005–1010

Rts W (2016) RTS , S (Mosquir ix ^{TM}) Frequently Asked Questions (FAQs). 1–5

Sabbagh A, Courtin D, Milet J, et al (2013) Association of HLA-G 3′ untranslated region polymorphisms with antibody response against Plasmodium falciparum antigens: Preliminary results. Tissue Antigens. doi: 10.1111/tan.12140

Sabbagh A, Luisi P, Castelli EC, et al (2014) Worldwide genetic variation at the 3′ untranslated region of the HLA-G gene: Balancing selection influencing genetic diversity. Genes Immun 15:95–106. doi: 10.1038/gene.2013.67

Sabbagh A, Sonon P, Sadissou I, et al (2018) The role of HLA-G in parasitic diseases. Hla 1–16. doi: 10.1111/tan.13196

Sadissou I, d'Almeida T, Cottrell G, et al (2014) High plasma levels of HLA-G are associated with low birth weight and with an increased risk of malaria in infancy. Malar J 13:312. doi: 10.1186/1475-2875-13-312

Sadun EH, Von Lichtenberg F, Cheever AW, et al (1970) Experimental infection with Schistosoma haematobium in chimpanzees. Am J Trop Med Hyg 19:427–458

Sampangi S, Wang X, Beagley KW, et al (2015) Human proximal tubule epithelial cells modulate autologous B-cell function. Nephrol Dial Transplant 30:1674–1683. doi: 10.1093/ndt/gfv242

Sangweme DT, Midzi N, Zinyowera-Mutapuri S, et al (2010) Impact of schistosome infection on plasmodium falciparum malariometric indices and immune correlates in school age children in burma valley, zimbabwe. PLoS Negl Trop Dis 4:e882. doi: 10.1371/journal.pntd.0000882

Sanogo B, Yuan D, Zeng X, et al (2018) Diversity and Compatibility of Human Schistosomes and Their Intermediate Snail Hosts. Trends Parasitol 0:. doi: 10.1016/j.pt.2018.03.004

Santos KE, Lima THA, Felício LP, et al (2013) Insights on the HLA-G evolutionary history provided by a nearby alu insertion. Mol Biol Evol 30:2423–2434. doi: 10.1093/molbev/mst142

Sartelet H, Schleiermacher D, Le-Hesran JY, et al (2005) Less HLA-G expression in plasmodium falciparum-infected third trimester placentas is associated with more natural killer cells. Placenta 26:505–511. doi: 10.1016/j.placenta.2004.08.006

Satayathum SA, Muchiri EM, Ouma JH, et al (2006) Factors affecting infection or reinfection with Schistosoma haematobium in Coastal Kenya: Survival analysis during a nine-year, school-based treatment program. Am J Trop Med Hyg 75:83–92. doi: 75/1/83 [pii]

Schipper J, Chanson JS, Chiozza F, Al. E (2008) The Status of the World ' s Land and Marine Mammals: Diversity, Threat, and Knowledge. Science (80-) 322:225–230. doi: 10.1126/science.1165115

Scholte RGC, Gosoniu L, Malone JB, et al (2014) Predictive risk mapping of schistosomiasis in brazil using bayesian geostatistical models. Acta Trop 132:57–63. doi: 10.1016/j.actatropica.2013.12.007

Schramm G, Mohrs K, Wodrich M, et al (2007) Cutting Edge: IPSE/alpha-1, a Glycoprotein from Schistosoma mansoni Eggs, Induces IgE-Dependent, Antigen-Independent IL-4 Production by Murine Basophils In Vivo. J Immunol 178:6023–6027. doi: 10.4049/jimmunol.178.10.6023

Schwartz BS, Mawhorter SD (2013) Parasitic infections in solid organ transplantation. Am J Transplant 13:280–303. doi: 10.1111/ajt.12120

Sharma S, DeOliveira RB, Kalantari P, et al (2011) Innate Immune Recognition of an AT-Rich Stem-Loop DNA Motif in the Plasmodium falciparum Genome. Immunity 35:194–207. doi: 10.1016/j.immuni.2011.05.016

Shi X, Xie X, Jia Y, Li S (2017) Maternal genetic polymorphisms and unexplained recurrent miscarriage: a systematic review and meta-analysis. Clin. Genet.

Shiina T, Suzuki S, Ozaki Y, et al (2012) Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. Tissue Antigens 80:305–316. doi: 10.1111/j.1399-0039.2012.01941.x

Shobu T, Sageshima N, Tokui H, et al (2006) The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. J Reprod Immunol 72:18–32. doi: 10.1016/j.jri.2006.02.001

Siddiqui AA, Siddiqui SZ (2017) Sm-p80-Based Schistosomiasis Vaccine: Preparation for Human Clinical Trials. Trends Parasitol 33:194–201. doi: 10.1016/j.pt.2016.10.010

Sizzano F, Testi M, Zito L, et al (2012) Genotypes and haplotypes in the 3′ untranslated region of the HLA-G gene and their association with clinical outcome of hematopoietic stem cell transplantation for beta-thalassemia. Tissue Antigens 79:326–332. doi: 10.1111/j.1399-0039.2012.01862.x

Slater H, Michael E (2012) Predicting the current and future potential distributions of lymphatic filariasis in africa using maximum entropy ecological niche modelling. PLoS One 7:. doi: 10.1371/journal.pone.0032202

Sogoba N, Vounatsou P, Doumbia S, et al (2007) Spatial analysis of malaria transmission parameters in the rice cultivation area of Office du Niger, Mali. Am J Trop Med Hyg 76:1009–1015. doi: 76/6/1009 [pii]

Sokolow SH, Jones IJ, Jocque M, et al (2017) Nearly 400 million people are at higher risk of schistosomiasis because dams block the migration of snail-eating river prawns. Philos Trans R Soc B Biol Sci 372:20160127. doi: 10.1098/rstb.2016.0127

Sokolow SH, Wood CL, Jones IJ, et al (2018) To Reduce the Global Burden of Human Schistosomiasis, Use 'Old Fashioned' Snail Control. Trends Parasitol 34:23–40. doi: 10.1016/j.pt.2017.10.002

Solomon S, Plattner G-K, Knutti R, Friedlingstein P (2009) Irreversible climate change due to carbon dioxide emissions. Proc Natl Acad Sci U S A 106:1704–9. doi: 10.1073/pnas.0812721106

Southgate VR, van Wijk HB, Wright CA (1976) Schistosomiasis at Loum, Cameroun; Schistosoma haematobium, S. intercalatum and their natural hybrid. Zeitschrift für Parasitenkd 49:145–159. doi: 10.1007/BF00382422

Sponaas AM, Do Rosario APF, Voisine C, et al (2009) Migrating monocytes recruited to the spleen play an important role in control of blood stage malaria. Blood 114:5522–5531. doi: 10.1182/blood-2009-04-217489

Steinmann P, Keiser J, Bos R, et al (2006) Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. Lancet Infect Dis 6:411–425. doi: 10.1016/S1473-3099(06)70521-7

Stephens M, Donnelly P (2003) A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. Am J Hum Genet 73:1162–1169. doi: 10.1086/379378

Stephens M, Smith NJ, Donnelly P (2001) A New Statistical Method for Haplotype Reconstruction from Population Data. Am J Hum Genet 68:978–989. doi: 10.1086/319501

Stephenson L (1993) The impact of schistosomiasis on human nutrition. Parasitology 107:S107–S123. doi: 10.1017/S0031182000075545

Sullivan LC, Clements CS, Rossjohn J, Brooks AG (2008) The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. Tissue Antigens 72:415–424

Sutherland CJ, Tanomsing N, Nolder D, et al (2010) Two Nonrecombining Sympatric Forms of the Human Malaria Parasite *Plasmodium ovale* Occur Globally. J Infect Dis 201:1544–1550. doi: 10.1086/652240

Sutherland GR, Baker E, Callen DF, et al (1988) Interleukin 4 is at 5q31 and interleukin 6 is at 7p15. Hum Genet 79:335–337. doi: 10.1007/BF00282171

Swartz JM, Dyer KD, Cheever AW, et al (2006) Schistosoma mansoni infection in eosinophil lineage-ablated mice. Blood 108:2420–2427. doi: 10.1182/blood-2006-04-015933

Tamouza R, Busson M, Fortier C, et al (2007) HLA-E*0101 allele in homozygous state favors severe bacterial infections in sickle cell anemia. Hum Immunol 68:849–853. doi: 10.1016/j.humimm.2007.08.260

Tan Z, Randall G, Fan J, et al (2008) Allele-Specific Targeting of microRNAs to HLA-G and Risk of Asthma (PII:S0002-9297(07)63059-6). Am J Hum Genet 82:251. doi: 10.1016/j.ajhg.2007.12.008

Tan Z, Shon AM, Ober C (2005) Evidence of balancing selection at the HLA-G promoter region. Hum Mol Genet 14:3619–3628. doi: 10.1093/hmg/ddi389

Tchuem Tchuenté LA, Southgate VR, Jourdane J, et al (2003) Schistosoma intercalatum: An endangered species in Cameroon? Trends Parasitol. 19:389–393

Teixeira-Carvalho A, Martins-Filho OA, Peruhype-Magalhaes V, et al (2008) Cytokines, chemokine receptors, CD4+CD25HIGH+ T-cells and clinical forms of human schistosomiasis. Acta Trop 108:139–149. doi: S0001-706X(08)00093-4 [pii]\r10.1016/j.actatropica.2008.04.010

Terrenato L, Shrestha S, Dixit KA, et al (1988) Decreased malaria morbidity in the Tharu people compared to sympatric populations in Nepal. Ann Trop Med Parasitol 82:1–11. doi: 10.1080/00034983.1988.11812202

