TISSUE ANTIGENS
IMMUNE RESPONSE GENETICS

Tissue Antigens ISSN 0001-2815

Non-classical *HLA-E* gene variability in Brazilians: a nearly invariable locus surrounded by the most variable genes in the human genome

L. C. Veiga-Castelli[1], E. C. Castelli[2], C. T. Mendes Jr[3], W. A. da Silva Jr[4], M.-C. Faucher[5], K. Beauchemin[5], M. Roger[5], P. Moreau[6] & E. A. Donadi[1]

1 Divisão de Imunologia Clínica, Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto (FMRP), Universidade de São Paulo (USP), Ribeirão Preto, SP, Brasil
2 Instituto de Ciências Biológicas, Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil
3 Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil
4 Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto, SP, Brasil
5 Laboratoire d'immunogénétique, Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal, Canada
6 Commissariat à l'Energie Atomique et aux Energies Alternatives, Université Paris 7/DSV/I2BM/Service de Recherches en Hémato-Immunologie, IUH, Hôpital Saint-Louis, Paris, France

Key words
Brazilians; evolution; haplotypes; HLA-E; polymorphism

Correspondence
Luciana C. Veiga-Castelli
Divisão de Imunologia Clínica
Departamento de Clínica Médica
Faculdade de Medicina de Ribeirão Preto (FMRP)
Universidade de São Paulo (USP)
Ribeirão Preto,
São Paulo 14049-900
Brasil
Tel: +55 16 3602 2566
Fax: +55 16 3633 6695
e-mail: luciana.veigacastelli@gmail.com

Received 21 July 2011; revised 4 October 2011; accepted 17 October 2011

doi: 10.1111/j.1399-0039.2011.01801.x

Abstract

The non-classical human leukocyte antigen (HLA) class I genes present a very low rate of variation. So far, only 10 *HLA-E* alleles encoding three proteins have been described, but only two are frequently found in worldwide populations. Because of its historical background, Brazilians are very suitable for population genetic studies. Therefore, 104 bone marrow donors from Brazil were evaluated for *HLA-E* exons 1–4. Seven variation sites were found, including two known single nucleotide polymorphisms (SNPs) at positions +424 and +756 and five new SNPs at positions +170 (intron 1), +1294 (intron 3), +1625, +1645 and +1857 (exon 4). Haplotyping analysis did show eight haplotypes, three of them known as *E*01:01:01*, *E*01:03:01* and *E*01:03:02:01* and five *HLA-E* new alleles that carry the new variation sites. The *HLA-E*01:01:01* allele was the predominant haplotype (62.50%), followed by *E*01:03:02:01* (24.52%). Selective neutrality tests have disclosed an interesting pattern of selective pressures in which balancing selection is probably shaping allele frequency distributions at an SNP at exon 3 (codon 107), sequence diversity at exon 4 and the non-coding regions is facing significant purifying pressure. Even in an admixed population such as the Brazilian one, the *HLA-E* locus is very conserved, presenting few polymorphic SNPs in the coding region.

Introduction

The human leukocyte antigen (HLA) complex encompasses a cluster of genes encoding glycoproteins that mediate cell-to-cell interactions and regulation of the immune response. The HLA class I genes are divided into classical (*HLA-A*, *HLA-B* and *HLA-C*) and non-classical loci (*HLA-E*, *HLA-F* and *HLA-G*) (1). The non-classical class I genes can be distinguished from the classical ones by their function, tissue distribution and degree of gene variation, because the classical class I genes are considered to be the most polymorphic in the human genome, whereas non-classical class I genes present a very low rate of variation. Like classical class I molecules, non-classical HLA-E, -F and -G molecules may complex with β2-microglobulin and may act as peptide presenters (2, 3), although their main role is not antigen presentation. HLA-E plays a role both as a modulator of natural killer (NK) cell activity by interacting with the CD94-NKG2A receptor in the innate immunity pathway (4) and as an antigen presenting molecule triggering a specific immune response (5). The CD94 glycoprotein is expressed on the membrane of most NK cells and on a subset of T lymphocytes, forming a heterodimer with the NKG2A/B, NKG2C, NKG2E and NKG2H glycoproteins. Interactions with such CD94 heterodimers can augment, inhibit, or have no effect on NK cell-mediated cytotoxicity and cytokine production (6). Apparently, the expression of HLA-E molecules may be increased by the upregulation

of the *HLA-G* gene as HLA-E preferentially binds to the monomeric leader peptide derived from HLA-G, and this complex interacts with the inhibitory NK receptor CD94/NKG2A. Therefore, the expression of HLA-G may influence the expression of HLA-E (4, 7, 8).

HLA-E is the least polymorphic of class I HLA genes, exhibiting only 10 alleles encoding three distinct proteins (The International Immunogenetics Database – IMGT, version 3.4.0 – http://www.ebi.ac.uk/imgt/hla/stats.html). A study on *HLA-E* typing (exons 2 and 3) performed on 371 individuals from six different ethnic groups detected only three distinct *HLA-E* alleles, encoding two distinct molecules (9). Overall, three distinct allele groups (*HLA-E*01:01*, *HLA-E*01:03* and *HLA-E*01:04*) (IMGT, database version 3.4.0) have been described in diverse human populations; however, only two allele groups (*HLA-E*01:01* and *HLA-E*01:03*), which differ by a single amino acid (arginine in the former and glycine in the latter at position 107 of the α2 heavy chain domain), are found in worldwide populations. Both allele groups are found at nearly equal frequencies, suggesting that balancing selection is acting on this gene, with putative functional differences between these two variants (10). Until now, the existence of the third allele group (*HLA-E*01:04*) has been the subject of much controversy with some authors suggesting that such allele group is in fact a consequence of sequencing artifacts (11), because it has not been found in any population so far studied, including Brazilians of the present series.

The genetic background of the Brazilian population is highly diverse as a consequence of five centuries of interethnic crosses of people from three continents, including European colonizers, primarily Portuguese, African slaves and autochthonous Amerindians (12). To date, the polymorphism of the *HLA-E* locus has not been characterized in the Brazilian population, in which more sequence variation would be predictable because of such interethnic admixture. Because of its historical background, the Brazilian population has been considered to be suitable for population genetic studies. On this basis, we evaluated the *HLA-E* sequence that encodes the external portion of the HLA-E molecule, searching for variation and assessing the haplotype structure of this region.

