

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA**

ROMÁRIO DE SOUSA MASCARENHAS

**Análise da expressão e caracterização funcional *in silico* de
lncRNAs em células epiteliais tímicas medulares nocauteadas
no gene Aire**

RIBEIRÃO PRETO - SP

2019

ROMÁRIO DE SOUSA MASCARENHAS

**Análise da expressão e caracterização funcional *in silico* de
lncRNAs em células epiteliais tímicas medulares nocauteadas
no gene Aire**

Dissertação apresentada ao Programa de Pós-Graduação em Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

Área de Concentração: Genética

Orientador: Dr. Geraldo Aleixo da Silva
Passos Júnior

RIBEIRÃO PRETO – SP

2019

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Catálogo da Publicação

Mascarenhas, Romário de Sousa

Análise da expressão e caracterização funcional *in silico* de lncRNAs em células epiteliais tímicas medulares nocauteadas no gene Aire.

Ribeirão Preto-SP, 2019.

116p.

Dissertação de Mestrado, apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo. Área de concentração: Genética

Orientador: PASSOS, Geraldo Aleixo

1. Células epiteliais tímicas medulares (mTECs). 2. Gene *Autoimmune regulator* (Aire). 3. RNAs longos não-codificantes (lncRNAs). 4. Sequenciamento de RNA (RNA-Seq).

APOIO E SUPORTE FINANCEIRO

Este trabalho foi realizado no Laboratório de Imunogenética Molecular, localizado no Departamento de Genética da Faculdade de Medicina de Ribeirão Preto (FMRP), da Universidade de São Paulo (USP), com o apoio ou suporte financeiro das seguintes instituições:

- Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP);
 - Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq);
 - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES);
 - Fundação de Apoio ao Ensino, Pesquisa e Assistência – FAEPA;
 - Faculdade de Medicina de Ribeirão Preto – FMRP – USP.
-

Nome: MASCARENHAS, Romário de Sousa

Título: Análise da expressão e caracterização funcional *in silico* de lncRNAs em células epiteliais tímicas medulares nocauteadas no gene Aire.

Dissertação apresentada ao Programa de Pós-graduação em Genética da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

Aprovada em: ___ / ___ / ___

Banca Examinadora

Prof. Dr. _____
Instituição: _____ Assinatura: _____

Prof. Dr. _____
Instituição: _____ Assinatura: _____

Prof. Dr. _____
Instituição: _____ Assinatura: _____

AGRADECIMENTOS

Agradeço a todos e a todas que acreditaram nesse trabalho e de alguma maneira contribuíram para a sua realização.

Agradeço à toda a minha família, em especial à minha mãe Marineide, minha avó Marina e meu avô Joaquim. Obrigado por serem meu porto seguro, por acreditarem nos meus sonhos e não medirem esforços para torná-lo real. Por se fazerem tão presentes apesar dos 1589 Km de distância que separam Ribeirão Preto – SP e Monte Alegre do Piauí - PI.

Ao meu orientador, Professor Dr. Geraldo Aleixo Passos, pela oportunidade, pelos ensinamentos, por fazer bem mais que simplesmente orientar; por me inspirar.

Aos meus amigos e amigas do Grupo de Imunogenética Molecular, pelo conhecimento compartilhado e por todo o apoio no desenvolvimento desse trabalho.

Agradeço a minha amiga e mentora de bioinformática, a Professora Dr^a Amanda Freire de Assis, por todo o suporte com as análises de bioinformática.

Agradeço aos meus amigos e amigas, de perto de e de longe, por se fazerem presentes e tornarem a caminhada mais leve.

Por fim, meus sinceros agradecimentos a todas as pessoas que acreditam na pesquisa científica realizada no Brasil, e que sabem que ciência não é gasto, é investimento.

“...quando se trata de questões científicas, não existe autoridade acima da soberania da ciência. Nem militar, nem política, nem religiosa.”

(GALVÃO, R., 2019)

RESUMO

MASCARENHAS, Romário de Sousa. **Análise da expressão e caracterização funcional *in silico* de lncRNAs em células epiteliais tímicas medulares nocauteadas no gene Aire.** 2019. Dissertação de Mestrado, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

A tolerância imunológica central refere-se à não responsividade aos auto-antígenos e esse mecanismo envolve a maturação e a seleção de linfócitos T no timo. Durante esse processo, os timócitos autorreativos contra antígenos restritos aos tecidos (TRAs) são eliminados por apoptose (seleção negativa) antes de chegar à periferia. Na homeostase, esse processo evita o desenvolvimento da autoimunidade agressiva. O principal gene envolvido no controle da expressão dos TRAs é o *Autoimmune regulator* (Aire), responsável pela modulação de cerca de 40% de todos os TRAs expressos em células epiteliais tímicas medulares (mTECs), pois sua proteína correspondente libera a RNA Pol II na etapa de alongação da transcrição. O Aire é considerado um gene pleiotrópico porque está associado a outras funções além da transcrição dos TRAs. Estudos demonstraram que Aire também atua no controle da expressão de RNAs não codificantes (por exemplo, microRNAs) em mTECs. Nossa hipótese é que o Aire também regula a expressão de RNAs longos não codificantes (lncRNAs) nessas células, já que são transcritos pela RNA Pol II. Este estudo teve como objetivo identificar e caracterizar lncRNAs diferencialmente expressos em células mTEC comparando células nocauteadas no gene Aire com células do tipo selvagem (WT) por meio de algoritmos de predição *in silico*. Utilizamos uma linhagem celular nocauteada no gene Aire (mTEC 3.10E6 Aire KO) obtida anteriormente em nosso laboratório por meio do sistema CRISPR-Cas9. O RNA total das células mTECs Aire KO e mTECs WT foi extraído e depois sequenciado através de RNA-seq (tecnologia Illumina HiSeq 2500). A análise bioinformática dos dados de RNA-Seq análise de qualidade (FASTQC), trimagem por meio do Trimomatic, montagem de genoma (*Mus musculus*) (STAR) e análise estatística no ambiente da plataforma R mostrou a expressão diferencial de lncRNAs que foram então caracterizados funcionalmente. Na ausência de Aire foram identificados 136 lncRNAs diferencialmente expressos (DE), sendo 72 hiperregulados e 64 hiporregulados. Após o ensaio de adesão com timócitos houve redução no número de lncRNAs diferencialmente expressos pelas mTECs Aire KO, foram encontrados somente 33

lncRNAs DE, 26 hiperregulados e 7 hiporregulados. Entre eles, destacamos os lncRNAs Ifi30, Morrbid, Malat1, Xist, Gas5 e Neat1, cujas respectivas funções estão relacionadas ao sistema imunológico. Estes resultados mostram que as células mTEC expressam lncRNAs e que o gene Aire está implicado em sua regulação.

Palavras-chave: Células epiteliais tímicas medulares (mTECs), Gene *Autoimmune regulator* (Aire), RNAs longos não-codificantes (lncRNAs), Sequenciamento de RNA (RNA-Seq).

ABSTRACT

MASCARENHAS, Romário de Sousa. ***In silico* expression analysis and functional characterization of lncRNAs in Aire gene knockout medullary thymic epithelial cells.** 2019. Dissertação de Mestrado, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

Central immunological tolerance concerns non-responsiveness to self-antigens and this mechanism involves maturation and selection of T lymphocytes within the thymus. During this process, those thymocytes autoreactive against self tissue-restricted antigens (TRAs) are eliminated by apoptosis (negative selection) before reaching the periphery. In homeostasis, this process avoids the development of aggressive autoimmunity. The main gene involved in the control of TRA expression is the Autoimmune regulator (Aire), responsible for the modulation of about 40% of all expressed TRAs in medullary thymic epithelial cells (mTECs) by pushing stalled RNA Pol II during the elongation phase of transcription. Aire is considered a pleiotropic gene because it is associated with functions other than TRA transcription. Studies have shown that Aire also acts in the control of the expression of non-coding RNAs (e.g. microRNAs) in mTECs. Our hypothesis is that Aire also regulates the expression of long non-coding RNAs (lncRNAs) in mTECs since they are also transcribed by RNA Pol II. This study aimed to identify and characterize lncRNAs differentially expressed in mTEC cells comparing Aire KO cells with Aire wild type (WT) using *in silico* prediction algorithms. We used an Aire KO mutant mTEC cell line previously obtained in our laboratory using the CRISPR-Cas9 system. The total RNA from Aire KO and WT mTECs was extracted and then sequenced through RNA-seq (Illumina HiSeq 2500 technology). Bioinformatics analysis of RNA-Seq data included quality analysis (FASTQC), trimming (Trimomatic), genome assemble (*Mus musculus*) and statistical analysis (R environment) showed the differential expression of lncRNAs that were further characterized. In the absence of Aire, 136 lncRNAs were differentially expressed (DE), being 72 up-regulated and 64 down-regulated. After thymocyte adhesion there was a reduction in the number of lncRNAs differentially expressed by Aire KO mTECs, only 33 lncRNAs DE, 26 upregulated and 7 downregulated lncRNAs were found. Among these, we highlight the lncRNAs lfi30, Morrbid, Malat1, Xist, Gas5 e Neat1, whose respective functions are related to the immune system. These results

show that mTEC cells express lncRNAs and that the Aire gene is implicated in its regulation.

Keywords: Medullary thymic epithelial cells (mTECs), Autoimmune regulator gene (Aire), Long non-coding RNAs (lncRNAs), RNA Sequencing (RNA-Seq).

LISTA DE FIGURAS

Figura	Descrição	Página
Figura 1	Localização anatômica do timo.	22
Figura 2	Corte histológico do timo.	22
Figura 3	Involução tímica.	23
Figura 4	Representação esquemática dos processos de seleção positiva e negativa.	26
Figura 5	Diversidade de antígenos próprios expressos em células mTECs.	27
Figura 6	Representação da atuação da proteína AIRE no núcleo das mTECs.	28
Figura 7	Representação gráfica da porcentagem de TRAs controlados por Aire e Fezf2 nas células mTECs.	29
Figura 8	Representação dos éxons do gene Aire em humanos e camundongos e os respectivos domínios da proteína AIRE.	30
Figura 9	Representação do mecanismo de ação de AIRE e suas proteínas parceiras na indução da PGE.	31
Figura 10	Descrição das principais manifestações apresentados por pacientes com APECED.	34
Figura 11	Classificação dos lncRNAs baseados em sua localização genômica.	37
Figura 12	Representação dos diferentes mecanismos de ação dos lncRNAs.	38
Figura 13	Classificação dos lncRNAs baseado nas funções.	40
Figura 14	Exemplo de qualidade das sequências.	62
Figura 15	Volcano plot mostrando a distribuição dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire ^{-/-}) antes do ensaio de adesão.	64

Figura 16	Heatmap do agrupamento dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire ^{-/-}) antes do ensaio de adesão.	65
Figura 17	Volcano plot mostrando a distribuição dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire ^{-/-}) após o ensaio de adesão.	67
Figura 18	Heatmap do agrupamento dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire ^{-/-}) após o ensaio de adesão.	68
Figura 19	Diagrama de Venn dos lncRNAs exclusivos e compartilhados antes e após adesão.	70
Figura 20	Proporção e tipos de lncRNAs encontrados nas células mTEC 3.10E6 (Aire ^{-/-}) antes do ensaio de adesão.	72
Figura 21	Proporção e tipos de lncRNAs encontrados nas células mTEC 3.10E6 (Aire ^{-/-}) após o ensaio de adesão.	72
Figura 22	Distribuição cromossômica dos lncRNAs diferencialmente expressos em mTEC 3.10E6 (Aire ^{-/-}) antes da adesão.	73
Figura 23	Distribuição cromossômica dos lncRNAs diferencialmente expressos em mTEC 3.10E6 (Aire ^{-/-}) após a adesão.	73
Figura 24	Enriquecimento funcional dos lncRNAs diferencialmente expressos em células mTEC 3.10E6 (Aire ^{-/-}) antes do ensaio de adesão.	83
Figura 25	Quantificação relativa dos níveis de expressão do lncRNA Peg13 por RT – qPCR.	84
Figura 26	Quantificação relativa dos níveis de expressão do lncRNA Dgkeos por RT – qPCR.	85
Figura 27	Quantificação relativa dos níveis de expressão do lncRNA Fendrr por RT – qPCR.	85

LISTA DE TABELAS

Tabela	Descrição	Página
Tabela 1	<i>Primers</i> utilizados nas reações de RT-qPCR.	60
Tabela 2	Total de <i>reads</i> geradas pelo RNA-seq das amostras de células mTEC 3.10 WT e mTEC 3.10E6 (<i>Aire</i> ^{-/-}).	63
Tabela 3	Total de reads de input e quantidade de <i>reads</i> mapeadas contra o genoma de referência.	63
Tabela 4	Top 10 lncRNAs hiperregulados nas células mTEC 3.10E6 (<i>Aire</i> ^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT sem ensaio de adesão com timócitos.	66
Tabela 5	Top 10 lncRNAs hiporregulados nas células mTEC 3.10E6 (<i>Aire</i> ^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT sem ensaio de adesão com timócitos.	66
Tabela 6	Top 10 lncRNAs hiperregulados nas células mTEC 3.10E6 (<i>Aire</i> ^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT após o ensaio de adesão com timócitos.	69
Tabela 7	lncRNAs hiporregulados nas células mTEC 3.10E6 (<i>Aire</i> ^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT após o ensaio de adesão com timócitos.	69
Tabela 8	lncRNAs diferencialmente expressos compartilhados antes e após adesão com timócitos.	71
Tabela 9	Genes codificadores de proteínas mais próximos dos lncRNAs hiperregulados antes da adesão.	74
Tabela 10	Genes codificadores de proteínas mais próximos dos lncRNAs hiporregulados antes da adesão.	76
Tabela 11	Genes codificadores de proteínas mais próximos dos lncRNAs hiperregulados após o ensaio de adesão mTECs-timócitos.	77
Tabela 12	Genes codificadores de proteínas mais próximos dos lncRNAs hiporregulados após o ensaio de adesão mTECs-timócitos.	78
Tabela 13	Potencial codificador dos lncRNAs hiperregulados antes do ensaio de adesão	79
Tabela 14	Potencial codificador dos lncRNAs hiporregulados antes do ensaio de adesão	80
Tabela 15	Potencial codificador dos lncRNAs hiperregulados após o ensaio de adesão mTECs-timócitos.	82
Tabela 16	Potencial codificador dos lncRNAs hiporregulados após o ensaio de adesão mTECs-timócitos.	83

LISTA DE SIGLAS E ABREVIATURAS

~: aproximadamente

°C: graus Celsius

5': extremidade cinco linha dos ácidos nucleicos

3': extremidade três linha dos ácidos nucleicos

A: adenina

T: timina

C: citosina

G: guanina

μl: microlitro

μg: micrograma

ng: nanograma

μM: micromolar

>: maior

<: menor

≥: maior ou igual

≤: menor ou igual

aa: aminoácido

Aire: gene *Autoimmune regulator*

AIRE: proteína codificada pelo gene *Autoimmune regulator*

APC: célula apresentadora de antígeno

APECED: (Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)

APS1: síndrome autoimune poliglandular do tipo 1

CARD: (Caspase recruitment domain): domínio recrutador de caspase

Cas9: (CRISPR associated protein 9), Proteína 9 associada a CRISPR

CBP: proteína ligadora de CREB

CD80: T-Lymphocyte Activation Antigen CD80

cDNA: DNA complementar

CPC: calculador de potencial codificador

cm²: centímetros quadrados

CRISPR: (Clustered Regularly Interspaced Short Palindromic Repeats) ou Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespaçadas.

cTEC: célula epitelial tímica cortical

DC: Célula dendrítica

DN: Duplo negativo

DNA: ácido desoxirribonucleico

DNA-PK: proteína quinase

DP: Duplo positivo

EpCAM (Epithelial cell adhesion molecule), Molécula de adesão celular epitelial

FASTQ: formato utilizado para representar sequências geradas por sequenciadores.

FDR (False Discovery Rate): Taxa de falsa descoberta

gRNA: RNA guia

H3K4me3: trimetilação da lisina 4 da histona H3

Indel: Inserções/deleções

IL: interleucina

IFN: interferon

kb: kilobase

kDa: kilodaltons

KO: *knockout*

LES: lúpus eritematoso sistêmico

lncRNA: RNA longo não codificante

mL: mililitro

mM: milimolar

MHC: complexo de histocompatibilidade principal

miRNA: MicroRNA

mRNA: RNA mensageiro

mTEC: célula epitelial tímica medular

ncRNA: RNA não codificante

NF- κ B (Nuclear Factor Kappa B): Fator Nuclear Kappa B

NLS (Nuclear signal localization): Sinal de Localização Nuclear

nt: Nucleotídeo

ORF: matriz de leitura aberta

PAGE (Polyacrylamide gel electrophoresis): Eletroforese em gel de poliacrilamida

pb: pares de bases

PBS: Solução salina tamponada com fosfato

PCR: reação em cadeia da polimerase

PGE: expressão gênica promíscua

PHD (plant homeodomain): Homeodominio de planta

PRR: Região rica em prolina

P-TEFb: fator b alongador positivo da transcrição

RT-qPCR: reação de polimerização em cadeia quantitativa e em tempo real

RANK (Receptor Activator of Nuclear Factor Kappa B): Receptor ativador de fator nuclear Kappa B

Read: fragmento sequenciado.

RIN: número de integridade do RNA

RNA pol II: RNA polimerase II

rRNA: RNA ribossômico

RNA-Seq: sequenciamento de RNA

RNAi: RNA de interferência

SAND: domínio denominado SP100, AIRE1, nucP41/P75 and DEAF1

SDS: Dodecil sulfato de sódio

SFB: soro fetal bovino

SP: Simples Positivo

TCR: receptor de célula T

TEC: célula tímica epitelial

Th: linfócitos T helper

TNF: fator de necrose tumoral

TRA: Antígeno restrito ao tecido

Treg: Célula T reguladora

tRNA: RNA transportador

U: unidade

V: volts

V(D)J: v = variável, d = diversidade, j = junção.

WT: Wild-type

SUMÁRIO

1. INTRODUÇÃO	21
1.1 O Sistema imunológico e o timo	21
1.2 Tolerância imunológica central: seleção positiva e negativa de timócitos	24
1.3 Expressão Gênica Promíscua (PGE)	26
1.4 O gene <i>Autoimmune regulator</i> (Aire): do DNA à Proteína	29
1.5 Funções moleculares e biológicas de Aire	31
1.6 APECED/APS-1	32
1.7 Aire e RNAs não-codificantes	35
1.8 RNAs longos não-codificantes (lncRNAs): histórico e características	35
1.9 lncRNAs e o sistema imunológico	40
2. HIPÓTESE	43
3. OBJETIVOS	45
3.1 Objetivo Geral	45
3.2 Objetivos específicos	45
4. DELINEAMENTO EXPERIMENTAL	47
5. MATERIAL E MÉTODOS	49
5.1 Linhagens celulares	49
5.1.1 Células mTEC 3.10 Wild type (WT)	49
5.1.2 Células mTEC 3.10E6 (<i>Aire</i>^{-/-})	49
5.2 Obtenção dos timócitos por timectomia	50
5.3 Ensaio de adesão mTECs-timócitos	50
5.4 Extração do RNA total	51
5.5 Quantificação e avaliação do grau de pureza das amostras de RNA total ..	52
5.6 Análise da integridade das preparações de RNA total	52
5.7 Construção da biblioteca e sequenciamento do RNA	53
5.8 Análise bioinformática dos lncRNAs	54
5.8.1 1ª análise de qualidade	54
5.8.2 Trimagem	54
5.8.3 2ª análise de qualidade	55
5.8.4 Indexação do genoma de referência	55
5.8.5 Montagem do transcriptoma	55
5.8.6 Contagem das reads mapeadas	55
5.8.7 Análise dos lncRNAs diferencialmente expressos	56
5.9 Caracterização e enriquecimento funcional dos lncRNAs	56

5.9.1 Identificação e posicionamento dos genes codificadores mais próximos dos lncRNAs	57
5.9.2 Análise do potencial codificador dos RNA longos não codificantes	57
5.9.3 Enriquecimento funcional dos lncRNAs	57
5.10 Validação dos lncRNAs por RT-qPCR	58
5.10.1 Transcrição reversa para síntese do cDNA.....	58
5.10.2 Confirmação da transcrição reversa por PCR convencional.....	58
5.10.3 Eletroforese em gel de agarose	58
5.10.4 PCR quantitativa em tempo real (RT-qPCR).....	59
6. RESULTADOS	62
6.1 Análise da qualidade	62
6.2 Trimagem	63
6.3 Montagem do genoma	63
6.4 Análise dos lncRNAs diferencialmente expressos	64
6.5 lncRNAs exclusivos e compartilhados antes e depois da adesão	70
6.6 Classificação e proporção dos tipos de lncRNAs	72
6.7 Distribuição Cromossômica dos lncRNAs	73
6.8 Identificação dos genes codificadores de proteínas mais próximos	74
6.9 Determinação do potencial codificador dos lncRNAs.....	78
6.10 Enriquecimento funcional.....	83
6.11 Validação da expressão dos lncRNAs por RT- qPCR.....	84
7. DISCUSSÃO	87
8. CONCLUSÃO.....	94
9. REFERÊNCIAS BIBLIOGRÁFICAS.....	96
10. APÊNDICE	108
10.1 Apêndice 1 – Lista completa dos lncRNAs hiperregulados antes do ensaio de adesão	108
10.2 Apêndice 2 – Lista completa dos lncRNAs hiporregulados antes do ensaio de adesão	111
10.3 Apêndice 3 – Lista completa dos lncRNAs hiperregulados após do ensaio de adesão	113
10.4 Apêndice 4 – Lista completa dos lncRNAs hiporregulados após do ensaio de adesão	114
11. ANEXOS	116
11.1 Anexo 1- Certificação da comissão de ética em experimentação animal	116

INTRODUÇÃO

1. INTRODUÇÃO

1.1 O Sistema imunológico e o timo

O sistema imunológico, ou sistema imune, é um sistema que tem como função primordial a defesa do organismo frente a invasão de agentes patogênicos; os processos biológicos de defesa que o sistema imunológico atua são realizados por órgãos e células específicas que interagem entre si, a fim de garantir a proteção e homeostasia do organismo (SATTLER, 2017).

Nesse contexto de defesa do organismo frente ao ataque de organismos causadores de doenças, é necessário que o sistema imune consiga fazer a distinção daquilo que é próprio, ou seja, antígenos inerentes ao organismo, que devem ser identificados e não sofrer ataque das células de defesa, daquilo que é considerado não-próprio, ou seja, os antígenos estranhos ao organismo, que devem ser reconhecidos como tal e conseqüentemente eliminados (KLEIN et al., 2014; TAKABA; TAKAYANAGI, 2017).

Para que aconteça a distinção próprio – não próprio, é crucial que células especializadas na expressão e/ou apresentação de antígenos próprios atuem corretamente, para que não ocorra o ataque contra esses antígenos, também conhecidos como antígenos restritos aos tecidos (TRAs, do inglês, *tissue restricted antigens*) (PASSOS et al., 2018). O processo supracitado ocorre em um órgão fundamental para correto funcionamento do sistema imune: o Timo (MUNOZ; ZAPATA, 2019).

O timo é um órgão linfoide primário, uma vez que é responsável pelo desenvolvimento e maturação de Linfócitos T. Está situado na região anterossuperior do mediastino e posterior ao osso esterno, possui dois lobos, e cada um deles é formado por múltiplos lóbulos menores que são divididos histologicamente em duas regiões, uma região mais externa denominada córtex tímico e uma região mais interna chamada de medula tímica (Figura 1 e Figura 2) (MUNOZ; ZAPATA, 2019; GEENEN; SAVINO, 2019).

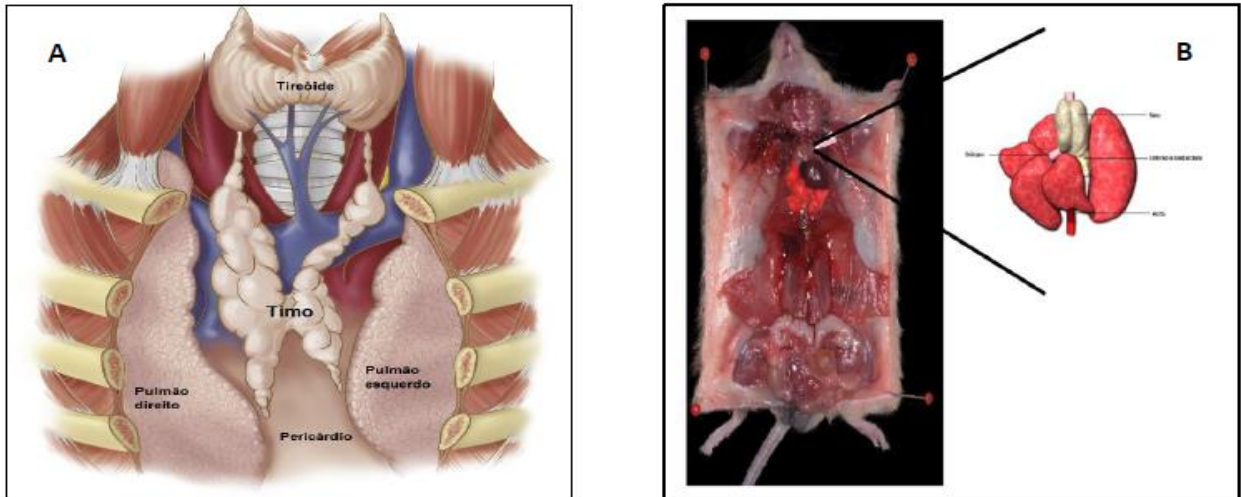


Figura 1. Localização anatômica do timo. A) Posição anatômica do timo em humanos. B) Posição anatômica do timo em camundongos. (RUEHL-FEHLERT *et al.*, 2003; SAFIEDDINE & KESHAVJEE, 2011)

Cada uma dessas regiões formam um microambiente especializado, com processos que permitem a maturação e seleção de linfócitos T. A correta manutenção do microambiente tímico exige que ocorra interações recíprocas entre os precursores de linfócitos T em maturação e as células estromais, processo denominado de "crosstalk" tímico (IRLA, 2019; PEZZI *et al.*, 2016).

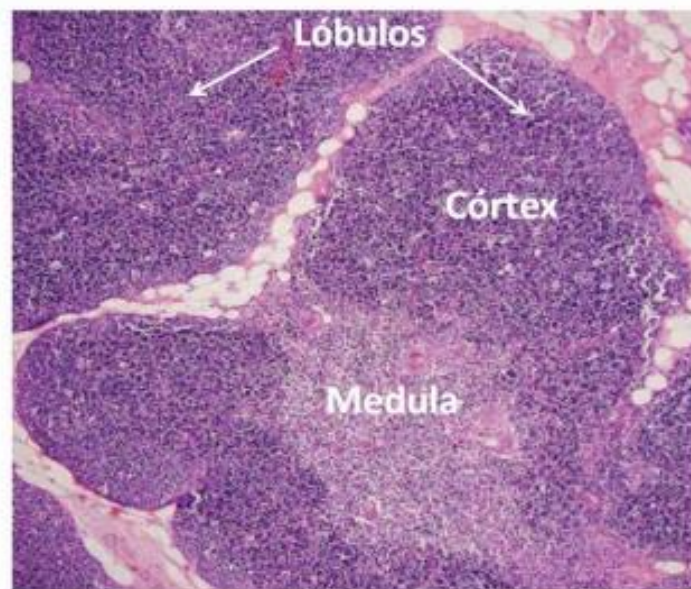


Figura 2. Corte histológico do timo. Foto histológica corada com hematoxilina-eosina demonstrando a região do córtex tímico, mais externamente, e a região da medula tímica, mais internamente (Adaptado de NISHINO *et al.*, 2006).

Em termos celulares, o timo é constituído majoritariamente por células epiteliais tímicas (TECs, do inglês, *thymic epithelial cells*), mas outras células também são encontradas, como as células dendríticas, macrófagos, fibroblastos, células mesenquimais, dentre outras (PASSOS et al., 2018).

Um das características mais marcantes do timo, é a sua involução, o que faz com ele seja pouco conhecido e acabe por ser negligenciado e pouco estudado, no Brasil e no mundo. A involução está diretamente relacionada a idade, ou seja, à medida que envelhecemos o estroma tímico é substituído por tecido adiposo e fibroblastos (MUNOZ & ZAPATA, 2019).

Em humanos a involução tímica se inicia após o primeiro ano de vida e a involução absoluta ocorre na puberdade, em camundongos a involução tem início no período de maturidade sexual (~40 dias de vida), porém mesmo após a involução, o timo de indivíduos adultos continua a receber células precursoras provenientes da medula óssea e a lançar células T maduras para as áreas periféricas (Figura 3). Como resultado do processo de involução temos a perda de citocinas e hormônios críticos para a manutenção do microambiente tímico, redução da imunidade adaptativa e da produção de células T *naive*, declínio no repertório do receptor de células T periféricas (TCR, do inglês, *T-cell receptor*), perda ou perturbação do *crosstalk* entre os linfócitos em desenvolvimento e o estroma tímico, depleção dos progenitores de células T, dentre outras (LYNCH, 2009; GEENEN; SAVINO, 2019).

Entre os fatores responsáveis pela involução do timo relacionada à idade temos a redução nos níveis de expressão do gene *Foxn1*, um fator de transcrição específico das células tímicas epiteliais. Foi demonstrado que a regulação positiva de *Foxn1* em camundongos com timo totalmente involuído resultou na regeneração total do timo (CHEN et al., 2009; BREDENKAMP et al., 2014).

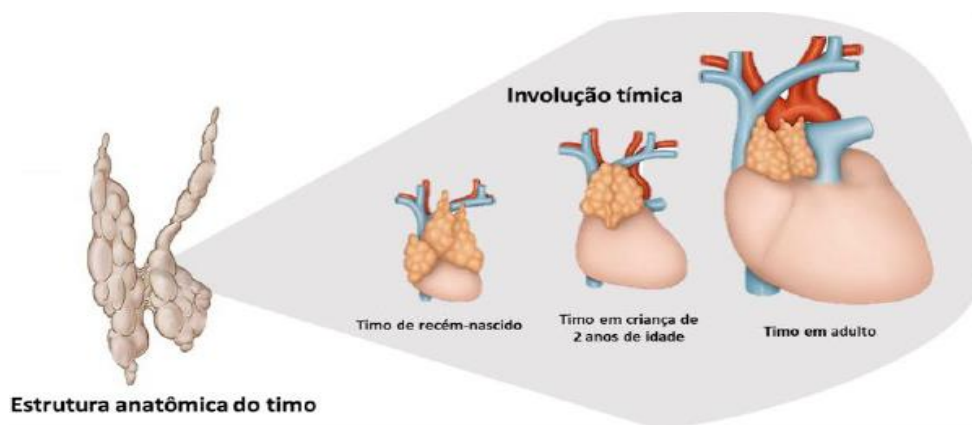


Figura 3. Involução tímica. Adaptado de Musculoskeletal Key, site: <https://musculoskeletalkey.com/>. Acesso em 05 de Agosto de 2019.

