



Contents lists available at ScienceDirect

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

Rapid Communication

Low variability at the *HLA-E* promoter region in the Brazilian populationLuciana Caricati Veiga-Castelli^{a,*}, João Marcos da Silveira Bulcão Bertuol^a, Erick C. Castelli^b, Eduardo Antonio Donadi^a^a Divisão de Imunologia Clínica, Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto (FMRP), Universidade de São Paulo (USP), 14049-900 Ribeirão Preto, SP, Brazil^b Departamento de Patologia, Faculdade de Medicina de Botucatu, Universidade Estadual Paulista – UNESP, 18618-970 Botucatu, SP, Brazil

ARTICLE INFO

Article history:

Received 26 May 2015

Revised 11 September 2015

Accepted 16 November 2015

Available online 17 November 2015

Keywords:

Brazilians

Haplotypes

HLA-E

Polymorphism

Promoter

ABSTRACT

Little information has been reported regarding the regulatory region of the *HLA-E* gene. Only a proximal segment (300 bp immediately before exon 1) has been described at IMGT/HLA database. We aimed to evaluate the genetic diversity of the promoter region by PCR amplification and DNA sequencing. We observed a pattern of sequencing interruption in the same position in all samples, suggesting the presence of a secondary structure hampering the DNA polymerase sliding during the sequencing process, which was confirmed after bioinformatics analysis. Considering the promoter region evaluated (nucleotide –440 to –1), only three variation sites were found, including one new variation site at position –104, and the concomitant deletions already described at positions –187 and –186. We concluded that the promoter region of the *HLA-E* gene presents few and rare variable sites in this population sample; however, the double-stranded branch formation hampered the evaluation of the distant promoter by conventional sequencing.

© 2015 American Society for Histocompatibility and Immunogenetics. Published by Elsevier Inc. All rights reserved.

The Brazilian population is one of the most heterogeneous worldwide, as a result of five centuries of integration among individuals from three continents: the European colonizers, mainly Portuguese, the African slaves and the Amerindians [1]. Hence, Brazil may be considered a rich repository of genetic variation and an excellent sampling for the characterization of the general variability of any potentially polymorphic genomic region. For example, most of the polymorphic sites described worldwide for the Human Leukocyte Antigen-G (*HLA-G*) and -E (*HLA-E*) genes considering 14 populations of the 1000Genomes project (phase 1) were also detected in Brazilians [2–6].

The non-classical class I human leukocyte antigen (HLA) genes, represented by the *HLA-G*, *HLA-E* and *HLA-F* loci, are located within the human Major Histocompatibility Complex and present a very low degree of variation in spite of being neighbor of the most variable genes at the human genome (the classical class I HLA genes) [7].

The non-classical *HLA-E* molecule 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 [8] and as an antigen-presenting molecule triggering a specific immune response [9]. 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. *HLA-E* interactions with such CD94 heterodimers can activate, inhibit, or have no effect on NK cell-mediated cytotoxicity and cytokine production [10].

HLA-E is the least polymorphic of all class I HLA genes, presenting a much more limited polymorphism than the classical loci and even other non-classical genes. A number of worldwide populations from all continents, including two West-African population samples [11], the French Southeastern, Congolese Teke and Tswa African pygmies [12], Chinese [13], Brazilian Southeastern [4] and others [14–17] have already been evaluated for the *HLA-E* gene variability and usually only two distinct molecules were detected, encoded by a series of coding alleles presenting synonymous or intronic mutations and just one non-synonymous polymorphism, supporting the idea of restrict polymorphism at *HLA-E* [18]. So far, only seventeen *HLA-E* alleles coding for six different proteins and a truncated one were recognized worldwide (The International Genetics Database [IMGT/HLA], database version 3.21.1).

* Corresponding author at: Avenida Bandeirantes, 3900, Monte Alegre, 14049-900 Ribeirão Preto, SP, Brazil.

E-mail address: luciana.veigacastelli@gmail.com (L.C. Veiga-Castelli).

