



**UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
FACULDADE DE MEDICINA**

**Viviane Aparecida de Oliveira Ciriaco**

**DIVERSIDADE GENÉTICA DO GENE *MICA* EM  
POPULAÇÕES MUNDIAIS**

Dissertação apresentada à Faculdade de Medicina, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Botucatu, para obtenção do título de Mestra em Patologia.

Orientador: Prof. Dr. Erick da Cruz Castelli  
Coorientadora: Profa. Dra. Camila Ferreira Bannwart Castro

**Botucatu  
2023**

Viviane Aparecida de Oliveira Ciriaco

DIVERSIDADE GENÉTICA DO GENE *MICA* EM  
POPULAÇÕES MUNDIAL

Dissertação apresentada à Faculdade de  
Medicina, Universidade Estadual Paulista  
“Júlio de Mesquita Filho”, Câmpus de  
Botucatu, para obtenção do título de  
Mestra em Patologia.

Orientador (a): Prof.Dr. Erick da Cruz Castelli  
Coorientadora: Profa.Dra. Camila Ferreira Bannwart Castro

Botucatu  
2023

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP

BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Ciriaco, Viviane Aparecida de Oliveira.

Diversidade genética do gene *Mica* em populações mundiais  
/ Viviane Aparecida de Oliveira Ciriaco. - Botucatu, 2023

Dissertação (mestrado) - Universidade Estadual Paulista  
"Júlio de Mesquita Filho", Faculdade de Medicina de Botucatu

Orientador: Erick da Cruz Castelli

Coorientador: Camila Ferreira Bannwart Castro

Capes: 40105008

1. Variação genética. 2. Genes. 3. Polimorfismo genético.

Palavras-chave: Diversidade; *MICA*; *MICA*\*del; NGS.

***Dedicatória***

Dedico este trabalho à minha Mãe, que sempre me apoiou e me incentivou a seguir meus sonhos, sempre acreditando em mim e me inspirando a ser uma pessoa melhor.

## ***Agradecimientos***

*“It matters not what someone is born, but what they grow to be.”*

*-Alvo Dumbledore*

Agradeço ao meu orientador, Prof. Dr. Erick C. Castelli, pela oportunidade e confiança em me aceitar como aluna de mestrado. Obrigada por todo apoio durante todo o mestrado, por todos os ensinamentos e por todo incentivo.

A minha coorientadora Camila Ferreira Bannwart Castro, por ter me apresentado ao GeMBio. Obrigada por ter me inspirado durante a faculdade a seguir a carreira da pesquisa e amar imunologia. Obrigada por todo carinho, dedicação, paciência e pela ajuda.

Agradeço a minha Mãe, uma mulher batalhadora que sempre fez de tudo por mim, obrigada por sempre me apoiar em tudo e me aceitar como eu sou, seu apoio foi essencial para que eu conseguisse levar a vida com mais leveza. Obrigada por além de ser minha mãe, ter uma relação de parceria e amizade. Sou muito grata a você por tudo. Te amo.

Ao meu padrasto, que sempre esteve ao meu lado e ao lado da minha mãe, muito obrigada pelo apoio.

A todos meus amigos, em especial as minhas amigas da escola Jéssica Oliveira, Raphaela Sales e Maria Eduarda, por todo apoio, companheirismo, por estarem presentes em todos os momentos da minha vida. Amo muito vocês.

As minhas amigas e “roommates”, Raphaela Neto Pereira e Gabriela Sato Paes, sou muito grata por ter conhecido vocês. Obrigada por toda paciência em me ouvir, por todas as risadas, conselhos, conversas aleatórias e filosóficas. Amo vocês, contem comigo para sempre, em todos os momentos.

Ao grupo MICA, especialmente a Amanda e Joyce por toda ajuda, conversas filosóficas sobre MICA e pela amizade e companheirismo de vocês. Contem comigo para tudo que precisarem, sou muito grata a vocês.

As minhas parceiras e parceiros de laboratório, Nayane Silva, Heloísa Andrade, Isabelle Silva, Lívia Ramos, Marcel Ferreira, Ícaro Scalisse, João Toloi, Ingrid Ferraz, Luise Tini, Nathaly Ogata, Brenda Tibúrcio, Douglas Nadur e Richard Ribeiro. Obrigada por toda a ajuda, pelas conversas e pela amizade de vocês.

Ao programa de Pós-Graduação em Patologia da Faculdade de Medicina de Botucatu (FMB-UNESP), pela oportunidade.

Aos membros da banca avaliadora, por aceitarem o convite e contribuírem com a minha formação.

À CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) 131710/2021-4, que possibilitou dedicação exclusiva ao trabalho desenvolvido através da bolsa de mestrado.

A todos, que contribuíram com os meus estudos e pesquisa, durante todos esses anos.

## Sumário

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Lista de Figuras .....                                                  | 11 |
| Lista de Tabelas .....                                                  | 11 |
| RESUMO .....                                                            | 15 |
| Capítulo I – Revisão de Literatura .....                                | 17 |
| INTRODUÇÃO .....                                                        | 18 |
| Estrutura e função das moléculas MICA e MICB .....                      | 19 |
| Expressão das proteínas MICA e MICB .....                               | 22 |
| Polimorfismos dos genes <i>MICA</i> e <i>MICB</i> .....                 | 23 |
| MICA/B em infecções, câncer e no transplante .....                      | 27 |
| JUSTIFICATIVA .....                                                     | 29 |
| OBJETIVOS .....                                                         | 30 |
| Objetivo Geral .....                                                    | 30 |
| Objetivos Específicos.....                                              | 30 |
| REFERÊNCIAS .....                                                       | 31 |
| Capítulo II – Artigo.....                                               | 37 |
| Introduction .....                                                      | 40 |
| Methods .....                                                           | 41 |
| Samples included in this study, sequencing strategy, and read alignment | 41 |
| Calling <i>MICA</i> and <i>MICB</i> copy numbers .....                  | 41 |
| Calling <i>MICA</i> , <i>MICB</i> , and <i>HLA-B</i> alleles.....       | 41 |

|                                                                                                  |    |
|--------------------------------------------------------------------------------------------------|----|
| Results .....                                                                                    | 42 |
| <i>Detecting MICA*del</i> .....                                                                  | 42 |
| <i>Other MICA null alleles</i> .....                                                             | 43 |
| The frequency of <i>MICA</i> null alleles in different biogeographical regions....               | 44 |
| Linkage disequilibrium between <i>MICA*del</i> and <i>MICB</i> , <i>HLA-B</i> , and <i>HLA-C</i> | 45 |
| Discussion .....                                                                                 | 47 |
| References .....                                                                                 | 50 |
| Supplementary .....                                                                              | 53 |

## Lista de Figuras

|                                                                                                                                                   |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Família dos genes <i>MIC</i> . Os genes em azul são pseudogenes, os verdes são genes funcionais e em vermelho o gene <i>HLA-B</i> ..... | 18 |
| Figura 2. Diferentes vias de indução da expressão de MICA/B: ligação de LPS no receptor TLR-4 e ATM/ATR.....                                      | 23 |

## Lista de Figuras do Artigo

|                                                                                             |    |
|---------------------------------------------------------------------------------------------|----|
| Figure 1: <i>MICA</i> copy number estimation by comparing read depth in two references..... | 43 |
| Figure 2: Frequency of MICA null alleles in different biogeographic regions.....            | 44 |
| Figure 3: Frequency of MICA null alleles in different population samples.....               | 45 |

## Lista de Figuras do Artigo - Material Suplementar

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| Figure S1: An overview of the sequencing coverage across the MICA and MICB loci.....                       | 53 |
| Figure S2: An overview of the sequencing coverage at the upstream border of the MICA deleted region.....   | 54 |
| Figure S3: An overview of the sequencing coverage at the downstream border of the MICA deleted region..... | 55 |

## Lista de Tabelas

### Lista de Tabelas - Capítulo I

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| Tabela 1: Classificação dos alelos de MICA de acordo com a região de microsatélite presente no éxon 5..... | 24 |
|------------------------------------------------------------------------------------------------------------|----|

## Lista de Tabelas do Artigo

|                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------|----|
| Table 1: The frequency of different haplotypes observed for chromosomes 6 carrying MICA*del or MICA*partial..... | 46 |
|------------------------------------------------------------------------------------------------------------------|----|

## Lista de Tabelas do Artigo - Material Suplementar

|                                                                                                                                        |    |
|----------------------------------------------------------------------------------------------------------------------------------------|----|
| Table S1: Frequency of MICA*del, <i>MICB</i> , <i>HLA-B</i> and <i>HLA-C</i> haplotypes.....                                           | 56 |
| Table S2: Frequency of <i>MICA</i> , <i>MICB</i> and <i>HLA-B</i> haplotype distribution in different biogeographical populations..... | 58 |

### Lista de Abreviaturas e Siglas

|                                |                                                                                          |
|--------------------------------|------------------------------------------------------------------------------------------|
| <b>ADAM</b>                    | Desintegrina e Metaloproteínase (do inglês, A Disintegrin and Metalloproteinase)         |
| <b>ATM</b>                     | Ataxia Telangiectasia Mutada                                                             |
| <b>ATR</b>                     | Ataxia Telangiectasia Rad-3relacionada                                                   |
| <b>BWA MEM</b>                 | Burrows-Wheeler Alignment Tool                                                           |
| <b>COVID-19</b>                | Coronavirus Disease of 2019                                                              |
| <b>DAP10</b>                   | DNAX-Activating Protein of 10KDa                                                         |
| <b>DNA</b>                     | Ácido Desoxirribonucleico (do inglês, Deoxyribonucleic Acid)                             |
| <b>ERp5</b>                    | Proteína 5 do retículo endoplasmático (do inglês, Endoplasmic Reticulum Protein 5)       |
| <b>GPI</b>                     | Glicosilfosfatidilinositol (do inglês, Glycosylphosphatidylinositol)                     |
| <b>GrB2</b>                    | Growth Factor Receptor-bound protein 2                                                   |
| <b>HCMV</b>                    | Citomegalovírus Humano (do inglês, Human Cytomegalovirus)                                |
| <b>HGDP</b>                    | Projeto de Diversidade do Genoma Humano (do inglês, Human Genome Diversity Project)      |
| <b>HLA</b>                     | Antígeno Leucocitário Humano (do inglês, Human Leukocyte Antigens)                       |
| <b>HSP70</b>                   | Proteína do choque térmico (do inglês, 70-kDa Heat Shock Protein)                        |
| <b>IFN-<math>\alpha</math></b> | Interferon Alfa                                                                          |
| <b>IFN-<math>\beta</math>1</b> | Interferon beta-1                                                                        |
| <b>IFN<math>\gamma</math></b>  | Interferon gamma                                                                         |
| <b>Ig</b>                      | Imunoglobulina                                                                           |
| <b>IL</b>                      | Interleucina (do inglês, Interleukin)                                                    |
| <b>IMGT</b>                    | International Immunogenetics Database                                                    |
| <b>LD</b>                      | Desequilíbrio de ligação (do inglês, Linkage Disequilibrium)                             |
| <b>LPS</b>                     | Lipopolissacarídeo                                                                       |
| <b>LTA</b>                     | Linfotoxina alfa (do inglês, Lymphotoxin Alpha)                                          |
| <b>Met</b>                     | Metionina                                                                                |
| <b>MHC</b>                     | Complexo Principal de Histocompatibilidade (do inglês, Major Histocompatibility Complex) |
| <b>MICA</b>                    | Major Histocompatibility Complex Class I-Chain Related A                                 |
| <b>MICA-TM</b>                 | Região transmembrana da proteína MICA                                                    |

**MICB** Major Histocompatibility Complex Class I-Chain Related B

**MICA\*del** Deleção de MICA

**miRNA** microRNA

**mMICA** MICA de membrana (do inglês, Membrane MICA)

**MMP** Metaloproteinases de Matriz (Matrix Metalloproteinase)

**mRNA** RNA Mensageiro (do inglês, Messenger RNA)

**NF-Y** Fator de Transcrição Nuclear Y (do inglês, Nuclear Transcription Factor Y)

**NK** Natural Killer

**NKG2** Natural Killer Group 2

**NKG2D** Natural Killer Group 2 molécula D

**PI3K** Fosfatidilinositol-3-Cinase (do inglês, Phosphatidylinositol 3-Kinase)

**RNA** Ácido Ribonucleico (do inglês, Ribonucleic Acid)

**SARS-CoV-2** Coronavirus 2 da Síndrome Respiratória Aguda Grave (do inglês, Severe Acute Respiratory Syndrome Coronavirus 2)

**sMICA** MICA solúvel (do inglês, Soluble MICA)

**SNP** Polimorfismo de base única (do inglês, Single Nucleotide Polymorphism)

**Sp1** Transcription Factor Specificity Protein 1

**Sp3** Transcription Factor Specificity Protein 3

**Sp4** Transcription Factor Specificity Protein 4

**TAP** Transportador associado ao processamento de antígeno (do inglês, Transporter Associated with Antigen Processing)

**TAPBP** TAP binding protein

**TGFβ** Fator de Crescimento Transformante Beta (do inglês, Transforming Growth Factor Beta)

**TLR4** Receptor do tipo Toll 4 (do inglês, Toll-like receptor 4)

**TM** Região Transmembrana (do inglês, Transmembrane)

**TNF** Fator de Necrose Tumoral (do inglês, Tumor Necrosis Factor)

**UTR** Região não traduzida (do inglês, Untranslated region)

**UV** Radiação Ultravioleta

**Val** Valina

**Vav1** Vav Guanine Nucleotide Exchange Factor 1

**WGS** Whole-Genome Sequencing

Impacto científico esperado: O desenvolvimento de uma ferramenta computacional que permite analisar genes de interesse para transplantes e imunoterapia tumoral na população mundial.

Impacto social: Os achados poderão servir de base para estudos futuros de associação com transplantes e doenças de acordo com as diferentes frequências encontradas, permitindo maiores chances de sucesso nos tratamentos.

Impacto econômico: os dados disponíveis nos bancos de dados mundiais poderão auxiliar no desenvolvimento de métodos de diagnósticos e tratamentos mais efetivos.

## RESUMO

O gene *MICA* codifica glicoproteínas expressas em situações de estresse celular, como no caso de infecções, câncer e estresse oxidativo. Quando *MICA* é expresso na membrana, a molécula interage com o receptor ativador NKG2D presente na superfície de células Natural Killer (NK), linfócitos T $\alpha\beta$  CD8 e linfócitos T $\gamma\delta$ , levando a morte das células-alvo. Porém, essa proteína pode ser liberada na forma solúvel, quando isso acontece, ela interage com o mesmo receptor ativador NKG2D, mas leva a regulação negativa das células NK. O gene *MICA* é altamente polimórfico, com mais 531 alelos descritos até o momento, alguns desses polimorfismos apresentam impacto funcional descritos. *MICA* também pode estar ausente em alguns indivíduos, devido a uma deleção de 100pb que acontece entre os genes *HLA-B* e *MICB*, deletando toda a região de *MICA*, dessa forma, não há produção da molécula de *MICA*. Assim, este trabalho utilizou um pipeline de bioinformática utilizando o hla-mapper para trabalhar com os polimorfismos dos genes *MICA* e criamos uma estratégia específica para analisar a deleção do gene. Analisamos a diversidade de *MICA* em mais de 5.000 amostras de populações do mundo todo, incluindo amostras brasileiras, população altamente miscigenada. Encontramos *MICA*\*del em todas as populações analisadas, exceto na Oceania, com maior frequência na população asiática e americana, especialmente em amostras com ancestralidade nativo americanas. A maioria das amostras com a deleção de *MICA*, são heterozigotas e apresentam ao menos uma cópia funcional da molécula de MIC, porém encontramos duas amostras com a deleção completa de *MICA* e *MICB* nulo, sugerindo que outros mecanismos devem compensar a ausência das moléculas MIC.