1.2 Tolerância imunológica central: seleção positiva e negativa de timócitos

A tolerância imunológica central diz respeito a não responsividade aos antígenos próprios, cujo mecanismo envolve a maturação e seleção de linfócitos T. Tem por objetivo selecionar linfócitos que sejam funcionais e que não gerem respostas contra autoantígenos, estando esses aptos a atuarem na periferia. É um processo essencial para a homeostase imunológica, sendo dividido em duas etapas, a seleção positiva e a seleção negativa (Figura 4) (PASSOS et al., 2018; PASSOS et al., 2019).

A seleção positiva ocorre na região do córtex tímico e busca selecionar timócitos que possuem TCR funcional, capaz de reconhecer e se ligar ao complexo peptídeo-MHC (MHC, do inglês, *major histocompatibility complex*), as principais células mediadoras desse processo são as células epiteliais tímicas corticais (cTECs, do inglês, *cortical thymic epithelial cells*). Nessa etapa as células precursoras de linfócitos T oriundas da medula óssea adentram o timo através da junção cortico-medular (JCM), no timo essas células passam a ser denominadas timócitos. (OHIGASHI; KOZAI; TAKAHAMA, 2019). Os timócitos recém-chegados ao timo não expressam as moléculas CD4 e CD8, sendo denominados duplo-negativos (DN). Os timócitos DN migram para o córtex do timo, onde inicialmente expressam os TCRs $\gamma\delta$ e $\alpha\beta$, lá ocorre um primeiro estágio de maturação, onde células T $\alpha\beta$ se diferenciam em células T CD4 e CD8, sendo CD4+ restrito ao MHC de classe II, e CD8+ restrito ao MHC de classe I, com isso passam a ser denominados timócitos duplo positivos (DP). (YOGANATHAN et al., 2019).

Além de expressar CD4 e CD8, os timócitos DP passam pelo processo de recombinação V(D)J, no qual ocorre a recombinação somática total dos genes das cadeias α e β dos seus receptores TCR, o que resulta num repertório diverso de TCRs (YOGANATHAN et al., 2019).

Durante a seleção positiva dos timócitos é necessário que haja uma interação física estável entre eles e as cTECs, essa interação ocorre por meio dos TCRs nos timócitos com uma molécula de MHC, que pode ser tanto de classe I quanto de classe II, associada a um antígeno próprio nas cTECs (OHIGASHI; KOZAI; TAKAHAMA, 2019; PASSOS et al., 2018).

Na seleção positiva, os timócitos cujos TCRs reconhecem e se ligam aos complexos peptídeos-MHC do hospedeiro com baixa avides, serão considerados

funcionais e selecionados positivamente, a ligação com pouca avidéz se faz necessária para garantir que os timócitos selecionados não sejam capazes de reagir fortemente contra antígenos próprios durante a apresentação dos TRAs; por outro lado aqueles timócitos que possuem TCRs incapazes de reconhecer ou se ligar aos complexos peptídeos-MHC do hospedeiro, serão eliminados por apoptose, processo denominado morte por negligência (TAKAHAMA et al., 2010; PASSOS et al., 2018).

Ainda no córtex tímico, os timócitos DP selecionados positivamente, passam por um novo processo de maturação, onde se diferenciam em timócitos CD4+ ou CD8+, sendo denominados timócitos simples positivo (SP), a partir disso essas células SP migram para a medula tímica, onde passarão pelo processo de seleção negativa (MENDES-DA-CRUZ et al., 2019).

Na seleção negativa dos timócitos, ocorre o reconhecimento dos TRAs, ou seja, dos antígenos próprios do organismo, nessa etapa aqueles timócitos que reconhecem os autoantígenos e se ligam fortemente a eles são eliminados, para evitar que esses atinjam a periferia e não ocorra reação imunológica agressiva contra os tecidos próprios (PASSOS *et al.*, 2018). Alguns timócitos que reconhecem autoantígenos ao invés de serem eliminados se diferenciam em células T reguladoras (Tregs) e passam a atuar na regulação da tolerância imunológica periférica, esse mecanismo de diferenciação ainda não foi elucidado (LI et al., 2016).

Como foi supracitado, a seleção negativa ocorre na medula tímica, o principal grupo de células responsáveis pela seleção negativa são as células epiteliais tímicas medulares (mTECs), essas possuem a capacidade de expressar e apresentar uma vasta gama de antígenos dos diferentes órgãos e tecidos do organismo, mecanismo denominado expressão gênica promíscua (PGE) (ST-PIERRE et al., 2015; TAKABA et al., 2015)

Outros tipos celulares, como as células dendríticas e os macrófagos também atuam no processo de seleção negativa. No entanto, essas células não têm capacidade de expressarem os TRAs, apenas de apresenta-los às mTECs, e são consideradas células apresentadoras de antígenos (APCs). De acordo com a literatura, os autoantígenos são transportados das mTECs até as células dendríticas e macrófagos por meio de moléculas de exossomas, esses antígenos são então englobados por essas células e apresentados aos timócitos (SKOGBERG; TELEMO; EKWALL, 2015).

Após o processo de apresentação de TRAs os timócitos SP CD4+ ou CD8+ que passaram pela seleção negativa e não se ligaram fortemente aos autoantígenos agora estão maduros, então eles migram para a periferia onde irão atuar, afim de garantir a proteção do organismo contra antígenos não próprios (KLEIN *et al.*, 2014; PASSOS *et al.*, 2018).

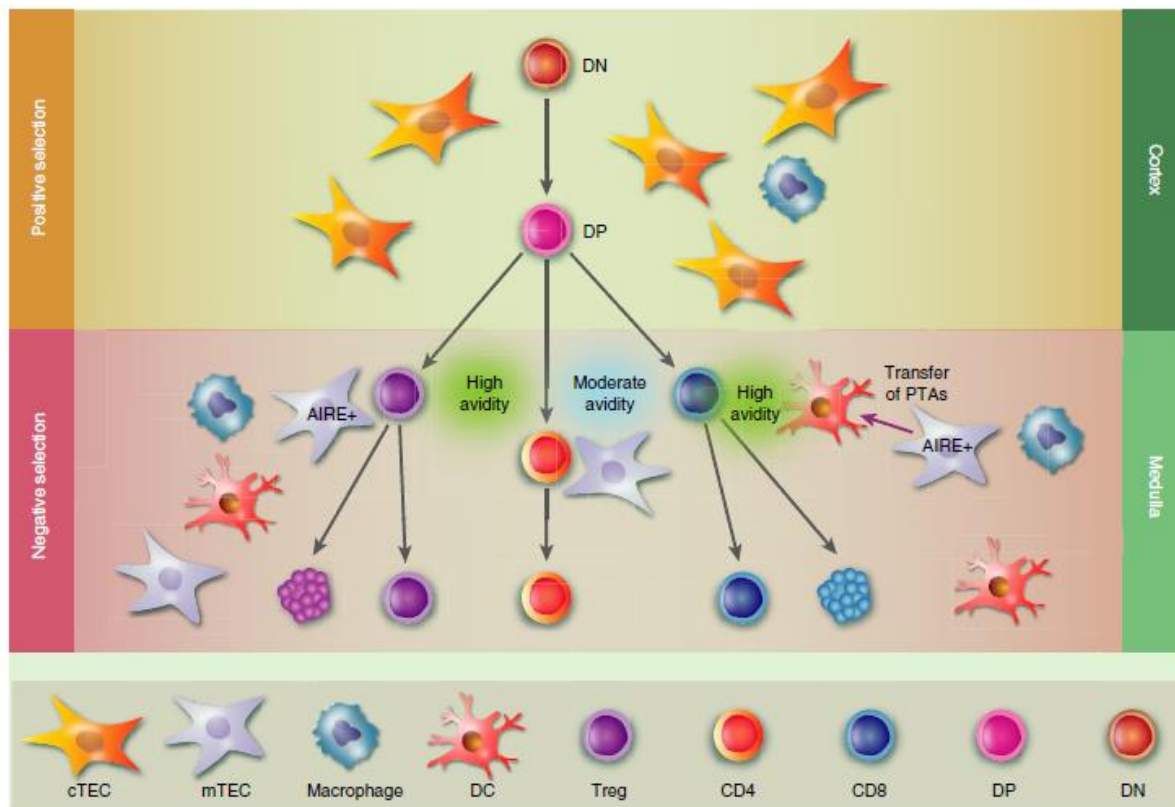


Figura 4. Representação esquemática dos processos de seleção positiva e negativa, que ocorrem no timo. Timócitos DN (CD4- e CD8-) passam pelo processo de seleção positiva no córtex do timo, onde são testados, pelas cTECs, quanto a sua capacidade de reconhecer fracamente o MHC próprio, os timócitos que falham nesse processo sofrem o processo apoptótico. Após a seleção positiva os timócitos que sobreviveram se tornam DP (CD4+ e CD8+) e migram para a medula do timo, onde passarão pela seleção negativa. Na medula as mTECs irão expressar uma variedade de TRAs e, com o auxílio de outras APCs, irão apresentá-los aos timócitos, os que reconhecem com alta afinidade algum desses autoantígenos, são eliminados por apoptose ou se diferenciam em células Tregs. Os timócitos que sobrevivem a esse processo se tornam SP (CD4+ e CD8- / CD4- e CD8+) e são liberados para a periferia, onde atuarão (PASSOS *et al.*, 2018).

1.3 Expressão Gênica Promíscua (PGE)

A principal característica das células mTECs é a capacidade de expressar e apresentar antígenos de outros órgãos e tecidos, a chamada PGE. Isso garante que os autoantígenos sejam apresentados aos timócitos, para que aqueles que forem

autorreativos durante a seleção negativa sejam eliminados, isso permite que a tolerância central se estenda para todo o organismo (SANSOM et al., 2014).

A expressão gênica promíscua é altamente conservada em humanos e camundongos. Em ambas as espécies, as células mTECs são consideradas o tipo celular que expressa o maior número de genes (Figura 5). Em camundongos, por exemplo, do total de cerca de 25 mil genes, as mTECs expressam aproximadamente 19 mil genes, o que representa mais de 70% do genoma murino (DERBINSKI et al, 2016).

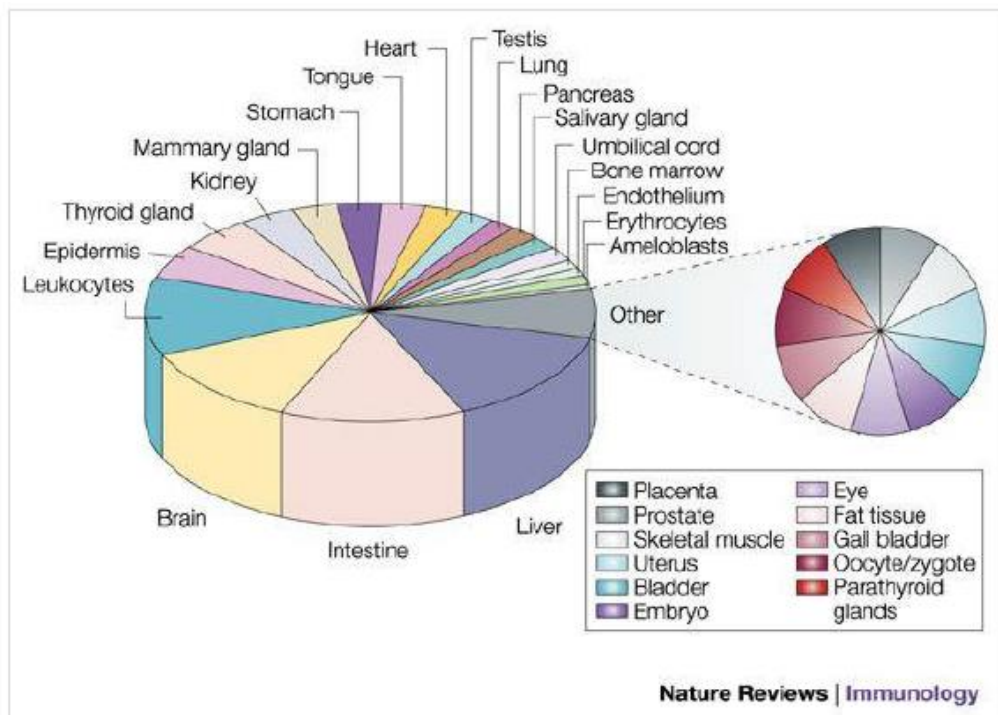


Figura 5. Diversidade de antígenos próprios expressos em células mTECs. Capacidade de expressão e apresentação de milhares de TRAs, que representam todos os tecidos do corpo (KYEWSKI; DERBINSKI, 2004).

O estágio de maturação das mTECs é diretamente proporcional a diversidade de TRAs e a PGE. Com isso as chamadas mTECs^{HI}, ou seja, aquelas que apresentam altos níveis de expressão de MHCII, CD80 e AIRE (CD80^{hi}MHCII^{hi}AIRE⁺) são mais aptas à PGE em comparação com mTECs^{LO} (CD80^{lo}MHCII^{lo}AIRE⁻) (TAKABA; TAKAYANAGI, 2017). É válido ressaltar que uma única mTEC não expressa todos os TRAs do organismo, cada TRA só é expresso por uma pequena fração (1-3%) das mTECs, assim na medula tímica, essas células se agrupam em subpopulações que

expressam TRAs diferentes, contemplando todo o repertório que deve ser expresso e apresentado aos timócitos (Figura 6) (SANSOM et al., 2014, ST-PIERRE et al., 2015).

Os mecanismos moleculares responsáveis pela regulação da tolerância imunológica central, sobretudo da expressão gênica promíscua, ainda não estão totalmente elucidados, sendo que diversos genes expressos pelas mTECs atuam na regulação desse processo (KLEIN et al., 2014).

A literatura aponta que o principal gene envolvido na PGE é o gene *Autoimmune regulator* (Aire), ele codifica uma proteína de mesmo nome, que é responsável pela modulação de cerca de 30% dos TRAs expressos, esses são denominados TRAs Aire dependentes (ANDERSON; SU, 2014; TAKABA; TAKAYANAGI, 2017).

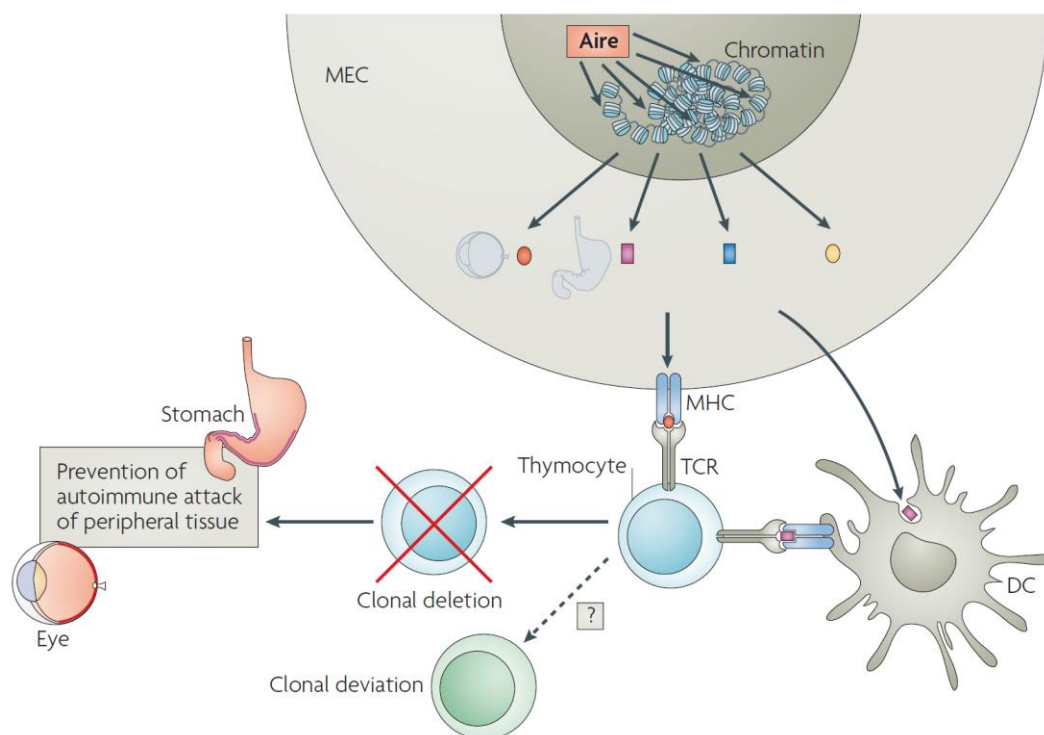


Figura 6. Representação da atuação da proteína AIRE no núcleo das mTECs. AIRE atua no núcleo das mTECs promovendo a expressão de TRAs dos diferentes tecidos e/ou órgãos. Esses TRAs são então apresentados via MHC ao timócitos durante a seleção negativa. Os clones que reconhecem com alta afinidade esses TRAs são eliminados por apoptose. (MATHIS; BENOIST, 2007).

Outro gene recentemente relacionado a esse processo, é o gene *Fezf2* responsável por modular a expressão de aproximadamente 22% dos TRAs, os *Fezf2* dependentes, e há ainda aqueles TRAs que tem sua expressão modulada por Aire e *Fezf2* concomitantemente, que correspondem a cerca de 12% dos TRAs expressos nas mTECs. Assim, Aire e *Fezf2* juntos são responsáveis pelo controle da expressão

de aproximadamente 60% dos TRAs. Por sua vez os responsáveis pela regulação dos 40% dos TRAs restantes ainda não são conhecidos sendo interesse de estudo de diversos grupos de pesquisas no mundo (Figura 7) (KLEIN et al., 2014; TAKABA et al., 2015).

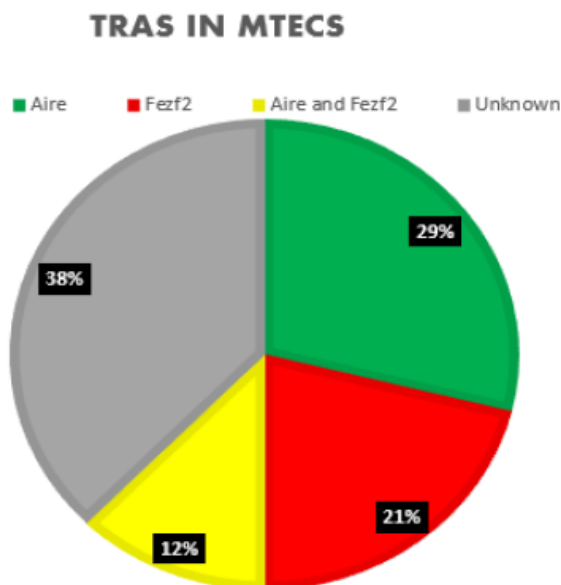


Figura 7. Representação gráfica da porcentagem de TRAs controlados por Aire e Fezf2 nas células mTEC (TAKABA et al., 2015).

1.4 O gene *Autoimmune regulator* (Aire): do DNA à Proteína

No fim da década de 90 variantes patogênicas de um determinado gene foram relacionadas com uma síndrome autoimune denominada *autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy syndrome* (APECED) ou autoimune polyglandular syndrome type 1 (APS1). Essa síndrome se caracteriza por uma grande invasão de linfócitos em diversos órgãos, principalmente os glandulares (PETERSON et al., 2004; CHENG; SHUM; ANDERSON, 2007).

Em 1997 dois grupos de pesquisa apontaram que essa síndrome era causada por mutações em um único gene, e que esse gene estava envolvido na regulação da autoimunidade, assim foi denominado *Autoimmune regulator* (*Aire*) (AALTONEN et al., 1997, NAGAMINE et al., 1997). O gene *Aire* em humanos está localizado no 21q22.3, possui 11.9 kb, distribuídos ao longo de 14 éxons com 1635 nucleotídeos. No genoma murino *Aire* localiza-se no cromossomo 10, posição 39_72 cM, com 13

kb e possui 14 éxons. Quando comparados, os genomas humano e murino possuem cerca de 80% de similaridade (BRUSERUD et al., 2016; PASSOS et al., 2018).

Em humanos a proteína AIRE possui 56kDa, é formada por 545 aminoácidos e 5 domínios: caspase-activation and recruitment domain (CARD), nuclear localization signal (NLS), SAND domain (Sp100, Aire1, Nupc41/p75, e Deaf1) e plant-homeodomains (PHD1 e PHD2) que estão separado por um região rica em prolina (PRR). A proteína AIRE em camundongos possui os mesmos domínios da AIRE em humanos (Figura 8) (ABRAMSON et al., 2010).

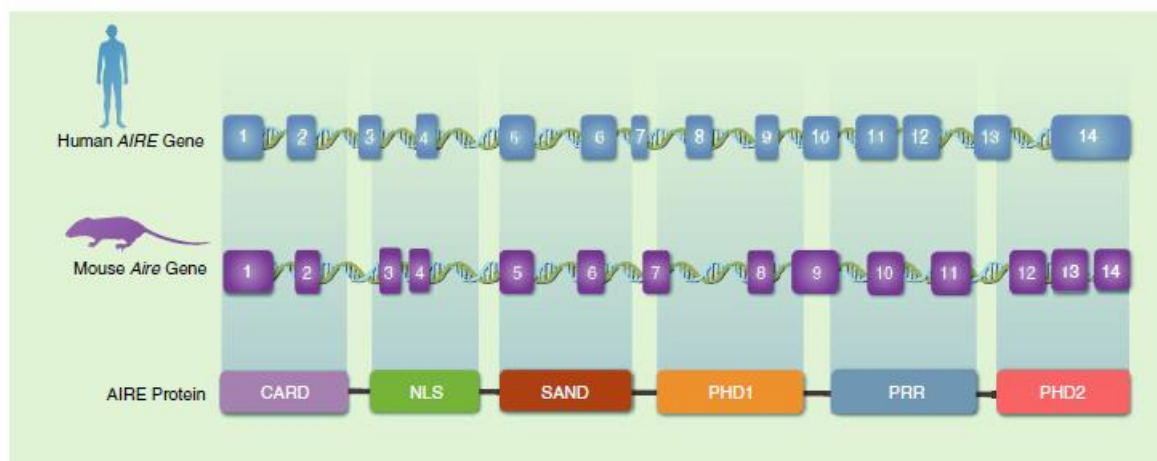


Figura 8. Representação dos éxons do gene Aire em humanos e camundongos e os respectivos domínios da proteína AIRE (PASSOS et al., 2018).

Os domínios de AIRE ainda não têm todas as suas funções conhecidas. Sabe-se que o domínio CARD atua na oligomerização proteica, pois AIRE atua como um tetrâmero; NLS como o nome sugere é um importante sinalizador nuclear e permite a migração de AIRE do citoplasma para o núcleo das células; o domínio SAND por sua vez medeia a interação entre proteínas; por fim, os domínios PHD1 e PHD2, que são leitores de histonas e auxiliam no afrouxamento da cromatina (PASSOS et al., 2018; PERNIOLA; MUSGO, 2014).

Aire é expresso preferencialmente no timo, apenas em alguns conjuntos de células tímicas, como as mTECs que possuem em sua superfície C80^{hi} e MHC-II^{hi} e as células dendríticas (DC). Além disso, a expressão de Aire foi observada em órgãos linfoides secundários como os linfonodos (ST-PIERRE et al., 2015; PERNIOLA, 2018).

1.5 Funções moleculares e biológicas de Aire

AIRE é um fator de transcrição, ou seja, atua juntamente com a RNA polimerase II (RNA Pol II) e permite que determinadas regiões do DNA sejam transcritas em um RNAm. No entanto, AIRE é considerada um fator transcricional não-canônico, pois ao invés de se ligar a região promotora de um gene e iniciar o processo transcricional, AIRE interage com a RNA polimerase II que já reconheceu a região que deve ser transcrita, no entanto está ancorada na cromatina nos sítios específicos T-boxes (TTATTA) e G-boxes (ATTGGTTA) das regiões promotoras dos genes AIRE dependentes. AIRE também recruta outras proteínas parceiras e esse complexo formado por AIRE e suas parceiras vai “empurrar” a RNA polimerase II que está ancorada e assim ocorre a elongação da transcrição gênica (Figura 9) (GIRAUD et al., 2012; PASSOS et al., 2018).

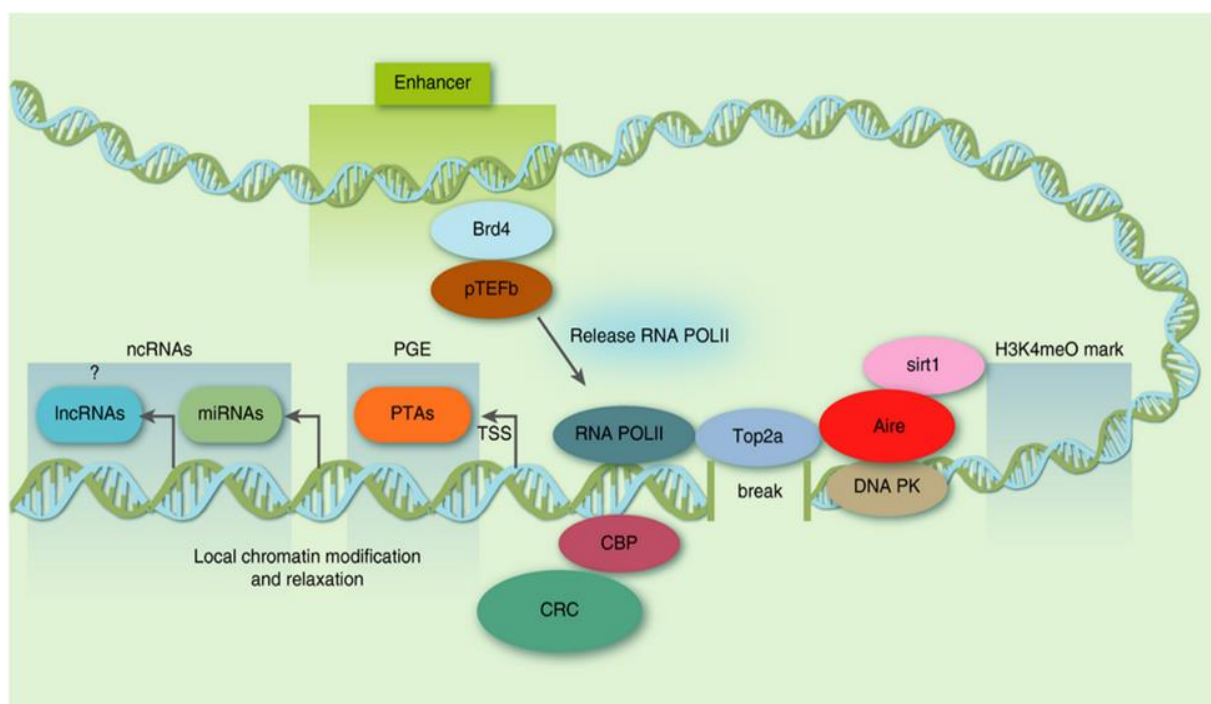


Figura 9. Representação do mecanismo de ação de AIRE e suas proteínas parceiras na indução da PGE. Sirt-1 deacetila resíduos de lisina de AIRE promovendo sua ativação, AIRE então reconhece marcas epigenéticas de repressão da transcrição (H3K4me0), se liga à essas marcas e recruta outras proteínas. TOP2a promove a quebra da fita dupla, DNA PKs em conjunto com suas parceiras promove o afrouxamento da fita e o complexo Brd4 e pTEFb libera a RNA pol II que está “ancorada” na cromatina, permitindo a transcrição de diversos genes. Por fim, a CBP acetila os resíduos de lisina de AIRE, inativando-a (PASSOS et al., 2018).

Dentre as proteínas parceiras de AIRE temos a Deacetylase sirtuin 1 (SIRT-1), que é responsável por sua ativação, por meio da deacetilação de resíduos de lisina em AIRE. Uma vez ativada, AIRE reconhece e se liga a marcas epigenéticas que atuam reprimindo a transcrição como a H3K4me0. A topoisomerase II (TOP2a) se junta a AIRE e promove a quebra da fita dupla do DNA, com isso DNA PKs e outras parceiras são ativadas e promovem o afrouxamento da cromatina. O complexo de proteínas *Bromodomain-containing 4* (Brd4) e *Positive transcription elongation factor b* (pTEFb) se ligam a AIRE e são responsáveis por liberar a RNA pol II que está estática na cromatina promovendo a transcrição. Por fim, a *CREB-binding protein* (CBP) é responsável por acetilar os resíduos de lisina de AIRE levando a sua inativação (ABRAMSON et al., 2010; GIRAUD; PERTERSON, 2019).

Aire é um gene pleotrópico, ou seja, possui várias funções que vão além da regulação da expressão de autoantígenos nas mTECs (BANSAL et al, 2017).

Foi demonstrado que a maturação e diferenciação das mTECs é dependente de AIRE, uma vez que essa proteína atua na constituição do microambiente tímico ideal para o desenvolvimento das mTECs (PASSOS et al., 2018). Além disso, durante o processo de seleção negativa AIRE atua na diferenciação das células T regulatórias, que atuam na periferia controlando reações autoimunes (LI et al., 2016).

A adesão entre mTECs e timócitos se faz necessária para que os autoantígenos sejam apresentados e reconhecidos durante a seleção negativa. Demonstrou-se por meio da técnica de RNA de interferência (RNAi) que Aire atua nesse processo, pois após sua perturbação houve um comprometimento nas funções de moléculas que medeiam a adesão célula-célula, como as integrinas e as claudinas (PEZZI et al, 2016; IRLA, 2019).

1.6 APECED/APS-1

Mutações no gene Aire estão relacionadas com diversas patologias, como é o caso da APECED ou APS1. Trata-se de uma síndrome genética com padrão de herança majoritariamente autossômico recessivo, que acomete tanto crianças quanto adultos, porém foi descrita recente uma forma de APECED com padrão de herança autossômico dominante (OFTEDAL et al., 2015). Possui um amplo espectro clínico pois de acordo com a região do gene Aire que sofreu mutação a mesma apresenta determinadas características. De forma geral APECED ocorre devido a alteração no controle da tolerância imune, levando a alterações nas funções de diversos órgãos,

principalmente dos órgãos glandulares (PETERSON et al., 2004; CHENG; SHUM; ANDERSON, 2007; WOLFF; OFTEDAL, 2019).

Já foram identificadas mais de 100 mutações em Aire que estão relacionadas com o desenvolvimento da APECED (WOLFF; OFTEDAL, 2019). Essa é uma síndrome rara, com prevalência estimada de 1:100.000 em muitos países. Algumas regiões do globo apresentam uma prevalência maior, como na Finlândia (1:25.000), na Ilha de Sardenha (1:14.000) e entre judeus de origem Persa que vivem em Israel (1:9.000) (HUSEBYE et al., 2009).

É caracterizada pelo aumento nos níveis de autoanticorpos circulantes associada a infiltração linfocitária, isso culmina na alteração da função dos órgãos afetados. Em casos mais graves leva a falência dos mesmos (BRUSERUD et al., 2016).

Como supracitado, a APECED causa reações autoimunes que afetam vários órgãos. A síndrome possui três sinais cardinais: candidíase mucocutânea crônica, hipoparatiroidismo e a insuficiência adrenal. O diagnóstico de APECED é confirmado pela presença de pelo menos dois desses sinais (HUSEBYE et al., 2009).

Devido à grande quantidade de órgãos afetados, além do sinais cardinais, a APECED pode causar muitas outras alterações no organismo tanto de homens quanto de mulheres, tais como insuficiência testicular, insuficiência ovariana primária, nefrite, queratite, pancreatite exócrina e endócrina que leva ao desenvolvimento de diabetes mellitus tipo 1, hepatite, vitiligo, dentre outras alterações (Figura 10) (FERRE et al., 2016; HUSEBYE; ANDERSON; KÄMPE, 2018).

