

RESSALVA

Atendendo solicitação do(a) autor(a),  
o texto completo desta dissertação será disponibilizado  
somente a partir de 26/02/2020.



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

**Iane de Oliveira Pires Porto**

**Estrutura e variabilidade da região 3' não-traduzida  
dos genes *HLA-A*, *HLA-C* e *HLA-G* e perfil de ligação  
de microRNAs**

Tese apresentada à Faculdade de  
Medicina, Universidade Estadual  
Paulista “Júlio de Mesquita Filho”,  
Campus de Botucatu, para obtenção  
do título de Doutora em Patologia.

Orientador: Prof. Dr. Erick da Cruz Castelli

**Botucatu**

**2018**

Iane de Oliveira Pires Porto

**Estrutura e variabilidade da região 3' não-traduzida dos genes *HLA-A*, *HLA-C* e *HLA-G* e perfil de ligação de microRNAs**

Tese apresentada à Faculdade de Medicina, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Botucatu, para obtenção do título de Doutora em Patologia.

Orientador: Prof. Dr. Erick da Cruz Castelli

Botucatu  
2018

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

Porto, Iane de Oliveira Pires.

Estrutura e variabilidade da região 3' não-traduzida dos genes *HLA-A*, *HLA-C* e *HLA-G* e perfil de ligação de microRNAs / Iane de Oliveira Pires Porto. - Botucatu, 2018

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Medicina de Botucatu

Orientador: Erick da Cruz Castelli

Capes: 40105008

1. Antígenos *HLA*. 2. Antígenos *HLA-C*. 3. Antígenos *HLA-G*.  
4. MicroRNAs.

Palavras-chave: *HLA-A*; *HLA-C*; *HLA-G*; microRNAs.



Agradecimentos

Agradeço inicialmente (e principalmente) à minha família, pelo apoio incondicional. Sem vocês, nada disso seria possível. Esse trabalho vem do esforço dos meus pais (José e Maria Socorro) e do apoio dos meus irmãos (Bruno, Igor e Clara).

Agradeço especialmente ao meu orientador, Dr. Erick Castelli, pessoa e profissional admirável a quem eu tive a sorte e o prazer de encontrar ainda nos tempos de iniciação científica. Obrigada por acreditar em mim e por me ajudar a realizar o sonho de um dia estar do outro lado da sala de aula.

Agradeço à melhor equipe que alguém poderia desejar: Thálitta Ayala (nossa Xuxu), Andreia Souza, Michelle Paz e Emiliana Weiss. Me faltam palavras para agradecer o papel de vocês não só nesse trabalho, mas na minha vida. Obrigada!

Aos amigos de todas as horas: Sarah Muricy, Suedio Meira, Gabriela Oliveira, Karla Polo, Monica Maffei, André Jim, companheiras de moradia, e amigos cujos nomes não foram citados – obrigada por todos os momentos felizes, pelos conselhos nos momentos de dúvida, por todos os ensinamentos compartilhados comigo. Meus dias são mais felizes com vocês.

E por último, mas não menos importante: Obrigada aos membros da banca, por dedicarem tempo e conhecimento para este trabalho. A contribuição de vocês é altamente apreciada.

Obrigada!!



Resumo

## RESUMO

O complexo gênico de Antígenos Leucocitário Humanos (HLA) é a região mais variável do genoma humano. Os genes HLA de classe I são divididos em clássicos e não clássicos, a depender de suas funções primárias, níveis de expressão e polimorfismos. Assim, *HLA-A*, *HLA-B* e *HLA-C* são loci clássicos e estão associados à apresentação antigênica para células T citotóxicas, enquanto *HLA-E*, *HLA-G* e *HLA-F* são loci não clássicos e estão associados à imunomodulação e inibição de células T e NK. No entanto, o *HLA-C* também apresenta propriedades imunomodulatórias a partir da interação com o TCR de células T CD8 (ativação) e com receptores KIR de células NK (principalmente inibição). A maioria dos estudos sobre variabilidade dos genes HLA têm como foco suas regiões codificadoras e negligenciam regiões regulatórias – região promotora e região 3' não-traduzida (3'NT). Esta última apresenta um papel essencial na estabilidade do mRNA e no controle pós-transcricional mediado pela ligação de microRNAs (miRNA). Neste trabalho, nós avaliamos a estrutura e a variabilidade das regiões 3'NTs dos genes *HLA-A*, *HLA-C* e *HLA-G* por sequenciamento de nova geração além do padrão de ligação de miRNAs observado para cada gene. A região 3'NT dos genes *HLA-A* e *HLA-C* são altamente variáveis, enquanto a de *HLA-G* apresenta poucos pontos de variação. No entanto, os três genes apresentaram números similares de haplótipos de 3'NT frequentes. Vários miRNAs possuem potencial de regular *HLA-A*, *HLA-C* e *HLA-G* tanto de modo específico quanto os três loci juntos. Este fato pode estar relacionado à regulação fina da expressão de HLA em tecidos diversos e em contextos imunológicos como a gravidez, onde *HLA-C* e *HLA-G* estão presentes, mas *HLA-A* está ausente. miRNAs capazes de regular mais de um gene HLA devem ser especialmente importante em sítios imunoprivilegiados. Os miRNAs com potencial de regular os genes estudados, individualmente ou em conjunto, foram compilados neste trabalho e devem nortear estudos funcionais a respeito da regulação pós-transcricional dos níveis de HLA.





Abstract

## ABSTRACT

The Human Leucocyte Antigen (HLA) gene complex is the most variable region of the human genome. The class I HLA genes are classified as classical or non-classical depending on their primary function, expression levels and polymorphism rate. Thus, *HLA-A*, *HLA-B*, and *HLA-C* are classical loci, associated with antigen presentation to cytotoxic T cells, while *HLA-E*, *HLA-G*, and *HLA-F* are non-classical loci associated with immunomodulation inhibition of both T and NK cells. Nevertheless, *HLA-C* also presents immunomodulatory features by interacting with TCR on CD8 T (activation) cells as well as KIR receptors on NK cells (mostly inhibition). Most studies on HLA variability focus on their coding sequences and neglects their regulatory sequences – promoter and 3'untranslated region (3'UTR). The last one plays a pivotal role on mRNA stability and post-transcriptional control by microRNA (miRNA) binding. Here we evaluated 3'UTR structure and variability for *HLA-A*, *HLA-C*, and *HLA-G* by using massively parallel sequencing in a Brazilian population sample. We also report the miRNA binding pattern observed for each locus. Both *HLA-A* and *HLA-C* 3'UTR segments are highly variable, while *HLA-G* 3'UTR presents few variation sites. However, the three loci present a similar number of frequent 3'UTR haplotypes. Several miRNAs are potential regulators of *HLA-A*, *HLA-C*, and *HLA-G* on a specific manner as well as the three of them together. This fact may be associated with the fine regulation of HLA expression on different tissues and immunological contexts, such as pregnancy, where *HLA-C* and *HLA-G* are present and *HLA-A* is absent. MiRNAs that are able to regulate more than one HLA locus may be important on immunoprivileged sites. miRNAs with potential to regulate the evaluated genes, both individually or together, were compiled on the present series and may guide functional studies regarding post-transcriptional regulation of HLA levels.

## LISTA DE ILUSTRAÇÕES

### CAPÍTULO I

<b>Figura 1</b> - Organização do Complexo Principal de Histocompatibilidade em humanos.....	19
<b>Figura 2</b> - Estrutura das moléculas de MHC de classes I e II.....	20
<b>Figura 3</b> - Processamento antigênico pelo proteassomo e montagem da molécula de MHC de classe I no Retículo Endoplasmático.....	23
<b>Figura 4</b> - Esquema da nomenclatura dos alelos de HLA.....	25
<b>Figura 5</b> - Esquema da biogênese dos microRNAs.....	39

### CAPÍTULO II

<b>Figure 1</b> - Conservation level across the <i>HLA-A</i> , <i>HLA-C</i> and <i>HLA-G</i> 3'UTR segment.....	70
<b>Figure 2</b> - Number of different miRNAs binding across the HLA genes with equal or lower energy values than the cutoff.....	78
<b>Figure 3</b> - Number of miRNAs interacting with <i>HLA-A</i> , <i>HLA-C</i> or <i>HLA-G</i> ...	79
<b>Figure 4</b> - Number of miRNAs interacting specifically with either <i>HLA-A</i> , <i>HLA-C</i> and <i>HLA-G</i> across their 3'UTR.....	80
<b>Figure 5</b> - Shared miRNAs across the <i>HLA-A</i> , <i>HLA-C</i> and <i>HLA-G</i> 3'UTR....	81
<b>Figure 6</b> - Number miRNAs shared between <i>HLA-A</i> and <i>HLA-C</i> 3'UTR and their distribution along the segment.....	82
<b>Figure 7</b> - <b>A</b> ) miRNAs highly specific (not shared) to <i>HLA-A</i> (red), <i>HLA-C</i> (green) and <i>HLA-G</i> (blue) 3'UTR and their target along the UTR segment. <b>B</b> ) Number of specific miRNAs (high S score) shared between the three loci and specific to each locus.....	83

## LISTA DE TABELAS

### CAPÍTULO I

<b>Tabela 1</b> - Quantidade de alelos e proteínas já descritos para os genes HLA de classe I clássicos e não clássicos.....	24
--	----

### CAPÍTULO II

<b>Table 1</b> - <i>HLA-A</i> 3'UTR haplotypes and associated coding alleles described at the IPD-IMGT/HLA database (version 3.30.0).....	71
<b>Table 2</b> - <i>HLA-A</i> 3'UTR haplotypes and associated coding allele groups found on a Brazilian population sample.....	72
<b>Table 3</b> - <i>HLA-C</i> 3'UTR haplotypes and associated coding alleles described at the IPD-IMGT/HLA database (version 3.30.0).....	73
<b>Table 4</b> - <i>HLA-C</i> 3'UTR haplotypes and associated coding allele groups found on a Brazilian population sample.....	75

## SIGLAS E ABREVIACES

**Clulas NK** – *Clulas Natural Killer*

**PAMPS** – *Patogen-Associated Molecular Patterns* - Padres Moleculares Associados a Patgenos

**PRRs** – *Pattern Recognition Receptors* - Receptores de Reconhecimento de Padro

**TLR** – *Toll-like Receptor* - Receptor Semelhante a Toll

**CD4/CD8** – *Cluster of differentiation*

**APC**- *Antigen-presentin Cell* – Clula apresentadora de Antgeno

**MHC** – *Major Histocompatibility Complex* – Complexo Principal de Histocompatibilidade

**HLA** – *Human Leucocyte Antigen* – Antgeno Leucocitrio Humano

**RE** – Retculo Endoplasmtico

**TAP** - *Transporter associated with Antigen Processing* - Protena transportadora associada ao processamento antignico

**KIR** - *Killer Immunoglobuline-like Receptor* – Receptor semelhante  Imunoglobulina de clulas Natural Killer

**TCR** – *T Cell Receptor* – Receptor de Clulas T

**DSA** – *Donor-specific Antibodies* – Anticorpos Especficos do Doador

**RNA** – *Ribonucleic Acid* – cido Ribonucleico

**mRNA** – *Messenger Ribonucleic Acid* – cido Ribonucleico mensageiro

**HPV** – *Human Papiloma Virus* – Papiloma Vrus Humano

**LD** – *Linkage Disequilibrium* – Desequilbrio de Ligao

**INF** - Interferon

**ISRE** - *IFN-stimulated response element* – Elemento Estimulado por Interferon

**3'NT/3'UTR** – 3' No Traduzida/3' *Untranslated Region*

**miRNA** - microRNA

**PAB** - *Poli(A) Binding Proteins* – Protenas de Ligao  Poli(A)

**RBPs** – *RNA Binding Proteins* – Protenas de Ligao ao RNA

**HER2** - *Human Epidermal growth factor Receptor-type 2* – Receptor do fator de crescimento Epidermal Humano tipo 2

**IRF-1** – *Interferon Regulatory Factor 1* – Fator Regulador de Interferon 1-

**DGCR8** - *DiGeorge syndrome chromosomal [critical] region 8* – Região 8 cromossomal [crítica] da Síndrome DiGeorge

**Xpo5** – *Exportin 5* – Exportina 5

**TRBP** - *Transactivation-responsive RNA Binding Protein* – Proteína de Ligação a RNA responsiva à transativação

**dNK** – *decidual Natural Killer* – Natural Killer decidual

**HIV** – *Human Immunodeficiency Virus* – Vírus da Imunodeficiência Human

**Ago2** – *Argonaut 2* – Argonauta 2

**RISC/miRISC** – *RNA-Induced Silencing Complex/microRNA-Induced Silencing Complex* – Complexo de Silenciamento Induzido por RNA/ Complexo de Silenciamento Induzido por microRNA

**MFE** – *Minimum Free Energy* – Energia Livre Mínima

## SUMÁRIO

### CAPÍTULO I

<b>CAPÍTULO I: REVISÃO DE LITERATURA</b>	15
<b>O Complexo Principal de Histocompatibilidade – processamento e apresentação de antígenos.....</b>	17
<b>Os genes HLA de classe I <i>HLA-A</i>, <i>HLA-C</i> e <i>HLA-G</i>.....</b>	23
<b><i>O gene HLA-A</i>.....</b>	30
<b><i>O gene HLA-C</i>.....</b>	34
<b><i>O gene HLA-G</i>.....</b>	36
<b>microRNAs e seu papel na regulação pós-transcricional do HLA de classe I.....</b>	37
<b>Métodos de detecção de alvos de microRNAs e uso do mirhunt.....</b>	41
<b>REFERÊNCIAS.....</b>	45
<b>JUSTIFICATIVA.....</b>	59
<b>OBJETIVOS.....</b>	61
<b>CAPÍTULO II: MANUSCRITO</b>	
<b>Abstract.....</b>	63
<b>Introduction.....</b>	64
<b>Methods.....</b>	66
<b>Results.....</b>	68
<b>Discussion.....</b>	83
<b>References.....</b>	90
<b>CONCLUSÃO.....</b>	97
<b>APÊNDICE.....</b>	99



# Capítulo I

## Revisão de Literatura



## CAPÍTULO I: REVISÃO DE LITERATURA

O Sistema Imunológico humano possui como função proteger o corpo contra os agentes infecciosos e o desenvolvimento de doenças ou qualquer outra ameaça capaz de causar lesão. Para isso, ele conta com as barreiras físicas, a pele e as mucosas, substâncias químicas e células capazes de reconhecer, com diferentes graus de especificidade, estruturas que não façam parte do organismo (tudo que é “não próprio”).

