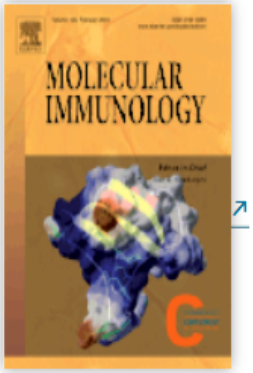
ELSEVIER



Visit journal homepage >

Submit your paper >

Guide for authors >

Track your paper >

Order journal >

Browse journals > Molecular Im... > Abstracting an...

Abstracting and Indexing

- BIOMED
- Medline/Index Medicus
- Excerpta Medica
- BIOSIS
- PASCAL/CNRS
- ACSA
- Cambridge Scientific Abstracts
- Chemical Abstracts
- Current Contents/Life Sciences
- Science Citation Index

….. etc.

https://www.elsevier.com/journals/molecular-immunology/0161-5890/abstracting-indexing

Facteur d'impact: 3,1

Molecular Immunology 104 (2018) 108–127

Contents lists available at ScienceDirect

Molecular Immunology

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

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, Brazil*
[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, Brazil*
[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, Benin*
[d] *UMR 216 MERIT, IRD, Faculté de Pharmacie de Paris, Université Paris Descartes, Sorbonne Paris Cité, Paris, France*
[e] *Université d'Abomey-Calavi, Cotonou, Benin*
[f] *Département de Zoologie, Faculté des Sciences et Techniques, Université d'Abomey-Calavi, Cotonou, Benin*
[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, Benin*
[k] *Sao Paulo State University (UNESP), Molecular Genetics and Bioinformatics Laboratory, Experimental Research Unit (UNIPEX), School of Medicine, Botucatu, State of Sao Paulo, Brazil*
[l] *Sao Paulo State University (UNESP), Department of Pathology, School of Medicine, Botucatu, State of Sao Paulo, Brazil*

ARTICLE INFO

Keywords:
HLA-E
HLA-F
HLA-G
Haplotypes
Massive parallel sequencing
African

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 in worldwide populations. *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

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

* Corresponding author at: Laboratório de Biologia Molecular, Universidade de São Paulo, Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (FMRP-USP), AV Bandeirantes, 3900, HC, Vila Monte Alegre, 14049900, Ribeirão Preto, SP, Brazil.

E-mail addresses: paulinsonon@gmail.com (P. Sonon), sadissouai@gmail.com (I. Sadissou), tokplonouleonidas@gmail.com (L. Tokplonou), kuumadox@yahoo.fr (K.K.G. M'po), glithosonya@yahoo.fr (S.S.C. Glitho), privatagniwo@yahoo.com (P. Agniwo), mibikounle2001@yahoo.fr (M. Ibikounlé), jmassaro@alumni.usp.br (J.D. Massaro), massougbodjiachille@yahoo.fr (A. Massougbodji), philippe.moreau@cea.fr (P. Moreau), audrey.sabbagh@ird.fr (A. Sabbagh), ctmendes@gmail.com (C.T. Mendes-Junior), kabirou.moutairou@gmail.com (K.A. Moutairou), erick.castelli@unesp.br (E.C. Castelli), david.courtin@ird.fr (D. Courtin), eadonadi@fmrp.usp.br (E.A. Donadi).

https://doi.org/10.1016/j.molimm.2018.08.016
Received 25 June 2018; Accepted 16 August 2018

(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.

1. Introduction

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). Differently from the HLA classical class I genes, *HLA-G*, *HLA-E* and *HLA-F* show limited genetic and protein variability. *HLA-E and HLA-F* are the most conserved ones, exhibiting 27 and 30 alleles, respectively, followed by *HLA-G* with 60 different alleles (IPD-IMGT/HLA v.3.32.0, 2018-04-16) (Robinson et al., 2015). In addition, non-classical HLA class I genes show a restricted expression pattern and are important regulators of 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 (Ishitani et al., 2006, 2003; Kovats et al., 1990).

HLA-G has a well-recognized checkpoint 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 (Allan et al., 1999; Baudhuin et al., 2013; Bian et al., 2016; Eguchi et al., 2016; Fournel et al., 2000; Grange et al., 2015; Huang et al., 2010; Köstlin et al., 2017; Lila et al., 2001; Morandi et al., 2010; Naji et al., 2014; Nazari and Farjadian, 2016; Petroff et al., 2002; Rouas-Freiss et al., 1997; Sampangi et al., 2015). Besides pregnancy, the immunomodulatory role of HLA-G has been reported in several disorders (Carosella et al., 2015; Dias et al., 2015; Donadi et al., 2011; Persson et al., 2017; Sabbagh et al., 2018). Many studies conducted in several worldwide populations have evaluated *HLA-G* variability at the 5' upstream regulatory (promoter), coding and/or 3' untranslated region (3′UTR) (Carlini et al., 2013; Castelli et al., 2017, 2011, 2010, Catamo et al., 2015, 2014; Consiglio et al., 2011; Courtin et al., 2013; de Albuquerque et al., 2016; Garcia et al., 2013; Garziera et al., 2015; Gineau et al., 2015; Hviid et al., 2006; Lucena-Silva et al., 2013, 2012; Martelli-Palomino et al., 2013; Nilsson et al., 2016; Porto et al., 2015; Sabbagh et al., 2014; Santos et al., 2013; Sizzano et al., 2012; Tan et al., 2005; Veit et al., 2014, 2012; Zambra et al., 2016). However, with the exception of a study reporting the *HLA-G* variability detected at the *1000 Genomes phase 1* data (Castelli et al., 2014a), which is characterized by low depth of coverage, there is no characterization of the complete *HLA-G* gene variability in African population samples.

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 HLA class I molecules (Persson et al., 2017) and binds to the inhibitory CD94-NKG2A receptor expressed by NK cells. 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; Braud et al., 1999; Lee et al., 1998). HLA-E is also a ligand for T CD8 cell receptor (Pietra et al., 2010, 2009), and may present non-self-antigens derived from cytomegalovirus, human immunodeficiency virus, Epstein-Barr virus, influenza and hepatitis C viruses (Pietra et al., 2010, 2009; Sullivan et al., 2008), 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 cellular stress contexts (Jørgensen et al., 2012). Currently, 27 *HLA-E* alleles encoding 8 full-length proteins and one null allele (*E*01:08N) are officially reported (IPD-IMGT/HLA v.3.32.0, 2018-04-16), but only two full-length proteins (E*01:01 and E*01:03) are frequently detected worldwide (Arnaiz-Villena et al., 2007; Carvalho dos Santos et al., 2013; Castelli et al., 2015; Felício et al., 2014; Liu et al., 2012; Ramalho et al., 2017; Veiga-Castelli et al., 2012). 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 (Ramalho et al., 2017; Tamouza et al., 2007; Veiga-Castelli et al., 2012), but the mechanisms underlying this phenomenon are unclear. Many studies have explored *HLA-E* exonic variability in different population samples, including Brazilians (Carvalho dos Santos et al., 2013; Felício et al., 2014; Ramalho et al., 2017; Veiga-Castelli et al., 2016, 2012), Africans (Guinea–Conakry and

Table 1
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									
		5′ upstream [a]							
Genes	Entire evaluated segment	Distal promoter [b]	Proximal promoter	Exons	Introns	Downstream the codon stop [c]	Region evaluated by IMGT [d]	Variable sites with MAF > 1% in region tracked by IMGT [d]	% of variable sites with MAF > 1% already described at IMGT/v.3.31.0 [e]
HLA-G	96 [f]	_	29	14	40	16	58	56	95.00 (53)
HLA-E	37 [f]	_	19	6	5	7	15	10	50.00 (5)
HLA-F	68 [f]	16	8	10	26	8	44	39	87.18 (34)

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

[a] All nucleotides upstream the +1 are considered 5′ upstream.

[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).

[c] "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*) (Tables S1–S3).

[d] 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).

[e] 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.

[f] Many of these variable sites were also included in *1000 Genome phase 3* project (Tables S1–S3).

Table 2
HLA-G promoter haplotypes encountered in a Beninese Toffin population sample.

	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 ($2n$ = 298)
PROMO-G010101a	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-G010101b	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-G010101d	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-G010101f	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-G010101h	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-G010101k	**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-G010102a	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-G010102f	**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-G0103a	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-G0103c	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-G0103e	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-G0103g	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-G0103h	**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-G0104a	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-G0104b	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-G0104c	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 gray and dark gray respectively.
[b]Haplotypes were named according to previous studies (Castelli et al., 2017, 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.

Burkina Faso) (Castelli et al., 2015), and worldwide population samples from the *1000 Genomes phase 1* panel (Felício et al., 2014). However, only few studies have evaluated the complete *HLA-E locus*, including the regulatory segments, in Asian Dutch or Hunan Han samples (Olieslagers et al., 2017), in Brazilians (Ramalho et al., 2017) and in cell lines from Asian, African American, and European origins (Pyo et al., 2006).

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 escaping from the NK cell cytotoxicity and consequently to poor prognosis of the disease (Harada et al., 2015; Xu et al., 2015). HLA-F might play an important role on the maintenance of immunotolerance in feto-maternal interface during pregnancy, 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 trimester (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) on NK cells and with ILT-2 on lymphomononuclear cells (Garcia-Beltran et al., 2016; Goodridge et al., 2013). It can also present peptides to T cells and regulate immunity through interactions with NK cell receptors (Dulberger et al., 2017). Being less polymorphic than *HLA-G*, the 30 officially recognized alleles encode five full-length HLA-F 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 evaluated only the well-described *HLA-F* polymorphisms in Korean and Chinese populations (Kim et al., 2011; Moscoso et al., 2007; Pan et al., 2013; Zhang et al., 2012) and few reports are available regarding *HLA-F* gene segments or the entire gene region in Brazilians (Lima et al., 2016; Manvailer et al., 2014) and in Asian, African American, and Caucasian samples (Pyo et al., 2006).

Although sub-Saharan African populations are characterized by high levels of intra-population nucleotide and haplotype diversity and inter-population genetic differentiation, little attention has been devoted to the study of immunomodulatory HLA genes in these populations. African countries 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 costumes.

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. Here, we evaluated the entire *HLA-G, -E and -F* gene variability, including the regulatory segments, all exons and introns, by using massively parallel sequencing technology, identifying variable sites and haplotypes in a Beninese Toffin (etymologically "people of water") population, which 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 in the same region for centuries (Principaud, 1995).

2. Material and methods

2.1. Study sample

Two hundred and eighty-eight children aged 4–8 years (mean age: 6.7 (± 0.8) years) of the Toffin ethnic group were sampled as part of a case/control survey about the risk factors of helminth infections and the immune consequences of these infections. The study was conducted

Table 3
HLA-G haplotypes encountered in a Beninese Toffin population sample considering the whole segment tracked by IPD-IMGT/HLA from −300 to +2838.

Coding haplotypes [a]	Notes	Frequency ($2n$ = 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 (+99G, +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

Notes:
(A) indicates the variable sites described in 1000 genome phase 3 data (Table S1).
(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.
[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.

between 2012 and 2014 at Sô-Ava, a lake area located in Southern region of Benin, 12 km North of Cotonou, the economical capital of Benin. Peripheral blood was collected from 187 (64.9%) boys and 101 (35.1%) girls. Based on recorded information, a total of 154 unrelated children (exclusion of close relatives up to the second degree) were enrolled in the present study, with the DNA samples successfully sequenced for *HLA-G* (n = 149), *HLA-E* (n = 150) and *HLA-F* (n = 152). The study was approved by the Ethics Committee of the ''Faculté des Sciences de la Santé (FSS)'' of Cotonou, Benin and registered at No.12/03/2012/CEIFSS/UAC. Informed consent was obtained from all individual participants included in the study.

2.2. DNA extraction and quantification

Genomic DNA was extracted using GeneJET Genomic DNA Purification kit (Thermo Fisher Scientific, Waltham, MA), according to manufacturer's instructions, and stored at −20 °C until use. The genomic DNA was quantified using Qubit® dsDNA HS Assay kit (Invitrogen, Carlsbad, CA) and normalized to 50 ng/μL.

2.3. DNA amplification and sequencing library preparation

Each DNA segment of interest was amplified as an unique amplicon, as described elsewhere (Castelli et al., 2017; Lima et al., 2016; Ramalho et al., 2017). PCR products were evaluated in 1% agarose gels, quantified using the Qubit® dsDNA HS Assay kit (Invitrogen), normalized to the same concentration and purified with ExoSap (GE Healthcare Life Sciences, Pittsburgh, PA). Each amplicon pool was then normalized at 0.2 ng/μL. After library preparation with Nextera® XT DNA Sample Preparation Kit (Illumina, San Diego, CA), quality checking was assessed by 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) for average length of library fragments and quantification was performed using qPCR (KAPA Library Quantification Kit, Kapa Biosystems, Wilmington, MA). Libraries were pooled together and sequenced at the MiSeq system (Illumina), using Illumina Reagent (V2 500 cycles) in a paired-end mode (2 × 250bp).