O diagnóstico muitas vezes é difícil devido ao amplo espectro patológico, em alguns casos só é feito após a morte do paciente. Há também situações em que os pacientes não apresentam dois dos sinais cardinais, e apresentam formas mais brandas de APECED que acabam não sendo diagnosticadas (FERRE et al., 2016).

O modelo eficiente para um diagnóstico eficaz de APECED, inclui inicialmente o teste para pesquisa de autoanticorpos anti-interferon 1, pois em 95% dos casos os pacientes apresentam altos níveis desses autoanticorpos, seguido de testes genéticos como o sequenciamento do gene Aire para identificação exata da região afetada, o que garante a escolha de um tratamento mais adequado (BRUSERUD et al., 2016; FERRE et al., 2016).

O tratamento da APECED consiste em minimizar sinais e sintomas ou curar o paciente de acordo com o quadro clínico. Assim, é comum a reposição hormonal, terapia antifúngica, terapia imunossupressora para hepatites, dentre outras abordagens focadas nas manifestações apresentadas (HUSEBYE; ANDERSON; KÄMPE, 2018).

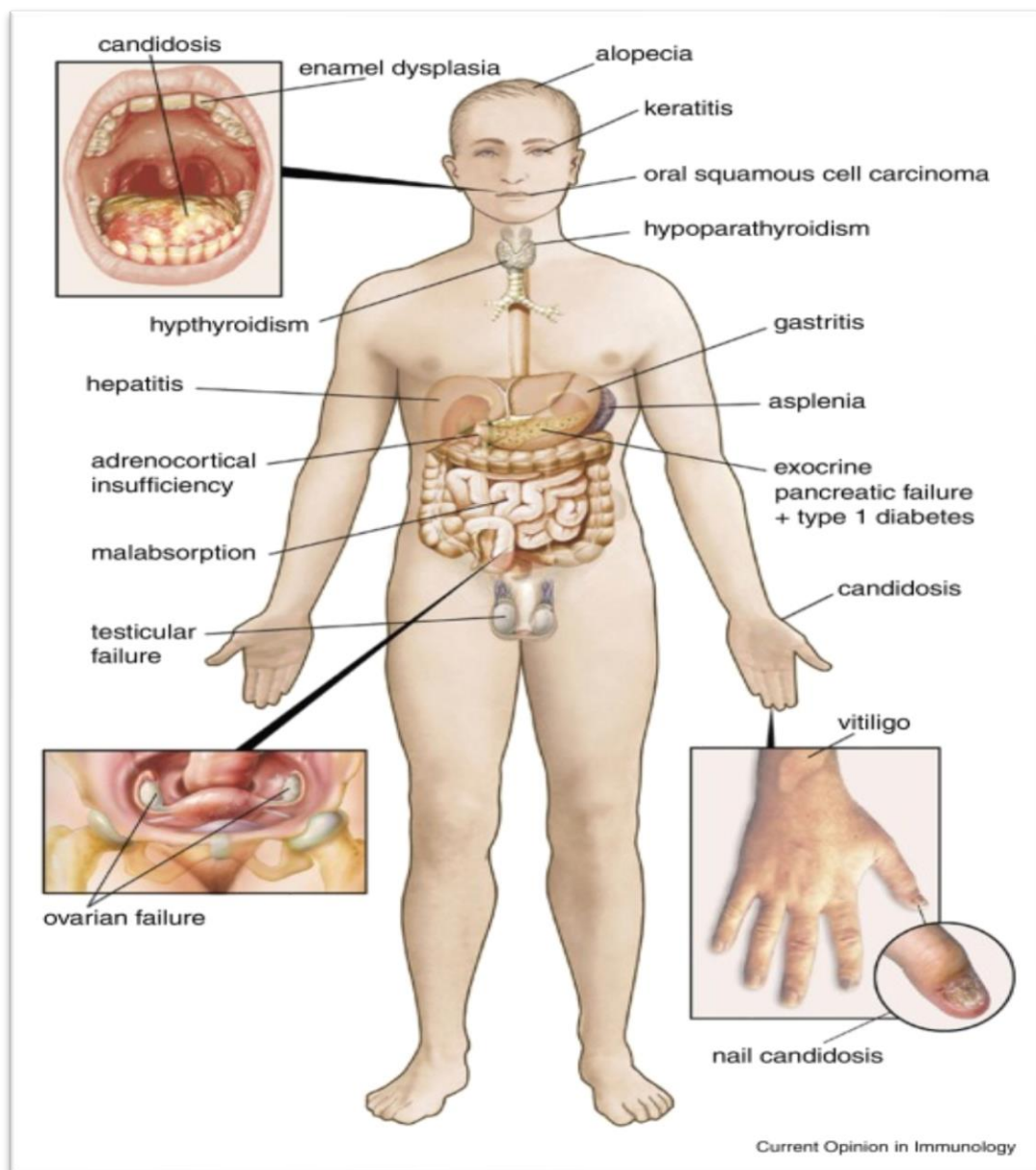


Figura 10. Descrição das principais manifestações apresentados por pacientes com APECED. Possui quadro clínico altamente variável, tendo como os 3 sinais cardinais a candidíase mucocutânea crônica, insuficiência adrenal e hipoparatiroidismo (BRUSERUD et al., 2016).

1.7 Aire e RNAs não-codificantes

Tendo em vista a pleiotropia do gene Aire, destaca-se outra função biológica importante, que é o controle da expressão de RNAs não-codificantes, como os MicroRNAs (miRNAs). Foi demonstrado que o silenciamento de Aire reduz o número de miRNAs transcritos nas mTECs. Essas pequenas moléculas de RNAs não codificantes (~22nt) tem papel fundamental na regulação da expressão gênica, e alterações na expressão de miRNAs podem interferir no processo de expressão e apresentação de TRAs. (MACEDO et al., 2013; PASSOS; MENDES-DA-CRUZ; OLIVEIRA, 2015; OLIVEIRA et al, 2016).

Com base na interação Aire x miRNAs, Passos et al, 2018, sugerem que Aire também pode influenciar na expressão de um outra classe de RNAs não codificantes que vem ganhando notoriedade, os RNAs longos não-codificantes (lncRNAs).

1.8 RNAs longos não-codificantes (lncRNAs): histórico e características

Durante muito tempo, o genoma foi dividido simploriamente em duas partes, uma parte que era formada por genes codificadores de proteínas, que corresponde a aproximadamente 2% de todo o genoma, sendo conhecida como a parte funcional, com basicamente todos os estudos focados nessa pequena fração; os outros 98% não é capaz de codificar proteínas, e acreditava-se que tinha apenas função estrutural, sendo chamado por muitos cientistas de “DNA lixo”, tanto que a definição de gene, era voltada totalmente para as regiões que codificavam proteínas, sendo gene denominado como uma porção do genoma capaz de codificar um produto proteico (BEADLE; TATUM, 1941; Huttenhofer et al., 2005; Djebali et al., 2012).

Todavia, alguns estudos demonstraram que esse “DNA lixo” tinha capacidade de ser transcrito, e a partir disso teve início a especulação de que esses que eram transcritos mas não eram codificantes pudessem ser mais que “DNA lixo”, e que provavelmente possuíam funções regulatórias. Atualmente prefere-se chamar esse tipo de DNA como “DNA sucata” ao invés de “lixo”, pois eventualmente as células poderão utilizar essas prções do genoma. Diversas classes de RNAs não-codificantes começaram a ser descritas, dentre muitas classes, destacam-se os lncRNAs (Quan et al., 2015). Tamanha é a importância dos RNAs não-codificantes que a definição de gene foi alterada, atualmente gene é denominado como uma porção do genoma capaz de ser transcrita e gerar um produto funcional, que não necessariamente seja uma proteína (Darien et al., 2012).

Os lncRNAs são definidos de forma ampla, como uma classe de RNAs que contém mais de 200 nucleotídeos de comprimento e não possuem capacidade de codificar proteínas, a maioria dos lncRNAs são transcritos pela RNA polimerase II, outros pelas RNAs polimerases III e IV; apresentam um padrão de expressão espaço-temporal muito específico, em relação a localização subcelular, eles podem se encontrar tanto no núcleo quanto no citoplasma, assim como mRNA podem ser poliadenilados, possuem a extremidade N terminal adicionada de um nucleotídeo de guanosina metilado (cap 5') e são capazes de realizar *splicing* (AUNE; SPURLOCK, 2016; FERNANDES et al., 2019). No início da década de 90, os primeiros lncRNAs começam a ser descritos, dentre esses destacam o H19 (*imprinted maternally expressed transcript*), envolvido no crescimento celular e o XIST (*X inactive specific transcript*), responsável pela inativação do cromossomo X e compensação de dose em mamíferos (PONTIER; GIRBNAU, 2011; KENIRY et al., 2012).

Sabe-se que os lncRNAs constituem uma classe relativamente nova de RNAs, assim eles estão sendo classificados paulatinamente, onde diversas abordagens são utilizadas para classificá-los, por exemplo temos a classificação fundamentada no comprimento do transcrito, onde são divididos em *small-lncRNA* (250-950 nt), *medium-lncRNA* (950-4.800 nt) e *large-lncRNA* (>4.800 nt); outras classificações são baseadas na localização intracelular, na estrutura secundária, nas funções, dentre outras, afim de se obter a maior quantidade de informações possíveis para caracterizá-los (MA; BAJIC; ZHANG, 2013; SIENA; BARROS-ICHIHARA, 2017).

A classificação mais utilizada como referência, tem como base a posição dos lncRNAs em relação aos genes codificadores de proteínas. Os mais comuns estão representados na figura abaixo, sendo divididos em intergênicos (lincRNAs): são os que estão localizados entre dois genes codificadores de proteínas, os do biotipo sense: lncRNAs que se sobrepõem com um ou vários éxons de um gene e são transcritos a partir da mesma fita; lncRNAs antisense: aqueles localizados na fita oposta de genes codificadores de proteínas e possuem intersecção com um ou mais éxons; os lncRNAs intrônicos: estão localizados dentro de íntrons de genes codificadores de proteínas; e os bidirecionais: lncRNAs que estão localizados na fita oposta a um gene codificador de proteínas, onde a transcrição do lncRNA e do gene codificador adjacente na fita oposta é iniciada em regiões genômicas próximas (Figura 11) (QUAN et al., 2015; SIENA; BARROS-ICHIHARA, 2017; HERMANS-BEIJNSBERGER; VAN BILSEN; SCHROEN, 2018).

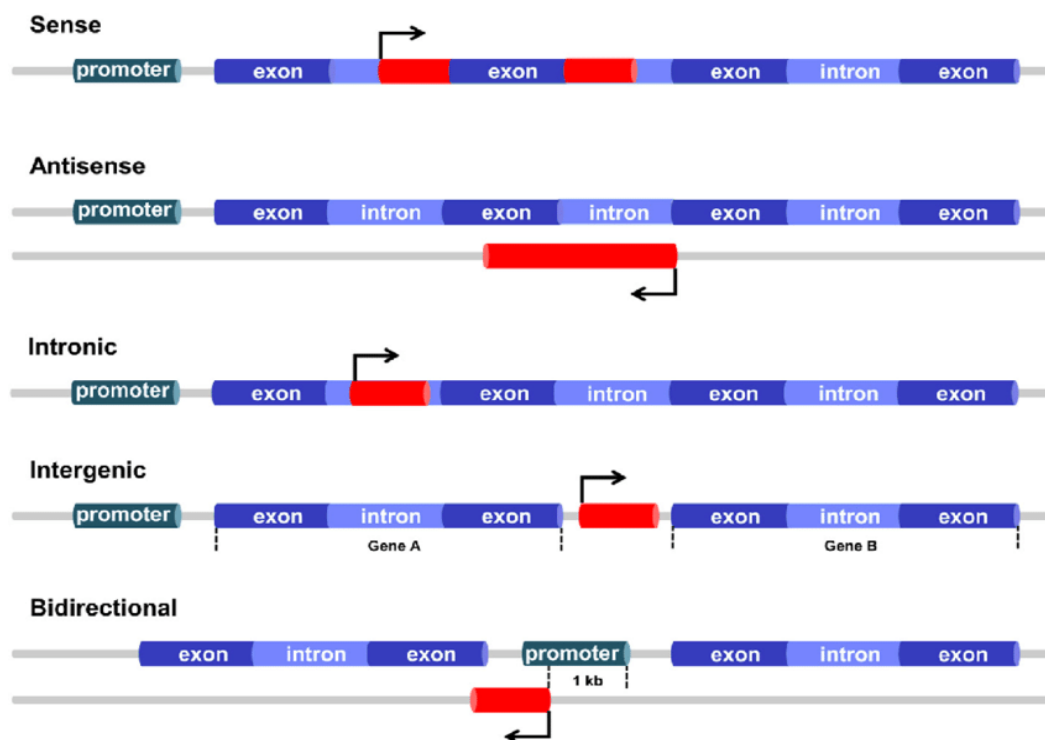


Figura 11. Classificação dos lncRNAs baseados em sua localização genômica.

Sense: lncRNAs que são transcritos a partir do mesmo filamento e na mesma direção que seus genes codificadores de proteínas circundantes; eles podem ser. Antisense: são os lncRNAs transcritos a partir da fita oposta dos genes codificadores de proteínas. Intrônicos: lncRNAs que estão localizados inteiramente em regiões intrônicas de um gene codificador de proteínas. Intergênicos: os lncRNAs intergênicos estão localizados entre dois genes codificadores de proteínas e são transcritos na mesma direção. Bidirecionais: aqueles lncRNAs que estão localizados dentro de 1 kb da região promotora de um gene codificador de proteínas, mas são transcritos a partir da fita oposta (HERMANS-BEIJNSBERGER; VAN BILSEN; SCHROEN, 2018)

Em relação aos mecanismos moleculares de regulação dos lncRNAs, atualmente estão divididos em 4 arquétipos, é válido ressaltar que um único *long non-coding* pode estar relacionado com mais de um mecanismo de regulação (WANG; CHANG, 2012; YANG; WEN; ZHU, 2015; MATHY; CHEN, 2017).

Assim, temos os lncRNAs sinalizadores (*signals*), a presença desses lncRNAs indica que pode ocorrer processos de sinalização celular, e é partir desse sinal que eles conseguem regular a expressão gênica, são considerados sinalizadores rápidos pois não carecem de ser traduzidos em proteínas para exercer sua função. Outro grupo, são os lncRNAs que atuam como chamarizes ou armadilhas (*decoys*), promovem a regulação negativa de um efetor funcional, esses são capazes de se ligar e capturar moléculas, sobretudo proteínas ou outros RNAs, inibindo a função dos

mesmos. O terceiro arquétipo é formado pelos lncRNAs guias (*guides*), esses recrutam, se ligam e direcionam proteínas para genes alvos específicos, modificando o perfil de expressão desses alvos. No último mecanismo temos os lncRNAs que servem como plataforma (*scaffolds*), permitem que outras moléculas se acoplem a eles, para formar complexos funcionais e estabilizar estruturas, onde os long non-coding são o elo entre essas moléculas formadoras de complexos (WANG; CHANG, 2012; YANG; WEN; ZHU, 2015; MATHY; CHEN, 2017).

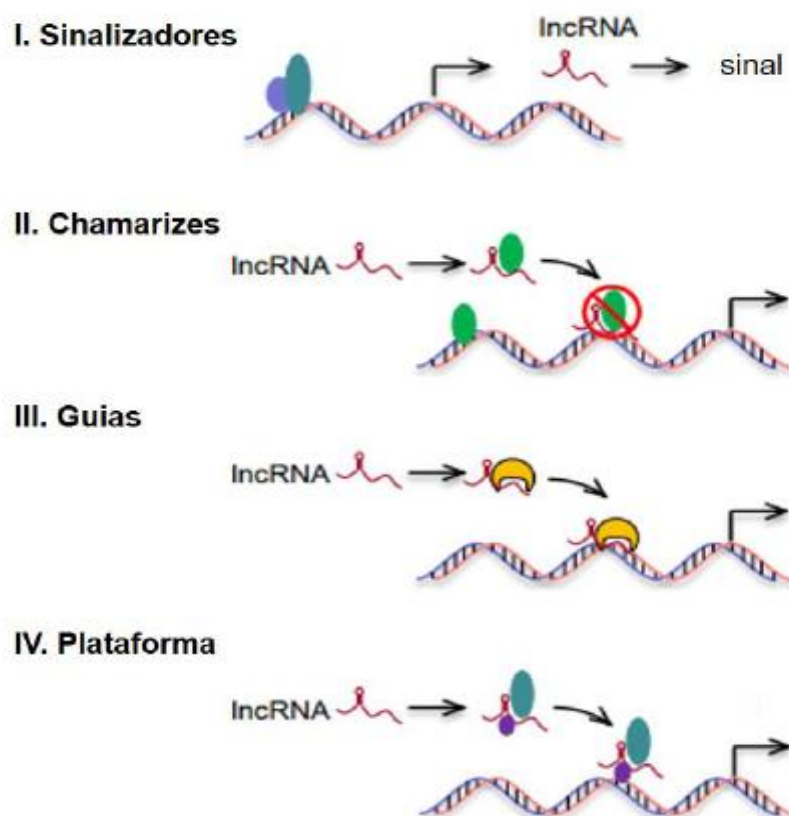


Figura 12. Representação dos diferentes mecanismos de ação dos lncRNAs. I – lncRNAs Sinalizadores (*signals*): aqueles que atuam em vias de sinalização celular. **II – lncRNAs Chamarizes (*decoys*):** os que capturam e afastam fatores de transcrição ou outras proteínas do seu alvo. **III – lncRNAs Guias (*guides*):** os que recrutam e direcionam proteínas para o gene alvo. **IV – lncRNAs Plataforma (*scaffolds*):** aqueles que servem de apoio para múltiplas proteínas e formam complexos de sinalização ou de estabilização. (Adaptado de MATHY; CHEN, 2017).

Os lncRNAs são moléculas altamente versáteis, pois são capazes de interagir a nível pré-transcricional, transcricional e pós-transcricional (Figura 13). A ação dos lncRNAs pode ocorrer em *cis*, quando estão no mesmo *locus* cromossômico que seus genes alvo, também podem atuar em *trans*, ou seja, regulando a expressão de genes localizados em domínios cromossômicos distantes (GUIL; ESTELLER, 2012).

A nível pré-transcricional, os lncRNAs interagem com o DNA no remodelamento da cromatina, por meio do recrutamento e seleção de moléculas que formam complexos remodeladores. No nível transcricional, certos lncRNAs são capazes de interagir com regiões promotoras do DNA, aumentando ou reduzindo os níveis de expressão gênica, outros lncRNAs são derivados de regiões com atividade *enhancer* (acentuadora), esses se assemelham a *enhancers* tradicionais, aumentando a atividade transcricional. Na regulação pós-transcricional, os lncRNAs podem atuar por meio de diversos mecanismos, inibição da tradução, modulação de *splicing* e degradação do mRNA são alguns exemplos, também podem agir facilitando a tradução do mRNA, uma vez que os lncRNAs podem interagir e funcionar como “esponjas” de miRNAs, lncRNAs que possuem sítios de ligação de miRNAs em sua sequência sequestram esses miRNAs e evitam que encontrem seus mRNAs-alvo e os silencie (SALEHI et al., 2017; SIENA; BARROS-ICHIHARA, 2017).

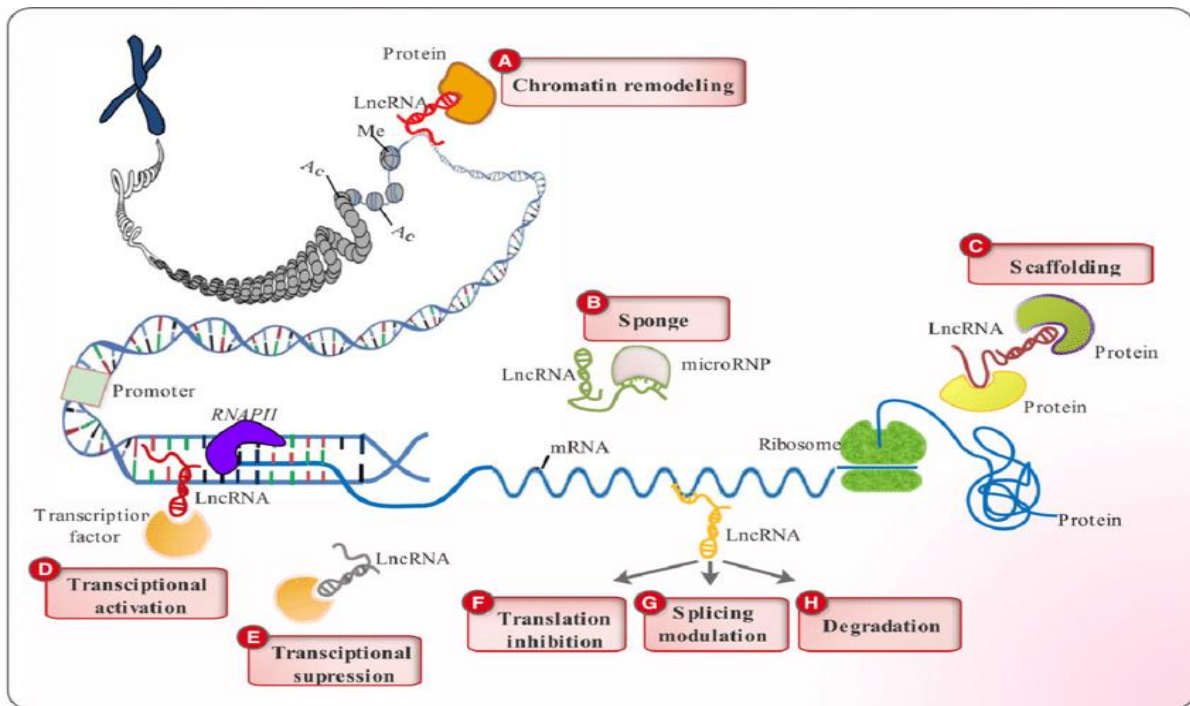


Figura 13. Classificação dos lncRNAs baseado nas funções. (A) Os lncRNAs recrutam componentes proteicos do complexo remodelador da cromatina para alterar os padrões organizacionais mesma. (B) Eles podem atuar como "esponjas" de miRNAs por meio do emparelhamento de base com seus miRNAs complementares e reduzindo seus efeitos. (C) lncRNAs podem desempenhar funções de andaimes (*scaffolds*) fornecendo locais de ancoragem para proteínas que atuam juntas na mesma via biológica. (D) Podem ativar a transcrição de determinados genes, orientando fatores de transcrição para seus promotores. (E) São capazes de suprimir a transcrição sequestrando fatores de transcrição e mantendo-os longe de seus promotores. Podem ainda modular o funcionamento do mRNA por meio do pareamento de bases com eles e (F) inibir sua tradução (G) Os lncRNAs também podem alterar padrões de *splicing* de mRNAs e (H) submeter mRNAs a vias degradativas (SALEHI et al., 2017).

1.9 lncRNAs e o sistema imunológico

A cada dia o envolvimento dos lncRNAs na regulação do desenvolvimento celular e no controle de vias é mais elucidado. Nesse contexto, diversos autores já demonstraram a importância dos lncRNAs na homeostase do sistema imunológico (XU et al, 2018; HADJICHARALAMBOUS; LINDSAY, 2019).

Evidenciou-se o papel dos mesmos na regulação do desenvolvimento de diversas células do sistema imune, tais como neutrófilos, linfócitos B e T, eosinófilos, monócitos, macrófagos, células dendríticas, dentre outras (ALMO et al., 2018; HADJICHARALAMBOUS; LINDSAY, 2019).

O lncRNA *Morrbid* (do inglês, *Myeloid RNA Regulator of Bim-Induced Death*), por exemplo, tem papel fundamental no desenvolvimento e regulação de monócitos, neutrófilos e eosinófilos, tanto em humanos e quanto em camundongos. *Morrbid*

interage com algumas citocinas envolvidas com a sobrevivência dessas células, como, IL-3, IL-5, GM-CSF (TIAN, 2016).

Outro *long non-coding* regulador do desenvolvimento de células imunológicas é o EGO (do inglês *Eosinophil Granule Ontogeny*), que como o próprio nome sugere, atua na ontogenia dos eosinófilos, durante a desenvolvimento dessas células EGO encontra-se hiperexpresso (HU; ALVAREZ-DOMINGUEZ; LODISH, 2012).

Em relação aos linfócitos T, já foi demonstrado que alguns lncRNAs atuam no desenvolvimento dessas células, ressalta-se que de acordo com a fase de maturação e diferenciação das células T, existem lncRNAs específicos. O *long non-coding* NRON é um exemplo, em condições normais, NRON integra um complexo proteico que mantém o fator de transcrição NFAT (do inglês, *Nuclear Factor of Activated T Cells*) fosforilado, assim NFAT permanece no citoplasma do Linfócito T, esse complexo proteico atua como mediador do processo de ativação de células T (ATIANAND; CAFFREY; FITZGERALD, 2017).

As células dendríticas, como citado anteriormente, são células apresentadoras de antígenos, etapa crucial da expressão gênica promíscua, e o lncRNA Inc-DC (ou LOC645638) atua na diferenciação celular de monócitos em células dendríticas. E foi demonstrado que o *knockdown* do lncRNA-DC culminou na perturbação de genes envolvidos na apresentação de antígenos pelas células dendríticas (WANG *et al.*, 2014).

Em relação ao timo, pouco se sabe sobre o papel dos lncRNAs no desenvolvimento e maturação das células tímicas. Um único trabalho publicado sobre a temática até o momento estudou o papel do hormônio sexual feminino 17-beta-estradiol no perfil de expressão de lncRNAs em células epiteliais tímicas medulares (WEI *et al.*, 2018).

Assim, reconhecendo a importância de Aire e do timo para o sistema imunológico, e o papel regulatório dos RNAs não codificantes sobre os mais variados processos biológicos, é importante elucidar se o gene Aire atua na regulação da expressão dos lncRNAs em células mTECs.

HIPÓTESE

2. HIPÓTESE

O gene Aire (*Autoimmune regulator*) regula a expressão de RNAs longos não codificantes (lncRNAs) em células epiteliais tímicas medulas (mTECs).

OBJETIVOS

3. OBJETIVOS

3.1 Objetivo Geral

Determinar o perfil de expressão e caracterizar os RNAs longos não codificantes diferencialmente expressos em células mTECs nocauteadas no gene Aire antes e após o ensaio de adesão com timócitos.

3.2 Objetivos específicos

- Identificar os lncRNAs diferencialmente expressos em células mTEC 3.10E6 (Aire^{-/-}) por meio de RNA-Seq e análise bioinformática dos dados, partindo de amostras de mTECs antes e depois da adesão com timócitos.

- Classificar funcionalmente os lncRNAs diferencialmente expressos em células mTEC 3.10E6 (Aire^{-/-}) antes e depois da adesão com timócitos por meio de ferramentas de predição *in silico*.

- Determinar a localização cromossômica dos lncRNAs diferencialmente expressos nas células mTEC 3.10E6 (Aire^{-/-}) antes e depois da com timócitos.

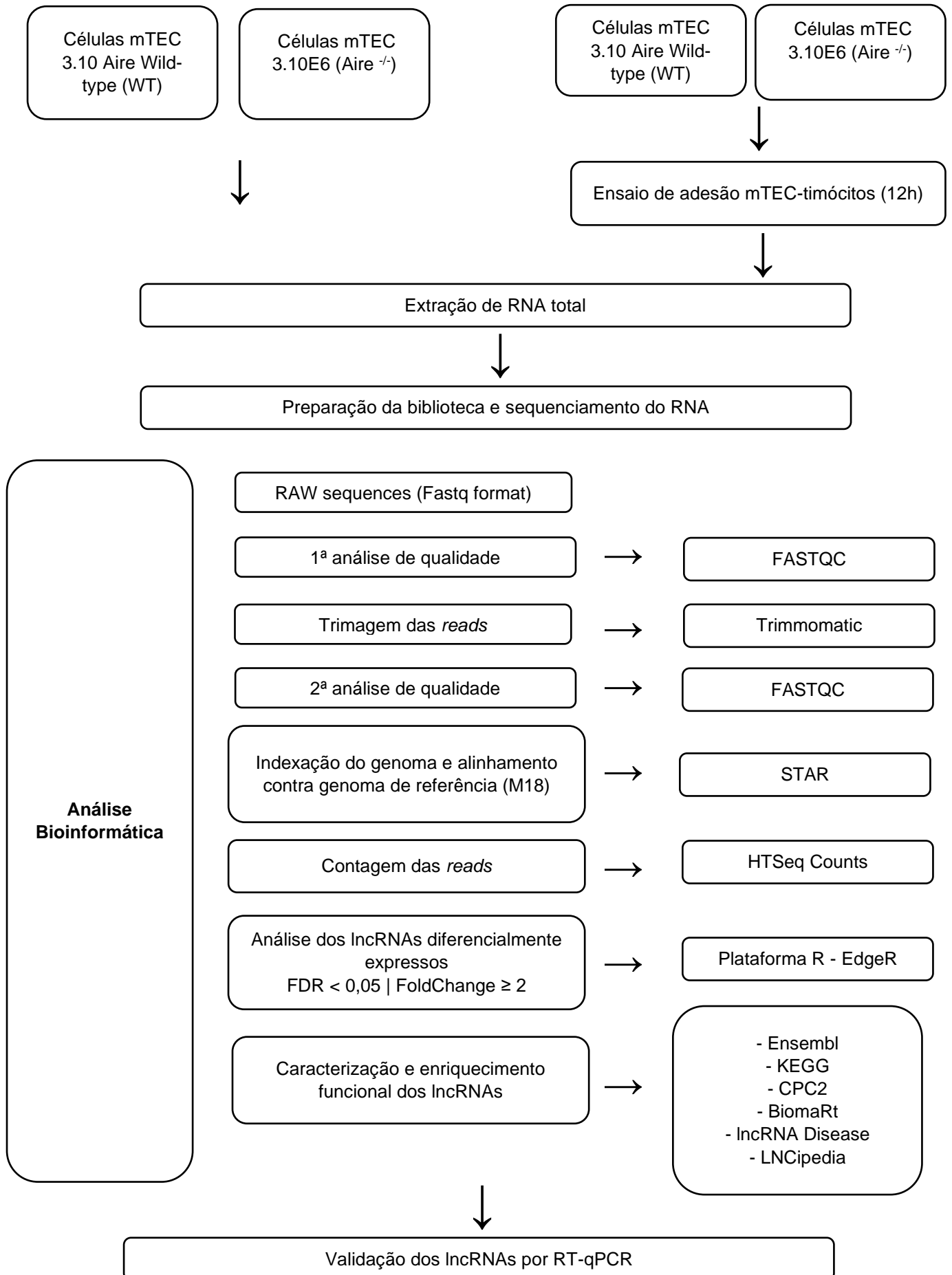
- Classificar os RNAs longos não codificantes de acordo com o seu potencial codificador.

- Identificar no genoma os genes codificadores de proteínas mais próximos dos lncRNAs diferencialmente expressos.

- Validar a expressão dos lncRNAs selecionados por meio de RT-qPCR.