Methods and samples

A total of 104 healthy bone marrow donors from the state of São Paulo, Southeastern Brazil were recruited, and genomic DNA was obtained from peripheral blood using a salting-out procedure (13). *HLA-E* alleles were defined by the nucleotide variation encompassing exons 1–4, including introns. Briefly, exons 1–4 were amplified by the polymerase chain reaction (PCR), using HE01F (5′-TCCTGGATACTCATGACGCAGACTC-3′) (11) and HE43 (5′-CCCATCCCCCTCCTTACT-3′) primers (14), yielding a 2041-bp product. PCR was performed in a final volume of 50 µl containing 0.7x PCR buffer [70 mM Tris–HCl pH 8.8, 20 mM $(NH_4)_2SO_4$, 3 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 1% Triton 100, 50 mg bovine serum albumin (BSA)], 0.2 mM of each dNTP, 15 pmol of each primer, 1.5 U of Taq DNA-polymerase Platinum (Invitrogen, Carlsbad, CA) and 200 ng of genomic DNA. The initial denaturation cycle was carried out at 94°C for 3 min, followed by 35 cycles at 94°C for 30 s, 62°C for 30 s and 72°C for 135 s, and by a final extension step at 72°C for 5 min. *HLA-E* variation was evaluated by direct sequencing of PCR products, which were evaluated using 1% agarose gel and directly sequenced using the following primers: HE01R (5′-CCTCTTACCCAGGTGAAGCAGCG-3′), HE532ALL (5′-CCTACGACGGCAAGGATTATC-3′), HE43 (5′-CCCATCC CCCTCCTTACT-3′) and HE45 (5′-TTTCTGACTCTTCCC CTC-3′), described by Grimsley and colleagues (11, 14), and HE02F (5′-TCTAGAGAAGCCAATCAGCG-3′), HEE2R (5′-CCTCGCTCTGATTGTAGTAG-3′) and HE03F (5′-CAC GACTCCCGACTATAAAG-3′), designed for this study. All sequences obtained from each sample were aligned with the genomic sequences of the official alleles [recognized by the World Health Organization (WHO) and the international ImMunoGeneTics (IMGT)], and each single nucleotide polymorphism (SNP) detected was individually annotated. The SNPs were evaluated using an inlab software, denominated SNPex (contact the corresponding author for availability), designed to define the phase of SNPs of each sample taking into account the sequences of the officially described alleles and, therefore, to determine the *HLA-E* alleles present in each sample. Besides the *HLA-E* coding region, these samples have also been previously typed for *HLA-A*, *-B*, *-C*, *-DR* and *-DQ* alleles using commercially available kits (One Lambda, Canoga Park, CA) and the coding region of *HLA-G*, as previously published (15–17).

The presence of a significant association between all *HLA-E* variation sites was evaluated by means of a likelihood ratio test of linkage disequilibrium (LD) (18), using the ARLEQUIN version 3.5.1.2 (19) and Haploview 4.2 (20) software. Given the positive association but unknown gametic phase, the PHASE method (21), implemented by the PHASE v2 package (Mac OS version) (21), as well as the EM (Expectation–Maximization) (18) and ELB (Excoffier–Laval–Balding) algorithms (22), implemented by the ARLEQUIN version 3.5.1.2 (19), were used to assign the most probable haplotype constitution of each sample. Set points used for haplotyping identifications included: i) EM algorithm: epsilon value: 1e–7, number of starting points: 100, maximum number of interactions: 5000, estimating standard deviation (SD) through bootstrap with the number of bootstraps to perform at 5000 and the number of starting points for SD estimations at 50; ii) ELB algorithm: alpha value: 0.01, epsilon value: 0.01, gamma value: 0, sampling interval: 500, burn-in steps: 100,000, recombination steps: 0;

iii) PHASE method, number of interactions: 1000, thinning interval: 1, burn-in value: 100, delta value for each locus: 1. For the PHASE method 10 independent runs performed yielded the same results for all samples. In fact, all methods used for haplotyping did show the same results for each sample. In addition, for samples in which new variation spots were detected, the PCR products were cloned using TOPO® (Invitrogen) and sequenced to define the phase of haplotypes harboring rare *HLA-E* alleles.

Allelic frequencies and observed heterozygosity (h_O) were computed by the direct counting method. Adherences of genotypic proportions to expectations under Hardy–Weinberg equilibrium were tested by the exact test of Guo and Thompson (23), using the GENEPOP 3.4 software (24). The expected heterozygosity values (h_{Sk}) and nucleotide (π) and haplotype diversity as well as their SDs were estimated by the ARLEQUIN 3.5.1.2 program (19).

Departure from selective neutrality was evaluated by three different methods. The Ewens–Watterson test (25, 26), implemented by Slatkin (27), where a Markov-Chain Monte Carlo resampling algorithm allows to obtain a null distribution and compares the observed homozygosity under Hardy–Weinberg proportions with the expected homozygosity under the hypothesis of neutrality for the same sample size and number of alleles. This procedure permits to test alternative hypotheses of either directional (observed homozygosity greater than expected homozygosity) or balancing selection (observed homozygosity lower than expected homozygosity). The test was carried out using the PyPop software (28). The Tajima's D test (29), which examines the relationship between the number of segregating sites and nucleotide diversity by comparing the sequence diversity statistics θ_W and π. Under the standard neutral model, the expectations of θ_W and of π are equal, and therefore, the expected value of Tajima's D is 0 under neutrality (30). A positive Tajima's D value reflects a relative excess of intermediate-frequency alleles and is evidence for heterozygous advantage, while a negative value reflects an excess of rare variants and is evidence for selection of one specific allele over alternate alleles (31). The significance of the D statistic is tested by generating 99,999 random samples under the hypothesis of selective neutrality and population equilibrium, using a coalescent simulation algorithm. The test was carried out using the ARLEQUIN version 3.1 program (19). The third approach included the synonymous and non-synonymous nucleotide substitution test, which evaluates the relative abundance of synonymous substitutions (no amino acid change) and non-synonymous substitutions (amino acid change) that have occurred in the gene sequence. For data sets containing more than two sequences, the test was carried out by first estimating the average number of synonymous substitutions per synonymous site (d_S) and the number of non-synonymous substitutions per non-synonymous site (d_N), and their variances. Then, the null hypothesis of neutrality ($d_N = d_S$) can be evaluated. The test was carried out using the Nei-Gojobori method implemented in the MEGA 5 program (32).

