



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU
GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS - MODALIDADE BACHARELADO

Isabelle Mira da Silva

**PERFIL TRANSCRICIONAL DO GENE *HLA-E*: TRANSCRITOS
ALTERNATIVOS PELA RETENÇÃO DE SEQUÊNCIAS
INTRÔNICAS**

Botucatu

2023

Isabelle Mira da Silva

**PERFIL TRANSCRICIONAL DO GENE *HLA-E*: TRANSCRITOS
ALTERNATIVOS PELA RETENÇÃO DE SEQUÊNCIAS
INTRÔNICAS**

Trabalho de Conclusão de Curso
apresentado à Universidade Estadual
Paulista, como requisito à obtenção do título
de Bacharel em Ciências Biológicas apoiado
pela FAPESP projeto 2021/14213 - 2.

Orientador: Prof. Dr. Erick da Cruz Castelli

Botucatu

2023

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Silva, Isabelle Mira da.

Perfil transcricional do gene HLA-E : transcritos
alternativos pela retenção de sequências intrônicas /
Isabelle Mira da Silva. - Botucatu, 2023

Trabalho de conclusão de curso (bacharelado - Ciências
Biológicas) - Universidade Estadual Paulista "Júlio de
Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Erick da Cruz Castelli
Capes: 21103003

1. Antígenos HLA. 2. Isoformas de proteínas.
3. Seqüenciamento de nucleotídeo. 4. Expressão gênica.

Palavras-chave: HLA-E; Isoformas; Perfil de expressão;
Perfil transcricional.

Isabelle Mira da Silva

**PERFIL TRANSCRICIONAL DO GENE *HLA-E*: TRANSCRITOS
ALTERNATIVOS PELA RETENÇÃO DE SEQUÊNCIAS
INTRÔNICAS**

BANCA EXAMINADORA

Prof. Dr. Erick da Cruz Castelli
Unidade de Pesquisa Experimental (UNIPLEX), FMB, UNESP

Profa. Dra. Patricia Pintor dos Reis
Unidade de Pesquisa Experimental (UNIPLEX), FMB, UNESP

Prof. Dr. Robson Francisco Carvalho
Departamento de Biologia Estrutural e Funcional, IBB, UNESP

DEDICATÓRIA

Dedico este trabalho à minha mãe,
Selma Mira e ao meu irmão, Gustavo
Mira da Silva.

AGRADECIMENTOS

Começo meus agradecimentos primeiramente a todas as pessoas maravilhosas que estiveram em minha vida, principalmente à minha mãe Selma, na qual sem ela não estaria aqui. Todo o apoio e incentivo me possibilitaram chegar e finalizar essa etapa tão importante e ao meu irmão Gustavo, por seu amor, companheirismo e amizade para a vida. Amo muito vocês. Agradeço a toda minha família, por tudo.

Agradeço à minha orientadora de Iniciação Científica, Profa. Dra. Flávia Karina Delella, minha amiga do coração, fada e inspiração pessoal e profissional para esse caminho.

Agradeço ao meu grande amor, Karine, por tudo o que você é, obrigada por tudo que estamos vivendo e temos, pelo apoio e força, te amo.

Agradeço à todas as minhas amigas e amigos especiais, Luana, Leonardo, Jefferson, Raphaela, Viviane, Gabriela, Joyce, Amanda e Nayane, por tudo, as risadas, choros, abraços, aprendizados (mais choros, brincadeira) e tudo de incrível que pudemos compartilhar e guardar. Que muito mais desses momentos venham, amo muito vocês.

Agradeço ao meu orientador, Prof. Dr. Erick da Cruz Castelli, por compartilhar um pouco do seu vasto conhecimento, saberes, conversas científicas e incentivo nesse caminho.

Agradeço a Profa. Camila, por compartilhar um pouco dos seus conhecimentos imunológicos em nossas conversas científicas e pelo apoio e ajuda.

Agradeço também à Fundação de Amparo à Pesquisa (FAPESP) pelo apoio financeiro e ao departamento de Unidade de Pesquisa Experimental (UNIPLEX)/FMB pela disponibilização do espaço para desenvolvimento desse projeto.

RESUMO

O Antígeno Leucocitário Humano (HLA)-E é a molécula de HLA de classe I menos polimórfica dentre outras do mesmo grupo, na qual encontra-se geralmente ancorada à superfície das APCs e estando envolvida em processos de imunorregulação de células *Natural Killer* (NK) e T, por meio da interação com os receptores ativadores (CD94/NKG2C) e inibidores (CD94/NKG2A). Para além da estrutura ancorada, diversos estudos relatam a presença de isoformas solúveis em contextos inflamatórios e relacionados à progressão de quadros clínicos ligados à desordens psiquiátricas como a Esquizofrenia e Bipolaridade, contudo não descrevendo como poderiam ser originadas. Analisando a sequência nucleotídica original e as com retenção intrônica, verificamos seu impacto em regiões da sequência peptídica ligadas à síntese de isoformas solúveis. Utilizamos as sequências genômicas, exônicas e intrônicas presentes no banco de dados IPD-IMGT/HLA e EMBL-EBI/Emboss Transeq para traduzir em sequências de aminoácidos e o banco de dados GEUVADIS para comparar e inferir a frequência dessas sequências em 460 amostras de 5 populações diferentes abrangidas pela base de dados. No final, pesquisamos na literatura quais aminoácidos estariam envolvidos na interação ligante-receptor da célula NK. Ao compararmos todas as informações, descobrimos que a retenção de íntron 1 leva à produção de um códon de parada prematuro no mesmo íntron, resultando em uma proteína não funcional. A retenção de íntron 2 leva a um códon de parada prematuro no início do íntron 2, produzindo uma proteína com o domínio alfa-1. A retenção do íntron 3 leva a produção de códon de parada no mesmo íntron e a produção de uma molécula com o domínio alfa-1 e -2. A retenção de íntron 4 leva à produção de uma molécula HLA-E maior, sem região transmembrana, mas com a presença de todos os domínios externos (alfa-1 ao alfa-3). Em todos os casos de retenção intrônica, o peptídeo líder é mantido e a isoforma pode migrar para a membrana celular. No entanto, como não existe domínio transmembrana, essas isoformas podem ser liberadas como solúvel. Tendo em vista as isoformas com alfa-1 e -2 produzidas pela retenção de íntron 3 e 4, há possibilidade de interação com o NKG2A e CD94, como habitualmente, contudo o domínio alfa-1 produzido por retenção de íntron 2, pode interagir apenas com o CD94. Com isso, a possível molécula isoforma solúvel produzida por estas diferentes sequências poderiam ter mais do que um tipo de estrutura.