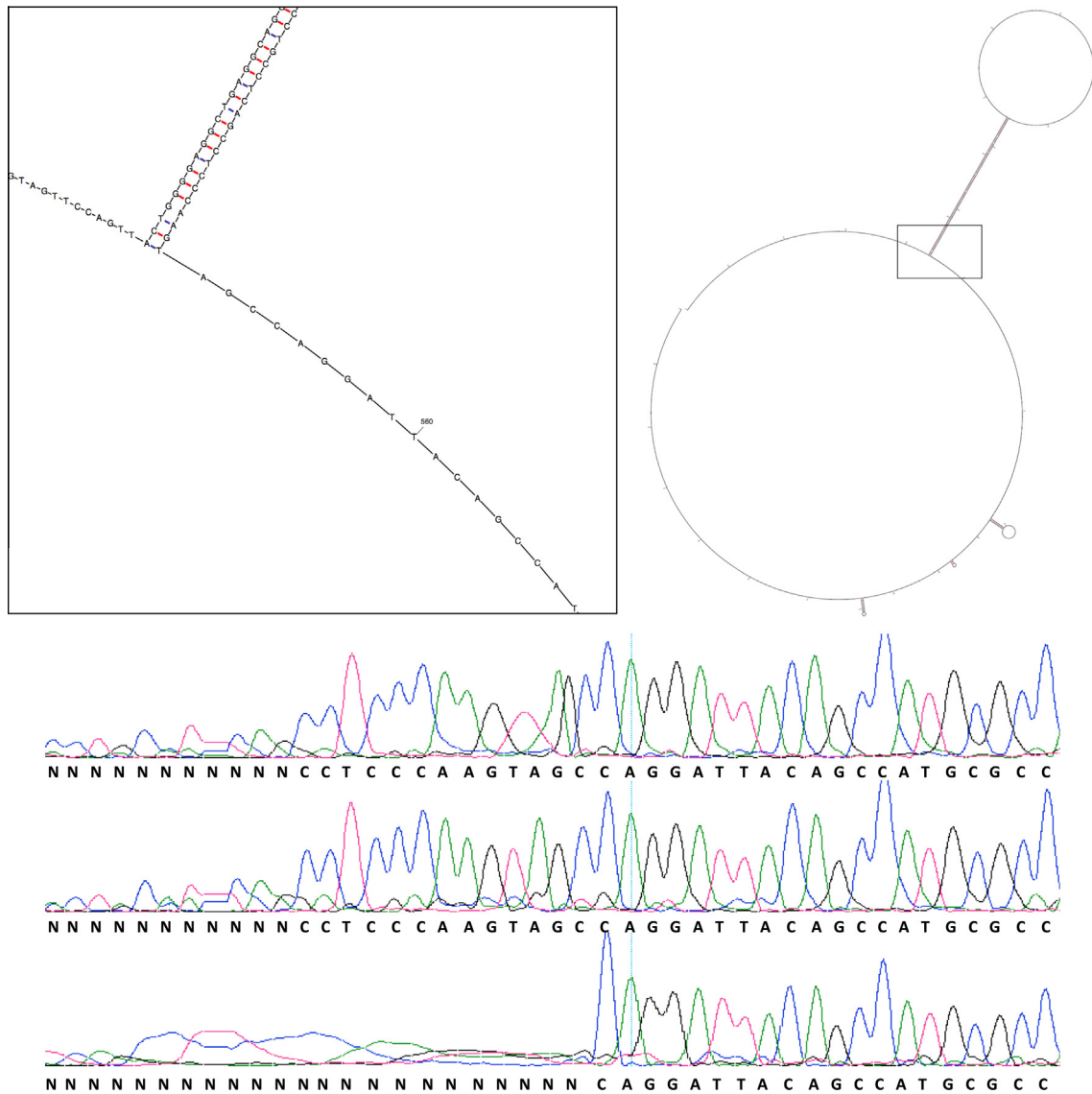


Fig. 1. Secondary structure of the *HLA-E* promoter region. The long annealed sequence shows that the molecule presents a repetitive and inverted sequence (Image generated by the mFold algorithm; Zuker, 2003 [25]). The large square is a zoom of the small square and do represent the exact point in which sequencing stops (bottom chromatograms).