Thomas B, Short E (2017) Climate Change Contribution to the Emergence or Re-Emergence of Parasitic Diseases. Infect Dis Res Treat 10:. doi: 10.1177/1178633617732296

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform 14:178–192. doi: 10.1093/bib/bbs017

Utzinger J, Bergquist R, Olveda R, Zhou XN (2010) Important Helminth Infections in Southeast Asia. Diversity, Potential for Control and Prospects for Elimination. Adv Parasitol 72:1–30. doi: 10.1016/S0065-308X(10)72001-7

Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma. doi: 10.1002/0471250953.bi1110s43

Veiga-Castelli LC, Bertuol JM da SB, Castelli EC, Donadi EA (2016) Low variability at the

HLA-E promoter region in the Brazilian population. Hum Immunol 77:172–5. doi: 10.1016/j.humimm.2015.11.014

Veiga-Castelli LC, Castelli EC, Mendes CT, et al (2012) Non-classical HLA-E gene variability in Brazilians: A nearly invariable locus surrounded by the most variable genes in the human genome. Tissue Antigens 79:15–24. doi: 10.1111/j.1399-0039.2011.01801.x

Veit TD, Cazarolli J, Salzano FM, et al (2012) New evidence for balancing selection at the HLA-G locus in South Amerindians. Genet Mol Biol 35:919–923. doi: 10.1590/S1415-47572012000600005

Veit TD, Chies JAB (2009) Tolerance versus immune response - MicroRNAs as important elements in the regulation of the HLA-G gene expression. Transpl Immunol 20:229–231. doi: 10.1016/j.trim.2008.11.001

Veit TD, De Lima CPS, Cavalheiro LC, et al (2014) HLA-G +3142 polymorphism as a susceptibility marker in two rheumatoid arthritis populations in Brazil. Tissue Antigens 83:260–266. doi: 10.1111/tan.12311

Verwaerde C, Joseph M, Capron M, et al (1987) Functional properties of a rat monoclonal IgE antibody specific for Schistosoma mansoni. J Immunol 138:4441–4446

Visser LG, Polderman AM, Stuiver PC (1995) Outbreak of schistosomiasis among travelers returning from Mali, West Africa. Clin Infect Dis 20:280–285. doi: 10.1093/clinids/20.2.280

Von Lichtenberg F, Edington GM, Nwabuebo I, et al (1971) Pathologic effects of schistomiasis in Ibadan Western State of Nigeria. II. Pathogenesis of lesions of the bladder and ureters. Am J Trop Med Hyg 20:244–254. doi: 10.4269/ajtmh.1971.20.244

Walther M, Woodruff J, Edele F, et al (2006) Innate Immune Responses to Human Malaria: Heterogeneous Cytokine Responses to Blood-Stage Plasmodium falciparum Correlate with Parasitological and Clinical Outcomes. J Immunol 177:5736–5745. doi: 10.4049/jimmunol.177.8.5736

Wang W, Wei Z, Lam TW, Wang J (2011) Next generation sequencing has lower sequence coverage and poorer SNP-detection capability in the regulatory regions. Sci. Rep. 1

Wang X, Jiang W, Zhang D (2013) Association of 14-bp insertion/deletion polymorphism of HLA-G gene with unexplained recurrent spontaneous abortion: A meta-analysis. Tissue Antigens. doi: 10.1111/tan.12056

Webster BL, Southgate VR, Tchuem Tchuenté L-A (2003) Isoenzyme analysis of Schistosoma haematobium, S. intercalatum and their hybrids and occurrences of natural hybridization in Cameroon. J Helminthol 77:269–74. doi: 10.1079/JOH2003166

Webster BL, Tchuem Tchuenté LA, Jourdane J, Southgate VR (2005) The interaction of <I>Schistosoma haematobium</I> and <I>S. guineensis</I> in Cameroon. J Helminthol 79:193–197. doi: 10.1079/JOH2005306

Wei X, Orr HT (1990) Differential expression of HLA-E, HLA-F, and HLA-G transcripts in human tissue. Hum Immunol 29:131–142. doi: 10.1016/0198-8859(90)90076-2

Weyher AH, Phillips-Conroy JE, Fischer K, et al (2010) Molecular Identification of Schistosoma mattheei from Feces of Kinda (Papio cynocephalus kindae) and Grayfoot Baboons (Papio ursinus griseipes) in Zambia. J Parasitol 96:184–190. doi: 10.1645/GE-2186.1

Whitty CJ, Mabey DC, Armstrong M, et al (2000) Presentation and outcome of 1107 cases of schistosomiasis from Africa diagnosed in a non-endemic country. Trans R Soc Trop Med Hyg 94:531–534. doi: 10.1016/S0035-9203(00)90077-4

WHO (2017) WHO | World malaria report 2017

WHO (2001) World Health Assembly resolution (WHA 54.19) Schistosomiasis and Soil transmitted helminth infections. Wkly Epidemiol Rec 3–5

WHO (2011) Helminth control in school-age children. A guide for managers of control programmes. Geneva World … 90

WHO (2006) Preventive chemotherapy in human helminthiasis: coordinated use of anthelminthic drugs in control interventions: a manual for health professionals and programme managers.

WHO (2015) For the treatment of malaria

Wilson MS, Mentink-Kane MM, Pesce JT, et al (2007) Immunopathology of schistosomiasis

Wiśniewski A, Biliń M, Klimczak A, et al (2010) Association of the HLA-G gene polymorphism with multiple sclerosis in a Polish population. Int J Immunogenet 37:307–311. doi: 10.1111/j.1744-313X.2010.00926.x

Xu Y, Han H, Zhang F, et al (2015) Lesion human leukocyte antigen-F expression is associated with a poor prognosis in patients with hepatocellular carcinoma. Oncol Lett 9:300–304. doi: 10.3892/ol.2014.2686

Y?? Y, Kyobutungi C, Louis VR, Sauerborn R (2007) Micro-epidemiology of Plasmodium falciparum malaria: Is there any difference in transmission risk between neighbouring villages? Malar J 6:. doi: 10.1186/1475-2875-6-46

Zaghloul MS (2012) Bladder cancer and schistosomiasis. J Egypt Natl Canc Inst 24:151–159. doi: 10.1016/j.jnci.2012.08.002

Zambra FMB, Biolchi V, de Cerqueira CCS, et al (2016) Immunogenetics of prostate cancer and benign hyperplasia--the potential use of an HLA-G variant as a tag SNP for prostate cancer risk. HLA 87:79–88. doi: 10.1111/tan.12741

Zhang J, Pan L, Chen L, et al (2012) Non-classical MHC-Iota genes in chronic hepatitis B and hepatocellular carcinoma. Immunogenetics 64:251–258. doi: 10.1007/s00251-011-0580-2

Zhang X, Li S, Zhang Y, et al (2014) Meta-analysis of the relationship between 14bp insertion/deletion polymorphism of HLA-G gene and susceptibility to systemic lupus erythematosus. Hum. Immunol.

Zheng H, Tan Z, Xu W (2014) Immune evasion strategies of pre-erythrocytic malaria parasites. Mediators Inflamm. 2014

(2002) Prevention and control of schistosomiasis and soil-transmitted helminthiasis. World Health Organ Tech Rep Ser 912:. doi: 10.1016/j.parint.2004.01.001

ANNEX

ANNEX I

Proof of biorepository.

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

Ao Comitê de Ética em Pesquisa do HCFMRP-USP e da FMRP-USP

Vem-se por meio desta solicitar a autorização do Comitê de Ética em Pesquisa do HCFMRP-USP e da FMRP-USP para a criação de um biorrepositório denominado: **AMOSTRAS PALUCO**. Esse biorrepositório será de responsabilidade do (s) **Paulin Sonon**, doutorando em Imunologia Básica e Aplicada, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brasil e do **Prof. Dr. Eduardo Antônio Donadi,** Professor da Divisão de Imunologia Clínica do Departamento de Medicina da Faculdade de Medicina de Ribeirão Preto - Universidade de São Paulo, Brasil, vinculado (s) ao Departamento de Bioquímica e Imunologia FMRP desta instituição, estando o referido Departamento ciente e de acordo com a criação.

O biorrepositório denominado **"AMOSTRAS PALUCO"** tem como objetivo o armazenamento de **duzentos e noventa e cinco (295) amostras do DNA do projeto PALUCO** para fins de pesquisa e análise científica, visto que **essas amostras estão necessárias para a realização do sequenciamento** vinculado ao projeto de pesquisa denominado **"COINFECÇÃO *PLASMODIUM FALCIPARUM* E *SCHISTOSOMA HAEMATOBIUM*: PAPEL DOS GENES *HLA (-A, -B, -C, -E, -F* e *-G)* NA SUSCETIBILIDADE À MALÁRIA"**.

Tal armazenamento será realizado da seguinte forma: **no laboratório da Biologia Molecular do Prof. Dr. Eduardo Antônio Donadi no freezer -20°C**.

Fica sob a responsabilidade do **Paulin Sonon e do Prof. Dr. Eduardo Antônio Donadi** o armazenamento, a guarda e o uso do material estocado. E se compromete (m) garantirem o sigilo e a confidencialidade dos doadores que já foram identificadas por meio de **números (1 a 295) ou códigos combinando o nome do centro do estudo, um número crescente a partir de 2013 (primeira amostra) a 2116 (última amostra), estado clínico**

do participante (2 = infetados, 0 ou 1 non-infetados), natureza da amostra e PV); ex: SZ2013-2/BC-PV2, SZ2013-0/BC-PV2, SZ2013-1/BC-PV2 (primeiras amostras) e SM2116-2/BC-PV2, SM2116-0/BC-PV2, SM2116-1/BC-PV2 (últimas amostras).