Palavras-chave: *HLA-E*, perfil transcricional, perfil de expressão, isoformas.

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVOS	4
3. MÉTODOS	4
3.1 <i>1000genomes</i> e <i>GEUVADIS</i>	4
3.2 Algoritmos de processamento e alinhamento de sequências: STAR e <i>hla-mapper</i>	5
3.3 Simulação de retenção intrônica e inferência de frequência	5
4. RESULTADOS	6
5. DISCUSSÃO	8
REFERÊNCIAS BIBLIOGRÁFICAS	10

LISTA DE FIGURAS

Figura 1: Estrutura proteica de MHC de Classe I e II	2
Figura 2: Proporção de transcritos de HLA-E contendo íntrons retidos	6
Figura 3: Estrutura do transcrito primário canônico e alternativos	7
Figura 4: Simulação tridimensional das sequências polipeptídicas	8

1. INTRODUÇÃO

Exercendo um papel fundamental na imunidade adaptativa regulando processos moleculares e celulares do sistema imunológico, o Complexo Principal de Histocompatibilidade (do inglês, *Major Histocompatibility Complex* - MHC), localizado no braço curto do cromossomo 6, em 6p21.3, possui aproximadamente 5 Megabases e mais de 200 genes bastante polimórficos. As variações genéticas ocorridas nesse complexo interferem diretamente na função e quantidade das proteínas codificadas nesta região, refletindo nos processos de carregamento e apresentação de antígenos e interação com os receptores presentes na superfície celular de células imunocompetentes, conseqüentemente, afetando a resposta imunológica. Diversas doenças vêm sendo associadas a essas variações e em diferentes genes dentro da região do MHC, como a Esclerose Múltipla, Doença Celíaca e Doença de Crohn (SHIINA et al., 2009; HANDUNNETTHI et al. 2010; DILTHEY, 2021).

Dentro da região do MHC, há um grupo de genes denominados genes HLA (do inglês, *Human Leucocyte Antigen*), com mais de 35.000 alelos diferentes considerando uma dezena de genes (IPD-IMGT/HLA, version 3.51, 01/2023) (ROBINSON et al., 2001). Esses genes estão relacionados principalmente com a apresentação de antígenos na superfície celular. Os genes HLA são divididos didaticamente em duas subclasses. Os genes de classe I abrangem os genes clássicos altamente polimórficos, como *HLA-A*, *HLA-C* e destacando-se, o *HLA-B*, e os não-clássicos, pouco polimórficos, como os genes *HLA-E*, *HLA-F* e *HLA-G*. A região de classe II inclui os genes que codificam as cadeias alfa e beta das moléculas clássicas *HLA-DP*, *HLA-DQ*, *HLA-DR*, e não clássicas *HLA-DO* e *HLA-DM* (YU et al., 2000; JONGSMA et al., 2017; DENDROU et al., 2018). A estrutura das moléculas de classe I e II estão representadas na Figura 1.

Os genes HLA de classe I clássicos são expressos em todas as células nucleadas. Essas moléculas apresentam peptídeos na superfície celular e interagem com os Receptores de Célula T (TCR), especificamente das células T CD8. Os genes de classe II são expressos em Células Apresentadoras de Antígenos (do inglês, *Antigen Presenting Cells*- APCs), como as células dendríticas, macrófagos e linfócitos B, e células epiteliais do Timo, sendo responsáveis por apresentar antígenos ao TCR de células T CD4. Além das suas diferentes vias de atuação, as moléculas HLA-I e HLA-II também diferenciam-se estruturalmente. As moléculas de classe I possuem

três domínios extracelulares ($\alpha 1$, $\alpha 2$ e $\alpha 3$), sendo dois deles, o $\alpha 1$ e $\alpha 2$, responsáveis pela interação com os peptídeos e sua apresentação aos receptores da família NKG2. Todos esses domínios, conjugam-se covalentemente a uma cadeia $\beta 2$ -microglobulina ($\beta 2$ -m), na qual promove estabilidade estrutural da proteína externalizada (**Figura 1.A**). Já as moléculas de classe II, apresentam uma estrutura formada por dois complexos proteicos de cadeia α e β que combinam-se, formando uma estrutura heterodímera (**Figura 1.B**). Ambas estruturas proteicas são ancoradas na superfície celular através do domínio transmembrana que se estende até o meio intracelular por meio da cauda citoplasmática. Diferentemente da classe I, os HLAs classe II estão, em sua maioria, relacionados a processos indiretos, como a edição de antígenos e a catalisação de troca entre peptídeos de cadeia invariável (do inglês, *Invariant Chain Peptide* - CLIP) pelos antígenos “editados” no sítio de ligação de peptídeo formado pelos domínios $\alpha 1$ e $\beta 1$ (ELSEN, 2011; WIECZOREK et al., 2017).