Our group has evaluated the *HLA-E* coding region in the Brazilian population and found a low variability of this gene, corroborating what has been reported in other worldwide populations [4,5].

Little information has been reported in the literature regarding the *HLA-E* regulatory region. Only the proximal part (approximately 300 bp immediately before the beginning of exon 1) is considered by the IMGT/HLA database and is little polymorphic, with only two consecutive base deletions at positions –187 and –186 (considering the Adenine of the first translated ATG as +1), both detected in a single haplotype named as *E*01:01:01:02*. In addition, this region has not been entirely evaluated in 10 of the *HLA-E* alleles (*E*01:01:01:03*, *E*01:01:02*, *E*01:03:01:03*, *E*01:03:03*, *E*01:03:05*, *E*01:04*, *E*01:05*, *E*01:06*, *E*01:07* and *E*01:08*) (IMGT/HLA, database 3.21.1). Other HLA genes, such as the *HLA-G*, have already been evaluated and a high variability at the promoter region has been observed, with balancing selection increasing heterozygosity [19–21,2,6,22,23]. Few data has been reported regarding the promoter region of the *HLA-E* locus. In fact, only a recent study described the variability between nucleotides –118 and –1 in two African population samples [11]. Therefore, it is not clear if high

heterozygosity at the regulatory regions is a characteristic of the non-classical genes or if it is restricted to *HLA-G*.

Considering that these molecules are potential natural immunosuppressors, the analysis of the variability and structure of the promoter region of *HLA-E* gene in an admixed population such as Brazilians can contribute for the understanding of the structure of its regulatory region and the mechanisms underlying its regulation. In this matter, we aim to evaluate the variability of the *HLA-E* promoter region in Brazilian samples.

A hundred and thirteen samples of healthy bone marrow donors from Ribeirão Preto, State of São Paulo, Brazil, with approximately 50% of men and women, and with no age restriction, were randomly selected and evaluated after approval of the local Ethics Committee (#9153/2008). These samples are expected to present ancestry contribution of 79% Europeans, 14% Africans and 7% Amerindians, according to previous studies conducted on similar population samples [24]. The *HLA-E* promoter region was amplified in a single PCR reaction, using primers EpromoF (5'-CCCGCCTA TACGTTGTT-3') (designed for this study) and HE01R (5'-CTCTTA CCCAGGTGAAGCAGCG-3') [16], producing an amplification

product of approximately 1972-bp, encompassing a segment from nucleotide –981 up to nucleotide +992 of the *HLA-E* gene.

PCR was performed in a final volume of 50 μ L containing 0.7 \times PCR buffer provided by the Polymerase vendor, 0.2 mM of each dNTP, 15 pmol of each primer, 1.5 units of Long PCR DNA Polymerase (Fermentas, Vilnius, Lithuania) and 100 ng of genomic DNA. The initial denaturation step was carried out at 94 °C for 3 min, followed by 30 cycles at 94 °C for 30 s, 65 °C for 30 s and 68 °C for 135 s, and by a final extension step at 68 °C for 5 min. *HLA-E* variation in the coding region was evaluated by direct sequencing of the PCR product by using the primer EpromoR (5'-GTAGAGGCCGTTCCATCCC-3'). The sequencing reaction was carried out for each amplicon by using the BigDye[®] Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California, USA) and the sequences were obtained in an automatic ABI PRISM[®] 3100 Genetic Analyzer sequencer (Applied Biosystems). All sequences obtained were aligned with the genomic reference for this gene (NC_000006) and each single nucleotide polymorphism (SNP) detected was individually annotated.