Paulin Sonon

Identificação e qualificação do responsável - doutorando

Ribeirão Preto, 10 de dezembro de 2015

Eduardo A. Donadi
Imunologia Clínica
CRM 28795

Identificação e qualificação do responsável - orientador

Ribeirão Preto, 10 de dezembro de 2015

Ciência e concordância da chefia do Departamento:

Identificação e qualificação do responsável

Prof. Dr. Jorge Elias Jr.
Suplente da Chefia do Depto de Clínica Médica

Ribeirão Preto, 11 de DEZEMBRO de 2015

ANNEX II

Letter to request the dispensation of ''free and informed consent'' from Brazil "Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto-Universidade de São Paulo" (HCRP e FMRP-USP)'', since the major project (PALUCO) had been already proved in Benin and the samples treated here were from PALUCO project.

Ribeirão Preto, dia 10 de dezembro de 2015

Paulin Sonon, Doutorando em Imunologia Básica e Aplicada
Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo
Email: paulinsonon@usp.br
Tél: +55 (16) 981810882
Tel laboratório : +55 (16) 3315-3373

A(o)

Senhor(a) Coordenadora do Comitê
de Ética em Pesquisa (CEP) do Hospital das
Clínicas da Faculdade de Medicina de
Ribeirão Preto

Solicitação de dispensa do Termo de Consentimento Livre e Esclarecido (TCLE)

Prezada Coordenadora,

Venho por meio desta solicitar a dispensa do Termo de Consentimento Livre e Esclarecido (TCLE) do projeto denominado **"COINFECÇÃO *PLASMODIUM FALCIPARUM* E *SCHISTOSOMA HAEMATOBIUM*: PAPEL DOS GENES *HLA (-A, -B, -C, -E, -F* e *-G)* NA SUSCETIBILIDADE À MALÁRIA".**

De fato, esse projeto é um componente genético de um projeto maior, denominado **PALUCO (malária e coinfecção).** O projeto **PALUCO** foi financiado em 2012, avaliado e aprovado pelo Comitê de Ética da Faculdade de Ciências da Saúde de Cotonou (capital econômica do Benin) sob o n°: 12/03/2012/CE/FSS/UAC (Carta do CEP do PALUCO anexado na Plataforma Brasil). Todo o recrutamento foi feito após assinatura do TCLE (TCLE PALUCO anexado na Plataforma Brasil).

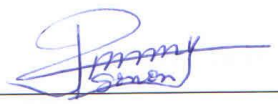
Assim, duzentos e noventa e cinco (295) amostras do DNA do projeto PALUCO foram transferidas para o Brasil e armazenadas no laboratório de Biologia Molecular do **Prof. Dr. Eduardo Antônio Donadi** (Carta de criação de biorrepositório anexada na Plataforma Brasil). As mesmas amostras serão utilizadas no quadro do projeto **"COINFECÇÃO *PLASMODIUM FALCIPARUM* E *SCHISTOSOMA HAEMATOBIUM*: PAPEL DOS GENES *HLA (-A, -B, -C, -E, -F* e *-G)* NA SUSCETIBILIDADE À MALÁRIA"** para o sequenciamento dos genes candidatos selecionados *HLA (-A, -B, -C, -E, -F* e *-G)*. Torna-se inviável o recrutamento e aplicação do TCLE novamente para os participantes que se encontram em outro pais – Benin:

Para esta nova pesquisa serão garantidos o sigilo e a confidencialidade dos dados.

Sem outro particular subscrevo-me colocando à disposição para quaisquer esclarecimentos.

Atenciosamente

Paulin Sonon



Paulin Sonon
Doutorando Direto, Imunologia Básica e Aplicada, Faculdade de Medicina de Ribeirão Preto,
Universidade de São Paulo, Brasil
Departamento de Bioquímica e Imunologia FMRP

ANNEX III

Written and informed consent: French version (original).

Université Nationale d'Abomey-Calavi
Laboratoire de l'institut de recherche pour le développement – Laboratoire d'immunologie des maladies infectieuses et allergiques
PROJET PALUCO

Formulaire de consentement éclairé.

Identité du représentant légal

Nom et prénom……

Lien de parenté ……

N° de tel. :

Identité de l'enfant

Nom et prénom……..

Adresse de la Personne/structure de référence :

Attestation de consentement (<J'> ici représente le participant ou son représentant)

J'ai lu (ou j'ai fait lire et interpréter les informations concernant l'étude du paludisme et de la schistosomose) et j'ai compris les informations sur le but de l'étude portant sur**:**

« Paludisme et coïnfection ou PALUCO»

J'ai eu l'occasion de poser toutes les questions avec les membres de l'équipe de recherche. Les réponses m'ont été fournies dans un langage que je comprends. J'ai en effet compris que:

- ma participation à l'étude est volontaire et que je peux me retirer à tout moment.
- les inconvénients et les avantages du fait de ma participation à cette étude
- les données qui me concernent seront gardées de manière entièrement anonyme, et je n'autorise leur consultation que par les personnes qui collaborent à cette recherche sous la responsabilité des investigateurs
- les chercheurs impliqués dans cette étude, aient accès aux données qui me concernent dans le respect de la plus stricte confidentialité.
- les informations collectées pourront être publiées dans des revues scientifiques ou présentées lors des réunions scientifiques.
- les dossiers de recherche pourraient être inspectés par le comité d'éthique du Bénin pour s'assurer du bon déroulement de l'étude.
- Les restes d'échantillons s'il y en a, seront préservés et pourraient être utilisés dans des études ultérieures après accord préalable d'un comité d'éthique.
- **Le projet prend en charge chaque participant en cas d'une manifestation de toux, de paludisme avec une somme maximale de 15000 francs CFA dans les cas graves.**
- **Le projet prend en charge le traitement en cas de schistosomose et les médicaments sont strictement pris au centre de santé Saint Joseph qui collabore avec les chercheurs sur cette étude**.

Mon consentement ne décharge pas les organisateurs de la recherche de leur responsabilité et je conserve tous mes droits tels que garantis par la loi.

Le répondant accepte de participer à cette étude de recherche et de répondre aux questions *(Répondant ici signifie le participant ou son représentant qui répond en son nom)*

OUI ☐ NON ☐

Date

Signature

Je soussigné, agissant en tant que témoin, atteste que l'interviewé(e) a compris et a donné son consentement pour participer à l'étude.

Date

Signature

Je soussigné avoir expliqué tous les détails de l'étude au participant indiqué plus haut et atteste qu'il a compris et donne son consentement.

Nom de l'enquêteur

Date

Signature de l'enquêteur

ANNEX III

Written and informed consent: Translated version (Portuguese).

Universidade Nacional de Abomey-Calavi
Laboratório do Instituto de Pesquisa para o Desenvolvimento - Laboratório de Imunologia de doenças infecciosas e alérgicas
PROJETO PALUCO

Formulário de Consentimento Livre e Esclarecido **(Versão original em francês)**

Identidade do representante legal

Nome completo……

Relacionamento parental……

N° de tel. :.....

Identidade da criança

Nome completo ……

Endereço da pessoa/estrutura de referência: .....

.....

Certificação de consentimento (<J'> aqui representa o participante ou seu representante)

Eu li (ou eu fiz ler e interpretar informações sobre o estudo da malária e esquistossomose) e eu entendi as informações sobre a finalidade do estudo:

«Malária e Coïnfection ou PALUCO»

Eu tive a oportunidade de fazer todas as perguntas com os membros da equipe de pesquisa. As respostas foram fornecidas para mim em uma língua que eu entendo. Na verdade, eu entendo que:

- minha participação no estudo é voluntária e que eu posso retirar a qualquer momento.
- vantagens e desvantagens da minha participação neste estudo.
- os meus dados serão mantidos completamente anônimos, e posso os consultar somente por intermediário das pessoas que colaboram nesta pesquisa sob a responsabilidade dos investigadores.
- os pesquisadores envolvidos neste estudo, têm acesso a meus dados com a mais estrita confidencialidade.
- as informações recolhidas poderão serem publicadas em revistas científicas ou apresentadas em reuniões científicas.
- os registros da pesquisa poderiam serem consultados por comitê de ética do Benin para garantir o bom andamento do estudo.
- as demais amostras, se houver, serão preservadas e poderiam serem usadas em estudos posteriores após aprovação prévia de uma comissão de ética.
- **o projeto apoia cada participante no caso de tosse, malária, com um montante máximo de 15 mil CFA nos casos graves.**
- **o projeto apoia o tratamento da esquistossomose e as drogas são estritamente tomadas no Centro de Saúde Saint Joseph, que trabalha em conjunto com os investigadores deste estudo.**

Meu consentimento não exime os organizadores da pesquisa das suas responsabilidades e eu mantenho todos meus direitos como garantidos por lei.

O entrevistado aceita em participar neste estudo de pesquisa e em responder às perguntas (o entrevistado aqui é o participante ou o seu representante)

SIM ☐ NÃO ☐

Data

Assinatura

Eu, abaixo assinado, como testemunha certifico que o(a) entrevistado(a) entendeu e consentiu em participar do estudo.

Data

Assinatura

Eu, abaixo assinado, certifico ter explicado todos os detalhes do estudo ao participante em acima referido e que ele entendeu e consentiu.