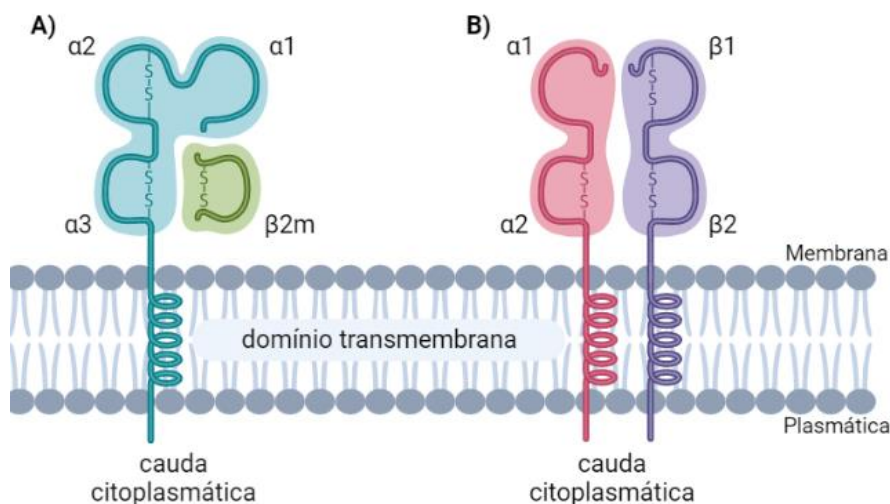


Figura 1. Estrutura proteica das proteínas de MHC de classe I (A) e II (B).

No conjunto de genes HLA classe I não-clássicos, estão presentes os genes *HLA-F*, *HLA-G* e *HLA-E*, menos polimórficos que os clássicos, porém importantes na ativação e regulação da resposta imune adaptativa e celular. O gene *HLA-E* apresenta 8 éxons, sendo o éxon 1 responsável por codificar o peptídeo sinal; o éxon 2 e 3, os domínios $\alpha 1$ e $\alpha 2$; o éxon 4, o domínio $\alpha 3$; e os éxons 5, 6 e 7 codificam a porção transmembrana e cauda citoplasmática, respectivamente. A presença de *HLA-E* é detectada em todos os tecidos e em níveis de expressão semelhantes, contudo é maior em tecidos pulmonares, trato urogenital (rins e bexiga), tonsilas, na medula óssea e também em diversas células imunes circulatórias

(<https://www.proteinatlas.org/ENSG00000204592-HLA-E/tissue>), como as APCs (IWASZKO & BOGUNIA-KUBIK, 2011; GAUDINO & KUMAR, 2019). Em certas condições fisiopatológicas, como hipóxia crônica, microambiente tumoral, alta secreção de Interferon- γ (IFN- γ) e interferência molecular de HLA-G, a expressão de HLA-E apresenta-se alterada (MORANDI & PISTOIA, 2014; SELIGER et al., 2016).

Segundo o banco de dados IPD-IMGT/HLA (<https://www.ebi.ac.uk/ipd/imgt/hla/about/statistics/>, versão 3.51, de 01/2023) (ROBINSON, et al. 2001), o banco de dados oficial que contém as sequências conhecidas de genes HLA, o gene *HLA-E* apresenta 346 alelos diferentes codificando 140 sequências proteicas. O número de sequências proteicas é menor que a de alelos pois muitas variações ocorrem em íntrons, ou são mutações sinônimas. As moléculas HLA-E estão envolvidas na interação com os receptores ativadores (CD94/NKG2C) e inibidores (CD94/NKG2A) de células *Natural Killers* (NK) e células T. *HLA-E* possui dois alelos frequentes em nível proteico, denominados HLA-E*01:01 e HLA-E*01:03 (GONZALEZ-GALARZA et al. 2020). Essas duas versões, com frequências muito semelhantes próximas de 50%, possuem diferenças estruturais relacionadas ao códon 107, onde E*01:01 possui uma Arginina e E*01:03 uma Glicina. E*01:03 apresenta maior afinidade e estabilidade com o peptídeo e potencial inibitório de ativação de NK quando comparada a E*01:01.

As moléculas HLA-E podem ser expressas tanto aderidas à membrana (mais comum) como em uma isoforma solúvel (sHLA-E). Esse mesmo padrão é observado para outro gene não clássico, o *HLA-G*. Porém, diferentemente das isoformas sHLA-G geradas pela presença de códons de parada prematuros ocasionados pela retenção de sequências intrônicas no mRNA, a sHLA-E é proveniente da ação de proteases e metaloproteinases (Mpsases) que clivam as moléculas presentes na superfície celular principalmente em ambiente com altos níveis de Fatores de Necrose Tumoral α (do inglês, *Tumor Necrosis Factors α* , TNF α), Interferon γ (IFN γ) e Interleucina-1 β (IL-1 β), como em Melanomas, Vasculites Associadas ao Anticorpo Citoplasmático Antineutrófilo (ANCA ou VAA) (KALLENBERG et al., 2002; COUPEL et al., 2007; ALLARD et al.; DONADI et al. 2011).