Com base em certos parâmetros, como tempo de ativação, células e mecanismos efetores, as respostas imunológicas podem ser divididas em duas classes: as respostas imunes inatas e as adaptativas. Enquanto as respostas inatas surgem pouco tempo após o estímulo nocivo e suas células reconhecem padrões moleculares característicos dos microrganismos, as respostas adaptativas levam mais tempo para acontecer, e apresentam um repertório de moléculas que garantem a sua especificidade ao patógeno em questão, sendo capazes de diferenciar as moléculas do hospedeiro de moléculas estranhas (Choo, 2007; Howell *et al.*, 2010; Abbas *et al.*, 2015).

Assim, a imunidade inata é tida como imediata e fornece a primeira linha de defesa do organismo, uma vez que é constituída pelas barreiras físicas e epitélios, células fagocíticas, proteínas circulantes (mediadores inflamatórios, sistema complemento, entre outros) e um tipo específico de linfócitos – as células *Natural Killer*, ou células *NK*. No caso de uma segunda exposição a determinado microrganismo, o mecanismo da resposta inata se mantém o mesmo da primeira exposição, não sendo capaz de distinguir pequenas diferenças entre microrganismos de um mesmo grupo. Por isso, ela é comumente chamada de resposta inespecífica (Murphy *et al.*, 2010; Abbas *et al.*, 2015). As células da resposta imune inata possuem em sua superfície receptores capazes de reconhecer padrões moleculares de diferentes tipos de microrganismos, conhecidos como Padrões Moleculares Associados a Patógenos (PAMPs, do inglês *Pathogen-Associated Molecular Pattern*). Por esse motivo, os receptores que reconhecem os PAMPs são chamados de Receptores de Reconhecimento Padrão (PRRs, do inglês

*Pattern Recognition Receptors*), entre os quais se destacam os receptores semelhantes a Toll (TLRs, do inglês *Toll-like Receptors*) (Murphy *et al.*, 2010).

Em contrapartida, a imunidade adaptativa é conhecida por seu alto grau de especificidade e capacidade de memória. Entre seus efetores podemos citar os linfócitos T e B e os anticorpos secretados por plasmócitos (linfócitos B ativados). A diferenciação entre as moléculas do hospedeiro e aquelas pertencentes a microrganismos que devem ser combatidos é feita por meio de moléculas especializadas de apresentação e reconhecimento de antígenos. Isso garante que mesmo microrganismos semelhantes entre si sejam reconhecidos de modo específico. As respostas adaptativas não são imediatas, mas em caso de exposição secundária, elas são rápidas e comumente exacerbadas (Murphy *et al.*, 2010; Abbas *et al.*, 2015). As respostas inatas podem ainda se dividir em celulares e humorais. As respostas celulares são mediadas pelas células T (CD4<sup>+</sup> e CD8<sup>+</sup>), enquanto a resposta humoral é mediada pelos anticorpos secretados pelas células B (Abbas *et al.*, 2015).

Para que uma resposta imune adaptativa seja deflagrada, o patógeno (quando extracelular) deve ser fagocitado por células capazes de processá-lo em peptídeos, conhecidas como células apresentadoras de antígenos (APCs, do inglês, *Antigen Presenting Cells*). Nas APCs, esses antígenos processados são acoplados a moléculas específicas – moléculas do Complexo Principal de Histocompatibilidade (MHC, do inglês *Major Histocompatibility Complex*) – que apresentarão esses antígenos aos linfócitos T, ativando-os (Abbas *et al.*, 2015). Em contrapartida, na presença de patógenos intracelulares, peptídeos antigênicos são produzidos no citosol por proteases que atuam em conjunto com o proteassoma. Esses antígenos são então transportados para o retículo endoplasmático, onde se juntam as moléculas de MHC em formação (Adiko *et al.*, 2015).

### **O Complexo Principal de Histocompatibilidade – processamento e apresentação de antígenos**

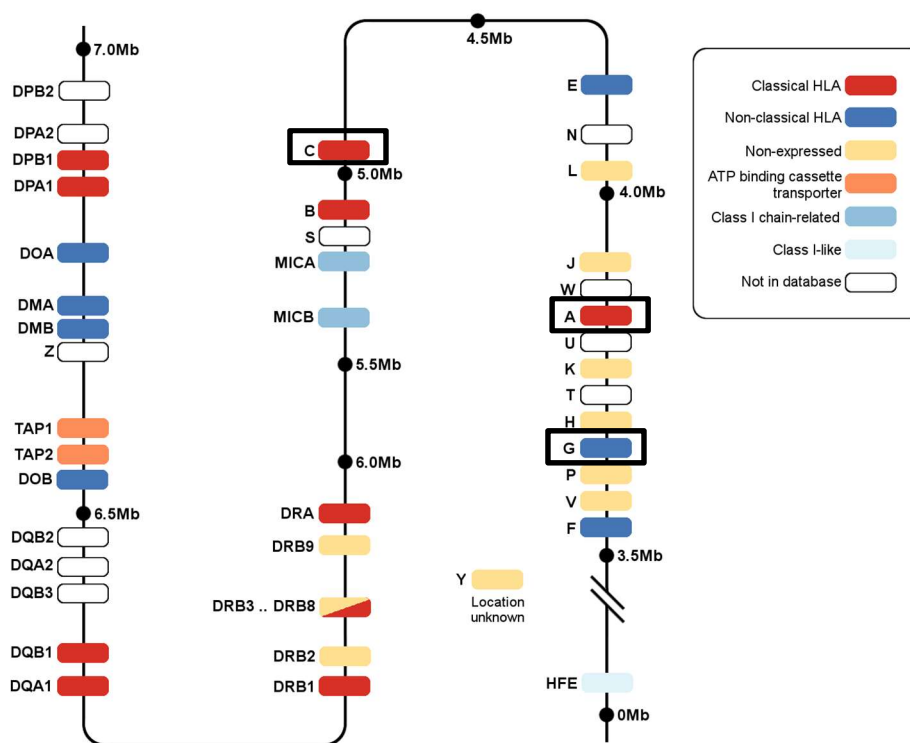
O MHC foi descoberto por meio de estudos de rejeição a transplantes em

camundongos ainda na década de 1930 e, por volta da década 1970, sabia-se que um grande complexo gênico era responsável pelo controle das respostas imunes – o Complexo Principal de Histocompatibilidade. Essa descoberta rendeu um prêmio Nobel em 1980 a George Davis Snell, Jean Baptiste Dausset e Baruj Benacerraf (Raju, 1999). Esses pesquisadores perceberam que enxertos de pele entre camundongos isogênicos não eram rejeitados, enquanto transplantes entre linhagens diferentes não eram aceitos pelo organismo do camundongo receptor.

Em humanos, essa descoberta foi feita a partir de três grupos de pacientes: (i) aqueles que haviam passado por múltiplas transfusões sanguíneas, (ii) pacientes que haviam passado por transplantes renais e (iii) mulheres múltiparas. Analisando o sangue desses pacientes, era possível encontrar, respectivamente, anticorpos que reconheciam as células sanguíneas e renais dos doadores e as células paternas. Por esse motivo, o MHC humano foi chamado de Antígenos Leucocitários Humanos (HLA, do inglês *Human Leucocyte Antigens*) (Terasaki, 2007; Thorsby, 2009; Abbas *et al.*, 2015).

Em humanos, os mais de 200 genes que compõem o MHC estão localizados no braço curto do cromossomo 6 (6p21.3), ocupando uma região que se estende por cerca de 3,5 megabases (Mb) (Figura 1) (Klein e Sato, 2000; Abbas *et al.*, 2015). Esses genes e as moléculas que eles codificam, bem como sua organização cromossômica, estão presentes em todos os vertebrados que possuem mandíbula, os gnatostomados, numa organização que vem sendo mantida conservada por mais de 500 milhões de anos (Klein e Sato, 1998). Certos genes desse complexo são particularmente variáveis e seus produtos são fundamentais ao desenvolvimento da resposta imunológica.

Com base em estrutura molecular e funções, esses genes foram convencionalmente divididos em classes I, II e III (Klein e Sato, 2000). Os genes que compõem as classes I e II também são conhecidos como Antígenos Leucocitários Humanos (HLA, do inglês *Human Leucocyte Antigens*), uma vez que atuam na apresentação antigênica a linfócitos. Os componentes do MHC de classe III incluem algumas citocinas e outros componentes das respostas imunológicas.



**Figura 1** – Organização do Complexo Principal de Histocompatibilidade em humanos. Os genes alvo desse trabalho estão em destaque.

Fonte: [http://hla.alleles.org/inc/images/hla\\_map\\_1000.png](http://hla.alleles.org/inc/images/hla_map_1000.png)

Os genes de classe I codificam as cadeias alfa das moléculas HLA. As cadeias alfa ligam-se não covalentemente a uma molécula de  $\beta$ 2-microglobulina, codificada por um gene localizado fora do complexo HLA, no cromossomo 15. A molécula HLA de classe I apresentam cinco domínios: dois domínios de ligação a peptídeos,  $\alpha$ 1 e  $\alpha$ 2, que juntos formam a fenda de ligação de peptídeo; um domínio de imunoglobulina (Ig),  $\alpha$ 3; um domínio transmembrana e a cauda citoplasmática (Figura 2). Dentro dessa classe, destacam-se os genes mais polimórficos do genoma humano e expressos em praticamente todas as células nucleadas – *HLA-A*, *HLA-B* e *HLA-C* (Klein e Sato, 2000; Murphy *et al.*, 2010; Abbas *et al.*, 2015). Já os genes de classe II codificam cadeias polipeptídicas  $\alpha$  e  $\beta$  e sua nomenclatura é derivada da estrutura da molécula e conta com o auxílio três letras que



os antígenos são mantidos na fenda de ligação a peptídeos por meio de interações não covalentes, entre a sequência do peptídeo e pontes de hidrogênio formadas entre aminoácidos conservados da molécula de HLA e aminoácidos do peptídeo (localizados nos finais da fenda) (Jensen, 2007).

A descoberta de que as moléculas do MHC de classe I e II eram necessárias para a ativação dos linfócitos – restrição por MHC – foi recompensada com um prêmio Nobel em 1996 a Peter C. Doherty e Rolf M. Zinkernagel (Adiko *et al.*, 2015). A partir de estudos nessa área, condicionou-se que as moléculas do MHC de classe I seriam responsáveis pela apresentação de antígenos de origem endógena (vírus, em especial), enquanto o MHC de classe II seria responsável por apresentar antígenos exógenos fagocitados. Porém, hoje é sabido que essa regra não é absoluta devido aos mecanismos de apresentação cruzada, onde antígenos exógenos são processados de modo a serem apresentados por moléculas de MHC de classe I e antígenos endógenos passam a ser apresentados por moléculas de MHC de classe II (Jensen, 2007; Adiko *et al.*, 2015; Brutkiewicz, 2016; Grotzke *et al.*, 2017).

