New evidence for balancing selection at the HLA-G locus in South Amerindians

Tiago Degani Veit1,2, Juciana Cazarolli1, Francisco Mauro Salzano1,2, Marion Schiengold1 and José Artur Bogo Chies1,2

1Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.
2Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

Abstract

HLA-G is a non-classical HLA (Human Leukocyte Antigen) molecule characterized by limited tissue distribution under normal physiological conditions and low variability at both DNA and protein levels. Several studies suggest that HLA-G could play a role, as an immunoregulatory molecule, in situations as diverse as transplantation, cancer, viral infections and inflammatory diseases. A total of 237 individuals from 21 South American tribes speaking nine different linguistic families were studied in relation to the 14 bp insertion/deletion polymorphism at the HLA-G gene. A consistent (seven in nine) excess of heterozygosity in samples classified by language was obtained. Our data supply evidences for balancing selection acting at the HLA-G 14 bp INDEL region. Enhanced fetal survival in a pathogen-rich environment may account for these findings.

Keywords: HLA-G 14 bp INDEL, balancing selection, Amerindians, pathogen resistance.

Introduction

HLA-G is a non-classical HLA (Human Leukocyte Antigen) molecule characterized by limited tissue distribution under normal physiological conditions and low polymorphism at both DNA and protein levels. HLA-G presents unique characteristics as compared to its classical HLA counterparts, such as the expression of multiple isoforms generated by alternative splicing, which can be membrane-bound (G1-G4) or secreted (G5-G7). Moreover, it is the only HLA molecule capable to form dimers. Unlike the classical HLA-I A, B and C molecules, HLA-G seems to be involved in immune modulation rather than in antigen presentation. It has been shown that HLA-G plays a major role in immunosuppression, interacting with cells of the immune system and suppressing the immune response by different mechanisms. These mechanisms include the inhibition of T lymphocytes (CTL) and NK cells cytotoxic activity, protection of class I-negative or allogeneic tumors from NK-mediated anti-tumor immunity (Riteau et al., 2001) and even tumor cells that express ligands for NK activator receptors such as MICA (Menier et al., 2002; Rouas-Freiss et al., 2003). Also, it was shown that the HLA-G molecule can inhibit CD4+ T cell alloproliferative responses (Riteau et al., 1999), the proliferation of T and peripheral blood NK cells (Bahri et al., 2006; Caumartin et al., 2007; LeMaoult et al., 2007), and can also act on antigen presenting cells (APC) by inhibiting their maturation and function (Horuzsko et al., 2001). Additionally, HLA-G may exert long-term immunotolerogenic effects through the generation of suppressor cells (reviewed by Carosella et al., 2008) and even cells which do not transcribe HLA-G may temporarily become HLA-G+ and suppressive through intercellular uptake of HLA-G-containing membrane patches, a mechanism called trogocytosis (LeMaoult et al., 2007). On the other hand, it was postulated that HLA-G can activate uterine NK cells, leading to the production of pro-inflammatory and pro-angiogenic factors, which would be important for placental development (Rajagopalan et al., 2006).

HLA-G expression was first described in the cytotrophoblast, and therefore the first studies on this molecule concentrated in its role during pregnancy. However, as soon as the first immunotolerogenic features of this molecule became known, the scientific community began to turn its attention to other immunological situations in which HLA-G could play a role, such as transplantation, cancer, viral infections, and inflammatory diseases (reviewed by Veit et al., 2010). Scientific interest on HLA-G has continuously increased since the description of the molecule in...
1987, as evidenced by the fact that half of the 1352 entries for the term “HLA-G” in Pubmed (until July 2011) were generated in the last six years. At the Universidade Federal do Rio Grande do Sul, the Immunogenetics Laboratory of the Genetics Department is developing research on the variability of the gene HLA-G, with special focus on autoimmune diseases but also approaching infectious diseases and materno-foetal tolerance (Vianna et al., 2007; Veit et al., 2008, 2009; Cordero et al., 2009; Consiglio et al., 2011).