Although we aimed to sequence the promoter segment from nucleotide –900 to –1, we observed a recurrent pattern of interruption of the sequencing always at the same position (approximately nucleotide –440) in every analyzed sample (Fig. 1). This interruption pattern is in agreement with the presence of a DNA secondary structure formed during the sequencing process, probably blocking the BigDye[®] DNA polymerase sliding during the sequencing process (Fig. 1). This secondary structure was inferred by using the mFold software [25] considering 1.5 mM of MgCl₂ and 60 °C, and it evidenced the presence of a large repetitive and inverted region, comprehending the repetition of a sequence of approximately 90 bases, creating a double-stranded branch (Fig. 1). This secondary structure coincides with the sequencing interruption point, showing that this structure is indeed responsible for such phenomenon. Strategies to avoid this structure, such as temperature increase, the use of DMSO and sequencing with other primers, including EpromoF and HE01R [4], were applied, but unsuccessfully. Because of this, all samples were sequenced and analyzed from the first base at exon 1 to position –440 of the promoter region, where the secondary structure begins. For future studies, Next Generation Sequencing (NGS) procedures can bypass this issue, since it relies on the parallel sequencing of thousands of single small molecules.

Considering the promoter region that was evaluated (nucleotide –440 up to –1), and the 113 samples, only three variation sites were found, including one new variation site at position –104 (Adenine to Guanine), occurring in five samples and confirmed after PCR and sequencing repetitions; and the concomitant deletions already described at the *HLA-E* promoter at positions –187 and –186 in just one sample. The polymorphic site at position –104 was recently found and described in two African samples with high frequency for the minor allele (around 11%) [11]. Therefore, we observed that the *HLA-E* promoter region between nucleotide –440 and –1, at least in this population sample, present a very low number of variable sites. The variable sites that were detected presented low frequencies for the minor allele, with the exception of the polymorphic site at position –104 (frequency of 2.2%). This is a relatively surprising fact since, in a very admixed population such as the Brazilian, we expected a higher genetic variability. In addition, a great number of polymorphic sites and high heterozygosity has already been observed for another non-classical HLA gene, *HLA-G*.

HLA-G gene has already been explored regarding its variability and structure at the regulatory regions, and also, the possible selective forces acting on it. In this context, at least 5 variation sites were found at the promoter region of the *HLA-G* in this very population (and same samples), at the same sequence interval analyzed

for *HLA-E* [2]. Of these 5 variation sites, all presented polymorphic frequencies with a minor allelic frequency higher than 8.5%. This high heterozygosity was previously associated with balancing selection acting on the *HLA-G* regulatory regions [2,26,6,22,23]. By these means, and considering that both *HLA-E* and *HLA-G* share many functions, we expected to observe the same variation pattern for *HLA-E*. However, opposite evidences were found. The *HLA-E* promoter region apparently follows the same conservation profile of the coding region [4]. The low variability at the promoter region could either be related with a hitchhiking effect of the purifying selection acting on the coding region or is a signal of the important role of this region in the *HLA-E* expression control.

The promoter region plays a mandatory role in the control of the transcriptional levels presented by a gene and consequently, in the control of the gene expression. For genes such as *HLA-G* and *HLA-E*, whose expression can be either beneficial or malefic depending on the situation, the high heterozygosity found in this region for *HLA-G* is justified, since haplotypes with different promoter structural features would be maintained. Nevertheless, such finding was not observed for the *HLA-E* gene, since only one invariable mechanism of control was observed, indicating therefore the importance of this regulatory function.

Strong selection acting at the promoter region, as it seems to occur based on our data, could result in the lack of variation at the coding region by a hitchhiking effect. Although previous studies have detected this lack of variability at the coding region, we should consider this low variability as a possible consequence of purifying selection acting at the *HLA-E* promoter region. This hypothesis regarding the influence of the promoter region over the coding region, although not proved yet, goes against previous literature reports. On the other hand, we can observe that the *HLA-E* gene variation increases according to the distance of the promoter region, since most of the alterations found at the *HLA-E* gene are located at exon 4, the most physically distant from the promoter already analyzed [4]. This phenomenon was also observed for two West-African population samples [11]. Further studies evaluating variability beyond exon 4 in different populations, including the transmembrane region and the 3' untranslated region, should be considered.