Nome do entrevistador

Data

Assinatura do entrevistador

ANNEX IV

Certificate from Ethics Committee of the ''Faculté des Sciences et de la Santé (FSS)'' of Cotonou, Benin, aproving the major project (PALUCO). **No.12/03/2012/CEIFSS/UAC, origginal version (French).**

UNIVERSITE d'ABOMEY - CALAVI
FACULTE DES SCIENCES DE LA SANTE
****
COMITE D'ETHIQUE

01 B.P 188 COTONOU REP. DU BENIN Tél.: (229) 21302513 Fax : (229) 21304096

Le Président du Comité d'Ethique
N° : 12/03/2012/CE/FSS/UAC

AVIS SUR LE PROTOCOLE « PALUDISME ET CO-INFECTION (PALUCO) »

Le Comité d'Ethique de la Faculté des Sciences de la Santé réuni en session restreinte le 05 mars 2012, a étudié la demande d'avis éthique du projet de recherche intitulé « Paludisme et co-infection (PALUCO) ».

Ce projet a pour objectif d'explorer le rôle de l'infection par *S. haematobium* sur la sensibilité au paludisme dans la commune de Sô-Ava dans le sud du Bénin. Il s'agira de réaliser un suivi parasitologique et clinique de 650 enfants âgés de 4 à 8 ans exposés ou non à la schistosomose durant la saison de transmission palustre et d'estimer l'impact de l'infection par *S. haematobium* sur la fréquence et le niveau d'infection par *P. falciparum*, sur le délai d'apparition et la fréquence des accès palustres et sur les niveaux de cytokines, de HLA-G et d'anticorps anti-palustres dirigés contre les antigènes candidats vaccins les plus prometteurs (AMA-1, MSP1, MSP2, MSP3, GLURP et CSP).

Les documents suivants ont été examinés et approuvés par le Comité d'Ethique :

- *le protocole, (version française du 12 décembre 2011)*
- *le cahier de collecte des données, (version française du 12 décembre 2011)*
- *la brochure de l'investigateur, (version française du 12 décembre 2011)*
- *le document de recueil du consentement éclairé, (version française du 11 janvier 2012)*

Le Comité d'Ethique de la Faculté des Sciences de la Santé a évalué les aspects scientifiques et éthiques de votre protocole de recherche conformément aux normes nationales en vigueur et vous donne un avis favorable au projet de recherche PALUCO.

Le Comité vous remercie de bien vouloir le tenir informé de l'issue de vos travaux.

Par ailleurs, le comité d'éthique de la Faculté des Sciences de la Santé vous demande de l'informer de toute nouvelle information qui surviendrait à une date ultérieure à cette approbation et qui impliquerait des changements dans le déroulement de cette recherche. La Faculté des Sciences de la Santé doit, en effet, dans ces cas, réévaluer et donner une nouvelle approbation avant l'entrée en vigueur desdites modifications;

Vous devez conserver dans vos dossiers, la version originale des consentements signés par les participants de recherche ou leurs témoins;

Cotonou, le 05 mars 2012

Le Président

Pr MEDJI Agbé P. Léon

ANNEX IV

Certificate from Ethics Committee of the ‘’Faculté des Sciences et de la Santé (FSS)’’ of Cotonou, Benin, aproving the major project PALUCO). **No.12/03/2012/CEIFSS/UAC, translated version (Portuguese).**

UNIVERSIDADE DE ABOMEY- CALAVI (BENIN)

FACULDADE DE CIÊNCIAS DA SAÚDE

****

COMITÊ DE ÉTICA

01 B.P 188 COTONOU REP. DU BENIN Tél.: (229) 21302513 Fax: (229) 21304096

Presidente do comitê de Ética

Núm : 12/03/2012/CE/FSS/UAC

PARECER SOBRE O PROTOCOLO "MALÁRIA E CO-INFECÇÃO (PALUCO)

O Comitê de Ética da Faculdade de Ciências da Saúde reunido na sessão restringida no dia 05 de março de 2012, avaliou o pedido de revisão ética do projeto de pesquisa intitulado **"Malária e Co-infecção (PALUCO)"**.

Este projeto tem como objetivo de explorar o papel da infecção por *S. haematobium* na susceptibilidade à malária na região de Sô-Ava no sul do Benin. O projeto realizará um acompanhamento parasitológico e clínico de 650 crianças de 4 a 8 anos de idade expostos ou não a esquistossomose durante o período de transmissão da malária e de estimar o impacto da infecção por *S. haematobium* sobre a frequência e o nível de infecção por *P. falciparum*, sobre o tempo de ocorrência do primeiro caso da malária e a frequência de ataques de malária e sobre os níveis de citocinas, do HLA-G e de anticorpos anti-malária dirigidos contra os antigénios candidatos a vacinas mais promissores (AMA-1, MSP1, MSP2, MSP3, GLURP e CSP).

Os seguintes documentos foram revistos e aprovados pela Comissão de Ética:

- protocolo (versão francesa de 12 de dezembro de 2011)
- caderna da recolha de dados, (versão francesa de 12 de dezembro de 2011)
- brochura do investigador (versão francesa de 12 de dezembro de 2011)
- documento de recolha de consentimento esclarecido (versão francesa de 11 de janeiro de 2012)

O Comitê de Ética da Faculdade de Ciências da Saúde avaliou os aspectos científicos e éticos do seu protocolo da pesquisa seguindo as normas nacionais em vigor e dá-lhe uma opinião favorável ao projeto de pesquisa PALUCO.

O Comitê vos agradece de mantê-lo informado sobre os resultados do seu trabalho.

Além disso, o Comitê de Ética da Faculdade de Ciências da Saúde pede-lhe para informá-lo de qualquer nova informação que poderia ocorrer em uma data posterior dessa aprovação e que envolveria das mudanças no curso desta pesquisa. Nesses casos, de fato, a Faculdade de Ciências da Saúde tem que re-avaliar e dar uma nova aprovação antes da entrada em vigor dessas alterações;

Vocês devem manter nos seus registros, a versão original dos consentimentos assinados pelos participantes da pesquisa ou suas testemunhas;

Feito a Cotonou, dia 05 de março de 2012

Presidente

Versão2 do certidão CE PALUCO do 23/02/2016

ANNEX V

Certificate from the "Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto-Universidade de São Paulo" (HCRP e FMRP-USP), aproving my doctorate project. No.CEP: FMRP-N°710/2016/CEP/MGV.

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

FMRP-USP
RIBEIRÃO PRETO

Ribeirão Preto, 02 de março de 2016

Ofício nº 710/2016
CEP/MGV

Prezados Senhores,

O trabalho intitulado **"COINFECÇÃO PLASMODIUM FALCIPARUM E SCHISTOSOMA HAEMATOBIUM: PAPEL DOS GENES HLA (-A, -B, -C, -E, -F e –G) NA SUSCETIBILIDADE À MALÁRIA"** – versão 2, de 23/02/2016, foi analisado pelo Comitê de Ética em Pesquisa, em sua 422ª Reunião Ordinária realizada em 29/02/2016 e enquadrado na categoria: APROVADO, bem como o Cronograma – versão 2, de 23/02/2016, o Biorrepositório e a solicitação de dispensa do Termo de Consentimento Livre e Esclarecido, de acordo com o Processo HCRP nº 17850/2015.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 466/12 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.

DRª. MARCIA GUIMARÃES VILLANOVA
Coordenadora do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores
PAULIN SONON
PROF.DR.EDUARDO ANTÔNIO DONADI(Orientador)
Depto. de Clínica Médica

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO
Campus Universitário – Monte Alegre
14048-900 Ribeirão Preto SP
Comitê de Ética em Pesquisa do HCRP e FMRP-USP
FWA-00002733; IRB-00002186 e Registro PB/CONEP nº 5440
(16)3602-2228/cep@hcrp.usp.br
www.hcrp.usp.br

ANNEX VI

Paper related to the doctorate project published during the doctorate course

Received: 6 December 2017 | Accepted: 20 December 2017

DOI: 10.1111/tan.13196

WILEY HLA Immune Response Genetics

REVIEW ARTICLE

The role of HLA-G in parasitic diseases

A. Sabbagh[1] | P. Sonon[2] | I. Sadissou[2] | C. T. Mendes-Junior[3] | A. Garcia[1,4] | E. A. Donadi[2] | D. Courtin[1]

[1]UMR 216 MERIT, Institut de Recherche pour le Développement, Faculté de Pharmacie de Paris, Université Paris Descartes, Paris, France

[2]Department of Medicine, School of Medicine of Ribeirão Preto, University of São Paulo, São Paulo, Brazil

[3]Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

[4]Centre d'Etude et de Recherche sur le Paludisme Associé à la Grossesse et à l'Enfance (CERPAGE), Faculté des Sciences de la Santé, Cotonou, Bénin

Correspondence
David Courtin, UMR 216 MERIT, IRD, Faculté de Pharmacie de Paris, Université Paris Descartes, 4 avenue de l'Observatoire, Paris 75006, France.
Email: david.courtin@ird.fr

Funding information
CNPq/Brazil; Brazil-France Research Cooperation Program USP/COFECUB, Grant/Award numbers: 304 931/2014-1, 467 157/2014-6, 309 572/2014-2, Uc Me 169-17

Little attention has been devoted to the role of HLA-G gene and molecule on parasitic disorders, and the available studies have focused on malaria, African and American trypanosomiasis, leishmaniosis, toxoplasmosis and echinococcosis. After reporting a brief description regarding the role of the cells of innate and adaptive immune system against parasites, we reviewed the major features of the HLA-G gene and molecule and the role of HLA-G on the major cells of immune system. Increased levels of soluble HLA-G (sHLA-G) have been observed in patients presenting toxoplasmosis and in the active phase of echinococcosis. In addition, increased sHLA-G has also been associated with increased susceptibility to malaria and increased susceptibility to develop human African trypanosomiasis (HAT). In contrast, decreased membrane-bound HLA-G has been reported in placenta of patients infected with *Plasmodium falciparum* and in heart and colon of patients presenting Chagas disease. The 3′ untranslated region of the *HLA-G* gene has been the main focus of studies on malaria, HAT and Chagas disease, exhibiting distinct patterns of associations. Considering that HLA-G is an immune checkpoint molecule, inhibiting the activity of several cells of the immune system, the excessive neoexpression and the increased sHLA-G levels together with the decreased constitutive tissue expression of membrane-bound HLA-G may be detrimental to the host infected with parasite agents.