The HLA-G gene is located at the Major Histocompatibility Complex (MHC) region, which comprises a collection of genes at the short arm of chromosome 6 (6p21.3). As previously mentioned, HLA-G presents limited polymorphism as compared to classic HLA molecules: only 47 alleles have been described to date, as compared to 1698, 2271 and 1213 alleles found at the A, B and C loci, respectively (IMGT/HLA Database), and it codifies 15 different proteins. Moreover, this limited variability is distributed along the three alpha domains, while in classic HLA molecules it is concentrated around the peptide binding groove.

Another striking difference between HLA-G and other HLA genes is its unique promoter region. While in classic HLA genes the promoter elements are located within 220 bp upstream the ATG start codon, the HLA-G regulatory elements are located on a region that spans ~1.5 kb upstream from its start codon. Many typical HLA promoter elements are deleted or modified in the HLA-G promoter region, rendering HLA-G expression unresponsive to classical HLA stimulator factors such as nF-kB, IRF1 and CIITA. Another remarkable HLA-G promoter characteristic is its relatively high variability. To date, more than 29 SNPs have been identified in this region, many present near known regulatory elements (Donadi et al., 2011). These polymorphisms may have an important impact on HLA-G expression, as reported by previous studies (Ober et al., 2003; 2006). Moreover, haplotype analyses pointed to the existence of two lineages which may be under balancing selection. These lineages probably present different promoter activity patterns (Tan et al., 2005).

The 3’ untranslated region (3’UTR) also seems to play an important role in HLA-G expression, mainly through post-transcriptional regulatory mechanisms. Castelli et al. (2010) described eight different haplotypes in a Brazilian population. In silico analysis of this region has identified numerous putative microRNA binding sites, which may influence HLA-G expression depending on the allele and/or biological context. Also, according to the latest information, eleven polymorphic positions have been identified at the HLA-G 3’UTR, many of which overlapping putative microRNA binding sites (Donadi et al., 2011). Among the 3’UTR polymorphisms, a 14 bp insertion/deletion (INDEL) (rs1704) has been extensively studied due to its potential involvement in alternative splicing processes and post-transcriptional regulation. Transcripts with the 14 nucleotide sequence (ins) can undergo an additional splicing step which removes 92 bases from the region in which this sequence is located. This deletion is thought to influence mRNA stability since, following actinomycin treatment, the HLA-G transcripts lacking these 92 bases were shown to be more stable than the “complete” mRNAs in placental cells, (Rousseau et al., 2003). On the other hand, several studies have repeatedly reported the association of the ins allele with lower soluble HLA-G levels, and even the lack of detectable HLA-G expression in the plasma of homzygotes for the ins allele (Hviid et al., 2004; Rizzo et al., 2005, 2006, 2008). Interestingly, it was also described that the 14 bp INDEL lies in a region which is a putative binding site for many microRNAs (Castelli et al., 2009), and regulation through microRNA binding was also postulated to contribute to HLA-G expression control (Veit and Chies, 2009). Nevertheless, it is evident that the net effect of microRNAs in HLA-G expression will be very difficult to access, as microRNA profiles may substantially vary among tissues and biological states.

Mendes-Junior et al. (2007) hypothesized that since the insertion allele is associated to recurrent miscarriages and other pregnancy complications, its frequency might be rather low in isolated populations such as the Amazonian Indians, and this was indeed observed in their study. However, their data did not support that, in the studied populations, the 14 bp variant frequencies would depart from neutrality. Here we present new data on Amerindian populations that provide evidence for balancing selection acting at the HLA-G 14 bp INDEL region.

Materials and Methods

Subjects

We obtained the frequencies of the allelic variants of the 14 bp polymorphism in 237 South Amerindians from 21 different tribes and nine different linguistic families, classified according to Campbell, 1997 (Table 1). They live all the way from 3°45’ N to 39°20’ S, and from 50°10’ W to 69°30’ W. These samples have been collected along five decades by members of our group and were conveniently stored for different types of testing. Informed oral consent was obtained from all participants, since they were illiterate. Consent was obtained according to the Helsinki Declaration, and this procedure was approved by the Brazilian National Ethics Commission (CONEP Resolution 123/98).