Anterior às contribuições de Coupel (2007), Dong e colaboradores (2003) demonstraram *in vitro* e em condições similares às fisiológicas normais, a presença de sHLA-G e sHLA-E em meio com baixo pH, apenas as cadeias pesadas livres de

β 2-m de ambas as moléculas, sendo elas provenientes da ação de proteinases e Mpases. Além disso, no mesmo estudo, foi atestada a possibilidade de expressão de sHLA-E via edição alternativa do transcrito primário, porém no estudo não há detalhes ou menção de como ocorre essa edição. Como relatado, HLA-E, em sua conformação original, apresenta um padrão de interação com os receptores NKG2 (principalmente, NKG2A) através dos resíduos de aminoácidos localizados na porção 65 a 89 da α -hélice 1 e 151 a 162 de α -hélice 2, porém não há relatos se outras configurações da mesma proteína compartilham o mesmo padrão de interação.

Neste estudo, avaliaremos a presença de transcritos alternativos de HLA-E formados pela retenção de sequências intrônicas e quais são os produtos gerados por esses transcritos, a partir de dados de RNA-seq de indivíduos de 5 populações distintas.

2. OBJETIVOS

Analisar a presença de retenção de íntron em transcritos de HLA-E em dados de RNA-seq e avaliar o perfil da molécula que seria gerada por esses transcritos alternativos.

3. MÉTODOS

3.1 Amostras utilizadas – o projeto *1000genomes* e GEUVADIS

O projeto *1000genomes* (DEVUYST, 2015) (<https://www.internationalgenome.org/>) é uma colaboração entre diferentes grupos de pesquisas dos Estados Unidos, Reino Unido, China e Alemanha, baseado no Projeto Internacional HapMap (The International HapMap Consortium, 2003), fundamentado no desenvolvimento de um repositório vasto e público contendo dados brutos de sequenciamento de amostras de diferentes populações mundiais, além de informações sobre variantes genéticas, estruturais e haplotípicas, direcionadas ao auxílio dos estudos de associação genômica relacionadas à diferentes doenças. Dentre os diversos dados presentes no projeto *1000genomes*, o GEUVADIS é um ramo especializado em informações sobre mRNA e miRNA de 465 amostras provenientes da combinação de sequenciamento genômico e transcriptômico de amostras da Toscana, Itália (TSI); Inglaterra e Escócia, Reino Unido (GBR); Ibadan, Nigéria (YRI); Utah, Estados Unidos (CEU); e Finlândia (FIN) (LAPPALAINEN et al.,

2013). Neste estudo, avaliamos o nível de retenção de íntrons no gene HLA-E utilizando os dados do projeto GEUVADIS. Para isso, baixamos os dados brutos de sequenciamento (FASTQ), que foi processado conforme descrito adiante.

3.2 Algoritmos de processamento e alinhamento de sequências: STAR e *hla-mapper*

O algoritmo STAR (do inglês, *Spliced Transcripts Alignment to a Reference*) (DOBIN et al., 2013) é uma ferramenta de alinhamento de sequências curtas obtidas via NGS (os *reads*), focada em sequenciamento de RNA. Este algoritmo detecta junções de *splicing*, alinhando as sequências ao genoma de referência. Utilizamos o programa STAR (parâmetros padrão) para alinhar os *reads* ao genoma de referência hg38 (última versão publicada do genoma de referência humano).

Os genes HLA possuem alta similaridade entre si e são extremamente polimórficos. Essas duas características prejudicam a análise de genes HLA utilizando o STAR, por duas razões: a similaridade entre genes causa alinhamentos cruzados, onde sequências de um gene alinham-se a outro; o elevado polimorfismo acarreta em falhas de alinhamento, onde o *read* possui sequência muito divergente em relação à referência e não é alinhado em nenhum lugar. Para resolver esses problemas, estamos trabalhando em uma versão do programa *hla-mapper* (CASTELLI, et al. 2018) que suporta dados de RNA-seq. Este estudo é o primeiro a utilizar esse protótipo.

Os dados de entrada do *hla-mapper* são os *reads* armazenados em formato FastQ já alinhados ao genoma de referência, ou que estão registrados como “unmapped”. O *hla-mapper* aplica um sistema de comparação e pontuação das sequências com base em um banco de dados de sequências conhecidas, modificando os alinhamentos conforme necessário (CASTELLI, et al. 2018). Isso permite uma avaliação mais confiável da presença de polimorfismos e, no caso do presente projeto, de sequências intrônicas retidas nos transcritos alternativos.

3.3 Análise da presença de íntrons retidos nos mRNA

Com base nas sequências já descritas para o gene HLA-E no banco de dados IPD-IMGT/HLA e dos polimorfismos observados nas mesmas amostras por conta do sequenciamento de DNA, detectamos regiões no início e fim de cada éxon e íntron que não são variáveis. Essas regiões foram usadas para desenhar motivos curtos (20 nucleotídeos) que contemplam sequências de éxon, íntron, e junções éxon/íntron.