Results

By evaluating the sequence of more than 1100 nucleotides of the *HLA-E* gene (exons 1–4, including introns) in 208 chromosomes, seven variation sites were found spread across the gene: one at intron 1 (position +170), one at exon 2 (position +424), one at exon 3 (position +756), one at intron 3 (position +1294) and three at exon 4 (positions +1625, +1645 and +1857) (Table 1). The position number considers the first base of the ATG codon at exon 1 as nucleotide +1. Three other previously described variation sites at the *HLA-E* gene were also evaluated by the current method, but they were not found in this Brazilian series (positions +887, +906 and +1691). The positions +887, +906 and +1691 are associated with the *E*01:03:03*, *E*01:04* and *E*01:03:04* alleles, respectively. Curiously, exception made for the *E*01:03:04*, these alleles have not been found in any other population evaluated so far, i.e. the allele *E*01:04* was found only in Japanese in the early 1990s (33), and the *E*01:03:03* was submitted to IMGT but not published, and some authors have suggested that they are in fact sequencing artifacts (34). Recently, another *HLA-E* allele was described with an intron 3 mutation (*E*01:03:02:02*) (IMGT, database version 3.4.0), which was not observed in the present series.

Of the seven variation sites observed in the present series, two had already been described for the *HLA-E* gene (positions +424 at codon 77 and +756 at codon 107) and five new mutations have not been defined in the IMGT database and have not been associated with a specific *HLA-E* allele or haplotype (Table 1). Interestingly, two out of the five new mutations have already been described as human SNPs (positions +1625 and +1857), one of them being recently described by the 1000 Genomes Project (www.1000genomes.org) (Table 1). Notwithstanding that, they were not associated with a specific *HLA-E* allele by IMGT.

Considering the variation sites observed at the *HLA-E* gene in Brazilians, only three reached polymorphic frequencies, i.e. the known variations at positions +424 and +756 and the new variation site at position +1645 (Table 1). These three SNPs correspond to a synonymous substitution (position +424) and missense mutations at exons 3 (position +756) and 4 (position +1645), leading to an amino acid substitution in the HLA-E molecule (Table 1). The non-synonymous mutation at the +756 position leads to an arginine (basic and hydrophilic amino acid) to glycine (polar and neutral amino acid) change, which theoretically does not drastically affect the final protein. However, the mutation at the +1645 position leads to a very different amino acid substitution: an aspartic acid (hydrophilic) to valine (apolar and hydrophobic). Another missense mutation at position +1857, which almost reached polymorphic frequencies, also leads to a substitution involving

Table 1 List of variation sites for the human leukocyte antigen (*HLA*) −*E* coding region (previously known and new ones), their effect on the HLA-E molecule and their frequency in Brazil

HLA-E region	SNP accession number	SNP effect	Associated with specific HLA-E alleles	Nucleotide position	Codon	Position on codon	Variation	Region where it was described	Frequency in Brazil ($2n = 208$)
Intron 1	None	None	No	170	—	—	G/T	Brazil	G = 0.9952
Exon 2	rs1059510	Synonymous AA**C** (Asn) AA**T** (Asn)	Yes	424	77	Third base	C/T	Spain	C = 0.7330
Exon 3	rs1264457	Missense **A**GG (Arg) **G**GG (Gly)	Yes	756	107	First base	A/G	Japan	A = 0.6298
Exon 3	rs41560815	Synonymous GC**C** (Ala) TC**T** (Ala)	Yes	887	150	Third base	C/T	Unknown	C = 1.0000
Exon 3	rs41562314	Missense **A**GA (Arg) **G**GA (Gly)	Yes	906	157	First base	A/G	East Asia	A = 1.0000
Intron 3	None	None	No	1294	—	—	G/A	Brazil	G = 0.9952
Exon 4	rs11548296	Synonymous GT**G** (Val) GT**C** (Val)	No	1625	189	Third base	G/C	West Africa (Nigeria), Brazil	G = 0.9952
Exon 4	None	Missense G**A**C (Asp) G**T**C (Val)	No	1645	196	Second base	A/T	Brazil	A = 0.9856
Exon 4	rs17875370	Synonymous GC**G** (Ala) GC**A** (Ala)	Yes	1691	211	Third base	G/A	Caucasoid (region unknown)	G = 1.0000
Exon 4	rs62621992 rs115331960	Missense **C**CC (Pro) **T**CC (Ser)	No	1857	267	First base	C/T	North America Brazil	C = 0.9904

amino acid residues with completely different properties: proline (apolar and hydrophobic) to serine (polar and neutral).

The presence of a significant association between all *HLA-E* variation sites was evaluated by using Haploview 4.2 (20). Figure 1 illustrates the pattern of LD along the *HLA-E* locus. Except for the SNP pair +424/+1857, all SNPs reached the maximum D' value. However, it can be noticed that only an SNP pair (positions +424 and +756) did present strong LD as evaluated by the log of the odds (LOD) score; such SNPs are the two most frequent variants in the present series. As all other SNPs presented the minor allele with low frequencies, i.e. lower than 1.5%, there was no power to detect such LD. In addition, the likelihood ratio test of LD (18), using the ARLEQUIN version 3.1 program (19), did also show LD only between the frequently polymorphisms (positions +424 and +756), corroborating the data obtained by Haploview. Nevertheless, considering such elevated LD among the frequent SNPs, computational haplotype inferences were performed as described earlier.