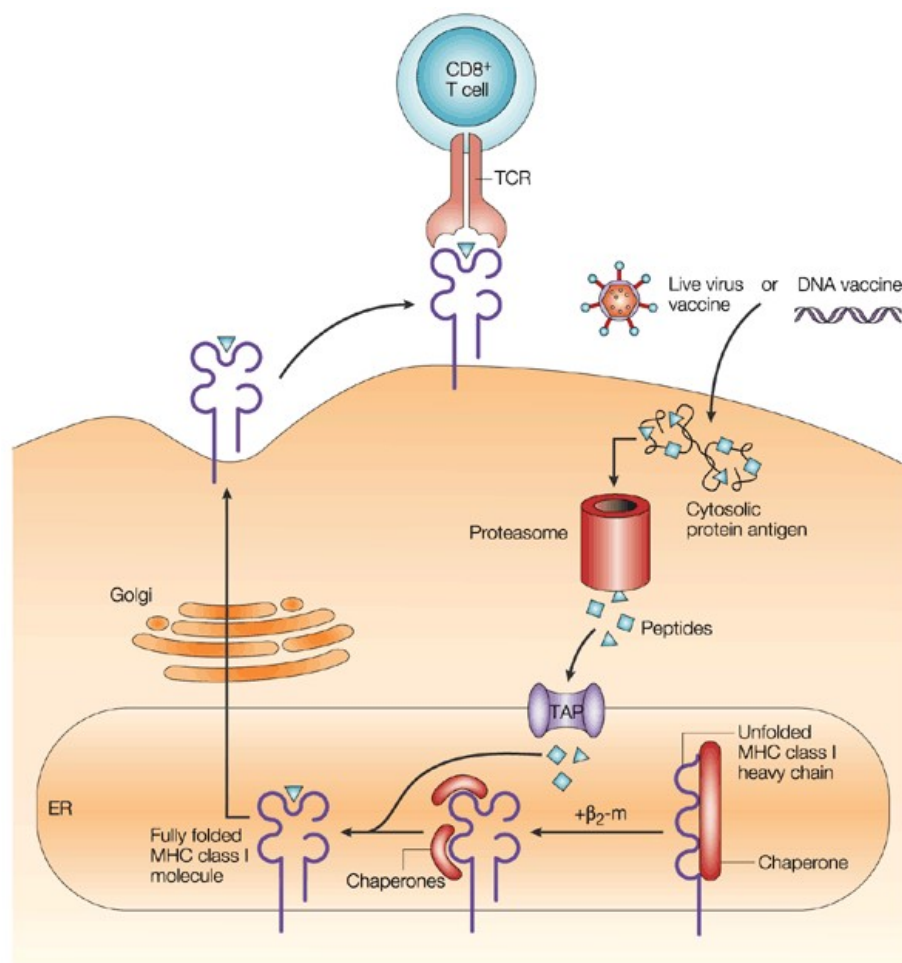
Os peptídeos que interagem com o HLA de classe I são convencionalmente gerados a partir do processamento de proteínas endógenas (normais ou alteradas) ou de patógenos (preferencialmente intracelulares) por meio da maquinaria do proteassomo e outras enzimas (Figura 3). Os proteassomos podem ser divididos a partir dos antígenos que eles processam. Quando o antígeno em questão é próprio, eles são chamados “proteassomos constitutivos”, enquanto aqueles que processam peptídeos não próprios são comumente chamados de “imunoproteassomos”. Tipos celulares diferentes possuem quantidades diferentes de cada tipo de proteassomo – células dendríticas são ricas em imunoproteassomos (Inaba e Inaba, 2005).

Após o processamento, os antígenos são transportados para o lúmen do Retículo Endoplasmático (RE) por meio da proteína transportadora associada ao processamento antigênico (TAP). Enquanto isso, no RE, a chaperona calnexina é responsável por moldar a cadeia pesada do MHC de classe I e por uni-la com a molécula de  $\beta$ 2-microglobulina. Quando a interação está estabelecida, a proteína tapasina mantém a ligação entre as duas (Figura 3). Somente quando a molécula

está completa – associada à  $\beta$ 2-microglobulina e com um antígeno em sua fenda – ela é transportada para a membrana pelo sistema de cisternas do Complexo de Golgi (Inaba e Inaba, 2005; Jensen, 2007). A montagem das moléculas de HLA corresponde a uma série de eventos finamente orquestrados que incluem um conjunto extenso de proteínas especializadas.

Apesar do MHC de classe I estar associado à apresentação de antígenos intracelulares, ocasionalmente, alguns antígenos extracelulares também podem ser apresentados aos linfócitos T CD8<sup>+</sup> por um processo chamado de apresentação cruzada. Este processo ocorre por meio de dois mecanismos distintos: um dependente de TAP e proteassomo (via citosólica) e um independente (via vacuolar) (Grotzke *et al.*, 2017). Na via citosólica, o antígeno internalizado por fagocitose precisa sair de sua vesícula e alcançar o citosol, onde ele é processado pelo proteassomo e transportado pela TAP. Esse mecanismo é mediado por uma proteína do retículo endoplasmático chamada Sec61 *translocon* (Ackerman *et al.*, 2006). Na via vacuolar, ocorre a endocitose do antígeno associado a uma proteína de membrana (como receptores Toll ou uma outra molécula de MHC), e o antígeno é degradado ainda no endossomo, sem a necessidade do proteassomo e, a protease catépsina S e o pH ácido endossomal participam desse mecanismo (Shen *et al.*, 2004).

Antígenos que se ligam ao HLA de classe II são, em geral, maiores que aqueles apresentados pelo HLA de classe I (limitados a 8-10 resíduos de aminoácidos) e, assim, as vias de processamento e acoplagem do peptídeo à fenda sofrem algumas diferenças. Como consequência disso, o processamento antigênico e sua interação com a fenda possuem um grande efeito na manutenção da estabilidade e conformação do domínio de ligação ao peptídeo (Jensen, 2007). As moléculas de MHC de classe II estão comumente associadas à apresentação de antígenos extracelulares e, por isso, sua expressão se limita às células apresentadoras de antígeno. Nessas células, os peptídeos antigênicos são gerados por proteólise nos endossomos e lisossomos.



Nature Reviews | Immunology

**Figura 3** – Processamento antigênico pelo proteassomo e montagem da molécula de MHC de classe I no Retículo Endoplasmático.

Fonte: Adaptado de Yewdell et al., 2003

### Os genes HLA de classe I *HLA-A*, *HLA-C* e *HLA-G*

Os genes que compõem o MHC de classe I em humanos podem ser classificados em clássicos e não clássicos (Beck *et al.*, 1999; Shiina *et al.*, 2009), quando levamos em consideração sua função na resposta imunológica e perfil de expressão. Os genes *HLA-A*, *HLA-B* e *HLA-C* são os chamados genes HLA de classe I clássicos. Eles são os genes mais variáveis de todo o genoma humano e juntos, contabilizam mais de 12.000 alelos descritos em um banco de dados



específico, o IPD-IMGT/HLA, disponível em <https://www.ebi.ac.uk/ipd/imgt/hla/> (Robinson *et al.*, 2015; Robinson *et al.*, 2016) (Tabela 1). As moléculas de classe I clássicas estão presentes em praticamente todas as células nucleadas do organismo e tem como função principal a apresentação antigênica aos linfócitos T CD8+. Por outro lado, os genes HLA de classe I não clássicos, *HLA-E*, *HLA-F* e *HLA-G*, são pouco variáveis, possuem expressão limitada a determinados tipos celulares (com exceção do gene *HLA-E*) e tem como função principal a imunomodulação – a apresentação antigênica aparece como uma função secundária e menos comum dessas moléculas (Wei e Orr, 1990; Geraghty *et al.*, 1992; Gobin e Van Den Elsen, 2000; Shobu *et al.*, 2006).

**Tabela 01** – Quantidade de alelos e proteínas já descritos para os genes HLA de classe I clássicos e não clássicos (dados disponíveis em IPD-IMGT/HLA Database – versão 3.34.0 <https://www.ebi.ac.uk/ipd/imgt/hla/stats.html>)

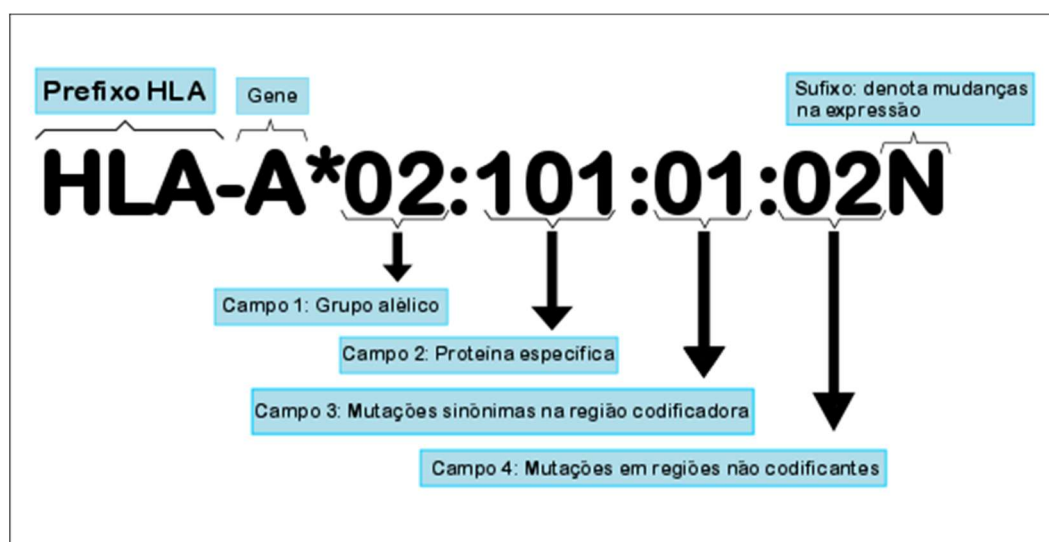
Gene	HLA-A	HLA-B	HLA-C	HLA-E	HLA-F	HLA-G
<b>Alelos</b>	4.638	5.590	4.374	27	31	61
<b>Proteínas</b>	3.172	3.923	2.920	8	6	19
<b>Alelos Nulos*</b>	224	169	171	1	0	3

\* apresentam mutações que impedem sua expressão na superfície celular

A grande quantidade de sequências descritas, em especial para os genes clássicos de classe I, exigiram que uma nomenclatura especial fosse desenvolvida para identificar os alelos já descritos e para facilitar a nomeação de novos alelos. Cada alelo é identificado por um prefixo (HLA) e o gene correspondente (A, B, C, E, F, G para os de classe I) separados por um hífen. Depois da identificação do gene, seguem-se quatro conjuntos de dígitos separados por dois pontos (:) (Figura 4).

O primeiro campo corresponde ao antígeno sorológico, representando um grupo de alelos que codificam moléculas com motivos semelhantes e que em geral são reconhecidos pelo mesmo anticorpo em um teste sorológico. O segundo campo determina subtipos dentro do mesmo grupo sorológico, gerados por

mutações não sinônimas (mudam um ou mais aminoácidos na proteína), e a numeração é designada na ordem em que as sequências foram descritas. O terceiro campo denota mutações sinônimas em éxons, que não mudam aminoácidos da proteína. O quarto e último campo designa mudanças nos segmentos que não fazem parte da proteína final (regiões regulatórias e íntrons). Além da numeração específica, os alelos também podem receber um sufixo que representam modos alternativos de expressão – N (alelo nulo, sem expressão na superfície celular), S (solúvel ou secretado), L (níveis baixos de expressão na superfície, derivado de “*Low*”), C (para alelos que ficam retidos no citoplasma), A (aberrante, quando há dúvidas quanto a expressão da proteína) e Q (questionável). Até o momento, nenhuma sequência foi descrita como C ou A (IPD-IMGT/HLA, disponível em <http://hla.alleles.org/nomenclature/naming.html>).



**Figura 04** – Esquema da nomenclatura dos alelos de HLA.

**Fonte:** Adaptado de IPD-IMGT/HLA (<http://hla.alleles.org/nomenclature/naming.html>)

As moléculas de MHC são expressas na superfície celular em codominância, ou seja, os dois alelos herdados dos pais são expressos. Assim, a presença de duas variantes alélicas aumentariam o repertório de antígenos apresentados. Além da codominância, observam-se altos níveis de heterozigose

para esse conjunto de genes e, uma vez que variantes proteicas diferentes apresentam capacidades diferentes de acoplamento e apresentação antigênica, maximiza-se a gama de antígenos a serem apresentados aos linfócitos T (Piertney e Oliver, 2006). As altas taxas de heterozigose também estão associadas à presença de seleção balanceadora atuando nesse locus.

Um outro fator que contribui para a grande quantidade de alelos dos genes clássicos é seleção mediada por parasitas, uma vez que as moléculas de MHC atuam diretamente no combate aos patógenos, a partir da apresentação de antígenos microbianos para as células T CD8<sup>+</sup>. Essa teoria ganha força quando observamos relações bem estabelecidas entre alelos específicos de MHC I e II e a suscetibilidade ou proteção a certas doenças humanas. Como exemplo, podemos citar o efeito protetor observado em pacientes com HIV portadores do alelo *HLA-B\*57:01*. A presença desse alelo está associada a menor viremia e maior contagem de células T CD4<sup>+</sup> (Leszczyszyn-Pynka *et al.*, 2015).

Algumas evidências apontam ainda a seleção sexual como sendo um dos fatores responsáveis por moldar o MHC até o modo como o conhecemos hoje (Edwards e Hedrick, 1998; Piertney e Oliver, 2006). Por mecanismos ainda não compreendidos, receptores sensoriais olfatórios específicos reconheceriam o odor de moléculas solúveis de MHC, de modo que os indivíduos são capazes de “cheirar a variabilidade” das moléculas de MHC do parceiro (Capittini *et al.*, 2008). Em humanos, esses receptores são codificados por genes localizados na região telomérica do complexo HLA no cromossomo 6, formando um conjunto de Receptores Olfatórios ligados ao HLA (Ehlers *et al.*, 2000). Esses receptores reconhecem os sinais químicos gerados pela ligação MHC/peptídeo e assim, reconhecem e diferenciam odores de familiares e possíveis parceiros não relacionados (Ehlers *et al.*, 2000; Capittini *et al.*, 2008). Nesse contexto, parceiros seriam escolhidos com base na diversidade entre o MHC de ambos de modo a ampliar a diversidade na prole e a produzir indivíduos heterozigotos capazes de combater uma ampla variedade de patógenos (Edwards e Hedrick, 1998; Capittini *et al.*, 2008).