Após a correção dos alinhamentos com o hla-mapper, filtramos os reads alinhados ao gene *HLA-E* e contamos o número de ocorrência de cada motivo. Essa contagem utilizou um script em Perl. Em seguida, calculamos a proporção entre o número de ocorrência dos motivos de junção éxon/íntron e o número de ocorrência dos motivos de éxon. Plotamos esses valores utilizando R e ggplot2.

4. RESULTADOS

A figura 5 ilustra a proporção de transcritos com retenção de íntrons, para cada um dos íntrons estudados. Nota-se que todos os íntrons são retidos em algum grau e que há uma grande variabilidade entre amostras, porém o nível de retenção é baixo, de cerca de 4% dos transcritos considerando todos os íntrons. Os íntrons 1 e 2 são mais comumente retidos do que os íntrons 3 a 5, podendo chegar a mais de 5% dos transcritos em algumas amostras.

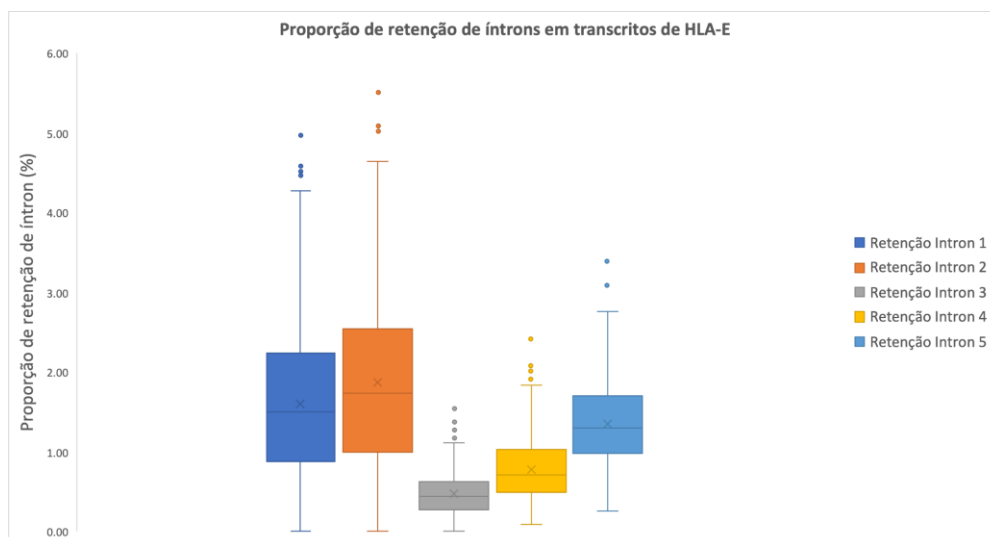


Figura 2. Proporção de transcritos de HLA-E contendo íntrons retidos

O processo de edição alternativa eleva a diversidade de proteínas. Como observamos que todos os íntrons testados são retidos em algum grau, avaliamos quais seriam os produtos proteicos codificados por esses transcritos. Para isso, simulamos a sequência de um transcrito alternativo que carrega todos os éxons e um dos íntrons retidos. Em seguida, traduzimos essa sequência de mRNA para observar qual seria o efeito dessa retenção, assumindo que toda a sequência de íntron é retida. Essa análise está demonstrada na **Figura 3**.

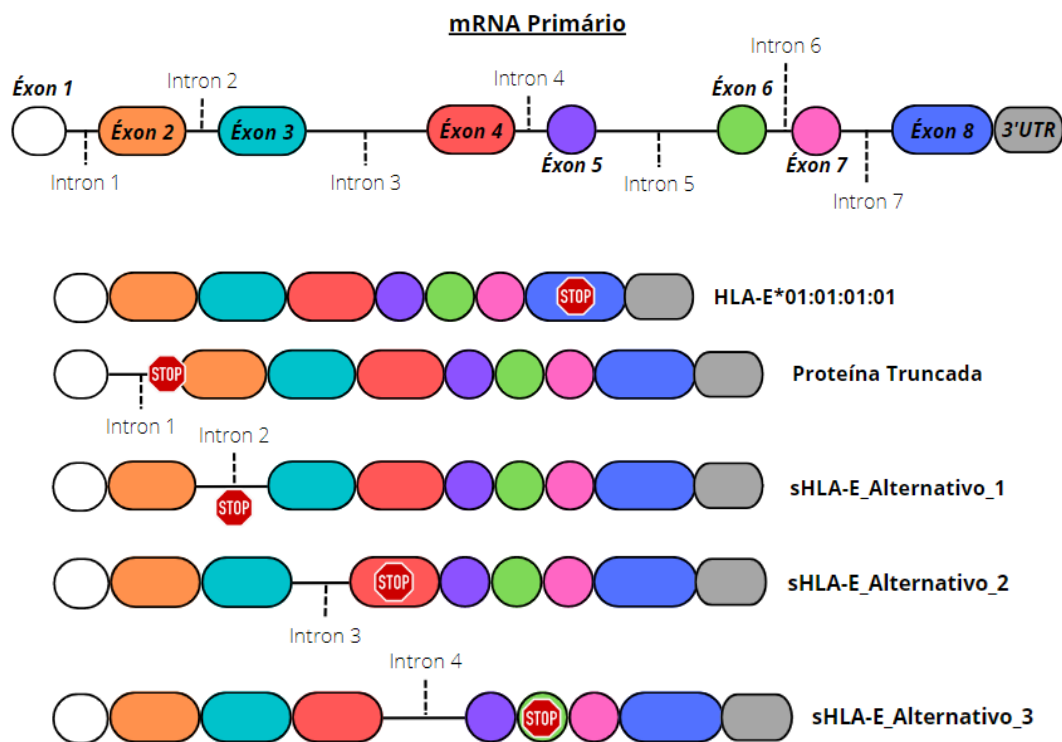


Figura 3. Estrutura do transcrito primário canônico de *HLA-E* e de transcritos alternativos com retenção de íntron. A sequência do transcrito canônico corresponde ao alelo E*01:01:01:01.