**Palavras-chave:** *MICA*, *MICA*\*del, Diversidade, NGS.

## ABSTRACT

The *MICA* gene encode glycoproteins expressed in situations of cellular stress, such as in the case of infections, tumor cells, oxidative stress, among others. *MICA* is expressed on the membrane, the molecule interacts with the activating receptor NKG2D present on the surface of NK cells, T $\alpha$  $\beta$  CD8 cells and T $\gamma$  $\delta$  cells, leading to the death of the target cells. However, this protein can be released in soluble isoform, when this happens, it interacts with the same activating receptor NKG2D, but leads to negative regulation of NK cells. The *MICA* gene is highly polymorphic, with over 531 alleles described to date, some of these polymorphisms have described functional impact. *MICA* may also be absent in some individuals, due to a 100pb deletion that occurs between the *HLA-B* and *MICB* genes, deleting the entire *MICA* region, thus, there is no production of the *MICA* molecule. Thus, this study used a specific methodology to work with *MICA* gene polymorphisms and we created a specific strategy to analyze gene deletion. We analyzed *MICA* diversity in more than 5,000 samples from populations around the world, including Brazilian samples (highly mixed population). We found *MICA*\*del in all analyzed populations, except in Oceania, more frequently in the Asian and American population, especially in samples with Native American ancestry. Most samples with *MICA* deletion are heterozygous and have at least one functional copy of the MIC molecule, but we found two samples with complete *MICA* deletion and null *MICB*, suggesting that another mechanism must compensate for the absence of MIC molecules.

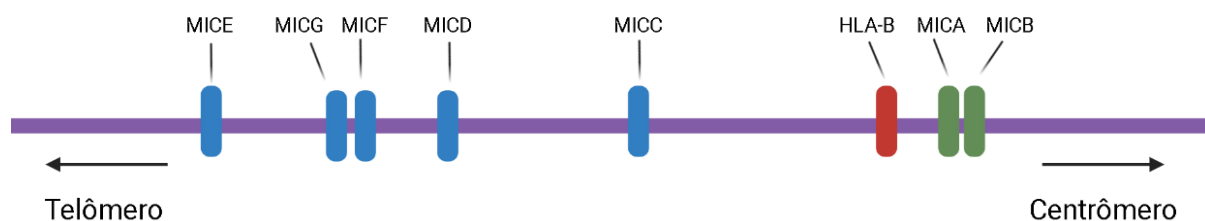
**Keywords:** *MICA*, *MICA*\*del, Diversity, NGS.

## **Capítulo I – Revisão de Literatura**

## INTRODUÇÃO

O complexo principal de histocompatibilidade humano (do inglês, *major histocompatibility complex* - MHC) é uma região do genoma altamente polimórfica, com diversos genes relacionados ao sistema imunológico <sup>1</sup>. O MHC está localizado em 6p21.3 e apresenta genes relacionados com apresentação antigênica (genes *HLA* clássicos de classe I e II), processamento de antígenos (TAP1, TAP2, TAPBP), modulação da ação de células T e *Natural Killer* (NK) (genes *HLA* não clássicos de classe I), citocinas (e.g., TNF, LTA) e sinalização de estresse celular (*MICA*, *MICB*) <sup>2-4</sup>.

Os genes *MICA* e *MICB* estão localizados, aproximadamente, 46kb e 141kb distantes do *HLA-B* em direção ao centrômero, respectivamente <sup>5</sup>. Foram descritos pela primeira vez em 1994 <sup>6,7</sup> e estão entre os últimos identificados dentro do MHC humano, na região de classe I, e incluem os loci: *MICA*, *MICB*, *MICC*, *MICD*, *MICE*, *MICF* e *MICG* (Figura 1). Apenas *MICA* e *MICB* codificam proteínas funcionais, enquanto os demais são pseudogenes <sup>5,6</sup>.



**Figura 1:** Família dos genes *MIC*. Os genes em azul são pseudogenes, os verdes são genes funcionais e em vermelho o gene *HLA-B*. Adaptado de Risti e Bicalho <sup>8</sup>.

As proteínas de *MICA/B* são expressas na superfícies de células endoteliais, queratinócitos, fibroblastos <sup>9</sup>, monócitos de sangue periférico <sup>10</sup> e no epitélio gástrico <sup>11</sup>, além de serem sintetizadas como marcador de estresse em condições como o câncer <sup>12</sup> e infecções <sup>13,14</sup>.

A molécula de *MICA* e *MICB* é expressa na superfície da membrana, mas pode ser liberada na isoforma solúvel. Quando expressa na superfície celular atua como ligante interagindo através dos domínios  $\alpha 1$ ,  $\alpha 2$  com o receptor

ativador *NKG2D*, expressos nas células *Natural killer*, células  $T\alpha\beta$  CD8 e células  $T\gamma\delta$ , ajudando na eliminação de células-alvos tumorais ou infectadas com patógenos intracelulares <sup>13,15</sup>. Quando são liberadas na isoforma solúvel, ao interagirem com *NKG2D*, levam à regulação negativa da sua expressão, inibindo ou diminuindo a atividade citotóxica dessas células <sup>16</sup>.

Diversos estudos já encontraram a associação entre polimorfismos de *MICA* e *MICB* com doenças. Por exemplo, os alelos *MICA*\*045 e *MICA*\*008 foram associados com a suscetibilidade a dengue, enquanto o alelo *MICB*\*002 com a proteção à doença <sup>17,18</sup>. Os alelos *MICA*\*009 e *MICA*\*A6 foram associados com a artrite reumatoide e *MICB*\*002 e *MICB*\*014 com a proteção <sup>19</sup>. Variantes de *MICA* associadas a maior expressão do mRNA, possivelmente de forma solúvel, e menor expressão de *MICB* foram relacionadas à susceptibilidade à infecção por SARS-CoV-2 <sup>20</sup>.

A proximidade dos genes *MICA* e *HLA-B* favorece um alto grau de desequilíbrio de ligação entre seus alelos <sup>21</sup>. Em muitos casos, o mesmo alelo de *MICA* aparece associado a diversos alelos de *HLA-B*, e vice-versa. Como exemplos dessas associações temos os alelos *HLA-B*\*07:02, *HLA-B*\*08:01, *HLA-B*\*40:01 e *HLA-B*\*40:02 com *MICA*\*008 e *HLA-B*\*51:01 com *MICA*\*006, *MICA*\*007, *MICA*\*009. Além disso, alguns alelos ocorrem sempre associados entre si, como *HLA-B*\*14:02 e *MICA*\*011 <sup>22</sup>. A associação desses alelos pode estar envolvida com a suscetibilidade a algumas doenças, como *MICA*\*002 e *HLA-B*\*35 está associado à hanseníase <sup>23</sup>.

Embora a diversidade dos genes *MICA* e *MICB* seja reduzida quando comparada aos genes *HLA* clássicos envolvidos com apresentação de antígenos, um grande número de alelos distintos foram descritos até o momento: 531 alelos para *MICA*, codificando 280 proteínas diferentes e 244 alelos para *MICB*, gerando 44 proteínas diferentes (HLA/IMGT version 3.51 - 2023-1) <sup>24</sup>. Alguns estudos mostraram que o alto polimorfismo dos genes *MIC* está relacionado com rejeição precoce do enxerto, além das associações com doenças descritas anteriormente <sup>25,26</sup>.

### **Estrutura e função das moléculas *MICA* e *MICB***

As moléculas *MIC* apresentam similaridade estrutural nos domínios extracelulares  $\alpha 1$ ,  $\alpha 2$  e  $\alpha 3$  com as moléculas *HLA* de classe I clássicas. No

entanto, as moléculas MIC não se associam a  $\beta$ 2-microglobulina e apresentam um sulco de ligação ao peptídeo extremamente estreito <sup>27</sup>, impedindo a apresentação de antígenos aos linfócitos T <sup>10,28</sup>. A molécula é constituída por três domínios extracelulares semelhantes a imunoglobulinas (Ig), codificados pelos éxons 2, 3 e 4 respectivamente, a região transmembrana codificada pelo éxon 5, a cauda citoplasmática e a região 3'NT (região não traduzida) codificada pelo éxon 6 <sup>6,29</sup>. O éxon 1 codifica o peptídeo líder, que direciona a molécula para a superfície celular.

Os genes *MICA/B* codificam glicoproteínas expressas em situações de estresse celular que irão interagir com o receptor ativador NKG2D resultando na ativação das células NK que leva a liberação de grânulos citotóxicos e morte da célula-alvo <sup>13,15,30</sup>.

As células NK são células da imunidade inata e contribuem efetivamente para a adaptativa <sup>31</sup> e possuem como principal função lisar células infectadas, malignas ou danificadas através de sinais ativadores que geram uma citotoxicidade contra a célula-alvo. As células NK são controladas através de um balanço de sinais fornecido por receptores ativadores e inibidores <sup>32</sup>. Diversos receptores regulam a ação das células NK, entre eles os receptores da família NKG2, aqui destacamos o receptor ativador NKG2D, devido a interação com as moléculas de MICA e MICB <sup>13,33</sup>. O NKG2D é um receptor ativador presente na superfície das células NK, linfócitos T  $\gamma\delta$  e linfócitos T  $\alpha\beta$  CD8 + <sup>13</sup>. A expressão de NKG2D pode ser estimulada pela ação de algumas citocinas, como a IL-2, IL-7, IL-12 <sup>34</sup>, IL-15 <sup>35</sup> e IFN- $\alpha$  <sup>36,37</sup> regulam positivamente a expressão de NKG2D, enquanto o TGF- $\beta$  <sup>38</sup>, IFN- $\beta$  <sup>37</sup>, IL-21 <sup>39</sup>, IL-4, IL-12 e IFN $\gamma$  <sup>37</sup> modulam negativamente a expressão de NKG2D. Essa regulação negativa também pode acontecer devido a superexposição a ligantes solúveis ou ligados a membrana <sup>8,36</sup>.

NKG2D não possui elementos em sua cauda citoplasmática que leve a citotoxicidade, por isso ele se associa com a proteína adaptadora DAP10 que possui um *motif* (YINM) <sup>40</sup>. Essa estrutura é formada por duas moléculas de NKG2D e quatro de DAP10 <sup>41</sup>. A interação do ligante ao complexo NKG2D/DAP10 leva a fosforilação do domínio YINM, assim, permite o recrutamento do complexo proteína 2 (Grb2)/Vav1 ligada ao receptor do fator de crescimento e a ativação da fosfatidil-inositol-3-quinase (PI3K) <sup>40,42-44</sup> e,

consequentemente, essa cascata de sinalização leva a ativação das células NK e morte da célula-alvo <sup>45</sup>.

A indução de estresse nas células leva a formação das glicoproteínas MICA e MICB, que irão se ancorar na membrana celular. Isso permitirá a interação com o receptor ativador NKG2D e consequente lise dessa célula-alvo <sup>13,33</sup>. Porém, essas glicoproteínas podem ser liberadas de forma solúvel devido a diferentes mecanismos, que serão descritos adiante.

A liberação de MICA solúvel pode acontecer devido a polimorfismos que modificam o quadro de leitura gerando um *stop códon* prematuro no éxon 5 <sup>46</sup>. Assim, a região transmembrana e a cauda citoplasmática não são traduzidas, gerando uma proteína que apresenta os domínios de interação com NKG2D mas não possui a estrutura que a âncora à membrana plasmática <sup>46,47</sup> (Figura 2). Essa proteína truncada pode se associar a uma âncora GPI para fixação na membrana, mas pode ser recrutada por exossomos e liberada de forma solúvel <sup>48</sup>.

Outro mecanismo de liberação de MICA solúvel, e o principal de MICB, está relacionado com clivagem da molécula pela ação de metaloproteinases de matriz (MMP) e a *disintegrin and metalloproteinase (ADAM) proteins* <sup>49,50</sup>. As MMPs levam a destruição da matriz extracelular, invasão da membrana basal e angiogênese. Muitas ADAMs são proteases que estão ligadas a membrana e sua principal função é a liberação proteolítica de ectodomínios de proteínas transmembranares da superfície celular <sup>51-54</sup>. A proteína 5 do retículo endoplasmático (ERp5) se liga ao domínio  $\alpha 3$  levando a uma mudança conformacional, permitindo assim a clivagem de MICA/B de superfície <sup>55</sup>. Dessa forma, polimorfismos que levam a substituição de aminoácidos no domínio  $\alpha 3$  inibem a liberação de MICA/B solúvel <sup>56</sup>. Após a mudança conformacional por ERp5, as metaloproteinases cortam MICA/B em algum ponto próximo ao domínio  $\alpha 3$ , liberando os domínios domínio  $\alpha 1$  e  $\alpha 2$ , local de ligação do NKG2D <sup>57</sup>.

Quando MICA/B é liberado na forma solúvel, ele interage com o mesmo receptor ativador NKG2D, levando a internalização do receptor, que poderá ser encaminhado até os lisossomos e degradado. Assim, os ligantes solúveis levam a regulação negativa das células NK, devido a diminuição de ligantes de estresse (moléculas MICA/B) e de receptor ativador disponível <sup>58-60</sup>.

### **Expressão das proteínas MICA e MICB**

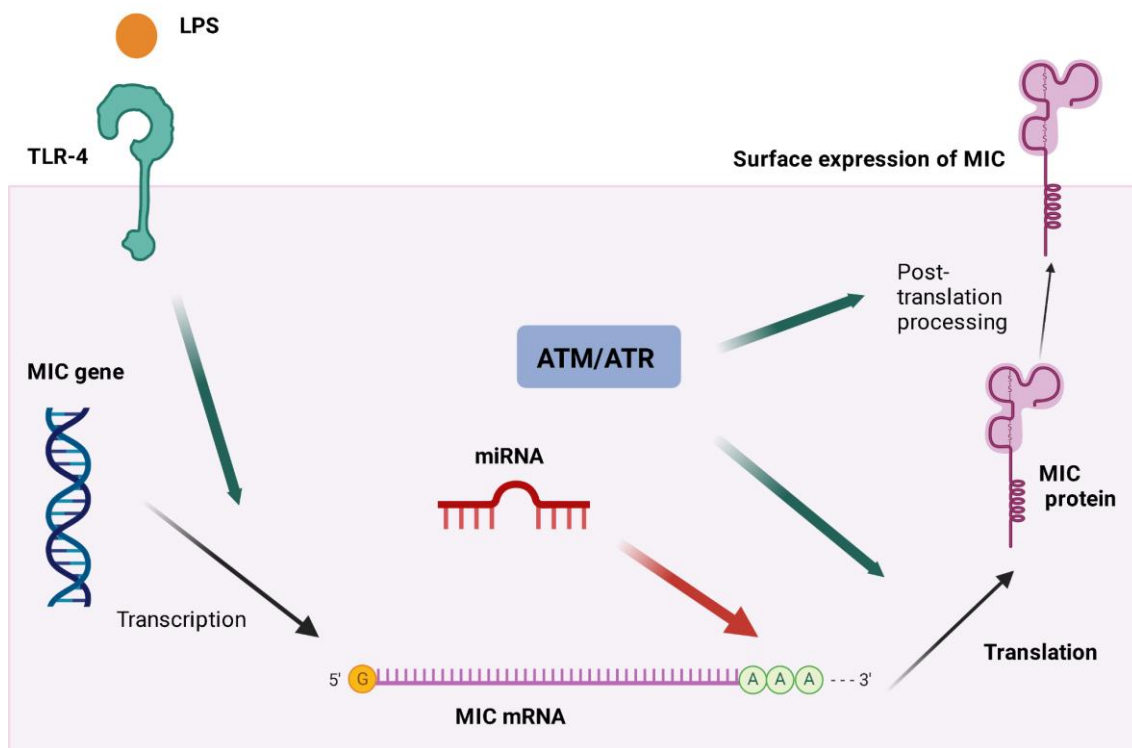
Como citado anteriormente, a expressão de MICA/B está limitada em células normais, porém pode ser regulada positivamente em situações de estresse celular, como, por exemplo, choque térmico, irradiação UV, infecções virais e estresse genotóxico/oncogênico <sup>61,62</sup>. Os genes *MICA* e *MICB* apresentam elementos de resposta ao choque térmico semelhante aos dos genes *HSP70* <sup>28</sup>, i.e, uma caixa CCAAT que se liga ao fator de transcrição nuclear Y (NF-Y) e uma caixa GC que se liga aos fatores de transcrição Sp1, Sp3 e Sp4 <sup>63,64</sup>.