Alguns estudos mostram também a co-evolução entre o MHC e seus

ligantes, KIR (*Killer Immunoglobuline-like Receptor*) e TCR (*T Cell Receptor*) (Rangarajan e Mariuzza, 2014; Guethlein *et al.*, 2015). Desse modo, TCR e MHC evoluíram juntamente para garantir a habilidade de interagirem entre si, garantindo a ativação eficiente de células T (Rangarajan e Mariuzza, 2014). Por outro lado, os receptores KIR interagem com as moléculas de MHC de classe I de modo alelo específico. Acredita-se que esse fato está associado a um melhor combate a infecções virais e ao sucesso reprodutivo, uma vez que células NK possuem um papel importante nas respostas imunológicas desses dois processos.

Embora, em geral, ambos os alelos de um indivíduo são expressos (co-dominância), não está claro se há expressão alélica diferencial, i.e., se alguns alelos expressam quantidades maiores ou menores da proteína em comparação aos outros. Uma vez que as moléculas HLA de classe I possuem um papel fundamental no desenvolvimento das respostas imunológicas, em especial na ativação de células T citotóxicas, a expressão inadequada dessas moléculas está intimamente associada ao desenvolvimento de certas doenças. Como exemplo, podemos citar certas infecções virais e o desenvolvimento de neoplasias, onde o principal mecanismo de evasão ao sistema imune adaptativo é redução da expressão de moléculas de HLA na superfície celular tumoral por mecanismos diversos, que envolvem desde o comprometimento das moléculas responsáveis pela montagem das moléculas de HLA à redução dos níveis de expressão gênica (Peltenburg *et al.*, 1993; Odeberg *et al.*, 2002; Shen *et al.*, 2005; Tang *et al.*, 2009; Apps *et al.*, 2016).

Em transplantes, a compatibilidade entre o HLA do doador e do receptor é avaliada para pelo menos cinco genes, os três genes clássicos de classe I (*HLA-A*, *HLA-B* e *HLA-C*) e dois de classe II (*HLA-DQB1* e *HLA-DRB1*). A compatibilidade entre os cinco *loci* está associada a uma maior taxa de sobrevivência do enxerto e do paciente. Porém, quando ela não é possível, a incompatibilidade entre algum desses genes resulta em taxas diversas de aceitação do enxerto a depender do gene incompatível e do órgão em questão. A rejeição mediada por respostas imunes adaptativas ocorre quando antígenos do doador são apresentados às células T CD4<sup>+</sup> ou CD8<sup>+</sup> do receptor por células dendríticas deste

(reconhecimento indireto), ou mesmo do doador (reconhecimento direto). Mudanças estruturais entre as moléculas de MHC do doador e do receptor também podem estimular respostas imunes por contato direto com o TCR do doador (Moreau *et al.*, 2013). Assim, quanto maior a compatibilidade entre os dois, melhor o resultado do transplante.

Um dos genes mais considerados quando a compatibilidade completa não é possível é o *HLA-C*, uma vez que ele parece ser expresso em níveis mais baixos na superfície celular. No entanto, os resultados encontrados são conflitantes – em transplantados renais, anticorpos doador-específico (DSA) anti-*HLA-C* são comumente encontrados em pacientes cujos enxertos são rejeitados (Duquesnoy e Marrari, 2011) ao mesmo tempo que de 15 a 30% dos transplantes de células tronco hematopoiéticas realizados apresentam incompatibilidade entre os alelos de *HLA-C* do doador e do receptor (Tiercy, 2014).

A eficiência das respostas imunes e o desfecho dos processos patológicos estão intimamente associados aos níveis de expressão das moléculas HLA. O aumento da expressão dessas moléculas resulta na melhora da eficácia da apresentação antigênica e conseqüentemente das respostas imunes (Faroudi *et al.*, 2003). Duas regiões gênicas estão intimamente relacionadas com os níveis de expressão de um determinado gene – a região promotora e a região 3' não traduzida (3'NT).

Na região promotora encontram-se os motivos gênicos responsáveis pela ligação dos fatores de transcrição e outras proteínas que influenciam na ligação e funcionamento do complexo da RNA polimerase e, assim, dando início à transcrição do RNA (Juven-Gershon *et al.*, 2008). Desse modo, a região promotora é fundamental no controle dos níveis de RNA.

Os genes HLA de classe I contam com uma série de reguladores transcricionais em seus promotores que garantem sua expressão, uma vez que eles devem estar presentes na superfície da grande maioria das células do organismo. Juntamente com os genes que codificam a cadeia pesada da molécula de HLA, é preciso também que haja a expressão coordenada de proteínas associadas ao processamento antigênico e montagem da molécula HLA como um todo, como as

proteínas do proteossoma, bem como da molécula de  $\beta_2$ -microglobulina. Alguns vírus, como o Papilomavirus Humano (HPV), são capazes de modular a expressão desses genes, reduzindo a produção das moléculas de HLA e facilitando seu escape do sistema imunológico. Para isso, proteínas virais, como a proteína E7 do HPV-16 e do HPV-6a são capazes de reprimir a transcrição das cadeias pesadas de HLA e da proteína TAP por meio de interações com as regiões promotoras desses genes (Georgopoulos *et al.*, 2000).

Apesar da expressão constitutiva desses genes, algumas moléculas específicas são capazes de aumentar os níveis de HLA de classe I, mesmo em linhagens celulares onde HLA é pouco expresso. Isso pode ser observado na presença dos Interferons (IFNs) de tipo 1 ( $\alpha$  e  $\beta$ ) e tipo 2 ( $\gamma$ ). Essa citocina é capaz de estimular a transcrição tanto dos genes clássicos quanto não clássicos do HLA de classe I devido à presença de uma sequência conhecida como Elemento de Resposta Estimulado por IFN ou ISRE (do inglês, *IFN-stimulated response element*) (Chen *et al.*, 1986; Browne *et al.*, 2006) no promotor dos genes HLA. O IFN- $\gamma$  é capaz de melhorar a resposta citotóxica contra tumores mediada por células T, por estimular a expressão dos genes clássicos do MHC de classe I. No entanto, um aumento da expressão de HLA-E, por exemplo, pode resultar na inibição da citotoxicidade mediada por células NK (Browne *et al.*, 2006).

A região 3'NT, por sua vez, tem um papel muito importante na manutenção da estabilidade e estrutura secundária da molécula de mRNA (Kuersten e Goodwin, 2003). Essa região é ainda a principal fonte de sequências-alvo para microRNAs (miRNAs), responsáveis pela redução do nível proteico via inibição da tradução e, assim, atuando como um importante regulador pós-transcricional (discutido adiante) (Bartel, 2009; Pasquinelli, 2012).

Em contrapartida, a região 3'NT de uma molécula de mRNA recém-sintetizada possui um papel importante no processo de tradução. Quando uma nova molécula de mRNA é sintetizada, uma sequência rica em Adenina é adicionada ao final 3' da molécula, formando a chamada cauda poli(A). Proteínas de Ligação a Poli(A) (PAB, do inglês *Poli(A) Binding protein*) ligam-se à cauda poliA, e interagem com complexos proteicos responsáveis por recrutar as

subunidades ribossomais para a ponta 5' do mRNA como o grupo de iniciadores da tradução eIF4F (Yamashita e Takeuchi, 2017).

Desse modo, podemos concluir que a região 3'NT de uma molécula de mRNA atua de modo homólogo às regiões promotoras dos genes: fatores (especialmente proteínas de ligação a RNA, além de miRNAs) se ligam a sequências específicas e orquestram a regulação pós-transcricional, determinando a quantidade de proteína gerada por essa molécula de mRNA. Em humanos, mais 1000 proteínas de ligação a RNA (RBPs, do inglês *RNA Binding Proteins*) (Gerstberger *et al.*, 2014) e mais de 2500 miRNAs (Kozomara e Griffiths-Jones, 2014) já foram descritos. Esses dois elementos podem atuar de modo cooperativo ou competitivo, mas estudos que tentam mostrar essa relação em escala de transcriptoma ainda são escassos (Preusse *et al.*, 2015; Hafezqorani *et al.*, 2016). Nesse contexto, mudanças na sequência de 3'NT de um gene são capazes de alterar os motivos de ligação de RBPs e miRNAs, alterando os mecanismos envolvidos na estabilização ou degradação da molécula de mRNA e tradução.

### ***O gene HLA-A***

O gene *HLA-A* é o segundo mais polimórfico do genoma humano. Ele é composto por oito éxons e sete íntrons. O éxon 1 é responsável por codificar o peptídeo líder, que não faz parte da estrutura proteica final, mas comumente se liga à fenda de ligação ao peptídeo de moléculas de HLA não-clássicas, como *HLA-E*. Os éxons dois e três são responsáveis pela codificação das cadeias  $\alpha 1$  e  $\alpha 2$ , que incluem a fenda de ligação ao peptídeo e são altamente variáveis (Malissen *et al.*, 1982). A cadeia  $\alpha 3$  é codificada pelo éxon 4 e, diferente das outras duas, possui uma cadeia de aminoácidos mais conservada. Da mesma forma, as regiões transmembrana (éxon 5) e citoplasmática (éxons 6 e 7) também são mais conservadas (Tragardh *et al.*, 1979). A conservação do domínio  $\alpha 3$  pode estar associada a sua função como ligante do co-receptor CD8 e, portanto, mudanças em sua sequência seriam capazes de afetar essa interação (Salter *et al.*, 1989; Hanvesakul *et al.*, 2007).

No domínio citoplasmático, os alelos de *HLA-A* apresentam uma cisteína

não pareada (não relacionada a uma ligação dissulfeto) no resíduo 339 que parece estar envolvida na reciclagem da molécula, sua degradação e ainda no reconhecimento de receptores de células NK e formação de dímeros de MHC de classe I em exossomos (Makhadiyeva *et al.*, 2012). Assim, variações na sequência desse gene apresenta repercussões funcionais variadas, a depender da região em questão.

A região mais bem caracterizada para o gene *HLA-A* compreende os éxons dois e três, pois ambos codificam o domínio de ligação ao peptídeo. Grande parte dos estudos populacionais focaram-se nesse segmento, e pouco se sabe sobre as regiões regulatórias – promotor e região 3'NT – apesar de seu papel na manutenção dos níveis de mRNA e proteicos (Kuersten e Goodwin, 2003). Uma investigação recente mostrou o papel de polimorfismos da região 3'NT do gene *HLA-A* na ligação de certos miRNAs capazes de reduzir sua expressão em nível pós-transcricional. Por exemplo, a presença de uma Adenina na posição +2969 (rs1061235) faz com que pelo menos cinco miRNAs deixem de atuar nessa região (miR526b, miR609, miR1290, miR342-5p, miR542-5p). Entretanto, ao mesmo tempo que uma variação elimina a possibilidade de ligação de certos miRNAs, essa mesma variação cria alvos para outros (miR-520f e miR-651) (Rene *et al.*, 2015). É preciso considerar, ainda, que a expressão de HLA-A deve ocorrer de modo constitutivo em praticamente todas as células humanas e, por isso, a regulação dos seus níveis de expressão deve ser finamente regulada independentemente do microambiente celular.

A molécula codificada por esse gene interage tanto com o receptor TCR (Rock e Benacerraf, 1983; Chouaib *et al.*, 1988) quanto com alguns receptores KIR, mas esta última ocorre de um modo alelo-específico (Thielens *et al.*, 2012; Shen *et al.*, 2014). Os alelos de *HLA-A* que possuem o motivo sorológico Bw4 (alelos dos grupos *HLA-A\*03* e *HLA-A\*11*) interagem com o receptor inibitório KIR3DL2 e KIR2DL3 (Jamil e Khakoo, 2011) e alelos dos grupo *HLA-A\*11* também são capaz de interagir com os receptores ativadores KIR2DS2 e KIR2DS4 (Graef *et al.*, 2009; Liu *et al.*, 2014). Essa interação entre KIR e certos alelos de *HLA-A* parece ser a razão pela qual alelos *HLA-A* que carregam o motivo Bw4



estarem associados ao desenvolvimento de hepatite B crônica, em especial na presença de alelos de *HLA-C* do alótipo C2 (discutido adiante). Nesse contexto, KIR2DL3, ao interagir com HLA-A Bw4 estaria inibindo a resposta citotóxica mediadas por células NK (Di Bona *et al.*, 2017).

Uma vez que essas interações alelo-específicas com receptores KIR ficam à mercê da frequência desses alelos de *HLA-A* na população, a molécula de HLA-A possui como principal ligante o TCR de células T CD8<sup>+</sup>, atuando diretamente na ativação de respostas imunes adaptativas. Desse modo, a capacidade de apresentação antigênica das moléculas HLA-A possui um importante papel em infecções virais. Além disso, vários alelos de *HLA-A* associam-se à uma série de contextos patológicos, como doenças autoimunes, doenças infecciosas, neoplasias e o resultado de transplantes (Ogahara *et al.*, 1998; Silva *et al.*, 2009; Vandiedonck *et al.*, 2009; Mosaad *et al.*, 2010).