On the basis of nucleotide sequence variations and the phase of alleles from those seven variation sites obtained by probabilistic models, eight different haplotypes (i.e. *HLA-E* alleles) were obtained. Of these, three haplotypes have already been described and matched with the IMGT

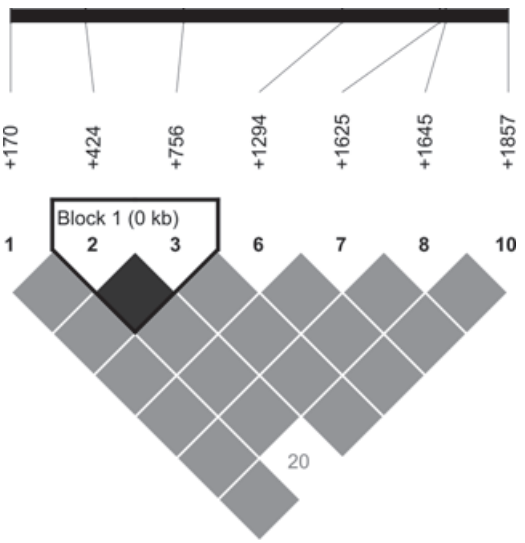

Figure 1 Linkage disequilibrium (LD) between pairs of SNPs at the human leukocyte antigen (*HLA)-E* locus. The image was generated using the Haploview software. Areas in black indicate strong LD (LOD ≥ 2, D′ = 1), gray indicates weak LD (LOD < 2, D′ = 1) and white indicates no LD (LOD < 2, D′ < 1). D′ values different from 1.00 are represented inside the squares as percentages. LOD, log of the odds; D′, pairwise LD parameter between SNPs.

Table 2 Human leukocyte antigen (*HLA*)-*E* allele frequency in Brazilians and other populations studied so far[a]

HLA-E alleles	African-Americans	African Shona	Hutterites	Hispanics	Australians	Asian Indians	Chinese	Thai[b]	Koreans[b]	Japanese	Brazilians
*E*01:01:01*[c]	56.58	50.93	41.70	56.20	57.00	0.6475	24.00	42.50	49.10	32.00	62.50
*E*01:03:01*[d]	18.42	29.63	56.50	43.80	43.00	0.1250	76.00	57.50	28.80	39.00	9.13
*E*01:03:02:01*	25.00	19.44				0.2025			22.10	29.00	24.52
*E*01:03:03*	—	—				—			—	—	0.00
*E*01:03:04*	—	—				—			—	—	0.00
*E*01:04*	0.00	0.00			0.00	0.00		0.00	0.00	0.00	0.00
New alleles[e]	—	—	—	—	—	—	—	—	—	—	3.84
$2n$[f]	76	216	1704	96	400	400	50	400	972	100	208
h_{Sk}[g]	0.5912	0.6179	—	—	—	0.5185	—	—	0.6279	0.6681	0.5432
References	(11)	(54)	(14)	(14)	(55)	(56)	(14)	(57)	(58)	(11)	This study

[a]The ' – ' denotes an allele that was not evaluated.
[b]Group of healthy controls.
[c]Alleles *E*01:01:01:01* to *E*01:01:01:03* were pooled together as *E*01:01:01*.
[d]Alleles *E*01:03:01:01* and *E*01:03:01:02* were pooled together as *E*01:03:01*.
[e]New *HLA-E* variants observed in the present study and specified in Table 3.
[f]Number of chromosomes sampled.
[g]Expected heterozygosity values.

criteria for the *E*01:01:01, E*01:03:01* and *E*01:03:02:01* alleles (Table 2). The *E*01:01:01:01*, *E*01:01:01:02* and *E*01:01:01:03* alleles were pooled together in this article as *E*01:01:01* because they differ by mutations outside the evaluated region. The *E*01:03:01:01* and *E*01:03:01:02* alleles were pooled together as *E*01:03:01* for the same reason. It should be mentioned that all frequency data presented in Table 2 did also use the same grouping policy described above.

The five remaining haplotypes were in fact new alleles, each one carrying one of the new variation sites described (Tables 1 and 3). As these new *HLA-E* alleles have not yet been recognized by IMGT, provisory names were given according to the original known allele in which the mutation seems to have occurred (Table 3). The *HLA-E* genotypes did fit the Hardy–Weinberg equilibrium expectations for all SNPs evaluated and for the *HLA-E* haplotypes ($P = 0.3548 \pm 0.0227$).

Table 3 New human leukocyte antigen (HLA)-E alleles found in the Brazilian population

HLA-E allele	Mutation	Frequency
*E*01:01:01*β	+170	0.48
*E*01:03:01*β	+1625	0.48
*E*01:03:01*$^{\mathbf{pro267ser}}$	+1857	0.96
*E*01:03:02:01*β	+1294	0.48
*E*01:03:02:01*$^{\mathbf{asp196val}}$	+1645	1.44

The Greek letter β indicates a new allele that does not result in new protein, thus sharing a preexistent allele group. This nomenclature option was used because of the fact that such alleles have not yet been recognized by the international ImMunoGeneTics database (IMGT). In cases of amino acid change, the former and the new amino acid is given with the codon number, which also indicates the existence of a new allele group.

The *HLA-E*01:01:01* allele was the predominant one in the Brazilian population, with a frequency of 62.50%, followed by the *E*01:03:02:01* (24.52%) and *E*01:03:01* (9.13%) alleles. Although the specific SNPs associated with the *E*01:03:03*, *E*01:03:04* and *E*01:04* alleles were evaluated, all the 104 samples analyzed were monomorphic for these variation sites.

Because of the paucity of frequency data of the *HLA-E* gene at a high resolution level in worldwide populations, it was not possible to perform an exhaustive population genetic approach involving the populations or ethnic groups that contributed to the Brazilian gene pool. Some preliminary clues, however, can already be obtained based on currently available data (Table 2). The Brazilian population presented the largest *HLA-E* allele diversity. However, together with Asian Indians, it presented the lowest expected heterozygosity (Table 2) among the populations with a more refined *HLA-E* evaluation (African-Americans, African Shona, Asian Indians, Koreans and Japanese).

The frequency of the *HLA-E* alleles, and as a consequence of each of the known SNPs at the *HLA-E* locus, varies substantially across the world (Table 2). An exact test of population differentiation showed that the Brazilian allele frequency distribution is significantly different from those of African Shona, Korean and Japanese populations (in all cases, $P = 0.0000 \pm 0.0000$). However, no significant differences in comparisons with African-Americans ($P = 0.2438 \pm 0.0233$) and Asian Indians ($P = 0.1149 \pm 0.0216$) were observed.