2.4. HLA-G, -E and -F NGS data processing, variant calling and haplotype inference

To avoid mapping bias because of the similar and repetitive nature of HLA genes, we used *hla-mapper* (available at www.castelli-lab.net) to map the reads, as described elsewhere (Castelli et al., 2017; Lima et al., 2016; Ramalho et al., 2017). The hla-mapper 2.0 was run using database version *002.1* and default parameters. BAM files were visualized with Integrative Genomics Viewer (IGV) v.2.3.81 (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). The genotype and haplotype calling approaches used here were similar to the ones described elsewhere (Castelli et al., 2017, 2015; Lima et al., 2016; Ramalho et al., 2017). Briefly, genotypes were inferred using the GVCF mode of the HaplotypeCaller algorithm (The Genome Analysis Toolkit, GATK, version 3.7) and combined as a multi-sample VCF file by using the GATK GenotypeGVCFs (DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2013). Genotypes were processed using *vcfx checkpl* (available at www.castelli-lab.net) to introduce missing alleles when the genotype likelihood was under 99.9%, assuring that only high-quality genotypes are passed forward to the imputation step using the PHASE algorithm (Stephens et al., 2001; Stephens and Donnelly, 2003). Before the phasing approach described below, singletons (variable site detected in a single heterozygous individual from the population sample) were removed and stored in a separate file.

Phasing was performed in a three-step approach. First, we used the algorithm GATK ReadBackedPhasing (*phaseQualityThresh* set to 2000) to detect the physical phase between variable sites that were close enough to be present in the same read or fragment. Because of limitations of the ReadBackedPhasing algorithm, indels and multi-allelic *loci* were ignored in this step. Second, the physical associations detected by ReadBackedPhasing were passed forward to the PHASE algorithm as. known files, using local scripts available at https://github.com/erickcastelli/phase-readbackedphasing. The PHASE algorithm was also used to impute the missing alleles introduced by *vcfx checkpl*. Third, singletons were manually introduced when they followed the criteria established elsewhere (Castelli et al., 2017): (a) no samples presented missing allele at this particular position, and (b) there was a clear relationship between the singleton and the next heterozygous site visually defined using the BAM file of that sample or it was the only heterozygous site detected in that particular sample.

The final phased VCF files were then converted into complete *HLA-G, -E and -F* sequences using *vcfx fasta* (www.castelli-lab.net/apps/vcfx) and the hg19 chromosome 6 sequence as reference. By using data from the IPD-IMGT/HLA version 3.31.0 database (Robinson et al., 2015) and a local BLAST server, we defined the *HLA-G, -E and -F* coding alleles for

Table 4
HLA-G 3′ UTR haplotypes encountered in a Beninese Toffin population sample.

	***HLA-G* 3' UTR haplotypes**[a]													
Haplotypes[b]	**14bp (+2960)**	**+3001**	**+3003**	**+3010**	**+3032**	**+3035**	**+3038**	**+3142**	**+3187**	**+3196**	**+3204**	**+3227**	**+3275**	**Frequency ($2n$ = 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 gray.
[b]Haplotypes were named according to previous studies (Nilsson et al., 2016; Sabbagh et al., 2014; Gineau et al., 2015; Castelli et al., 2010, 2011; Lucena-Silva et al., 2012; de Albuquerque et al., 2016; Santos et al., 2013; Castelli et al., 2014b, 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.

each sample and, when the sequences here obtained were not identical to any one already described, we defined the closest known allele followed by any variation observed. In addition, the promoter and 3'UTR haplotypes were defined and named according to previous studies (Castelli et al., 2010, 2011, 2014a, 2015, 2017; de Albuquerque et al., 2016; Gineau et al., 2015; Lima et al., 2016; Lucena-Silva et al., 2012; Nilsson et al., 2016; Ramalho et al., 2017; Sabbagh et al., 2014; Santos et al., 2013; Tan et al., 2005), since no official nomenclature has been assigned to these HLA gene regions yet.

2.5. Other statistical analyses

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 and Thompson, 1992), using the ARLEQUIN v3.5.2 software. The linkage disequilibrium (LD) pattern was evaluated between polymorphic sites by computing the correlation coefficient r^2 (Devlin and Risch, 1995), and 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%.

3. Results

Considering the *HLA-G* (from position −1377 to +3275), *HLA-E* (from −2143 to +4420) and *HLA-F* (from −1709 to +3537) *loci*, we detected 96, 37 and 68 variable sites in the Beninese Toffin sample, respectively. The full list of variable sites, along with their location, SNPid, and MAF is provided in Tables S1 to S3 for *HLA-G*, *HLA-E* and *HLA-F*, respectively. The majority of these variable sites have already been reported either by the *1000 Genomes phase 3* project (Auton et al., 2015) (indicated at Tables S1 to S3) or by the IPD-IMGT/HLA v3.31.0 database (Robinson et al., 2015) (Table S1 to S3: lines shaded in gray) or by specific population studies (Castelli et al., 2017, 2015; Lima et al., 2016; Ramalho et al., 2017), exception made for: **i)** three singletons at the *HLA-G locus* at positions −90, +192 and +3204, **ii)** four variable sites at the *HLA-E locus* at positions −1262, +758, +1994 and +3824 and **iii)** six variable sites at the *HLA-F locus*, at positions −1558, −1207, −1114, +2804, +3061 and +3096. Most of the variable sites here detected were polymorphic, since 92.7%, 73.0% and 86.8% of them exhibited frequencies higher than 1% for *HLA-G, -E and -F,* respectively (Tables S1–S3).

For *HLA-G* and *HLA-E*, all variable sites did fit the Hardy-Weinberg expectations (Tables S1 and S2). For *HLA-F*, 26 variable sites did not fit Hardy-Weinberg expectations. These variations occurred mainly in the *F*distal promoter (positions upstream −1013) and in some introns (Table S3). These segments usually harbor low read coverage (Castelli et al., 2017, 2015; Lima et al., 2016) due to the errors produced by the Nextera Kit transposase during the DNA library preparation (Shiina et al., 2012; Wang et al., 2011). However, the *hla-mapper* and *vcfx* software (available at www.castelli-lab.net) had been proposed and recommended by our group (Castelli et al., 2017, 2015; Lima et al., 2016) to overcome the genotype errors that can occur during the genotype inference or imputation steps at these positions. We used the same strategies of analysis (detailed at Section 2.4) and obtained an average read of 1228 covering the bases in each specific segment mentioned above. In addition, only 0.31% of missing alleles introduced by *vcfx* that was after imputed by PHASE program (Stephens et al., 2001; Stephens and Donnelly, 2003) (detailed at Section 2.4). Moreover, most of the variable sites (24 out of 26) had been already described either at *1000 Genomes phase 3* project (Table S3) or at IPD-IMGT/HLA v3.31.0 database or at specific population (Lima et al., 2016).

Table 5
HLA-G extended region (from −1377 to +3275) haplotypes encountered in a Beninese Toffin population sample.

Promoter haplotypes [a]	Coding haplotypes [b]	3' UTR haplotypes [c]	Notes	Frequency ($2n$ = 298) [d]
PROMO-G010102a	G*01:01:02:01	HLA-G-3UTR-02		0.1946
PROMO-G0104a	G*01:04:04	HLA-G-3UTR-03		0.1846
PROMO-G010102a	G*01:05N	HLA-G-3UTR-02		0.1141
PROMO-G010101a	G*01:01:01:01	HLA-G-3UTR-01		0.0772
PROMO-G010101f	G*01:01:01:04	HLA-G-3UTR-06		0.0604
PROMO-G0104a	G*01:04:01	HLA-G-3UTR-03		0.0537
PROMO-G010101b	G*01:01:01:05	HLA-G-3UTR-04		0.0336
PROMO-G010101h	G*01:01:09 (nodel at +615)	HLA-G-3UTR-04		0.0302
PROMO-G0103e	G*01:03:01:02	HLA-G-3UTR-05		0.0302
PROMO-G0103c	G*01:03:01:02	HLA-G-3UTR-05		0.0302
PROMO-G010101d	G*01:01:01:08	HLA-G-3UTR-01		0.0235
PROMO-G010101b	G*01:01:15-compatible	HLA-G-3UTR-06		0.0168
PROMO-G0103e	G*01:03:01:02	HLA-G-3UTR-17		0.0168
PROMO-G010101d	G*01:01:01:05 (+99G, +1147C, +2412A)	HLA-G-3UTR-01		0.0134
PROMO-G010102a	G*01:01:02:02	HLA-G-3UTR-02		0.0134
PROMO-G0103a	G*01:03:01:02	HLA-G-3UTR-05		0.0134
PROMO-G010101f	G*01:01:01:04	HLA-G-3UTR-18		0.0101
PROMO-G010102a	G*01:01:02:01(new1452C)	HLA-G-3UTR-02	New haplotype	0.0101
PROMO-G0104a	G*01:04:01 (new1523C)	HLA-G-3UTR-03	New haplotype	0.0101
PROMO-G0104c	G*01:04:05-compatible	HLA-G-3UTR-03		0.0101
PROMO-G010101a	G*01:01:01:06	HLA-G-3UTR-04		0.0067
PROMO-G0103e	G*01:03:01:02	HLA-G-3UTR-48	New haplotype	0.0067
PROMO-G0104a	G*01:04:04	HLA-G-3UTR-46		0.0067
PROMO-G010101a	G*01:01:01:01 (new192T)	HLA-G-3UTR-01	New haplotype	0.0034
PROMO-G010101f	G*01:01:01:04	HLA-G-3UTR-47	New haplotype	0.0034
PROMO-G010101f	G*01:01:01:04	HLA-G-3UTR-20		0.0034
PROMO-G010101k	G*01:01:01:05 (new-90A)	HLA-G-3UTR-04	New haplotype	0.0034
PROMO-G010101b	G*01:01:01:05	HLA-G-3UTR-03		0.0034
PROMO-G010102f	G*01:01:02:01	HLA-G-3UTR-02	New haplotype	0.0034
PROMO-0103g	G*01:03:01:02	HLA-G-3UTR-05		0.0034
PROMO-G0103h	G*01:03:01:02	HLA-G-3UTR-17	New haplotype	0.0034
PROMO-0104b	G*01:04:04	HLA-G-3UTR-03		0.0034
PROMO-G0104a	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 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.

[a] The list of promoter haplotypes are depicted at Table 2.
[b] The list of coding haplotypes are depicted at Table 3.
[c] The list of 3′ UTR haplotypes are depicted at Table 4.
[d] Haplotypes were classified in descending order of their frequency.

3.1. HLA-G variability and haplotypes

3.1.1. HLA-G 5' upstream regulatory region

Considering the region upstream the first transcribed ATG, encompassing the positions −1 to −1377, we detected 29 variable sites for *HLA-G* (Table 1), among which 27 (93.1%) are polymorphic (MAF ≥ 1%) (Table S1). These variable sites were found arranged into 16 promoter haplotypes, clustered into four groups or lineages (PROMO-G010101, PROMO-G010102, PROMO-G0103 and PROMO-G0104), according to previous studies (Castelli et al., 2011, 2014a, 2017; Gineau et al., 2015; Nilsson et al., 2016; Tan et al., 2005) (Table 2). Indeed, the PROMO-G010101 lineage grouped PROMO-G010101a to PROMO-G010101k, the PROMO-G010102 lineage grouped PROMO-G010102a and PROMO-G010102f, the PROMO-G0103 lineage grouped PROMO-G0103a to PROMO-G0103h and the PROMO-G0104 lineage grouped the PROMO-G0104a to PROMO-G0104c. Within each group or lineage, the most frequent haplotype is PROMO-G010101a (8.7%), PROMO-G010102a (33.2%), PROMO-G0103e (5.4%) and PROMO-G0104a (25.8%), respectively. The PROMO-G010101k, PROMO-G010102f and PROMO-G0103 h represent new haplotypes, sampled as single copies in our population (Table 2).

3.1.2. HLA-G IPD-IMGT/HLA segment

Fifty-eight variable sites were identified in the IPD-IMGT/HLA region (−300 to +2838) (Tables 1 and S1: lines shaded in gray), among which 56 occurred with a MAF ≥ 1% (Table 1). Fifty-three (95%) of these polymorphisms have already been described in the IPD-IMGT/HLA v3.31.0 database (Tables 1 and S1). The three remaining variable sites, at positions −256, +1452 and +1523, occurring either in an intron (intron 3) or in regulatory regions, may be considered as new variable sites (Table S1). Note that 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).