Polymerase chain reaction (PCR) amplification of exon 8 of the HLA-G gene and genotyping

Genotyping was performed by PCR as previously described (Hviid et al., 2002). Briefly, 100 ng of genomic DNA was amplified in a 25 μL reaction, with final concentrations as follows: PCR buffer 1X, dNTP 0.2 mM, MgCl₂ 1.5 mM, Taq DNA polymerase 0.75 U, and 10 pmol of each primer (GE14HLAG: 5’-
GTGATGGGGCTTTAAGGTTCACC, RHG4: 5'—GGAAGGATGCAGTCTCA GCATGA). Thermocycling conditions were 94 °C for 2 min, 35 cycles of 94 °C for 30 s, 64 °C for 60 s and 72 °C for 60 s, and a final extension step of 72 °C for 10 min. The amplified PCR products were visualized in 6% polyacrylamide gels stained with ethidium bromide. The amplicon sizes for the 14 bp polymorphism were: 224 bp for the +14 bp allele and 210 bp for the -14 bp allele.

### Statistical analysis

Allele frequencies and observed heterozygosity (Ho) were computed by the direct counting method, and adherence of phenotypical proportions to expectations under Hardy-Weinberg equilibrium were tested by the complete enumeration method using the GENEPOP 3.4 software (Rousset, 2008). Expected heterozygosity values (He) were estimated by the ARLEQUIN 3.0 software (Excoffier et al., 2007). Departure from selective neutrality was tested by Slatkin’s implementation of the Ewens-Watterson homozygosity neutrality test (Slatkin, 1994); and the calculations were carried out using the PyPop software (Lancaster et al., 2003).

### Results and Discussion

Due to the relatively low number of individuals studied per tribe, we opted for grouping tribes according to their linguistic group, since it has previously been suggested that, among South American Indians, populations whose members speak the same language are genetically homogeneous and may be viewed as the ultimate evolutionary unit (Fagundes et al., 2002). The 14 bp deletion allele was the most common one in seven of nine linguistic groups, while in two others (Karib and Chapacura) the 14 bp insertion prevailed. Two Amazonian linguistic samples deviated significantly from Hardy-Weinberg expectations (Tupi and Mura), both with an excess of heterozygotes (Table 2). Seven of the nine linguistic group samples also presented higher observed heterozygote frequencies (Ho) as compared to those expected (He), and four of these groups (Mapugungu, Mura, Tupi and Zamuco) presented Ho frequencies above 0.6. To investigate whether natural selection could be reflected in these values, the Ewens-Watterson’s test of neutrality was performed. All linguistic groups presented negative normalized F values (Ho lower than He), with P-values ranging from 0.0024 to 0.1958 (Table 2). Among them, four reached statistical significance, significantly deviating from the neutral evolution hypothesis; curiously, three were members of the Macro-Tupi linguistic family (Kariri Tupi, Tupi Mondé and Tupi) and the Karib (Table 2). This finding is remarkable, considering that Macro-Tupi encompasses a wide group of languages spoken by tribes which have a widespread distribution in the Brazilian Amazon. Mendes-Junior et al. (2007), studied seven different Amazonian tribes and none presented a genotype distribution that departed from neutrality, in spite of the fact that 14 out of the 16 villages (87%) presented negative F values. It is worthy of note that there is no overlap between the tribes or linguistic groups of the two studies. Our
deviation was observed for haptoglobin (data not shown, served for acid phosphatase, and only among the Zamuco respectively). As expected, no significant deviation was observed for the number of individuals (188 and 171 individuals, respectively). The same analyses were performed on data already available for two apparently neutral markers (haptoglobin and acid phosphatase) in a subset of the individuals tested for polymorphisms. These two biallelic markers were chosen due to the availability of data on a representative number of individuals (188 and 171 individuals, respectively). As expected, no significant deviation was observed for acid phosphatase, and only among the Zamuco a deviation was observed for haptoglobin (data not shown, available on request).