The presence of a significant association between the *HLA* loci was evaluated by means of a likelihood ratio test of LD (18), using the ARLEQUIN version 3.1 program (19). The likelihood ratio test of LD involving the *HLA-A*, *HLA-B*,

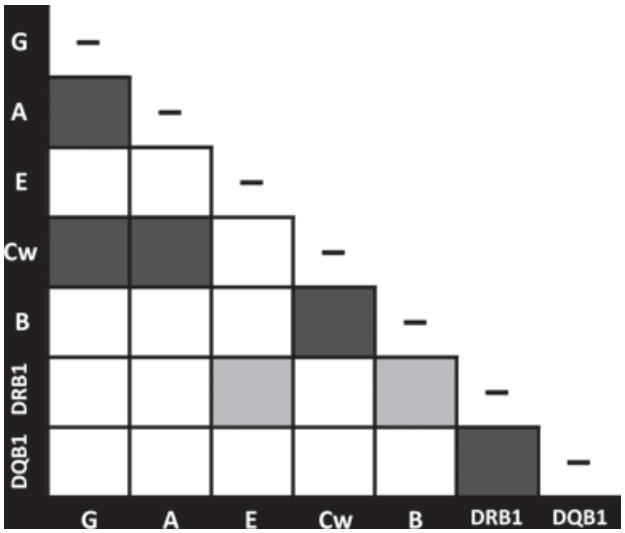

Figure 2 Linkage disequilibrium (LD) among human leukocyte antigen (*HLA)* loci. Dark gray squares indicate statistical LD with $P < 0.05$, light gray squares represent LD with P between 0.5 and 0.1 and white squares represent absence of LD. This image was generated manually.

HLA-C, *HLA-DRB1* and *HLA-DQB1* loci (low-resolution allele groups), *HLA-G* and *HLA-E* (haplotypes) evidenced the absence of a significant association between *HLA-E* and other *HLA* loci (Figure 2). Because of their proximity, the loci *HLA-G, HLA-A* and *HLA-C*, *HLA-B* and *HLA-C*, *HLA-DRB1* and *HLA-DQB1* did present significant LD (Figure 1). In spite of the LD among *HLA-G, HLA-A* and *HLA-C*, it was not possible to detect such interaction with *HLA-E*, which lies between *HLA-A* and *HLA-C*.

The nucleotide diversity at the *HLA-E* locus was 0.00050, about 66% of the human genome average (0.00075) (35, 36) (Table 4). Considering the nucleotide diversity for each exon separately, only exons 2 and 3 presented a higher diversity, about two times the human average. In addition, exon 1 was monomorphic and exon 4 presented very low nucleotide diversity. Interestingly, the non-coding regions (introns) were practically monomorphic (nucleotide diversity of 0.00002) in a sequence even bigger than the coding sequence evaluated in the present sample. The coding sequences presented 50 times the nucleotide diversity of the non-coding sequences.

By using the polymorphic data obtained in this study, the Ewans–Watterson neutrality test, which takes into account heterozygosis, was performed to investigate whether natural selection may exert significant pressure on the allele and genotype frequencies at the *HLA-E* locus (Table 4), but statistical significance was not reached. However, negative normalized F values exhibiting low P-values were observed for exons 2 and 3 ($P = 0.1706$ and $P = 0.0091$, respectively), and positive normalized values were observed for exon 4 and for non-coding sequences. The second neutrality test used, Tajima's D test, which takes into account sequence diversity, resulted in the observation that exons 2 and 3 presented positive D values, but with no statistical significance. On the other hand, exon 4 and the non-coding sequences did present negative D values ($P = 0.0220$ and $P = 0.02400$, respectively). Taken together, these negative Ewens–Waterson's normalized F values and positive Tajima's D values observed in exons 2 and 3 reflect an excess of heterozygosis that may be due to balancing selection (35). On the other hand, the positive Ewens–Waterson's normalized F values and negative Tajima's D values observed in exon 4 and non-coding sequences may indicate a directional selection toward a specific haplotype or negative selection, especially considering the low nucleotide and haplotype diversity presented by exon 4 and non-coding sequences and their

Table 4 Neutrality tests performed in the human leukocyte antigen (HLA)-E locus

Region	Sequence length	Nucleotide diversity (π)	No. of segregating sites[a]	Number of haplotypes (k)	Haplotype diversity	Tajima's D (*P*)	Ewens-Watterson normalized *F*
HLA-E genomic[a]	1881	0.00050	7	8	0.543	−0,4428 $P = 0.3870$	0.5665 $P = 0.7608$
HLA-E exon 1	64	0.00000	—	0	0.000	—	—
HLA-E exon 2	270	0.00145	1	2	0.391	1.3178 $P = 0.9170$	−1.3488 $P = 0.1706$
HLA-E exon 3	276	0.00170	1	2	0.469	1.7793 $P = 0.9490$	−1.7835 $P = 0.0909$
HLA-E exon 4	276	0.00021	3	4	0.057	**−1.4132** ***P* = 0.0220**	1.7813 $P = 0.9693$
HLA-E coding[b]	886	0.00104	5	6	0.535	0.1603 $P = 0.6440$	−0.0371 $P = 0.5727$
HLA-E non-coding[c]	995	0.00002	2	3	0.019	**−1.2800** ***P* = 0.0240**	1.4482 $P = 1.0000$

Significant results ($P < 0.05$) are given in bold.

[a]Exons 1–4, including introns.

[b]Exons 1–4 (introns not included).

[c]Introns 1–3 (exons not included).

Table 5 Synonymous and non-synonymous nucleotide substitution test of neutrality, positive and purifying selection for analysis averaging human leukocyte antigen (*HLA*)-*E* sequence pairs (the official ones plus the new ones described here)[a]

d_N/d_S neutrality test	Number of sequences	Number of codons	H_A = neutrality ($d_N \neq d_S$)	H_A = positive selection ($d_N > d_S$)	H_A = purifying selection ($d_N < d_S$)
Exon 1	14	21	$d_N - d_S = 0.0000$; $P = 1.0000$	$d_N - d_S = 0.0000$; $P = 1.0000$	$d_N - d_S = 0.0000$; $P = 1.0000$
Exon 2	14	89	$d_N - d_S = -0.9845$; $P = 0.3269$	$d_N - d_S = -0.9899$; $P = 1.0000$	$d_S - d_N = 0.9992$; $P = 0.1599$
Exon 3	15	91	$d_N - d_S = 0.1889$; $P = 0.8505$	$d_N - d_S = 0.1952$; $P = 0.4228$	$d_S - d_N = -0.1945$; $P = 1.0000$
Exon 4	14	91	$d_N - d_S = -0.9367$; $P = 0.3508$	$d_N - d_S = -0.9328$; $P = 1.0000$	$d_S - d_N = 0.9336$; $P = 0.1762$
Exons 2 and 3	14	181	$d_N - d_S = -0.9150$; $P = 0.3620$	$d_N - d_S = -0.9335$; $P = 1.0000$	$d_N - d_S = 0.9828$; $P = 0.1638$
Exons 3 and 4	14	183	$d_N - d_S = -0.7263$; $P = 0.4691$	$d_N - d_S = -0.7224$; $P = 1.0000$	$d_N - d_S = 0.7164$; $P = 0.2376$
Exons 2–4	14	273	$d_N - d_S = -1.2118$; $P = 0.2280$	$d_N - d_S = -1.2244$; $P = 1.0000$	$d_N - d_S = 1.2885$; $P = 0.1000$
Exons 1–4	14	295	$d_N - d_S = -1.2159$; $P = 0.2264$	$d_N - d_S = -1.2145$; $P = 1.0000$	$d_S - d_N = 1.2083$; $P = 0.1147$