A total of 19 coding haplotypes (or *HLA-G* alleles) were inferred in our population sample (Table 3). Thirteen of them (including the compatible ones) have already been described at the IPD-IMGT/HLA v3.31.0 database. Of the six remaining haplotypes, two have already been identified in a Brazilian sample (Castelli et al., 2017). The remaining four can be considered as new haplotypes since their variable sites have not been described in the IPD-IMGT/HLA database. Two of them [G*01:01:02:01$^{(new1452C)}$ (1.01%) and G*01:04:01$^{(new1523C)}$ (1.01%)] carried the two new variable sites mentioned above, while the last ones [G*01:01:01:01$^{(new192T)}$ (0.34%) and G*01:01:01:05$^{(new-90A)}$

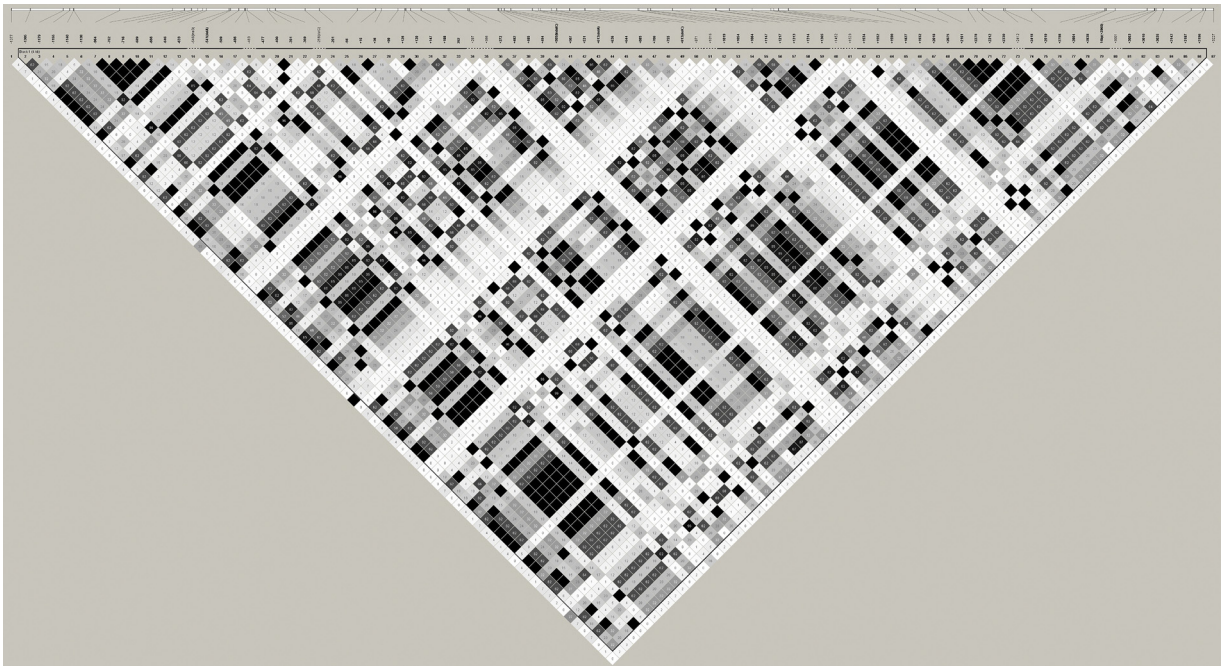


Fig. 1. LD patterns for the entire *HLA-G* region encompassing positions -1377 to +3275 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (×100) 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%.

(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 3).

3.1.3. HLA-G 3'UTR

In the 3'UTR segment (from +2960 to +3275), we identified 16 variable sites (Table 1), among them, 12 (75%) were polymorphic (Table S1). These sites were arranged into 12 haplotypes, named in accordance with previous studies (Castelli et al., 2010, 2011, 2014a, 2017; de Albuquerque et al., 2016; Lucena-Silva et al., 2012; Nilsson et al., 2016; Sabbagh et al., 2014; Santos et al., 2013) (Table 4). 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%.

3.1.4. HLA-G extended haplotypes

We found 33 extended haplotypes in the whole *HLA-G* gene region, encompassing positions −1377 to +3275, which are presented in decreasing order of frequency in Table 5. Among these haplotypes, 20 (60.61%) 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 (Fig. 1), as already reported elsewhere (Castelli et al., 2017; Santos et al., 2013), strong associations were noted between haplotypes of different gene segments: **a)** all copies of *G**01:01:01:01 (or derivatives) are associated with PROMO-G010101a and *HLA-G*-3UTR-01, **b)** all copies of *G**01:01:01:04 are associated to PROMO-G010101f and *HLA-G*-3UTR-06 or similar (*HLA-G*-3UTR-18, -20 and -47), **c**) *G**01:01:02:01 is associated to PROMO-G010102a or similar (PROMO-G010102f) and *HLA-G*-3UTR-02, and **d)** the *G**01:04:01 or *G**01:04:04 is associated to PROMO-G0104a and *HLA-G*-3UTR-03 or similar (*HLA-G*-3UTR-46).

3.2. HLA-E variability and haplotypes

3.2.1. HLA-E 5' upstream regulatory region

In the region upstream the first transcribed ATG, encompassing positions −1 to −2143, we identified 19 variable sites (Table 1), of which 15 (78.95%) are polymorphic (Table S2). They were arranged into 19 promoter haplotypes, named as E-Promo-01 to E-Promo-39, according to a previous *HLA-E* study (Ramalho et al., 2017). Most of them (16 out of 19) were present at a minimum frequency of 1%. The E-Promo-14 (0.67%), E-Promo-23 (0.67) and E-Promo-38 (0.33%) were the least frequent. The E-Promo-38 (0.33%) and E-Promo-39 (1%) represent new haplotypes detected in this population sample (Table 6).

3.2.2. HLA-E IPD-IMGT/HLA segment

Fifteen variable sites were described in the IPD-IMGT/HLA region for *HLA-E* (−300 to +3522) (Tables 1 and S2: lines shaded in gray), 10 of which are polymorphic (Table 1). Five (50%) of these polymorphic sites were described in the IPD-IMGT/HLA v3.31.0 database, exception made for 5 sites, at positions −113, −104, +2269, +2626 and +3468 (Tables 1 and S2). However, these sites have also been described in the *1000 Genomes phase 3* panel. They occurred in either intronic or regulatory regions (Table S2). We found two new substitutions: +758 G > C in codon 107 of exon 3 (MAF = 0.33%) and +3042 A > C in codon 336 of exon 7 (0.67%) (Table 7). In addition, we identified one variable site at position +1994 (intron 4), carried by the *E**01:01:01:09 allele, that corresponds to a deletion in IPD-IMGT/HLA v3.31.0 database, while it corresponded to a T (0.67%) in the present study and was carried by the *E**01:01:01:01 allele (Tables 7 and S2).

In this IPD-IMGT/HLA region, 15 different coding haplotypes (or *HLA-E* alleles) were inferred (Table 7). Six of them were officially recognized by IPD-IMGT/HLA v.3.31.0 database. Of the remaining 9 alleles: **i)** two (*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) and **ii)** one (*E**01:01:01:01$^{(new3468C)}$) was also observed in the Brazilian population (Ramalho et al., 2017). The six remaining haplotypes can be considered as new alleles and they present a summed frequency of 4.66%. They carried one or more variable sites not described in the IPD-IMGT/HLA v.3.31.0 database. Of them, the *E**01:01:01:01$^{(new758C)}$ and *E**01:01:01:01$^{(new3042C)}$ carried variable sites at exons 3 and 7, respectively (Tables 7 and S2). Theses

Table 6
HLA-E 5′ upstream regulatory region haplotypes encountered in a Beninese Toffin population sample.

	***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 (2*n* = 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.

variations (+758 C and +3042 C) lead to a G > C 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 7). 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 7).

3.2.3. HLA-E 3'UTR

In the region downstream the stop codon (3'UTR segment, from +3468 to +4420), we identified seven variable sites, all polymorphic except one (Tables 1 and S2). They were arranged into seven haplotypes (Table 8) and named as *HLA-E*-3UTR-01 to *HLA-E*-3UTR-14 in accordance with previous studies (Castelli et al., 2015; Ramalho et al., 2017). 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 8).

3.2.4. HLA-E extended haplotypes

Thirty-one *HLA-E* extended haplotypes were detected in the region encompassing positions −143 to +4420 (Table 9). 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 combinations can be observed. For example: **a)** *E*-Promo-01 and similar haplotypes (*E*-Promo-05, -06, -14, -31 and -39) are associated to the coding allele *E**01:01:01 (or derivatives) and to the *HLA-E*-3UTR-01 haplotype, and **b)** *E*-Promo-02 and similar haplotypes (E-Promo-17) are associated to the coding allele *E**01:03 (or derivatives) and to *HLA-E*-3UTR-02. However, these associations are not straightforward because of the weak LD among the different *HLA-E* segments (Fig. 2), as also reported elsewhere (Castelli et al., 2017; Felício et al., 2014; Ramalho et al., 2017).

3.3. HLA-F variability and haplotypes

3.3.1. HLA-F 5' upstream regulatory region

The region upstream the first transcribed ATG was evaluated in two segments, a proximal promoter (from −556 to −1), which was previously evaluated in a Brazilian sample (Lima et al., 2016), and a distal promoter, considering nucleotides upstream position −556. However, we did not call genotypes at the segment encompassing positions −547 (hg 19 position 29690694) to −1013 (hg 19 position 29690228), because it consisted of extended short tandem repeat (STR), with low sequence complexity. After genotype quality checking with *vcfx checkpl* function (detailed at Section 2.4), this region presented high level of missing alleles. For this reason, this region was excluded from the analysis. Then, we detected 16 and 8 variable sites in the distal and proximal promoter, respectively (Table 1). Taken together, 20 out of the 24 (83.3%) variable sites identified in these *HLA-F* promoter regions presented a MAF > 1% (Table S3).

The distal promoter variable sites were found arranged into eight different haplotypes (Table 10). As 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 10), in accordance with the pattern followed 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

Table 7
HLA-E haplotypes encountered in a Beninese Toffin population sample considering the whole segment tracked 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).
(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) leading to a substitution of A nucleotide into C nucleotide in codon 12.
(E) Synonymous mutation on exon 3 (Table S2) 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 al., 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.

greater than 1%, with *F**Distal-A being the most frequent (45.1%) one. In the same Table 10, we identified and presented seven F-proximal promoter haplotypes. They were named as *F**upstream-A to *F**upstream-H according to a previous *HLA-F* study (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%).

3.3.2. *HLA-F IPD-IMGT/HLA segment*

Forty-four variable sites were described in the IPD-IMGT/HLA region (-300 to +3250) for *HLA-F* (Tables 1 and S3: lines shaded in gray). Among them, 39 were polymorphic with a MAF > 1% (Table 1). Thirty-four (87.2%) of these polymorphic sites were also detected and described in IPD-IMGT/HLA v3.31.0 database (Tables 1 and S3). The remaining five, at positions +205, +1378, +1946, +2804 and +3097, can be considered as new variable sites since they were not described in the IPD-IMGT/HLA database. The variable site at position +205 [*+205 A (2.3%)*] carried by *F**01:01:01:09$^{(new205A)}$ is a synonymous mutation in exon 2, leading to a C > A substitution at codon 4 (Tables 11 and S3). Except for *+2804 T > TG (23.68%),* all the new variable sites have also been described by the *1000 Genomes phase 3* project (Table S3).

A total of 29 haplotypes were inferred in this IPD-IMGT/HLA coding region (Table 11). Of these, 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 carried variable sites already described in IPD-IMGT/HLA v.3.31.0 database or shared with the Brazilian population sample (Lima et al., 2016), and the 5 remaining haplotypes can be considered as new ones (Table 11), since they carried variants that were not described in IPD-IMGT/HLA v.3.31.0 database. Altogether, 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 11). Interestingly, the F*01:02 protein (produced by F*01:02$^{(1193C,\ 1943G)}$ and F*01:02$^{(1193C,\ 1943G,\ new1946C)}$) was absent in a previous Brazilian population sample (Lima et al., 2016).

Table 8
HLA-E 3′ UTR haplotypes encountered in a Beninese Toffin population sample.

***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 gray.
[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.

3.3.3. HLA-F 3'UTR

We identified eight variable sites in the region downstream the stop codon, encompassing the positions +3061 to +3537 (Table 1), and six presented a MAF ≥ 1% (Table S3). As described by Lima et al., 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 our 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 12).

3.3.4. HLA-F extended haplotypes

Considering the entire gene from position −1709 to +3250, 32 different extended haplotypes were inferred (Table 12). Fifteen of them presented polymorphic frequencies ranging from 1.3% to 15.1%, whereas 17 had frequencies below 1% (Table 12). Generally, we can observe the following associations: **i)** *F**01:01:01:01 or derivatives are associated to *F**Distal-B, *F**upstream-B and *F**3UTR-A, **ii)** *F**01:01:01:09 or *F**01:02 or derivatives are associated to *F**Distal-A, *F**upstream-A and *F**3UTR-A, and **iii)** *F**01:03:01:01 or derivatives are associated to *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 Fig. 3, already described elsewhere (Lima et al., 2016), may explain these associations.