A retenção de íntron 1 introduz um códon de parada prematuro no próprio íntron. Se traduzido, teremos a produção de apenas o peptídeo líder, que não é funcional. A retenção do íntron 2 introduz um códon de parada no meio do íntron 2. Sua tradução geraria uma proteína com o domínio $\alpha 1$ e peptídeo sinal, que seria conduzida para a membrana, porém secretada pela ausência do domínio transmembrana. Essa versão gera uma possível versão solúvel de HLA-E, que mantém o domínio $\alpha 1$ (**Figura 4.I**). A retenção no íntron 3 introduz um códon de parada prematuro no éxon seguinte, produzindo uma molécula possivelmente secretada composta domínios $\alpha 1$ e $\alpha 2$ (**Figura 4.II**). A retenção do íntron 4 também introduz um códon de parada prematuro, porém a molécula codificada possui os 3 domínios externos da molécula HLA-E canônica (**Figura 4.III**).

Simulamos as estruturas resultantes da tradução desses transcritos alternativos usando o servidor SWISS-MODEL (<https://swissmodel.expasy.org/>) e o software PyMOL (YUAN et al. 2017), modelando as estruturas proteicas em 3D. Também comparamos as sequências proteicas com as documentadas em bancos de dados públicos, utilizando o BLAST (*Nucleotide* e *Protein*) (BORATYN, et al. 2013).

Observamos que a estrutura dos domínios $\alpha 1$, $\alpha 2$ e $\alpha 3$ é mantida mesmo quando presente nas isoformas solúveis produzidas por esses transcritos alternativos.

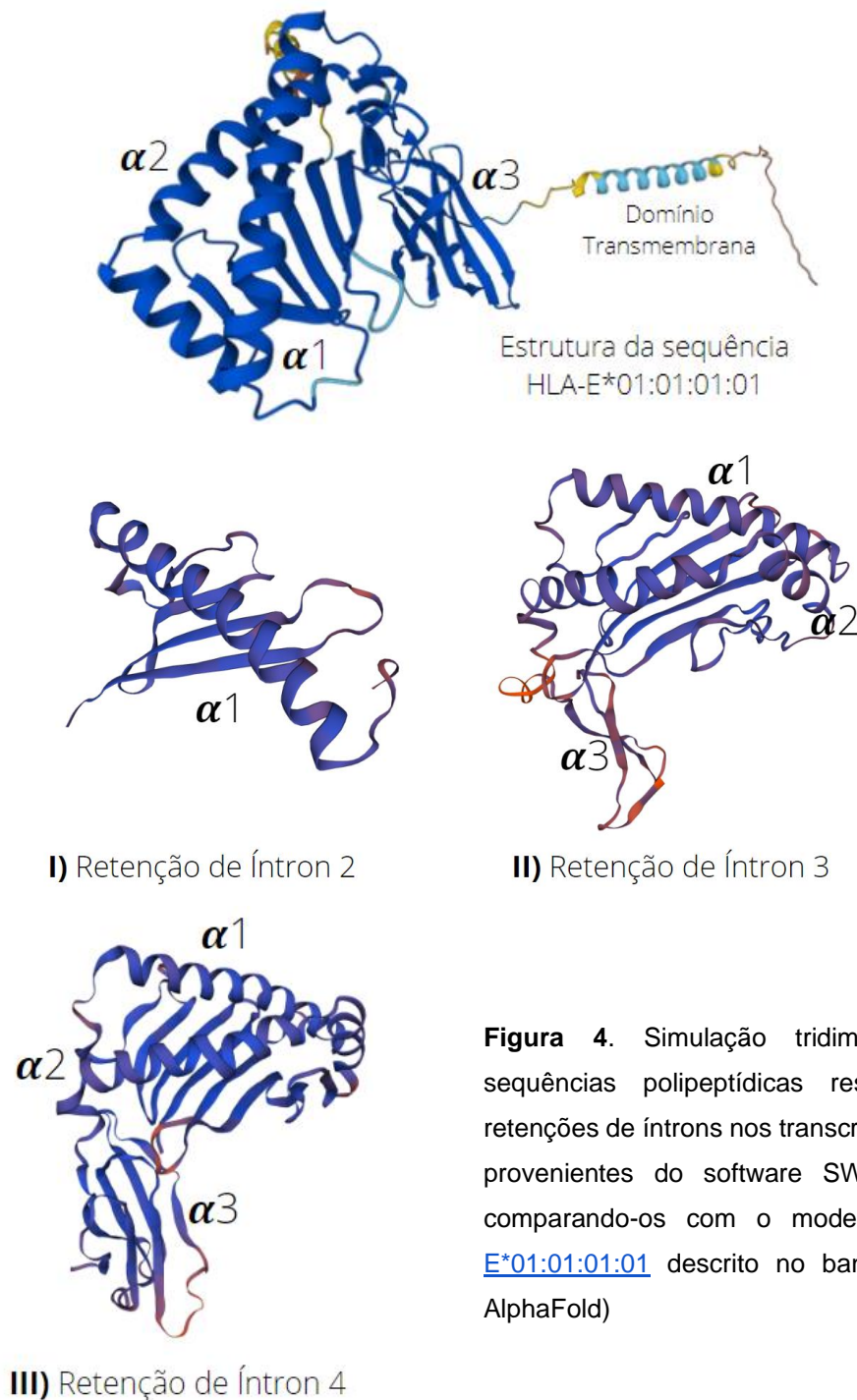


Figura 4. Simulação tridimensional das sequências polipeptídicas resultantes das retenções de íntrons nos transcritos de HLA-E, provenientes do software SWISS MODEL, comparando-os com o modelo de [HLA-E*01:01:01:01](#) descrito no banco de dados AlphaFold)