KEYWORDS
echinococcosis, genetics, HLA-G, immune system, leishmaniosis, malaria, parasite, toxoplasmosis, trypanosomiasis

ANNEX VI

Paper related to the doctorate project submitted and accepted

HLA-G, -E and *-F* regulatory and coding region variability and haplotypes in the Beninese Toffin population sample

Paulin Sonon[a], Ibrahim Sadissou[b], Léonidas Tokplonou[c,d,e], Kuumaaté K. G. M'po[c,f], Sonya S. C. Glitho[c,f], Privat Agniwo[c,f], Moudachirou Ibikounlé[e,f], Juliana Doblas Massaro[b], Achille Massougbodji[c], Philippe Moreau[g,h], Audrey Sabbagh[d], Celso T. Mendes-Junior[i], Kabirou A. Moutairou[j], Erick C. Castelli[k, l], David Courtin[d], Eduardo A. Donadi[b]

[a] Laboratório de Biologia Molecular. Universidade de São Paulo, Programa de Imunologia Básica e Aplicada (IBA), Faculdade de Medicina de Ribeirão Preto (FMRP-USP), Estado de São Paulo (SP), Brasil.

[b] Laboratório de Biologia Molecular. Universidade de São Paulo, Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (FMRP-USP), Estado de São Paulo (SP), Brasil.

[c] Centre d'Etude et de Recherche sur le Paludisme Associé à la Grossesse et à l'Enfance (CERPAGE), Faculté des Sciences de la Santé, Cotonou, Bénin.

[d] UMR 216 MERIT, IRD, Faculté de Pharmacie de Paris, Université Paris Descartes, Sorbonne Paris Cité, Paris, France.

[e] Université d'Abomey-Calavi, Cotonou, Bénin

[f] Département de Zoologie, Faculté des Sciences et Techniques, Université d'Abomey-Calavi, Cotonou, Bénin

[g] Commissariat à l'Energie Atomique et aux Energies Alternatives, Direction de la Recherche Fondamentale, Institut de Biologie François Jacob, Service de Recherches en Hémato-Immunologie, Hôpital Saint-Louis, Paris, France

[h] Université Paris-Diderot, Sorbonne Paris-Cité, UMR_E5, Institut Universitaire d'Hématologie, Hôpital Saint-Louis, Paris, France

[i] Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, State of São Paulo, Brazil.

[j] Laboratoire de Biologie et Physiologie Cellulaire, Université d'Abomey-Calavi, Cotonou, Bénin

[k] São Paulo State University (UNESP), Molecular Genetics and Bioinformatics Laboratory, Experimental Research Unit (UNIPEX), School of Medicine, Botucatu, State of São Paulo, Brazil

[l] São Paulo State University (UNESP), Department of Pathology, School of Medicine, Botucatu, State of São Paulo, Brazil

*Corresponding author at: Laboratório de Biologia Molecular. Universidade de São Paulo, Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (FMRP-USP), Estado de São Paulo (SP), Brasil. AV Bandeirantes, 3900, HC, Vila Monte Alegre, 14049900 - Ribeirão Preto, SP – Brasil. Phone: (+55) 16 3602-2566, E-mail address: eadonadi@fmrp.usp.br

Email addresses: paulinsonon@gmail.com (Paulin Sonon), sadissouai@gmail.com (Ibrahim Sadissou), tokplonouleonidas@gmail.com (Léonidas Tokplonou), kuumadox@yahoo.fr (Kuumaaté K. G. M'po), glithosonya@yahoo.fr (Sonya S. C. Glitho), privatagniwo@yahoo.com (Privat Agniwo), mibikounle2001@yahoo.fr (Moudachirou Ibikounlé), jmassaro@alumni.usp.br (Juliana Doblas Massaro), massougbodjiachille@yahoo.fr (Achille Massougbodji), philippe.moreau@cea.fr (Philippe Moreau), audrey.sabbagh@ird.fr (Audrey Sabbagh), ctmendes@gmail.com (Celso T. Mendes-Junior), kabirou.moutairou@gmail.com (Kabirou A. Moutairou), erick.castelli@unesp.br (Erick C. Castelli), david.courtin@ird.fr (David Courtin), eadonadi@fmrp.usp.br (Eduardo A. Donadi).

Abstract

HLA-G/E/F genes exhibit immunomodulatory properties and are expressed in placenta. Little attention has been devoted to the study of these genes in sub-Saharan African populations, which are yet the most diverse. To fill this gap, we evaluated the complete gene variability, approximately 5.1 kb for *HLA-G* (n=149), 7.7 kb for *HLA-E* (n=150) and 6.2 kb for *HLA-F* (n=152) in the remote Beninese Toffin population, using massive parallel sequencing. Overall, 96, 37 and 68 variable sites were detected along the entire *HLA-G, -E and -F,* respectively, arranged into region-specific haplotypes; i.e., promoter haplotypes (16, 19, and 15 respectively), coding haplotypes (19, 15, and 29 respectively), 3' untranslated region (3'UTR) haplotypes (12, 7 and 2, respectively) and extended haplotypes (33, 31 and 32 respectively). All promoter/coding/3'UTR haplotypes followed the patterns already described. *HLA-E* was the most conserved, exhibiting mainly two full-length encoded-molecules (E*01:01 and E*01:03), followed by *HLA-F*, three full-length proteins (F*01:01, F*01:02 and F*01:03) and *HLA-G,* four proteins: three full-length (G*01:01, G*01:03 and G*01:04) and one truncated (G*01:05N). Although *HLA-G/E/F* alleles in the Toffin population were the most frequently observed worldwide, the frequencies of the coding haplotypes were closely similar to those described for other African populations (Guinea-Conakry and Burkina-Faso), when compared to non-African ones (Brazilian), indicating that variable sites along these genes were present in Africa before human dispersion.

Keywords: HLA-E, *HLA-F*, *HLA-G*, haplotypes, massive parallel sequencing, African.

Abbreviations: HLA-G/E/F, Human leukocyte antigen *(G, E, F)*; UTR, Untranslated region; MHC, Human major histocompatibility complex; DCs, Dendritic cells; NK, Natural killer cells; ILT-(2, 4),

Ig-like transcript receptor 2 and 4; DNA, Deoxyribonucleic acid; PCR, Polymerase chain reaction; qPCR, Quantitative real-time PCR; BAM, Binary alignment map; IGV, Integrative genomics viewer; VCF, Variant call format; GATK, Genome analysis toolkit; LD, Linkage disequilibrium; MAF, Minor allele frequency; SNP, Single nucleotide polymorphism; STR, Short tandem repeat; *P. f, Plasmodium falciparum*; Kb, Kilobases (10^3 bases); mRNA, Messenger RNA (ribonucleic acid); SINE, short interspersed nuclear element

Decision Letter - Accept: 16 August 2018

Ref: MIMM_2018_50
Title: HLA-G, -E and -F regulatory and coding region variability and haplotypes in the Beninese Toffin population sample
Journal: Molecular Immunology

Dear Dr. Donadi,

I am pleased to inform you that your paper has been accepted for publication. My own comments as well as any reviewer comments are appended to the end of this letter. Now that your manuscript has been accepted for publication it will proceed to copy-editing and production.

Thank you for submitting your work to Molecular Immunology. We hope you consider us again for future submissions.

Kind regards,

Jose Villadangos
Editor-in-Chief
Molecular Immunology

Comments from the editors and reviewers:
- **Reviewer 1**

-

The manuscript Sonon et al is a well performed and well described study. It has a complete overview of the current literature on HLA-E,F and G and provide inportant new information. The discussion contains some repeats of earlier described findings making the paper lengthly but interesting. Lay out and organisation of the tables makes the data easy readable and figures and tables are well described in the text.

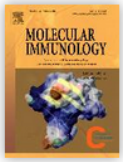
Have questions or need assistance?
For further assistance, please visit our Customer Support site. Here you can search for

Article:	**HLA-G, -E and -F regulatory and coding region variability and haplotypes in the Beninese Toffin population sample**
Corresponding author:	**Dr. Eduardo A. Donadi**
E-mail address:	**eadonadi@fmrp.usp.br**
Journal:	**Molecular Immunology**
Our reference	**MIMM5567**
PII:	**S0161-5890(18)30513-3**
DOI:	**10.1016/j.molimm.2018.08.016**

My Co-authored Submissions

Molecular Immunology

Overview



Full Length Article | MIMM_2018_50

HLA-G, -E and -F regulatory and coding region variability and haplotypes in the Beninese Toffin population sample

Eduardo Donadi, Achille Massougbodji, Audrey Sabbagh, Celso T. Mendes-Junior, David Courtin, Erick C. Castelli, Ibrahim Sadissou, Juliana Massaro, Kabirou A. Moutairou, Kuumaaté K. G. M'po, Léonidas Tokplonou, Moudachirou IBIKOUNLE, Paulin Sonon, Philippe Moreau, Privat Agniwo, Sonya S. C. Glitho

Submitted 25 Jun 2018

Sent to Production 16 Aug 2018

View PDF

Alert: keep me informed about the submission status for this manuscript

APPENDIX

APPENDIX I

Table S1

HLA-G variable sites detected by NGS in entire evaluated segment in Beninese Toffin population.