4. Discussion

HLA-G, -E and *-F* (or HLA class Ib genes) are primarily involved in the modulation of the immune response rather than in antigen presentation, which is a typical feature of the classical class I genes *(HLA-A, -B* and *-C)*. Although expressed in many tissues (except *HLA-G* which has tissue-restricted expression) in physiological conditions, *HLA-G and -E* are primarily expressed in the placenta, where they may be partially responsible for the acceptance of the semi-allogeneic fetus by the mother immune system cells. However, yet, it is not clear whether *HLA-F* do influence mother immune system cells. Besides the immunomodulatory effect at the fetus-mother interface, the physiological and pathological expression of HLA class Ib molecules may modulate the function of several cells of the innate and adaptive immune system. Due to these important functions, high selection pressures have acted on these genes maintaining HLA class Ib protein sequence and molecule structure as opposed to HLA classical class I genes, that have evolved under the pressure of a myriad of environmental pathogens to be adapted to millions of different antigens. Gene variable sites at coding regions of HLA class Ib genes, particularly the non-synonymous substitutions, may lead to protein modifications that may alter the molecule binding to leukocyte receptors, whereas variation sites at regulatory regions (enhancers, promoters and 3′UTRs) may modify the binding sites to transcription factors (mainly the promoter region) and posttranscriptional factors like microRNAs (3′UTR). A full description of the nucleotide diversity along the major HLA class Ib genes, considered either as gene segments or as whole gene haplotype, may provide a better understanding of how gene variation may account for protein modification or for the magnitude of protein production. In this context, the study of HLA class Ib genes in African populations, of whom other populations are supposed to descend from, may shed light into the evolution of these immune checkpoint genes. Of particular importance, the Toffin is a population that fled from Godomé Coast to escape from the war at Aja-Tado to establish their villages in the Northern shore of Nokoué lake, living as simple fishermen in the Beninese Sô-Ava region (Principaud, 1995). Since Toffins built their houses on stilts in Nokoué lake to give them safety from the attacks of the Dahomey soldiers, they are highly exposed to diseases such as schistosomiasis due to their lake activities (Ibikounlé et al., 2018, 2013, 2009), representing an interesting population at a genetic point of view.

4.1. Variability at the entire HLA-G, -E and -F genes

Most of the variable sites detected within the entire *HLA-G* (90.6%), *HLA-E* (86.5%) and *HLA-F* (88.2%) genes are also shared with other populations included in the *1000 Genomes phase 3* project (Auton et al., 2015) (Tables S1–S3). Although few remaining variable sites were not included in this public database, technical errors may be discarded, since the massive parallel sequencing data analysis strategies used here have proven to be accurate and reliable, with a low false positive rate (Castelli et al., 2017, 2015; Lima et al., 2016; Ramalho et al., 2017). Therefore, it is possible that these missing variable sites were not correctly detected by the *1000 Genomes consortium* or the consortium failed to detect them. A typical example in the MHC region refers to tri-allelic variable sites that the 1000 *Genomes Consortium* describes them as bi-allelic ones.

Table 9
HLA-E extended region (from −2143 to +4420) haplotypes encountered in a Beninese Toffin population sample.

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 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.
[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.

4.2. *HLA-G* variability

HLA-G is by far the most studied gene among the *HLA Ib* family. However, few studies have evaluated the complete gene variability, from promoter to 3'UTR, in worldwide populations (Castelli et al., 2011, 2014a, 2017; Ober et al., 2003; Pyo et al., 2006). Here, we evaluated the promoter region encompassing the positions −1 to −1377 (Table 2), the same already evaluated for the Beninese Tori sample (Gineau et al., 2015). Consistent with the results of previous studies, the promoter haplotype lineage PROMO-G010102a (33.2%) was the most frequent, followed by PROMO-G0104a (25.8%), PROMO-G010101a (8.7%) and PROMO-G0103a (1.3%) (Table 2), and usually, the PROMO-G0104a is not so frequent (Castelli et al., 2017, 2014a, 2014b, 2011; Gineau et al., 2015; Nilsson et al., 2016; Tan et al., 2005). As in the Tori ethnic group (Gineau et al., 2015), the *HLA-G* promoter region appeared to be highly conserved in Toffin, with a similar level of nucleotide diversity (0.0060 ± 0.0031) [Table 2 and Ref. (Gineau et al., 2015)] within this gene region. This is corroborated by the suggestion of balancing selection acting in this region [Tajima's *D* value (2.5522, p = 0.0087) (Table 2)], as also observed in many other populations (Castelli et al., 2017, 2011; Gineau et al., 2015; Ober et al., 2003; Santos et al., 2013; Tan et al., 2005).

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 3). Interestingly, 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*01:05 N (11.41%)] can be encoded by these haplotypes, reinforcing the hypothesis of conserved coding region or by 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**01:05 N (11.41%) and *G**01:03:01:02 (10.40%) (Table 3), two of them (*G**01:04:04 and *G**01:05 N) appeared at frequencies relatively higher than that observed worldwide (Castelli et al., 2014a). Although these two alleles are observed in Occidental regions like Europe and Asia at low frequency, they were always higher in regions on natural pressure (i.e., endemic for parasite infection), including African regions (Carlini et al., 2013; Castelli et al., 2014a) or Amerindian populations from 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).

The *HLA-G* 3'UTR region evaluated here revealed well-documented haplotypes (Carlini et al., 2013; Castelli et al., 2017, 2014a, 2011, 2010, Catamo et al., 2015, 2014; Consiglio et al., 2011; Courtin et al., 2013; de Albuquerque et al., 2016; Garcia et al., 2013; Garziera et al., 2015; Gineau et al., 2015; Hviid et al., 2006; Lucena-Silva et al., 2013,

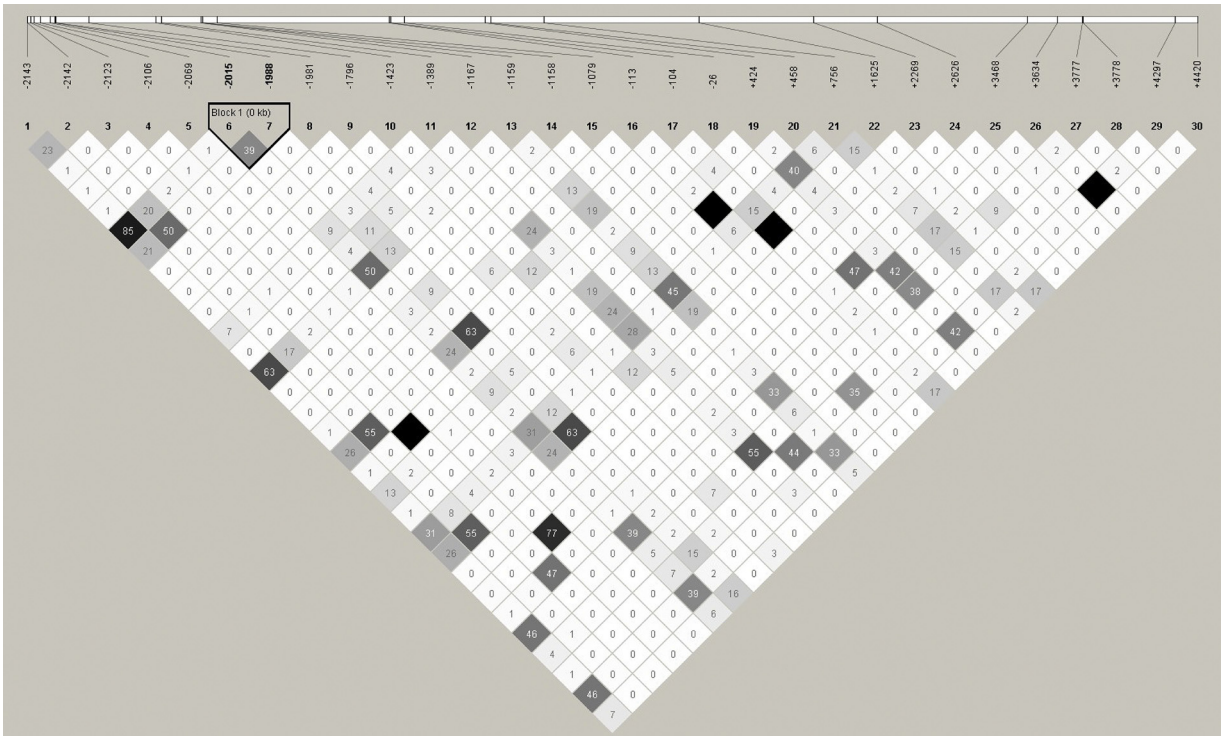


Fig. 2. LD patterns for the entire *HLA-E* region encompassing the positions -2143 to +4420 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (×100) 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%.

2012; Martelli-Palomino et al., 2013; Nilsson et al., 2016; Porto et al., 2015; Sabbagh et al., 2014; Santos et al., 2013; Sizzano et al., 2012; Tan et al., 2005; Veit et al., 2014, 2012; Zambra et al., 2016) (Table 4). The first six and 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 4), but were absent in Beninese Tori-Bossito (Tori ethnic group) sample (Sabbagh et al., 2014). The first possible explanation of this observation may be in 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). 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). Due to its endemicity for malaria, Tori-Bossito region has become the locality for various malaria research projects (D'Almeida et al., 2017, 2016, Dechavanne et al., 2017, 2016; Sadissou et al., 2014). Some studies revealed the effect of soluble HLA-G (sHLA-G) expression in mother/child pair susceptibility to *Plasmodium falciparum* (*P. f*) infection (D'Almeida et al., 2017, 2016; Sadissou et al., 2014), while other disclosed the natural humoral immunity acquisition against *P. f* in early life (Dechavanne et al., 2016), due to IgG3 transfer from mother to child (Dechavanne et al., 2017). Since there is no study focusing on the association between malaria and the *HLA-G* gene in the Toffin population, it would be important to implement one to show whether the *HLA-G* 3'UTR may be influenced by *P. f* malaria pressure. Another explanation may be related to a possible difference in genetic background between the two ethnic groups and genetic drift. This hypothesis may be justified by the presence of other haplotypes such as *HLA-G*-3UTR-20, -46, -47 and -48 in the Toffin group studied here, but absent in the Tori group (Table 4).

Regarding the *HLA-G* extended haplotypes, we found a strong association between the 5' UTR/coding/3'UTR segments (Table 5), 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 previously observed elsewhere (Castelli et al., 2017, 2014a, 2011, 2010; Donadi et al., 2011; Lucena-Silva et al., 2012; Nilsson et al., 2016; Santos et al., 2013). The balancing selection acting in the entire segment of the gene, as shown by the significant Tajima's *D* (Table 5), 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, our population missed *G*01:01:03/UTR-07 allele/haplotypes (Tables 3 and 4), which are frequent in Europe and highly frequent in Asia, but absent in African countries (Castelli et al., 2014a; Santos et al., 2013), suggesting that the haplotypes observed here and around the world may be quite old.

4.3. HLA-E variability

The *HLA-E* promoter region evaluated here was the same evaluated in the Brazilian sample (Ramalho et al., 2017). The Brazilian population has an important African contribution, but the one evaluated by Ramalho et al. was primarily composed by individuals of European ancestry (approximately 80%), which is typical from the Southeastern Brazilian population. Thus, the Brazilian study and ours evaluated very different samples in terms of ethnical composition. Contrary to Ramalho et al., most of the variable sites we detected in this gene segment

Table 10
HLA-F 5′ distal (A) and proximal (B) promoter haplotypes encountered in a Beninese Toffin population sample.

(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	T	T	T	A	G	T	G	A	C	G	T	C	T	G	T	C	0.1118
F*Distal-C	G	G	T	AG	T	T	G	A	T	T	C	C	C	A	T	G	0.2303
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). 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.

were polymorphic (i.e., with MAF ≥ 1%) and the number of promoter haplotypes that presented at least 1% of frequency was higher than those identified in Brazil (Table 6). This discrepancy may be explained by: **i)** the difference in nucleotide variability within promoter region between both populations [0.0013 ± 0.0008 (here) *versus* 0.0010 ± 0.0006 (in Brazil, (Ramalho et al., 2017), and **ii)** the negative Tajima's *D* value (−0.1764, $p = 0.5000$) (Table 6), reflecting a possible assumption of new alleles or population expansion (demography). In addition, some variants detected here with higher frequency were less frequent in Brazil (e.g. −104G with 2.38% in Brazil but with 12% in the present study). This is in agreement to a previous *HLA-E* study, evaluating African Guinea-Conakry and Burkina Faso samples (−104G with 11.39%) (Castelli et al., 2015). Moreover, while in Brazil the variants with lower frequency were between -440 to -1 (except position −104) (Ramalho et al., 2017; Veiga-Castelli et al., 2016), the variation sites we identified in the same region were more frequent (Table S2). However, the pattern of variability along the *HLA-E* gene between Brazil and the present studied population is similar. Indeed, likewise in Brazil, the promoter region variability (Table 6) is more pronounced than the coding (Table 7) and 3'UTR segments (Table 8). In addition, 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 (*E*-Promo-38 and *E*-Promo-39) (Table 6). Similarly to the Brazilian population, the frequent promoter variants identified here were also concentrated upstream position −1389 (Ramalho et al., 2017; Veiga-Castelli et al., 2016). In addition, the most frequent haplotype detected here was also *E*-Promo-01 (37.67%). Nevertheless, additional functional studies are needed to further show whether the

Table 11
HLA-F haplotypes encountered in a Beninese Toffin population sample considering the whole segment tracked 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

Notes:
(A) indicates the variable sites described in 1000 genome phase 3 data base (Table S3).
(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) 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.