5. DISCUSSÃO

É notável a relevância de HLA-E na intermediação de processos biológicos, na detecção e no auxílio à contenção de infecções, principalmente virais. Sabe-se que HLA-E, ao ser expressa na superfície das células trofoblásticas, exerce funções essenciais para a construção do limiar de tolerância entre o organismo materno e fetal

e inibição das ações das células NKs contra células fetais, como estudado por King e colaboradores (2000). Portanto, é de extrema importância a compreensão do modo de ação do gene *HLA-E* e a definição das possíveis isoformas da molécula HLA-E, bem como dos mecanismos subjacentes à produção dessas isoformas (que podem ser tecido-específicas, relacionadas a polimorfismos, entre outros).

Já foram descritas isoformas solúveis para a molécula HLA-E, em geral associada com patologias, como a presença de sHLA-E em pacientes com quadros psiquiátricos de Esquizofrenia e Bipolaridade avançados (RIPKE et al. 2014; BOUKOUACI et al. 2021). No entanto, não estão claros os mecanismos (e a origem) dessas moléculas. Neste estudo, a análise do perfil de expressão de HLA-E demonstrou possíveis origens alternativas para produção de sHLA-E, além da possibilidade de clivagem da molécula HLA-E por Mpases.

As análises computacionais demonstraram retenção de todos os íntrons em uma pequena proporção dos transcritos. A retenção dos íntrons introduz códons de parada prematuros, que, por sua vez, previne a tradução do domínio citoplasmático e transmembrana. Essas isoformas alternativas poderiam ser encaminhadas para a superfície celular, pois apresentam o peptídeo líder, e serão secretadas, pois não apresentam o domínio transmembrana. Em todas as situações, a estrutura dos domínios externos ($\alpha 1$, $\alpha 2$ ou $\alpha 3$) é mantida. Portanto, essas possíveis versões solúveis poderiam desempenhar seu papel de modulação da ação de células T e NK.

Embora tenhamos detectado transcritos com capacidade de produção de isoformas solúveis, esta análise foi realizada apenas no nível transcricional, e não está claro se estes transcritos são de fato traduzidos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARKER, D. J. et al. 2023. The IPD-IMGT/HLA Database. **Nucleic Acids Research**. Vol. 51, Issue D1, Pages D1053-D1060. DOI: <https://doi.org/10.1093/nar/gkac1011>.
- BORATYN, G. M. et al. 2013. BLAST: a more efficient report with usability improvements. **Nucleic Acids Res.** 41(Web Server issue): W29-W33. DOI: [10.1093/nar/gkt282](https://doi.org/10.1093/nar/gkt282).
- BOUKOUACI, W. et al. 2021. HLA-E circulating and genetic determinants in schizophrenia and bipolar disorder. **Sci Rep.** 11, 20260. DOI: <https://doi.org/10.1038/s41598-021-99732-9>.
- CASTELLI, E. C. et al. 2018. *Hla-mapper*: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. **Human Immunology**. Vol. 79, Issue 9, pages 678-684. DOI: <https://doi.org/10.1016/j.humimm.2018.06.010>.
- COUPEL, S. et al. 2007. Expression and release of soluble HLA-E is an immunoregulatory feature of endothelial cell activation. **Blood**. 109(7):2806-2814. DOI: <https://doi.org/10.1182/blood-2006-06-030213>.
- DENDROU, C. A. et al. 2018. HLA variation and disease. **Nature reviews immunology**. v. 18, p. 325-339. DOI: <https://doi.org/10.1038/nri.2017.143>.
- DEVUYST, O. 2015. The 1000 Genomes Project: Welcome to a New World. **Perit Dial Int.** 35(7): 676-677. DOI: [10.3747/pdi.2015.00261](https://doi.org/10.3747/pdi.2015.00261).
- DILTHEY, A. T. State-of-the-art genome inference in the human MHC. **The International Journal of Biochemistry & Cell Biology**. v. 131, 105882. 2021. DOI: <https://doi.org/10.1016/j.biocel.2020.105882>.
- DOBIN, A. et al. 2013. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**. 29(1): 15-21. DOI: [10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635).
- DONADI, E. A. et al. 2011. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. **Cell Mol Life Sci.** 68(3): 369-395. DOI: [10.1007/s00018-010-0580-7](https://doi.org/10.1007/s00018-010-0580-7).
- DONG, Y. et al. 2003. Soluble nonclassical HLA generated by the metalloproteinase pathway. **Human Immunology**. v. 64, Issue 8, pages 802-810. DOI: [https://doi.org/10.1016/S0198-8859\(03\)00093-4](https://doi.org/10.1016/S0198-8859(03)00093-4).
- van den ELSEN, P. J. 2011. Expression regulation of Major Histocompatibility complex class I and II encoding genes. **Front Immunol.** v. 2, p. 48. DOI: [10.3389/fimmu.2011.00048](https://doi.org/10.3389/fimmu.2011.00048).
- GAUDINO, S. J. & KUMAR, P. Cross-Talk Between Antigen Presenting Cells and T Cells Impacts Intestinal Homeostasis, Bacterial Infections, and Tumorigenesis. **Front Immunol.** 10: 360. DOI: [10.3389/fimmu.2019.00360](https://doi.org/10.3389/fimmu.2019.00360).
- GONZALEZ-GALARZA, F. F. et al. 2020. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. **Nucleic Acids Res.** 48(D1): D783-d788. DOI: [10.1093/nar/gkz1029](https://doi.org/10.1093/nar/gkz1029).