Chr.	Gene	hg19/GRCh 37 position	Is it detected by 1Kgen phase3?	ID_SNP	Gene segment	IMGT/HLA relative position	Ref. allele	Ref. Freq (2n = 298)	ALT1 allele	ALT1 Freq (2n = 298)	ALT2	ALT2 Freq. (2n = 298)	HW *p-Value* (α = 0.01)	Variable sites with MAF > 1%	Observations
6	*HLA-G*	29794245	yes	rs113076482	5'UTR	-1377	T	0.9698	G	0.0302	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794317	yes	rs1736936	5'UTR	-1305	G	0.3926	A	0.6074	_	_	0.1210	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794443	yes	rs1736935	5'UTR	-1179	A	0.2886	G	0.7114	_	_	0.2327	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794467	yes	rs3823321	5'UTR	-1155	G	0.7282	A	0.2718	_	_	0.3007	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794482	yes	rs1736934	5'UTR	-1140	A	0.6644	T	0.3356	_	_	0.0989	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794484	yes	rs17875389	5'UTR	-1138	A	0.8960	G	0.1040	_	_	0.3685	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794501	yes	rs3115630	5'UTR	-1121	T	0.0000	C	1.0000	_	_	NA	yes	monomorphic alternative/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794658	yes	rs1632947	5'UTR	-964	G	0.3926	A	0.6074	_	_	0.1226	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794860	yes	rs1632946	5'UTR	-762	C	0.3926	T	0.6074	_	_	0.1221	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794897	yes	rs1233334	5'UTR	-725	G	0.0570	C	0.8389	T	0.1040	0.5707	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794906	yes	rs2249863	5'UTR	-716	T	0.3926	G	0.6074	_	_	0.1221	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794933	yes	rs2735022	5'UTR	-689	A	0.3926	G	0.6074	_	_	0.1201	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794956	yes	rs35674592	5'UTR	-666	G	0.3926	T	0.6074	_	_	0.1198	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794976	yes	rs17875391	5'UTR	-646	A	0.9128	G	0.0872	_	_	0.6024	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29794989	yes	rs1632944	5'UTR	-633	G	0.3926	A	0.6074	_	_	0.1237	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795076	no	rs201221694	5'UTR	-546	A	0.9832	AG	0.0168	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795081	no	rs368205133	5'UTR	-541	GA	0.9228	G	0.0772	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795083	yes	rs112940953	5'UTR	-539	A	0.9966	G	0.0034	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795113	yes	rs17875393	5'UTR	-509	C	0.8993	G	0.1007	_	_	0.3636	yes	not tracked by IPD-IMGT/HLA v.3.31.0

6	*HLA-G*	29795136	yes	rs1736933	5'UTR	-486	A	0.3926	C	0.6074	_	_	0.1213	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795139	yes	rs149890776	5'UTR	-483	A	0.9631	G	0.0369	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795145	yes	rs1736932	5'UTR	-477	C	0.2886	G	0.7114	_	_	0.2329	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795222	yes	rs17875395	5'UTR	-400	G	0.9262	A	0.0738	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795231	yes	rs17875396	5'UTR	-391	G	0.9262	A	0.0738	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795253	yes	rs1632943	5'UTR	-369	C	0.2886	A	0.7114	_	_	0.2330	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795366	no	rs551170211	5'UTR	-256	TC	0.9899	T	0.0101	_	_	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795421	yes	rs1233333	5'UTR	-201	G	0.3926	A	0.6074	_	_	0.1185	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795532	no	.	5'UTR	-90	G	0.9966	A	0.0034	_	_	1.0000	no	singleton/not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795566	yes	rs17875397	5'UTR	-56	C	0.8960	T	0.1040	_	_	0.3695	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795636	yes	rs1630223	Exon 1	15	G	0.3926	A	0.6074	_	_	0.1233	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795657	yes	rs1630185	Exon 1	36	G	0.3926	A	0.6074	_	_	0.1214	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795720	yes	rs56388903	Intron 1	99	A	0.0940	G	0.9060	_	_	0.6171	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795747	yes	rs6932888	Intron 1	126	G	0.8960	C	0.1040	_	_	0.6579	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795751	yes	rs6932596	Intron 1	130	C	0.8960	T	0.1040	_	_	0.6577	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795768	yes	rs1629329	Intron 1	147	T	0.2886	C	0.7114	_	_	0.2335	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795809	yes	rs1628628	Intron 1	188	C	0.3926	T	0.6074	_	_	0.1224	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795813	no	.	Intron 1	192	C	0.9966	T	0.0034	_	_	1.0000	no	singleton/not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795913	yes	rs41551813	Exon 2	292	A	0.8960	T	0.1040	_	_	0.3695	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795918	yes	rs80153902	Exon 2	297	G	0.9899	A	0.0101	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795987	yes	rs78627024	Exon 2	366	G	0.9832	A	0.0168	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29795993	yes	rs1130355	Exon 2	372	G	0.3926	A	0.6074	_	_	0.1217	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796103	yes	rs1626038	Intron 2	482	T	0.2517	C	0.7483	_	_	0.0268	yes	included in IPD-IMGT/HLA v.3.31.0

6	*HLA-G*	29796106	yes	rs17875399	Intron 2	485	G	0.8960	T	0.1040	_	_	0.3676	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796115	yes	rs1736927	Intron 2	494	A	0.2517	C	0.7483	_	_	0.0279	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796126	no	rs3215482	Intron 2	505	A	0.3557	AC	0.6443	_	_	0.0197	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796128	yes	rs17875400	Intron 2	507	C	0.8960	A	0.1040	_	_	0.3682	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796152	yes	rs1625907	Intron 2	531	G	0.3557	C	0.6443	_	_	0.0196	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796234	no	rs9278333	Intron 2	613	CA	0.3926	C	0.6074	_	_	0.1210	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796257	yes	rs1625035	Intron 2	636	C	0.2886	T	0.7114	_	_	0.2338	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796265	yes	rs17875401	Intron 2	644	G	0.8960	T	0.1040	_	_	0.3678	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796306	yes	rs1624337	Intron 2	685	G	0.3926	A	0.6074	_	_	0.1195	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796327	yes	rs1130356	Exon 3	706	C	0.6644	T	0.3356	_	_	0.1001	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796376	yes	rs12722477	Exon 3	755	C	0.7282	A	0.2718	_	_	0.3001	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796434	yes	rs41557518	Exon 3	813	AC	0.8859	A	0.1141	_	_	0.6947	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796492	yes	rs17875402	Exon 3	871	G	0.9698	A	0.0302	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796637	yes	rs17875403	Intron 3	1016	C	0.9866	T	0.0134	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796640	yes	rs1632942	Intron 3	1019	T	0.2886	C	0.7114	_	_	0.2315	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796675	yes	rs17875404	Intron 3	1054	G	0.8960	T	0.1040	_	_	0.3682	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796685	yes	rs1632941	Intron 3	1064	T	0.3926	C	0.6074	_	_	0.1239	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796768	yes	rs1632940	Intron 3	1147	T	0.2752	C	0.7248	_	_	0.0393	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796838	yes	rs1736923	Intron 3	1217	A	0.3926	G	0.6074	_	_	0.1214	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796934	yes	rs114041958	Intron 3	1313	G	0.8960	A	0.1040	_	_	0.3679	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796935	yes	rs1632939	Intron 3	1314	G	0.3926	A	0.6074	_	_	0.1202	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29796986	yes	rs1632938	Intron 3	1365	G	0.3926	A	0.6074	_	_	0.1198	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797073	yes	rs188836562	Intron 3	1452	G	0.9899	C	0.0101	_	_	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797144	yes	rs372858913	Intron 3	1523	G	0.9899	C	0.0101	_	_	1.0000	yes	not included in IPD-

															IMGT/HLA v.3.31.0
6	*HLA-G*	29797155	yes	rs17875405	Intron 3	1534	G	0.8960	C	0.1040	_	_	0.3700	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797173	yes	rs1736920	Intron 3	1552	A	0.2886	G	0.7114	_	_	0.2311	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797211	yes	rs41562616	Exon 4	1590	C	0.8960	T	0.1040	_	_	0.3684	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797448	yes	rs17875406	Exon 4	1827	G	0.8020	A	0.1980	_	_	0.4451	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797553	yes	rs1632937	Intron 4	1932	G	0.3926	C	0.6074	_	_	0.1230	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797639	yes	rs1049033	Exon 5	2018	C	0.6644	T	0.3356	_	_	0.1004	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797696	yes	rs1130363	Exon 5	2075	A	0.2886	G	0.7114	_	_	0.2344	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797782	yes	rs1611627	Intron 5	2161	T	0.3926	C	0.6074	_	_	0.1227	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797899	yes	rs1632934	Intron 5	2278	T	0.3926	C	0.6074	_	_	0.1233	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797933	yes	rs1632933	Intron 5	2312	C	0.3926	T	0.6074	_	_	0.1192	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29797951	yes	rs1736912	Intron 5	2330	A	0.3926	G	0.6074	_	_	0.1190	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798033	yes	rs17179080	Intron 5	2412	G	0.0369	A	0.9631	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798039	yes	rs1632932	Intron 5	2418	G	0.3926	A	0.6074	_	_	0.1220	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798140	yes	rs915667	Intron 5 (codon stop at +2536)	2519	A	0.3926	G	0.6074	_	_	0.1193	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798419	yes	rs915670	Intron 6	2798	G	0.6644	A	0.3356	_	_	0.0982	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798425	yes	rs915669	Intron 6	2804	G	0.2886	T	0.7114	_	_	0.2314	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798459	yes	rs915668	Intron 6	2838	C	0.2886	G	0.7114	_	_	0.2306	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798581	no	rs371194629	3'UTR/ Exon8 (14bp)	2960	G	0.5604	GATTTGTTCATGCCT	0.4396	_	_	0.4080	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798608	yes	rs567747015	3'UTR/ Exon8	3001	C	0.9799	T	0.0201	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798610	yes	rs1707	3'UTR/ Exon8	3003	C	0.0772	T	0.9228	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798617	yes	rs1710	3'UTR/ Exon8	3010	G	0.2886	C	0.7114	_	_	0.2330	yes	not tracked by IPD-IMGT/HLA v.3.31.0