DELINEAMENTO EXPERIMENTAL

4. DELINEAMENTO EXPERIMENTAL



MATERIAL E MÉTODOS

5. MATERIAL E MÉTODOS

5.1 Linhagens celulares

5.1.1 Células mTEC 3.10 Wild type (WT)

Essas células foram estabelecidas a partir do timo de camundongos recém-nascidos da linhagem C57BL/6. A linhagem primária de células tímicas epiteliais medulares (mTEC 3.10 wild-type) que utilizamos nesse projeto, foi cedida pelo Dr. Wilson Savino (Laboratório de Pesquisas sobre o Timo, FioCruz, Rio de Janeiro).

As células mTEC 3.10 foram fenotipadas e classificadas como Epcam⁺ Ly51⁻ UEA1⁺ CD80^{low} MHC-II^{low} e que expressam o gene Autoimmune regulator (*Aire*) (Pezzi et al., 2016). As células mTEC 3.10 são aderentes ao plástico poliestireno das garrafas de cultura e foram cultivadas em garrafas T25 (25 cm²) ou T75 (75 cm²), contendo, respectivamente, 5 e 10 mL de meio RPMI 1640 (Gibco – El Paso, TX, USA) suplementado com 10% de soro fetal bovino (SFB) (Cultilab – Campinas, SP, Brasil), e mantidas a 37°C em estufa com atmosfera de 5% de CO₂.

5.1.2 Células mTEC 3.10E6 (*Aire*^{-/-})

As células mutantes mTEC 3.10E6 (*Aire*^{-/-}) foram isoladas durante o trabalho de mestrado de Cesar Augusto Speck-Hernandez (orientado do Prof. Dr. Geraldo A. Passos). A mutação dessa linhagem foi introduzida por meio do sistema CRISPR-Cas9 tendo como alvo o exon 3 de *Aire*, e ocorreu especificamente na região que codifica o domínio NLS (Sinalizador de Localização Nuclear) da proteína AIRE.

Na caracterização das células mutantes mTEC 3.10E6 (*Aire*^{-/-}), foram identificadas mutações do tipo *indel* afetando ambos os alelos de *Aire* (nocaute heterozigoto composto). No alelo 1 do gene *Aire*, há dois tipos de mutações: substituição T > G (na posição 351) seguida por uma deleção de nove pares de base (GCTGGTCCC, posições 352-360). Já no alelo 2, houve uma única deleção de G na posição 352.

A análise *in silico* da tradução da proteína AIRE no clone mTEC 3.10E6 (*Aire*^{-/-}) e o alinhamento das sequências polipeptídicas por meio do programa Clustal Omega mostraram que a substituição T > G no alelo 1 resultou numa mutação silenciosa L118L, enquanto que a deleção de nove pares de base provocou um erro de leitura do tipo *frameshift* causando a deleção de três resíduos de aminoácidos (A119_P121 del).

Na caracterização dos efeitos funcionais causados por variações na sequência de aminoácidos, utilizou-se o programa Provean, observou-se que a deleção dos 9 nucleotídeos no alelo 1 foi deletéria para o domínio NLS da proteína AIRE, já para o alelo 2, a deleção de um G resultou num códon de parada TGA na posição 352 e, conseqüentemente, uma proteína AIRE truncada de 158 resíduos de aminoácidos. De acordo com essas análises, o clone mTEC 3.10E6 (Aire^{-/-}) é considerado um Aire nocaute heterozigoto composto que expressa uma proteína AIRE truncada, que não é carregada para o núcleo da célula (SPECK-HERNANDEZ et al., 2018).

As condições de cultivo das células mTEC 3.10E6 (Aire^{-/-}) foram as mesmas das células mTEC 3.10 WT.

5.2 Obtenção dos timócitos por timectomia

Para obtenção dos timócitos camundongos C57BL/6 com cinco a seis semanas de idade foram sacrificados em câmara de CO₂ e os timos foram removidos por cirurgia torácica. Os órgãos foram transferidos para uma placa de Petri com meio de cultura RPMI 1640 onde foram desintegrados resultando numa suspensão de células tímicas. Os timócitos foram separados pela passagem da suspensão celular através de uma membrana de nylon (Sefar Inc. Depew, NY, USA) com porosidade de 10 µm.

A suspensão de timócitos foi centrifugada por 5 minutos a 1000 x g e as células foram lavadas duas vezes com PBS, e depois ressuspensas em meio RPMI 1640. Os timócitos obtidos foram utilizados no ensaio de adesão celular com as células mTEC 3.10 WT e as células mTEC 3.10E6 (Aire^{-/-}). Este procedimento foi realizado durante do mestrado do aluno César Speck e seu projeto foi previamente aprovado pelo Comitê de Ética CEUA da Faculdade de Medicina de Ribeirão Preto, USP, recebendo o número 006/2016-1 (anexo 1).

5.3 Ensaio de adesão mTECs-timócitos

O protocolo do ensaio de adesão mTECs-timócitos é baseado em trabalhos prévios do Laboratório de Pesquisas sobre o Timo (Fiocruz, Rio de Janeiro) (LINHARES-LACERDA et al., 2010) com algumas modificações introduzidas por nosso grupo (PEZZI et al., 2016).

Os timócitos recém coletados de timos de camundongos adultos-jovens da linhagem C57BL/6 foram adicionados às culturas de mTEC 3.10WT e mTEC 3.10E6 (Aire^{-/-}) na proporção de 50:1 (timócitos:mTEC). As co-culturas permaneceram em

estufa a 37° C em atmosfera com 5 % de CO₂ durante 12 horas. Depois, os timócitos não aderidos e que permanecem em suspensão foram removidos com cuidado por meio de aspiração do meio de cultura e subsequente lavagem cuidadosa com PBS a 37° C. Os timócitos que estavam aderidos às mTECs foram agora removidos por meio de lavagem vigorosa com PBS a 4°C para contagem.

Posteriormente foi feita a remoção das células mTECs aderidas às garrafas por meio de tratamento com tripsina/EDTA convencional, utilizando 5 ml de solução de tripsina 1x (2.5g de tripsina suína por litro em HBSS) + 0.5ml de EDTA pH 8.0 por 5 minutos, seguido de inativação da tripsina com meio de cultura RPMI + SBF e colhidas por centrifugação a 1000 x g durante 5 minutos. Tanto os timócitos que estavam aderidos como as mTECs foram contados num contador de células automático da marca Cellometer Auto T4 Cell Viability Counter (Nexcelom Bioscience, Lawrence, MA, USA).

O índice de adesão (IA) foi calculado com base na fórmula abaixo:

$$IA = \text{Número de timócitos aderidos} / \text{Número de células mTEC}$$

5.4 Extração do RNA total

A extração de RNA total das células mTEC 3.10WT e mTEC 3.10E6 (Aire ^{-/-}) antes e após o ensaio de adesão mTCEs-timócitos foi realizada utilizando o kit *mirVana miRNA Isolation Kit Ambion* (Grand Island, NY, USA). As células foram destacadas das garrafas de cultura por meio da adição de 3 mL de solução de tripsina/EDTA por 5 minutos. A tripsina foi inativada com meio de cultura RPMI contendo 10% de SFB, e as células colhidas por centrifugação a 1000 x g durante 5 minutos. Os botões de células resultantes foram lavados duas vezes com PBS estéril e adicionou-se 1ml de solução de lise. Agitou-se vigorosamente (1 minuto no vortex) para homogeneizar e lisar completamente as células. Adicionou-se 100µl de *miRNA Homogenate Additive* e misturou-se bem (vortex por 20 segundos).

As misturas provenientes mTEC 3.10WT e mTEC 3.10E6 (Aire ^{-/-}) foram deixadas em banho de gelo por 10 minutos. Em seguida foi adicionado 1mL de fenol ácido-clorofórmio, e as misturas foram agitadas vigorosamente em vórtex por mais 1 minuto. Os lisados celulares foram centrifugados a 10.000 x g por 5 minutos em temperatura ambiente. A fase aquosa resultante de cada amostra foi cuidadosamente transferida para novos tubos e então foi adicionado etanol 100% (pelo menos 1,5 vezes do volume da fase aquosa recuperado após a centrifugação) à temperatura

ambiente. Para cada amostra foi montada uma coluna com tubo coletor, adicionou-se 700µL da mistura (fase aquosa + etanol) nas colunas e elas foram centrifugadas por 15 segundos a 10.000 x g. O líquido foi descartado e o procedimento repetido até que toda a mistura fosse filtrada.

Na sequência, em cada coluna foram adicionados 700µL de *miRNA Wash Solution 1*, posteriormente elas foram centrifugadas por 10 segundos a 10.000 x g. O líquido foi descartado, e então 500µL de *miRNA Wash Solution 2/3* foi aplicado nas colunas e em seguida elas foram centrifugadas por 10 segundos a 10.000 x g. O último passo foi repetido e, após descartar o líquido, as amostras foram centrifugadas por 1 minuto para remover resíduos líquidos do filtro.

Por fim, as colunas foram transferidas para novos tubos coletores, e a elas foram adicionados 100µL de água MilliQ/DEPC pré-aquecida a 95°C e os tubos foram centrifugados por 30 segundos para a recuperação do RNA total.

5.5 Quantificação e avaliação do grau de pureza das amostras de RNA total

As amostras de RNA total foram quantificadas por espectrofotometria UV no aparelho NanoDrop ND-1000 (NanoDrop Products, Wilmington, DE, USA). Neste trabalho considerou-se que uma unidade (1U) de absorbância A260 corresponde a 40µg de RNA/mL.

A pureza das preparações também foi avaliada por meio de cálculo das razões absorbância em diferentes comprimentos de onda de UV. E só foram utilizadas amostras livres de fenol que apresentaram razão A260/A230 entre 1.8 a 2.0 e livres de contaminação com proteínas razão A260/A280 entre 1.8 e 2.0.

5.6 Análise da integridade das preparações de RNA total

A análise da integridade do RNA total extraído das células mTEC 3.10WT e mTEC 3.10E6 (Aire^{-/-}) foi realizada por eletroforese microfluídica através de RNA 6000 Nano chips no aparelho Agilent 2100 Bioanalyzer® Agilent Technologies (Santa Clara, CA, USA).

Inicialmente todos os reagentes, que anteriormente estavam a 4°C, foram mantidos por 30 minutos a temperatura ambiente. Decorrido o tempo, iniciou-se o preparo do gel pipetando 550µl do RNA 6000 Nano gel em uma coluna com filtro e centrifugou-se a 1.500 x g durante 10 minutos em temperatura ambiente. Uma alíquota de 65µl foi colocada em um tubo de 0,5 ml livre de nucleases e em seguida, foi

adicionado 1µl de RNA 6000 Nano Dye. A amostra foi submetida a agitação em vórtex por 10 segundos, seguida de centrifugação a 13.000 x g por 10 minutos em temperatura ambiente.

Posteriormente, iniciou-se o preparo do RNA 6000 Nano chip que foi colocado na *priming station* com os ajustes corretos para a leitura de RNAs. Primeiramente foram pipetados 9µl da mistura gel/dye na região G indicada no chip e com o auxílio de uma seringa acoplada à *priming station* o gel foi distribuído por todo o chip. Logo depois, outros 9µl da mistura foram pipetados nos demais pontos indicados com a letra G. Foram pipetados 5µl do RNA 6000 Nano Marker em cada uma das amostras, bem como na posição do marcador e 1µl do marcador na posição indicada. Por último, pipetou-se µl de cada amostra de RNA nos respectivos poços marcados de 1 a 12 e com a ajuda de um vórtex IKA MS 3 (Manca, Hong Kong, CHN), agitou-se o chip horizontalmente a 2.200 rpm por 1 minuto e em seguida, colocou-se o chip no bioanalisador para proceder-se a corrida eletroforética.

O resultado das corridas (eletroforograma e densitometria dos géis) foi obtido pelo Agilent 2100 Expert Software após 30 minutos de corrida. Somente as amostras que apresentaram picos correspondentes aos padrões das espécies de RNAr 28S, 18S, 5S e RNAt 4S e com RNA Integrity Number (RIN) ≥ 9.0 foram utilizadas nos experimentos seguintes.

5.7 Construção da biblioteca e sequenciamento do RNA

Foram enviadas amostras de RNA total provenientes de replicatas biológicas (n=2) de células mTEC 3.10WT e mTEC 3.10E6 (Aire^{-/-}) tanto antes quanto após o ensaio de adesão com timócitos ao Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD), parte da Universidade Estadual de Campinas (UNICAMP) (Lactad, Unicamp, Campinas-SP) para sequenciamento utilizando a plataforma HiSeq 2500 (Illumina).

Para cada amostra foram enviados 2µg de RNA total, no LaCTAD a integridade das mesmas foi verificada novamente pelo sistema Bioanalyser 2100 (Agilent) Technologies (Santa Clara, CA, USA), sendo utilizadas apenas amostras com RIN maior ou igual a 9,0. O RNA ribossomal (RNAr) foi então depletado (riblodepleção) com uso de oligos alvo-específicos biotinilados combinados com esferas de remoção de rRNA citoplasmático Ribo-zero (Illumina Inc., San Diego, CA, USA), e o preparo

das bibliotecas foi realizado com o *TruSeq Stranded Total RNA Library Prep Kit* (Illumina Inc., San Diego, CA, USA), seguindo as instruções do fabricante.

As bibliotecas foram sequenciadas utilizando a plataforma HiSeq 2500 (Illumina) (Illumina Inc., San Diego, CA, USA) seguindo as instruções do fabricante, onde obteve-se um sequenciamento do tipo *paired-end stranded specific* com fragmentos de 100 pb (2 x 100 pb) e profundidade de 50 milhões de *reads*.

5.8 Análise bioinformática dos lncRNAs

5.8.1 1ª análise de qualidade

Após o sequenciamento os arquivos referentes as amostras sequenciadas foram disponibilizadas em formato Fastq pelo LaCTAD, essas foram submetidas a primeira análise de qualidade com objetivo de avaliar a qualidade do preparo da biblioteca e do sequenciamento. O programa utilizado para análise de qualidade do material sequenciado foi o FASTQC (ANDREWS, 2010), muitos dos sequenciadores geram um relatório de qualidade, porém este geralmente se limita em identificar problemas que foram gerados pelo próprio sequenciador. O programa FastQC por sua vez apresenta um relatório de controle de qualidade que pode detectar problemas advindos do sequenciador ou na biblioteca.

Foram levados em consideração o tamanho das sequências, a qualidade de cada base e da *read*, que são dadas pelos valores de *Phred*, que considera bons resultados de qualidade quando as bases estão acima de 30 na escala *Phred*. A cor de fundo do gráfico divide o eixo das ordenadas em: chamadas de muito boa qualidade (verde), chamadas de qualidade razoável (laranja), e chamadas de má qualidade (vermelho).

5.8.2 Trimagem

Após a primeira análise de qualidade as amostras passaram pelo processo de trimagem, por meio da ferramenta de linhas de comando Trimmomatic versão 0.36 (BOLGER et al., 2014). A trimagem objetiva remover os adaptadores que foram utilizados durante o sequenciamento bem como *reads* de baixa qualidade e *reads* muito curtas (< 36pb), deixando o material sequenciado “limpo”, pois se não forem removidas tais “sujeiras” podem causar erros na análise dos dados e na mensuração dos níveis de expressão.

5.8.3 2ª análise de qualidade

As amostras trimadas passaram novamente pela análise de qualidade pelo FASTQC, para garantir que o processo de trimagem removeu apenas os adaptadores e que as *reads* a serem montadas contra o genoma de referência estão íntegras e com boa qualidade.

5.8.4 Indexação do genoma de referência

Utilizou-se a ferramenta de linhas de comando STAR Spliced Aligner) 2.5 (<https://github.com/alexdobin/STAR>) (DOBIN et al., 2013) para a indexação do genoma de referência da espécie *Mus musculus* versão M18 (Ano 2018) (GRCm38.p6), obtida no GENCODE (https://www.genencodegenes.org/mouse/release_M18.html). A indexação é realizada antes da montagem das *reads* sequenciadas contra o genoma de referência, essa etapa permite que o genoma seja separado em grupos com características parecidas, formando um índice. Com isso o processo de mapeamento é acelerado, pois o índice de uma grande sequência de DNA permite que as *reads* encontrem de forma mais rápida a sequência do genoma que elas devem se alinhar.

5.8.5 Montagem do transcriptoma

Para a montagem do transcriptoma das amostras sequenciadas, assim como na indexação utilizou-se a ferramenta STAR Spliced Aligner 2.5 (<https://github.com/alexdobin/STAR>) (DOBIN ET AL., 2013) que permitiu a montagem do transcriptoma a partir das *reads* trimadas contra o genoma de referência da espécie *Mus musculus* versão M18 (Ano de 2018) (GRCm38.p6), obtida no GENCODE (https://www.genencodegenes.org/mouse/release_M18.html) que foi indexado anteriormente.

5.8.6 Contagem das reads mapeadas

Após a montagem do transcriptoma, utilizou-se a ferramenta HTSeq counts (ANDERS; PYL; HUBER, 2014), para contagem das *reads* que se alinharam ao genoma de referência para cada gene, aquelas que se alinharam em várias posições ou que se sobrepuseram a mais de um gene foram desconsideradas. É a partir desse

conjunto de dados que são mensurados os níveis de expressão gênica. São considerados satisfatórios porcentagens de *reads* alinhadas > 60% (considerando amostras de boa qualidade).

5.8.7 Análise dos lncRNAs diferencialmente expressos

O processamento e análise dos lncRNAs diferencialmente expressos foi realizada por meio da plataforma R (versão 3.5) (<https://www.r-project.org/>), para isso utilizou-se o pacote EdgeR, que está disponível na plataforma de pacotes de bioinformática Bioconductor, as amostras foram normalizadas e na determinação dos lncRNAs diferencialmente expressos considerou-se os seguintes parâmetros: FoldChange (FC) ≥ 2 e *False Discovery Rate* (FDR ou p ajustado) $\leq 0,05$.

O EdgeR fornece de um ambiente estatístico-matemático para se calcular e construir uma visualização gráfica dos dados, sendo recomendando para análise de dados com no máximo 12 amostras. O arquivo de saída com as informações dos transcritos diferencialmente expressos contém as estatísticas, com valores de expressão (FC em escala log2), os valores de P, tanto brutos como corrigidos para múltiplos testes por FDR, além de informar o nome, a localização cromossômica e a posição no genômica dos genes diferencialmente expressos (ROBINSON; MCCARTHY; SMYTH, 2009). Foi feita então uma intersecção (*merge*) entre o arquivo de saída do EdgeR contendo todos os transcritos diferencialmente expressos com uma tabela específica para lncRNAs validados e preditos proveniente do GENCODE, que resultou numa tabela contendo apenas informações dos lncRNAs diferencialmente expressos.

5.9 Caracterização e enriquecimento funcional dos lncRNAs

A caracterização e o enriquecimento funcional dos lncRNAs foi realizada com base em estudos acerca da função dessas moléculas na literatura e utilizando bases de dados que permitem o estudo dos lncRNAs no transcriptoma murino, como por exemplo: 1) Ensembl (<https://www.ensembl.org/index.html>), GeneOntology (<http://geneontology.org/>), GENCODE (<https://www.gencodegenes.org/>), NONCODE (<http://www.noncode.org/>), RNACentral (<https://rnacentral.org/>), lncRNADisease (<http://www.cuilab.cn/lncrnadisease>): esses bancos de dados permitem identificar o nome, tipo e os processos biológicos em que os lncRNAs estão envolvidos. 2) CPC2: utilizado para calcular o potencial codificador dos lncRNAs. 3) biomaRt: permite

identificar o gene codificador mais próximo de cada lncRNA diferencialmente expresso.

5.9.1 Identificação e posicionamento dos genes codificadores mais próximos dos lncRNAs

A plataforma R foi utilizada para identificar o nome e a posição ao longo do genoma murino dos genes codificadores mais próximos aos lncRNAs diferencialmente expressos. Para isso utilizou-se dentro da plataforma R o pacote biomaRt (<https://www.ensembl.org/biomart/martview/>) (biblioteca baixada pelo Bioconductor). O pacote biomaRt permite acesso aos principais bancos de dados, inclusive Ensembl, onde a busca foi restrita aos genes codificadores de proteínas mais próximos dos lncRNAs diferencialmente expressos (DURINCK et al., 2009).

5.9.2 Análise do potencial codificador dos RNA longos não codificantes

A análise do potencial codificador dos lncRNAs diferencialmente expressos foi realizada através do software CPC2 (<http://cpc2.cbi.pku.edu.cn/>) (KANG et al., 2017). O CPC2 permite avaliar o potencial codificador de um determinado lncRNA, ou seja, calcula as chances de um determinado lncRNA codificar ou não peptídeos. Utiliza um modelo de regressão logística baseado em quatro características: 1) O tamanho da *Open Reading Frame* (ORF, ou matriz de leitura aberta), uma vez que ORFs longas dificilmente são observadas em sequências não codificadoras; 2) a cobertura da ORF, isto é, a razão entre os comprimentos da ORF e do transcrito; 3) o Fickett TESTCODE score, que reflete o efeito combinatório possível da composição de nucleotídeos e a tendência do uso de determinado códon; e por fim o 4) hexâmero score, que determina o grau relativo de viés no uso de um hexâmero em determinada sequência. Os valores desse score variam de 0 a 1. Onde valores < 0,5 indicam que os lncRNAs não são codificadores, e valores > 0,5 indicam que os lncRNAs possuem potencial para codificar peptídeos (KANG et al., 2017).

5.9.3 Enriquecimento funcional dos lncRNAs

Para analisar as interações e os processos biológicos que os lncRNAs diferencialmente estão envolvidos, foi realizada a predição *in silico* da função dos mesmos, a análise foi feita na plataforma R, usando como referência a coleção de

bancos de dados Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.genome.jp/kegg/>).

5.10 Validação dos lcnRNAs por RT-qPCR

5.10.1 Transcrição reversa para síntese do cDNA

As reações de transcrição reversa para síntese de cDNA a partir do RNA das células mTEC 3.10WT e mTEC 3.10E6 (*Aire*^{-/-}) antes e após adesão com timócitos foram realizadas com *High Capacity cDNA Rverse Transcription Kit* da Applied Biosystems™ (Foster City, CA, EUA) seguindo as instruções do fabricante. Para isso, foram utilizados 500 ng de RNA total de cada amostra.

5.10.2 Confirmação da transcrição reversa por PCR convencional

Um total de 200 ng de DNA total foi utilizado como *input* para uma reação de PCR convencional do mRNA (cDNA) de *Gapdh* para confirmar a reação de transcrição reversa. Os *primers* específicos *Foward*: 5' GGGTGTGAACACGAGAAAT 3' e *Reverse*: 5' CCTTCCACAATGCCAAAGTT3' foram desenhados utilizando o *Primer3 Web Tool* (<http://frodo.wi.mit.edu/primer3/>). Foi feito um mix contendo *primers* a uma concentração final de 0.2 μM, dNTPs (0.2 μM), MgCl₂ (1,25 mM), Buffer (1x), e 1.25U da enzima *Gotaq G2 Flexi DNA polimerase* Promega (Wisconsin, USA) foi feita para um volume final de 23 μL, e foi adicionado ao mix 2 μL de cDNA totalizando de 25 μL. As amostras foram amplificadas usando um termociclador *GeneAmp PCR System 9700* (Applied Biosystem, Foster City, CA, USA) com o seguinte programa: Ativação da enzima a 95°C por 2 minutos e 30 ciclos de denaturação a 95°C por 20 segundos, pareamento dos *primers* a 60°C por 30 segundos e alongação da fita de DNA a 72°C por 30 segundos, um último passo de alongação foi feito a 72°C por 5 minutos.

5.10.3 Eletroforese em gel de agarose

Aos produtos da PCR foram adicionados 2,5 μL de *Gel Loading Buffer II* da Ambion, e mais 2,5 μL de solução diluída de brometo de etídeo. A eletroforese das amostras ocorreu em gel de agarose 2% (10 x 7 cm) por 1 hora a 100 V. O tampão da cuba foi o 1X TAE (pH 8.6). O gel foi então visualizado sob luz ultravioleta de transiluminador e fotografado com auxílio do aparelho *ImageQuant™ LAS 500 Ge Life*

Sciences (Piscataway, NJ, USA) para posterior análise das bandas e confirmação da transcrição reversa.

5.10.4 PCR quantitativa em tempo real (RT-qPCR)

Foram selecionados aleatoriamente 3 lncRNAs, 1 de cada condição, o lncRNA Peg13 estava diferencialmente expresso apenas antes da adesão (hiporregulado), o lncRNA Dgkeos diferencialmente expresso somente após a adesão com timócitos (hiperregulado) e o lncRNA Fendrr com expressão diferencial antes e após adesão (hiporregulado em ambas ocasiões) para validação dos níveis de expressão dos mesmos nas células mTEC 3.10WT e mTEC 3.10E6 (Aire^{-/-}) antes e após o ensaio de adesão mTECs-timócitos, por meio da técnica de RT-qPCR (*Reverse Transcription Quantitative Real-Time Polymerase Chain Reaction*).

As reações de RT-qPCR foram realizadas em triplicatas utilizando o reagente SYBR® Green PCR Master Mix Applied Biosystems (Life Technologies Corporation, Carlsbad, CA, USA), seguindo as instruções do fabricante, em placas de 48 poços, com ciclagem padrão de aproximadamente duas horas no aparelho StepOne® Real-Time PCR System Applied Biosystems (Life Technologies Corporation, Carlsbad, CA, USA) e software de leitura StepOne® Real-Time PCR System Software (versão 2.1).

A normalização dos níveis de expressão dos lncRNAs selecionados foi feita através da expressão do gene constitutivo Hprt, o qual é normalmente utilizado com referência.

Os pares de *primers* para os lncRNAs Fendrr, Dgkeos e Peg13 e do constitutivo Hprt (Tabela 1) foram desenhados utilizando o *Primer3 Web Tool* (<http://frodo.wi.mit.edu/primer3/>) e sintetizados pela empresa Integrated DNA Technologies – IDT (Coralville, Iowa, EUA).

Tabela 1. *Primers* utilizados nas reações de RT-qPCR e suas respectivas sequências

Nome do Gene	ID do Gene	Sequência
Dgkeos (lncRNA)	ENSMUSG00000063109	F- 5' TCCAGATCATTTGCCACCT 3' R- 5' GCTGACAAAACCTTCTGAGACCA 3'
Fendrr (lncRNA)	ENSMUSG00000097336	F- 5' CTCTCCAGTTCCCACCACC 3' R- 5' TGGTCTGGTTCTGGTCACTC 3'
Peg13 (lncRNA)	ENSMUSG00000106847	F- 5' TGTGACCACGAACCGAAGAG 3' R- 5' TCGTCTACATAGCACCAGCG 3'
Hprt (RNAm controle)	ENSMUSG00000025630	F- 5' GCCCAAATGGTTAAGGTT 3' R- 5' CAAGGGCATATCCAACAACA 3'

A quantificação relativa da expressão gênica foi determinada utilizando o método $2^{-\Delta\Delta CT}$. Para análise estatística dos dados, foram utilizados os testes *t-student* e *One-way ANOVA* de acordo com a necessidade por meio do software estatístico *GraphPad Prism*.

RESULTADOS

6. RESULTADOS

A parte inicial (cultivo celular, ensaio de adesão mTECs-timócitos, extração e sequenciamento do RNA) desse trabalho foi realizada por César Speck-Hernández, durante o seu mestrado no Laboratório de Imunogenética Molecular e o resultado foi publicado em 2018 na revista *Frontiers in Immunology* (SPECK-HERNANDEZ et al., 2018).

6.1 Análise da qualidade

A análise de qualidade dos pares de base foi realizada com uso do programa FASTQC, verificou-se que as *reads* apresentaram altos valores de qualidade. Tanto na primeira análise (figura 1A), quanto na segunda análise, após a trimagem (figura 1B), sendo que após a trimagem obteve-se uma melhora na qualidade das amostras. Foram obtidos valores de Phred sempre acima de 30 para todas as bases das *reads*, isso significa que a probabilidade da chamada de bases estar correta é de 99,9%.

Os resultados do controle de qualidade também demonstraram que não houve uma queda na qualidade geralmente observada no final das sequências obtidas por meio de metodologias de sequenciamento Illumina.

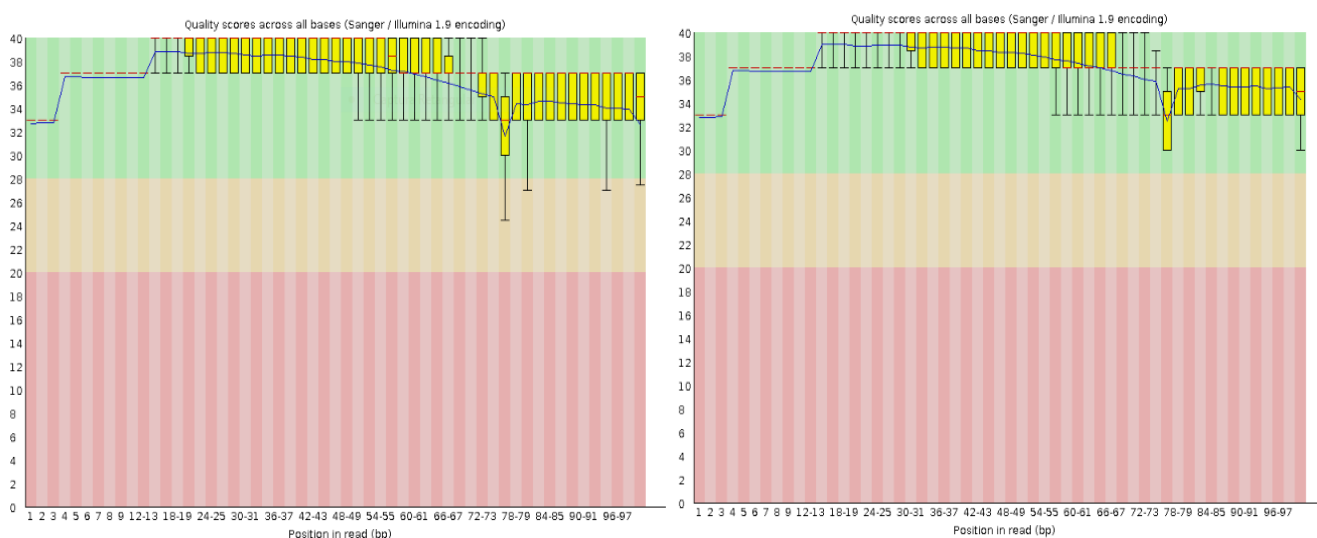


Figura 14. Exemplo de qualidade das sequências, obtidas na plataforma Illumina.

A: mTEC 3.10 WT sem adesão amostra 2 sem trimagem. **B:** mTEC 3.10 WT sem a adesão amostra 2 após a trimagem. Verde: Boa qualidade (Phred > 30); laranja: qualidade razoável (20 > Phred < 30); vermelho: qualidade ruim (Phred < 20). No eixo das abscissas estão as posições das bases das *reads* e no eixo das ordenadas as notas de qualidade Phred.

6.2 Trimagem

O processo de trimagem foi realizado para remoção dos adaptadores utilizados no sequenciamento, remoção das *reads* muito curtas (< 36pb) e das *reads* de baixa qualidade. Ao término do processo de trimagem das 8 amostras sequenciadas, obteve-se uma média percentual de 89,75% de *reads* “sobreviventes”, ou seja, *reads* sem adaptadores, maiores que 36pb e de boa qualidade (Tabela 2).