A regulação positiva da expressão de *MICA/B* pode ocorrer por diferentes vias. Os agentes genotóxicos levam a indução da expressão de MIC pela via de dano ao DNA iniciada por ATR (ataxia telangiectasia e Rad3 relacionado) e/ou ATM (ataxia telangiectasia mutada). Essas quinases possuem um papel na detecção de quebras no DNA de fita dupla ou simples, respectivamente. Esta mesma via também pode ser ativada devido a infecções virais <sup>65</sup> e células tumorais <sup>66</sup>. Nos macrófagos, a via ATR/ATM regula MIC por um mecanismo translacional ou pós-transcricional, uma vez que a inibição dessa via por inibidores prejudicou a expressão de MIC de superfície celular mas não afetou a regulação positiva do mRNA de MIC <sup>67</sup>. Outros trabalhos mostram o envolvimento da via ATR/ATM na expressão do mRNA <sup>68,69</sup>, indicando que o envolvimento desta via no controle de expressão do mRNA depende do ligante, do tipo de célula e/ou do estímulo específico <sup>67</sup>.

O LPS, componente da parede celular de bactérias gram-negativas, se liga no TLR4 <sup>70</sup>, um receptor de reconhecimento de padrões que reconhece padrões moleculares associados a patógenos, como o LPS que é liberado durante a infecção. Esta ligação leva a liberação de mediadores inflamatórios, como citocinas e quimiocinas <sup>71,72</sup>. Porém, é importante controlar a resposta imunológica para evitar inflamação excessiva, que pode levar a quadros graves de sepse <sup>62,71</sup>. Assim, macrófagos podem expressar proteínas MICA e MICB pode regular a resposta inflamatória <sup>62,71,72</sup> (Figura 2).

A regulação da expressão das proteínas MIC pode acontecer de forma pós-transcricional, devido a ação de microRNAs (miRNAs) que tem como alvo a região 3' não-traduzida do mRNA de *MICA* e *MICB*. Em condições de

homeostase, esses miRNAs estabelecem um limiar de expressão para impedir a expressão de MICA e MICB e consequentemente levar a morte celular de células saudáveis <sup>73</sup> (Figura 3).



**Figura 2:** Diferentes vias de indução da expressão de MICA/B: ligação de LPS no receptor TLR-4 e ATM/ATR. As setas verdes indicam estímulos ativadores e a seta vermelha representa a inibição da tradução de MICA/B através da ação do miRNA. Adaptado de Eissmann et al., 2010 <sup>62</sup>.

### Polimorfismos dos genes *MICA* e *MICB*

O gene *MICA* é altamente polimórfico, com 531 alelos descritos até o momento segundo o banco de dados IPD/IMGT/HLA <sup>24</sup>. Estudos demonstraram que alguns polimorfismos do gene *MICA* pode afetar a estrutura da proteína produzida, sua localização (membrana ou solúvel) <sup>74</sup> e a afinidade de ligação de MICA com seus receptores <sup>75</sup>.

Uma das principais variações de *MICA* relacionada com mudanças na estrutura da proteína é um microsatélite presente no éxon 5, que codifica a região transmembrana da molécula de MICA, também conhecido como MICA-TM. Esse microsatélite apresenta repetições de trinucleotídeos (GCT), que codificam o aminoácido Alanina. Portanto, a proteína MICA terá um número de

Alaninas correspondente ao número de repetições neste microssatélite. Os alelos do gene *MICA* (as diferentes sequências conhecidas) podem ser classificados de acordo com a quantidade de Alaninas em: A4, A5, A6, A7, A8, A9, A10 e A5.1 (Tabela 1) <sup>76</sup>. O MICA\*A5.1 se difere dos demais devido a inserção de uma Guanina (rs67841474), que resulta em uma mudança no quadro de leitura e introdução de um stop códon prematuro, que impede a tradução do domínio transmembrana e a cauda citoplasmática. Assim, gera uma proteína com âncora GPI o que resulta no recrutamento de MICA para exossomos e consequente liberação na forma de vesículas extracelulares, sendo liberada de forma solúvel <sup>46–48,77</sup>.

A separação dos alelos de MICA de acordo com a quantidade de Alaninas, é considerada uma segunda nomenclatura para os alelos, apesar disso, pouco se sabe sobre o impacto funcional dessa região de microssatélite. Diversos estudos relacionam a quantidade de Alaninas presente no MICA-TM a doenças <sup>78,79</sup>, mas como a quantidade de Alaninas afeta a molécula ainda não foi descrito.

**Tabela 1:** Classificação dos alelos de *MICA* de acordo com o número de Alaninas do microssatélite presente no éxon 5.

| STR | Alelos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| A4  | MICA*001, MICA*007, MICA*012, MICA*018, MICA*029, MICA*043, MICA*045, MICA*051, MICA*059, MICA*060, MICA*061, MICA*063N, MICA*079, MICA*100, MICA*109, MICA*112, MICA*142, MICA*162, MICA*168, MICA*172, MICA*177, MICA*190, MICA*209, MICA*213, MICA*222N, MICA*225, MICA*232, MICA*243, MICA*244, MICA*245, MICA*250, MICA*253, MICA*258, MICA*266, MICA*270, MICA*272, MICA*279, MICA*281                                                                                                     |
| A5  | MICA*010, MICA*016, MICA*019, MICA*022, MICA*027, MICA*033, MICA*048, MICA*054, MICA*056, MICA*057, MICA*062, MICA*064N, MICA*065, MICA*066, MICA*069, MICA*076, MICA*089, MICA*099, MICA*104, MICA*114, MICA*118, MICA*121, MICA*124, MICA*135, MICA*159, MICA*165, MICA*170, MICA*174, MICA*176, MICA*191, MICA*194, MICA*197, MICA*198, MICA*206, MICA*212, MICA*216, MICA*218, MICA*228, MICA*231, MICA*235, MICA*236, MICA*238, MICA*239, MICA*249, MICA*251, MICA*252, MICA*273, MICA*274, |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|      | MICA*280, MICA*286N, MICA*291                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| A5.1 | MICA*008, MICA*023, MICA*028, MICA*053, MICA*058, MICA*070, MICA*073, MICA*080, MICA*082, MICA*085, MICA*087, MICA*088N, MICA*094, MICA*096N, MICA*097, MICA*103, MICA*113, MICA*117, MICA*128, MICA*130, MICA*137, MICA*140, MICA*145, MICA*147, MICA*150, MICA*154, MICA*155, MICA*156, MICA*175, MICA*180, MICA*181, MICA*185, MICA*192, MICA*193, MICA*196, MICA*201, MICA*202, MICA*203, MICA*204, MICA*217, MICA*224, MICA*237, MICA*242, MICA*247, MICA*248, MICA*261, MICA*262, MICA*264, MICA*268, MICA*271, MICA*277, MICA*278, MICA*282, MICA*283, MICA*284, MICA*285                                                                                                                |
| A6   | MICA*004, MICA*006, MICA*009, MICA*011, MICA*024, MICA*026, MICA*030, MICA*047, MICA*049, MICA*067, MICA*074, MICA*077, MICA*095, MICA*101, MICA*102, MICA*105Q, MICA*107N, MICA*108, MICA*115, MICA*116, MICA*119, MICA*120, MICA*122, MICA*123, MICA*126, MICA*127, MICA*134, MICA*136, MICA*138, MICA*141, MICA*144, MICA*146, MICA*148, MICA*149, MICA*151, MICA*152, MICA*153, MICA*157, MICA*160, MICA*163, MICA*164, MICA*166, MICA*167, MICA*169, MICA*171Q, MICA*173, MICA*182, MICA*186, MICA*188, MICA*199, MICA*205, MICA*208, MICA*219, MICA*221, MICA*223, MICA*241, MICA*254, MICA*255, MICA*259, MICA*263, MICA*265, MICA*267, MICA*269, MICA*275, MICA*287, MICA*289, MICA*290 |
| A7   | MICA*050, MICA*210                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| A8   | MICA*055, MICA*091, MICA*092, MICA*211                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| A9   | MICA*002, MICA*015, MICA*017, MICA*020, MICA*041, MICA*046, MICA*052, MICA*068, MICA*072, MICA*075, MICA*078, MICA*081, MICA*083, MICA*084, MICA*086, MICA*090, MICA*093, MICA*098, MICA*106, MICA*110, MICA*111, MICA*125, MICA*129, MICA*131, MICA*132, MICA*133, MICA*139, MICA*143, MICA*158, MICA*161, MICA*178, MICA*179, MICA*183, MICA*184, MICA*187, MICA*189,                                                                                                                                                                                                                                                                                                                         |

|  |                                                                                                                                                                                               |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | MICA*195N, MICA*200, MICA*207, MICA*214, MICA*215, MICA*220, MICA*226, MICA*227, MICA*229, MICA*230, MICA*233, MICA*234, MICA*240, MICA*246, MICA*256, MICA*257, MICA*260, MICA*276, MICA*288 |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**Fonte:** Material elaborado pelo autor utilizando os dados de sequência de nucleotídeos fornecidos pelo IPD-HLA/IMGT version 3.51 - 2023-1 <sup>24</sup>.

Outro polimorfismo com impacto funcional é a troca de Metionina por Valina na posição 129 da proteína MICA, no domínio  $\alpha 2$ . Esse dimorfismo acontece devido ao SNP rs1051792 G/A no éxon 3 <sup>80</sup>. MICA-129 afeta a ligação de NKG2D. Essa alteração está possivelmente relacionada à uma alteração conformacional, pois esta posição não está diretamente envolvida na ligação com o receptor <sup>15,81</sup>. MICA-129 met ligado ao NKG2D presente nas células NK apresenta sinalização mais forte e rápida, resultando em uma maior citotoxicidade de células NK e liberação aumentada de interferon- $\gamma$  <sup>81</sup>. Nas células TCD8+, esta variante levou a uma co-estimulação mais rápida. Porém, em ambos os tipos de células, MICA-129 met induziu rapidamente a regulação negativa de NKG2D <sup>81</sup>.

*MICA* também pode ter variação no número de cópias. Alguns indivíduos apresentam uma deleção de 100kb que acontece entre os genes *HLA-B* e *MICB*, deletando todo o gene *MICA*. Essa deleção é chamada de *MICA\*del* <sup>82,83</sup>. A deleção de *MICA* está em desequilíbrio de ligação (LD) com o alelo *HLA-B\*048* e com *MICB\*009N* <sup>84</sup>, um alelo nulo de *MICB*. Assim, alguns indivíduos podem carregar um cromossomo 6 que não é capaz de produzir proteínas MIC funcionais.

O gene *MICB* apresenta até o momento 244 alelos, gerando 44 proteínas diferentes (IPD-HLA/IMGT version 3.51 - 2023-1) <sup>24</sup>. *MICB* é menos polimórfico do que *MICA*, porém, existem menos trabalhos com foco em *MICB* quando comparado a *MICA* e os poucos estudos analisam *MICB* em pequenas amostras e em populações específicas. Dessa forma, o menor grau de polimorfismo de *MICB* pode ser um viés relacionado aos poucos estudos realizados. *MICA* e *MICB* apresentam 91% de homologia em suas sequências. Porém, apresentam diferenças em seus polimorfismos e diferenças funcionais <sup>85</sup>, além de *MICB*

apresentar o íntron 1 e uma 3' não traduzida muito maior quando comparado com *MICA* <sup>86</sup>.

*MICB* apresenta uma região de microssatélite no íntron 1 conhecida como C1\_2\_A <sup>87</sup>. Esse microssatélite pode variar de 14-28 repetições de dinucleotídeos (CA/TG), cujos alelos são denominados CA14-CA28 <sup>88</sup>. Apesar de ter sido identificado a muito tempo, poucos estudos exploraram este polimorfismo em associação com doenças <sup>88,89</sup> e o impacto desse microssatélite na função do gene ainda não foi discutido.

Diferente do gene *MICA*, *MICB* apresenta um dimorfismo de inserção ou deleção de elementos ALU no intron 1. A presença ou ausência de elementos ALU pode ser estudada na genética comparativa de humanos e primatas. Outra diferença entre os dois genes é que *MICB* se liga à proteína UL16, produzida durante a infecção pelo citomegalovírus humano como mecanismo de escape imunológico <sup>85</sup>. *MICB* não apresenta a região de microssatélite no éxon 5 que codifica Alaninas, assim como acontece em *MICA*.

O alelo *MICB*\*009N é um dos alelos com maior impacto funcional. Este alelo contém um stop códon prematuro no éxon 3, região que codifica o domínio  $\alpha 2$ , levando ao truncamento da molécula na posição 170 <sup>29,90</sup>. Esse alelo está em desequilíbrio de ligação com a deleção do gene *MICA*. Desta forma, alguns cromossomos 6 carregam tanto *MICA*\*del quanto *MICB*\*009N e não codificam uma proteína MIC funcional.

### **MICA/B em infecções, câncer e no transplante**

Devido a função das moléculas de MICA/B em ativar as células NK e levar a morte de células estressadas, estas moléculas acabam sendo alvo de mecanismos de escape imunológico.

Os polimorfismos de *MICA* e *MICB* podem estar associados a susceptibilidade ou proteção a determinadas doenças. Já é bem estabelecido o papel das moléculas MICA/B no escape imunológico do citomegalovírus humano (HCMV). Na infecção por esse vírus, as proteínas virais levam à expressão de MICA/B. No entanto, o HCMV inibe a função de MIC por diferentes mecanismos, que incluem o sequestro de ligante *MICB* pela proteína UL16 <sup>91</sup> e/ou retenção de *MICA* pela proteína UL142 dentro do cis-Golgi. Porém vale ressaltar que *MICA*\*008 (que é um alelo *MICA*\*5.1 e, portanto, produz uma versão truncada

da molécula de MICA) não é afetada por esse mecanismo de escape <sup>92</sup>. Além disso, o mRNA de *MICB* pode ser alvo de miRNA que impedem sua tradução <sup>93</sup>. A proteína truncada de MICA era considerada não afetada pelo mecanismo de escape deste vírus, porém recentemente Seidel e colaboradores mostraram um mecanismo de escape diferente do HCMV que impede a associação de MICA truncada com a âncora GPI, leva o desvio de MICA para o citosol, onde acaba sendo degradada no proteossoma e assim o HCMV prejudicando a citotoxicidade das células NK <sup>94</sup>.

As proteínas de MICA e MICB também já foram associadas à COVID-19. A maior expressão de MICB pode influenciar positivamente no controle da infecção, ativando as células NK precocemente. Enquanto os indivíduos com maior produção de mRNA de MICA, possivelmente da forma solúvel e a baixa expressão de MICB, foram associadas à suscetibilidade à doença <sup>95</sup>.

Diversos estudos já abordaram o mecanismo pelo qual as células tumorais escapam da citotoxicidade mediada por células NKs. As células tumorais liberam os ligantes de estresse (MICA e MICB) da superfície, assim, as proteínas solúveis levam à regulação negativa do ligante NKG2D, escapando da citotoxicidade mediada pelas células NKs <sup>96–99</sup>.

O polimorfismo de *MICA* e *MICB* pode levar a incompatibilidade entre o doador e receptor e rejeição do tecido transplantado <sup>26</sup>. Além disso, essas moléculas são sinalizadoras de estresse celular e podem ser reguladas positivamente devido a fatores clínicos relacionados ao doador e ao receptor, favorecendo a rejeição do enxerto. Além disso, a incompatibilidade entre as moléculas do doador e do receptor pode levar à formação de anticorpos anti-MICA e anti-MCB <sup>26,100,101</sup>.

## **JUSTIFICATIVA**

*MICA* é altamente polimórfico e codifica uma glicoproteína sinalizadora de estresse celular e que interage com o receptor ativador NKG2D, ativando as células NK e levando a morte das células-alvo. A expressão de *MICA* é estimulada em situações como o câncer, infecções virais e pode estar relacionada com a rejeição precoce a enxertos. Poucos trabalhos abordam a diversidade desses genes em populações verdadeiras, principalmente em populações miscigenadas como a brasileira. Portanto, a caracterização da variabilidade de *MICA* poderia favorecer uma melhor compreensão sobre as diferenças genótípicas a nível mundial, na busca de diagnósticos e tratamentos de doenças.