Estudos que avaliam a afinidade das moléculas de HLA aos peptídeos a serem apresentados (tanto próprios quanto não próprios), mostram a HLA-A interagindo preferencialmente a motivos conservados de proteínas humanas e de vírus que apresentam DNA como material genético (Hertz *et al.*, 2011). HLA-A parece mostrar afinidade especial a trechos conservados de proteínas virais dos *Herpesviridae* e *Adenoviridae*, e assim, sendo fundamental na formação de células T CD8<sup>+</sup> específicas. De fato, epítomos de Herpesvírus Humano 6B (HHV-6B) são capazes de gerar expansão clonal em células T CD8<sup>+</sup> e seus peptídeos são associados aos grupos alélicos *HLA-A\*02* e *HLA-A\*03* (os mais frequentes em indivíduos caucasianos e com altas frequências – superiores a 40% - em várias populações do mundo) (Iampietro *et al.*, 2014).

Em células neoplásicas, várias vias moleculares estão alteradas, e muitas vezes, o mecanismo de escape à imunovigilância se dá pela regulação negativa da expressão do HLA de classe I. Nesse contexto, Herrmann e colaboradores (2004) demonstraram que a glicoproteína HER2 (super-expressa em até 30% dos cânceres mamários) é capaz de reduzir a expressão de HLA de classe I *in vitro*, mas que esse efeito pôde ser revertido após tratamento com IFN- $\gamma$  (Herrmann *et al.*, 2004). Sabe-se que essa regulação é mediada pela via da MAPK, ativada por

HER2, mas o mecanismo exato dessa regulação ainda é desconhecido (Mimura *et al.*, 2013).

Em contraste, no carcinoma hepatocelular, não há redução da expressão de HLA de classe I e ainda mostram níveis elevados de HLA-A, mediada pelo aumento da expressão dos fatores de transcrição IRF-1 e p56, que regulam a expressão de *HLA-A* ao interagirem com os motivos ISRE e Enhancer A. Porém, não foi observada relação entre a quantidade de HLA-A e o tamanho do tumor, sua progressão e relação com a infecção por hepatite B ou C e cirrose (Shen *et al.*, 2009).

O gene *HLA-A* está cercado por genes de classe I não clássicos – *HLA-E*, *HLA-F*, *HLA-G*, além do pseudogene *HLA-H*. Acredita-se que uma possível associação entre esses genes atue na indução de alorreatividade ou de imunotolerância, uma vez que moléculas não clássicas possuem importantes papéis imunomodulatórios (Carlini *et al.*, 2016). Os estudos envolvendo *mismatches* de *HLA-A* em transplantes têm gerado resultados órgão-dependentes e alelo-específicos. Em transplantes hepáticos, a compatibilidade entre HLA-A está associada a melhores resultados com impacto na sobrevivência do enxerto (Muro *et al.*, 2012). Transplantes renais entre receptores A\*02 e doadores A\*68 ou A\*69 (sorotipo A\*28) mostram pouca produção de anticorpos e boas taxas de sobrevivência do enxerto (Dankers *et al.*, 2003). Assim, acredita-se que a associação entre *HLA-A* e outros genes não clássicos possa atuar na indução de alorreatividade ou tolerância ao órgão transplantado (Carlini *et al.*, 2016).

Anteriormente, já foram observados altos níveis de Desequilíbrio de Ligação (LD) entre alelos de *HLA-A* com alelos de *HLA-E*, de modo que alelos desses dois loci se encontram totalmente associados (*HLA-A\*01* e *HLA-E\*01:01*; *HLA-A\*23*, *\*24* e *\*29* e *HLA-E\*01:03*) (Geraghty *et al.*, 1992). O mesmo também já foi observado entre alelos de *HLA-A* e *HLA-G* (Ober *et al.*, 1996; Kolte *et al.*, 2010). Recentemente, alto LD foi observado entre *HLA-A*, *HLA-H* e *HLA-G* (em especial, alelos de região 3' não traduzida – 3'NT), mas não com *HLA-E*, o que é possivelmente explicado pela distância física entre este *locus* e outros genes (Carlini *et al.*, 2016).

### ***O gene HLA-C***

O gene *HLA-C*, quando comparado aos seus homólogos *-A* e *-B*, apresenta a mesma estrutura molecular e a mesma função primária de apresentação antigênica. Porém, levando os três genes em consideração, ele é o menos variável, mesmo contando com mais de 4.300 alelos, e costuma ser expresso em menores níveis (Mccutcheon *et al.*, 1995). Os mecanismos relacionados com essa menor expressão envolvem a meia vida mais curta do mRNA de *HLA-C*, responsável por menores níveis de tradução das cadeias pesadas (Mccutcheon *et al.*, 1995), a retenção da molécula no RE e sua eventual degradação (Neisig *et al.*, 1998).

Mesmo possuindo a mesma propriedade de apresentação antigênica e ser altamente polimórfico, o *HLA-C* é o único gene HLA clássico a ser amplamente expresso em tecidos placentários. Acredita-se que isso ocorra devido ao seu importante papel como modulador de células NK (King *et al.*, 2000), uma vez que o sucesso da implantação depende de uma resposta inflamatória leve (mediada pelas células dNK, pouco citotóxicas) que logo deve ser controlada e inibida (Male *et al.*, 2011). Juntamente com a *HLA-C*, as moléculas *HLA-E* e *HLA-G* (ambas com propriedades de imunomodulação e inibição de célula NK) constituem a expressão de MHC de classe I na placenta. Assim, esses três *loci* estão associados ao sucesso da gestação e à manutenção da tolerância materno-fetal (Varla-Leftherioti, 2004; Tripathi *et al.*, 2006; Betz, 2012; Klitkou *et al.*, 2015).

Uma vez que durante a gestação, as principais células de origem linfóide presentes no útero são as células NK decíduais (dNK), as moléculas *HLA-C*, *HLA-E* e *HLA-G*, possuem um papel fundamental na orquestração dos eventos necessários ao sucesso da implantação e da não rejeição materna ao feto (Parham e Guethlein, 2010). As moléculas *HLA-C* que interagem com os receptores KIR maternos podem ser divididas em dois grupos, devido a diferenças em aminoácidos fundamentais à interação – *HLA-C1*, que contém uma asparagina na posição 80 (ou 80N) e *HLA-C2*, que apresentam uma lisina na posição 80 (ou 80K) (The IPD-IMGT/HLA database, disponível em

<https://www.ebi.ac.uk/ipd/kir/ligand.html>). As moléculas HLA-C1 interagem com os receptores KIR2DL2, KIR2DL3 e KIR2DS2, sendo os dois primeiros capazes de inibir a resposta mediada por células NK, e o último capaz de ativá-las. As moléculas C2, por sua vez, são específicas para os receptores KIR2DL1, KIR2DS1 e KIR2DS2 (Vilches e Parham, 2002; Deng e Mariuzza, 2006). Além de atuar em receptores de células NK, a molécula HLA-C também desempenha um papel importante na apresentação antigênica para células T CD8<sup>+</sup> e na deflagração de respostas imunes celulares.

Por atuar tanto na resposta imune inata quanto na adaptativa e por seu papel de imunomodulação, o gene *HLA-C* tem sido bastante estudado em uma série de contextos patológicos. Entre eles podemos destacar as complicações gestacionais, como eclampsia e abortos de recorrência e espontâneos (Flores *et al.*, 2007; Yu *et al.*, 2014), e as infecções virais, com atenção especial dada ao vírus HIV (do inglês, *Human Immunodeficiency Virus*) (Qi *et al.*, 2006; Kulkarni *et al.*, 2013; Apps *et al.*, 2016). Algumas variantes de *HLA-C* têm sido associadas à infecção por HIV e desenvolvimento da doença, entre elas, uma deleção na região 3'NT que influencia o alvo do miRNA miR-148a (Kulkarni *et al.*, 2011). Níveis mais elevados de HLA-C se associam a uma maior atividade de linfócitos T citotóxicos, mas ao mesmo tempo uma maior inibição de células NK (a depender do haplótipo do hospedeiro) (Alcami e Koszinowski, 2000).

Nesse contexto, um dos mecanismos de escape usado pelo vírus HIV é a redução dos níveis de HLA-A e HLA-B (mas não de HLA-C) pela proteína Nef. Desse modo, células T CD8<sup>+</sup> deixam de ser ativadas ao mesmo tempo que células NK seriam moduladas. Em contrapartida, a proteína Vpu do HIV é capaz de reduzir a expressão de HLA-C, mantendo sob controle a resposta imune mediada por células T citotóxicas (Apps *et al.*, 2016). Altos níveis de HLA-C em infecção por HIV estão associados a maiores respostas de células T ativadas por esta molécula, porém, observam-se também maiores taxas de mutação viral, como uma tentativa do vírus de escapar à vigilância imunológica (Apps *et al.*, 2013). Assim, a regulação dos níveis de HLA-C (e outras moléculas HLA) é muito importante na determinação do resultado de uma infecção.

No que diz respeito à importância de HLA-C em transplantes, não é recomendável que o enxerto possua incompatibilidade entre nenhuma das moléculas comparadas, mas uma vez que entre os três loci clássicos o *HLA-C* é o menos expresso, alguns autores demonstram a possibilidade de transplantes bem-sucedidos mesmo quando as moléculas de HLA-C são incompatíveis (Petersdorf *et al.*, 2014; Thus *et al.*, 2014).

### ***O gene HLA-G***

O gene *HLA-G* é um HLA de classe I não clássico, pouco variável e com expressão restrita a poucos tecidos, como córnea, pâncreas e tecidos fetais (Ellis *et al.*, 1986; Le Discorde *et al.*, 2003). Até o momento, somente 61 alelos, responsáveis por codificar 19 variantes proteicas foram descritos (IPD-IMGT), porém apenas 5 delas são encontradas frequentemente (HLA-G\*01:01, G\*01:02, G\*01:03, G\*01:04 e G\*01:06) (Castelli, Ramalho, *et al.*, 2014; Castelli *et al.*, 2017). A presença de HLA-G foi primeiramente descrita no trofoblasto extraviloso e em linhagens celulares de coriocarcinoma (Ellis *et al.*, 1986). Sua estrutura molecular é semelhante à das moléculas de HLA clássicas, mas com uma redução da cauda citoplasmática devido à presença de códon de parada prematuro localizado no éxon 6 (Geraghty *et al.*, 1987).

Diferentemente dos HLA clássicos, a apresentação antigênica é uma função secundária das moléculas não-clássicas. Sua função primária é a imunomodulação, geralmente mediada por inibição da atividade de células T CD8<sup>+</sup>, via receptores ILT2 e ILT4 (receptores semelhantes à imunoglobulina transcritos 2 e 4) e de células NK. A molécula de HLA-G, em particular, interage com o receptor KIR2DL4 de células NK. Este, apesar da nomenclatura indicar ser um receptor inibitório (cauda citoplasmática longa devido à presença de motivos de tirosina inibitórios), sua interação com HLA-G resulta na indução da produção de citocinas, mas inibe respostas de citotoxicidade (Khalil-Daher *et al.*, 1999; Shiroishi *et al.*, 2003; Favier *et al.*, 2010).

Devido seu papel na imunomodulação e sua presença em tecidos fetais, o

HLA-G possui um papel importante no desenvolvimento da tolerância materno-fetal, na aceitação de tecidos transplantados e ainda no escape à vigilância imunológica por tumores (Adrian Cabestre *et al.*, 1999; Rouas-Freiss *et al.*, 1999; Carosella *et al.*, 2003).

Durante a gravidez, HLA-C, HLA-E e HLA-G participam dos fenômenos imunológicos gestacionais, da implantação ao parto (Gregori *et al.*, 2015; Hackmon *et al.*, 2017). Isso ocorre na interface materno-fetal, onde os tecidos fetais, mais especificamente o trofoblasto, estão em contato direto com leucócitos, estroma uterino e células endoteliais maternas que compreendem a decídua. A expressão de HLA-G na superfície do trofoblasto extraviloso ocorre primariamente durante o primeiro trimestre de gestação e, após este período, ela se limita ao citoplasma dessas células. No entanto, ele permanece presente nas membranas fetais amniótica e coriônica durante toda a gravidez (Hackmon *et al.*, 2017).

O *HLA-G* é o gene HLA cujas regiões regulatórias são mais bem estudadas quanto a estrutura e função (Castelli *et al.*, 2010; Donadi *et al.*, 2011; Sabbagh *et al.*, 2013; Castelli, Veiga-Castelli, *et al.*, 2014). Sua região 3'NT está associada à manutenção da estabilidade da molécula de mRNA (Rousseau *et al.*, 2003) e à regulação pós-transcricional dos níveis de HLA-G por meio da ligação de miRNAs (Tan *et al.*, 2007; Manaster *et al.*, 2012; Porto *et al.*, 2015).