Tabela 2. Total de *reads* geradas pelo RNA-seq das amostras de células mTEC 3.10 WT e mTEC 3.10E6 (Aire^{-/-}) e quantidade de *reads* sobreviventes após a trimagem das amostras.

Amostra	Nº de reads antes da trimagem	Nº de reads após a trimagem	Reads “sobreviventes” (%)
WT Sem Adesão 1	21.842.170	19.537.448	89,45
WT Sem Adesão 2	20.288.575	18.176.342	89,59
WT Após Adesão 1	23.459.757	20.570.971	87,69
WT Após Adesão 2	25.575.705	23.314.049	91,16
E6 Sem Adesão 1	19.788.406	17.169.731	86,77
E6 Sem Adesão 2	24.613.831	22.271.704	90,48
E6 Após Adesão 1	21.161.427	19.514.908	92,22
E6 Após Adesão 2	21.732.873	19.700.249	90,65
			Média= 89,75

6.3 Montagem do genoma

Após as análises de qualidade as amostras foram mapeadas contra o genoma de referência murino versão M18 (Ano 2018) (GRCm38.p6), obtendo-se uma média de mapeamento de 90,20% (Tabela 3).

Tabela 3. Total de *reads* de input e quantidade de *reads* mapeadas contra o genoma de referência versão M18 (GRCm38.p6) das células mTEC 3.10 WT e mTEC 3.10E6 (Aire^{-/-}).

Amostra	Total de reads	Reads mapeadas	Reads mapeadas (%)
WT Sem Adesão 1	19.537.448	18.059.610	92,44
WT Sem Adesão 2	18.176.342	14.702.382	80,89
WT Após Adesão 1	20.570.971	17.791.375	86,49
WT Após Adesão 2	23.314.049	21.566.986	92,51
E6 Sem Adesão 1	17.169.731	15.845.135	92,29
E6 Sem Adesão 2	22.271.704	20.534.757	92,20
E6 Após Adesão 1	19.514.908	18.029.272	92,39
E6 Após Adesão 2	19.700.249	18.203.981	92,40
			Média= 90,20

Após o mapeamento as amostras foram divididas em um grupo composto pelas linhagens celulares antes do ensaio de adesão e um outro grupo composto pelas

amostras de células que passaram pelo ensaio de adesão. A partir disso realizou-se as análises dos lncRNAs diferencialmente expressos.

6.4 Análise dos lncRNAs diferencialmente expressos

Utilizou-se o pacote edgeR para analisar o lncRNAs diferencialmente expressos nas células mTEC 3.10E6 (Aire^{-/-}) em comparação com a linhagem mTEC 3.10 WT antes e após o ensaio de adesão, foi definido como diferencialmente expresso genes que nas análises tivessem um Fold Change de no mínimo 2.0 com um valor de p ajustado (FDR) $\leq 0,05$.

Na análise dos lncRNAs antes do ensaio de adesão mTECs-timócitos foram identificados 136 lncRNAs diferencialmente expressos, sendo que 72 lncRNAs estavam hiperregulados e 64 hiporregulados, conforme é demonstrado no *volcano plot* (Figura 15) e no *heatmap* de agrupamento hierárquico não supervisionado (Figura 16); além disso os 10 lncRNAs mais hiper e hipoexpressos foram destacados nas tabelas 4 e 5. Podemos considerar esses 136 lncRNAs Aire-dependentes, uma vez que a ausência de Aire resultou na expressão diferencial dos mesmos.

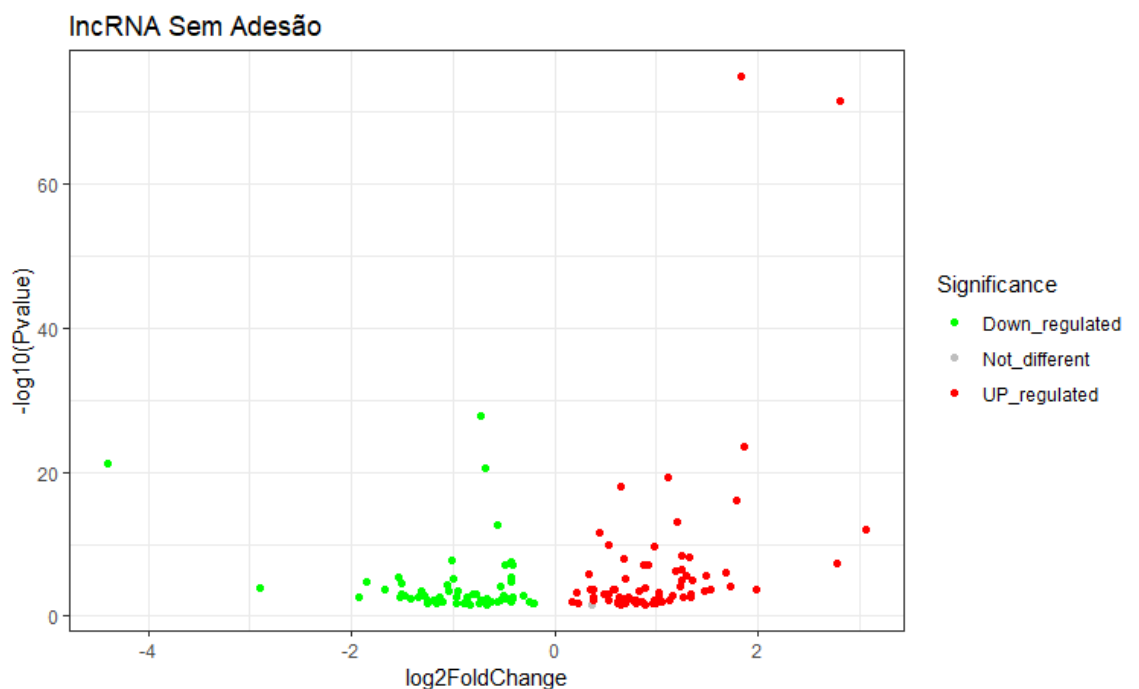


Figura 15. *Volcano plot* mostrando a distribuição dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire^{-/-}) e comparação com células mTECs 3.10 WT antes do ensaio de adesão (fold change ≥ 2.0 , FDR ≤ 0.05 , Hiperregulados (UP_regulated), Hiporregulados (Down_regulated) .

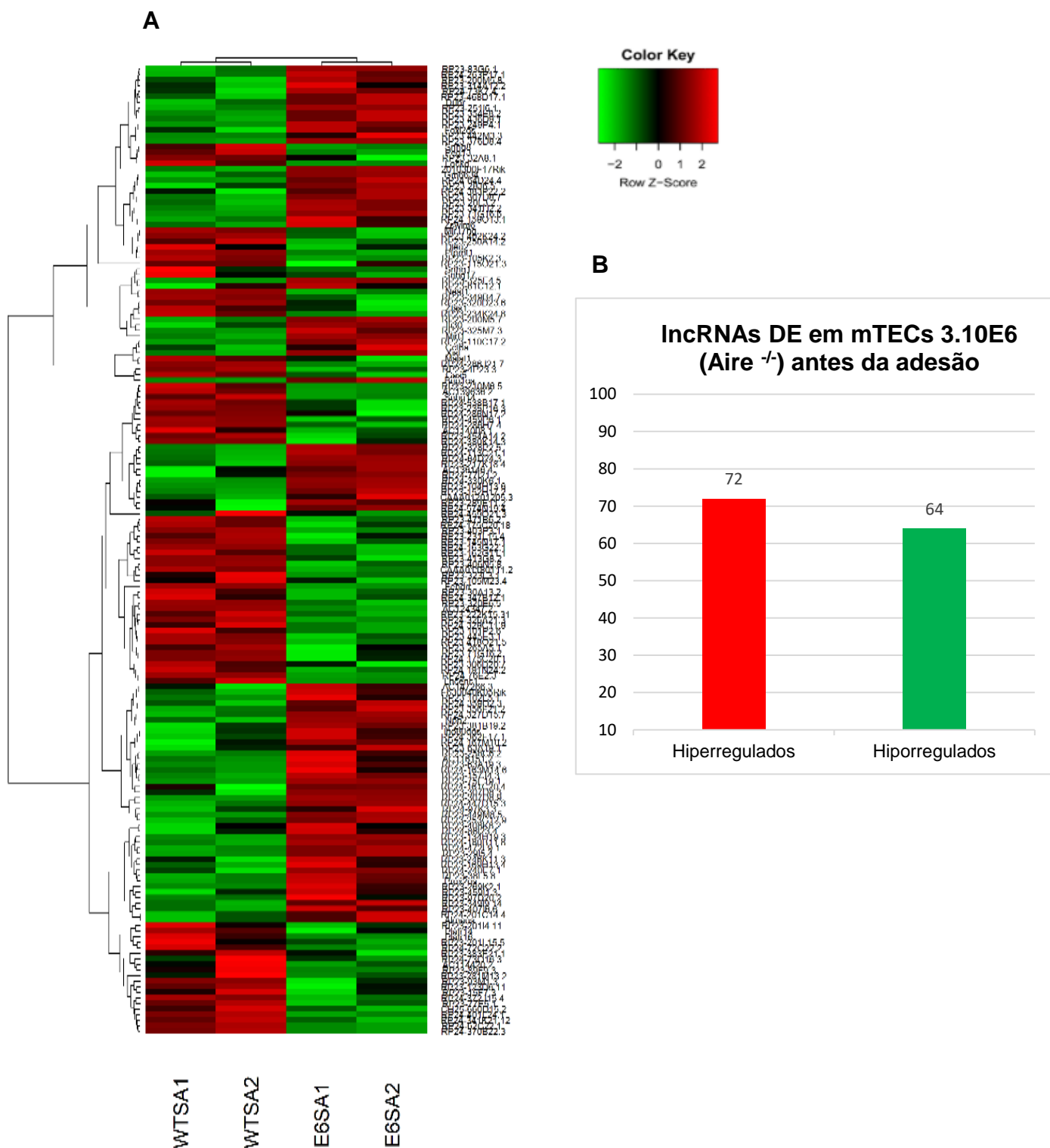


Figura 16. A) *Heatmap* do agrupamento hierárquico não supervisionado dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire^{-/-}) e comparação com células mTECs 3.10 WT antes do ensaio de adesão (fold change \geq 2.0, FDR \leq 0.05 (WTSA1: mTECs 3.10 WT sem adesão amostra 1, WTSA2: mTECs 3.10 WT sem adesão amostra 2, E6SA1: mTECs 3.10E6 (Aire^{-/-}) sem adesão amostra 1, E6SA2: mTECs 3.10E6 (Aire^{-/-}) sem adesão amostra 2). **B)** dos 136 lncRNAs diferencialmente expressos, 72 lncRNAs estavam hiperregulados e 64 hiporregulados.

Tabela 4. Top 10 lncRNAs hiperregulados nas células mTEC 3.10E6 (Aire^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT sem ensaio de adesão com timócitos.

ID_Ensembl	lncRNA	Log2 (Fold Change)	FDR
ENSMUSG00000104811.1	Gm33093	3.05697094892126	1.68887769298415e-11
ENSMUSG00000097289.1	2010300F17Rik	2.80000087711663	1.49799978763693e-69
ENSMUSG00000113185.1	Gm47127	2.76977635492375	4.10197679502256e-07
ENSMUSG00000109021.1	Gm44812	1.98691880403124	0.000662818348209971
ENSMUSG00000104554.1	Gm4610	1.86186727758227	1.39654733697852e-22
ENSMUSG00000097636.7	Mirt1	1.82963943266738	6.24891059179398e-73
ENSMUSG00000104982.1	Gm32554	1.7898999306723	2.03458606660714e-15
ENSMUSG00000106121.1	Gm42679	1.72570384462814	0.000251621874711925
ENSMUSG00000087107.8	AI662270	1.687331514335	5.97330879447768e-06
ENSMUSG00000106652.1	Gm43358	1.52897408641536	0.000789113531882971

Tabela 5. Top 10 lncRNAs hiporregulados nas células mTEC 3.10E6 (Aire^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT sem ensaio de adesão com timócitos.

ID_Ensembl	lncRNA	Log2 (Fold Change)	FDR
ENSMUSG00000097336.8	Fendrr	-4.39841211281411	2.9813879363468e-20
ENSMUSG00000086324.8	Gm15564	-2.8992396959009	0.000588293004549254
ENSMUSG00000110537.1	Gm4316	-1.92139045902179	0.00778888982304384
ENSMUSG00000051832.7	E230016K23Rik	-1.84972000470977	7.52183360620377e-05
ENSMUSG00000087132.7	A930001C03Rik	-1.6690622794737	0.000788398326160964
ENSMUSG00000112627.1	4933412E12Rik	-1.54359798447323	2.4157528207743e-05
ENSMUSG00000048603.8	Gm9828	-1.52886402885372	0.00707536255075073
ENSMUSG00000086629.8	2810403D21Rik	-1.51183828152075	0.000119022734938128
ENSMUSG00000057359.7	Gm17494	-1.50537074849227	0.00250079518114827
ENSMUSG00000085147.1	Gm12609	-1.4788977971886	0.0037369013076587

Na análise da expressão diferencial dos lncRNAs nas mTEC 3.10E6 (Aire^{-/-}) após a co-cultura com timócitos durante 12 horas (adesão) identificou-se 33 lncRNAs expressos diferencialmente, sendo 26 hiperregulados e 7 hiporregulados. Para analisar o padrão de expressão dos lncRNAs diferencialmente expressos assim como no grupo sem adesão, foi elaborado um *volcano plot* (Figura 17) e um *heatmap* de agrupamento hierárquico não supervisionado (Figura 18), além das tabelas 06 e 07 contendo 10 lncRNAs mais hiperregulados e os lncRNAs hiporregulados.

Os resultados mostraram que houve uma redução no número de lncRNAs diferencialmente expressos durante o processo de adesão com timócitos na ausência de Aire, indicando que durante o processo de adesão há uma alteração no padrão de expressão dos lncRNAs.

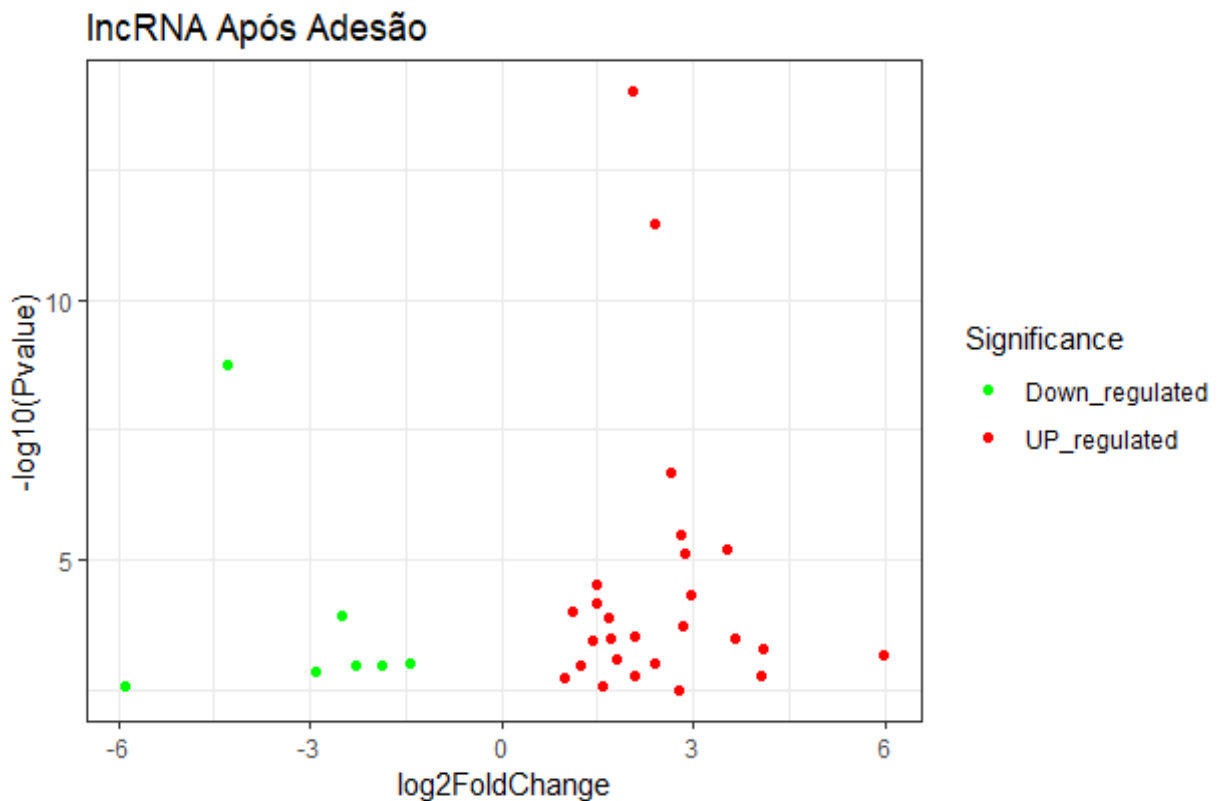


Figura 17. *Volcano plot* mostrando a distribuição dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire^{-/-}) e comparação com células mTECs 3.10 WT após o ensaio de adesão com timócitos (fold change \geq 2.0, $FDR \leq$ 0.05, Hiperreguados (UP_regulated), Hiporregulados (Down_regulated)).

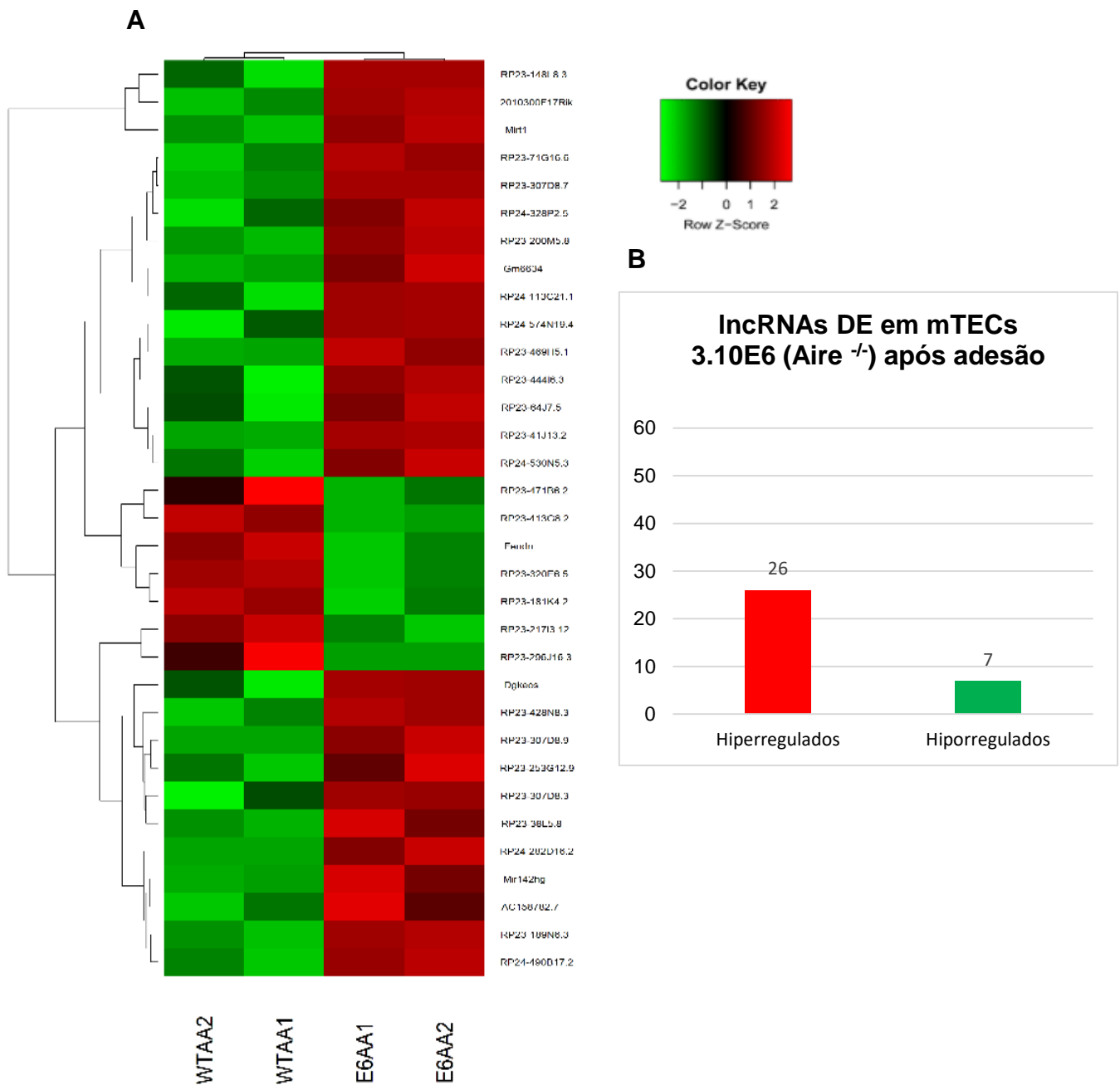


Figura 18. A) Heatmap do agrupamento hierárquico não supervisionado dos lncRNAs diferencialmente expressos em mTECs 3.10E6 (Aire^{-/-}) e comparação com células mTECs 3.10 WT após o ensaio de adesão (fold change \geq 2.0, FDR \leq 0.05 (WTSA1: mTECs 3.10 WT sem adesão amostra 1, WTSA2: mTECs 3.10 WT sem adesão amostra 2, E6SA1: mTECs 3.10E6 (Aire^{-/-}) sem adesão amostra 1, E6SA2: mTECs 3.10E6 (Aire^{-/-}) sem adesão amostra 2). **B)** dos 33 lncRNAs diferencialmente expressos, 26 lncRNAs estavam hiperregulados e 7 hiporregulados.

Tabela 6. Top 10 lncRNAs hiperregulados nas células mTEC 3.10E6 (Aire^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT após o ensaio de adesão com timócitos.

ID_Ensembl	lncRNA	Log2 (Fold Change)	FDR
ENSMUSG00000098959.1	Gm27230	5.9655969914345	0.0153160676415
ENSMUSG00000103358.1	Gm37593	4.09199945251259	0.0116725385928821
ENSMUSG00000085971.4	Gm15411	4.05899934931405	0.030137938877361
ENSMUSG00000084796.1	Mir142hg	3.66200316824931	0.00864720939550989
ENSMUSG00000063109.13	Dgkeos	3.51854929686628	0.0003396721470543
ENSMUSG00000113908.1	Gm10811	2.97052406437679	0.00183207179026624
ENSMUSG00000104811.1	Gm33093	2.86206663521232	0.00037556158158385
ENSMUSG00000109002.1	Gm38405	2.84112571329894	0.00527155218745538
ENSMUSG00000113049.1	Gm48544	2.78821575342399	0.000203126823303089
ENSMUSG00000106757.1	Gm43482	2.78217860226	0.0497819180882571

Tabela 7. Top 7 lncRNAs hiporregulados nas células mTEC 3.10E6 (Aire^{-/-}) em comparação com a linhagem de células mTEC 3.10 WT após o ensaio de adesão com timócitos.

ID_Ensembl	lncRNA	Log2 (Fold Change)	FDR
ENSMUSG00000115762.1	Gm34907	-5.89257169219668	0.0201938154702461
ENSMUSG00000097336.8	Fendrr	-4.28041421183493	0.0212035262397667
ENSMUSG00000086163.1	Gm14206	-2.91898491759021	0.0215481311130741
ENSMUSG00000100094.1	1810008I18Rik	-2.50524442052799	0.00375298653288945
ENSMUSG00000087143.9	A830082K12Rik	-2.28834716027512	0.0255770284702173
ENSMUSG00000051832.7	E230016K23Rik	-1.87340794765576	2.90299935067954e-07
ENSMUSG00000112627.1	4933412E12Rik	-1.44061656832222	0.0434567518724243

6.5 LncRNAs exclusivos e compartilhados antes e depois da adesão

Na identificação dos lncRNAs diferencialmente expressos exclusivos e compartilhados antes e depois da adesão com timócitos, todos os lncRNAs DE nas células mTEC 3.10E6 (Aire^{-/-}) foram analisados e partir dessa análise foi produzido um diagrama de Venn (Figura 19).

O resultado após análise de todos os lncRNAs mostrou que dos 136 lncRNAs expressos diferencialmente antes da adesão 121 são exclusivos dessa etapa, nas amostras de células após a adesão dos 33 lncRNAs DE 18 aparecem somente após o processo e que 15 lncRNAs diferencialmente expressos são compartilhados entre os grupos, sendo DE tanto antes quanto depois da adesão.

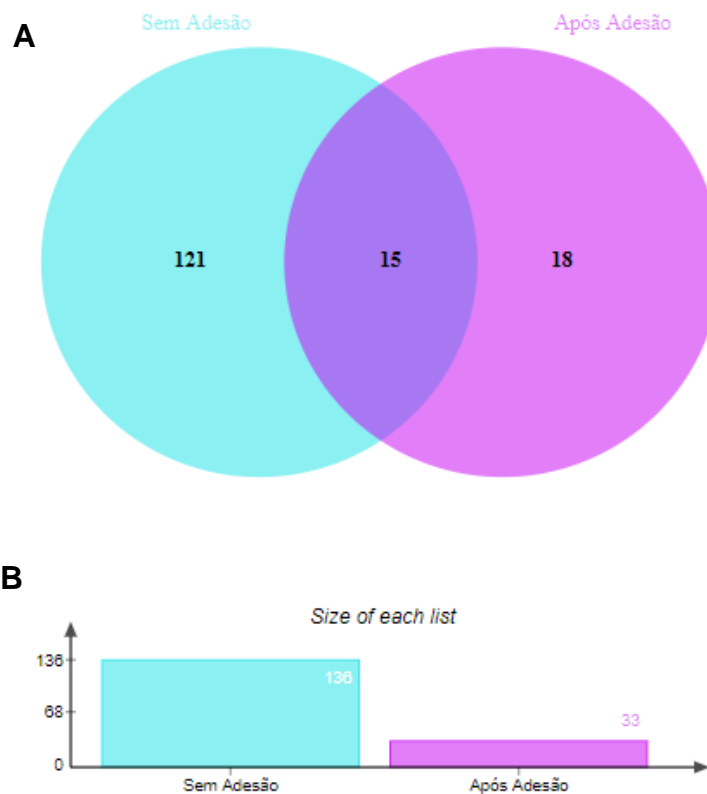


Figura 19. A) Diagrama de Venn mostrando a quantidade dos lncRNAs diferencialmente expressos exclusivos e compartilhados nas células mTEC 3.10E6 (Aire^{-/-}) antes e após adesão com timócitos. **B)** Total de lncRNAs diferencialmente expressos antes e depois da adesão.

Dos 15 lncRNAs compartilhados, descritos na tabela 8, os 11 lncRNAs que estavam hiperregulados e os 4 que estavam hiporregulados antes da adesão mantiveram o padrão de expressão após a adesão, não havendo influência da adesão sobre o padrão de expressão deles.

Tabela 8. Tabela dos lncRNAs diferencialmente expressos compartilhados antes e após adesão com timócitos.

ID_Ensembl	LncRNA
Hiperregulados	
ENSMUSG00000091422	Gm6455
ENSMUSG00000097252	Gm6634
ENSMUSG00000097289	2010300F17Rik
ENSMUSG00000097636	Mirt1
ENSMUSG00000104554	Gm4610
ENSMUSG00000104811	Gm33093
ENSMUSG00000104982	Gm32554
ENSMUSG00000106652	Gm43358
ENSMUSG00000106734	Gm20559
ENSMUSG00000108179	Gm44274
ENSMUSG00000113049	Gm48544
Hiporregulados	
ENSMUSG00000087143	A830082K12Rik
ENSMUSG00000051832	E230016K23Rik
ENSMUSG00000097336	Fendrr
ENSMUSG00000112627	4933412E12Rik

6.6 Classificação e proporção dos tipos de lncRNAs

Os lncRNAs diferencialmente expressos antes e depois do ensaio de adesão foram classificados de acordo com a divisão do GENCODE, sendo divididos nas categorias descritas nas figuras 6 e 7. Para ambos os grupos os lncRNAs mais frequentes foram os do tipo intergênico (“long intergenic non-coding RNA” - lincRNA). Tendo em vista os poucos estudos sobre lncRNAs no genoma murino outra classe que chama atenção é a de lncRNAs que foram preditos mas ainda precisam ser validados experimentalmente, que pertencem ao grupo “to be experimentally confirmed transcript” (TEC).

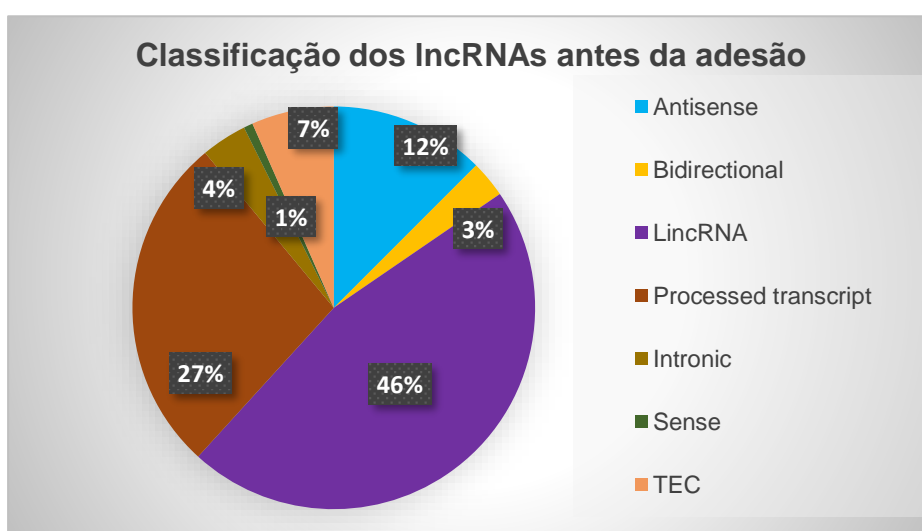


Figura 20. Proporção e tipos de lncRNAs encontrados nas células mTEC 3.10E6 (*Aire*^{-/-}) antes do ensaio de adesão.

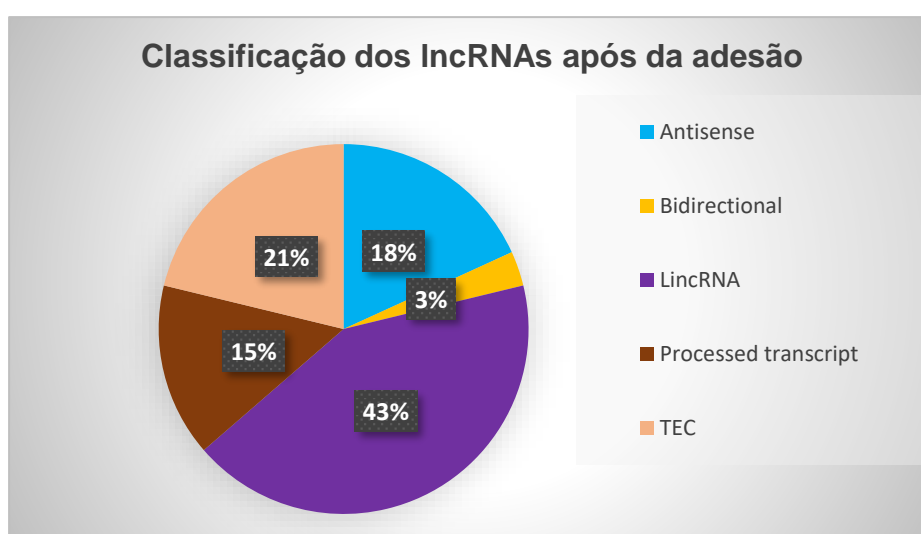


Figura 21. Proporção e tipos de lncRNAs encontrados nas células mTEC 3.10E6 (*Aire*^{-/-}) após o ensaio de adesão.