GUZMÁN-FULGENCIO, M. et al. 2013. *HLA-E* variants are associated with sustained virological response in HIV/hepatitis C virus - coinfecting patients on hepatitis C virus therapy. **AIDS**. v. 27(8), p. 1231-1238. DOI: [10.1097/QAD.0b013e32835f5b9c](https://doi.org/10.1097/QAD.0b013e32835f5b9c).

HANDUNNETTHI, L. et al. 2010. Regulation of MHC class II gene expression, genetic variation and disease. **Genes Immun**. v. 11(2), p. 99-112. DOI: [10.1038/gene.2009.83](https://doi.org/10.1038/gene.2009.83).

IWASZKO, M. & BOGUNIA-KUBIK, K. 2011. Clinical Significance of the HLA-E and CD94/NKG2 Interaction. **Arch. Immunol. Ther. Exp.** 59, 353. DOI: <https://doi.org/10.1007/s00005-011-0137-y>.

JONGSMA, M. L. M., GUARDA, G. & SPAAPEN, R. M. 2019. The regulatory network behind MHC class I expression. **Molecular Immunology**. v. 113, p. 16-21. 2019. DOI: <https://doi.org/10.1016/j.molimm.2017.12.005>.

KALLENBERG, C. G. M. 2002. New insights into the pathogenesis of antineutrophil cytoplasmic autoantibody-associated vasculitis. **Autoimmun Rev.** v.1, Issues 1-2, Pages 61-66. DOI: [https://doi.org/10.1016/S1568-9972\(01\)00008-8](https://doi.org/10.1016/S1568-9972(01)00008-8).

KRAEMER, T., BLASCZYK, R. & BADE-DOEDING, C. 2014 HLA-E: A Novel Player for Histocompatibility. **Journal of Immunology Research**. v. 2014, article ID 352160. DOI: <https://doi.org/10.1155/2014/352160>.

LAPPALAINEN, T. et al. 2013. Transcriptome and genome sequencing uncovers functional variation in humans. **Nature**. 501, 506-511. DOI: <https://doi.org/10.1038/nature12531>.

MANLIK, O. et al. 2015. Fitness & major histocompatibility complex diversity of two bottlenecked dolphin populations. **Thesis project**. DOI: <https://doi.org/10.26190/unsworks/18957>.

MORANDI, F. & PISTOIA, V. 2014. Interactions between HLA-G and HLA-E in physiological and pathological conditions. **Front. Immunol.** v. 5: 394. DOI: [10.3389/fimmu.2014.00394](https://doi.org/10.3389/fimmu.2014.00394).

RIPKE, S. et al. 2014. Biological insights from 108 schizophrenia-associated genetic loci. **Nature**. 511, 421-427. DOI: <https://doi.org/10.1038/nature13595>.

ROBINSON, J. et al. 2001. IMGT/HLA Database - a sequence database for the human major histocompatibility complex. **HLA**. Vol. 55, Issue 3. DOI: <https://doi.org/10.1034/j.1399-0039.2000.550314.x>.

SELIGER, B. et al. 2016. HLA-E expression and its clinical relevance in human renal cell carcinoma. **Oncotarget**. v. 7(14): 67360–67372. DOI: [10.18632/oncotarget.11744](https://doi.org/10.18632/oncotarget.11744).

SHIINA, T. et al. 2009. The HLA genomic loci map: expression, interaction, diversity and disease. **J. Hum. Genet.** v. 54(1), p. 15-39. DOI: <https://doi.org/10.1038/jhg.2008.5>.

The 1000 Genomes Project Consortium. A global reference for human genetic variation. **Nature** 526, 68–74 (2015). DOI: <https://doi.org/10.1038/nature15393>.

The International HapMap Consortium. The International HapMap Project. **Nature**. **426**, 789-796 (2003). DOI: <https://doi.org/10.1038/nature02168>.

WIECZOREK, M. et al. 2017. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. **Frontiers in Immunology**. v. 8, 292. DOI: <https://doi.org/10.3389/fimmu.2017.00292>.

YU, C. Y. et al. 2000. The human and mouse MHC class III region: a parade of 21 genes at the centromeric segment. **Trends in Immunology**. v. 21 (7), p. 320-328. DOI: [DOI:https://doi.org/10.1016/S0167-5699\(00\)01664-9](https://doi.org/10.1016/S0167-5699(00)01664-9).

YUAN, S., CHAN, H.C. & HU, Z. 2017. Using PyMOL as a platform for computational drug design. **WIREs Computational Molecular Science**. Vol. 7, Issue 2, e1298. DOI: <https://doi.org/10.1002/wcms.1298>.