[a]Probabilities of rejecting the null hypothesis of strict-neutrality ($d_N = d_S$) in favor of the alternative hypotheses are shown.

number of segregation sites. The outcome of the third neutrality test, the synonymous and non-synonymous nucleotide substitution test, suggests a non-significant but at least different behaviors of exon 3 (possibility of positive selection) in comparison with the remaining evaluated regions (possibility of purifying selection), which is compatible with the increased Ewens–Watterson's normalized *F* values obtained for the Brazilian population. However, such test is completely different from the previous ones because it does not evaluate the dynamics of the *HLA-E* gene in the dataset obtained (i.e. Brazilians), but it evaluates natural selection over the worldwide identified *HLA-E* alleles, including the new ones described here (Table 5).

Discussion

The classical HLA proteins, especially HLA-A, -B and -C, are fundamental for T-cell mediated responses, as they are highly involved in antigen presentation processes. The genes encoding such proteins are known to be the most polymorphic in the human genome as well as in most vertebrates. Several lines of evidence support the idea of microorganism-mediated balancing selection acting in such genes enhancing their variability, because such variations would be essential to an efficient antigen presentation of peptides derived from thousands of microorganisms (37–41). Two mechanisms may be involved in the maintenance of variability in such genes, the heterozygote advantage in which heterozygous individuals would be more adapted, and the sexual selection, in which a preferred mating would occur between individuals with different histocompatibility molecules (39–43). In addition, a pathogenic co-evolution theory has recently emerged, whereby the most common alleles are placed under the greatest pathogenic pressure, leading to a positive selection of the least common alleles. Irrespective of the most plausible theory to explain variability, classical HLA genes are in fact highly variable, with approximately 4828 different alleles already detected, considering the *HLA-A, -B* and *-C* loci together (IMGT database version 3.4.0).

In contrast to classical *HLA* genes, limited *HLA-E* variability has been observed in worldwide populations. Despite the recognized role of the non-classical HLA molecules as modulators of immune responses, the exact mechanisms by which the HLA-E molecule does influence T and NK cell responses is not quite understood. Moreover, it is interesting to remark that all non-classical *HLA* loci that modulate immune responses are quite invariable at the coding region compared with their classical counterparts. Although few studies have evaluated the variability of the *HLA-E* locus, a consistent data is the low *HLA-E* variability observed in worldwide populations (Table 2).

Besides the low-resolution typing methods used in the few available studies regarding *HLA-E* variability (Table 2), only a few different alleles were found coding for only two different proteins (*E*01:01* and *E*01:03*). As the methodology used in the present work allowed high-resolution typing, any *HLA-E* allele with nucleotide differences between exons 1 and 4 would be detected. In addition to the two worldwide HLA-E molecules currently reported, only two others were detected, both encoded by *HLA-E* alleles presenting nucleotides changes in exon 4 (mutation at positions +1645 and +1857, Table 1), but none achieving polymorphic frequencies (Table 1). Another recent study evaluating *HLA-E* variability at a high-resolution level was reported in a Canadian population, confirming the presence of alleles that encode the same two worldwide most frequent HLA-E molecules (44). In the previous study, the authors did evaluate only exon 3, so some degree of variation could be missing. Even so, more than 600 patients were evaluated, and no new variation was found in exon 3, with both *E*01:01* and *E*01:03* alleles showing similar frequencies.

Interestingly, even in a population such as the Brazilian one, in which a high degree of variability would be expected because of its genetic background, only two HLA-E molecules were also frequently found. Regarding another non-classical *HLA* gene, *HLA-G*, Brazilians did present the highest *HLA-G* variability so far described (45), supporting the idea of the high variability expected in such population. Those two

HLA-E molecules found worldwide are different for a substitution of glycine to arginine at codon 107, which may alter the structure of the molecule. In addition, relative differences in peptide affinity, complex stability and ligand interaction may account for putative distinct biological functions. The only structural difference between these two molecules is the presence of an additional hydrogen bond involving the side chain of arginine 107 (1). Although different surface levels of HLA-E may modulate ligand interaction, the substitution at residue 107 does not seem to affect the interaction of HLA-E with CD94-NKG2 (46). However, the allele bearing codon 107G (glycine) is associated with higher HLA-E expression levels, higher affinity for available peptides and higher stability of the HLA-E^{107Gly}/peptide complex (47–49). Therefore, the HLA-E molecule coded by the *E*01:03* allele has been associated with greater stability, higher expression and a potentially stronger inhibitory effect on NK cells (50, 51).

Curiously, the *HLA-E* locus is located between the *HLA-C* and *HLA-A* loci, two of the most polymorphic genes in the human genome, encompassing a region of approximately 1200 kb at chromosome 6p21.3. This region does present a high degree of variability at the *HLA-C* and *HLA-A* loci, probably because of balancing selection, and in between these loci a strong pressure against variation for the *HLA-E* locus, probably because of purifying selection.

As HLA-E may modulate immune system cells, a certain degree of invariability is in fact expected for the coding region of the gene; however, it is interesting to observe that different selective pressures are acting along this region. Exons 2 and 3, for instance, do present a trend for balancing selection given the observed negative Ewens–Watterson's normalized *F* values and positive Tajima's *D* values. This balancing selection is probably maintaining the polymorphism at codon 107 in exon 3, which is associated with the two common HLA-E proteins found worldwide. Evidence for balancing selection at exon 2 is probably a hitchhiking effect, because the polymorphic site at codon 77 in exon 2 is a silent mutation, which is in LD with codon 107. As codon 107 polymorphism is found worldwide, it may be considered a very old mutation. In addition, the frequency of both alleles in this polymorphism is somehow elevated (Table 2), which is compatible with balancing selection, as evidenced in this study and in a previous report (10).