[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). The total number of dinucleotide TG repeats is indicated when different from the known *HLA-F* allele. The word The term **"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.

promoter variability may influence the *HLA-E* function and expression.

Regarding the coding region, we identified 15 different alleles that mainly encoded the two most frequent and worldwide distributed proteins, known as E*01:01 and E*01:03 (Arnaiz-Villena et al., 2007; Carvalho dos Santos et al., 2013; Felício et al., 2014; Liu et al., 2012; Ramalho et al., 2017; Veiga-Castelli et al., 2012). These findings may be due to gene conservation and its implication on immune tolerance (Pietra et al., 2010; Sullivan et al., 2008). Indeed, the *HLA-E**01:01 was observed in 53.66% and *E**01:03 in 40.67% of the Toffin individuals, percentages that are closely similar to several worldwide studies (Arnaiz-Villena et al., 2007; Carvalho dos Santos et al., 2013; Felício et al., 2014; Liu et al., 2012; Ramalho et al., 2017; Veiga-Castelli et al., 2012). Altogether, the grouped frequency of these alleles accounted for approximately 94–95% of *HLA-E* encoded-molecules, a finding that is similar to that observed for our population (Table 7). This gene conservation may justify the few (6 out of 15) new alleles identified in our population, and the intronic or synonymous mutations occurring in this gene region was also shown elsewhere (Auton et al., 2015; Castelli et al., 2015; Ramalho et al., 2017). *HLA-E* coding region is the most conserved among 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 3, 7 and 11). In addition, the nucleotide diversity observed at the coding region (Table 7) was lower than those observed at the promoter (Table 6) and 3′UTR (Table 8) *HLA-E* segments. This observation agrees with Ramalho et al., who found the same pattern of *HLA-E* variability between the different gene segments (Ramalho et al., 2017).

Of the seven *HLA-E* 3'UTR haplotypes (Table 8), 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, where they were detected for the

Table 12
HLA-F extended region (from −1709 to +3250) haplotypes encountered in a Beninese Toffin population sample.

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-C	F*upstream-C	F*01:01:02:05(1943G, $[TG]_{11}$)	F*3UTR-A	0.1020
F*Distal-B	F*upstream-B	F*01:01:01:01($[TG]_{12}$)	F*3UTR-A	0.0954
F*Distal-C	F*upstream-C	F*01:01:02:03(1943G, 2208C, $[TG]_{12}$)	F*3UTR-A	0.0691
F*Distal-C	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-B	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-B	F*upstream-B	F*01:01:01:01(new145A)	F*3UTR-A	0.0033
F*Distal-B	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-C	F*upstream-F	F*01:01:02:03(1943G, $[TG]_{12}$)	F*3UTR-A	0.0033
F*Distal-C	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 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.

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

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

[c] As the 3'UTR haplotypes are represented by an unique variable site at position +3061 (Table S3), we contented to cite it in the text (at Sections 3.3.3 and 4.4). The extended sequences are available upon request.

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

first time (Ramalho et al., 2017). The *HLA-E* 3′ UTR segment is also very conserved (Castelli et al., 2015; Felício et al., 2014; Ramalho et al., 2017). In fact, this region showed the lowest nucleotide diversity when compared to the promoter and 3'UTR of other HLA Ib genes (Tables 2, 4, 8, and 10). This conservation at 3'UTR segment may reflect the low variability of HLA-E expression by post-transcriptional mechanisms, modulated by microRNA binding (Bartel, 2009, 2004) and mRNA secondary/tertiary structure (Chen et al., 2006). Noteworthy, the presence of the highly polymorphic variable site at position +3777 (Tables 8 and 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).

In agreement to previous *HLA-E* studies (Castelli et al., 2015; Felício et al., 2014; Ramalho et al., 2017), we did not find a strong association between the promoter/coding/3'UTR haplotypes (Table 9). This may be explained by the weak LD between these gene segments (Castelli et al., 2015; Felício et al., 2014; Ramalho et al., 2017) (Fig. 2). However, some associations detected here followed the same pattern already found by Ramalho et al., (Ramalho et al., 2017) (Table 9). 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 found along the *HLA-E* gene (Deininger, 2011; Kent et al., 2002; Ramalho et al., 2017).

4.4. HLA-F variability

The *HLA-F* promoter region was evaluated separately; i. e, the distal (from −1709 to −1013) and proximal (from −547 to −101) promoters (Table 10). Considering the *HLA-F* structure described in Section 3.3.1 (Lima et al., 2016), the F-distal promoter had already been evaluated in DNA samples from Asian, African American, and Caucasian populations (Pyo et al., 2006). We 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 10). This observation could be explained by the strong LD between *F**Distal and *F**proximal promoters (Fig. 3). 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 10 and 11). 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

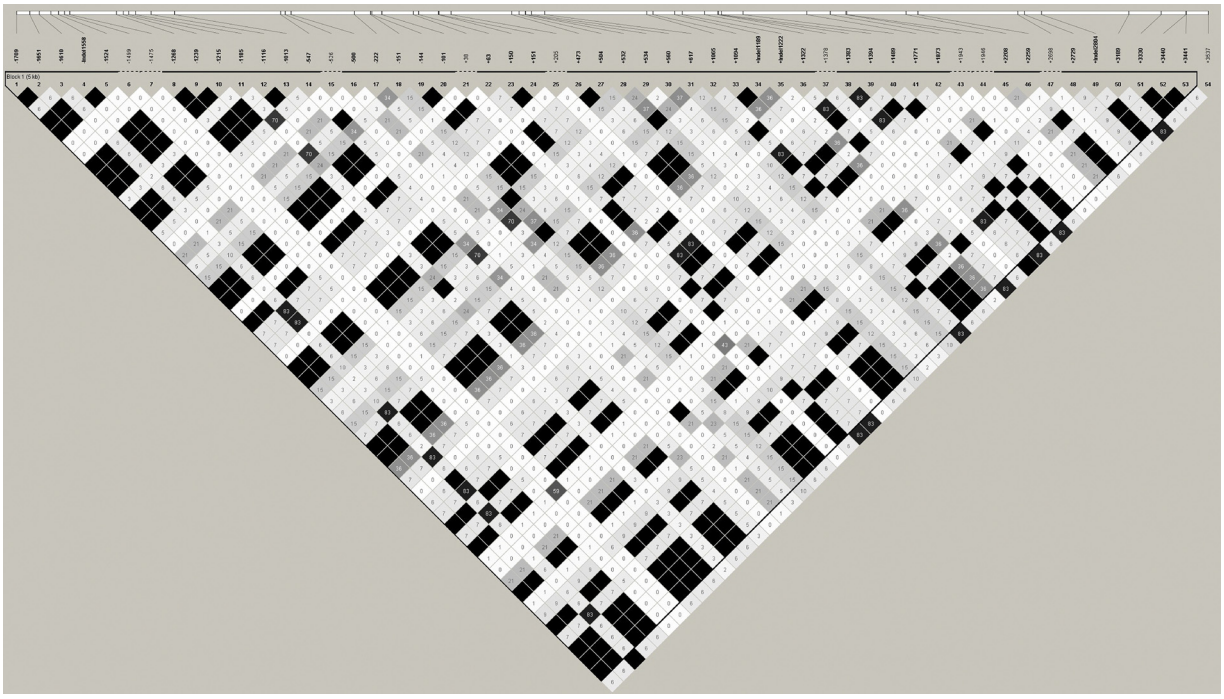


Fig. 3. LD patterns for the entire *HLA-F* region encompassing positions -1709 to +3537 in Beninese Toffin population. The Linkage Disequilibrium (LD) pattern was evaluated by calculating r^2. LD plot was generated and visualized by Haploview v4.2 (Barrett et al., 2005). High pairwise LD (r^2) between variants is illustrated with dark shading. The r^2 values (×100) 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%.

(Pyo et al., 2006), as also speculated elsewhere (Lima et al., 2016).

As reported for Brazilians, we did not observe some frequent polymorphisms described in IPD-IMGT/HLA, exemplifying, the deletion of G at position +1943 found at IPD-IMGT/HLA v.3.31.0 database. Considering all the 29 coding haplotypes detected (Table 11), they encoded only three different proteins: F*01:01, F*01:02 and F*01:03 (Table 11). 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 (Lima et al., 2016; Pyo et al., 2006) and by *1000 Genomes phase 3* project (Auton et al., 2015). The protein F*01:01 was present at 77.63% of the individuals, accompanied by F*01:03 (19.41%) and F*01:02 (2.96%) (Table 11). 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 high frequency and F*01:02, when detected, was always at low frequency (Lima et al., 2016; Manvailer et al., 2014; Moscoso et al., 2007; Pan et al., 2013; Pyo et al., 2006; Zhang et al., 2012). The presence of dinucleotide repeats at position +3097, detected here (Tables S3 and 11) and also elsewhere (Auton et al., 2015; Lima et al., 2016), confirmed the existence of microsatellites and consequently the existence of the new haplotypes harboring them. Notwithstanding, a proper nomenclature must be assigned by the WHO Nomenclature Committee for Factors of the HLA System.

As only one variable site was detected at 3' UTR at position +3061 (Table S3), only two 3'UTR haplotypes were detected for *HLA-F*. 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) with the *HLA-F**3UTR-B, detected in two unrelated individuals, was also associated to the same extended haplotype (Table 12). Likewise, the two alleles ($F^*01:02^{(1193C,\ 1943G)}$ and $F^*01:02^{(1193C,\ 1943G,\ new1946C)}$) detected here and absent in Brazil, were also associated to the same extended haplotype (Table 12).

Finally, it is interesting to notice that some *HLA-G, -E and -F* coding alleles (Tables 3, 7 and 11), which were shared with the Brazilian population (Castelli et al., 2017; Lima et al., 2016; Ramalho et al., 2017), were more frequent in our population than in Brazilians. For instance, in the present study, 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 7) were more frequent than in Brazilians (Ramalho et al., 2017), but exhibited closely 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 by Brazilians (Table 3 and 11). In addition, the difference in nucleotide diversity along the coding (Tables 3, 7 and 11) and extended (Tables 5, 9 and 12) 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.

4.5. Conclusion

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 (Tables 3 and 7) than at promoter (Tables 2 and 6) and at 3' UTR (Tables 4 and 8). *HLA-F* was more conserved at the regulatory [proximal promoter (Table 10) and 3'UTR (Table 12)] than in the coding (Table 11) and distal promoter (Table 10) 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.

Declaration of interest

None.

Acknowledgements

This work was supported by Institut de Recherche pour le Développement (IRD)/France (JEAI PALUCO); by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Brazil (Grant numbers PVE406 594/2013-9, 304 931/2014-1 and 466036/2013-5); by Brazil-France Research Cooperation Program USP/COFECUB (Grant numbers Uc Me 169-17). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001/CAPES/PROEX0055042/1768/2013. A PhD scholarship was awarded by IRD to Léonidas Tokplonou. The authors thank the school children and parents for their free consent and the teachers of the Sô-Ava district for their great collaboration.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2018.08.016.

References

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

Allan, D.S., Colonna, M., Lanier, L.L., Churakova, T.D., Abrams, J.S., Ellis, S.A., McMichael, A.J., Braud, V.M., 1999. Tetrameric complexes of human histocompatibility leukocyte antigen (HLA)-G bind to peripheral blood myelomonocytic cells. J. Exp. Med. 189, 1149–1156. https://doi.org/10.1084/jem.189.7.1149.

Arnaiz-Villena, A., Vargas-Alarcon, G., Serrano-Vela, J.I., Reguera, R., Martinez-Laso, J., Silvera-Redondo, C., Granados, J., Moscoso, J., 2007. HLA-E polymorphism in Amerindians from Mexico (Mazatecans), Colombia (Wayu) and Chile (Mapuches): evolution of MHC-E gene. Tissue Antigens 132–135. https://doi.org/10.1111/j.1399-0039.2006.763_2.x.

Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A., Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G., Zhu, Y., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Lan, T., Li, G., Li, J., Li, Y., Liu, S., Liu, X., Lu, Y., Ma, X., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Xu, X., Yin, Y., Zhang, D., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Gupta, N., Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., Barker, J., Clarke, L., Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., McLaren, W.M., Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thormann, A., Toneva, I., Vaughan, B., Zheng-Bradley, X., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.L., Fulton, L., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C., Church, D., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Davies, C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Campbell, C.L., Kong, Y., Marcketta, A., Yu, F., Antunes, L., Bainbridge, M., Sabo, A., Huang, Z., Coin, L.J.M., Fang, L., Li, Q., Li, Z., Lin, H., Liu, B., Luo, R., Shao, H., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Garrison, E.P., Kural, D., Lee, W.P., Leong, W.F., Stromberg, M., Ward, A.N., Wu, J., Zhang, M., Daly, M.J., DePristo, M.A., Handsaker, R.E., Banks, E., Bhatia, G., Del Angel, G., Genovese, G., Li, H., Kashin, S., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Yoon, S.C., Lihm, J., Makarov, V., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., Rausch, T., Fritz, M.H., Stütz, A.M., Beal, K., Datta, A., Herrero, J., Ritchie, G.R.S., Zerbino, D., Sabeti, P.C., Shlyakhter, I., Schaffner, S.F., Vitti, J., Cooper, D.N., Ball, E.V., Stenson, P.D., Barnes, B., Bauer, M., Cheetham, R.K., Cox, A., Eberle, M., Kahn, S., Murray, L., Peden, J., Shaw, R., Kenny, E.E., Batzer, M.A., Konkel, M.K., Walker, J.A., MacArthur, D.G., Lek, M., Herwig, R., Ding, L., Koboldt, D.C., Larson, D., Ye, K., Gravel, S., Swaroop, A., Chew, E., Lappalainen, T., Erlich, Y., Gymrek, M., Willems, T.F., Simpson, J.T., Shriver, M.D., Rosenfeld, J.A., Bustamante, C.D., Montgomery, S.B., De La Vega, F.M., Byrnes, J.K., Carroll, A.W., DeGorter, M.K., Lacroute, P., Maples, B.K., Martin, A.R., Moreno-Estrada, A., Shringarpure, S.S., Zakharia, F., Halperin, E., Baran, Y., Cerveira, E., Hwang, J., Malhotra, A., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Hyland, F.C.L., Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A.A., Sinari, S.A., Squire, K., Xiao, C., Sebat, J., Antaki, D., Gujral, M., Noor, A., Ye, K., Burchard, E.G., Hernandez, R.D., Gignoux, C.R., Haussler, D., Katzman, S.J., Kent, W.J., Howie, B., Ruiz-Linares, A., Dermitzakis, E.T., Devine, S.E., Kang, H.M., Kidd, J.M., Blackwell, T., Caron, S., Chen, W., Emery, S., Fritsche, L., Fuchsberger, C., Jun, G., Li, B., Lyons, R., Scheller, C., Sidore, C., Song, S., Sliwerska, E., Taliun, D., Tan, A., Welch, R., Wing, M.K., Zhan, X., Awadalla, P., Hodgkinson, A., Li, Y., Shi, X., Quitadamo, A., Lunter, G., Marchini, J.L., Myers, S., Churchhouse, C., Delaneau, O., Gupta-Hinch, A., Kretzschmar, W., Iqbal, Z., Mathieson, I., Menelaou, A., Rimmer, A., Xifara, D.K., Oleksyk, T.K., Fu, Y., Liu, X., Xiong, M., Jorde, L., Witherspoon, D., Xing, J., Browning, B.L., Browning, S.R., Hormozdiari, F., Sudmant, P.H., Khurana, E., Tyler-Smith, C., Albers, C.A., Ayub, Q., Chen, Y., Colonna, V., Jostins, L., Walter, K., Xue, Y., Gerstein, M.B., Abyzov, A., Balasubramanian, S., Chen, J., Clarke, D., Fu, Y., Harmanci, A.O., Jin, M., Lee, D., Liu, J., Mu, X.J., Zhang, J., Zhang, Y., Hartl, C., Shakir, K., Degenhardt, J., Meiers, S., Raeder, B., Casale, F.P., Stegle, O., Lameijer, E.W., Hall, I., Bafna, V., Michaelson, J., Gardner, E.J., Mills, R.E., Dayama, G., Chen, K., Fan, X., Chong, Z., Chen, T., Chaisson, M.J., Huddleston, J., Malig, M., Nelson, B.J., Parrish, N.F., Blackburne, B., Lindsay, S.J., Ning, Z., Zhang, Y., Lam, H., Sisu, C., Challis, D., Evani, U.S., Lu, J., Nagaswamy, U., Yu, J., Li, W., Habegger, L., Yu, H., Cunningham, F., Dunham, I., Lage, K., Jespersen, J.B., Horn, H., Kim, D., Desalle, R., Narechania, A., Sayres, M.A.W., Mendez, F.L., Poznik, G.D., Underhill, P.A., Mittelman, D., Banerjee, R., Cerezo, M., Fitzgerald, T.W., Louzada, S., Massaia, A., Yang, F., Kalra, D., Hale, W., Dan, X., Barnes, K.C., Beiswanger, C., Cai, H., Cao, H., Henn, B., Jones, D., Kaye, J.S., Kent, A., Kerasidou, A., Mathias, R., Ossorio, P.N., Parker, M., Rotimi, C.N., Royal, C.D., Sandoval, K., Su, Y., Tian, Z., Tishkoff, S., Via, M., Wang, Y., Yang, H., Yang, L., Zhu, J., Bodmer, W., Bedoya, G., Cai, Z., Gao, Y., Chu, J., Peltonen, L., Garcia-Montero, A., Orfao, A., Dutil, J., Martinez-Cruzado, J.C., Mathias, R.A., Hennis, A., Watson, H., McKenzie, C., Qadri, F., LaRocque, R., Deng, X., Asogun, D., Folarin, O., Happi, C., Omoniwa, O., Stremlau, M., Tariyal, R., Jallow, M., Joof, F.S., Corrah, T., Rockett, K., Kwiatkowski, D., Kooner, J., Hien, T.T., Dunstan, S.J., ThuyHang, N., Fonnie, R., Garry, R., Kanneh, L., Moses, L., Schieffelin, J., Grant, D.S., Gallo, C., Poletti, G., Saleheen, D., Rasheed, A., Brooks, L.D., Felsenfeld, A.L., McEwen, J.E., Vaydylevich, Y., Duncanson, A., Dunn, M., Schloss, J.A., 2015. A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263–265. https://doi.org/10.1093/bioinformatics/bth457.

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116 (2), 281–297. https://doi.org/10.1016/S0092-8674(04)00045-5.

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136 (2), 212–233. https://doi.org/10.1016/j.cell.2009.01.002.

Baudhuin, J., Migraine, J., Faivre, V., Loumagne, L., Lukaszewicz, A.-C., Payen, D., Favier, B., 2013. Exocytosis acts as a modulator of the ILT4-mediated inhibition of neutrophil functions. Proc. Natl. Acad. Sci. 110, 17957–17962. https://doi.org/10.1073/pnas.1221535110.

Bian, X., Si, Y., Zhang, M., Wei, R., Yang, X., Ren, H., Zheng, G., Wang, C., Zhang, Y., 2016. Down-expression of miR-152 lead to impaired anti-tumor effect of NK via upregulation of HLA-G. Tumor Biol. 37, 3749–3756. https://doi.org/10.1007/s13277-015-3669-7.

Borrego, F., Ulbrecht, M., Weiss, E.H., Coligan, J.E., Brooks, A.G., 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, 813–818. https://doi.org/10.1084/jem.187.5.813.

Braud, V.M., Allan, D.S., McMichael, A.J., 1999. Functions of nonclassical MHC and non-MHC-encoded class I molecules. Curr. Opin. Immunol. 11 (1), 100–108. https://doi.org/10.1016/S0952-7915(99)80018-1.

Carlini, F., Traore, K., Cherouat, N., Roubertoux, P., Buhler, S., Cortey, M., Simon, S., Doumbo, O., Chiaroni, J., Picard, C., Di Cristofaro, J., 2013. HLA-G UTR haplotype conservation in the Malian population: association with soluble HLA-G. PLoS One 8. https://doi.org/10.1371/journal.pone.0082517.

Carosella, E.D., Rouas-Freiss, N., Le Roux, D.T., Moreau, P., LeMaoult, J., 2015. HLA-G. An immune checkpoint molecule. Advances in Immunology, 1st ed. Elsevier Inc.https://doi.org/10.1016/bs.ai.2015.04.001.

Carvalho dos Santos, L., Tureck, L.V., Wowk, P.F., Mattar, S.B., Gelmini, G.F., Magalhães, J.C.M., da G. Bicalho, M., Roxo, V.M.M.S., 2013. HLA-E polymorphisms in an Afro-descendant Southern Brazilian population. Hum. Immunol. 74, 199–202. https://doi.org/10.1016/j.humimm.2012.10.020.

Castelli, E.C., Mendes-Junior, C.T., Deghaide, N.H.S., de Albuquerque, R.S., Muniz, Y.C.N., Simões, R.T., Carosella, E.D., Moreau, P., Donadi, E.A., 2010. The genetic structure of 3′ untranslated region of the HLA-G gene: polymorphisms and haplotypes. Genes Immun. 11, 134–141. https://doi.org/10.1038/gene.2009.74.

Castelli, E.C., Mendes-Junior, C.T., Veiga-Castelli, L.C., Roger, M., Moreau, P., Donadi, E.A., 2011. A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution. Mol. Biol. Evol. 28, 3069–3086. https://doi.org/10.1093/molbev/msr138.

Castelli, E.C., Ramalho, J., Porto, I.O.P., Lima, T.H.A., Felício, L.P., Sabbagh, A., Donadi, E.A., Mendes-Junior, C.T., 2014a. Insights into HLA-G genetics provided by worldwide haplotype diversity. Front. Immunol. 5. https://doi.org/10.3389/fimmu.2014.00476.

Castelli, E.C., Veiga-Castelli, L.C., Yaghi, L., Moreau, P., Donadi, E.A., 2014b. Transcriptional and posttranscriptional regulations of the HLA-G gene. J. Immunol. Res. 2014. https://doi.org/10.1155/2014/734068.

Castelli, E.C., Mendes-Junior, C.T., Sabbagh, A., Porto, I.O.P., Garcia, A., Ramalho, J., Lima, T.H.A., Massaro, J.D., Dias, F.C., Collares, C.V.A., Jamonneau, V., Bucheton, B., Camara, M., Donadi, E.A., 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. https://doi.org/10.1016/j.humimm.2015.06.016.

Castelli, E.C., Gerasimou, P., Paz, M.A., Ramalho, J., Porto, I.O.P., Lima, T.H.A., Souza, A.S., Veiga-Castelli, L.C., Collares, C.V.A., Donadi, E.A., Mendes-Junior, C.T., Costeas, P., 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. https://doi.org/10.1016/j.molimm.2017.01.020.

Catamo, E., Addobbati, C., Segat, L., Sotero Fragoso, T., Domingues Barbosa, A., Tavares Dantas, A., de Ataíde Mariz, H., da Rocha, L.F., Branco Pinto Duarte, A.L., Monasta, L., Sandrin-Garcia, P., Crovella, S., 2014. HLA-G gene polymorphisms associated with susceptibility to rheumatoid arthritis disease and its severity in Brazilian patients. Tissue Antigens 84, 308–315. https://doi.org/10.1111/tan.12396.

Catamo, E., Addobbati, C., Segat, L., Sotero Fragoso, T., Tavares Dantas, A., de Ataide Mariz, H., Ferreira da Rocha Junior, L., Branco Pintoduarte, A.L., Coelho, A.V.C., de Moura, R.R., Polesello, V., Crovella, S., Sandrin Garcia, P., 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. https://doi.org/10.1111/tan.12545.

Chen, J.M., Férec, C., Cooper, D.N., 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 (3), 301–333. https://doi.org/10.1007/s00439-006-0218-x.

Consiglio, C.R., Veit, T.D., Monticielo, O.A., Mucenic, T., Xavier, R.M., Brenol, J.C.T., Chies, J.A.B., 2011. Association of the HLA-G gene +3142C&G polymorphism with systemic lupus erythematosus. Tissue Antigens 77, 424–430. https://doi.org/10.1111/j.1399-0039.2011.01635.x.

Courtin, D., Milet, J., Sabbagh, A., Massaro, J.D., Castelli, E.C., Jamonneau, V., Bucheton, B., Sese, C., Favier, B., Rouas-Freiss, N., Moreau, P., Donadi, E.A., Garcia, A., 2013. HLA-G 3' UTR-2 haplotype is associated with Human African trypanosomiasis susceptibility. Infect. Genet. Evol. 17, 1–7. https://doi.org/10.1016/j.meegid.2013.03.004.

D'Almeida, T.C., Sadissou, I., Cottrell, G., Tahar, R., Moreau, P., Favier, B., Moutairou, K., Donadi, E.A., Massougbodji, A., Rouass-Freiss, N., Courtin, D., Garcia, A., d'Almeida, T.C., Sadissou, I., Cottrell, G., Tahar, R., Moreau, P., Favier, B., Moutairou, K., Donadi, E.A., Massougbodji, A., Rouass-Freiss, N., Courtin, D., Garcia, A., 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. https://doi.org/10.1186/s12936-016-1131-y.