## **OBJETIVOS**

### **Objetivo Geral**

Analisar a diversidade genética do gene *MICA* em populações mundiais, incluindo amostras brasileiras.

### **Objetivos Específicos**

- Desenvolver um método computacional para inferir adequadamente os genótipos de *MICA*;
- Determinar as frequências alélicas de *MICA* observadas;
- Determinar a frequência de *MICA*\*del de acordo com as regiões biogeográficas;
- Avaliar o desequilíbrio de ligação entre os genes *MICA* e *MICB* e entre *MICA* e *HLA-B*.

## REFERÊNCIAS

1. Ellis SA, Ballingall KT. Cattle MHC: evolution in action? *Immunol Rev* 1999;167(1):159–168.
2. Cruvinel W de M, Mesquita Júnior D, Araújo JAP, et al. Sistema imunitário: Parte I. Fundamentos da imunidade inata com ênfase nos mecanismos moleculares e celulares da resposta inflamatória. *Rev Bras Reumatol* 2010;50:434–447.
3. Nepom GT, Erlich H. MHC Class-II Molecules and Autoimmunity. *Annu Rev Immunol* 1991;9(1):493–525.
4. Radosavljevic M, Bahram S. In vivo immunogenetics: from MIC to RAET1 loci. *Immunogenetics* 2003;55(1):1–9.
5. BAHRAM S. MIC genes: from genetics to biology. *MIC Genes Genet Biol* 2000;76:1–60.
6. Bahram S et al. A second lineage of mammalian major histocompatibility complex class I genes. *Proceedings of the National Academy of Sciences* 1994; 91(14): 6259-6263.
7. Leelayuwat C, Degli-Esposti MA, Abraham LJ, Dawkins RL, Townsend DC. A new polymorphic and multicopy MHC gene family related to nonmammalian class I. *Immunogenetics* 1994;40(5):339–351.
8. Risti M, Bicalho M da G. MICA and NKG2D: Is There an Impact on Kidney Transplant Outcome? *Front Immunol* 2017;8:179.
9. Zwirner NW, Fernández-Viña MA, Stastny P. MICA, a new polymorphic HLA-related antigen, is expressed mainly by keratinocytes, endothelial cells, and monocytes. *Immunogenetics* 1997;47(2):139–148.
10. Zwirner NW, Dole K, Stastny P. Differential surface expression of MICA by endothelial cells, fibroblasts, keratinocytes, and monocytes. *Hum Immunol* 1999;60(4):323–330.
11. Groh V, Steinle A, Bauer S, Spies T. Recognition of stress-induced MHC molecules by intestinal epithelial  $\gamma\delta$  T cells. *Science* 1998; 279(5357): 1737-1740.
12. DING W et al. MICA\* 012: 01 Allele Facilitates the Metastasis of KRAS-Mutant Colorectal Cancer. *Frontiers in Genetics* 2020; 11: 51.
13. Raulet DH, Gasser S, Gowen BG, Deng W, Jung H. Regulation of Ligands for the NKG2D Activating Receptor. *Annu Rev Immunol* 2013;31(1):413–441.
14. Castelli EC, Castro MV de, Naslavsky MS, et al. MHC Variants Associated With Symptomatic Versus Asymptomatic SARS-CoV-2 Infection in Highly Exposed Individuals. *Front Immunol* 2021;12:742881.
15. Li P, Morris DL, Willcox BE, Steinle A, Spies T, Strong RK. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nat Immunol* 2001;2(5):443–451.
16. Chitadze G, Bhat J, Lettau M, Janssen O, Kabelitz D. Generation of Soluble NKG2D Ligands: Proteolytic Cleavage, Exosome Secretion and Functional Implications. *Scand J Immunol* 2013;78(2):120–129.

17. García G, Puerto F del, Pérez AB, et al. Association of MICA and MICB alleles with symptomatic dengue infection. *Hum Immunol* 2011;72(10):904–907.
18. Luangtrakool P, Vejbaesya S, Luangtrakool K, et al. Major Histocompatibility Complex Class I Chain–Related A and B (MICA and MICB) Gene, Allele, and Haplotype Associations With Dengue Infections in Ethnic Thais. *J Infect Dis* 2020;222(5):840–846.
19. Wang Y, Li S, Chen C, et al. MICB\*002 and MICB\*014 protect against rheumatoid arthritis, whereas MICA\*009 and MICA\*A6 are associated with rheumatoid arthritis in a Hainan Han Chinese population. *Int J Rheum Dis* 2019;22(1):90–95.
20. CASTELLI EC et al. Immunogenetics of resistance to SARS-CoV-2 infection in discordant couples. *Medrxiv* 2021;21255872.
21. GAO X et al. Diversity of MICA and linkage disequilibrium with HLA-B in two North American populations. *Human immunology* 2006; 67(3):152-158.
22. FODIL, Nassima et al. MICA haplotypic diversity. *Immunogenetics* 1999; 49: 557-560.
23. Jarduli LR, Alves HV, Souza VH, et al. Association of *MICA* and HLA-B alleles with leprosy in two endemic populations in Brazil. *Int J Immunogenet* 2021;48(1):25–35.
24. Dominic JB et al. The IPD-IMGT/HLA Database, *Nucleic Acids Research* 2023; 51: Issue D1, D1053–D1060.
25. Hankey KG, Drachenberg CB, Papadimitriou JC, et al. MIC EXPRESSION IN RENAL AND PANCREATIC ALLOGRAFTS. *Transplantation* 2002;73(2):304–306.
26. Zwirner NW, Marcos CY, Mirbaha F, Zou Y, Stastny P. Identification of MICA as a new polymorphic alloantigen recognized by antibodies in sera of organ transplant recipients. *Hum Immunol* 2000;61(9):917–924.
27. Li P, Willie ST, Bauer S, Morris DL, Spies T, Strong RK. Crystal Structure of the MHC Class I Homolog MIC-A, a  $\gamma\delta$  T Cell Ligand. *Immunity* 1999;10(5):577–584.
28. Groh V, Bahram S, Bauer S, Herman A, Beauchamp M, Spies T. Cell stress-regulated human major histocompatibility complex class I gene expressed in gastrointestinal epithelium. *Proc Natl Acad Sci* 1996;93(22):12445–12450.
29. Ando H, Mizuki N, Ota M, et al. Allelic variants of the human MHC class I chain-related B gene (MICB). *Immunogenetics* 1997;46(6):499–508.
30. Crosby CM, Kronenberg M. Tissue-specific functions of invariant natural killer T cells. *Nat Rev Immunol* 2018;18(9):559–574.
31. VIVIER E et al. Innate or adaptive immunity? The example of natural killer cells. *science* 2011; 331(6013): 44-49.
32. Cooper MA, Fehniger TA, Caligiuri MA. The biology of human natural killer-cell subsets. *Trends Immunol* 2001;22(11):633–640.
33. Li P, Morris DL, Willcox BE, Steinle A, Spies T, Strong RK. Complex structure of the activating immunoreceptor NKG2D and its MHC class I–like ligand MICA. *Nat Immunol* 2001;2(5):443–451.
34. Park YP, Choi S-C, Kiesler P, et al. Complex regulation of human NKG2D-DAP10 cell surface expression: opposing roles of the  $\gamma\delta$  cytokines and TGF- $\beta$ 1. *Blood* 2011;118(11):3019–3027.
35. Roberts A I et al. Cutting edge: NKG2D receptors induced by IL-15

- costimulate CD28-negative effector CTL in the tissue microenvironment. *The Journal of Immunology* 2001; 167(10): 5527-5530.
36. Fernández-Sánchez A, Baragaño Raneros A, Carvajal Palao R, et al. DNA demethylation and histone H3K9 acetylation determine the active transcription of the NKG2D gene in human CD8<sup>+</sup> T and NK cells. *Epigenetics* 2013;8(1):66–78.
  37. Muntasell A, Magri G, Pende D, Angulo A, López-Botet M. Inhibition of NKG2D expression in NK cells by cytokines secreted in response to human cytomegalovirus infection. *Blood* 2010;115(25):5170–5179.
  38. CRANE C A et al. TGF- $\beta$  downregulates the activating receptor NKG2D on NK cells and CD8<sup>+</sup> T cells in glioma patients. *Neuro-oncology* 2010; 12(1): 7-13.
  39. Burgess SJ, Marusina AI, Pathmanathan I, Borrego F, Coligan JE. IL-21 Down-Regulates NKG2D/DAP10 Expression on Human NK and CD8<sup>+</sup> T Cells1. *J Immunol* 2006;176(3):1490–1497.
  40. Wu J et al. An activating immunoreceptor complex formed by NKG2D and DAP10. *Science* 1999; 285(5428): 730-732.
  41. Garrity D et al. The activating NKG2D receptor assembles in the membrane with two signaling dimers into a hexameric structure. *Proceedings of the National Academy of Sciences* 2005; 102(21): 7641-7646.
  42. Billadeau DD, Upshaw JL, Schoon RA, Dick CJ, Leibson PJ. NKG2D-DAP10 triggers human NK cell-mediated killing via a Syk-independent regulatory pathway. *Nat Immunol* 2003;4(6):557–564.
  43. Chitadze G, Bhat J, Lettau M, Janssen O, Kabelitz D. Generation of Soluble NKG2D Ligands: Proteolytic Cleavage, Exosome Secretion and Functional Implications. *Scand J Immunol* 2013;78(2):120–129.
  44. Molfetta R, Quatrini L, Santoni A, Paolini R. Regulation of NKG2D-Dependent NK Cell Functions: The Yin and the Yang of Receptor Endocytosis. *Int J Mol Sci* 2017;18(8):1677.
  45. Mizuki N, Ota M, Kimura M, et al. Triplet repeat polymorphism in the transmembrane region of the MICA gene: A strong association of six GCT repetitions with Behcet disease. *Proc Natl Acad Sci* 1997;94(4):1298–1303.
  46. Zou Y, Han M, Wang Z, Stastny P. MICA Allele-Level Typing by Sequence-based Typing with Computerized Assignment of Polymorphic Sites and Short Tandem Repeats within the Transmembrane Region. *Hum Immunol* 2006;67(3):145–151.
  47. Ashiru O et al. A GPI anchor explains the unique biological features of the common NKG2D-ligand allele MICA\* 008. *Biochemical Journal* 2013; 454(2): 295-302.
  48. Waldhauer I, Goehlsdorf D, Gieseke F, et al. Tumor-Associated MICA Is Shed by ADAM Proteases. *Cancer Res* 2008;68(15):6368–6376.
  49. Matusali G, Tchidjou HK, Pontrelli G, et al. Soluble ligands for the NKG2D receptor are released during HIV-1 infection and impair NKG2D expression and cytotoxicity of NK cells. *FASEB J* 2013;27(6):2440–2450.
  50. Blobel C P ADAMs: key components in EGFR signalling and development. *Nature reviews Molecular cell biology* 2005; 6(1) 32-43.
  51. Reiss K, Ludwig A, Saftig P. Breaking up the tie: Disintegrin-like metalloproteinases as regulators of cell migration in inflammation and invasion. *Pharmacol Ther* 2006;111(3):985–1006.
  52. White JM. ADAMs: modulators of cell–cell and cell–matrix interactions. *Curr*

- Opin Cell Biol 2003;15(5):598–606.
53. Waldhauer I, Goehlsdorf D, Gieseke F, et al. Tumor-Associated MICA Is Shed by ADAM Proteases. *Cancer Res* 2008;68(15):6368–6376.
  54. Kaiser BK, Yim D, Chow I-T, et al. Disulphide-isomerase-enabled shedding of tumour-associated NKG2D ligands. *Nature* 2007;447(7143):482–486.
  55. Wang X, Lundgren AD, Singh P, Goodlett DR, Plymate SR, Wu JD. An six-amino acid motif in the  $\alpha 3$  domain of MICA is the cancer therapeutic target to inhibit shedding. *Biochem Biophys Res Commun* 2009;387(3):476–481.
  56. Xing S, Ferrari A L NKG2D and MICA/B shedding: A ‘tag game’ between NK cells and malignant cells. *Clinical & translational immunology* 2020; 9(12):1230.
  57. Roda-Navarro P, Reyburn HT. The Traffic of the NKG2D/Dap10 Receptor Complex during Natural Killer (NK) Cell Activation\*. *J Biol Chem* 2009;284(24):16463–16472.
  58. Molfetta R, Quatrini L, Capuano C, et al. c-Cbl regulates MICA- but not ULBP2-induced NKG2D down-modulation in human NK cells. *Eur J Immunol* 2014;44(9):2761–2770.
  59. Molfetta R, Quatrini L, Zitti B, et al. Regulation of NKG2D Expression and Signaling by Endocytosis. *Trends Immunol* 2016;37(11):790–802.
  60. Gasser S, Raulet DH. Activation and self-tolerance of natural killer cells. *Immunol Rev* 2006;214(1):130–142.
  61. Eissmann P, Evans JH, Mehrabi M, Rose EL, Nedvetzki S, Davis DM. Multiple Mechanisms Downstream of TLR-4 Stimulation Allow Expression of NKG2D Ligands To Facilitate Macrophage/NK Cell Crosstalk. *J Immunol* 2010;184(12):6901–6909.
  62. Rodríguez-Rodero S, González S, Rodrigo L, et al. Transcriptional regulation of MICA and MICB: A novel polymorphism in MICB promoter alters transcriptional regulation by Sp1. *Eur J Immunol* 2007;37(7):1938–1953.
  63. Venkataraman GM, Suci D, Groh V, Boss JM, Spies T. Promoter Region Architecture and Transcriptional Regulation of the Genes for the MHC Class I-Related Chain A and B Ligands of NKG2D1. *J Immunol* 2007;178(2):961–969.
  64. Weitzman MD, Carson CT, Schwartz RA, Lilley CE. Interactions of viruses with the cellular DNA repair machinery. *DNA Repair* 2004;3(8):1165–1173.
  65. Gasser S, Orsulic S, Brown EJ, Raulet DH. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. *Nature* 2005;436(7054):1186–1190.
  66. Eissmann P, Evans JH, Mehrabi M, Rose EL, Nedvetzki S, Davis DM. Multiple Mechanisms Downstream of TLR-4 Stimulation Allow Expression of NKG2D Ligands To Facilitate Macrophage/NK Cell Crosstalk. *J Immunol* 2010;184(12):6901–6909.
  67. Tang K-F, Ren H, Cao J, et al. Decreased Dicer expression elicits DNA damage and up-regulation of MICA and MICB. *J Cell Biol* 2008;182(2):233–239.
  68. Cerboni C, Zingoni A, Cippitelli M, Piccoli M, Frati L, Santoni A. Antigen-activated human T lymphocytes express cell-surface NKG2D ligands via an ATM/ATR-dependent mechanism and become susceptible to autologous NK- cell lysis. *Blood* 2007;110(2):606–615.
  69. O’Neill LAJ. The interleukin-1 receptor/Toll-like receptor superfamily: 10