To maintain both alleles worldwide, which would explain such high heterozygosis, one of the following circumstances may occur. First, this amino acid change does have a functional impact, and the presence of both encoded proteins is profitable. This is particularly difficult to infer at the present time, as few studies have evaluated such functional relevance. Apparently, the *E*01:03* molecule (glycine at codon 107) is more stable and would result in a stronger inhibitory response (47–51), but with essentially the same affinity to CD94 as compared with its counterpart (46). It is not clear why two different HLA-E proteins would be profitable in humans, unless the presence of both molecules, one highly inhibitory and more stable and the other less effective, would prepare individuals for situations in which a high or low HLA-E availability would be required. The second hypothesis is that alleles from this SNP are in high LD with other SNPs at the regulatory regions, which in fact are suffering balancing selection to maintain high-expressing and low-expressing haplotypes in the regulatory regions. This phenomenon apparently has been observed for the *HLA-G* gene, in which balancing selection maintains a high degree of variability in the regulatory regions, probably maintaining heterozygosis of high-expressing and low-expressing haplotypes (17, 35).

However, it should be mentioned that both mechanisms could be operating in different degrees, i.e. balancing selection maintaining both two functional different molecules and regulatory haplotypes at the same time. Nevertheless, it is important to point out that only exon 3 did present a slight suggestion of positive selection by the d_N/d_S test, which is compatible with both selection mechanisms (Table 5).

Interestingly, exons 1 and 4, and introns 1–3 did present a high pressure against variance. Exon 1 encodes the leader peptide and exon 4 encodes the α3 external domain of the HLA-E molecule. The lack of variation at exon 4 is somehow expected, but the lack of variation in *HLA-E* introns is quite intriguing. The nucleotide diversity of *HLA-E* introns was approximately zero (0.00002), with only two segregating sites (with just one occurrence in 208 chromosomes) in 995 nucleotides. This lack of variation is quite unexpected, because its non-classical counterpart, *HLA-G*, presents several non-coding variations (mainly in intron 1) with elevated frequencies (17), and the classical HLA genes do present hundreds of non-coding polymorphic sites already described (IMGT, version 3.4.0). This apparent negative selection acting on the non-coding regions of the *HLA-E* locus may be a consequence of a functional relevance of such introns (52); however, major known functions for introns such as synthesis of microRNAs and alternative splicing are not observed in these regions of the gene (53). In conclusion, the *HLA-E* locus is quite invariant even in an admixed population such as the Brazilian one because only two frequent encoded proteins were found due to a missense mutation at exon 3 codon 107. In addition to the known alleles, five new *HLA-E* haplotypes were found in Brazilians, two of them associated with a new HLA-E protein, but with low frequencies. The variance at exon 3 apparently has been maintained in worldwide populations because of balancing selection, which was confirmed by our data, but a strong pressure for invariance is also acting at exon 4 and introns, probably reflecting its functional relevance, deserving further studies.

Acknowledgments

This study was supported by the Brazilian National Research Council (CNPq/Brazil – Grants 475670/2007-8 and 558476/

2008-0) and the binational collaborative research program CAPES-COFECUB (project # 653/09). L. C. V. C. is supported by a postdoctoral fellowship (150175/2009-4) from CNPq/Brazil. M. R. is supported by a career award from Fonds de la Recherche en Santé du Québec (FRSQ). We thank Flavia Tremeschin de Almeida and Sandra Rodrigues for their invaluable help.

Conflict of interest

The authors declare that the findings have not been published, have not been submitted, or are not being submitted elsewhere for publication. All contributors to this manuscript have read and approved the submission to this journal and there is no financial conflict of interest.

References

1. Strong RK, Holmes MA, Li P, Braun L, Lee N, Geraghty DE. HLA-E allelic variants. Correlating differential expression, peptide affinities, crystal structures, and thermal stabilities. *J Biol Chem* 2003: **278**: 5082–90.
2. Carosella ED, Paul P, Moreau P, Rouas-Freiss N. HLA-G and HLA-E: fundamental and pathophysiological aspects. *Immunol Today* 2000: **21**: 532–4.
3. Diehl M, Munz C, Keilholz W et al. Nonclassical HLA-G molecules are classical peptide presenters. *Curr Biol* 1996: **6**: 305–14.
4. Braud VM, Allan DS, O'Callaghan CA et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* 1998: **391**: 795–9.
5. Tomasec P, Braud VM, Rickards C et al. Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. *Science* 2000: **287**: 1031.
6. Perez-Villar JJ, Melero I, Rodriguez A et al. Functional ambivalence of the Kp43 (CD94) NK cell-associated surface antigen. *J Immunol* 1995: **154**: 5779–88.
7. Braud V, Jones EY, McMichael A. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9. *Eur J Immunol* 1997: **27**: 1164–9.
8. Lee N, Llano M, Carretero M, et al. HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. *Proc Natl Acad Sci U S A* 1998: **95**: 5199–204.
9. Arnaiz-Villena A, Vargas-Alarcon G, Serrano-Vela JI, et al. HLA-E polymorphism in Amerindians from Mexico (Mazatecans), Colombia (Wayu) and Chile (Mapuches): evolution of MHC-E gene. *Tissue Antigens* 2007: **69** (Suppl 1): 132–5.
10. Tamouza R, Busson M, Fortier C, et al. *HLA-E*0101* allele in homozygous state favors severe bacterial infections in sickle cell anemia. *Hum Immunol* 2007: **68**: 849–53.
11. Grimsley C, Kawasaki A, Gassner C, et al. Definitive high resolution typing of HLA-E allelic polymorphisms: identifying potential errors in existing allele data. *Tissue Antigens* 2002: **60**: 206–12.
12. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci U S A* 2003: **100**: 177–82.
13. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988: **16**: 1215.
14. Grimsley C, Ober C. Population genetic studies of HLA-E: evidence for selection. *Hum Immunol* 1997: **52**: 33–40.
15. Castelli EC, Mendes-Junior CT, Deghaide NH, et al. The genetic structure of 3′untranslated region of the HLA-G gene: polymorphisms and haplotypes. *Genes Immun* 2010: **11**: 134–41.
16. Castelli EC, Mendes-Junior CT, Donadi EA. HLA-G alleles and HLA-G 14 bp polymorphisms in a Brazilian population. *Tissue Antigens* 2007: **70**: 62–8.
17. Castelli EC, Mendes-Junior CT, Veiga-Castelli LC, Roger M, Moreau P, Donadi EA. A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution. *Mol Biol Evol* 2011: **28**: 3069–86.
18. Excoffier L, Slatkin M. Incorporating genotypes of relatives into a test of linkage disequilibrium. *Am J Hum Genet* 1998: **62**: 171–80.
19. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 2005: **1**: 47–50.
20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005: **21**: 263–5.
21. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001: **68**: 978–89.
22. Excoffier L, Laval G, Balding D. Gametic phase estimation over large genomic regions using an adaptive window approach. *Hum Genomics* 2003: **1**: 7–19.
23. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992: **48**: 361–72.
24. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 1995: **86**: 248–9.
25. Ewens WJ. The sampling theory of selectively neutral alleles. *Theor Popul Biol* 1972: **3**: 87–112.
26. Watterson GA. The homozygosity test of neutrality. *Genetics* 1978: **88**: 405–17.
27. Slatkin M. An exact test for neutrality based on the Ewens sampling distribution. *Genet Res* 1994: **64**: 71–4.
28. Lancaster A, Nelson MP, Meyer D, Thomson G, Single RM. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. *Pac Symp Biocomput* 2003: 514–25.
29. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989: **123**: 585–95.
30. Nei M, Kumar S. *Molecular Evolution and Phylogenetics.*, *New York*: *Oxford University Press*, *2000.*
31. Stephens JC, Schneider JA, Tanguay DA, et al. Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 2001: **293**: 489–93.
32. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007: **24**: 1596–9.