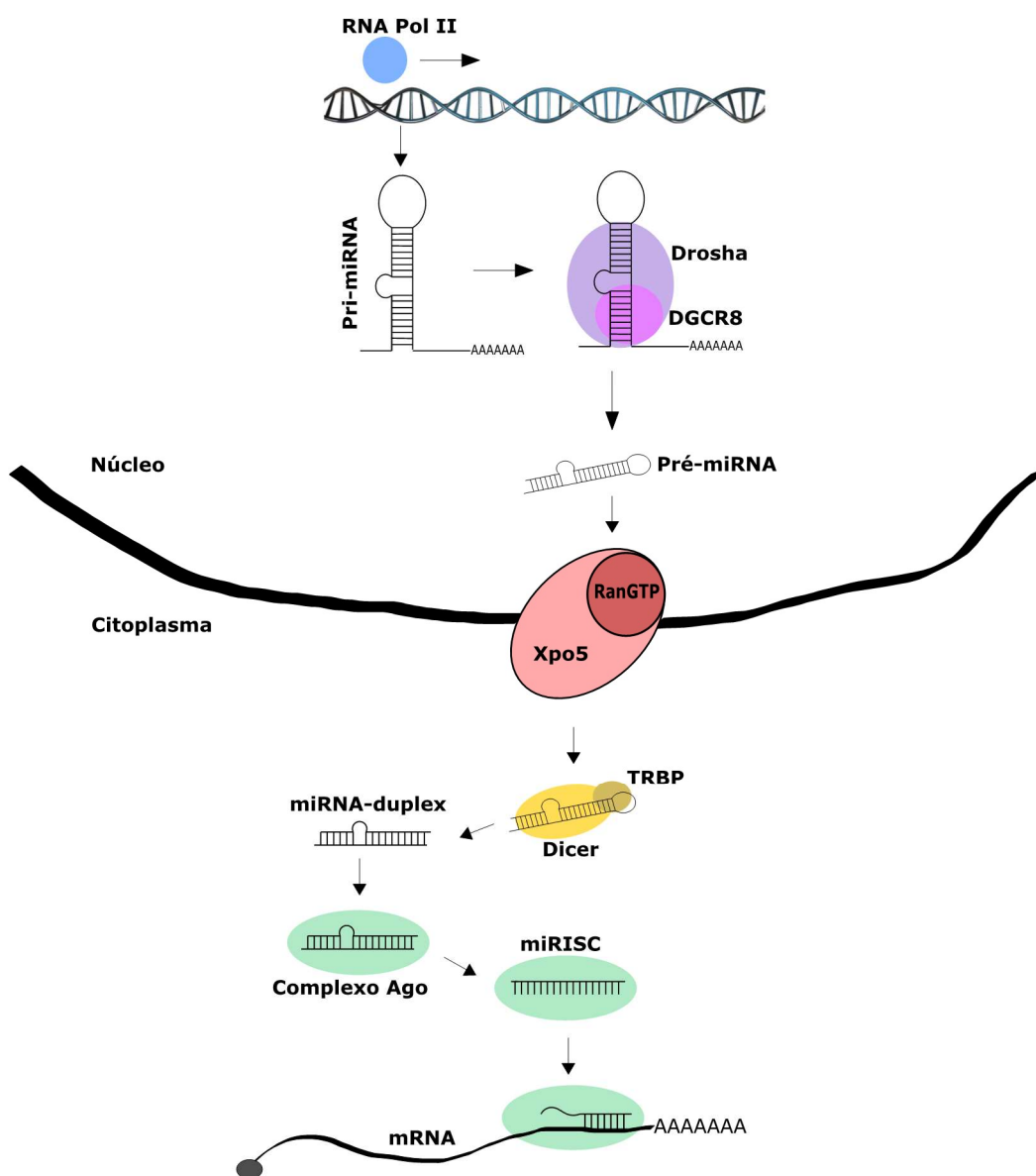
### **microRNAs e seu papel na regulação pós-transcricional do HLA de classe I**

Os miRNAs são pequenos RNA não codificantes (21-24 nucleotídeos), que atuam no mRNA alvo a partir da complementaridade de bases com sua sequência alvo. Geralmente eles são transcritos a partir de genes próprios, mas cerca de 30% deles são derivados de regiões intrônicas ou até mesmo de éxons de outros genes (Bartel, 2009). Essas pequenas moléculas de RNA já foram descritas em uma grande variedade de organismos – humanos, algas, insetos, plantas, mamíferos, protozoários, entre outros (Griffiths-Jones *et al.*, 2006; Kozomara e Griffiths-Jones, 2014). Sugere-se que miRNAs cuja transcrição ocorra dentro de “genes

hospedeiros” podem ter sua função relacionada à regulação desses mesmos genes ou genes relacionados, mas os resultados até o momento são conflitantes (Barik, 2008; Bell *et al.*, 2010).

Os miRNAs são transcritos pela RNA polimerase II e seus transcritos primários, os pri-miRNAs, podem chegar a 1kb formando estruturas de grampo. Eles podem gerar uma única molécula ou um *cluster* de miRNAs. O pri-miRNA é clivado pelo complexo proteico Drosha, com auxílio da proteína DGCR8, responsável pela ligação ao pri-miRNA e ativação da RNase III contida na Drosha. Assim, forma-se uma segunda estrutura em grampo, porém menor, contendo 70-100 nucleotídeos, o pré-miRNA. Este é transportado para o citoplasma pela proteína Exportina5 (Xpo5), que forma um complexo de transporte quando complexada à proteína nuclear Ran GTPase (RanGTP). No citoplasma, o pré-miRNA sofre uma segunda clivagem mediada pela enzima Dicer, e finalmente gera o miRNA maduro em fita dupla (Shukla *et al.*, 2011; Dong *et al.*, 2013). A precisão de clivagem da Dicer é aumentada na presença de um cofator, proteína TRBP (*Transactivation-responsive RNA Binding Protein*), que também atua na estabilização da Dicer e na determinação da fita de miRNA a interagir com a proteína Argonata, um dos componentes do RISC (Wilson *et al.*, 2015). O miRNA maduro (~ 22nt) e funcional deve ainda se soltar da sua fita complementar e se juntar ao Complexo de Silenciamento Induzido por RNA, ou RISC. O complexo miRNA/RISC interage geralmente com as regiões 3'NT do mRNA alvo, guiado pela sequência do miRNA (Dong *et al.*, 2013; Daugaard e Hansen, 2017). Um esquema da biogênese dos miRNAs está representado na Figura 5.

A interação miRNA/MRNA se baseia na complementaridade de bases entre as duas moléculas. Esta complementaridade pode ser perfeita (comumente observada em plantas) ou imperfeita. Desse modo, a sequência de RNA carregada no miRISC serve de “guia” até o mRNA alvo, no qual os resultados da interação incluem repressão da maquinaria de tradução em diferentes estágios do processo de tradução, deadenilação do mRNA e sua degradação (Bartel, 2009; Huntzinger e Izaurralde, 2011; Eichhorn *et al.*, 2014).



**Figura 5-** Esquema da biogênese dos microRNAs.

Fonte: Adaptado de Daugaard & Hansen, 2017

Em animais, o pareamento imperfeito entre miRNAs e seus alvos é suficiente para desempenhar seu papel na regulação pós-transcricional. Até pouco tempo atrás, por conta dessas diferenças entre níveis de pareamento, acreditava-se que o controle pós-transcricional por clivagem e degradação do mRNA via miRISC era um mecanismo exclusivo de plantas, enquanto em animais os miRNAs atuavam somente interferindo na maquinaria de tradução. Hoje, sabe-se



que essas duas vias de controle pós-transcricional mediada por miRNAs atuam tanto em animais quanto plantas (Huntzinger e Izaurralde, 2011). Uma das vias de inibição da tradução via miRISC tem como alvo o processo de reconhecimento do “Cap” presente na ponta 5’ do mRNA, impedindo, assim o início da tradução (Mathonnet *et al.*, 2007). A proteína Argonauta 2 (Ago2), se liga diretamente ao *cap* e reprime o estágio inicial da tradução. A degradação do mRNA costuma ocorrer por clivagem direta por meio da Ago2 ou por desestabilização da molécula via deadenilação que tem como consequência, a perda do *cap* e a degradação do mRNA (Behm-Ansmant *et al.*, 2006)

As interações miRNA/mRNA ocorrem preferencialmente na região 3’NT dos mRNAs, mas ocasionalmente elas ocorrem na região codificadora (Gaidatzis *et al.*, 2007). Nas regiões 3’NT, os miRNAs se ligam preferencialmente nas extremidades iniciais e terminais (Gaidatzis *et al.*, 2007; Didiano e Hobert, 2008), mas uma vez que a porção final dos mRNAs são compostas por sequências pouco complexas e ricas em adenina, essas interações são escassas e geralmente de baixa afinidade (Majoros e Ohler, 2007; Didiano e Hobert, 2008).

Os resultados da interação de um miRNA com seu alvo podem incluir alterações na homeostase celular, o que torna os miRNAs importantes marcadores de condições normais ou patológicas e suas funções como supressores de tumor ou “oncoMIRs” (Dong *et al.*, 2013). Sua função também já foi descrita em uma série de processos biológicos, entre eles proliferação, morte celular, desenvolvimento, diferenciação celular, entre outros (Pasquinelli e Ruvkun, 2002; Fiorenza e Barco, 2016; Roy, 2016).

Apesar do tamanho pequeno (miRNAs são os menores carreadores de informação de regulação genética tanto em animais quanto em plantas), os miRNAs representam as sequências de DNA mais conservadas já conhecidas, de modo que a família de miRNAs 854 (miR-854a-e), por exemplo, pode ser encontrada em plantas (*Arabidopsis thaliana*), no nematódeo *Caenorhabditis elegans* e em mamíferos como o camundongo e humanos (Arteaga-Vazquez *et al.*, 2006). Além da sequência, a estrutura de diversos miRNAs, incluindo *loops*, bolhas de RNA não pareado e estrutura secundária das moléculas mostram sinais

de conservação. A região mais conservada dos miRNAs é a porção final da ponta 5' (nucleotídeos 2 a 8), conhecida como *seed*. A *seed* é fundamental ao funcionamento do miRISC, de modo que o não pareamento dessa região ao mRNA alvo é suficiente para inativar os mecanismos de silenciamento, enquanto o mesmo fato não é observado para a ponta 3' do miRNA (Doench e Sharp, 2004).

No que diz respeito às regiões 3'NTs de genes HLA, sabe-se que as regiões 3'NT dos genes *HLA-A* (Rene *et al.*, 2015), *HLA-C* (Kulkarni *et al.*, 2011) e *HLA-G* (Tan *et al.*, 2007; Porto *et al.*, 2015) são alvos de miRNAs. Esses miRNAs são capazes de alterar seus níveis proteicos. Vale ressaltar que um mesmo miRNA (miR-148a-3p) foi descrito como sendo capaz de regular tanto os níveis de HLA-G e HLA-C em nível pós transcricional e, para este último, apresenta um papel importante na progressão do HIV (Kulkarni *et al.*, 2011; Manaster *et al.*, 2012; Celsi *et al.*, 2013).

### **Métodos de detecção de alvos de microRNAs e uso do miRhunt**

A detecção de alvos de miRNAs e os mecanismos que governam a interação miRNA/mRNA tornaram-se importantes áreas de estudo nas últimas décadas. Com os avanços computacionais, estudos envolvendo as características estruturais e de conservação das sequências alvos de miRNAs contribuíram bastante para o entendimento da regulação mediada por miRNAs.

Os algoritmos de predição de alvos de miRNAs usam estratégias e parâmetros variados. No entanto, um fator em comum a todos os algoritmos é a necessidade de um banco de dados com as sequências de miRNAs, obtidas em bancos de dados especializados, como o miRbase (disponível em <http://mirbase.org/>) e a sequência de mRNA alvo, geralmente uma 3'NT (Riffo-Campos *et al.*, 2016). Uma vez que o pareamento da região *seed* é fundamental à interação miRNA/mRNA, certos algoritmos, como RNAhybrid (Rehmsmeier *et al.*, 2004; Kruger e Rehmsmeier, 2006), miRanda (Enright *et al.*, 2003) e PITA (Kertesz *et al.*, 2007), usam o pareamento desse segmento do miRNA à sequência alvo de mRNA em seus cálculos. Além da *seed*, a acessibilidade ao alvo,

influenciada pela estrutura secundária do mRNA também é levada em consideração. Isso é observado em algoritmos como PITA (Kertesz *et al.*, 2007), IntaRNA (Busch *et al.*, 2008) e Vienna RNA (Lorenz *et al.*, 2011). De modo geral, os algoritmos costumam usar combinações de parâmetros (conservação de local de ligação, complementaridade à região *seed*, contribuição de regiões próximas à ligação e sua estrutura, termodinâmica, entre outros) para gerar seus resultados (Wagner *et al.*, 2014).

A ligação entre miRNA e mRNA ocorre via complementaridade de bases. No entanto, quando somente este parâmetro é avaliado, observa-se um número excessivo de falsos positivos (Bartel, 2009). Assim, os algoritmos de predição de alvos de miRNAs geralmente levam em consideração os princípios da termodinâmica que governam a interação miRNA/mRNA. Nesse contexto, eles calculam a Energia Livre Mínima (MFE, do inglês *Minimum Free Energy*) para determinar o quão forte é essa interação entre miRNA e alvo. O MFE tem como unidade o “kcal/mol” e é expresso em um valor negativo que quanto menor, melhor é a interação miRNA/mRNA (Doench e Sharp, 2004; Riffo-Campos *et al.*, 2016). Quando a acessibilidade é levada em consideração, calcula-se o “delta G” ( $\Delta G$ ), ou a diferença entre a energia necessária para romper a ligação de bases que formam uma estrutura secundária do mRNA e a energia gerada na ligação miRNA/mRNA (Kertesz *et al.*, 2007; Busch *et al.*, 2008).