6.7 Distribuição Cromossômica dos lncRNAs

As análises de bioinformática permitem designar a localização cromossômica dos genes de interesse, assim foi possível determinar a distribuição dos lncRNAs diferencialmente expressos antes (Figura 22) e depois da adesão (Figura 23) ao longo dos cromossomos do genoma murino.

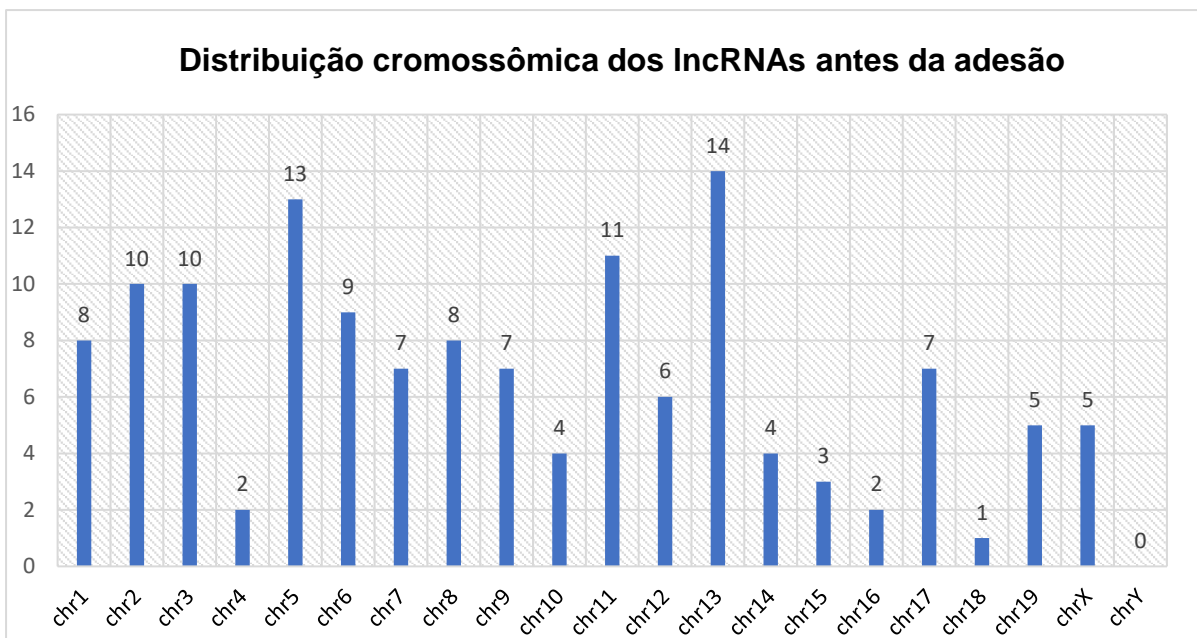


Figura 22. Distribuição cromossômica dos 136 lncRNAs diferencialmente expressos em mTEC 3.10E6 (Aire^{-/-}) ao longo do genoma do camundongo antes da adesão.

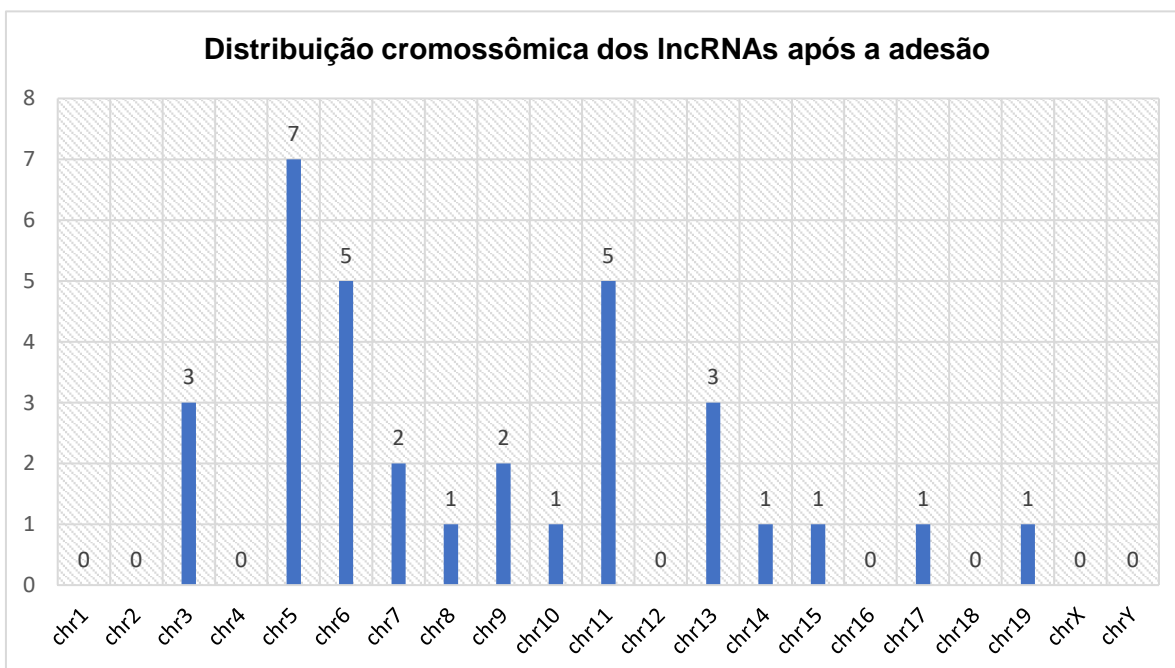


Figura 23. Distribuição cromossômica dos 33 lncRNAs diferencialmente expressos em mTEC 3.10E6 (Aire^{-/-}) ao longo do genoma do camundongo após a adesão.

6.8 Identificação dos genes codificadores de proteínas mais próximos

Por meio da ferramenta biomaRt foi possível identificar os genes codificadores de proteínas mais próximos, tanto os que estão a montante (*upstream*) quanto os que estão a jusante (*downstream*) em relação aos lncRNAs diferencialmente expressos, conforme mostram as tabelas 9, 10, 11 e 12.

Tabela 9. Genes codificadores de proteínas mais próximos dos lncRNAs hiperregulados antes da adesão.

lncRNA	Upstream	Posição	Downstream	Posição
Cct6a	Psph	7825	Sumf2	59988
AI506816	Srpk2	82742	Pus7	84435
Gm26917	Esp15	202059	Esp38	107507
Brip1os	Brip1	177	Ints2	56180
Ifi30	Mpv17l2	5717	Pik3r2	10050
D930016D06Rik	BC005561	17383	Zfp951	334335
Gm49654	Ttc9	88803	Map3k9	61123
5430416N02Rik	Sec31a	50622	Lin54	71104
Foxl2os	7420426K07Rik	52180	Foxl2	-11
Xist	Gm26992	61244	Zcchc13	147332
Dubr	Ccdc54	141619	Cblb	1298452
1110002L01Rik	Kif3c	61879	Asxl2	-154
Gm28875	Clmn	61613	Syne3	83120
Gm44275	Pbp2	9156	Emp1	42984
Gm20544	Dym	348307	Smad7	441
Gm19589	Glrp1	41451	Arl4c	150359
Gm44274	Pbp2	1306	Emp1	50834
B130055M24Rik	Zfp689	-522	Prr14	10975
F730311O21Rik	Nuak2	25897	Tmcc2	49258
Zswim6	Kif2a	867946	Ndufaf2	268559
Tmem250-ps	Nacc2	48344	Lhx3	67768
Gm10501	Hspa1l	9085	Lsm2	74
Morrbid	Bcl2l11	303327	Spdye4c	161843
4930558J18Rik	Satb2	405042	1700066M21Rik	76
2610035D17Rik	Sox9	419614	Slc39a11	363918
Gm6455	Gm5152	618537	Gm8871	175086
Gm47813	Net1	6670	Tuba13	18026
Gm20554	Irx1	664841	Irx2	256
C030037D09Rik	Msi2	416	Akap1	132858
Gm14221	Mafb	252906	Top1	25917
Gm47507	Klf6	25627	Pitrm1	661040
2810455O05Rik	Rab4a	34800	Ccsap	18745
2700038G22Rik	Tomm7	6436	Fam126a	72307
Gm44040	Itpr1	312591	Bhlhe40	134942

Gm20559	Gm26463	-148356397	Samd9l	53501
Gm20744	Bnc1	78095	Sh3gl3	103453
Gm26651	Gadd45g	126536	Diras2	558065
Gm15706	Kras	290	Lmntd1	363744
3300005D01Rik	Zdhhc14	306100	Snx9	42672
Gm42979	Sfswap	54926	Mmp17	28022
Ino80dos	Ino80d	21149	Ndufs1	31883
F630040K05Rik	1600014C23Rik	662	Vegfa	297309
Gm10516	Rcor3	-1114	Kcnh1	53878
Gm35940	Ubash3b	647350	Sorl1	318885
Gm36377	Ice1	82594	Adamts16	186024
2900005J15Rik	Prkag2	816	Galntl5	80002
Gm37145	Pigm	14443	Slamf9	84369
Gm20712	Fer1l6	271903	Tmem65	41687
Gm26532	Nanos3	23703	Zswim4	28135
Gm12022	Etaa1	539380	Meis1	525730
Gm48544	Prl2c2	29475	Prl2c5	147857
9430018G01Rik	Tpk1	-198299	Cntnap2	1591378
Gm37352	Nanos3	15652	Zswim4	36186
Gm14547	Cald1	136889	Agbl3	45043
Gm10138	Gm47996	133951	Ivns1abp	395
Gm35330	Akr1e1	1115336	Klf6	136972
Gm6634	Otol1	799678	Sis	2160572
Gm19619	Areg	143822	Btc	119573
Gm13986	Rasgrp1	768201	Thbs1	674
Gm29093	Abhd17a	-551	Scamp4	13082
A230001M10Rik	Vangl1	57712	Ngf	207514
Npn2	Cald1	128406	Agbl3	53526
Gm43358	Gm5152	700079	Gm8871	93544
Al662270	Slfn3	32246	Slfn14	63150
Gm42679	Ngf	21423	Tspan2	243187
Gm32554	Gm8879	54191	Gm5861	1247
Mirt1	Smndc1	75361	Dusp5	64326
Gm4610	Bcar3	-120411	Fnbp1l	172070
Gm44812	Nr2f2	214241	Gm10295	776651
Gm47127	Slc22a23	267795	Pxdc1	41240
2010300F17Rik	Skap1	282817	Snx11	30121
Gm33093	Gm5861	71681	Gm8890	1280

Tabela 10. Genes codificadores de proteínas mais próximos dos lncRNAs hiporregulados antes da adesão.

lncRNA	Upstream	Posição	Downstream	Posição
Firre	Olf1324	136385	Stk26	276164
Snhg17	Lbp	50350	Snhg11	18795
Snhg1	Slc3a2	762	Wdr74	11696
Zfas1	Znfx1	1010	Kcnb1	125169
Lockd	Cdkn1b	8691	Apold1	52626
Peg13	Trappc9	-245486	Chrac1	280068
Gas5	Zbtb37	-427	Dars2	36236
Malat1	Scyl1	31253	Frmd8	72602
1110038B12Rik	Neu1	19991	Hspa1b	7994
Gm49759	Pkm	16446	Gramd2	7330
5530601H04Rik	Magee2	212857	Pbdc1	9632
Mir17hg	Tpm3-rs7	1728271	Gpc5	49336
Gm37494	Zfp619	26636	Vstm2b	1353926
2410006H16Rik	Trpv2	28391	Lrrc75a	45646
1810026B05Rik	Chd2	16649	Fam174b	181907
Snhg8	Prss12	61491	Ndst3	182449
Ptpmt1	Kbtbd4	13117	Celf1	22525
Gm10076	Ndfip2	423256	Spry2	214990
Dleu2	Trim13	84126	Kcnrg	-74916
4933407K13Rik	Rab39b	186468	Pls3	110483
Gm15708	Baiap2l1	-81051	Nptx2	268841
4933439C10Rik	Zkscan17	-982	Nlrp3	35906
Gm16046	Tcp10c	328925	Tcte2	77881
Neat1	Scyl1	74059	Frmd8	29796
Gm28229	Gm20537	-12964	Syncrip	38907
Gm42047	Klhdc4	3678	Slc7a5	74447
Gm16136	Zfp706	225907	Grhl2	-273
2700033N17Rik	Meis2	2397	Tmco5	811264
Snhg14	Ube3a	95399	Gm38393	680915
Gm21781	Zbtb2	4112	Rmnd1	40801
Gdap10	Nampt	2327	Sypl	132019
Zfp469	Banp	319819	Zfpm1	12572
Gm47204	Upk2	24783	Bcl9l	3066
Gm19605	Ppp2r5c	39084	Dync1h1	115248
Gm45737	Uevld	-29103	Gm9999	54156
E130202H07Rik	Tsc22d1	111631	Serp2	30080
1110046J04Rik	Serpib6a	-631	Nqo2	29120
A830082K12Rik	Nr2f1	-1740	Arrdc3	2685367
Gm39469	Ccr5	160596	2010315B03Rik	30557
2810429I04Rik	Gm27236	-116522297	Gdi2	59833
4632415L05Rik	Crh	199502	Cp	62156
Gm4117	Vcan	-2862	Xrcc4	350063
Gm13431	Mrrf	80388	Ptgs1	13649

Gm11515	Dlx4	79369	Tac4	35897
Gm20707	Zfp984	34007	Zfp933	38526
Gm43668	Dnajc19	-10653	Sox2	579746
4833417C18Rik	Gngt2	21649	B4galnt2	56026
1110018N20Rik	Kcnb1	843	Ptgis	50943
Gm17249	Tcp11l2	45952	Polr3b	-286
2810029C07Rik	Eif5	36386	Mark3	121
A530020G20Rik	Slc44a3	70708	4930432M17Rik	139143
E330034L11Rik	Ensa	14311	Mcl1	19484
Gm19412	Zbtb22	4581	Rgl2	9058
C030014I23Rik	Upk2	13368	Bcl9l	14481
Gm12609	Cdkn2b	70918	Dmrta1	297479
Gm17494	Gclm	47990	Dnttip2	-19159
2810403D21Rik	Brwd3	-548	Hmgn5	179604
Gm9828	Btf3	102	Foxd1	37173
4933412E12Rik	Rab3ip	-674	Gm10271	22852
A930001C03Rik	Rhod	-429	Syt12	38444
E230016K23Rik	Ccl9	3420	Ccl6	11031
Gm4316	Sult5a1	6985	Dpep1	20987
Gm15564	Parp9	42611	Kpna1	-2331
Fendrr	Irf8	346674	Foxf1	1354

Tabela 11. Genes codificadores de proteínas mais próximos dos lncRNAs hiperregulados após o ensaio de adesão mTECs-timócitos.

lncRNA	Upstream	Posição	Downstream	Posição
Gm44274	Pbp2	1306	Emp1	50834
Gm20186	Cttnbp2	418532	Lsm8	1816
Gm6634	Otol1	799678	Sis	2160572
Gm6455	Gm5152	618537	Gm8871	175086
Gm4610	Bcar3	-120411	Fnbp1l	172070
Gm32554	Gm8879	54191	Gm5861	1247
AU022754	Wnt7b	-1330	Ppara	154731
Gm20559	Gm26463	-148356397	Samd9l	53501
Gm38077	Mrpl3	11538	Nudt16	66535
Gm42141	Lrrc8b	79695	Lrrc8c	23918
Mirt1	Smndc1	75361	Dusp5	64326
2310001H17Rik	Clec2d	57333	Cd69	37488
Gm43197	Gm26463	-148363114	Samd9l	60218
2010300F17Rik	Skap1	282817	Snx11	30121
Gm43358	Gm5152	700079	Gm8871	93544
AI504432	Kcna3	3335	Kcna2	61642
Gm43482	Vkorc1l1	22396	Gusb	38683
Gm48544	Prl2c2	29475	Prl2c5	147857
Gm38405	Xrra1	30242	Rnf169	91090

Gm33093	Gm5861	71681	Gm8890	1280
Gm10811	Prl2c2	165794	Prl2c5	11538
Dgkeos	Dgke	-762	Gm525	7753
Mir142hg	Supt4a	18025	Tspoap1	4964
Gm15411	Gm42791	20666	Hmgb1	37227
Gm37593	Satb1	10189	Kcnh8	759848
Gm27230	Zfp280d	121114	Mns1	42037

Tabela 12. Genes codificadores de proteínas mais próximos dos lncRNAs hiporegulados após o ensaio de adesão mTECs-timócitos.

lncRNA	Upstream	Posição	Downstream	Posição
4933412E12Rik	Rab3ip	-674	Gm10271	22852
E230016K23Rik	Ccl9	3420	Ccl6	11031
A830082K12Rik	Nr2f1	-1740	Arrdc3	2685367
1810008I18Rik	Tm2d3	111942	Pcsk6	58623
Gm14206	Krt42	25308	Eif1	28599
Fendrr	Irf8	346674	Foxf1	1354
Gm34907	Mycbp2	308656	Scel	49423

6.9 Determinação do potencial codificador dos lncRNAs

De maneira geral, lncRNAs não possuem capacidade de codificar proteínas, no entanto alguns lncRNAs podem codificar pequenos peptídeos, mas não codificam uma proteína completa.

A fim de melhor caracterizar os lncRNAs DE, utilizou-se a ferramenta Coding Potential Calculator 2 - CPC2 para determinar o potencial de codificação dos mesmos, onde lncRNAs com potencial codificador < 0.5 são classificados como *noncoding*, ou seja, sem potencial de codificar moléculas, e lncRNAs com potencial codificador > 0.5 são classificados como *coding*, indicando que esses podem codificar pequenos peptídeos.

Tabela 13. Potencial codificador dos lncRNAs hiperregulados antes do ensaio de adesão.

ID lncRNA	Classificação	Potencial codificador
ENSMUSG00000031838 lfi30	noncoding	0.0862842
ENSMUSG00000029447 Cct6a	coding	0.934241
ENSMUSG00000097971 Gm26917	noncoding	0.429843
ENSMUSG00000085208 Brip1os	noncoding	0.410905
ENSMUSG00000105987 Al506816	noncoding	0.123182
ENSMUSG00000086503 Xist	noncoding	0.196558
ENSMUSG00000097072 Foxl2os	noncoding	0.234836
ENSMUSG00000097392 D930016D06Rik	noncoding	0.152485
ENSMUSG00000097772 5430416N02Rik	noncoding	0.144062
ENSMUSG00000113184 Gm49654	noncoding	0.143305
ENSMUSG00000092569 Gm20544	noncoding	0.0243612
ENSMUSG00000022639 Dubr	noncoding	0.209164
ENSMUSG00000100975 Gm28875	noncoding	0.00934646
ENSMUSG00000071456 1110002L01Rik	noncoding	0.200819
ENSMUSG00000107768 Gm44275	noncoding	0.0764748
ENSMUSG00000100798 Gm19589	noncoding	0.113099
ENSMUSG00000032846 Zswim6	noncoding	0.107888
ENSMUSG00000086502 B130055M24Rik	noncoding	0.10458
ENSMUSG00000100658 F730311O21Rik	noncoding	0.161147
ENSMUSG00000108179 Gm44274	noncoding	0.00468821
ENSMUSG00000073415 Gm10501	noncoding	0.381583
ENSMUSG00000087679 Tmem250-os	noncoding	0.292669
ENSMUSG00000097519 4930558J18Rik	noncoding	0.0881016
ENSMUSG00000074813 Morrbid	noncoding	0.287118
ENSMUSG00000087259 2610035D17Rik	noncoding	0.440043
ENSMUSG00000087574 C030037D09Rik	noncoding	0.220116
ENSMUSG00000093587 Gm20554	noncoding	0.151178
ENSMUSG00000091422 Gm6455	noncoding	0.0449664
ENSMUSG00000113128 Gm47813	noncoding	0.0315369
ENSMUSG00000078956 Gm14221	noncoding	0.161748
ENSMUSG00000107458 Gm44040	noncoding	0.0945394
ENSMUSG00000106734 Gm20559	noncoding	0.0676802
ENSMUSG00000097180 2700038G22Rik	noncoding	0.125934
ENSMUSG00000085030 2810455O05Rik	noncoding	0.0513749
ENSMUSG00000113757 Gm47507	noncoding	0.376405
ENSMUSG00000104814 Gm42979	noncoding	0.0316422
ENSMUSG00000096965 3300005D01Rik	noncoding	0.132757
ENSMUSG00000097495 Gm26651	noncoding	0.251243
ENSMUSG00000086013 Gm15706	noncoding	0.0240817
ENSMUSG00000109145 Gm20744	noncoding	0.20815
ENSMUSG00000084799 lno80dos	noncoding	0.080607
ENSMUSG00000097316 Gm10516	noncoding	0.0906968
ENSMUSG00000113909 Gm36377	noncoding	0.0744511

ENSMUSG00000110981 Gm35940	noncoding	0.309178
ENSMUSG00000097727 F630040K05Rik	noncoding	0.206769
ENSMUSG00000097296 Gm26532	noncoding	0.0399538
ENSMUSG00000043833 2900005J15Rik	noncoding	0.265626
ENSMUSG00000093677 Gm20712	noncoding	0.019848
ENSMUSG00000102278 Gm37145	noncoding	0.0534146
ENSMUSG00000085675 Gm12022	noncoding	0.0961452
ENSMUSG00000103593 Gm37352	noncoding	0.129761
ENSMUSG00000086834 Gm14547	noncoding	0.205799
ENSMUSG00000107709 9430018G01Rik	noncoding	0.0285769
ENSMUSG00000100954 Gm10138	noncoding	0.0934414
ENSMUSG00000113049 Gm48544	noncoding	0.0963495
ENSMUSG00000100201 Gm29093	noncoding	0.0153925
ENSMUSG00000107198 Gm19619	noncoding	0.15874
ENSMUSG00000097252 Gm6634	noncoding	0.165308
ENSMUSG00000087203 Gm13986	noncoding	0.100344
ENSMUSG00000114053 Gm35330	noncoding	0.112498
ENSMUSG00000106121 Gm42679	noncoding	0.0360897
ENSMUSG00000087107 Al662270	noncoding	0.260206
ENSMUSG00000106652 Gm43358	noncoding	0.10157
ENSMUSG00000060411 Npn2	noncoding	0.00306937
ENSMUSG00000105891 A230001M10Rik	noncoding	0.0550794
ENSMUSG00000109021 Gm44812	noncoding	0.0914372
ENSMUSG00000097289 2010300F17Rik	noncoding	0.296135
ENSMUSG00000104982 Gm32554	noncoding	0.10804
ENSMUSG00000097636 Mirt1	noncoding	0.20149
ENSMUSG00000113185 Gm47127	noncoding	0.0436181
ENSMUSG00000104811 Gm33093	noncoding	0.0741184
ENSMUSG00000104554 Gm4610	noncoding	0.128126

Tabela 14. Potencial codificador dos lncRNAs hiporregulados antes do ensaio de adesão.

ID lncRNA	Classificação	Potencial codificador
ENSMUSG00000085396 Firme	noncoding	0.200936
ENSMUSG00000098318 Lockd	noncoding	0.466468
ENSMUSG00000074578 Zfas1	noncoding	0.0989879
ENSMUSG00000085385 Snhg17	noncoding	0.232932
ENSMUSG00000108414 Snhg1	noncoding	0.0546895
ENSMUSG00000053332 Gas5	noncoding	0.196687
ENSMUSG00000092203 1110038B12Rik	noncoding	0.0616174
ENSMUSG00000106847 Peg13	noncoding	0.0720553
ENSMUSG00000092341 Malat1	noncoding	0.14466
ENSMUSG00000111394 Gm49759	noncoding	0.0606007
ENSMUSG00000089726 Mir17hg	noncoding	0.0509525

ENSMUSG00000101970 1810026B05Rik	noncoding	0.317742
ENSMUSG00000086841 2410006H16Rik	noncoding	0.0340274
ENSMUSG00000087174 5530601H04Rik	noncoding	0.482448
ENSMUSG00000108621 Gm37494	noncoding	0.110088
ENSMUSG00000063235 Ptpmt1	noncoding	0.104815
ENSMUSG00000097589 Dleu2	noncoding	0.31624
ENSMUSG00000060143 Gm10076	noncoding	0.0126527
ENSMUSG00000087396 4933407K13Rik	noncoding	0.114618
ENSMUSG00000104960 Snhg8	noncoding	0.0132288
ENSMUSG00000072893 4933439C10Rik	noncoding	0.274861
ENSMUSG00000085705 Gm16046	noncoding	0.0631734
ENSMUSG00000086894 Gm15708	noncoding	0.127058
ENSMUSG00000092274 Neat1	noncoding	0.178003
ENSMUSG00000101335 Gm28229	noncoding	0.0320261
ENSMUSG00000110631 Gm42047	coding	1
ENSMUSG00000089804 Gm16136	noncoding	0.107214
ENSMUSG00000086328 2700033N17Rik	noncoding	0.273697
ENSMUSG00000095123 Gm21781	noncoding	0.161664
ENSMUSG00000100826 Snhg14	noncoding	0.0509525
ENSMUSG00000096954 Gdap10	noncoding	0.189302
ENSMUSG00000110980 Gm47204	noncoding	0.240217
ENSMUSG00000043903 Zfp469	coding	0.995812
ENSMUSG00000109231 Gm45737	noncoding	0.0238138
ENSMUSG00000113585 Gm19605	noncoding	0.0847596
ENSMUSG00000097709 2810429I04Rik	noncoding	0.241033
ENSMUSG00000110834 Gm39469	noncoding	0.068416
ENSMUSG00000087143 A830082K12Rik	coding	0.859375
ENSMUSG00000085457 1110046J04Rik	noncoding	0.089534
ENSMUSG00000115092 E130202H07Rik	noncoding	0.07083
ENSMUSG00000085992 Gm11515	noncoding	0.175652
ENSMUSG00000048106 4632415L05Rik	noncoding	0.192054
ENSMUSG00000089940 Gm4117	noncoding	0.181073
ENSMUSG00000087187 Gm13431	noncoding	0.0857193
ENSMUSG00000093594 Gm20707	noncoding	0.238579
ENSMUSG00000085151 1110018N20Rik	noncoding	0.190141
ENSMUSG00000097309 Gm17249	noncoding	0.120235
ENSMUSG00000086015 4833417C18Rik	noncoding	0.0653277
ENSMUSG00000096995 2810029C07Rik	noncoding	0.112825
ENSMUSG00000105176 Gm43668	noncoding	0.0940917
ENSMUSG00000101438 Gm19412	noncoding	0.00543772
ENSMUSG00000085147 Gm12609	coding	0.867102
ENSMUSG00000060380 C030014I23Rik	noncoding	0.0934543
ENSMUSG00000097124 A530020G20Rik	noncoding	0.121943
ENSMUSG00000106383 E330034L11Rik	noncoding	0.0154079
ENSMUSG00000087132 A930001C03Rik	noncoding	0.0640057
ENSMUSG00000057359 Gm17494	noncoding	0.338265

ENSMUSG00000086629 2810403D21Rik	noncoding	0.0931799
ENSMUSG00000048603 Gm9828	noncoding	0.0894544
ENSMUSG00000112627 4933412E12Rik	noncoding	0.22101
ENSMUSG00000086324 Gm15564	noncoding	0.18726
ENSMUSG00000110537 Gm4316	noncoding	0.143081
ENSMUSG00000097336 Fendrr	noncoding	0.240908
ENSMUSG00000051832 E230016K23Rik	noncoding	0.243372

Tabela 15. Potencial codificador dos lncRNAs hiperregulados após o ensaio de adesão mTECs-timócitos.

ID lncRNA	Classificação	Potencial codificador
ENSMUSG00000097252 Gm6634	noncoding	0.165308
ENSMUSG00000091422 Gm6455	coding	0.593718
ENSMUSG00000106874 Gm20186	noncoding	0.142522
ENSMUSG00000104554 Gm4610	coding	0.738779
ENSMUSG00000108179 Gm44274	noncoding	0.00468821
ENSMUSG00000105826 Gm42141	noncoding	0.260594
ENSMUSG00000104982 Gm32554	noncoding	0.10804
ENSMUSG00000097209 AU022754	noncoding	0.254906
ENSMUSG00000104344 Gm38077	noncoding	0.336198
ENSMUSG00000106734 Gm20559	noncoding	0.0676802
ENSMUSG00000097636 Mirt1	noncoding	0.20149
ENSMUSG00000097289 2010300F17Rik	noncoding	0.296135
ENSMUSG00000106652 Gm43358	noncoding	0.10157
ENSMUSG00000097354 2310001H17Rik	noncoding	0.207166
ENSMUSG00000107215 Gm43197	noncoding	0.0086279
ENSMUSG00000106757 Gm43482	noncoding	0.055349
ENSMUSG00000056145 AI504432	noncoding	0.130808
ENSMUSG00000109002 Gm38405	noncoding	0.107023
ENSMUSG00000104811 Gm33093	noncoding	0.0741184
ENSMUSG00000113049 Gm48544	noncoding	0.0963495
ENSMUSG00000103358 Gm37593	noncoding	0.0397225
ENSMUSG00000063109 Dgkeos	noncoding	0.111296
ENSMUSG00000085971 Gm15411	noncoding	0.319148
ENSMUSG00000098959 Gm27230	noncoding	0.146837
ENSMUSG00000084796 Mir142hg	noncoding	0.0533265
ENSMUSG00000113908 Gm10811	noncoding	0.0335535

Tabela 16. Potencial codificador dos lncRNAs hiporregulados após o ensaio de adesão mTECs-timócitos.

ID lncRNA	Classificação	Potencial codificador
ENSMUSG00000097336 Fendrr	noncoding	0.240908
ENSMUSG00000087143 A830082K12Rik	coding	0.696578
ENSMUSG00000100094 1810008I18Rik	noncoding	0.213922
ENSMUSG00000112627 4933412E12Rik	noncoding	0.0881181
ENSMUSG00000051832 E230016K23Rik	noncoding	0.243372
ENSMUSG00000086163 Gm14206	noncoding	0.077804
ENSMUSG00000115762 Gm34907	noncoding	0.275368

6.10 Enriquecimento funcional

O enriquecimento funcional dos lncRNAs diferencialmente expressos foi realizado com base no banco de dados KEGG. No enriquecimento dos lncRNAs DE antes da adesão, somente um lncRNA foi relacionado a processos biológicos já descritos para camundongos, que foi o lncRNA lfi30 (Figura 24), esse *long non-coding* participa do processamento e apresentação de antígenos. Na análise após adesão não foi encontrada relação entre os lncRNAs diferencialmente e os processos biológicos descritos no KEGG.