D'Almeida, T.C., Sadissou, I., Milet, J., Cottrell, G., Mondière, A., Avokpaho, E., Gineau, L., Sabbagh, A., Massougbodji, A., Moutairou, K., Donadi, E.A., Favier, B., Carosella, E., Moreau, P., Rouas-Freiss, N., Courtin, D., Garcia, A., d'Almeida, T.C., Sadissou, I., Milet, J., Cottrell, G., Mondière, A., Avokpaho, E., Gineau, L., Sabbagh, A., Massougbodji, A., Moutairou, K., Donadi, E.A., Favier, B., Carosella, E., Moreau, P., Rouas-Freiss, N., Courtin, D., Garcia, A., 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. https://doi.org/10.1371/journal.pone.0171117.

de Albuquerque, R.S., Mendes-Junior, C.T., Lucena-Silva, N., da Silva, C.L.L., Rassi, D.M., Veiga-Castelli, L.C., Foss-Freitas, M.C., Foss, M.C., Deghaide, N.H.S., Moreau, P., Gregori, S., Castelli, E.C., Donadi, E.A., 2016. Association of HLA-G 3' untranslated region variants with type 1 diabetes mellitus. Hum. Immunol. 77, 358–364. https://doi.org/10.1016/j.humimm.2016.02.001.

Dechavanne, C., Sadissou, I., Bouraima, A., Ahouangninou, C., Amoussa, R., Milet, J., Moutairou, K., Massougbodji, A., Theisen, M., Remarque, E.J., Courtin, D., Nuel, G., Migot-Nabias, F., Garcia, A., 2016. Acquisition of natural humoral immunity to P. falciparum in early life in Benin: impact of clinical, environmental and host factors. Sci. Rep. 6, 1–9. https://doi.org/10.1038/srep33961.

Dechavanne, C., Dechavanne, S., Sadissou, I., Lokossou, A.G., Alvarado, F., Dambrun, M., Moutairou, K., Courtin, D., Nuel, G., Garcia, A., Migot-Nabias, F., King, C.L., 2017. Associations between an IgG3 polymorphism in the binding domain for FcRn, transplacental transfer of malaria-specific IgG3, and protection against Plasmodium falciparum malaria during infancy: a birth cohort study in Benin. PLoS Med. 14, 1–19. https://doi.org/10.1371/journal.pmed.1002403.

Deininger, P., 2011. Alu elements: know the SINEs. Genome Biol. 12, 236. https://doi.org/10.1186/gb-2011-12-12-236.

DePristo, M.A., Banks, E., Poplin, R.E., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498. https://doi.org/10.1038/ng.806.

Devlin, B., Risch, N., 1995. A comparison of linkage disequilibrium measures for fine-scale mappingitle. Genomics 29, 311–322. https://doi.org/10.1006/geno.1995.9003. S0888-7543(85)79003-9 [pii].

Dias, F.C., Castelli, E.C., Collares, C.V.A., Moreau, P., Donadi, E.A., 2015. The role of HLA-G molecule and HLA-G gene polymorphisms in tumors, viral hepatitis, and parasitic diseases. Front. Immunol. 6 (Feb), 1–10. https://doi.org/10.3389/fimmu.2015.00009.

Djenontin, A., Bio-Bangana, S., Moiroux, N., Henry, M.C., Bousari, O., Chabi, J., Osse, R., Koudenoukpo, S., Corbel, V., Akogbeto, M., Chandre, F., 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. https://doi.org/10.1186/1756-3305-3-83.

Donadi, E.A., Castelli, E.C., Arnaiz-Villena, A., Roger, M., Rey, D., Moreau, P., 2011. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. Cell. Mol. Life Sci. 68, 369–395. https://doi.org/10.1007/s00018-010-0580-7.

Dulberger, C.L., McMurtrey, C.P., Hölzemer, A., Neu, K.E., Liu, V., Steinbach, A.M., Garcia-Beltran, W.F., Sulak, M., Jabri, B., Lynch, V.J., Altfeld, M., Hildebrand, W.H., Adams, E.J., 2017. Human leukocyte antigen F presents peptides and regulates immunity through interactions with NK cell receptors. Immunity 46, 1018–1029. https://doi.org/10.1016/j.immuni.2017.06.002. e7.

Eguchi, H., Maeda, A., Lo, P.C., Matsuura, R., Esquivel, E.L., Asada, M., Sakai, R., Nakahata, K., Yamamichi, T., Umeda, S., Deguchi, K., Ueno, T., Okuyama, H., Miyagawa, S., 2016. HLA-G1, but not HLA-G3, suppresses human monocyte/macrophage-mediated swine endothelial cell lysis. Transplant. Proc. 48, 1285–1287. https://doi.org/10.1016/j.transproceed.2015.10.070.

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. https://doi.org/10.1016/j.ympev.2005.04.031.

Felício, L.P., Porto, I.O.P., Mendes-Junior, C.T., Veiga-Castelli, L.C., Santos, K.E., Vianello-Brondani, R.P., Sabbagh, A., Moreau, P., Donadi, E.A., Castelli, E.C., 2014. Worldwide HLA-E nucleotide and haplotype variability reveals a conserved gene for coding and 3′ untranslated regions. Tissue Antigens 83, 82–93. https://doi.org/10.1111/tan.12283.

Fournel, S., Aguerre-Girr, M., Huc, X., Lenfant, F., Alam, A., Toubert, A., Bensussan, A., Le Bouteiller, P., 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. https://doi.org/10.4049/jimmunol.164.12.6100.

Garcia, A., Milet, J., Courtin, D., Sabbagh, A., Massaro, J.D., Castelli, E.C., Migot-Nabias, F., Favier, B., Rouas-Freiss, N., Donadi, E.A., Moreau, P., 2013. Association of HLA-G 3'UTR polymorphisms with response to malaria infection: a first insight. Infect. Genet. Evol. 16, 263–269. https://doi.org/10.1016/j.meegid.2013.02.021.

Garcia-Beltran, W.F., Hölzemer, A., Martrus, G., Chung, A.W., Pacheco, Y., Simoneau, C.R., Rucevic, M., Lamothe-Molina, P.A., Pertel, T., Kim, T.E., Dugan, H., Alter, G., Dechanet-Merville, J., Jost, S., Carrington, M., Altfeld, M., 2016. Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. Nat. Immunol. 17, 1067–1074. https://doi.org/10.1038/ni.3513.

Garziera, M., Bidoli, E., Cecchin, E., Mini, E., Nobili, S., Lonardi, S., Buonadonna, A., Errante, D., Pella, N., D'Andrea, M., De Marchi, F., De Paoli, A., Zanusso, C., De Mattia, E., Tassi, R., Toffoli, G., 2015. HLA-G 3'UTR polymorphisms impact the prognosis of stage II-III CRC patients in fluoropyrimidine-based treatment. PLoS One 10. https://doi.org/10.1371/journal.pone.0144000.

Geraghty, D.E., Koller, B.H., Hansen, J.A., Orr, H.T., 1992. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. J. Immunol. 149, 1934–1946. https://doi.org/10.4049/jimmunol.171.10.5264.

Gineau, L., Luisi, P., Castelli, E.C., Milet, J., Courtin, D., Cagnin, N., Patillon, B., Laayouni, H., Moreau, P., Donadi, E.A., Garcia, A., Sabbagh, A., 2015. Balancing immunity and tolerance: genetic footprint of natural selection in the transcriptional regulatory region of HLA-G. Genes Immun. 16, 57–70. https://doi.org/10.1038/gene.2014.63.

Goodridge, J.P., Burian, A., Lee, N., Geraghty, D.E., 2010. HLA-F complex without peptide binds to MHC class I protein in the open conformer form. J. Immunol. 184, 6199–6208. https://doi.org/10.4049/jimmunol.1000078.

Goodridge, J.P., Burian, A., Lee, N., Geraghty, D.E., 2013. HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. J. Immunol. 191, 3553–3562. https://doi.org/10.4049/jimmunol.1300081.

Grange, C., Tapparo, M., Tritta, S., Deregibus, M.C., Battaglia, A., Gontero, P., Frea, B., Camussi, G., 2015. Role of HLA-G and extracellular vesicles in renal cancer stem cell-induced inhibition of dendritic cell differentiation. BMC Cancer 15, 1009. https://doi.org/10.1186/s12885-015-2025-z.

Guo, S.W., Thompson, E.A., 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48, 361–372. https://doi.org/10.2307/2532296.

Hackmon, R., Pinnaduwage, L., Zhang, J., Lye, S.J., Geraghty, D.E., Dunk, C.E., 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. https://doi.org/10.1111/aji.12643.

Harada, A., Ishigami, S., Kijima, Y., Nakajo, A., Arigami, T., Kurahara, H., Kita, Y., Yoshinaka, H., Natsugoe, S., 2015. Clinical implication of human leukocyte antigen (HLA)-F expression in breast cancer. Pathol. Int. 65, 569–574. https://doi.org/10.1111/pin.12343.

Huang, J., Burke, P., Yang, Y., Seiss, K., Beamon, J., Cung, T., Toth, I., Pereyra, F., Lichterfeld, M., Yu, X.G., 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. https://doi.org/10.1128/JVI.01292-10.

Hviid, T.V.F., Rizzo, R., Melchiorri, L., Stignani, M., Baricordi, O.R., 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. https://doi.org/10.1016/j.humimm.2005.12.003.

Ibikounlé, M., Mouahid, G., Sakiti, N.G., Massougbodji, A., Moné, H., 2009. Freshwater snail diversity in Benin (West Africa) with a focus on human schistosomiasis. Acta Trop. 111, 29–34. https://doi.org/10.1016/j.actatropica.2009.02.001.

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

Ibikounlé, M., Onzo-Aboki, A., Doritchamou, J., Tougoué, J.J., Boko, P.M., Savassi, B.S., Siko, E.J., Daré, A., Batcho, W., Massougbodji, A., Kindé-Gazard, D.A., Kaboré, A., 2018. Results of the first mapping of soil-transmitted helminths in Benin: evidence of countrywide hookworm predominance. PLoS Negl. Trop. Dis. 12, 1–17. https://doi.org/10.1371/journal.pntd.0006241.