Uma vez que vários algoritmos com abordagens distintas continuam a ser desenvolvidos, a seleção da melhor estratégia de predição torna-se um grande desafio. Assim, sugere-se o uso de mais de uma ferramenta, seguida da integração e comparação dos resultados. Outra dificuldade encontrada na predição *in silico* de alvos de miRNAs é determinar quais valores obtidos pelos algoritmos são relevantes – quais valores de energia seriam compatíveis à uma interação miRNA/mRNA funcional. Nesse contexto, afim de sanar a dificuldade em determinar valores de MFE ou  $\Delta G$  que sirvam de *cutoff*, um trabalho anterior do nosso grupo criou um programa que automatiza a corrida de três algoritmos distintos – RNAhybrid (Kruger e Rehmsmeier, 2006), miRanda (Enright *et al.*, 2003) e IntaRNA (Busch *et al.*, 2008) – chamado miRhunt (disponível em

[www.castelli-lab.net/apps/mirhunt](http://www.castelli-lab.net/apps/mirhunt)).

O algoritmo miRanda identifica locais de complementaridade entre miRNA e mRNA e identifica possíveis sítios de ligação. A ligação da região 5' do miRNA (*seed*) resulta em escores maiores do que aqueles cuja complementaridade ocorre somente na sua região 3'. Assim, quanto maior a complementaridade entre *seed* e alvo de ligação, melhores os resultados. Estes ainda levam em consideração estabilidade termodinâmica da ligação miRNA/mRNA e conservação evolutiva da sequência alvo (Enright *et al.*, 2003). Já o RNAhybrid permite ao usuário determinar o tamanho da *seed* a ser considerado, o uso ou não da oscilação G:U (comum em ligações RNA/RNA) e a formação de *loops* na sequência alvo, permitindo a complementaridade com regiões menos próximas à *seed*. Uma grande vantagem do seu uso é a velocidade de análise, uma vez que somente miRNAs em que há pareamento da *seed* são avaliados (mas ele permite mudanças no tamanho e na posição da *seed*) (Kruger e Rehmsmeier, 2006). O IntaRNA, por sua vez, acrescenta aos cálculos de estabilidade da interação (hibridização miRNA/mRNA) a energia necessária para tornar essa interação acessível (energia necessária para tornar o sítio alvo em uma região de fita simples), considerando que mRNAs assumem conformações secundárias. Assim como o RNAhybrid ele incorpora parâmetros relacionados à região *seed*, como o número de bases que devem ser pareadas perfeitamente e o número máximo de bases não pareadas (Busch *et al.*, 2008).

Além de automatizar as corridas dos três algoritmos, miRhunt compara seus resultados para a sequência alvo com os obtidos com milhares de sequências de 3'UTR humanas, para calcular um sistema de escores que facilita o *ranking* dos melhores resultados com base não só na afinidade ou força de ligação de um miRNA a uma região 3'NT, mas também em sua especificidade a ela (Porto *et al.*, 2015).

Para isso, o miRhunt conta com um banco de dados onde se encontram os resultados obtidos com cada um dos três algoritmos para todas as sequências de 3'NT humanas descritas até 2015 contra todos os miRNAs maduros descritos no miRbase *release 21*. Assim, ele compara a energia de ligação observada entre um

miRNA e o mRNA alvo à energia desse mesmo miRNA quando interagindo com as outras 3'NTs humanas. Para se calcular a afinidade (estabilidade), escore E, de um certo miRNA ao alvo em questão, o menor valor de MFE ou  $\Delta G$  encontrado na matriz (de cada algoritmo) é levado em consideração. Desta forma, o score E seria a razão entre o MFE ou  $\Delta G$  da interação em questão e o menor valor encontrado em toda matriz ( $E_{\text{algoritmo}} = \text{MFE}_{\text{miRNA}}$  ou  $\Delta G_{\text{miRNA}}/\text{MFE}_{\text{mínimo}}$ ). O escore E final é calculado a partir da soma dos escores dos três algoritmos ( $E_{\text{final}} = E_{\text{hybrid}} + E_{\text{miranda}} + E_{\text{intaRNA}}$ ). Quanto maior este valor, maior a probabilidade deste miRNA interagir fortemente com a região 3'NT estudada.

Para avaliar a especificidade de um miRNA à região 3'NT em questão, o MFE ou  $\Delta G$  desse miRNA é comparado aos valores deste mesmo miRNA distribuídos ao longo da matriz. O número de valores de energia menores que o encontrado para a sequência alvo é dividido pelo número total de 3'NTs disponíveis na matriz (cerca de 27 mil valores), gerando um valor (P) que determina a probabilidade empírica desse miRNA se ligar à sequência alvo estudada sem muita interferência em outros mRNAs. Calculando-se  $1 - p$ , obtem-se o escore S de especificidade e a soma do valor S dos três algoritmos gera o  $S_{\text{final}}$ . Quanto maior este valor, maior a especificidade de um dado miRNA à sequência alvo estudada (Porto *et al.*, 2015).

Esse sistema de escores minimiza o uso arbitrário de valores de MFE ou  $\Delta G$  na determinação de quais miRNAs são relevantes quanto ao seu potencial de regular um mRNA alvo. Essa estratégia é particularmente útil quando observamos que miRNAs cujos valores de MFE não são particularmente baixos, mas que comprovadamente são capazes de regular um determinado alvo. Isso é observado, por exemplo, na interação entre o miRNA *miR-148a-3p* e a região 3'NT do gene *HLA-G*. A afinidade observada entre eles é relativamente baixa devido ao alto valor de MFE dessa interação, porém, raras são as sequências de 3'NT humanas que produziram um valor de especificidade (S) maior do que o observado para *HLA-G* e este miRNA em particular (Porto *et al.*, 2015).

Assim, miRNAs que apresentam valores de MFE ou  $\Delta G$  altos que seriam excluídos em um estudo baseado somente nesses valores, podem ser considerados

pela estratégia aqui proposta caso apresentem altos valores de especificidade. Além disso, miRNAs que apresentam alta especificidade a uma determinada sequência alvo, apresentariam pouca interferência em outras moléculas de mRNA. Contudo, estudos funcionais ainda se fazem necessários para comprovar se os miRNAs detectados pelo miRhunt como específicos ou de alta afinidade são realmente capazes de regular o mRNA alvo.

No entanto, é preciso salientar que o uso de mais de um algoritmo de predição de alvos de miRNAs não exclui a presença de falsos positivos. É preciso considerar que até o momento não se observa uma correlação entre os valores de MFE ou  $\Delta G$  e a capacidade funcional de controle pós-transcricional real de um miRNA. Em contrapartida, estes valores são boas medidas da afinidade entre um determinado miRNA e uma possível sequência alvo. Quanto maior essa afinidade, maior a possibilidade deste miRNA atuar no controle pós-transcricional do mRNA alvo.

A predição computacional de miRNAs que se ligam a um determinado mRNA alvo é necessária para nortear estudos sobre a regulação pós-transcricional. Uma vez que mais de 2000 sequências de miRNAs maduros estão disponíveis nos bancos de dados, seria totalmente inviável testá-los funcionalmente um a um. O uso de ferramentas computacionais vem então para filtrar miRNAs com alto potencial de regulação de um determinado alvo, especialmente quando múltiplos algoritmos são usados. Ao mesmo tempo, o principal obstáculo no estudo de predição de alvos de miRNAs é a falta de validação experimental das interações putativas entre miRNA e mRNA preditas pelos métodos computacionais. Nesse contexto, não existe nenhum método de predição de alvos de miRNAs que possa ser considerado “padrão-ouro” (Wagner *et al.*, 2014).

In conclusion, the present series provides an important set of data for future functional studies, which may focus on specific miRNA regulation of HLA levels, and how 3'UTR variability may influence miRNA binding and allelic specific differential levels.

## References

- [1] N. Anfossi, P. Andre, S. Guia, C.S. Falk, S. Roetynck, C.A. Stewart, V. Bresò, C. Frassati, D. Reviron, D. Middleton, F. Romagne, S. Ugolini, E. Vivier, Human NK cell education by inhibitory receptors for MHC class I, *Immunity*, 25 (2006) 331-342.
- [2] D.K. Cole, N.J. Pumphrey, J.M. Boulter, M. Sami, J.I. Bell, E. Gostick, D.A. Price, G.F. Gao, A.K. Sewell, B.K. Jakobsen, Human TCR-binding affinity is governed by MHC class restriction, *Journal of immunology*, 178 (2007) 5727-5734.
- [3] Dausset, [The HLA complex. I. Immunogenetics of the HLA system. The HLA-A, B and C allelic series], *La Nouvelle presse medicale*, 5 (1976) 1301-1304.
- [4] P. Parham, HLA, anthropology, and transplantation, *Transplantation proceedings*, 25 (1993) 159-161.
- [5] M. Muro, M.R. Lopez-Alvarez, J.A. Campillo, L. Marin, M.R. Moya-Quiles, J.M. Bolarin, C. Botella, G. Salgado, P. Martinez, F. Sanchez-Bueno, R. Lopez-Hernandez, F. Boix, A. Bosch, H. Martinez, J.M. de la Pena-Moral, N. Perez, R. Robles, A.M. Garcia-Alonso, A. Minguela, M. Miras, M.R. Alvarez-Lopez, Influence of human leukocyte antigen mismatching on rejection development and allograft survival in liver transplantation: is the relevance of HLA-A locus matching being underestimated?, *Transplant immunology*, 26 (2012) 88-93.
- [6] J.A. McCutcheon, J. Gumperz, K.D. Smith, C.T. Lutz, P. Parham, Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA, *The Journal of experimental medicine*, 181 (1995) 2085-2095.
- [7] A. Neisig, C.J. Melief, J. Neeffjes, Reduced cell surface expression of HLA-C molecules correlates with restricted peptide binding and stable TAP interaction, *Journal of immunology*, 160 (1998) 171-179.
- [8] A. King, T.D. Burrows, S.E. Hiby, J.M. Bowen, S. Joseph, S. Verma, P.B. Lim, L. Gardner, P. Le Bouteiller, A. Ziegler, B. Uchanska-Ziegler, Y.W. Loke,

Surface expression of HLA-C antigen by human extravillous trophoblast, *Placenta*, 21 (2000) 376-387.

[9] P. Tripathi, S. Naik, S. Agrawal, HLA-E and immunobiology of pregnancy, *Tissue antigens*, 67 (2006) 207-213.

[10] M. Dahl, L. Klitkou, O.B. Christiansen, S. Djuricic, Z.M. Piosik, P. Skovbo, A.M. Moller, R. Steffensen, T.V. Hviid, Human leukocyte antigen (HLA)-G during pregnancy part II: associations between maternal and fetal HLA-G genotypes and soluble HLA-G, *Human immunology*, 76 (2015) 260-271.

[11] H.G. Hilton, L.A. Guethlein, A. Goyos, N. Nemat-Gorgani, D.A. Bushnell, P.J. Norman, P. Parham, Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes, *Journal of immunology*, DOI 10.4049/jimmunol.1501358(2015).

[12] J. Klein, A. Sato, The HLA system. First of two parts, *The New England journal of medicine*, 343 (2000) 702-709.

[13] D. Didiano, O. Hobert, Molecular architecture of a miRNA-regulated 3' UTR, *Rna*, 14 (2008) 1297-1317.

[14] A. Yamashita, O. Takeuchi, Translational control of mRNAs by 3'-Untranslated region binding proteins, *BMB reports*, 50 (2017) 194-200.

[15] H.S. Yeh, J. Yong, Alternative Polyadenylation of mRNAs: 3'-Untranslated Region Matters in Gene Expression, *Molecules and cells*, 39 (2016) 281-285.

[16] E.C. Castelli, P. Gerasimou, M.A. Paz, J. Ramalho, I.O. Porto, T.H. Lima, A.S. Souza, L.C. Veiga-Castelli, C.V. Collares, E.A. Donadi, C.T. Mendes-Junior, P. Costeas, HLA-G variability and haplotypes detected by massively parallel sequencing procedures in the geographically distinct population samples of Brazil and Cyprus, *Molecular immunology*, 83 (2017) 115-126.

[17] E.C. Castelli, L.C. Veiga-Castelli, L. Yaghi, P. Moreau, E.A. Donadi, Transcriptional and posttranscriptional regulations of the HLA-G gene, *Journal of immunology research*, 2014 (2014) 734068.

[18] E.C. Castelli, C.T. Mendes-Junior, N.H. Deghaide, R.S. de Albuquerque, Y.C. Muniz, R.T. Simoes, E.D. Carosella, P. Moreau, E.A. Donadi, The genetic structure of 3'untranslated region of the HLA-G gene: polymorphisms and haplotypes, *Genes and immunity*, 11 (2010) 134-141.