6	*HLA-G*	29798639	yes	rs146339774	3'UTR/ Exon8	3032	G	0.9966	C	0.0034	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798642	yes	rs17179108	3'UTR/ Exon8	3035	C	0.8960	T	0.1040	_	_	0.3687	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798659	yes	rs569057854	3'UTR/ Exon8	3038	C	0.9933	T	0.0067	_	_	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798749	yes	rs1063320	3'UTR/ Exon8	3142	C	0.2886	G	0.7114	_	_	0.2311	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798794	yes	rs9380142	3'UTR/ Exon8	3187	A	0.8826	G	0.1175	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798803	yes	rs1610696	3'UTR/ Exon8	3196	C	0.6644	G	0.3356	_	_	0.0993	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798825	no	.	3'UTR/ Exon8	3204	G	0.9966	C	0.0034	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798834	yes	rs1233331	3'UTR/ Exon8	3227	G	0.9866	A	0.0134	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-G*	29798896	yes	rs541542414	3'UTR/ Exon8	3275	C	0.9933	T	0.0067	_	_	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0

Notes: Chr. = chromosome, 1kgen = 1000 genome, Ref. allele = reference allele, Ref. Freq = reference frequency, ALT allele = alternative allele, HW = Hardy-Weinberg.

APPENDIX II

Table S2

HLA-E variable sites detected by NGS in entire evaluated segment in Beninese Toffin population.

Chr.	Gene	hg19/GRCh 37 position	Is it detected by 1Kgen phase3?	ID_SNP	Gene segment	IMGT/HLA relative position	Ref. allele	Ref. Freq (2n = 300)	ALT1 allele	ALT1 Freq (2n = 300)	HW *p-Value* (α = 0.01)	Variable site with MAF > 1%	Observations
6	*HLA-E*	30455166	Yes	rs2078675	5'Upstream	-2143	T	0.7567	C	0.2433	0.8265	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455167	Yes	rs188844729	5'Upstream	-2142	G	0.9300	A	0.0700	0.5308	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455186	Yes	rs139422860	5'Upstream	-2123	G	0.9500	A	0.0500	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455203	Yes	rs17875359	5'Upstream	-2106	G	0.9700	A	0.0300	0.1164	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455240	Yes	rs28780108	5'Upstream	-2069	G	0.9600	C	0.0400	0.2053	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455294	Yes	rs12207974	5'Upstream	-2015	C	0.7267	G	0.2733	0.6855	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455321	Yes	rs17875360	5'Upstream	-1988	T	0.8700	C	0.1300	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455328	Yes	rs555000642	5'Upstream	-1981	G	0.9900	A	0.0100	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455513	Yes	rs576666054	5'Upstream	-1796	A	0.9900	G	0.0100	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455886	Yes	rs762324	5'Upstream	-1423	G	0.9833	A	0.0167	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30455920	Yes	rs1264459	5'Upstream	-1389	G	0.1967	A	0.8033	0.2959	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30456047	No	.	5'Upstream	-1262	GT	0.9933	G	0.0067	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30456142	Yes	rs116253207	5'Upstream	-1167	A	0.9833	G	0.0167	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30456150	Yes	rs17875364	5'Upstream	-1159	A	0.6933	G	0.3067	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30456151	No	rs17875365	5'Upstream	-1158	T	0.9900	C	0.0100	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30456230	Yes	rs146647219	5'Upstream	-1079	G	0.9867	T	0.0133	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30457196	Yes	rs141224659	5'UTR	-113	T	0.9433	C	0.0567	0.3808	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30457205	Yes	rs61356961	5'UTR	-104	A	0.8800	G	0.1200	0.6954	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30457283	Yes	rs76971248	5'UTR	-26	G	0.9500	T	0.0500	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30457732	Yes	rs1059510	Exon2	424	T	0.7400	C	0.2600	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0

6	*HLA-E*	30457766	Yes	rs150949676	Exon2	458	G	0.9433	A	0.0567	0.3812	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30458064	Yes	rs1264457	Exon3	756	G	0.4667	A	0.5333	0.5097	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30458066	No	.	Exon3	758	G	0.9967	C	0.0033	1.0000	no	singleton/not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30458630	Yes	rs116563630	Intron3	1322	G	0.9933	A	0.0067	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30458933	Yes	rs11548296	Exon4	1625	G	0.8800	C	0.1200	0.6966	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30459302	No	.	Intron4	1994	C	0.9933	T	0.0067	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30459577	Yes	rs183165297	Intron5	2269	T	0.9867	C	0.0133	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30459934	Yes	rs192032415	Intron5	2626	C	0.9767	T	0.0233	0.0696	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30460252	Yes	rs73727419	Intron6	2944	C	0.9967	T	0.0033	1.0000	no	singleton/not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30460350	Yes	rs201843188	Exon7	3042 (stop codon at +3048)	A	0.9933	C	0.0067	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30460776	Yes	rs74295295	3'UTR/Exon8	3468	A	0.9000	C	0.1000	0.6416	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30460942	Yes	rs17195376	3'UTR/Exon8	3634	G	0.8700	A	0.1300	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30461085	Yes	rs1059655	3'UTR/Exon8	3777	G	0.1200	A	0.8800	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30461086	Yes	rs115051198	3'UTR/Exon8	3778	A	0.9633	G	0.0367	0.1723	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30461132	No	.	3'UTR/Exon8	3824	T	0.9967	TG	0.0033	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30461605	Yes	rs9283	3'UTR/Exon8	4297	G	0.8700	A	0.1300	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-E*	30461728	Yes	rs566930407	3'UTR/Exon8	4420	C	0.9767	T	0.0233	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0

Notes: Chr. = chromosome, 1kgen = 1000 genome, Ref. allele = reference allele, Ref. Freq = reference frequency, ALT allele = alternative allele, HW = Hardy-Weinberg.

APPENDIX III

Table S3

HLA-F variable sites detected by NGS in entire evaluated segment in Beninese Toffin population.

Chr.	Gene	hg19/GRCh37 position	Is it detected by 1Kgen phase3?	ID_SNP	Gene segment	IMGT/HLA relative position	Ref. allele	Ref. Freq. (2n = 304)	ALT1 allele	ALT1 Freq. (2n = 304)	ALT2 allele	ALT2 Freq. (2n = 304)	ALT3 allele	ALT3 Freq. (2n = 304)	ALT4 allele	ALT3 Freq. (2n = 304)	HW *p-Value* (α = 0.01)	Variable sites with MAF > 1%	Observations
6	*HLA-F*	29689532	yes	rs9258169	F-Distal Promotor	-1709	T	0.7632	G	0.2368	_	_	_	_	_	_	0.0029	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689590	yes	rs9258170	F-Distal Promotor	-1651	T	0.7632	G	0.2368	_	_	_	_	_	_	0.0029	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689631	yes	rs17875375	F-Distal Promotor	-1610	T	0.8322	C	0.1678	_	_	_	_	_	_	0.5746	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689683	no	rs78739406	F-Distal Promotor	-1558	A	0.7632	AG	0.2368	_	_	_	_	_	_	0.0032	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689717	yes	rs9258171	F-Distal Promotor	-1524	G	0.7632	T	0.2368	_	_	_	_	_	_	0.0029	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689742	yes	rs183540921	F-Distal Promotor	-1499	T	0.9868	G	0.0132	_	_	_	_	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689766	yes	rs573592415	F-Distal Promotor	-1475	G	0.9803	A	0.0197	_	_	_	_	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689825	yes	rs558910896	F-Distal Promotor	-1416	A	0.9934	G	0.0066	_	_	_	_	_	_	1.0000	no	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29689973	yes	rs9258172	F-Distal Promotor	-1268	C	0.7632	T	0.2368	_	_	_	_	_	_	0.0030	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690002	yes	rs9258173	F-Distal Promotor	-1239	G	0.7632	T	0.2368	_	_	_	_	_	_	0.0027	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690026	yes	rs9258174	F-Distal Promotor	-1215	T	0.7632	C	0.2368	_	_	_	_	_	_	0.0033	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690034	no	.	F-Distal Promotor	-1207	C	0.9967	T	0.0033	_	_	_	_	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690056	yes	rs17875376	F-Distal Promotor	-1185	C	0.8882	T	0.1118	_	_	_	_	_	_	0.4011	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690125	yes	rs9258175	F-Distal Promotor	-1116	G	0.7632	A	0.2368	_	_	_	_	_	_	0.0030	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690127	no	.	F-Distal Promotor	-1114	T	0.9967	C	0.0033	_	_	_	_	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690228	yes	rs9258176	F-Distal Promotor	-1013	C	0.7632	G	0.2368	_	_	_	_	_	_	0.0031	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690694	yes	rs3998799	5' Upstream	-547	C	0.8487	G	0.1513	_	_	_	_	_	_	0.7512	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690715	yes	rs114795943	5' Upstream	-526	T	0.9737	A	0.0263	_	_	_	_	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29690741	yes	rs17875377	5' Upstream	-500	G	0.9375	A	0.0625	_	_	_	_	_	_	1.0000	yes	not tracked by IPD-IMGT/HLA v.3.31.0