33. Ohya K, Kondo K, Mizuno S. Polymorphism in the human class I MHC locus HLA-E in Japanese. *Immunogenetics* 1990: **32**: 205–9.
34. Pyo CW, Williams LM, Moore Y, et al. HLA-E, HLA-F, and HLA-G polymorphism: genomic sequence defines haplotype structure and variation spanning the nonclassical class I genes. *Immunogenetics* 2006: **58**: 241–51.
35. Tan Z, Shon AM, Ober C. Evidence of balancing selection at the HLA-G promoter region. *Hum Mol Genet* 2005: **14**: 3619–28.
36. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001: **409**: 928–33.
37. Apanius V, Penn D, Slev PR, Ruff LR, Potts WK. The nature of selection on the major histocompatibility complex. *Crit Rev Immunol* 1997: **17**: 179–224.
38. Hughes AL, Yeager M. Natural selection and the evolutionary history of major histocompatibility complex loci. *Front Biosci* 1998: **3**: d509–16.
39. Meyer D, Thomson G. How selection shapes variation of the human major histocompatibility complex: a review. *Ann Hum Genet* 2001: **65**: 1–26.
40. Penn DJ, Damjanovich K, Potts WK. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci U S A* 2002: **99**: 11260–4.
41. Bernatchez L, Landry C. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years?. *J Evol Biol* 2003: **16**: 363–77.
42. Potts WK, Wakeland EK. Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends Genet* 1993: **9**: 408–12.
43. Jordan WC, Bruford MW. New perspectives on mate choice and the MHC. *Heredity* 1998: **81**: 127–33.
44. Ferguson R, Ramanakumar AV, Richardson H, et al. Human leukocyte antigen (HLA)-E and HLA-G polymorphisms in human papillomavirus infection susceptibility and persistence. *Hum Immunol* 2011: **72**: 337–41.
45. Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moreau P. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cell Mol Life Sci* 2011: **68**: 369–95.
46. Kaiser BK, Barahmand-Pour F, Paulsene W, Medley S, Geraghty DE, Strong RK. Interactions between NKG2x immunoreceptors and HLA-E ligands display overlapping affinities and thermodynamics. *J Immunol* 2005: **174**: 2878–84.
47. Ulbrecht M, Couturier A, Martinozzi S, et al. Cell surface expression of HLA-E: interaction with human beta2-microglobulin and allelic differences. *Eur J Immunol* 1999: **29**: 537–47.
48. O'Callaghan CA, Tormo J, Willcox BE, et al. Structural features impose tight peptide binding specificity in the nonclassical MHC molecule HLA-E. *Mol Cell* 1998: **1**: 531–41.
49. Pietra G, Romagnani C, Manzini C, Moretta L, Mingari MC. The emerging role of HLA-E-restricted CD8+ T lymphocytes in the adaptive immune response to pathogens and tumors. *J Biomed Biotechnol* 2010: **2010**: 907092.
50. Julie DC, Buhler S, Frassati C, et al. Linkage disequilibrium between *HLA-G*0104* and *HLA-E*0103* alleles in Tswa Pygmies. *Tissue Antigens* 2011: **77**: 193–200.
51. Sullivan LC, Clements CS, Rossjohn J, Brooks AG. The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. *Tissue Antigens* 2008: **72**: 415–24.
52. Chamary JV, Parmley JL, Hurst LD. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat Rev Genet* 2006: **7**: 98–108.
53. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res* 2002: **12**: 996–1006.
54. Matte C, Lacaille J, Zijenah L, Ward B, Roger M. HLA-G and HLA-E polymorphisms in an indigenous African population. The ZVITAMBO Study Group. *Hum Immunol* 2000: **61**: 1150–6.
55. Danzer M, Polin H, Proll J, et al. Clinical significance of *HLA-E*0103* homozygosity on survival after allogeneic hematopoietic stem-cell transplantation. *Transplantation* 2009: **88**: 528–32.
56. Tripathi P, Naik S, Agrawal S. *HLA-E*0101* associated with recurrent spontaneous abortion. *J Turkish-German Gynecol Assoc* 2007: **8**: 278.
57. Hirankarn N, Kimkong I, Mutirangura A. HLA-E polymorphism in patients with nasopharyngeal carcinoma. *Tissue Antigens* 2004: **64**: 588–92.
58. Park KS, Park JS, Nam JH, Bang D, Sohn S, Lee ES. *HLA-E*0101* and *HLA-G*010101* reduce the risk of Behcet's disease. *Tissue Antigens* 2007: **69**: 139–44.