The occurrence of balancing selection at the HLA-G promoter region has previously been described (Tan et al., 2005). Considering that HLA-G expression and levels were already related to different situations (either in physiological or pathological conditions) and that this molecule, depending on the context, may be deleterious or advantageous, balancing selection at this locus seems to be a plausible possibility. Recently, Castelli et al. (2011) made a comprehensive review of the HLA-G gene polymorphism and haplotypes in a Brazilian urban cohort, evidencing a high linkage disequilibrium along the whole length of the gene. In this same work, the authors revealed evidence for balancing selection acting on the regulatory regions only (5′ and 3′ UTRs) and on the HLA-G locus as a whole. We cannot rule out that the evidence of balancing selection observed in our data could be due to a hitchhiking effect caused by a linkage disequilibrium between the 14 bp locus and the HLA-G promoter region. Nevertheless, the compelling evidence for the functionality of the 14 bp insertion in alternative splicing and its potential role in post-transcriptional regulation by microRNA binding make us believe that the 14 bp INDEL might also be an adaptive factor, influencing HLA-G expression patterns and probably is related to survival of heterozygous fetuses due to resistance to pathogens (Mendes-Junior et al., 2007). In conclusion, our data corroborate the evidence of balancing selection at the HLA-G gene, highlighting important regulatory roles of this molecule in the immune system.

Acknowledgments

We would like to thank Dr. Sidia Maria Callegari-Jacques for her help with the statistical analysis. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) 135450/2009-8, 479438/2009-9 and 160530/2011-3; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) 10/1516-6.

References

Bahi R, Hirsch F, Josse A, Rouas-Freiss N, Bidere N, Vasquez A, Carosella ED, Charpentier B and Durrbach A (2006) Soluble HLA-G inhibits cell cycle progression in human alloreactive T lymphocytes. J Immunol 176:1331-1339.

Campbell L (1997) American Indian Languages. The Historical Linguistics of Native America. Oxford University Press, Oxford, 512 pp.

Carosella ED, HoWangYin K-Y, Favier B and LeMaoult J (2008) HLA-G-dependent suppressor cells: Diverse by nature, function, and significance. Hum Immunol 69:700-707.

Castelli EC, Moreau P, Oya e Chioromatto A, Mendes-Junior CT, Veiga-Castelli LC, Yaghi L, Giuliani S, Carosella ED and Donadi EA (2009) In silico analysis of microRNAS targeting the HLA-G 3′ untranslated region alleles and haplotypes. Hum Immunol 70:1020-1025.

Castelli EC, Mendes-Junior CT, Deghaide NH, de Albuquerque RS, Muniz YC, Simões RT, Carosella ED, Moreau P and Donadi EA (2010) The genetic structure of 3′ untranslated region of the HLA-G gene: Polymorphisms and haplotypes. Genes Immun 11:134-141.

Castelli EC, Mendes-Junior CT, Veiga-Castelli LC, Roger M, Moreau P and Donadi EA (2011) A comprehensive study of polymorphic sites along the HLA-G gene: Implication for gene regulation and evolution. Mol Biol Evol 28:3069-3086.

Caumartin J, Favier B, Daouya M, Guillard C, Moreau P, Carosella ED and LeMaoult J (2007) Trogocytosis-based generation of suppressive NK cells. EMBO J 26:1423-1433.