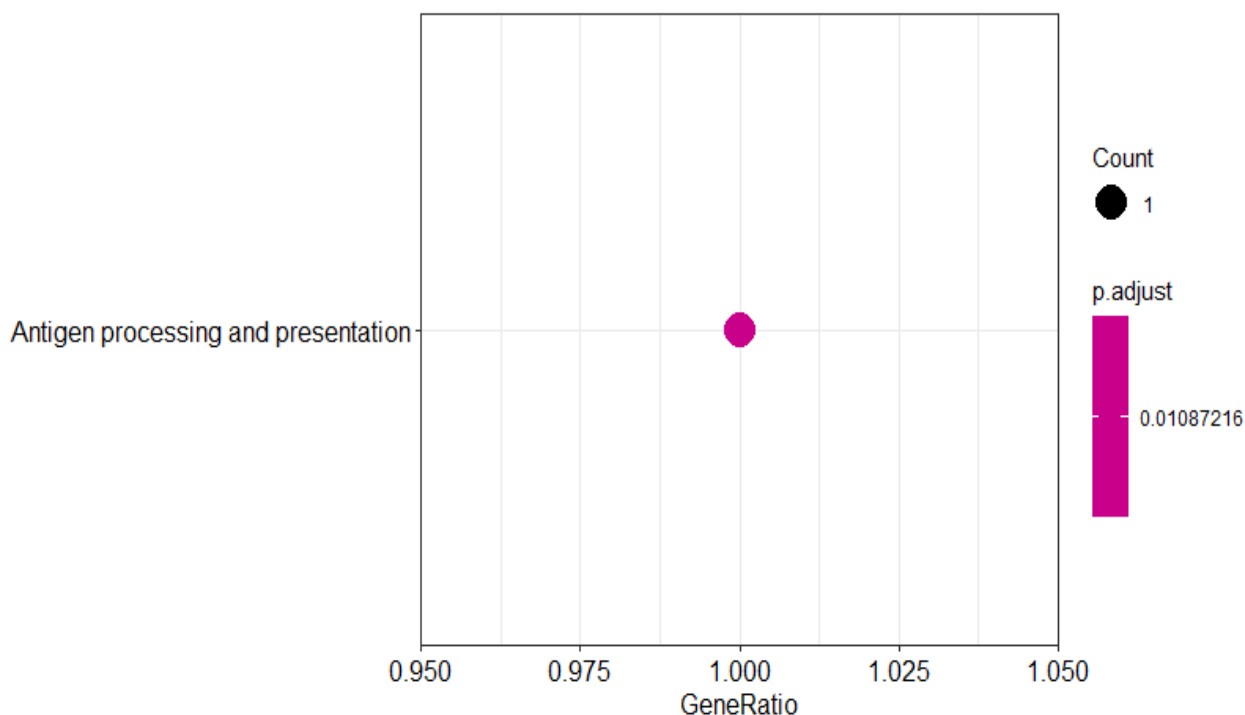


Figura 24. Enriquecimento funcional dos lncRNAs diferencialmente expressos em células mTEC 3.10E6 (*Aire*^{-/-}) antes do ensaio de adesão, realizado por meio do banco de dados KEGG.

6.11 Validação da expressão dos lncRNAs por RT- qPCR

Os níveis relativos de expressão dos lncRNAs foram validados por reações de RT-qPCR. Foram escolhidos 3 lncRNAs (Peg13, Dgkeos e Fendrr) que na análise *in silico* foram classificados como diferencialmente expressos, sendo que o lncRNA Peg13 estava diferencialmente expresso apenas antes da adesão (hiporregulado), o lncRNA Dgkeos diferencialmente expresso somente após a adesão com timócitos (hiperregulado) e o lncRNA Fendrr com expressão diferencial antes e após adesão (hiporregulado em ambas ocasiões). Os resultados obtidos corroboraram com os achados *in silico*, conforme mostram as figuras 25, 26 e 27.

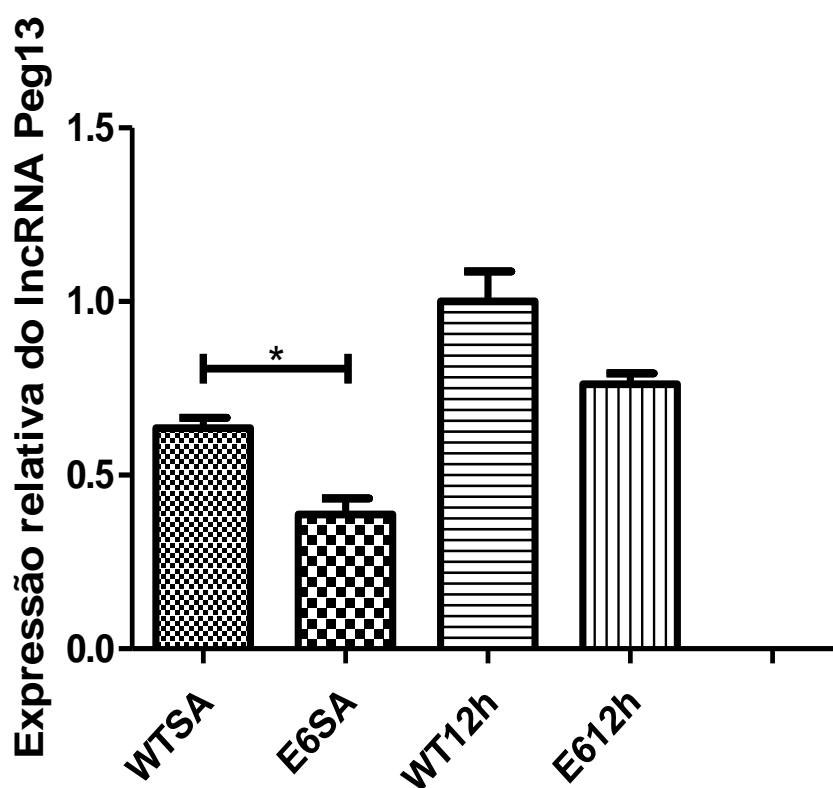


Figura 25. Quantificação relativa dos níveis de expressão do lncRNA Peg13 por RT – qPCR, onde o mesmo apresenta diferença significativa entre as células mTEC 3.10 WT e mTEC 3.10E6 (Aire^{-/-}) antes do ensaio de adesão, Teste t-student. $p = 0,0104$. Após o ensaio de adesão não houve diferença estatisticamente significativa entre as amostras, Teste t-student. $p = 0,0607$.

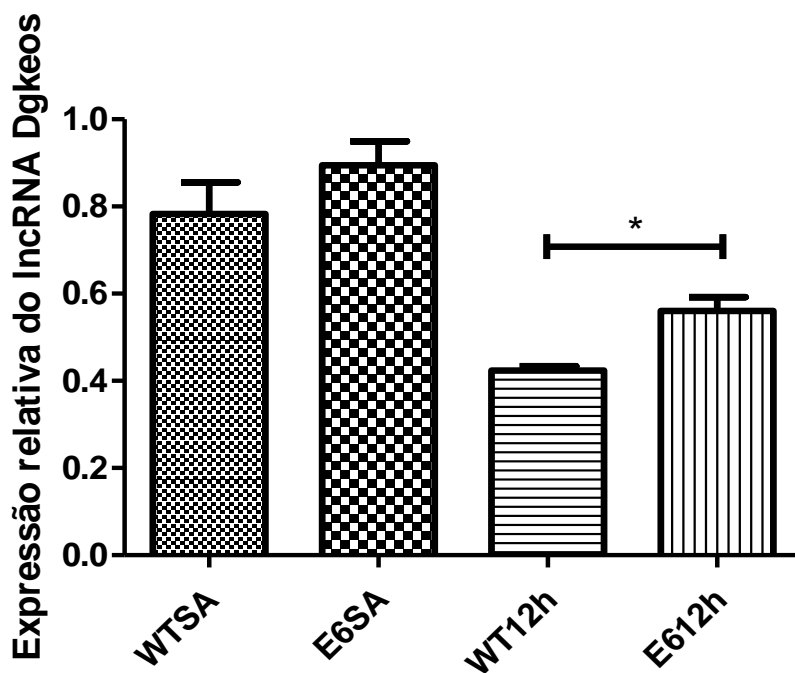


Figura 26. Quantificação relativa dos níveis de expressão do lncRNA Dgkeos por RT – qPCR, sem diferença significativa entre as células mTEC 3.10 WT e mTEC 3.10E6 (Aire^{-/-}) antes do ensaio de adesão, Teste t-student. $p = 0,2883$. Após o ensaio de adesão houve diferença estatisticamente significativa entre as amostras, Teste t-student. $p = 0,0148$.

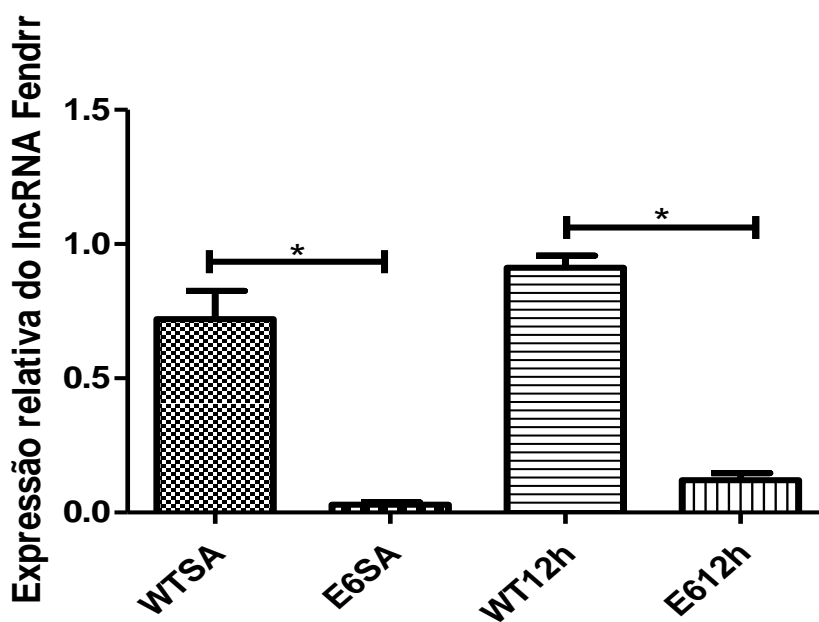


Figura 27. Quantificação relativa dos níveis de expressão do lncRNA Fendrr por RT – qPCR, o mesmo apresenta diferença estatisticamente significativa entre as células mTEC 3.10 WT e mTEC 3.10E6 (Aire^{-/-}) antes (Teste t-student. $p = 0,0029$) e após o ensaio de adesão mTEC-timócitos (Teste t-student. $p = 0,0001$).

DISCUSSÃO

7. DISCUSSÃO

A capacidade de distinção entre o próprio e o não-próprio é uma característica primordial do sistema imunológico, dentre as células envolvidas nesse processo, destacam-se as células mTECs, que durante a seleção negativa de Linfócitos T expressam e apresentam TRAs aos timócitos. Essa capacidade de expressar virtualmente os antígenos próprios relativos a todos os tecidos e órgãos é denominada expressão gênica promíscua (ST-PIERRE et al., 2015; PASSOS et al., 2019)

A expressão gênica promíscua nas mTCEs é regulada majoritariamente pelo gene *Autoimmune regulator* (Aire), porém recentemente relatou-se a participação do gene *Fezf2* na expressão dos TRAs (TAKABA; TAKAYANAGI, 2017). Aire é um gene pleiotrópico, sua atuação já foi descrita em processos como a promoção do microambiente tímico, diferenciação e maturação de mTECs, desenvolvimento de linfócitos T reguladores, entre outras funções (PASSOS et al., 2019).

Além dos TRAs as mTCEs expressam diversos genes que codificam moléculas de adesão como as claudinas, integrinas e selectinas, para garantir que ocorra o “*crosstalk* tímico”, ou seja, a adesão mTEC-timócitos para apresentação dos TRAs (PEZZI et al., 2016).

Nosso grupo de pesquisa avaliou a atuação de Aire no controle da expressão dos genes de adesão e foi possível evidenciar que o silenciamento de Aire perturba a interação mTEC-timócitos (PEZZI et al., 2016; SPECK-HERNANDEZ et al., 2018).

Para que ocorra a apresentação dos TRAs, é necessário que aconteça uma interação recíproca entre as células mTEC e os timócitos, essa etapa é essencial para a maturação dos linfócitos T e para o estabelecimento da tolerância imunológica central, se houver comprometimento do processo de adesão, os TRAs expressos pelas mTECs não serão apresentados aos timócitos, e isso pode resultar no desenvolvimento de doenças autoimunes (PEZZI et al., 2016; SPECK-HERNANDEZ et al., 2018; PASSOS et al., 2019)).

Outra função de Aire nas mTECs estudada por nosso grupo, diz respeito a atuação desse gene na regulação da expressão de RNAs não codificadores de proteínas. Foi demonstrado que Aire participa do controle a expressão de microRNAs em células mTECs (MACEDO et al., 2013; PASSOS et al., 2018).

Nesse contexto, surgiu a hipótese de que Aire também é responsável pela controle da expressão de RNAs longos não-codificantes nas células mTECs, isso porque os lncRNAs assim como os mRNAs são transcritos pela RNA Pol II, uma vez

que Aire interage com a RNA POL II para regular a transcrição de TRAs, provavelmente ele também atua na expressão dos lncRNAs.

Para elucidar a hipótese levantada, adotou-se a estratégia de perda de função gênica, e por meio da técnica CRISPR foram geradas no laboratório de Imunogenética Molecular clones de células mTECs nocauteadas no gene Aire (SPECK-HERNANDEZ et al., 2018). Esses clones tiveram seu RNA sequenciado e análise do perfil do lncRNAs de expressão foi realizada com uso de ferramentas de bioinformática.

Com o advento da tecnologia de sequenciamento de nova geração (NGS, do inglês, *Next Generation Sequencing*), ferramentas de bioinformática foram desenvolvidas e atualmente permitem a análise de dados biológicos em larga escala, possibilitando o desenvolvimento de estudos mais completos e abrangentes.

RNA-Seq é uma técnica indicada para quantificação dos níveis de expressão de transcritos, é extremamente sensível e consegue detectar e quantificar transcritos com baixos níveis de expressão, como é o caso dos lncRNAs. Outra vantagem que o RNA-Seq oferece em relação a outras técnicas, como por exemplo o *microarray*, é que sua capacidade de detecção não se limita apenas aos transcritos que possuem sequência genômica correspondente nas sondas de hibridação (MAEKAWA et al., 2014).

A presente dissertação buscou identificar e caracterizar os lncRNAs diferencialmente expressos nas células mTECs na ausência de Aire. Para isso os dados brutos gerados no nosso laboratório passaram por um rigoroso controle de qualidade. Isso garante que os resultados apresentados sobre a caracterização dos *long non-codings* são altamente confiáveis. Todas as amostras apresentaram score de Phred superior a 30, valores que são considerados excelentes, o que indica que a probabilidade de erro na chamada de bases foi de 1/1000 (EWING; GREEN, 1998; LIAO; SATTEN; HU; 2017).

A análise dos lncRNAs foi realizada em duas situações diferentes, um grupo com amostras de células mTECs nocauteadas no gene Aire que não tiveram contato com timócitos e outro grupo formado por células mTECs Aire KO que foram submetidas ao ensaio de adesão mTEC-timócitos durante 12 horas. A análise dos lncRNAs antes da adesão indicou 136 lncRNAs diferencialmente expressos, sendo 72 hiperregulados e 64 hiporregulados. Após a adesão houve uma redução no número de lncRNAs com expressão diferencial, 33 no total, 26 hiperregulados e 7 hiporregulados.

A análise do diagrama de Venn mostrou que do total de lncRNAs diferencialmente expressos, 121 são exclusivos do grupo sem adesão, 18 exclusivos do grupo após a adesão e 15 lncRNAs são compartilhados entre os dois grupos. Entre os lncRNAs compartilhados notou-se que o padrão de expressão antes da adesão se manteve após o ensaio de adesão, o que indica que a interação com timócitos não influenciou no perfil de expressão deles.

Como esse é o primeiro estudo acerca da relação Aire-lncRNAs, a inferência que fizemos em relação a redução da expressão de lncRNAs após a adesão, é que Aire regula a expressão de lncRNAs nas mTECS, porém como essas células são especializadas na expressão de TRAs, durante a adesão mTEC-timócitos ocorre uma remodelação do perfil transcricional, onde há uma redução no número de lncRNAs transcritos e um aumento da expressão de TRAs e moléculas de adesão, para que aconteça o *crosstalk* tímico e a apresentação dos TRAs.

O único trabalho envolvendo lncRNAs e mTECs foi publicado em 2018 por Wei e colaboradores, onde os pesquisadores analisaram o efeito do hormônio beta estradiol na expressão de lncRNAs em mTECs, na presença do hormônio mais de 900 lncRNAs foram diferencialmente expressos, esses estavam envolvidos em processos biológicos como regulação imunológica, controle do metabolismo e expressão de citocinas.

No presente estudo diversas abordagens foram utilizadas para caracterizar os lncRNAs diferencialmente expressos. Em relação a classificação do GENCODE tanto antes quanto após a adesão os lncRNAs do tipo intergênicos foram os mais abundantes. Esse tipo é o mais bem descrito até o momento, e o papel dos lincRNAs na regulação da expressão gênica vem sendo melhor elucidado a cada dia. Outro tipo que chama a atenção são os lncRNAs do tipo TEC, esses são preditos, porém não validados. O estudo e validação desses lncRNAs podem resultar em descobertas importantes em relação ao papel deles em processos biológicos.

Os lncRNAs são moléculas que não possuem capacidade de codificar proteínas, na análise do potencial codificador a maioria dos lncRNAs foram classificados como noncoding como era esperado, no entanto oito lncRNAs foram classificados como coding. Isso se justifica porque embora os lncRNAs não possuam capacidade de codificar proteínas, alguns long non-codings possuem pequenas janelas abertas de leitura e conseguem codificar pequenos peptídeos ativos, mas

ainda sim são considerados lncRNAs como foi descrito por Morlando; Ballarino; Fatica em 2015.

Em relação enriquecimento funcional realizado, a plataforma KEGG indicou a participação do lncRNA lfi30 nas vias de processamento e apresentação de antígenos, isso vai de encontro com a principal função das células mTEC, que é a expressão e apresentação de TRAs durante o processo de seleção negativa dos tímocitos. Ressalta-se que no enriquecimento funcional todos os lncRNAs foram testados, porém a maioria não foi reconhecida, isso ocorre porque os bancos de dados com informações sobre lncRNAs em camundongos ainda estão em fase inicial de desenvolvido, o que dificulta a caracterização dessas moléculas.

Em relação a localização dos lncRNAs diferencialmente expressos, foi possível determinar a localização cromossômica e os genes codificadores mais próximos. É sabido que lncRNAs podem atuar em *cis*, ou seja, em genes próximos do seu *locus*, e em *trans*, quando regula genes em regiões distantes, assim essa caracterização é importante na determinação de possíveis mRNAs alvos dos lncRNAs.

Os lncRNAs diferencialmente expressos Fendrr, Gas5, Malat1, Neat1 e Xist atuam em *cis*, pois tem como alvos seus codificadores de proteínas mais próximos, tanto a montante quanto a jusante, além de atuarem em outros RNAm que estão localizados no mesmo cromossomo que o lncRNA se encontra, conforme é evidenciado pelo banco de dados lncRNA Disease v2.0 (BAO et al., 2019). Além disso, Yan em colaboradores em 2017 apontaram que o lncRNA Fendrr tem como alvo genes que estão em localizados em *loci* e/ou cromossomos distantes, demonstrando que Fendrr também possui ação regulatória em *trans*.

Embora não se tenha relatos até o momento de lncRNAs exclusivos de células mTECs, muitos lncRNAs já foram associados com a regulação do sistema imunológico. Dentre os lncRNAs que estavam diferencialmente expressos no presente trabalho, os lncRNAs Firre, Malat1, Morrbid e Neat1 são descritos como reguladores do sistema imune (HADJICHARALAMBOUS; LINDSAY; 2019).

O lncRNA Firre tem sua expressão dependente de NF- κ B. Foi demonstrado que tanto em macrófagos de humanos quanto de camundongos Firre atua na regulação de genes envolvidos em processos inflamatórios (LU et al., 2017). O lncRNA Malat1 em macrófagos também interage com NF- κ B para controlar a expressão de genes relacionados a inflamação como IL-6 e TNF- α ; nas células dendríticas Malat1 tem função de *enhancer* acentuando a expressão de IL-10, além

de servir como esponja para o microRNA mir-155, isso resulta no aumento da capacidade tolerogênica das células dendríticas (CUI et al., 2019).

O *long non coding* Morrbid atua no desenvolvimento de eosinófilos, monócitos e neutrófilos, em camundongos e humanos, e é capaz de interagir com citocinas envolvidas na sobrevivência dessas células como IL-3 e IL-5. Também foi visto que em infecções virais Morrbid promove a regulação de Linfócitos T CD8+ através da interação com o fator apoptótico Bcl2l11, controlando a crescimento, sobrevivência e função efetora dessas células (KOTZIN et al., 2019). Já o lncRNA Neat1 participa da formação de complexos importantes na regulação resposta imune inata e na resposta imune adaptativa atua no processo de diferenciação de células Th2 (MORCHIKH et al. 2017; HUR; KIM; KIM, 2019).

O envolvimento de lncRNAs em doenças autoimunes também já foi descrito na literatura (XU et al., 2018). Tendo em vista essa problemática, destacamos alguns *long noncodings* diferencialmente expressos na ausência de Aire, como Gas5, Neat1 e o Xist.

Gas5 interage com o receptor de glicocorticoides, bloqueando a via de sinalização desse hormônio, o que pode culminar no desenvolvimento de doenças autoimunes (WANG; CHANG, 2012; WU et al., 2017). No presente estudo Gas5 encontra-se hipoexpresso. Mayama, Mar e Kino em 2016, demonstraram que em camundongos com lúpus eritematoso sistêmico (LES) e/ou artrite reumatoide o lncRNA Gas5 está hiporregulado.

A hiperexpressão de Neat1 foi relatada em células mononucleares do sangue periférico de pacientes com LES, essa expressão aumentada tem correlação positiva com o *Systemic Lupus Erythematosus Disease Activity Index* – SLEDAI, índice utilizado para avaliar a atividade da doença, assim quanto mais altos são os níveis de expressão de NEAT1 mais grave são os sinais e sintomas apresentados (ZHANG et al., 2016; HUR; KIM; KIM, 2019). No presente estudo, no entanto, nas células mTECs nocauteadas no gene Aire, o lncRNA Neat1 apresentou-se hipoexpresso.

O lncRNA Xist é notável pela sua capacidade de atuação no mecanismo de compensação de dose em mamíferos. Além disso, Huang e colaboradores em 2014, apontaram que camundongos com glomerulonefrite membranosa, uma patologia renal geralmente associada a doenças autoimunes, apresentam altos níveis de Xist na urina, sendo esse um possível biomarcador para diagnóstico dessa patologia.

Os níveis de expressão dos lncRNAs Dgkeos, Fendrr e Peg13 foram validados por meio de RT- qPCR. Os resultados obtidos corroboraram com os resultados da análise bioinformática, ratificando a qualidade e a precisão das análises *in silico*.

Dentre esses três lncRNAs, Fendrr é o que está mais bem caracterizado. Foi demonstrado que Fendrr atua no desenvolvimento do coração de camundongos e que sua hipoexpressão leva à falência cardíaca (GROTE et al., 2014). No câncer gástrico e no carcinoma renal, relatou-se que a hipoexpressão de Fendrr está associada a um pior prognóstico, pois promove o aumento da migração, proliferação e invasão de células cancerosas (XU et al., 2014; HE et al., 2019). No presente trabalho, Fendrr está hiporregulado antes e depois da adesão, o que indica que ele é um lncRNA Aire dependente. Até o momento não há estudos que associem Fendrr a regulação do sistema imunológico ou a doenças autoimunes.

Em suma, foi possível identificar e caracterizar lncRNAs diferencialmente expressos na ausência de Aire. Nossos achados mostram que Aire atua no controle da expressão de lncRNAs em células mTEC, e que diversos *long non-codings* diferencialmente expressos nessas células estão relacionados com a regulação do sistema imunológico e/ou com a patogênese de doenças autoimunes. Em relação aos lncRNAs preditos, ressalta-se que o estudo e a validação dos mesmos pode resultar em descobertas importantes para o entendimento do papel desses lncRNAs nas mTECs.

CONCLUSÃO

8. CONCLUSÃO

Foi possível determinar o perfil de expressão, identificar e caracterizar os lncRNAs diferencialmente expressos nas mTECs 3.10E6 (Aire \pm) antes e após a adesão com timócitos. A partir dos resultados obtidos nesse trabalho, infere-se que:

- Aire atua no controle expressão de lncRNAs em células mTEC;
- A adesão celular mTECs-Timócitos promove a remodelação do perfil transcricional das células mTEC nocauteadas no gene Aire, e isso reduz a número de lncRNAs diferencialmente expressos.
- Dentre os lncRNAs diferencialmente expressos nas mTECs que já estão validados e caracterizados, vários deles estão relacionados com a regulação do sistema imunológico e/ou com a patogênese de doenças autoimunes.
- Muitos lncRNAs são apenas preditos, e ainda não estão validados e caracterizados experimentalmente.

REFERÊNCIAS BIBLIOGRÁFICAS

9. REFERÊNCIAS BIBLIOGRÁFICAS

AALTONEN, J.; BJÖRSES, P.; PERHEENTUPA, J.; *et al.* An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. **Nature Genetics**, v. 17, n. 4, p. 399–403, 1997.

ABRAMSON, J.; GIRAUD, M.; BENOIST, C.; MATHIS, D. Aire's Partners in the Molecular Control of Immunological Tolerance. **Cell**, v. 140, n. 1, p. 123–135, 2010.

ALMO, M.M.; SOUSA, I.G.; MARANHÃO, A.Q.; BRIGIDO, M. M. The role of long noncoding RNAs in human T CD3+ cells. **Journal of Immunological Sciences**, v. 2, n. 1, p. 32–36, 2018.

ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq - A Python framework to work with high-throughput sequencing data. **Bioinformatics**, v. 31, n. 2, p. 166-169, 2014.

ANDERSON, M. S.; SU, M. A. AIRE expands: New roles in immune tolerance and beyond. **Nature Reviews Immunology**, v. 16, n. 4, p. 247, 2016.

ANDREWS, S. FastQC A Quality Control tool for High Throughput Sequence Data. (2010). Disponível em: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

ATIANAND, M. K.; CAFFREY, D. R.; FITZGERALD, K. A. Immunobiology of Long Noncoding RNAs. **Annual Review of Immunology**, v. 35, n. January, p. 177–198, 2017.

AUNE, T. M.; SPURLOCK, C. F. Long non-coding RNAs in innate and adaptive immunity. **Virus Research**, v. 212, p. 146–160, 2016.

BANSAL, K; YOSHIDA, H; BENOIST, C; MATHIS, D. The transcriptional regulator Aire binds and activates super-enhancers. **Nature Immunology**, 18:263–73, 2017.

BAO, Z.; YANG, Z.; HUANG, Z.; ZHOU, Y.; CUI, Q.; DONG, D. LncRNADisease 2.0: An updated database of long non-coding RNA-associated diseases. **Nucleic Acids Research**, v. 47, n. D1, p. D1034–D1037, 2019.

BEADLE, G. W.; TATUM, E. L. Genic Control of Biochemical Reactions in *Neurospora*. **The American Naturalist**, v. 79, n. 783, p. 304–317, 1941.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics** (Oxford, England), p. 1–7, 2014.

BREDENKAMP, N.; ULYANCHENKO, S.; O'NEILL, K. E.; MANLEY, N. R.; VAIDYA, H. J.; BLACKBURN, C. C. An organized and functional thymus generated from Foxn1-reprogrammed fibroblast. **Nat. Cell Biol.**v.16(9), 902-908, 2014.

BRUSERUD, Ø.; OFTEDAL, B. E.; WOLFF, A. B.; *et al.* AIRE-mutations and autoimmune disease. **Current Opinion in Immunology**, v. 43, p. 8–15, 2016.

CHEN, L., XIAO, S., MANLEY, N. R. Foxn1 is required to maintain the postnatal thymic microenvironment in a dosage-sensitive manner. *Blood*. v.15:113(3), p.567-574, 2009.

CHENG, M. H.; SHUM, A. K.; ANDERSON, M. S. What's new in the Aire? **Trends in Immunology**, v. 28, n. 7, p. 321–327, 2007.

CUI, H.; BANERJEE, S.; GUO, S.; XIE, N.; GE, J.; JIANG, D.; ZÖRNIG, M.; THANNICKAL, V.J.; LIU, G. Long noncoding RNA Malat1 regulates differential activation of macrophages and response to lung injury. **JCI Insight**, 4, 176, 2019.

DERBINSKI, J. *et al.* Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. **Journal of Immunology**, v. 196, n. 7, p. 2915–2922, 2016.

DERRIEN, T.; JOHNSON, R.; BUSSOTTI, G.; TANZER, A.; DJEBALI, S.; TILGNER, H.; GUERNEC, G.; MARTIN, D.; MERKEL, A.; KNOWLES, D. G.; LAGARDE, J.; VEERAVALLI, L.; RUAN, X.; RUAN, Y.; LASSMANN, T.; CARNINCI, P.; BROWN, J. B.; LIPOVICH, L.; GONZALEZ, J. M.; THOMAS, M.; DAVIS, C. A.; SHIEKHATTAR, R.; GINGERAS, T. R.; HUBBARD, T. J.; NOTREDAME, C.; HARROW, J.; GUIGÓ, R. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. **Genome Research**, v. 22, n. 9, p. 1775–1789, 2012.

DJEBALI S., DAVIS C. A., MERKEL A., DOBIN A., LASSMANN T., MORTAZAVI A., TANZER A., LAGARDE J., LIN W., SCHLESINGER F., XUE C., MARINOV G. K., KHATUN J., WILLIAMS B. A., ZALESKI C., ROZOWSKY J., RODER M., KOKOCINSKI F., ABDELHAMID R. F., ANTOSHECHKIN I., BAER M. T., BAR N. S., BATUT P., BELL K., BELL I., CHAKRABORTTY S., CHEN X., CHRAST J., CURADO J., DERRIEN T., DRENKOW J., DUMAIS E., DUMAIS J., DUTTAGUPTA R., FALCONNET E., FASTUCA M., FEJES-TOTH K., FERREIRA P., FOISSAC S., FULLWOOD M. J., GAO H., GONZALEZ D., GORDON A., GUNAWARDENA H., HOWALD C., JHA S., JOHNSON R., KAPRANOV P., KING B., KINGSWOOD C., LUO O. J., PARK E., PERSAUD K., PREALL J. B., RIBECA P., RISK B., ROBYR D., SAMMETH M., SCHAFFER L., SEE L. H., SHAHAB A., SKANCKE J., SUZUKI A. M., TAKAHASHI H., TILGNER H., TROUT D., WALTERS N., WANG H., WROBEL J., YU Y., RUAN X., HAYASHIZAKI Y., HARROW J., GERSTEIN M., HUBBARD T., REYMOND A., ANTONARAKIS S. E., HANNON G., GIDDINGS M. C., RUAN Y., WOLD B., CARNINCI P., GUIGO R., GINGERAS T. R., 2012. Landscape of transcription in human cells. **Nature**. v. 489, n° 7414, p. 101-108.

DOBIN, A.; DAVIS, C. A.; SCHLESINGER, F.; DRENKOW, J.; ZALESKI, C.; JHA, S. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**, 29(1), 15–21, 2013. <https://doi.org/10.1093/bioinformatics/bts635>

DURINCK, S.; SPELLMAN, P.; BIRNEY, E.; HUBER, W. “Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt.” **Nature Protocols**, 4, 1184–1191, 2009.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. **Genome Research**, v. 8, p. 186-194, 1998.