- years of progress. *Immunol Rev* 2008;226(1):10–18.
70. Rittirsch D, Flierl MA, Ward PA Harmful molecular mechanisms in sepsis. *Nature Reviews Immunology* 2008; 8(10): 776-787.
  71. Medzhitov R. Origin and physiological roles of inflammation. *Nature* 2008;454(7203):428–435.
  72. Stern-Ginossar N, Gur C, Biton M, et al. Human microRNAs regulate stress-induced immune responses mediated by the receptor NKG2D. *Nat Immunol* 2008;9(9):1065–1073.
  73. Suemizu H, Radosavljevic M, Kimura M, et al. A basolateral sorting motif in the MICA cytoplasmic tail. *Proc Natl Acad Sci* 2002;99(5):2971–2976.
  74. Steinle A, Li P, Morris DL, et al. Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics* 2001;53(4):279–287.
  75. Frigoul, A., Lefranc, M. P. MICA: standardized IMGT allele nomenclature, polymorphisms and diseases, Recent Research Developments in Human Genetics. Kerala: Research signpost, 2005; 3: 95-145.
  76. Ashiru O, Boutet P, Fernández-Messina L, et al. Natural Killer Cell Cytotoxicity Is Suppressed by Exposure to the Human NKG2D Ligand MICA\*008 That Is Shed by Tumor Cells in Exosomes. *Cancer Res* 2010;70(2):481–489.
  77. Zhang J, Liao D, Yang L, Hou S. Association between Functional MICA-TM and Behcet's Disease: A Systematic Review and Meta-analysis. *Sci Rep* 2016;6(1):21033.
  78. Gutiérrez-Bautista JF, Martínez-Chamorro A, Rodríguez-Nicolas A, et al. Major Histocompatibility Complex Class I Chain-Related  $\alpha$  (MICA) STR Polymorphisms in COVID-19 Patients. *Int J Mol Sci* 2022;23(13):6979.
  79. Steinle A et al. Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics* 2001; 53: 279-287.
  80. Isernhagen A, Malzahn D, Viktorova E, et al. The MICA-129 dimorphism affects NKG2D signaling and outcome of hematopoietic stem cell transplantation. *EMBO Mol Med* 2015;7(11):1480–1502.
  81. Komatsu-Wakui M, Tokunaga K, Ishikawa Y, et al. Wide distribution of the *MICA-MICB null* haplotype in East Asians: The *MICA-MICB null* haplotype in East Asians. *Tissue Antigens* 2001;57(1):1–8.
  82. Komatsu-Wakui M, Tokunaga K, Ishikawa Y, et al. MIC-A polymorphism in Japanese and a MIC-A-MIC-B null haplotype. *Immunogenetics* 1999;49(7):620–628.
  83. Ota M, Bahram S, Katsuyama Y, et al. On the MICA deleted-MICB null, HLA-B\*4801 haplotype. *Tissue Antigens* 2000;56(3):268–271.
  84. Cosman D, Müllberg J, Sutherland CL, et al. ULBPs, Novel MHC Class I–Related Molecules, Bind to CMV Glycoprotein UL16 and Stimulate NK Cytotoxicity through the NKG2D Receptor. *Immunity* 2001;14(2):123–133.
  85. Wang WY, Tian W, Zhu F m., Liu XX, Li LX, Wang F. MICA, MICB Polymorphisms and Linkage Disequilibrium with HLA-B in a Chinese Mongolian Population. *Scand J Immunol* 2016;83(6):456–462.
  86. Tamiya G, Ota M, Katsuyama Y, et al. Twenty-six new polymorphic microsatellite markers around the HLA-B, -C and -E loci in the human MHC class I region. *Tissue Antigens* 2008;51(4):337–346.
  87. Kimura T, Goto K, Yabuki K, et al. Microsatellite polymorphism within the

- MICB gene among Japanese patients with Behçet's disease. *Hum Immunol* 1998;59(8):500–502.
88. Lü M, Xia B, Li J, Ye M, Zhang X, Tan Q. MICB microsatellite polymorphism is associated with ulcerative colitis in Chinese population. *Clin Immunol* 2006;120(2):199–204.
  89. Stephens HAF. MICA and MICB genes: can the enigma of their polymorphism be resolved? *Trends Immunol* 2001;22(7):378–385.
  90. Wu J, Chalupny NJ, Manley TJ, Riddell SR, Cosman D, Spies T. Intracellular Retention of the MHC Class I-Related Chain B Ligand of NKG2D by the Human Cytomegalovirus UL16 Glycoprotein1. *J Immunol* 2003;170(8):4196–4200.
  91. Ashiru O, Bennett NJ, Boyle LH, Thomas M, Trowsdale J, Wills MR. NKG2D Ligand MICA Is Retained in the cis-Golgi Apparatus by Human Cytomegalovirus Protein UL142. *J Virol* 2009;83(23):12345–12354.
  92. Stern-Ginossar N, Elefant N, Zimmermann A, et al. Host Immune System Gene Targeting by a Viral miRNA. *Science* 2007;317(5836):376–381.
  93. Seidel E, Dassa L, Schuler C, et al. The human cytomegalovirus protein UL147A downregulates the most prevalent MICA allele: MICA\*008, to evade NK cell-mediated killing. *PLOS Pathog* 2021;17(5):e1008807.
  94. Castelli EC, Castro MV de, Naslavsky MS, et al. MHC Variants Associated With Symptomatic Versus Asymptomatic SARS-CoV-2 Infection in Highly Exposed Individuals. *Front Immunol* 2021;12:742881.
  95. Balsamo M, Vermi W, Parodi M, et al. Melanoma cells become resistant to NK-cell-mediated killing when exposed to NK-cell numbers compatible with NK-cell infiltration in the tumor. *Eur J Immunol* 2012;42(7):1833–1842.
  96. Siemens DR, Hu N, Sheikh AK, et al. Hypoxia Increases Tumor Cell Shedding of MHC Class I Chain-Related Molecule: Role of Nitric Oxide. *Cancer Res* 2008;68(12):4746–4753.
  97. Ivanova M, Al Hadra B, Yordanov S, et al. Associations of high-resolution-typing-defined MICA and MICB polymorphisms, and the levels of soluble MICA and MICB with Oral Squamous Cell Carcinoma in Bulgarian patients. *J Oral Pathol Med* 2021;50(8):758–765.
  98. Holdenrieder S, Stieber P, Peterfi A, Nagel D, Steinle A, Salih HR. Soluble MICB in malignant diseases: analysis of diagnostic significance and correlation with soluble MICA. *Cancer Immunol Immunother* 2006;55(12):1584–1589.
  99. Zou Y, Mirbaha F, Lazaro A, Zhang Y, Lavingia B, Stastny P. MICA is a target for complement-dependent cytotoxicity with mouse monoclonal antibodies and human alloantibodies. *Hum Immunol* 2002;63(1):30–39.
  100. Quiroga I, Salio M, Koo DDH, et al. Expression of MHC Class I-Related Chain B (MICB) Molecules on Renal Transplant Biopsies. *Transplantation* 2006;81(8):1196.

## **Capítulo II – Artigo**

## **The MICA deletion across different populations**

### **Authors**

Viviane Aparecida de Oliveira Ciriaco<sup>1,2</sup>, Amanda Muniz Rodrigues<sup>1,3</sup>, Brenda Caroline da Silva Tibúrcio<sup>1,4</sup>, Joyce Machado Silva<sup>1,5</sup>, Camila Ferreira Bannwart Castro<sup>1,4,5</sup>, Erick C. Castelli<sup>1,2</sup>

### **Filiations**

<sup>1</sup>São Paulo State University, Molecular Genetics and Bioinformatics Laboratory, School of Medicine, Botucatu, State of São Paulo, Brazil. <sup>2</sup>Pathology Program, School of Medicine, Unesp, Botucatu, SP, Brazil. <sup>3</sup>Bioscience Institute, Unesp, Botucatu, SP, Brazil. <sup>4</sup>Centro Universitário Sudoeste Paulista, Avaré, SP, Brazil. <sup>5</sup>General and Applied Biology Program, Bioscience Institute, Unesp, Botucatu, SP, Brazil.

### **Contact**

Erick da Cruz Castelli  
Departamento de Patologia, Faculdade de Medicina de Botucatu, Unesp –  
Botucatu, SP  
CEP: 18618970, Brazil  
Phone: +55 14 3880-1696  
E-mail address: [erick.castelli@unesp.br](mailto:erick.castelli@unesp.br)

## ABSTRACT

The complete deletion of major histocompatibility complex class I-chain related A (MICA), known as MICA\*del, may be an important point of weakness for the immune response, since this gene encodes a NKG2D receptor-ligand, highly polymorphic and responsible for activating an immune response during intracellular infections, neoplastic cells development and after allograft transplantation. However, data on the deletion frequencies in different populations are limited. To evaluate the distribution of the MICA\*del in different biogeographic regions, including a large sample of admixed individuals from Brazil we use a specific bioinformatic pipeline and hla-mapper software to correct the alignments, we addressed MICA\*del and others *MICA* and *MICB* null alleles in eight different populations of biogeographical regions. Different deletion patterns were found MICA deletion complete, called MICA\*del and the partial MICA deletion, which involves the promoter region, exon 1 and intron 1, called MICA\*partial. MICA\*del was detected in all biogeographical regions, except Oceania. It's more frequent in the Asian population. Interestingly, America was the second region where MICA\*del were most frequent, specially in samples from Mexican, Peruvians and Native Americans individuals from Brazil (Surui). MICA\*partial was detected only in America and the Middle East, with low frequency. Most alleles with MICA deletion were in LD with B\*48:01 and MICB\*009:01N in Asia and America. Therefore, the presence of different patterns of deletion suggests independent deletion events. Moreover, we found two samples with MICA deletion in both chromosomes and one of these samples did not have functional MICB, which may indicate that another mechanism compensates for the lack of functional MIC.

## Introduction

The major histocompatibility complex class I-chain related (*MIC*) genes are among the last loci identified within the human Major Histocompatibility Complex (MHC) in the class I region. They include two functional genes, *MICA* and *MICB*, and five pseudogenes, *MICC*, *MICD*, *MICE*, *MICF*, and *MICG*<sup>5,10,13,102</sup>. *MICA* is 46kb upstream of the human leukocyte antigen (*HLA*)-*B* gene, with strong Linkage Disequilibrium (LD) among them<sup>103</sup>. *MICA* presents six exons, large introns, and a 3' untranslated region (UTR) longer than most class I genes<sup>29,102</sup>.

*MICA* encodes a glycoprotein that is constitutively expressed in epithelial cells of the gastrointestinal tract, endothelial cells, and monocytes. It is upregulated upon cellular stress, particularly in oxidative stress<sup>104</sup>, intracellular infections<sup>17,18,20,105</sup>, and tumorigenesis<sup>106</sup>. This stress-signaling molecule interacts with the activating receptor NKG2D from Natural Killer (NK), Tαβ and Tγδ CD8+ lymphocytes<sup>10,13,102</sup>.

*MICA* can be expressed in two isoforms: bound to the membrane surface (m*MICA*) or released as soluble molecules (s*MICA*). Membrane-bound isoforms activate a cytotoxic function of immune cells after being bound to the receptor. However, the soluble isoforms may lead to NKG2D downregulation, preventing immune cell cytotoxic activity<sup>107</sup>. Two mechanisms can produce the soluble isoform: due to polymorphisms that prevent the translation of the transmembrane domain<sup>47,108</sup>, or shed by metalloproteinases such as ADAM10 and ADAM17<sup>109–111</sup>. The production of s*MICA* configures an efficient mechanism of target cells to escape NK cell-mediated cytotoxicity in innate immunity<sup>9,13,15,16</sup>.

Moreover, *MICA* might be absent from chromosome 6 due to a large deletion of approximately 100Kb between *HLA-B* and *MICB*<sup>83,112</sup>; therefore, some individuals may not produce *MICA* at all. This deletion is known as the *MICA*\*deletion or *MICA*\*del. *MICA*\*del has been associated with susceptibility to nasopharyngeal carcinoma in Taiwan<sup>113</sup> but not in Southern China<sup>114</sup>.

Previous studies reported that *MICA*\*del exhibits an LD with *HLA-B*\*48 and *MICB*\*009N, a null allele for the *MICB*<sup>82,84</sup>. *MICB* is encoded close to *MICA* and produces a molecule with similar functions. The linkage disequilibrium between *MICA*\*del and *MICB*\*009N suggests that these individuals do not produce functional *MIC* proteins from this chromosome<sup>82</sup>. The distribution of this phenotype may vary among different populations.

Here, we evaluate the distribution of the *MICA*\*del and other *MICA* null alleles in different biogeographic regions and the linkage between this allele and *HLA-B* and *MICB*. This data may contribute to a better understanding of *MICA* function and how these alleles distribute globally and influence disease susceptibility.

## Methods

### Samples included in this study, sequencing strategy, and read alignment

We evaluated the *MICA\*del* and other null alleles from *MICA* in 5,347 samples from different biogeographic regions: 3,197 individuals from the 1000genomes project <sup>115</sup>; 828 individuals from the Human Genome Diversity Project (HGDP) <sup>116</sup>, and 1,322 Brazilians from the city of São Paulo, Brazil <sup>117</sup>. For all these samples, we downloaded the raw sequence data (Fastq) from high-coverage whole-genome sequencing (WGS), or a BAM file with reads aligned to the human reference genome hg38. For raw Fastq, we aligned all reads to the reference genome using BWA MEM to produce BAM files. We used these BAM files as inputs for hla-mapper version 4.3, which minimizes alignment bias in *HLA* and *MIC* genes ([www.castelli-lab.net/apps/hla-mapper](http://www.castelli-lab.net/apps/hla-mapper)) <sup>118</sup>.

### Calling *MICA* and *MICB* copy numbers

To evaluate the number of copies of *MICA* and *MICB*, we compared the read depth between a reference region in chromosome 6 that does not present copy number variation, genes *HLA-G* and *TNF*, and those observed in *MICA* and *MICB*. To do that, we used samtools depth to calculate read depth across the selected regions, using the BAM files from hla-mapper as input. After, we calculated the ratio between the read depth at the target and the reference. We also tracked heterozygous sites across *MICA* and *MICB*. We considered only heterozygous sites from regions with coverage > 15x and with a balanced proportion of reads supporting both alleles (from 45 to 55%). After, we plotted these ratios (target / reference) using ggplot2 version ggplot2\_3.4.1.

### Calling *MICA*, *MICB*, and *HLA-B* alleles

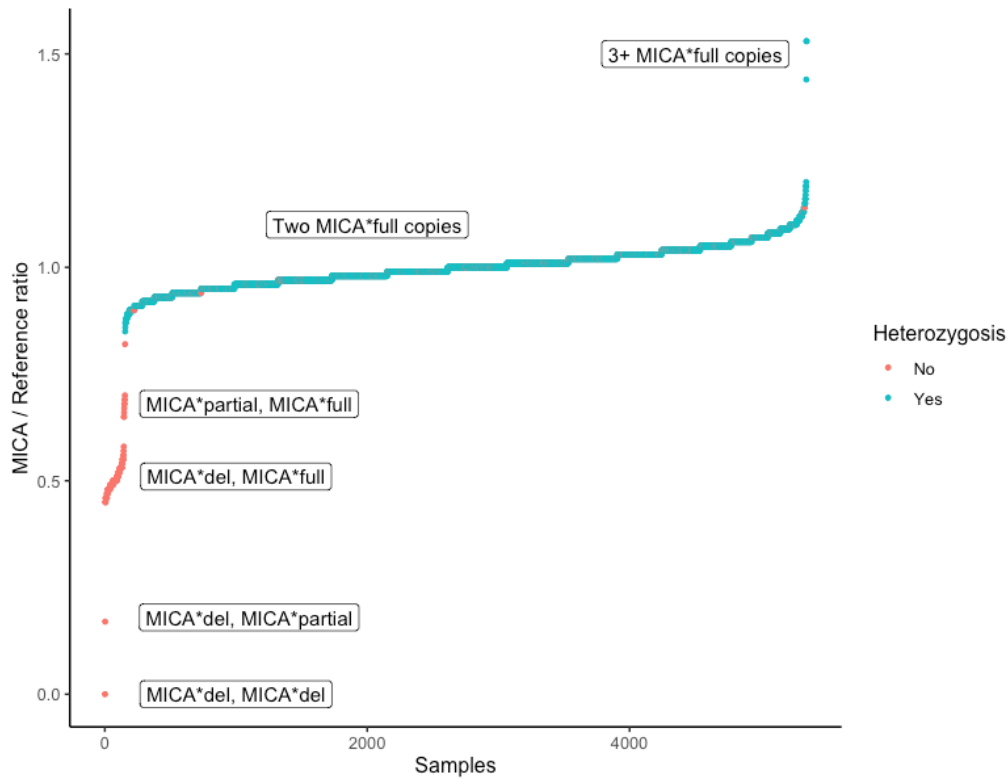
We used the hla-mapper workflow available at [https://github.com/erickcastelli/HLA\\_genotyping](https://github.com/erickcastelli/HLA_genotyping) (version 2) to call SNPs and alleles for *MICA* and *MICB*. The *HLA-B* alleles were already available from another study using the same samples and the same workflow <sup>119</sup>.