Table 2 - Sample sizes (n), 14-bp polymorphism allele frequencies, observed (Ho) and expected (He) heterozygosities, Hardy-Weinberg Equilibrium (HWE) probability values (significant values in boldface), and Ewens-Watterson’s neutrality test results on Amerindsians samples classified in nine linguistic groups.

| Linguistic group | n  | -14 bp | +14 bp | Ho   | He   | HWE p-value | Ewens-Watterson test |
|------------------|----|--------|--------|------|------|-------------|----------------------|
| Mapudungu        | 8  | 0.625  | 0.375  | 0.7500 | 0.50000 | 0.4420 | -1.3139; p = 0.1958 |
| Karib            | 28 | 0.482  | 0.518  | 0.46429 | 0.50844 | 0.7149 | -1.7484; p = 0.0219 |
| Chapacura        | 15 | 0.467  | 0.533  | 0.53333 | 0.51494 | 1.0000 | -1.6273; p = 0.00517 |
| Macro-Jê         | 49 | 0.642  | 0.357  | 0.46939 | 0.46392 | 1.0000 | -1.6078; p = 0.01173 |
| Mura             | 25 | 0.620  | 0.380  | 0.68000 | 0.48082 | 0.0446 | 0.1554; p = 0.177 |
| Kariri-Tupi      | 45 | 0.522  | 0.478  | 0.55556 | 0.50462 | 0.5585 | -1.8317; p = 0.0202 |
| Tupi-Mondé       | 24 | 0.500  | 0.500  | 0.50000 | 0.51064 | 1.0000 | -1.7247; p = 0.0074 |
| Tupi             | 32 | 0.581  | 0.419  | 0.77419 | 0.49498 | 0.0021 | 0.1691; p = 0.0024 |
| Zamuco           | 11 | 0.545  | 0.455  | 0.72727 | 0.51948 | 0.2565 | -1.5644; p = 0.0755 |


Consiglio CR, Veit TD, Monticielo OA, Mucenic T, Xavier RM, Brenol JC and Chies JA (2011) Association of the HLA-G gene +3142C > G polymorphism with systemic lupus erythematosus. Tissue Antigens 77:540-545.

Cordero EA, Veit TD, Silva MA, Jacques SM, Silla LM and Chies JA (2009) HLA-G polymorphism influences the susceptibility to HCV infection in sickle cell disease patients. Tissue Antigens 74:308-313.

Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D and 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.

Excoffier L, Laval G and Schneider S (2007) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47-50.

Fagundes NJ, Bonatto SL, Callegari-Jacques SM and Salzano FM (2002) Genetic, geographic, and linguistic variation among South American Indians: Possible sex influence. Am J Phys Anthropol 117:68-78.

Horuzsko A, Lenfant F, Munn DH and Mellor AL (2001) Maturation of antigen-presenting cells is compromised in HLA-G transgenic mice. Int Immunol 13:385-394.

Hviid TV, Hylenius S, Hoegh AM, Kruse C and Christiansen OB (2002) HLA-G polymorphisms in couples with recurrent spontaneous abortions. Tissue Antigens 60:122-132.

Hviid TV, Rizzo R, Christiansen OB, Melchiorri L, Lindhard A and Baricordi OR (2004) HLA-G and IL-10 in serum in relation to HLA-G genotype and polymorphisms. Immunogenetics 56:135-141.

Lancaster A, Nelson MP, Meyer D, Thomson G and Single RM (2003) PyPop: A software framework for population genomics: Analyzing large-scale multi-locus genotype data. Pac Symp Biocomput 2003:514-525.

LeMaoult J, Caumartin J, Daouya M, Favier B, Le Rond S, Gonzalez A and Carosella ED (2007) Immune regulation by pretenders: Cell-to-cell transfers of HLA-G make effector T cells act as regulatory cells. Blood 109:2040-2048.

Mendes-Junior CT, Castelli EC, Simes AC, Simes AL and Donadi EA (2007) HLA-G 14-bp polymorphism at exon 8 in Amerindian populations from the Brazilian Amazon. Tissue Antigens 69:255-260.

Menier C, Riteau B, Carosella ED and Rouas-Freiss N (2002) MICA triggering signal for NK cell tumor lysis is counteracted by HLA-G1-mediated inhibitory signal. Int J Cancer 100:63-70.