6	*HLA-F*	29690838	yes	rs17875378	5' Upstream	-403	G	0.9967	A	0.0033	_	_	_	_	_	_	1.0000	no	singleton/not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691019	yes	rs1362126	5' Upstream	-222	G	0.8487	A	0.1513	_	_	_	_	_	_	0.7506	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691090	yes	rs1362125	5' Upstream	-151	T	0.6612	A	0.3388	_	_	_	_	_	_	0.3665	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691097	yes	rs2075682	5' Upstream	-144	A	0.7632	T	0.2368	_	_	_	_	_	_	0.0030	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691140	yes	rs2072896	5'UTR	-101	C	0.7632	G	0.2368	_	_	_	_	_	_	0.0027	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691278	yes	rs17875379	Exon 1	38	C	0.9704	T	0.0296	_	_	_	_	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691303	yes	rs2076183	Exon 1	63	G	0.7632	A	0.2368	_	_	_	_	_	_	0.0030	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691385	yes	rs373876297	Intron 1	145	G	0.9967	A	0.0033	_	_	_	_	_	_	1.0000	no	singleton/not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691390	yes	rs1632955	Intron 1	150	C	0.1941	G	0.8059	_	_	_	_	_	_	0.2002	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691391	yes	rs1632954	Intron 1	151	T	0.1941	C	0.8059	_	_	_	_	_	_	0.2011	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691445	yes	rs61739958	Exon 2	205	C	0.9770	A	0.0230	_	_	_	_	_	_	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691713	yes	rs2076182	Exon 2	473	A	0.7632	C	0.2368	_	_	_	_	_	_	0.0030	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691744	yes	rs2076181	Exon 2	504	C	0.7632	G	0.2368	_	_	_	_	_	_	0.0029	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691772	yes	rs2076179	Exon 2	532	T	0.6612	C	0.3388	_	_	_	_	_	_	0.3657	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691774	yes	rs2076178	Exon 2	534	A	0.8882	C	0.1118	_	_	_	_	_	_	0.4013	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691800	yes	rs17875381	Exon 2	560	C	0.8388	G	0.1612	_	_	_	_	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29691857	yes	rs2072895	Exon 2	617	C	0.6612	G	0.3388	_	_	_	_	_	_	0.3675	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692305	yes	rs1632953	Intron 3	1065	A	0.1941	G	0.8059	_	_	_	_	_	_	0.2028	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692334	yes	rs9258186	Intron 3	1094	A	0.7632	G	0.2368	_	_	_	_	_	_	0.0032	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692429	no	rs140848774	Intron 3	1189	AAATTTCTGAGGG	0.7632	A	0.2368	_	_	_	_	_	_	0.0030	yes	included in IPD-IMGT/HLA v.3.31.0

6	*HLA-F*	29692433	yes	.	Intron 3	1193	TTCTGAGGGAATG	0.3421	CTCTGAGGGAATG	0.5691	*	0.0888	_	_	_	_	0.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692462	no	rs142275427	Intron 3	1222	TGGAATACCGATCCGC	0.8980	T	0.1020	_	_	_	_	_	_	0.3661	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692465	yes	rs368156595	Intron 3	1225	A	0.8980	*	0.1020	_	_	_	_	_	_	0.3672	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692470	yes	rs13437375	Intron 3	1230	C	0.7697	A	0.1349	*	0.0954	_	_	_	_	0.0021	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692562	yes	rs1736926	Intron 3	1322	A	0.1941	G	0.8059	_	_	_	_	_	_	0.2013	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692618	yes	rs144376525	Intron 3	1378	C	0.9868	T	0.0132	_	_	_	_	_	_	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692622	yes	rs2072899	Intron 3	1382	C	0.4243	A	0.3388	G	0.2368	_	_	_	_	0.0087	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692623	yes	rs17178385	Intron 3	1383	C	0.8322	G	0.1678	_	_	_	_	_	_	0.5730	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692634	yes	rs1736925	Intron 3	1394	T	0.1941	G	0.8059	_	_	_	_	_	_	0.2010	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692729	yes	rs2072898	Intron 3	1489	T	0.7632	G	0.2368	_	_	_	_	_	_	0.0030	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29692737	yes	rs181298082	Intron 3	1497	T	0.9901	C	0.0099	_	_	_	_	_	_	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693011	yes	rs1736924	Exon 4	1771	C	0.1941	T	0.8059	_	_	_	_	_	_	0.2015	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693113	yes	rs2076177	Intron 4	1873	C	0.7632	T	0.2368	_	_	_	_	_	_	0.0031	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693183	yes	rs17184813	Intron 4	1943	G	0.9671	A	0.0329	_	_	_	_	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693186	yes	rs367892074	Intron 4	1946	T	0.9868	C	0.0132	_	_	_	_	_	_	1.0000	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693363	yes	rs371435287	Intron 5	2123	G	0.9901	T	0.0099	_	_	_	_	_	_	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693448	yes	rs17875383	Intron 5	2208	C	0.9375	T	0.0625	_	_	_	_	_	_	1.0000	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693499	yes	rs2235383	Intron 5	2259	A	0.7632	G	0.2368	_	_	_	_	_	_	0.0029	yes	included in IPD-IMGT/HLA v.3.31.0

6	*HLA-F*	29693938	yes	rs2735061	Intron 5	2698	G	0.9605	A	0.0395	_	_	_	_	_	_	0.2022	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29693969	yes	rs1736922	Intron 6	2729	C	0.7697	T	0.2303	_	_	_	_	_	_	0.1749	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694044	no	rs3835204	Intron 6 (stop codon at +2942)	2804	T	0.7632	TG	0.2368	_	_	_	_	_	_	0.0031	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694301	no	.	3' UTR	3061	A	0.9934	T	0.0066	_	_	_	_	_	_	1.0000	no	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694336	no	.	3' Downstream	3096	A	0.9967	G	0.0033	_	_	_	_	_	_	1.0000	no	singleton/included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694337	yes	.	3' Downstream	3097	TG_3	0.4572	TG_1	0.2336	TG2	0.1941	TG4	0.1086	TG5	0.0066	0.0143	yes	not included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694427	yes	rs1059174	3' Downstream	3189	C	0.1941	T	0.8059	_	_	_	_	_	_	0.2020	yes	included in IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694570	yes	rs3734813	3' Downstream	3330	A	0.7632	G	0.2368	_	_	_	_	_	_	0.0029	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694680	yes	rs3734814	3' Downstream	3440	A	0.7632	C	0.2368	_	_	_	_	_	_	0.0029	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694681	yes	rs3734815	3' Downstream	3441	A	0.7632	T	0.2368	_	_	_	_	_	_	0.0031	yes	not tracked by IPD-IMGT/HLA v.3.31.0
6	*HLA-F*	29694777	yes	rs17875384	3' Downstream	3537	G	0.8322	A	0.1678	_	_	_	_	_	_	0.5738	yes	not tracked by IPD-IMGT/HLA v.3.31.0

Notes: Chr. = chromosome, 1kgen = 1000 genome, Ref. allele = reference allele, Ref. Freq = reference frequency, ALT allele = alternative allele, HW = Hardy-Weinberg.

APPENDIX IV

Table S4. Extended haplotypes analysed (frequency ≥ 5%) and their specific alleles

(A) *HLA-G*					
Extended haplotypes	**Promo**	**Coding**	**3'UTR**	**Freq. Absolute (2n = 298)**	**Relative freq. (2n = 298)**
extHG1	PROMO-G010102a	G*01:01:02:01	UTR-02	58	0.1946
extHG2	PROMO-G0104a	G*01:04:04	UTR-3	55	0.1846
extHG10	PROMO-G010102a	G*01:05N	UTR-02	34	0.1141
extHG5	PROMO-G010101a	G*01:01:01:01	UTR-01	23	0.0772
(B) *HLA-E*					
Extended haplotypess	**Promo**	**Coding**	**3'UTR**	**Freq. Absolute (2n = 300)**	**Relative freq. (2n = 300)**
extHE1	E-Promo-15	E*01:05-compatible	HLA-E-3UTR-1	17	0.0567
extHE3	E-Promo-6	E*01:01:01:01	HLA-E-3UTR-1	12	0.0400
extHE4	E-Promo-7	E*01:03:02:01	HLA-E-3UTR-4	39	0.1300
extHE6	E-Promo-1	E*01:01:01:01	HLA-E-3UTR-1	97	0.3233
extHE7	E-Promo-20	E*01:03:05-compatible	HLA-E-3UTR-1	21	0.0700
extHE8	E-Promo-2	E*01:03:02:01	HLA-E-3UTR-2	29	0.0967
(C) *HLA-F*					
Extended haplotypes	**Distal Promoter**	**Upstream Promoter**	**Coding**	**3'UTR**	**Freq. Absolute (2n = 304)**	**Relative freq. (2n = 304)**
extHF1	F*Distal-A	F*upstream-A	F*01:01:01:09	F*3UTR-A	38	0.1250
extHF15	F*Distal-A	F*Upstream-D	F*01:01:01:05(1943G,TG12)	F*3UTR-A	46	0.1513
extHF19	F*Distal-B	F*upstream-C	F*01:01:02:03(1943G,2208C,TG12)	F*3UTR-A	21	0.0691
extHF24	F*Distal-B	F*upstream-C	F*01:01:02:05(1943G,TG11)	F*3UTR-A	31	0.1020
extHF35	F*Distal-C	F*upstream-B	F*01:01:01:01(TG12)	F*3UTR-A	29	0.0954
extHF38	F*Distal-D	F*upstream-A	F*01:03:01:01(1383G)	F*3UTR-A	45	0.1480