Hence, we conclude that the promoter region of the *HLA-E* gene in this particular segment is well conserved (few and rare variable sites) in this Brazilian population sample, but the presence of repetitive and inverted regions leading to the double-stranded branch formation hamper the evaluation of the distant promoter by conventional sequencing. Further studies using next generation sequencing methodologies should help in the investigation of this region.

Contributors

J.M.S.B.B. performed PCR amplifications, DNA purification and quantification and DNA sequencing. J.M.S.B.B. and L.C.V.C. performed the analysis. E.C.C. participated in the design of the study and performed statistical analysis. L.C.V.C., E.C.C. and E.A.D. conceived, designed and coordinated the study. L.C.V.C. wrote the manuscript. All authors have approved the final article.

Conflict of interest

I 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 conflict of interest.

Acknowledgments

This study was supported by the Brazilian Research Council (CNPq/Brazil – Grant 558476/2008. L.C.V.C. was supported by a postdoctoral fellowship (150329/2011-3) from CNPq/Brazil. J.M.S. B.B. was supported by fellowships (507474/2010-4 and 501350/2010-1) from CNPq/Brazil.

References

- [1] F.C. Parra, R.C. Amado, J.R. Lambertucci, J. Rocha, C.M. Antunes, S.D. Pena, Color and genomic ancestry in Brazilians, *Proc. Natl. Acad. Sci. USA* 100 (2003) 177–182.
- [2] E.C. Castelli, C.T. Mendes-Júnior, L.C. Veiga-Castelli, M. Roger, P. Moreau, E.A. Donadi, A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution, *Mol. Biol. Evol.* 28 (11) (2011) 3069–3086.
- [3] 1000 Genomes Project Consortium, G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, R.E. Handsaker, H.M. Kang, G.T. Marth, G.A. McVean, An integrated map of genetic variation from 1,092 human genomes, *Nature* 491 (2012) 56–65.
- [4] L.C. Veiga-Castelli, E.C. Castelli, C.T. Mendes Jr., W.A. da Silva Jr., M.C. Faucher, K. Beauchemin, M. Roger, P. Moreau, E.A. Donadi, 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 (2012) 15–24.
- [5] L.P. Felício, I.O. Porto, C.T. Mendes-Junior, L.C. Veiga-Castelli, K.E. Santos, R.P. Vianello-Brondani, A. Sabbagh, P. Moreau, E.A. Donadi, E.C. Castelli, Worldwide HLA-E nucleotide and haplotype variability reveals a conserved gene for coding and 3′ untranslated regions, *Tissue Antigens* 83 (2) (2014) 82–93.
- [6] E.C. Castelli, J. Ramalho, I.O. Porto, T.H. Lima, L.P. Felício, A. Sabbagh, E.A. Donadi, C.T. Mendes-Junior, Insights into HLA-G genetics provided by worldwide haplotype diversity, *Front. Immunol.* 6 (5) (2014) 476.
- [7] E.A. Donadi, E.C. Castelli, A. Arnaiz-Villena, M. Roger, D. Rey, P. Moreau, Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association, *Cell Mol Life Sci* 68 (2011) 369–395.
- [8] V.M. Braud, D.S. Allan, C.A. O’Callaghan, et al., HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C, *Nature* 391 (1998) 795–799.
- [9] P. Tomasec, V.M. Braud, C. Rickards, et al., Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40, *Science* 287 (2000) 1031.
- [10] J.J. Perez-Villar, I. Melero, A. Rodriguez, et al., Functional ambivalence of the Kp43 (CD94) NK cell-associated surface antigen, *J. Immunol.* 154 (1995) 5779–5788.