To evaluate linkage disequilibrium among *MICA* and other genes at the MHC region by using Arlequin 3.5.

## Results

### ***Detecting MICA\*del***

We plotted the depth ratios between *MICA/B* and the references, as illustrated in Figure 1. The pattern suggests six groups of samples. We also plotted read depth across the *MICA* and *MICB* for one sample from each group (Figure 1S). The first group represents one sample with no reads for *MICA* (sample 5 in Figure 1S), carrying a full deletion of *MICA* in both chromosomes, i.e., homozygous for the *MICA\*del* allele. The second group represents one sample with *MICA\*del* at one chromosome and a partial deletion of *MICA* in the other (sample 4 in Figure S1). In this case, the partial deletion of *MICA* comprehends the promoter region, exon 1, and most intron 1. We called this allele *MICA\*partial*. The third group represents samples with the *MICA\*del* allele in one chromosome and a full copy of *MICA* (*MICA\*full*) in the other (sample 2 in Figure S1), with a depth ratio of around 0.5. The fourth group includes samples with *MICA\*partial* and *MICA\*full* (sample 3 in Figure S1). The fifth group is the most common one and includes samples with two copies of *MICA\*full* (sample 1 in Figure S1), with a depth ratio of around 1. The last group might represent samples with 3 copies of *MICA\*full*, with a depth ratio of around 1.5. We detected heterozygous sites only in samples with two or more copies of *MICA\*full*.



**Figure 1:** *MICA* copy number estimation by comparing read depth in two references, *TNF* and *HLA-G* demonstrates one, two and three or more copies of *MICA* gene. There are samples carrying three possible alleles: the complete *MICA* deletion (*MICA\*del*), a partial *MICA* deletion lacking the promoter, exon 1 and part of intron 1 (*MICA\*partial*), and the presence of full *MICA* copy (*MICA\*full*).

The presence of two chromosomes carrying a *MICA\*del*, *MICA\*partial*, or both is a rare condition. Assuming that *MICA\*partial* is not functional, there were only two samples out of 5,347 with a double deletion of *MICA*. The first is NA19761 from the 1000 Genomes dataset (Mexican Ancestry from Los Angeles, USA), with the genotype *MICA\*del/MICA\*partial*. This individual presents at least one functional copy of *MICB*. The second individual was an older man from Brazil, more than 70 years old, with African (56%), European (28%), and Native American (16%) ancestries. This individual is homozygous for the *MICB* null allele *MICB\*009N*, *HLA-B\*48:02*, and *C\*04:01*.

### **Other *MICA* null alleles**

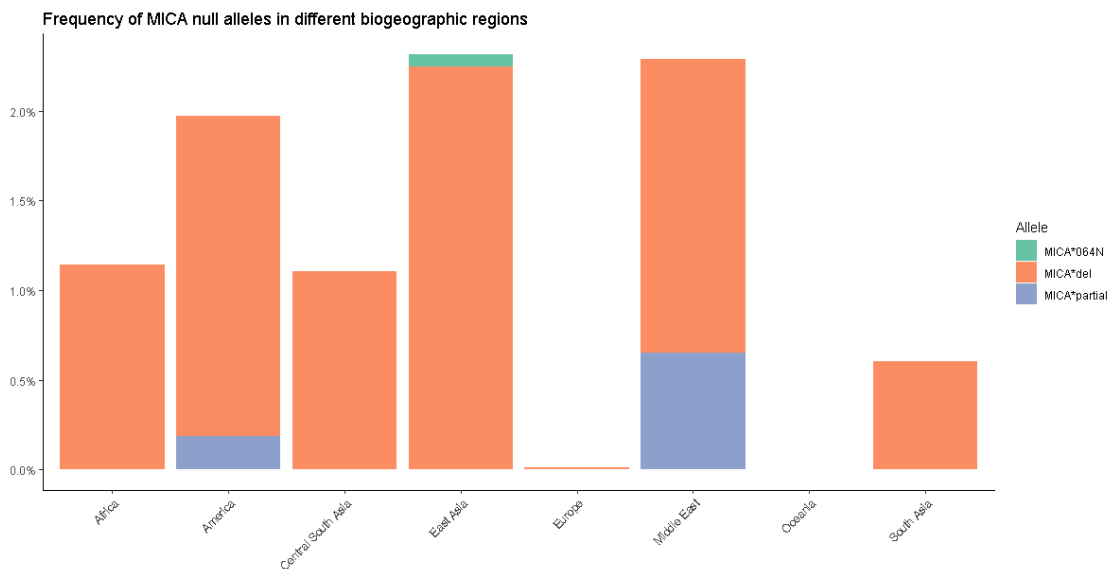
Besides *MICA\*del* and *MICA\*partial*, we detected another null allele called *MICA\*064N*. This allele produces a truncated *MICA* molecule due to a premature stop codon, and it was detected in China, East Asia, sample HGDP01291.

## The frequency of *MICA* null alleles in different biogeographical regions

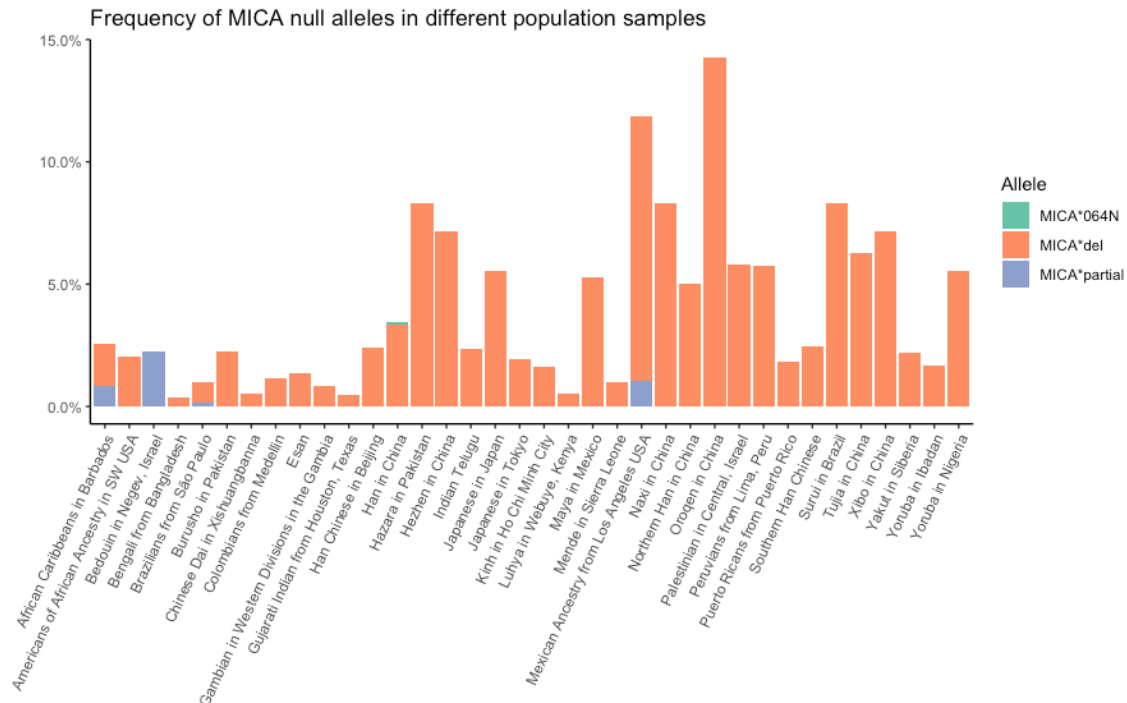
We tracked the frequency of *MICA\*del*, *MICA\*partial*, and *MICA\*064N* across different biogeographic regions (Figure 2).

*MICA\*del* was not detected in Oceania. *MICA\*del* is frequent in most Asian populations (Figure 3). In America, *MICA\*del* is frequent among individuals with Mexican ancestry in Los Angeles, Native Americans from Brazil (Surui), and Peruvians from Peru, but not in São Paulo/Brazil and in Colombia (Figure 3). Apparently, *MICA\*del* is frequent in populations with higher Native American backgrounds, such as Mexico and Peru. However, the Brazilian and Colombian samples included in this study have a major European ancestry<sup>120</sup>, and *MICA\*del* is rare in Europe.

*MICA\*partial* was detected only in America and in the Middle East in low frequencies (Figure 2). In the Middle East, it is the only null allele from Bedoin in Negev (Figure 3). In America, it was detected among Brazilians, Caribbeans, and Americans with Mexican ancestry (Figure 3). Interestingly, the frequency of *MICA* null alleles in East Asia and the Middle East are very similar since the higher frequency of *MICA\*partial* at the Middle East compensates the lower frequency of *MICA\*del* (Figure 2).



**Figure 2:** Frequency of *MICA* null alleles in different biogeographic regions. *MICA\*064N* presents a premature stop-codon with no translation of a functional *MICA*. *MICA\*del* represents the complete deletion of *MICA*. *MICA\*partial* represents a partial deletion in *MICA* encompassing the promoter, exon1, and part of intron 1.



**Figure 3:** Frequency of *MICA* null alleles in different population samples. *MICA*\*064N presents a premature stop-codon with no translation of a functional *MICA*. *MICA*\*del represents the complete deletion of *MICA*. *MICA*\*partial represents a partial deletion in *MICA* encompassing the promoter, exon1, and part of intron 1.

### Linkage disequilibrium between *MICA*\*del and *MICB*, *HLA-B*, and *HLA-C*

We evaluated the haplotypes observed in chromosomes carrying *MICA*\*del or *MICA*\*partial, considering *MICB*, *HLA-B*, and *HLA-C* alleles (Table 1 and Table S1). We noticed a strong LD among *MICA*\*del, *HLA-B*, and *MICB*, but not with *HLA-C* (Table S1). In general, more than 50% of all haplotypes carrying *MICA*\*del also carry B\*48:01 and *MICB*\*009:01N. This haplotype can be detected in America and in South and East Asia but not in Africa, Europe, and the Middle East (TS2).

There are other frequent haplotypes associated with *MICA*\*del (Table S2). For instance, the second most frequent haplotype is B\*15:16/*MICA*\*del/*MICB*\*002:01, present in Africa, America, and the Middle East. The haplotype B\*39:06/*MICA*\*del/*MICB*\*004:01 is also frequent and exclusive from America, possibly among Amerindians since B\*39:06 is a Native American allele <sup>121</sup>. Interestingly, the most frequent *MICA*\*del haplotype in the Middle East, B\*52:01/*MICA*\*del/*MICB*\*005:02, only occurs in that region. *HLA-B*\*52:01 is frequent in the Middle East and among Native Americans <sup>121</sup>. *MICA*\*partial is mainly linked to *HLA-B*\*41:01 and two different *MICB* alleles (TS2).

**Table 1:** The frequency of different haplotypes observed for chromosomes 6 carrying MICA\*del or MICA\*partial

| Frequency<br><br>N = 158 | HLA-B      | MICA         | MICB            |
|--------------------------|------------|--------------|-----------------|
| 0.3608                   | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.1646                   | B*15:16:01 | MICA*del     | MICB*002:01:01  |
| 0.1139                   | B*48:02:01 | MICA*del     | MICB*009:01:01N |
| 0.0633                   | B*39:06:02 | MICA*del     | MICB*004:01:01  |
| 0.038                    | B*48:03:01 | MICA*del     | MICB*009:01:01N |
| 0.0316                   | B*15:16:01 | MICA*del     | MICB*005:02:01  |
| 0.0316                   | B*41:01:01 | MICA*partial | MICB*005:02:01  |
| 0.0253                   | B*48:01:01 | MICA*del     | MICB*004:01:01  |
| 0.0253                   | B*48:04:01 | MICA*del     | MICB*003:01:01  |
| 0.0253                   | B*52:01:01 | MICA*del     | MICB*005:02:01  |
| 0.019                    | B*41:01:01 | MICA*partial | MICB*002:01:01  |
| 0.0127                   | B*15:10:01 | MICA*del     | MICB*004:01:01  |
| 0.0127                   | B*48:01:01 | MICA*del     | MICB*005:02:01  |
| 0.0063                   | B*15:11:01 | MICA*del     | MICB*005:02:01  |
| 0.0063                   | B*15:15:01 | MICA*del     | MICB*004:01:01  |

|        |            |              |                |
|--------|------------|--------------|----------------|
| 0.0063 | B*15:18:01 | MICA*del     | MICB*004:01:01 |
| 0.0063 | B*27:05:02 | MICA*del     | MICB*002:01:01 |
| 0.0063 | B*41:01:01 | MICA*partial | MICB*008:01:01 |
| 0.0063 | B*48:04:01 | MICA*del     | MICB*005:08,L  |
| 0.0063 | B*48:43    | MICA*del     | MICB*005:02:01 |
| 0.0063 | B*55:01:01 | MICA*del     | MICB*003:01:01 |
| 0.0063 | B*58:01:01 | MICA*del     | MICB*004:01:01 |
| 0.0063 | B*58:02:01 | MICA*del     | MICB*002:01:01 |
| 0.0063 | newcds_3   | MICA*partial | MICB*005:02:01 |

## Discussion

The *HLA* and *MIC* genes are challenging to analyze using NGS short reads because of their polymorphic nature and their paralogy. These genes are very similar to each other, which jeopardizes read alignment mostly due to cross-alignments. *MICA* and *MICB*, for instance, present a sequence homology of 91%<sup>85 86</sup>. Because of that, we may obtain many genotyping errors when using conventional tools for short-read alignment. This issue may jeopardize detecting the correct alleles for *MICA* and *MICB*.

In this study, we focused on the deletion of the *MICA* gene (*MICA\*del*) and the *HLA-B* and *MICB* haplotypes associated with it. To get accurate coverage estimations (used for detecting *MICA\*del*) and accurate genotypes for *MICB* and *HLA-B*, we used a pipeline that minimizes cross-alignments for *HLA* and *MIC* genes. We explored samples from all continents, including samples from Brazil, a highly admixed population not represented in global databases such as the international genome<sup>116,122</sup>. This study is the first one to address *MICA* deletion in worldwide populations.

The deletion at the *MICA* region was described previously<sup>82,83</sup> and comprehended about 100Kb between *HLA-B* and *MICB*. This deletion, named *MICA\*del*, was associated with *MICB\*009N* and *HLA-B\*048* in several Asian

populations<sup>82,83,123</sup>. We confirm that this haplotype is frequent in Asia, and we also detected it in America but not in Europe, Africa, and Oceania. The lack of *MICA\*del* in Oceania might result from the low sample size we have studied for this region. Although this corresponds to 50% of the sequences carrying *MICA\*del*, we also detected different haplotypes than B\*48/*MICA\*del*/*MICB\*009N* (Table S2) involving other *HLA-B* and *MICB* alleles. In the Middle East, *MICA\*del* is associated mainly with B\*52:01 and *MICB\*005:01*; this haplotype only occurs in that region. Likewise, the haplotype B\*39:06/*MICA\*del*/*MICB\*004:01* was detected only in America and it is related to Native Americans.

Different haplotypes carrying *MICA\*del* suggests that *MICA* deletion might have occurred in different chromosomes as an independent event due to *denovo* deletions involving recombination hotspots. When zooming at the coverage pattern at the borders of the deleted region, we can notice at least three different patterns for the deletion starting point (samples 1, 2, and 3 in Figure S2). Each sample presents a different *HLA-B* and *MICB* allele associated with *MICA\*del*. Likewise, the deletion ending point seems also to be different (samples 1, 2, and 3 in Figure S3). Therefore, although we called all these deletions *MICA\*del*, we may have three different deletions with specific starting and ending points. Nevertheless, the effect would be the same: no functional *MICA* from these chromosomes. It is unclear whether these deletions also influence *MICB* expression patterns since they also delete part of the distal *MICB* promoter region.