[19] E.A. Donadi, E.C. Castelli, A. Arnaiz-Villena, M. Roger, D. Rey, P. Moreau, Implications of the polymorphism of HLA-G on its function, regulation, evolution



and disease association, *Cellular and molecular life sciences : CMLS*, 68 (2011) 369-395.

[20] T.V. Hviid, R. Rizzo, L. Melchiorri, M. Stignani, O.R. Baricordi, Polymorphism in the 5' upstream regulatory and 3' untranslated regions of the HLA-G gene in relation to soluble HLA-G and IL-10 expression, *Human immunology*, 67 (2006) 53-62.

[21] C. Ober, C.L. Aldrich, HLA-G polymorphisms: neutral evolution or novel function?, *Journal of reproductive immunology*, 36 (1997) 1-21.

[22] C. Ober, C.L. Aldrich, I. Chervoneva, C. Billstrand, F. Rahimov, H.L. Gray, T. Hyslop, Variation in the HLA-G promoter region influences miscarriage rates, *Am J Hum Genet*, 72 (2003) 1425-1435.

[23] P. Rousseau, M. Le Discorde, G. Mouillot, C. Marcou, E.D. Carosella, P. Moreau, The 14 bp deletion-insertion polymorphism in the 3' UT region of the HLA-G gene influences HLA-G mRNA stability, *Human immunology*, 64 (2003) 1005-1010.

[24] Z. Tan, G. Randall, J. Fan, B. Camoretti-Mercado, R. Brockman-Schneider, L. Pan, J. Solway, J.E. Gern, R.F. Lemanske, D. Nicolae, C. Ober, Allele-specific targeting of microRNAs to HLA-G and risk of asthma, *Am J Hum Genet*, 81 (2007) 829-834.

[25] Z. Tan, A.M. Shon, C. Ober, Evidence of balancing selection at the HLA-G promoter region, *Human molecular genetics*, 14 (2005) 3619-3628.

[26] E.C. Castelli, C.T. Mendes-Junior, L.C. Veiga-Castelli, M. Roger, P. Moreau, E.A. Donadi, A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution, *Molecular biology and evolution*, 28 (2011) 3069-3086.

[27] A. Sabbagh, P. Luisi, E.C. Castelli, L. Gineau, D. Courtin, J. Milet, J.D. Massaro, H. Laayouni, P. Moreau, E.A. Donadi, A. Garcia, Worldwide genetic variation at the 3' untranslated region of the HLA-G gene: balancing selection influencing genetic diversity, *Genes and immunity*, DOI 10.1038/gene.2013.67

gene201367 [pii](2013).

[28] K.E. Santos, T.H. Lima, L.P. Felicio, J.D. Massaro, G.M. Palomino, A.C. Silva, S.F. Oliveira, A. Sabbagh, A. Garcia, P. Moreau, E.A. Donadi, C.T. Mendes-Junior, E.C. Castelli, Insights on the HLA-G evolutionary history provided by a nearby Alu insertion, *Molecular biology and evolution*, 30 (2013) 2423-2434.

- [29] C. Rene, C. Lozano, M. Villalba, J.F. Eliaou, 5' and 3' untranslated regions contribute to the differential expression of specific HLA-A alleles, *European journal of immunology*, 45 (2015) 3454-3463.
- [30] S. Kulkarni, V. Ramsuran, M. Rucevic, S. Singh, A. Lied, V. Kulkarni, C. O'HUigin, S. Le Gall, M. Carrington, Posttranscriptional Regulation of HLA-A Protein Expression by Alternative Polyadenylation Signals Involving the RNA-Binding Protein Syncrip, *Journal of immunology*, 199 (2017) 3892-3899.
- [31] S. Kulkarni, R. Savan, Y. Qi, X. Gao, Y. Yuki, S.E. Bass, M.P. Martin, P. Hunt, S.G. Deeks, A. Telenti, F. Pereyra, D. Goldstein, S. Wolinsky, B. Walker, H.A. Young, M. Carrington, Differential microRNA regulation of HLA-C expression and its association with HIV control, *Nature*, 472 (2011) 495-498.
- [32] S.A. Miller, D.D. Dykes, H.F. Polesky, A simple salting out procedure for extracting DNA from human nucleated cells, *Nucleic acids research*, 16 (1988) 1215.
- [33] E.C. Castelli, C.T. Mendes-Junior, A. Sabbagh, I.O. Porto, A. Garcia, J. Ramalho, T.H. Lima, J.D. Massaro, F.C. Dias, C.V. Collares, V. Jamonneau, B. Bucheton, M. Camara, E.A. Donadi, HLA-E coding and 3' untranslated region variability determined by next-generation sequencing in two West-African population samples, *Human immunology*, DOI 10.1016/j.humimm.2015.06.016(2015).
- [34] T.H. Lima, R.V. Buttura, E.A. Donadi, L.C. Veiga-Castelli, C.T. Mendes-Junior, E.C. Castelli, HLA-F coding and regulatory segments variability determined by massively parallel sequencing procedures in a Brazilian population sample, *Human immunology*, 77 (2016) 841-853.
- [35] M.A. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A.A. Philippakis, G. del Angel, M.A. Rivas, M. Hanna, A. McKenna, T.J. Fennell, A.M. Kernysky, A.Y. Sivachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler, M.J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data, *Nature genetics*, 43 (2011) 491-498.
- [36] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, *Genome research*, 20 (2010) 1297-1303.
- [37] G.A. Van der Auwera, M.O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K.V. Garimella, D. Altshuler, S. Gabriel, M.A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline,

Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.], 11 (2013) 11 10 11-11 10 33.

[38] M. Stephens, N.J. Smith, P. Donnelly, A new statistical method for haplotype reconstruction from population data, *Am J Hum Genet*, 68 (2001) 978-989.

[39] I.O. Porto, C.T. Mendes-Junior, L.P. Felicio, R.C. Georg, P. Moreau, E.A. Donadi, J.A. Chies, E.C. Castelli, MicroRNAs targeting the immunomodulatory HLA-G gene: a new survey searching for microRNAs with potential to regulate HLA-G, *Molecular immunology*, 65 (2015) 230-241.

[40] I. Manaster, D. Goldman-Wohl, C. Greenfield, D. Nachmani, P. Tsukerman, Y. Hamani, S. Yagel, O. Mandelboim, MiRNA-mediated control of HLA-G expression and function, *PloS one*, 7 (2012) e33395.

[41] E.C. Castelli, J. Ramalho, I.O. Porto, T.H. Lima, L.P. Felicio, A. Sabbagh, E.A. Donadi, C.T. Mendes-Junior, Insights into HLA-G Genetics Provided by Worldwide Haplotype Diversity, *Frontiers in immunology*, 5 (2014) 476.

[42] T. Shiina, G. Tamiya, A. Oka, N. Takishima, T. Yamagata, E. Kikkawa, K. Iwata, M. Tomizawa, N. Okuaki, Y. Kuwano, K. Watanabe, Y. Fukuzumi, S. Itakura, C. Sugawara, A. Ono, M. Yamazaki, H. Tashiro, A. Ando, T. Ikemura, E. Soeda, M. Kimura, S. Bahram, H. Inoko, Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region, *Proceedings of the National Academy of Sciences of the United States of America*, 96 (1999) 13282-13287.

[43] M. Le Discorde, P. Moreau, P. Sabatier, J.M. Legeais, E.D. Carosella, Expression of HLA-G in human cornea, an immune-privileged tissue, *Human immunology*, 64 (2003) 1039-1044.

[44] S. Lefebvre, F. Adrian, P. Moreau, L. Gourand, J. Dausset, S. Berrih-Aknin, E.D. Carosella, P. Paul, Modulation of HLA-G expression in human thymic and amniotic epithelial cells, *Human immunology*, 61 (2000) 1095-1101.

[45] Y.Q. Wang, S.L. Chen, F.Q. Xing, [Expression of HLA-G protein in trophoblast cells], *Di 1 jun yi da xue xue bao = Academic journal of the first medical college of PLA*, 25 (2005) 1488-1490.

[46] D.R. Johnson, Differential expression of human major histocompatibility class I loci: HLA-A, -B, and -C, *Human immunology*, 61 (2000) 389-396.

[47] G. Martelli-Palomino, J.A. Pancotto, Y.C. Muniz, C.T. Mendes-Junior, E.C. Castelli, J.D. Massaro, I. Krawice-Radanne, I. Poras, V. Rebmann, E.D. Carosella, N. Rouas-Freiss, P. Moreau, E.A. Donadi, Polymorphic sites at the 3' untranslated

region of the HLA-G gene are associated with differential hla-g soluble levels in the Brazilian and French population, *PloS one*, 8 (2013) e71742.

[48] X.H. Wei, H.T. Orr, Differential expression of HLA-E, HLA-F, and HLA-G transcripts in human tissue, *Human immunology*, 29 (1990) 131-142.

[49] C.H. Chou, S. Shrestha, C.D. Yang, N.W. Chang, Y.L. Lin, K.W. Liao, W.C. Huang, T.H. Sun, S.J. Tu, W.H. Lee, M.Y. Chiew, C.S. Tai, T.Y. Wei, T.R. Tsai, H.T. Huang, C.Y. Wang, H.Y. Wu, S.Y. Ho, P.R. Chen, C.H. Chuang, P.J. Hsieh, Y.S. Wu, W.L. Chen, M.J. Li, Y.C. Wu, X.Y. Huang, F.L. Ng, W. Buddhakosai, P.C. Huang, K.C. Lan, C.Y. Huang, S.L. Weng, Y.N. Cheng, C. Liang, W.L. Hsu, H.D. Huang, miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions, *Nucleic acids research*, DOI 10.1093/nar/gkx1067(2017).

[50] B. Panwar, G.S. Omenn, Y. Guan, miRmine: a database of human miRNA expression profiles, *Bioinformatics*, 33 (2017) 1554-1560.

[51] S.A. Ellis, Immune status: normal expression of MHC class I in the placenta and what is expected in clones, *Cloning Stem Cells*, 6 (2004) 121-125.

[52] D. Fiszer, M. Kurpisz, Major histocompatibility complex expression on human, male germ cells: a review, *American journal of reproductive immunology*, 40 (1998) 172-176.

[53] C. Rene, C. Lozano, J.F. Eliaou, Expression of classical HLA class I molecules: regulation and clinical impacts, *Hla*, 87 (2016) 338-349.

[54] G. Kaur, S. Gras, J.I. Mobbs, J.P. Vivian, A. Cortes, T. Barber, S.B. Kuttikkatte, L.T. Jensen, K.E. Attfield, C.A. Dendrou, M. Carrington, G. McVean, A.W. Purcell, J. Rossjohn, L. Fugger, Structural and regulatory diversity shape HLA-C protein expression levels, *Nature communications*, 8 (2017) 15924.

[55] A. King, C. Boocock, A.M. Sharkey, L. Gardner, A. Beretta, A.G. Siccardi, Y.W. Loke, Evidence for the expression of HLA-C class I mRNA and protein by human first trimester trophoblast, *Journal of immunology*, 156 (1996) 2068-2076.

[56] R. Hackmon, L. Pinnaduwege, J. Zhang, S.J. Lye, D.E. Geraghty, C.E. Dunk, Definitive class I human leukocyte antigen expression in gestational placentation: HLA-F, HLA-E, HLA-C, and HLA-G in extravillous trophoblast invasion on placentation, pregnancy, and parturition, *American journal of reproductive immunology*, 77 (2017).

[57] E.D. Carosella, P. Moreau, J. Le Maoult, M. Le Discorde, J. Dausset, N. Rouas-Freiss, HLA-G molecules: from maternal-fetal tolerance to tissue acceptance, *Advances in immunology*, 81 (2003) 199-252.

- [58] A. Lycoudi, D. Mavreli, A. Mavrou, N. Papantoniou, A. Kolialexi, miRNAs in pregnancy-related complications, *Expert review of molecular diagnostics*, 15 (2015) 999-1010.
- [59] D.M. Morales-Prieto, S. Ospina-Prieto, W. Chaiwangyen, M. Schoenleben, U.R. Markert, Pregnancy-associated miRNA-clusters, *Journal of reproductive immunology*, 97 (2013) 51-61.
- [60] I. Hromadnikova, K. Kotlabova, M. Ondrackova, P. Pirkova, A. Kestlerova, V. Novotna, L. Hympanova, L. Krofta, Expression profile of C19MC microRNAs in placental tissue in pregnancy-related complications, *DNA and cell biology*, 34 (2015) 437-457.
- [61] M.J. Chen, Y.M. Cheng, C.C. Chen, Y.C. Chen, C.J. Shen, MiR-148a and miR-152 reduce tamoxifen resistance in ER<sup>+</sup> breast cancer via downregulating ALCAM, *Biochemical and biophysical research communications*, 483 (2017) 840-846.
- [62] F. Wang, H. Ying, B. He, Y. Pan, H. Sun, S. Wang, Circulating miR-148/152 family as potential biomarkers in hepatocellular carcinoma, *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, 37 (2016) 4945-4953.
- [63] A.S. Daar, S.V. Fuggle, J.W. Fabre, A. Ting, P.J. Morris, The detailed distribution of HLA-A, B, C antigens in normal human organs, *Transplantation*, 38 (1984) 287-292.