- [11] E.C. Castelli, C.T. Mendes-Junior, A. Sabbagh, I.O. Porto, A. Garcia, J. Ramalho, T. H. Lima, J.D. Massaro, F.C. Dias, C.V. Collares, E.A. Donadi, HLA-E coding and 3′ untranslated region variability determined by next-generation sequencing in two West-African population samples, *Hum. Immunol.* (2015).
- [12] D.C. Julie, S. Buhler, C. Frassati, A. Basire, V. Galicher, C. Baier, A. Essautier, A. Regnier, T. Granier, A.D. Lepfoundzou, et al., Linkage disequilibrium between HLA-G*0104 and HLA-E*0103 alleles in Tswa Pygmies, *Tissue Antigens* 77 (2011) 193–200.
- [13] X.X. Liu, F.H. Pan, W. Tian, Characterization of HLA-E polymorphism in four distinct populations in Mainland China, *Tissue Antigens* 80 (1) (2012) 26–35.
- [14] C. Grimsley, C. Ober, Population genetic studies of HLA-E: evidence for selection, *Hum. Immunol.* 52 (1997) 33–40.
- [15] C. Matte, J. Lacaille, L. Zijenah, B. Ward, M. Roger, HLA-G and HLA-E polymorphisms in an indigenous African population. The ZVITAMBO Study Group, *Hum. Immunol.* 61 (2000) 1150–1156.
- [16] C. Grimsley, A. Kawasaki, C. Gassner, N. Sageshima, Y. Nose, K. Hatake, D.E. Geraghty, A. Ishitani, Definitive high resolution typing of HLA-E allelic polymorphisms: identifying potential errors in existing allele data, *Tissue Antigens* 60 (2002) 206–212.
- [17] S. Metcalfe, M. Roger, M.C. Faucher, F. Coutlée, E.L. Franco, P. Brassard, The frequency of HLA alleles in a population of Inuit women of northern Quebec, *Int. J. Circumpolar Health* 5 (2013) 72.
- [18] A. Arnaiz-Villena, G. Vargas-Alarcon, J.I. Serrano-Vela, R. Reguera, J. Martinez-Laso, C. Silvera-Redondo, J. Granados, J. Moscoso, HLA-E polymorphism in Amerindians from Mexico (Mazatecos), Colombia (Wayu) and Chile (Mapuches): evolution of MHC-E gene, *Tissue Antigens* 69 (Suppl. 1) (2007) 132–135.
- [19] T.V. Hviid, S. Sørensen, N. Morling, Polymorphism in the regulatory region located more than 1.1 kilobases 5′ to the start site of transcription, the promoter region, and exon 1 of the HLA-G gene, *Hum. Immunol.* 60 (12) (1999 Dec) 1237–1244.
- [20] C. Ober, C.L. Aldrich, I. Chervoneva, C. Billstrand, F. Rahimov, H.L. Gray, T. Hyslop, Variation in the HLA-G promoter region influences miscarriage rates, *Am J Hum Genet* 72 (6) (2003 Jun) 1425–1435.
- [21] Z. Tan, A.M. Shon, C. Ober, Evidence of balancing selection at the HLA-G promoter region, *Hum. Mol. Genet.* 14 (23) (2005) 3619–3628.
- [22] E.C. Castelli, L.C. Veiga-Castelli, L. Yaghi, P. Moreau, E.A. Donadi, Transcriptional and posttranscriptional regulations of the HLA-G gene, *J. Immunol. Res.* 2014 (2014) 734068.
- [23] L. Gineau, P. Luisi, E.C. Castelli, J. Milet, D. Courtin, N. Cagnin, B. Patillon, H. Laayouni, P. Moreau, E.A. Donadi, et al., Balancing immunity and tolerance: genetic footprint of natural selection in the transcriptional regulatory region of HLA-G, *Genes Immun.* 16 (1) (2015) 57–70.
- [24] L.M.O. Ferreira, C.T. Mendes-Junior, C.E.V. Wiesel, M.R. Luizon, A.L. Simões, Genomic ancestry of a sample population from the State of São Paulo, Brazil, *Am. J. Hum. Biol.* 18 (2006) 702–705.
- [25] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res.* 31 (13) (2003) 3406–3415.
- [26] A. Sabbagh, P. Luisi, E.C. Castelli, L. Gineau, D. Courtin, J. Milet, J.D. Massaro, H. Laayouni, P. Moreau, E.A. Donadi, et al., Worldwide genetic variation at the 3′ untranslated region of the HLA-G gene: balancing selection influencing genetic diversity, *Genes Immun.* 15 (2) (2014) 95–106.