Besides the *MICA\*del* allele, we also detected a different deletion that removes the promoter region, exon 1, and part of intron 1. We called this allele *MICA\*partial*. Because the promoter region is deleted, and exon 1 encodes the leader peptide that addresses the molecule to the cell surface, we considered that *MICA\*partial* and *MICA\*del* might have similar effects, canceling *MICA* expression producing a phenotype that we will call the absence of *MICA*. *MICA\*partial* was detected only in America and the Middle East.

The chromosome with the absence of *MICA* is uncommon, with its higher frequency of around 2.5% in East Asia and the Middle East (Figure 2). It is almost absent from Europe. We notice a high frequency of this phenotype in America, which is incompatible with the expected frequencies when considering the observed for Africa and Europe. This pattern does not occur for all the American populations we have studied. While the absence of *MICA* is frequent among Mexicans (Mayas and from Los Angeles with Mexican ancestry), Native Americans (Surui from Brazil), and Peruvians, it is not frequent among Brazilians from São Paulo and Colombians from Medellin. These two last groups present major European ancestry, and this phenotype is almost absent from Europe. Therefore, the chromosome with the absence of *MICA* is more common in Asians and Native Americans. Nevertheless, most samples presenting the deletion of *MICA* are heterozygous and present at least one copy of a functional *MICA*.

Although rare, we detected two individuals that do not produce MICA. The first individual is NA19761 from the 1000 Genomes dataset (Mexican Ancestry from Los Angeles, USA), with the genotype *MICA*\*del/*MICA*\*partial. (Sample 4 in Figure S1). This individual also presents MICB\*009:01N in one chromosome but has a functional *MICB* in the other. Therefore, this individual might produce *MICB* from one chromosome. A second individual is an older man (> 70 y) homozygous B\*48:02/*MICA*\*del/MICB\*009:01N. Unless there is a mechanism to bypass the stop codon from MICB\*009:01N, this individual cannot produce a functional *MIC*, and other mechanisms to activate the cytotoxic response of NK cells might be in place.

Assuming that the Brazilian individual mentioned above does not produce a functional MIC, the nonexpression of both *MICA* and *MICB* is not critical regarding immunological fitness. Still, the chromosome with a deleted MICA (\*del or \*partial) may be selectively neutral in immune defense or present some advantage. Since the MICA-MICB null haplotype (*MICA*\*del/MICB\*009N) was widely observed with almost the same frequencies among many Asian populations, positive selection might be in place. The present observations would be useful for understanding the essential function of MIC molecules.

In conclusion, we have explored the deletion of the MICA allele and the HLA-B and MICB haplotypes associated with it in different biogeographic regions, including admixed samples from Brazil. We detected at least four different patterns of deletion, 3 with full deletion but different starting and ending points and associated with different HLA-B/MICB haplotypes and one partial MICA deletion. The presence of different patterns of deletion suggests independent deletion events. The *MICA*\*del allele is mostly associated with B\*48 and MICB\*009N in Asia and America (particularly among samples with higher Native American ancestries), but other haplotypes also occur. While most samples present the complete partial MICA deletion are heterozygous and present at least one functional MICB gene, we detected two samples with no functional MICA and one sample with no functional MIC gene. Therefore, other mechanisms might be in place to compensate for the absence of MIC molecules.

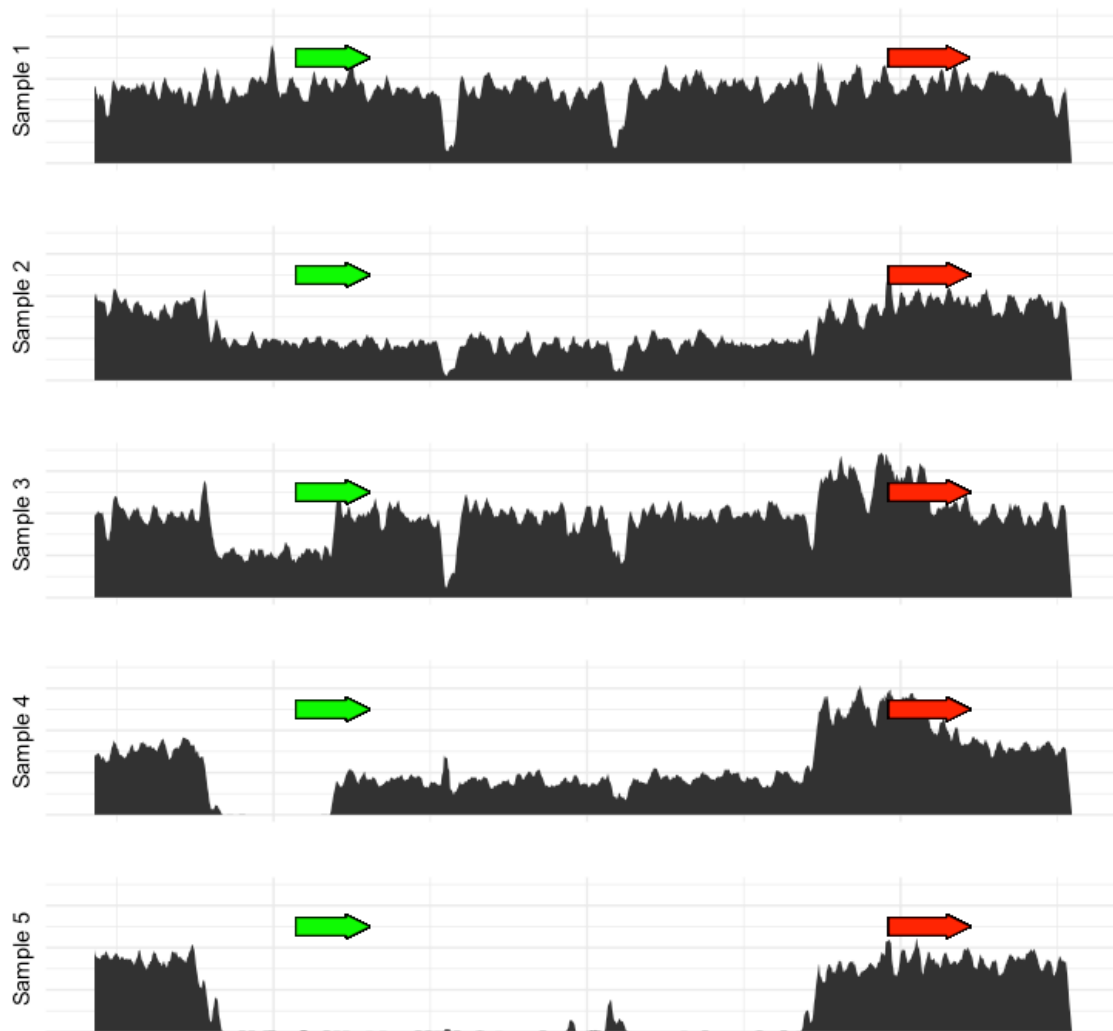
## References

1. Bahram S, Bresnahan M, Geraghty DE, Spies T. A second lineage of mammalian major histocompatibility complex class I genes. *Proc Natl Acad Sci* 1994;91(14):6259–6263.
2. Raulet DH, Gasser S, Gowen BG, Deng W, Jung H. Regulation of Ligands for the NKG2D Activating Receptor. *Annu Rev Immunol* 2013;31(1):413–441.
3. Zwirner NW, Dole K, Stastny P. Differential surface expression of MICA by endothelial cells, fibroblasts, keratinocytes, and monocytes. *Hum Immunol* 1999;60(4):323–330.
4. BAHRAM S. MIC genes: from genetics to biology. *MIC Genes Genet Biol* 2000;76:1–60.
5. Tian W, Boggs DA, Ding W-Z, Chen D-F, Fraser PA. MICA genetic polymorphism and linkage disequilibrium with HLA-B in 29 African-American families. *Immunogenetics* 2001;53(9):724–728.
6. Ando H, Mizuki N, Ota M, et al. Allelic variants of the human MHC class I chain-related B gene (MICB). *Immunogenetics* 1997;46(6):499–508.
7. Yamamoto K, Fujiyama Y, Andoh A, Bamba T, Okabe H. Oxidative stress increases MICA and MICB gene expression in the human colon carcinoma cell line (CaCo-2). *Biochim Biophys Acta BBA - Gen Subj* 2001;1526(1):10–12.
8. García G, Puerto F del, Pérez AB, et al. Association of MICA and MICB alleles with symptomatic dengue infection. *Hum Immunol* 2011;72(10):904–907.
9. Luangtrakool P, Vejbaesya S, Luangtrakool K, et al. Major Histocompatibility Complex Class I Chain–Related A and B (MICA and MICB) Gene, Allele, and Haplotype Associations With Dengue Infections in Ethnic Thais. *J Infect Dis* 2020;222(5):840–846.
10. Chen E, Chen C, Chen F, Yu P, Lin L. Positive association between MIC gene polymorphism and tuberculosis in Chinese population. *Immunol Lett* 2019;213:62–69.
11. Castelli EC et al. Immunogenetics of resistance to SARS-CoV-2 infection in discordant couples. *medRxiv* 2021.
12. Groh V, Rhinehart R, Secrist H, Bauer S, Grabstein KH, Spies T. Broad tumor-associated expression and recognition by tumor-derived  $\gamma\delta$  T cells of MICA and MICB. *Proc Natl Acad Sci* 1999;96(12):6879–6884.
13. Groh V, Wu J, Yee C, Spies T. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature* 2002;419(6908):734–738.
14. Mizuki N, Ota M, Kimura M, et al. Triplet repeat polymorphism in the transmembrane region of the MICA gene: A strong association of six GCT repetitions with Behçet disease. *Proc Natl Acad Sci* 1997;94(4):1298–1303.
15. Zou Y, Han M, Wang Z, Stastny P. MICA Allele-Level Typing by Sequence-based Typing with Computerized Assignment of Polymorphic Sites and Short Tandem Repeats within the Transmembrane Region. *Hum Immunol* 2006;67(3):145–151.
16. Salih HR, Rammensee H-G, Steinle A. Cutting Edge: Down-Regulation of

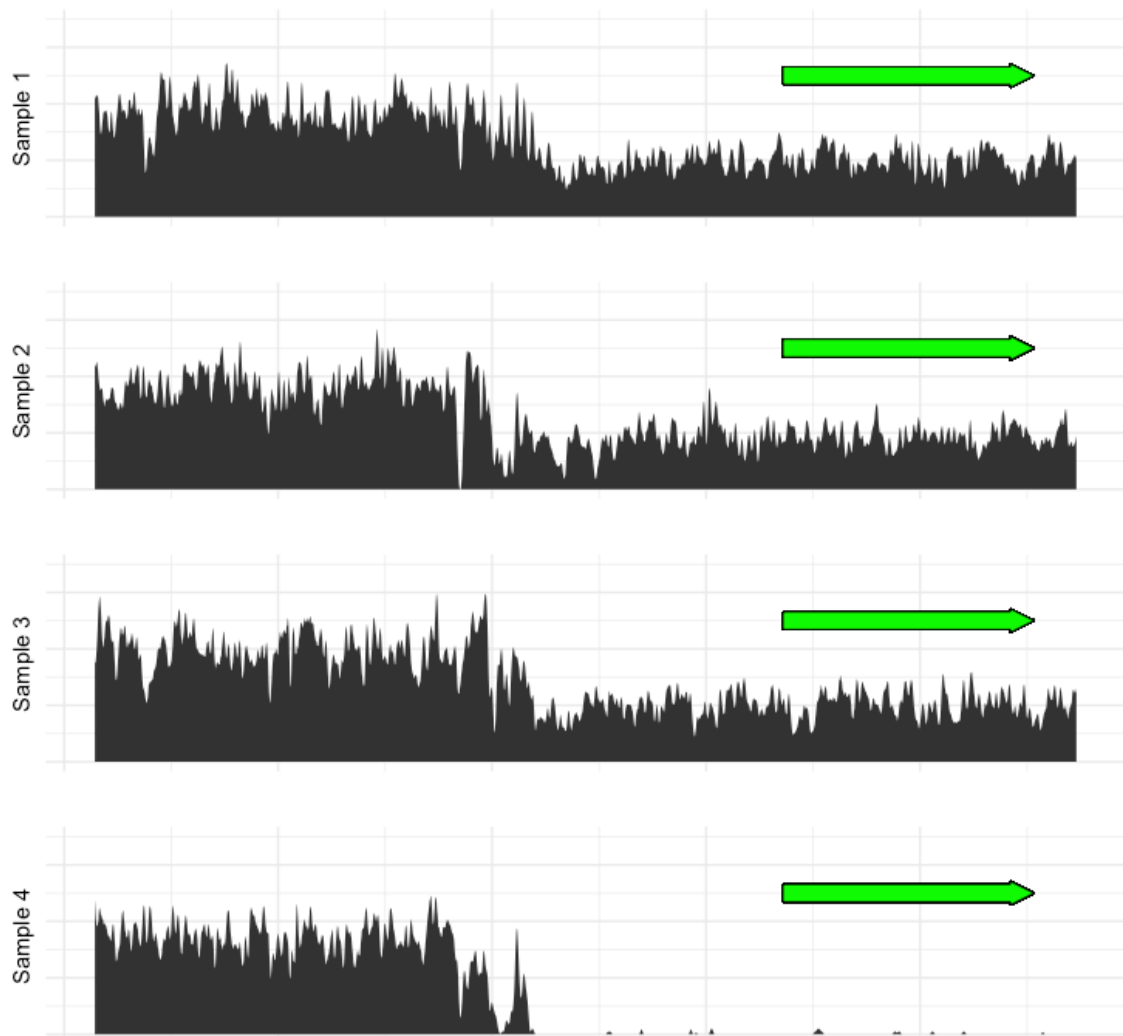
- MICA on Human Tumors by Proteolytic Shedding. *J Immunol* 2002;169(8):4098–4102.
17. Waldhauer I, Goehlsdorf D, Gieseke F, et al. Tumor-Associated MICA Is Shed by ADAM Proteases. *Cancer Res* 2008;68(15):6368–6376.
  18. Salih HR, Goehlsdorf D, Steinle A. Release of MICB Molecules by Tumor Cells: Mechanism and Soluble MICB in Sera of Cancer Patients. *Hum Immunol* 2006;67(3):188–195.
  19. Li P, Morris DL, Willcox BE, Steinle A, Spies T, Strong RK. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nat Immunol* 2001;2(5):443–451.
  20. Chitadze G, Bhat J, Lettau M, Janssen O, Kabelitz D. Generation of Soluble NKG2D Ligands: Proteolytic Cleavage, Exosome Secretion and Functional Implications. *Scand J Immunol* 2013;78(2):120–129.
  21. Zwirner NW, Fernández-Viña MA, Stastny P. MICA, a new polymorphic HLA-related antigen, is expressed mainly by keratinocytes, endothelial cells, and monocytes. *Immunogenetics* 1997;47(2):139–148.
  22. Komatsu-Wakui M, Tokunaga K, Ishikawa Y, et al. Wide distribution of the MICA-MICB null haplotype in East Asians. *Tissue Antigens* 2001;57(1):1–8.
  23. Komatsu-Wakui M, Tokunaga K, Ishikawa Y, et al. MIC-A polymorphism in Japanese and a MIC-A-MIC-B null haplotype. *Immunogenetics* 1999;49(7):620–628.
  24. Tse K-P, Su W-H, Yang M, et al. A gender-specific association of CNV at 6p21.3 with NPC susceptibility†. *Hum Mol Genet* 2011;20(14):2889–2896.
  25. Wang W, Tian W, Zhu F, et al. *MICA* Gene Deletion in 3411 DNA Samples from Five Distinct Populations in Mainland China and Lack of Association with Nasopharyngeal Carcinoma (NPC) in a Southern Chinese Han population: MICA Gene Deletion in 3411 DNA Samples. *Ann Hum Genet* 2016;80(6):319–326.
  26. Komatsu-Wakui M, Tokunaga K, Ishikawa Y, et al. Wide distribution of the *MICA-MICB* null haplotype in East Asians: The *MICA-MICB* null haplotype in East Asians. *Tissue Antigens* 2001;57(1):1–8.
  27. Ota M, Bahram S, Katsuyama Y, et al. On the MICA deleted-MICB null, HLA-B\*4801 haplotype. *Tissue Antigens* 2000;56(3):268–271.
  28. 1000 GENOMES PROJECT CONSORTIUM et al. A global reference for human genetic variation. *Nature* 2015; 526 (7571):68.
  29. Cann HM, Toma C de, Cazes L, et al. A Human Genome Diversity Cell Line Panel. *Science* 2002;296(5566):261–262.
  30. NASLAVSKY, Michel S. et al. Whole-genome sequencing of 1,171 elderly admixed individuals from the largest Latin American metropolis (São Paulo, Brazil). *BioRxiv* 2020.
  31. Castelli EC, Paz MA, Souza AS, Ramalho J, Mendes-Junior CT. Hla-mapper: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. *Hum Immunol* 2018;79(9):678–684.
  32. Silva N dos SB, Souza A da S, Andrade H de S, et al. Immunogenetics of HLA-B: SNP, allele, and haplotype diversity in populations from different continents and ancestry backgrounds. *HLA* 2023;101(6):634–646.
  33. Naslavsky MS, Scliar MO, Yamamoto GL, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nat Commun* 2022;13(1):1004.