Ishitani, A., Sageshima, N., Lee, N., Dorofeeva, N., Hatake, K., Marquardt, H., Geraghty, D.E., 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. https://doi.org/10.4049/jimmunol.171.3.1376.
Ishitani, A., Sageshima, N., Hatake, K., 2006. The involvement of HLA-E and -F in pregnancy. J. Reprod. Immunol. 69 (2), 101–113. https://doi.org/10.1016/j.jri.2005.10.004.
Jørgensen, P.B., Livbjerg, A.H., Hansen, H.J., Petersen, T., Höllsberg, P., 2012. Epstein-Barr virus peptide presented by HLA-E is predominantly recognized by CD8bright cells in multiple sclerosis patients. PLoS One 7. https://doi.org/10.1371/journal.pone.0046120.
Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, A.D., 2002. The human genome browser at UCSC. Genome Res. 12, 996–1006. https://doi.org/10.1101/gr.229102.
Kim, S.K., Hong, M.S., Shin, M.K., Uhm, Y.K., Chung, J.H., Lee, M.H., 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. https://doi.org/10.1007/s00403-011-1160-x.
Kochan, G., Escors, D., Breckpot, K., Guerrero-Setas, D., 2013. Role of non-classical MHC class I molecules in cancer immunosuppression. Oncoimmunology 2 (11), 1–8. https://doi.org/10.4161/onci.26491.
Koller, B.H., Geraghty, D.E., Shimizu, Y., DeMars, R., Orr, H.T., 1988. HLA-E. A novel HLA class I gene expressed in resting T lymphocytes. J. Immunol. 141, 897–904.
Köstlin, N., Ostermeir, A.L., Spring, B., Schwarz, J., Marmé, A., Walter, C.B., Poets, C.F., Gille, C., 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. https://doi.org/10.1002/eji.201646564.
Kovats, S., Main, E.K., Librach, C., Stubblebine, M., Fisher, S.J., DeMars, R., 1990. A class I antigen, HLA-G, expressed in human trophoblasts. Science 248, 220–223. https://doi.org/10.1126/science.2326636.
Kulski, J.K., Martinez, P., Longman-Jacobsen, N., Wang, W., Williamson, J., Dawkins, R.L., Shiina, T., Naruse, T., Inoko, H., 2001. The association between HLA-A alleles and an Alu dimorphism near HLA-G. J. Mol. Evol. 53, 114–123. https://doi.org/10.1007/s002390010199.
Le Port, A., Cottrell, G., Martin-Prevel, Y., Migot-Nabias, F., Cot, M., Garcia, A., 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. https://doi.org/10.1136/bmjopen-2011-000342.
Lee, N., Goodlett, D.R., Ishitani, A., Marquardt, H., Geraghty, D.E., 1998. HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. J. Immunol. 160, 4951–4960. https://doi.org/10.1038/nature01413.
Lila, N., Rouas-Freiss, N., Dausset, J., Carpentier, a, Carosella, E.D., 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, 12150–12155. https://doi.org/10.1073/pnas.201407398.
Lima, T.H.A., Buttura, R.V., Donadi, E.A., Veiga-Castelli, L.C., Mendes-Junior, C.T., Castelli, E.C., 2016. HLA-F coding and regulatory segments variability determined by massively parallel sequencing procedures in a Brazilian population sample. Hum. Immunol. 77, 841–853. https://doi.org/10.1016/j.humimm.2016.07.231.
Liu, X.X., Pan, F.H., Tian, W., 2012. Characterization of HLA-E polymorphism in four distinct populations in Mainland China. Tissue Antigens 80, 26–35. https://doi.org/10.1111/j.1399-0039.2012.01873.x.
Lucena-Silva, N., Monteiro, A.R., de Albuquerque, R.S., Gomes, R.G., Mendes-Junior, C.T., Castelli, E.C., Donadi, E.A., 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. https://doi.org/10.1111/j.1399-0039.2012.01842.x.
Lucena-Silva, N., de Souza, V.S.B., Gomes, R.G., Fantinatti, A., Muniz, Y.C.N., de Albuquerque, R.S., Monteiro, A.L.R., Diniz, G.T.N., Coelho, M.R.C.D., Mendes-Junior, C.T., Castelli, Ed.C., Donadi, E.A., 2013. HLA-G 3' untranslated region polymorphisms are associated with systemic lupus erythematosus in 2 Brazilian populations. J. Rheumatol. 40, 1104–1113. https://doi.org/10.3899/jrheum.120814.
Manvailer, L.F.S., Wowk, P.F., Mattar, S.B., da Siva, J.S., da Graça Bicalho, M., Roxo, V.M.M.S., 2014. HLA-F polymorphisms in a Euro-Brazilian population from Southern Brazil. Tissue Antigens 84, 554–559. https://doi.org/10.1111/tan.12444.
Marchesi, M., Andersson, E., Villabona, L., Seliger, B., Lundqvist, A., Kiessling, R., Masucci, G.V., 2013. HLA-dependent tumour development: a role for tumour associate macrophages? J. Transl. Med. 11. https://doi.org/10.1186/1479-5876-11-247.
Martelli-Palomino, G., Pancotto, J.A., Muniz, Y.C., Mendes-Junior, C.T., Castelli, E.C., Massaro, J.D., Krawice-Radanne, I., Poras, I., Rebmann, V., Carosella, E.D., Rouas-Freiss, N., Moreau, P., Donadi, E.A., 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. https://doi.org/10.1371/journal.pone.0071742.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110.
Mendes-Junior, C.T., Castelli, E.C., Meyer, D., Simões, A.L., Donadi, E.A., 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. https://doi.org/10.1038/gene.2013.47.
Morandi, F., Ferretti, E., Bocca, P., Prigione, I., Raffaghello, L., Pistoia, V., 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. https://doi.org/10.1371/journal.pone.0011763.
Moscoso, J., Serrano-Vela, J.I., Arnaiz-Villena, A., 2007. MHC-F polymorphism and evolution. Tissue Antigens. 136–139. https://doi.org/10.1111/j.1399-0039.2006.763_3.x.
Naji, A., Menier, C., Morandi, F., Agaugue, S., Maki, G., Ferretti, E., Bruel, S., Pistoia, V., Carosella, E.D., Rouas-Freiss, N., 2014. Binding of HLA-G to ITIM-bearing Ig-like transcript 2 receptor suppresses B cell responses. J. Immunol. 192, 1536–1546. https://doi.org/10.4049/jimmunol.1300438.
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.
Nilsson, L.L., Djurisic, S., Andersen, A.M.N., Melbye, M., Bjerre, D., Ferrero-Miliani, L., Hackmon, R., Geraghty, D.E., Hviid, T.V.F., 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. https://doi.org/10.1111/tan.12871.
Ober, C., Aldrich, C.L., Chervoneva, I., Billstrand, C., Rahimov, F., Gray, H.L., Hyslop, T., 2003. Variation in the HLA-G promoter region influences miscarriage rates. Am. J. Hum. Genet. 72, 1425–1435. https://doi.org/10.1086/375501.
Olieslagers, T.I., Voorter, C.E.M., Groeneweg, M., Xu, Y., Wieten, L., Tilanus, M.G.J., 2017. New insights in HLA-E polymorphism by refined analysis of the full-length gene. HLA 89, 143–149. https://doi.org/10.1111/tan.12965.
Pan, F.H., Liu, X.X., Tian, W., 2013. Characterization of HLA-F polymorphism in four distinct populations in Mainland China. Int. J. Immunogenet. 40, 369–376. https://doi.org/10.1111/iji.12047.
Persson, G., Melsted, W.N., Nilsson, L.L., Hviid, T.V.F., 2017. HLA class Ib in pregnancy and pregnancy-related disorders. Immunogenetics. 69 (8–9), 581–595. https://doi.org/10.1007/s00251-017-0988-4.
Petroff, M.G., Sedlmayr, P., Azzola, D., Hunt, J.S., 2002. Decidual macrophages are potentially susceptible to inhibition by class Ia and class Ib HLA molecules. J. Reprod. Immunol. 56, 3–17. https://doi.org/10.1016/S0165-0378(02)00024-4.
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. https://doi.org/10.2174/138161209789105207.
Pietra, G., Romagnani, C., Manzini, C., Moretta, L., Mingari, M.C., 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. https://doi.org/10.1155/2010/907092.
Porto, I.O.P., Mendes-Junior, C.T., Felício, L.P., Georg, R.C., Moreau, P., Donadi, E.A., Chies, J.A.B., Castelli, E.C., 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. https://doi.org/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. pp. 519–546.
Pyo, C.W., Williams, L.M., Moore, Y., Hyodo, H., Li, S.S., Zhao, L.P., Sageshima, N., Ishitani, A., Geraghty, D.E., 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. https://doi.org/10.1007/s00251-005-0076-z.
Ramalho, J., Veiga-Castelli, L.C., Donadi, E.A., Mendes-Junior, C.T., Castelli, E.C., 2017. HLA-E regulatory and coding region variability and haplotypes in a Brazilian population sample. Mol. Immunol. 91, 173–184. https://doi.org/10.1016/j.molimm.2017.09.007.
Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat. Biotechnol. 29 (1), 24–26. https://doi.org/10.1038/nbt.1754.
Robinson, J., Halliwell, J.A., Hayhurst, J.D., Flicek, P., Parham, P., Marsh, S.G.E., 2015. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res. 43, D423–D431. https://doi.org/10.1093/nar/gku1161.
Rouas-Freiss, N., Goncalves, R.M.-B., Menier, C., Dausset, J., Carosella, E.D., 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. https://doi.org/10.1073/pnas.94.21.11520.
Sabbagh, A., Luisi, P., Castelli, E.C., Gineau, L., Courtin, D., Milet, J., Massaro, J.D., Laayouni, H., Moreau, P., Donadi, E.A., Garcia, A., 2014. Worldwide genetic variation at the 3′ untranslated region of the HLA-G gene: balancing selection influencing genetic diversity. Genes Immun. 15, 95–106. https://doi.org/10.1038/gene.2013.67.
Sabbagh, A., Sonon, P., Sadissou, I., Mendes-Junior, C.T., Garcia, A., Donadi, E.A., Courtin, D., 2018. The role of HLA-G in parasitic diseases. HLA 1–16. https://doi.org/10.1111/tan.13196.
Sadissou, I., d'Almeida, T., Cottrell, G., Luty, A., Krawice-Radanne, I., Massougbodji, A., Moreau, P., Moutairou, K., Garcia, A., Favier, B., Rouas-Freiss, N., Courtin, D., 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. https://doi.org/10.1186/1475-2875-13-312.
Sampangi, S., Wang, X., Beagley, K.W., Klein, T., Afrin, S., Healy, H., Wilkinson, R., Kassianos, A.J., 2015. Human proximal tubule epithelial cells modulate autologous B-cell function. Nephrol. Dial. Transplant. 30, 1674–1683. https://doi.org/10.1093/ndt/gfv242.
Santos, K.E., Lima, T.H.A., Felício, L.P., Massaro, J.D., Palomino, G.M., Silva, A.C.A., Oliveira, S.F., Sabbagh, A., Garcia, A., Moreau, P., Donadi, E.A., Mendes-, C.T., Castelli, E.C., 2013. Insights on the HLA-G evolutionary history provided by a nearby alu insertion. Mol. Biol. Evol. 30, 2423–2434. https://doi.org/10.1093/molbev/mst142.
Shiina, T., Suzuki, S., Ozaki, Y., Taira, H., Kikkawa, E., Shigenari, A., Oka, A., Umemura, T., Joshita, S., Takahashi, O., Hayashi, Y., Paumen, M., Katsuyama, Y., Mitsunaga, S., Ota, M., Kulski, J.K., Inoko, H., 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. https://doi.org/10.1111/j.1399-0039.2012.01941.x.

Shobu, T., Sageshima, N., Tokui, H., Omura, M., Saito, K., Nagatsuka, Y., Nakanishi, M., Hayashi, Y., Hatake, K., Ishitani, A., 2006. The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. J. Reprod. Immunol. 72, 18–32. https://doi.org/10.1016/j.jri.2006.02.001.

Sizzano, F., Testi, M., Zito, L., Crocchiolo, R., Troiano, M., Mazzi, B., Turchiano, G., Torchio, M., Pultrone, C., Gregori, S., Chiesa, R., Gaziev, J., Sodani, P., Marktel, S., Amoroso, A., Roncarolo, M.G., Lucarelli, G., Ciceri, F., Andreani, M., Fleischhauer, K., 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. https://doi.org/10.1111/j.1399-0039.2012.01862.x.

Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162–1169. https://doi.org/10.1086/379378.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A New statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989. https://doi.org/10.1086/319501.

Sullivan, L.C., Clements, C.S., Rossjohn, J., Brooks, A.G., 2008. The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. Tissue Antigens 72 (5), 415–424. https://doi.org/10.1111/j.1399-0039.2008.01138.x.

Tamouza, R., Busson, M., Fortier, C., Diagne, I., Diallo, D., Sloma, I., Contouris, H., Krishnamoorthy, R., Labie, D., Girot, R., Charron, D., 2007. HLA-E*0101 allele in homozygous state favors severe bacterial infections in sickle cell anemia. Hum. Immunol. 68, 849–853. https://doi.org/10.1016/j.humimm.2007.08.260.

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

Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192. https://doi.org/10.1093/bib/bbs017.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo, M.A., 2013. From fastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr. Protoc. Bioinform. 43 (SUPL. 43), 11.10.1–11.10.33. https://doi.org/10.1002/0471250953.bi1110s43.

Veiga-Castelli, L.C., Castelli, E.C., Mendes, C.T., Da Silva, W.A., Faucher, M.C., Beauchemin, K., Roger, M., Moreau, P., Donadi, E.A., 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. https://doi.org/10.1111/j.1399-0039.2011.01801.x.

Veiga-Castelli, L.C., da S.B. Bertuol, J.M., Castelli, E.C., Donadi, E.A., 2016. Low variability at the HLA-E promoter region in the Brazilian population. Hum. Immunol. 77, 172–175. https://doi.org/10.1016/j.humimm.2015.11.014.

Veit, T.D., Cazarolli, J., Salzano, F.M., Schiengold, M., Chies, J.A.B., 2012. New evidence for balancing selection at the HLA-G locus in South Amerindians. Genet. Mol. Biol. 35, 919–923. https://doi.org/10.1590/S1415-47572012000600005.

Veit, T.D., De Lima, C.P.S., Cavalheiro, L.C., Callegari-Jacques, S.M., Brenol, C.V., Brenol, J.C.T., Xavier, R.M., Da Cunha Sauma, M.F.L., Dos Santos, E.J.M., Chies, J.A.B., 2014. HLA-G +3142 polymorphism as a susceptibility marker in two rheumatoid arthritis populations in Brazil. Tissue Antigens 83, 260–266. https://doi.org/10.1111/tan.12311.

Wang, W., Wei, Z., Lam, T.W., Wang, J., 2011. Next generation sequencing has lower sequence coverage and poorer SNP-detection capability in the regulatory regions. Sci. Rep. 1 (55), 1–7. https://doi.org/10.1038/srep00055.

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

Xu, Y., Han, H., Zhang, F., Lv, S., Li, Z., Fang, Z., 2015. Lesion human leukocyte antigen-F expression is associated with a poor prognosis in patients with hepatocellular carcinoma. Oncol. Lett. 9, 300–304. https://doi.org/10.3892/ol.2014.2686.

Zambra, F.M.B., Biolchi, V., de Cerqueira, C.C.S., Brum, I.S., Castelli, E.C., Chies, J.A.B., 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. https://doi.org/10.1111/tan.12741.

Zhang, J., Pan, L., Chen, L., Feng, X., Zhou, L., Zheng, S., 2012. Non-classical MHC-Iota genes in chronic hepatitis B and hepatocellular carcinoma. Immunogenetics 64, 251–258. https://doi.org/10.1007/s00251-011-0580-2.