Ober C, Aldrich CL, Chervoneva I, Billstrand C, Rahimov F, Gray HL and Hyslop T (2003) Variation in the HLA-G promoter region influences miscarriage rates. Am J Hum Genet 72:1425-1435.

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

Rajagopalan S, Bryceson YT, Kuppusamy SP, Geraghty DE, van der Meer A, Joosten I and Long EO (2006) Activation of NK cells by an endocytosed receptor for soluble HLA-G. PLoS Biology 4:e9.

Riteau B, Menier C, Khalil-Daher I, Sedlick C, Dausset J, Rouas-Freiss N and Carosella ED (1999) HLA-G inhibits the allogeneic proliferative response. J Reprod Immunol 43:203-11.

Riteau B, Menier C, Khalil-Daher I, Martinozzi S, Pla M, Dausset J, Carosella ED and Rouas-Freiss N (2001) HLA-G1 co-expression boosts the HLA class I-mediated NK lysis inhibition. Int Immunol 13:193-201.

Rizzo R, Hviid TV, Stignani M, Balboni A, Grappa MT, Melchiorri L and Baricordi OR (2005) The HLA-G genotype is associated with IL-10 levels in activated PBMCs. Immunogenetics 57:172-181.

Rizzo R, Rubini M, Govoni M, Padovan M, Melchiorri L, Stignani M, Carturian S, Ferretti S, Totta F and Baricordi OR (2006) HLA-G 14-bp polymorphism regulates the methotrexate response in rheumatoid arthritis. Pharmacogenet Genomics 16:615-623.

Rizzo R, Hviid TV, Govoni M, Padovan M, Rubini M, Melchiorri L, Stignani M, Carturian S, Grappa MT, Fotinidi M, et al. (2008) HLA-G genotype and HLA-G expression in systemic lupus erythematosus: HLA-G as a putative susceptibility gene in systemic lupus erythematosus. Tissue Antigens 71:520-529.

Rouas-Freiss N, Moreau P, Menier C and Carosella ED (2003) HLA-G in cancer: A way to turn off the immune system. Semin Cancer Biol 13:325-336.

Rousseau P, Le Discorde M, Mouillot G, Marcou C, Carosella ED and Moreau P (2003) The 14 bp Deletion-Insertion polymorphism in the 3’UT region of the HLA-G gene influences HLA-G mRNA stability. Hum Mol Genet 64:1005-1010.

Rousset F (2008) genepop’007: A complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8:103-106.

Slatkin M (1994) An exact test for neutrality based on the Ewens sampling distribution. Genet Res 64:71-74.

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

Veit TD and Chies JA (2009) Tolerance vs. immune response — microRNAs as important elements in the regulation of the HLA-G gene expression. Transpl Immunol 20:229-231.

Veit TD, Vianna P, Scheibel I, Brenol CV, Brenol JC, Xavier RM, Delgado-Cañedo A, Gutierrez JE, Brindizzi AP, Schuler-Faccini L, et al. (2008) Association of the HLA-G 14-bp insertion/deletion polymorphism with juvenile idiopathic arthritis and rheumatoid arthritis. Tissue Antigens 71:440-446.

Veit TD, Cordero EA, Mucenic T, Monticielo OA, Brenol JC, Xavier RM, Delgado-Cañedo A and Chies JA (2009) Association of the HLA-G 14 bp polymorphism with systemic lupus erythematosus. Lupus 18:424-430.

Veit TD, Vianna P and Chies JA (2010) HLA-G - From fetal tolerance to a regulatory molecule in inflammatory diseases. Curr Immunol Rev 6:1-15.

Vianna P, Dalmáz CA, Veit TD, Tedoldi C, Reisenberg I and Chies JA (2007) Immunogenetics of pregnancy: Role of a 14-bp deletion in the maternal HLA-G gene in primiparous pre-eclamptic Brazilian women. Hum Immunol 68:668-674.

Internet Resources

IMGT/HLA Database, June 2011, http://www.ebi.ac.uk/imgt/hla/stats.html (June 2011).

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.