34. Solberg OD, Mack SJ, Lancaster AK, et al. Balancing selection and heterogeneity across the classical human leukocyte antigen loci: A meta-analytic review of 497 population studies. *Hum Immunol* 2008;69(7):443–464.
35. Cosman D, Müllberg J, Sutherland CL, et al. ULBPs, Novel MHC Class I–Related Molecules, Bind to CMV Glycoprotein UL16 and Stimulate NK Cytotoxicity through the NKG2D Receptor. *Immunity* 2001;14(2):123–133.
36. Wang WY, Tian W, Zhu F m., Liu XX, Li LX, Wang F. MICA, MICB Polymorphisms and Linkage Disequilibrium with HLA-B in a Chinese Mongolian Population. *Scand J Immunol* 2016;83(6):456–462.
37. A global reference for human genetic variation. *Nature* 2015;526(7571):68–74.
38. Liu X, Tian W, Li L, Cai J. Characterization of the major histocompatibility complex class I chain-related gene B (MICB) polymorphism in a northern Chinese Han population: The identification of a new MICB allele, MICB\*023. *Hum Immunol* 2011;72(9):727–732.

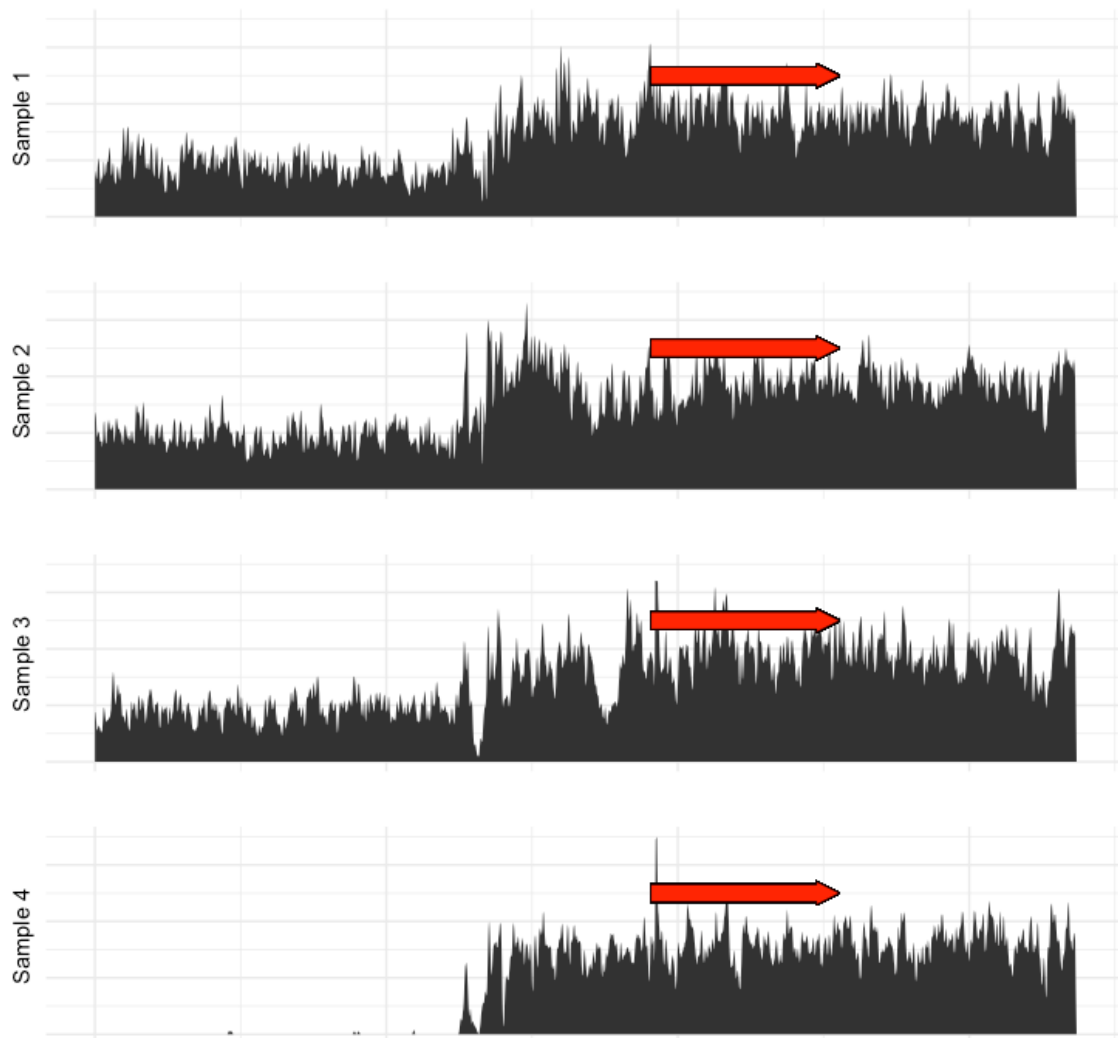
## Supplementary



**Figure S1:** An overview of the sequencing coverage across the *MICA* and *MICB* loci. Sample 1 indicates a sample with two copies of *MICA*. Sample 2 indicates a sample with one copy of *MICA* and a full deletion. Sample 3 indicates a sample with a full copy of *MICA* and one with a partial deletion evolving exon 1 and intron 1. Sample 4 represents a sample with a partial deletion and a full deletion. Sample 5 represents a sample with a full deletion in both chromosomes. The green and red arrows indicated the *MICA* and *MICB* genes, respectively.



**Figure S2:** An overview of the sequencing coverage at the upstream border of the MICA deleted region. Sample 1 indicates an heterozygous sample carrying B\*15:16/MICA\*del/MICB\*002:01. Sample 2 indicates an heterozygous sample carrying B\*39:06/MICA\*del/MICB\*004:01. Sample 3 indicates an heterozygous sample carrying B\*48:01/MICA\*del/MICB\*009:01N. Sample 4 indicates an homozygous sample for B\*48:02/MICA\*del/MICB\*009:01N. The green arrow indicates the *MICA* gene.



**Figure S3:** An overview of the sequencing coverage at the downstream border of the MICA deleted region. Sample 1 indicates an heterozygous sample carrying B\*15:16/MICA\*del/MICB\*002:01. Sample 2 indicates an heterozygous sample carrying B\*39:06/MICA\*del/MICB\*004:01. Sample 3 indicates an heterozygous sample carrying B\*48:01/MICA\*del/MICB\*009:01N. Sample 4 indicates an homozygous sample for B\*48:02/MICA\*del/MICB\*009:01N. The red arrow indicates the *MICB* gene.

**Table S1:** Frequency of MICA\*del, MICB, HLA-B and HLA-C haplotypes.

| Frequency | HLA-C      | HLA-B      | MICA*del     | MICB            |
|-----------|------------|------------|--------------|-----------------|
| 0.0063    | C*01:02:01 | B*15:15:01 | MICA*del     | MICB*004:01:01  |
| 0.0063    | C*01:02:01 | B*27:05:02 | MICA*del     | MICB*002:01:01  |
| 0.0063    | C*01:02:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0253    | C*01:02:01 | B*48:04:01 | MICA*del     | MICB*003:01:01  |
| 0.0063    | C*01:02:01 | B*55:01:01 | MICA*del     | MICB*003:01:01  |
| 0.0063    | C*01:64,L  | newcds_39  | MICA*del     | MICB*002:01:01  |
| 0.0063    | C*03:02:02 | B*58:01:01 | MICA*del     | MICB*004:01:01  |
| 0.0063    | C*03:03:01 | B*15:11:01 | MICA*del     | MICB*005:02:01  |
| 0.019     | C*03:03:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0127    | C*03:04:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*03:04:01 | B*48:43    | MICA*del     | MICB*005:02:01  |
| 0.0063    | C*03:04:02 | B*15:10:01 | MICA*del     | MICB*004:01:01  |
| 0.0063    | C*04:01:01 | B*15:10:01 | MICA*del     | MICB*004:01:01  |
| 0.0127    | C*04:01:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.1076    | C*04:01:01 | B*48:02:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*04:81,L  | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*04:82:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*06:02:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*06:02:01 | B*58:02:01 | MICA*del     | MICB*002:01:01  |
| 0.019     | C*07:01:01 | B*41:01:01 | MICA*partial | MICB*002:01:01  |
| 0.0316    | C*07:01:01 | B*41:01:01 | MICA*partial | MICB*005:02:01  |
| 0.0063    | C*07:01:01 | B*41:01:01 | MICA*partial | MICB*008:01:01  |
| 0.0633    | C*07:02:01 | B*39:06:02 | MICA*del     | MICB*004:01:01  |
| 0.0127    | C*07:02:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*07:04:01 | B*15:18:01 | MICA*del     | MICB*004:01:01  |
| 0.0253    | C*08:01:01 | B*48:01:01 | MICA*del     | MICB*004:01:01  |
| 0.1456    | C*08:01:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.038     | C*08:01:01 | B*48:03:01 | MICA*del     | MICB*009:01:01N |
| 0.0063    | C*08:01:32 | B*48:01:01 | MICA*del     | MICB*009:01:01N |

|        |            |            |              |                 |
|--------|------------|------------|--------------|-----------------|
| 0.0063 | C*08:02:01 | B*48:02:01 | MICA*del     | MICB*009:01:01N |
| 0.0127 | C*08:03:01 | B*48:01:01 | MICA*del     | MICB*005:02:01  |
| 0.1203 | C*08:03:01 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.0063 | C*12:02:02 | B*48:01:01 | MICA*del     | MICB*009:01:01N |
| 0.1203 | C*14:02:01 | B*15:16:01 | MICA*del     | MICB*002:01:01  |
| 0.0316 | C*14:02:01 | B*15:16:01 | MICA*del     | MICB*005:02:01  |
| 0.0443 | C*16:01:01 | B*15:16:01 | MICA*del     | MICB*002:01:01  |
| 0.0253 | C*16:02:01 | B*52:01:01 | MICA*del     | MICB*005:02:01  |
| 0.0063 | newcds_23  | newcds_3   | MICA*partial | MICB*005:02:01  |
| 0.0063 | newcds_35  | B*48:04:01 | MICA*del     | MICB*005:08,L   |

**Table S2:** Frequency of *MICA*, *MICB* and *HLA-B* haplotype distribution in different biogeographical populations.

| DATA       |              |                 | AFR    | AMR    | CSA   | EAS    | EUR    | MEA    | SAS   | Global |
|------------|--------------|-----------------|--------|--------|-------|--------|--------|--------|-------|--------|
| 2N         |              |                 | 1578   | 4310   | 362   | 1556   | 1540   | 306    | 992   | 10694  |
| B*48:01:01 | MICA*del     | MICB*009:01:01N | -      | 0.0058 | 0.011 | 0.0174 | -      | -      | 0.001 | 0.0053 |
| B*15:16:01 | MICA*del     | MICB*002:01:01  | 0.0076 | 0.003  | -     | -      | -      | 0.0033 | -     | 0.0024 |
| B*48:02:01 | MICA*del     | MICB*009:01:01N | -      | 0.0042 | -     | -      | -      | -      | -     | 0.0017 |
| B*39:06:02 | MICA*del     | MICB*004:01:01  | -      | 0.0023 | -     | -      | -      | -      | -     | 0.0009 |
| B*48:03:01 | MICA*del     | MICB*009:01:01N | -      | 0.0002 | -     | 0.0032 | -      | -      | -     | 0.0006 |
| B*15:16:01 | MICA*del     | MICB*005:02:01  | 0.0013 | 0.0007 | -     | -      | -      | -      | -     | 0.0005 |
| B*41:01:01 | MICA*partial | MICB*005:02:01  | -      | 0.0009 | -     | -      | -      | 0.0033 | -     | 0.0005 |
| B*48:01:01 | MICA*del     | MICB*004:01:01  | -      | 0.0009 | -     | -      | -      | -      | -     | 0.0004 |
| B*48:04:01 | MICA*del     | MICB*003:01:01  | -      | -      | -     | -      | -      | -      | 0.004 | 0.0004 |
| B*52:01:01 | MICA*del     | MICB*005:02:01  | -      | -      | -     | -      | -      | 0.0131 | -     | 0.0004 |
| B*41:01:01 | MICA*partial | MICB*002:01:01  | -      | 0.0005 | -     | -      | -      | 0.0033 | -     | 0.0003 |
| B*15:10:01 | MICA*del     | MICB*004:01:01  | 0.0013 | -      | -     | -      | -      | -      | -     | 0.0002 |
| B*48:01:01 | MICA*del     | MICB*005:02:01  | -      | -      | -     | 0.0013 | -      | -      | -     | 0.0002 |
| B*15:11:01 | MICA*del     | MICB*005:02:01  | -      | -      | -     | 0.0006 | -      | -      | -     | 0.0001 |
| B*15:15:01 | MICA*del     | MICB*004:01:01  | -      | 0.0002 | -     | -      | -      | -      | -     | 0.0001 |
| B*15:18:01 | MICA*del     | MICB*004:01:01  | -      | -      | -     | 0.0006 | -      | -      | -     | 0.0001 |
| B*27:05:02 | MICA*del     | MICB*002:01:01  | -      | -      | -     | -      | 0.0006 | -      | -     | 0.0001 |

|            |              |                |        |        |   |        |   |   |       |        |
|------------|--------------|----------------|--------|--------|---|--------|---|---|-------|--------|
| B*41:01:01 | MICA*partial | MICB*008:01:01 | -      | 0.0002 | - | -      | - | - | -     | 0.0001 |
| B*48:04:01 | MICA*del     | MICB*005:08,L  | -      | 0.0002 | - | -      | - | - | -     | 0.0001 |
| B*48:43    | MICA*del     | MICB*005:02:01 | -      | -      | - | 0.0006 | - | - | -     | 0.0001 |
| B*55:01:01 | MICA*del     | MICB*003:01:01 | -      | -      | - | -      | - | - | 0.001 | 0.0001 |
| B*58:01:01 | MICA*del     | MICB*004:01:01 | 0.0006 | -      | - | -      | - | - | -     | 0.0001 |
| B*58:02:01 | MICA*del     | MICB*002:01:01 | 0.0006 | -      | - | -      | - | - | -     | 0.0001 |
| newcds_39  | MICA*del     | MICB*002:01:01 | -      | -      | - | 0.0006 | - | - | -     | 0.0001 |
| newcds_3   | MICA*partial | MICB*005:02:01 | -      | 0.0002 | - | -      | - | - | -     | 0.0001 |

Africa (AFR), America (AMR), Central South Asia (CSA), East Asia (EAS), Europa (EUR) Middle East (MEA), Oceania (OCE), South Asia (SAS). (–) indicates that this allele is absent in this biogeographic region considering the samples we have studied.