FERNANDES, J. et al. Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. **Non-Coding RNA**, v. 5, n. 1, p. 17, 2019.

FERRE, E. M. N.; ROSE, S. R.; ROSENZWEIG, S. D.; BURBELO, P. D.; ROMITO, K. R.; NIEMELA, J. E.; ROSEN, L. B.; BREAK, T. J.; GU, W.; HUNSBERGER, S.; BROWNE, S. K.; HSU, A. P.; RAMPERTAAP, S.; SWAMYDAS, M.; COLLAR, A. L.; KONG, H. H.; LEE, C.-C. R.; CHASCSA, D.; SIMCOX, T.; PHAM, A.; BONDICI, A.;

NATARAJAN, M.; MONSALE, J.; KLEINER, D. E.; QUEZADO, M.; ALEVIZOS, I.; MOUTSOPOULOS, N. M.; YOCKEY, L.; FREIN, C.; SOLDATOS, A.; CALVO, K. R.; ADJEMIAN, J.; SIMILUK, M. N.; LANG, D. M.; STONE, K. D.; UZEL, G.; KOPP, J. B.; BISHOP, R. J.; HOLLAND, S. M.; OLIVIER, K. N.; FLEISHER, T. A.; HELLER, T.; WINER, K. K.; LIONAKIS, M. S. Redefined clinical features and diagnostic criteria in autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy. **JCI Insight**, v. 1, n. 13, p. 1–18, 2016.

GEENEN, V & SAVINO, W. History of the Thymus: From a Vestigial Organ to the Programming of Immunological Self-Tolerance . *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 1, p. 1-18.

GIRAUD, M. & PETERSON, P. The Autoimmune Regulator (*AIRE*) Gene, the Master Activator of Self-Antigen Expression in the Thymus.*In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 7, p. 169-190.

GROTE, P; WITTLER, L; WÄHRISCH, S; HENDRIX, D; BEISAW, A; MACURA, K; BLÄSS, G; KELLIS, M; WERBER, M; HERRMANN, B. G. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. **Development Cell**, v. 24, n. 2, p. 206–214, 2014.

GUIL, S.; ESTELLER, M. *Cis*-acting noncoding RNAs: friends and foes. **Nature Structural & Molecular Biology**, v. 19, p. 1068–1075, 2012.

HADJICHARALAMBOUS, M. R.; LINDSAY, M. Long Non-Coding RNAs and the Innate Immune Response. **Non-Coding RNA**, v. 5, n. 2, p. 34, 2019.

HE, W.; ZHONG, G.; WANG, P.; JIANG, C.; JIANG, N.; HUANG, J. Downregulation of long noncoding RNA FENDRR predicts poor prognosis in renal cell carcinoma. **Oncology Letters**, v. 17, n. 1, p. 103–112, 2019.

HERMANS-BEIJNSBERGER, S.; VAN BILSEN, M.; SCHROEN, B. Long non-coding RNAs in the failing heart and vasculature. **Non-coding RNA Research**, v. 3, n. 3, p. 118–130, 2018.

HU, W.; ALVAREZ-DOMINGUEZ, J. R.; LODISH, H. F. Regulation of mammalian cell differentiation by long non-coding RNAs. **EMBO reports**, v. 13, n. 11, p. 971–983, 2012.

HUANG, Y. S.; HSIEH, H. Y.; SHIH, H. M.; SYTWU, H. K.; WU, C. C. Urinary Xist is a potential biomarker for membranous nephropathy. **Biochemical and Biophysical Research Communications**, v. 452, n. 3, p. 415–421, 2014.

HUR, K.; KIM, S. H.; KIM, J. M. Potential implications of long noncoding RNAs in autoimmune diseases. **Immune Network**, v. 19, n. 1, p. 1–16, 2019.

HUSEBYE, E. S.; PERHEENTUPA, J.; RAUTEMAA, R.; KÄMPE, O. Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. **Journal of Internal Medicine**, v. 265, n. 5, p. 514–529, 2009.

HUSEBYE, E.S; ANDERSON M.S; KÄMPE O. Autoimmune Polyendocrine Syndromes. **New England Journal of Medicine**. ;378(12):1132-1141. 2018 Mar

HUTTENHOFER, A., SCHATNER, P. & POLACEK, N. Non-coding RNAs: hope or hype? **Trends in genetics** : TIG, v. 21, n° 5, p. 289–97., 2005.

IRLA, M. Thymic Crosstalk: An Overview of the Complex Cellular Interactions That Control the Establishment of T-Cell Tolerance. *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 6, p. 149-168.

KANG, Y. J.; YANG, D. C.; KONG, L.; HOU, M.; MENG, Y. Q.; WEI, L.; GAO, G. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. **Nucleic Acids Research 45(Web Server issue): W12–W16, 2017.**

KENIRY, A.; OXLEY, D.; MONNIER, P.; KYBA, M.; DANDOLO, L.; SMITS, G.; REIK, KLEIN, L.; KYEWSKI, B.; ALLEN, P. M.; HOGQUIST, K. A. Positive and negative selection of the T cell repertoire: what thymocytes see and don't see. **Nature Reviews Immunology**, v. 14, n. 6, p. 377–391, 2014.

KOTZIN, J. J.; ISEKA, F.; WRIGHT, J.; BASAVAPPA, M. G.; CLARK, M. L.; ALI, M. A.; ABDEL-HAKEEM, M. S.; ROBERTSON, T. F.; MOWEL, W. K.; JOANNAS, L.; NEAL, V. D.; SPENCER, S. P.; SYRETT, C. M.; ANGUERA, M. C.; WILLIAMS, A.;

WHERRY, E. J.; HENAO-MEJIA, J. The long noncoding RNA Morrbid regulates CD8 T cells in response to viral infection. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 24, p. 11916–11925, 2019.

KYEWSKI, B.; DERBINSKI, J. Self-representation in the thymus: An extended view. **Nature Reviews Immunology**, v. 4, n. 9, p. 688, 2004.

LI, J.; MA, W.; ZENG, P.; WANG, J.; GENG, B.; YANG, J.; CUI, Q. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. **Briefings in Bioinformatics** 16(5) 2015.

LI, X.; ZHENG, Y.; SALK, T.; JOLLA, L. Regulatory T cell identity: formation and maintenance Xudong. **Trends Immunol. Author**, v. 36, n. 6, p. 344–353, 2016.

LINHARES-LACERDA, L.; RIBEIRO-ALVES, M.; NOGUEIRA, A. C.; MENDES-DA-CRUZ, D. A.; MAGALHÃES, D. A.; DARDENNE, N.; PASSOS, G. A.; SAVINO, W. RNA interferencemediated knockdown of CD49e ($\alpha 5$ integrin chain) in human thymic epithelial cells modulates the expression of multiple genes and decreases thymocyte adhesion. **BMC Genomics**. v.22:11, 2010.

LU, Y.; LIU, X.; XIE, M.; LIU, M.; YE, M.; LI, M.; CHEN, X.-M.; LI, X.; ZHOU, R. The NF- κ B-Responsive Long Noncoding RNA FIRRE Regulates Posttranscriptional Regulation of Inflammatory Gene Expression through Interacting with hnRNPU. **J. Immunol**, 199, 3571–3582, 2017.

MA, L.; BAJIC, V. B.; ZHANG, Z. On the classification of long non-coding RNAs. **RNA Biology**, v. 10, n. 6, p. 924–933, 2013.

MACEDO, C.; EVANGELISTA, A. F.; MARQUES, M. M.; OCTACÍLIO-SILVA, S.; DONADI, E. A.; SAKAMOTO-HOJO, E. T.; PASSOS, G. A. Autoimmune regulator (Aire) controls the expression of microRNAs in medullary thymic epithelial cells. **Immunobiology**, v. 218, n. 4, p. 554–560, 2013.

MAEKAWA S., SUZUKI A., SUGANO S., SUZUKI Y. RNA Sequencing: From Sample Preparation to Analysis. In: Miyamoto-Sato E., Ohashi H., Sasaki H., Nishikawa J.,

Yanagawa H. (eds) Transcription Factor Regulatory Networks. **Methods in Molecular Biology (Methods and Protocols)**. Humana Press, New York, NY, 2014, vol 1164.

MATHIS, D.; BENOIST, C. A decade of AIRE. **Nature Reviews Immunology**, v. 7, n. August, p. 645–650, 2007.

MATHY, N. W.; CHEN, X.-M. LncRNAs and Their Transcriptional Control of Inflammatory Responses. **Journal of Biological Chemistry**, p. jbc.R116.760884, 2017.

MAYAMA, T.; MARR, A. K.; KINO, T. Differential Expression of Glucocorticoid Receptor Noncoding RNA Repressor Gas5 in Autoimmune and Inflammatory Diseases. **Hormone and Metabolic Research**, v. 48, n. 8, p. 550–557, 2016.

MENDES-DA-CRUZ, D. A. et al. Intrathymic Cell Migration: Implications in Thymocyte Development and T-Cell Repertoire Formation . *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 5, p. 117-148.

MORCHIKH, M.; CRIBIER, A.; RA_EL, R.; AMRAOUI, S.; CAU, J.; SEVERAC, D.; DUBOIS, E.; SCHWARTZ, O.; BENNASSER, Y.; BENKIRANE, M. HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response. **Mol. Cell**, 67, 387–399, 2017.

MUNOZ, J. J. & ZAPATA, A. G. Thymus Ontogeny and Development . *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 2, p. 19-34.

NAGAMINE, K.; PETERSON, P.; SCOTT, H. S.; *et al.* Positional cloning of the APECED gene. **Nature Genetics**, v. 17, n. 4, p. 393–398, 1997.

NISHINO M, ASHIKU SK, KOCHER ON, THURER RL, BOISELLE MP & HATABU H. The Thymus: a comprehensive review. (2006). **RadioGraphics**, 26, 335-348.

OFTEDAL, B. E. et al. Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. **Immunity**, v. 42, n. 6, p. 1185–1196, 2015.

OHIGASHI, I.; KOZAI, M.; TAKAHAMA, Y. Development and developmental potential of cortical thymic epithelial cells. **Immunological Reviews**, v. 271, n. 1, p. 10–22, 2016.

OLIVEIRA, E. H., MACEDO, C., COLLARES, C. V., FREITAS, A. C., DONATE, P. B., SAKAMOTO-HOJO, E. T., PASSOS, G. A. Aire Downregulation Is Associated with Changes in the Posttranscriptional Control of Peripheral Tissue Antigens in Medullary Thymic Epithelial Cells. **Frontiers in Immunology**, v. 7, n. 526. 2016.

PASSOS, G. A. et al. The Thymus as a Mirror of the Body's Gene Expression. *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 9, p. 215-234.

PASSOS, G. A.; MENDES-DA-CRUZ, D. A.; OLIVEIRA, E. H. The thymic orchestration involving Aire, miRNAs and cell-cell interactions during the induction of central tolerance. **Frontiers in Immunology**, v. 6, n. JUN, p. 1–7, 2015.

PASSOS, G. A.; SPECK-HERNANDEZ, C. A.; ASSIS, A. F.; MENDES-DA-CRUZ, D. A. Update on *Aire* and thymic negative selection. **Immunology**, p. 10–20, 2018.

PERNIOLA, R. Twenty years of AIRE. **Frontiers in Immunology**, v. 9, n. FEB, 2018.

PERNIOLA, R.; MUSCO, G. The biophysical and biochemical properties of the autoimmune regulator (AIRE) protein. **Biochimica et Biophysica Acta - Molecular Basis of Disease**, v. 1842, n. 2, p. 326–337, 2014.

PETERSON, P.; PITKÄNEN, J.; SILLANPÄÄ, N.; *et al.* Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED): A model disease to study molecular aspects of endocrine autoimmunity. **Clinical and Experimental Immunology**, v. 135, n. 3, p. 348–357, 2004.

PEZZI, N.; ASSIS, A. F.; COTRIM-SOUSA, L. C.; LOPES, G. S.; MOSELLA, M. S.; LIMA, D. S.; BOMBONATO-PRADO, K. F.; PASSOS, G. A. Aire knockdown in medullary thymic epithelial cells affects Aire protein, deregulates cell adhesion genes and decreases thymocyte interaction. **Molecular Immunology**, v. 77, p. 157–173, 2016.

PONTIER, D. B.; GRIBNAU, J. Xist regulation and function eXplored. **Human Genetics**, v. 130, n. 2, p. 223–236, 2011.

QUAN, M., CHEN, J. & ZHANG, D. Exploring the Secrets of Long Noncoding RNAs. **International Journal of Molecular Sciences**. v. 16, n° 3, p. 5467–5496, 2015.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 2009.

RUEHL-FEHLERT, C., KITTEL, B., MORAWIETZ, G., DESLEX, P., KEENAN, C., MAHRT, C. R; et al. NACAD Group. Revised guides for organ sampling and trimming in rats and mice--part 1. **Experimental and Toxicologic Pathology : Official Journal of the Gesellschaft Fur Toxikologische Pathologie**, 55(2–3), 91–106, 2003.

SAFIEDDINE, N., & KESHAVJEE, S. Anatomy of the Thymus Gland. **Thoracic Surgery Clinics**, 21(2), 191–195, 2011.

SALEHI, S.; TAHERI, M. N.; AZARPIRA, N.; ZARE, A.; BEHZAD-BEHBAHANI, A. State of the art technologies to explore long non-coding RNAs in cancer. **Journal of Cellular and Molecular Medicine**, v. 21, n. 12, p. 3120–3140, 2017.

SANSOM, S. N.; PONTING, C. P.; NUSSPAUMER, G.; *et al.* Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. **Genome Research**, v. 24, n. 12, p. 1918–1931, 2014.

SATTLER, S. The Role of the Immune System Beyond the Fight Against Infection. In: Sattler, S.; Kennedy-Lydon, T. (Eds). *The Immunology of Cardiovascular Homeostasis and Pathology*. **Advances in Experimental Medicine and Biology**. Springer, Cham 2017. vol 1003. Cap. 1, p. 3-14.

SHANNON, P. et al. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. **Genome Research**, v. 13, n. 11, p. 2498-2504, 2003.

SIENA, A. D. D; BARROS-ICHIHARA, I. RNAs Longos Não Codificantes. *In*: PEREIRA, T. C. (Ed.). **Introdução ao universo dos Non-Coding RNAs**. Brasil: SBG, 2017. Cap. 2, seção 16, p. 95-132.

SKOGBERG, G.; TELEMO, E.; EKWALL, O. Exosomes in the thymus: Antigen transfer and vesicles. **Frontiers in Immunology**, v. 6, n. JUL, p. 1–8, 2015.

SPECK-HERNANDEZ, C. A.; ASSIS, A. F.; FELICIO R. F.; COTRIM-SOUSA, L.; PEZZI N.; LOPES, G. S.; BOMBONATO-PRADO, K. F.; GIULIATTI, S. AND PASSOS, G. A. Aire Disruption Influences the Medullary Thymic Epithelial Cell Transcriptome and Interaction With Thymocytes. **Frontiers in Immunology**, v.9, n.964. 2018.

ST-PIERRE, C.; TROFIMOV, A.; BROCHU, S.; LEMIEUX, S.; PERREAULT, C. Differential Features of AIRE-Induced and AIRE-Independent Promiscuous Gene Expression in Thymic Epithelial Cells. **The Journal of Immunology**, v. 195, n. 2, p. 498–506, 2015.

TAKABA, H.; MORISHITA, Y.; TOMOFUJI, Y.; DANKS, L.; NITTA, T.; KOMATSU, N.; KODAMA, T.; TAKAYANAGI, H. Fezf2 Orchestrates a Thymic Program of Self-Antigen Expression for Immune Tolerance. **Cell**, v. 163, n. 4, p. 975–987, 2015.

TAKABA, H.; TAKAYANAGI, H. The Mechanisms of T Cell Selection in the Thymus. **Trends in Immunology**, v. 38, n. 11, p. 805–816, 2017.

TAKAHAMA, Y.; NITTA, T.; MAT, A.; NITTA, S. Seminars in Immunology Role of thymic cortex-specific self-peptides in positive selection of T cells. **Seminars in Immunology**, v. 22, n. 5, p. 287–293, 2010.

TIAN, X. *et al.* Long non-coding RNAs in the regulation of myeloid cells. **Journal of Hematology & Oncology**, v. 9, n. 1, p. 99, 2016.

KENIRY, A.; OXLEY, D.; MONNIER, P.; KYBA, M.; DANDOLO, L.; SMITS, G.; REIK, W. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. **Nat. Cell Biol.** v. 14, n° 7, p. 659–665, 2012.

WANG, K. C.; CHANG, H. Y. Molecular mechanisms of long noncoding RNAs. **Mol Cell**, v. 43, n. 6, p. 904–914, 2012.

WANG, P. *et al.* The STAT3-binding long noncoding RNA Inc-DC controls human dendritic cell differentiation. **Science (New York, N.Y.)**, v. 344, n. April, p. 310–3, 2014.

WEI, C.; GUO, D.; LI, Y.; ZHANG, K.; LIANG, G.; LI, Y.; MA, Y.; LIU, J.; LI, Y. Profiling analysis of 17 β -estradiol-regulated lncRNAs in mouse thymic epithelial cells. **Physiological Genomics**, v. 50, n. 8, p. 553–562, 2018.

WOLFF, A. S. B; OFTEDAL, B. E. Aire Mutations and Autoimmune Diseases . *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 8, p. 191-214.

WU GC, LI J, LENG RX, LI XP, LI XM, WANG DG, PAN HF, YE DQ. Identification of long non-coding RNAs Gas5, linc0597 and lnc-DC in plasma as novel biomarkers for systemic lupus erythematosus. **Oncotarget**, 8:23650-23663, 2017.

XU, F.; JIN, L.; JIN, Y.; NIE, Z.; ZHENG, H. Review Article Long noncoding RNAs in autoimmune diseases. **J Biomed Mater Res Part A.**, p. 1–8, 2018.

XU, T. P.; HUANG, M. De; XIA, R.; LIU, X. X.; SUN, M.; YIN, L.; CHEN, W. M.; HAN, L.; ZHANG, E. B.; KONG, R.; DE, W.; SHU, Y. Q. Decreased expression of the long non-coding RNA FENDRR is associated with poor prognosis in gastric cancer and FENDRR regulates gastric cancer cell metastasis by affecting fibronectin1 expression. **Journal of Hematology and Oncology**, v. 7, n. 1, p. 1–15, 2014.

YAN, P.; LUO, S.; LU, J. Y.; SHEN, X. Cis- and trans-acting lncRNAs in pluripotency and reprogramming. **Current Opinion in Genetics and Development**, v. 46, p. 170–178, 2017.

YANG, Y.; WEN, L.; ZHU, H. Unveiling the hidden function of long non-coding RNA by identifying its major partner-protein. **Cell & Bioscience**, v. 5, n. 1, p. 1–10, 2015.

YOGANATHAN, K. et al. T-Cell Development: From T-Lineage Specification to Intrathymic Maturation. *In*: PASSOS, G. A. (Ed.). **Thymus Transcriptome and Cell Biology**. Suíça: Springer Nature, 2019. Cap. 4, p. 67-116.

ZHANG F, WU L, QIAN J, QU B, XIA S, LA T, WU Y, MA J, ZENG J, GUO Q, ET AL. Identification of the long noncoding RNA NEAT1 as a novel inflammatory regulator acting through MAPK pathway in human lupus. **J Autoimmun**, 75:96-104, 2016.

APÊNDICE

10. APÊNDICE

10.1 Apêndice 1 – Lista completa dos lncRNAs hiperregulados antes do ensaio de adesão

ID_Ensembl	Nome_Ensembl	Nome_Gencode	Tipo	Chr	log(2)FC
ENSMUSG00000104811.1	Gm33093	RP23-307D8.9	lincRNA	chr5	3.05697094892126
ENSMUSG00000097289.1	2010300F17Rik	2010300F17Rik	lincRNA	chr11	2.80000087711663
ENSMUSG00000113185.1	Gm47127	RP24-447D15.3	lincRNA	chr13	2.76977635492375
ENSMUSG00000109021.1	Gm44812	RP23-449M8.5	lincRNA	chr7	1.98691880403124
ENSMUSG00000104554.1	Gm4610	RP23-71G16.6	antisense	chr3	1.86186727758227
ENSMUSG00000097636.7	Mirt1	Mirt1	lincRNA	chr19	1.82963943266738
ENSMUSG00000104982.1	Gm32554	RP23-307D8.7	lincRNA	chr5	1.7898999306723
ENSMUSG00000106121.1	Gm42679	RP23-63A19.3	sense_intronic	chr3	1.72570384462814
ENSMUSG00000087107.8	AI662270	RP23-381B19.2	processed_transcript	chr11	1.687331514335
ENSMUSG00000106652.1	Gm43358	RP23-307D8.3	lincRNA	chr5	1.52897408641536
ENSMUSG00000060411.3	Npn2	Npn2	antisense	chr6	1.49444727073916
ENSMUSG00000105891.4	A230001M10Rik	RP23-63A19.1	lincRNA	chr3	1.47903457751734
ENSMUSG00000100201.1	Gm29093	RP24-327D15.7	sense_intronic	chr10	1.34691629527569
ENSMUSG00000087203.7	Gm13986	RP23-75E19.1	processed_transcript	chr2	1.3455848737091
ENSMUSG00000107198.1	Gm19619	RP23-258C6.2	lincRNA	chr5	1.33380581405073
ENSMUSG00000097252.1	Gm6634	Gm6634	lincRNA	chr3	1.32908462093271
ENSMUSG00000114053.1	Gm35330	RP24-330K6.1	lincRNA	chr13	1.29068402091438
ENSMUSG00000100954.4	Gm10138	RP24-167M10.2	antisense	chr1	1.25867515179835
ENSMUSG00000086834.1	Gm14547	RP23-104H13.6	antisense	chr6	1.24983772616283
ENSMUSG00000103593.1	Gm37352	RP24-64D24.3	TEC	chr8	1.24566674056038
ENSMUSG00000107709.1	9430018G01Rik	RP23-20L3.2	antisense	chr6	1.2421382941668
ENSMUSG00000113049.1	Gm48544	RP24-574N19.4	lincRNA	chr13	1.23663742802406
ENSMUSG00000085675.1	Gm12022	RP23-436D9.1	lincRNA	chr11	1.20408910062542
ENSMUSG00000097296.1	Gm26532	RP24-64D24.4	lincRNA	chr8	1.18258021768828
ENSMUSG00000093677.1	Gm20712	RP23-102L5.1	sense_overlapping	chr15	1.15976863698207

ENSMUSG00000102278.1	Gm37145	RP23-157J4.3	TEC	chr1	1.13572329348447
ENSMUSG00000043833.6	2900005J15Rik	AC116151.2	processed_transcript	chr5	1.1252602588244
ENSMUSG00000113909.1	Gm36377	RP23-376D8.4	lincRNA	chr13	1.11695920002197
ENSMUSG00000110981.1	Gm35940	RP24-362F17.1	processed_transcript	chr9	1.0605019780447
ENSMUSG00000097316.2	Gm10516	CAA01201205.3	antisense	chr1	1.0275984874863
ENSMUSG00000097727.7	F630040K05Rik	F630040K05Rik	lincRNA	chr17	1.01905270244104
ENSMUSG00000084799.7	Ino80dos	Ino80dos	lincRNA	chr1	1.00137344475405
ENSMUSG00000104814.1	Gm42979	RP24-161C20.4	TEC	chr5	0.994327403817658
ENSMUSG00000096965.7	3300005D01Rik	RP23-251I6.1	lincRNA	chr17	0.985490651997845
ENSMUSG00000086013.1	Gm15706	RP24-359O2.3	lincRNA	chr6	0.976647915480917
ENSMUSG00000097495.1	Gm26651	RP23-356F21.2	lincRNA	chr13	0.963164522768757
ENSMUSG00000109145.1	Gm20744	RP23-341H2.2	lincRNA	chr7	0.924680880828776
ENSMUSG00000106734.3	Gm20559	RP24-328P2.5	processed_transcript	chr6	0.886331427449879
ENSMUSG00000107458.1	Gm44040	RP24-266P6.8	antisense	chr6	0.882791881019041
ENSMUSG00000097180.5	2700038G22Rik	RP23-354E8.2	lincRNA	chr5	0.871829451602272
ENSMUSG00000085030.1	2810455O05Rik	AC147266.3	lincRNA	chr8	0.861067666251148
ENSMUSG00000113757.1	Gm47507	RP23-28J6.5	lincRNA	chr13	0.831402274651288
ENSMUSG00000078956.3	Gm14221	RP23-152H17.2	lincRNA	chr2	0.801766436133876
ENSMUSG00000087574.1	C030037D09Rik	RP23-393B19.1	lincRNA	chr11	0.798922632237689
ENSMUSG00000093587.3	Gm20554	RP23-289E11.2	lincRNA	chr13	0.775659976581808
ENSMUSG00000113128.1	Gm47813	RP24-383P22.2	TEC	chr13	0.722455486426885
ENSMUSG00000091422.8	Gm6455	RP24-113C21.1	lincRNA	chr5	0.698278682102634
ENSMUSG00000087259.7	2610035D17Rik	RP23-468D17.1	lincRNA	chr11	0.691379614178706
ENSMUSG00000097519.2	4930558J18Rik	RP24-77I21.2	lincRNA	chr1	0.690016938118697
ENSMUSG00000074813.13	Morrbid	RP23-249P4.1	lincRNA	chr2	0.677112100650472
ENSMUSG00000073415.4	Gm10501	RP23-349B4.5	antisense	chr17	0.652671578083652
ENSMUSG00000087679.2	Tmem250-ps	RP23-325E4.5	lincRNA	chr2	0.652578300114685
ENSMUSG00000032846.10	Zswim6	Zswim6	processed_transcript	chr13	0.632720844443469
ENSMUSG00000100658.1	F730311O21Rik	RP23-217K18.4	lincRNA	chr1	0.625194963977381
ENSMUSG00000086502.1	B130055M24Rik	AC136146.4	processed_transcript	chr7	0.618543517767385

ENSMUSG00000108179.1	Gm44274	RP23-200M5.8	TEC	chr6	0.587801436378275
ENSMUSG00000100798.1	Gm19589	RP24-73K7.4	lincRNA	chr1	0.572946076366558
ENSMUSG00000092569.1	Gm20544	RP24-263P17.1	processed_transcript	chr18	0.533341171626104
ENSMUSG00000107768.1	Gm44275	RP23-200M5.7	lincRNA	chr6	0.532765220932104
ENSMUSG00000100975.2	Gm28875	RP24-159O13.1	lincRNA	chr12	0.528535249894406
ENSMUSG00000071456.6	1110002L01Rik	RP23-83G6.1	antisense	chr12	0.516662782162896
ENSMUSG00000022639.14	Dubr	Dubr	lincRNA	chr16	0.481941499178748
ENSMUSG00000086503.3	Xist	Xist	lincRNA	chrX	0.446012439564391
ENSMUSG00000097072.7	Foxl2os	Foxl2os	processed_transcript	chr9	0.38733666853129
ENSMUSG00000097772.7	5430416N02Rik	RP23-325M7.3	lincRNA	chr5	0.382381981214918
ENSMUSG00000113184.1	Gm49654	RP23-442M3.3	lincRNA	chr12	0.380769298346065
ENSMUSG00000097392.2	D930016D06Rik	RP23-314A12.2	lincRNA	chr5	0.36142080614016
ENSMUSG00000031838.8	Ifi30	Ifi30	processed_transcript	chr8	0.354722396340203
ENSMUSG00000085208.3	Brip1os	Brip1os	antisense	chr11	0.329050142008989
ENSMUSG00000097971.3	Gm26917	RP23-81C12.1	lincRNA	chr17	0.229461574038311
ENSMUSG00000105987.4	AI506816	RP23-110C17.2	processed_transcript	chr5	0.211823470629223
ENSMUSG00000029447.11	Cct6a	Cct6a	processed_transcript	chr5	0.167349589647712

10.2 Apêndice 2 – Lista completa dos lncRNAs hiporregulados antes do ensaio de adesão

ID_Ensembl	Nome_Ensembl	Nome_Gencode	Tipo	Chr	Log(2)FC
ENSMUSG00000097336.8	Fendrr	Fendrr	processed_transcript	chr8	-4.39841211281411
ENSMUSG00000086324.8	Gm15564	RP24-459O21.3	antisense	chr16	-2.8992396959009
ENSMUSG00000110537.1	Gm4316	RP23-115O21.3	lincRNA	chr8	-1.92139045902179
ENSMUSG00000051832.7	E230016K23Rik	RP23-320E6.5	processed_transcript	chr11	-1.84972000470977
ENSMUSG00000087132.7	A930001C03Rik	AC124347.2	processed_transcript	chr19	-1.6690622794737
ENSMUSG00000112627.1	4933412E12Rik	RP23-413G8.2	bidirectional	chr10	-1.54359798447323
ENSMUSG00000048603.8	Gm9828	RP24-76E2.3	bidirectional	chr13	-1.52886402885372
ENSMUSG00000086629.8	2810403D21Rik	RP23-145N17.1	processed_transcript	chrX	-1.51183828152075
ENSMUSG00000057359.7	Gm17494	RP23-71G16.2	processed_transcript	chr3	-1.50537074849227
ENSMUSG00000085147.1	Gm12609	RP23-445E3.1	processed_transcript	chr4	-1.4788977971886
ENSMUSG00000060380.3	C030014I23Rik	RP24-175C20.1	lincRNA	chr9	-1.42278454775158
ENSMUSG00000101438.1	Gm19412	RP23-222K15.31	antisense	chr17	-1.34751270849937
ENSMUSG00000106383.1	E330034L11Rik	RP23-231L15.4	TEC	chr3	-1.30822140186272
ENSMUSG00000097124.2	A530020G20Rik	RP23-30A13.2	antisense	chr3	-1.27569344193324
ENSMUSG00000096995.2	2810029C07Rik	RP23-3K10.3	bidirectional	chr12	-1.24779473939718
ENSMUSG00000097309.2	Gm17249	RP24-326A21.3	bidirectional	chr10	-1.24618082354846
ENSMUSG00000085151.7	1110018N20Rik	RP23-105M23.4	processed_transcript	chr2	-1.19463424466126
ENSMUSG00000086015.1	4833417C18Rik	RP23-151M22.6	processed_transcript	chr11	-1.16889650371503
ENSMUSG00000105176.1	Gm43668	RP23-418O21.5	TEC	chr3	-1.16647224173831
ENSMUSG00000093594.1	Gm20707	RP23-406N5.8	lincRNA	chr4	-1.13016847676318
ENSMUSG00000085992.1	Gm11515	RP23-265A5.1	lincRNA	chr11	-1.10209603917335
ENSMUSG00000087187.1	Gm13431	RP23-454A14.2	lincRNA	chr2	-1.06088148047096
ENSMUSG00000089940.1	Gm4117	RP24-163G22.1	antisense	chr13	-1.03888168979845
ENSMUSG00000048106.5	4632415L05Rik	RP23-105K2.3	lincRNA	chr3	-1.01054326842167
ENSMUSG00000097709.8	2810429I04Rik	RP24-459P6.1	lincRNA	chr13	-0.99043245400406
ENSMUSG00000110834.1	Gm39469	CAA01180111.2	lincRNA	chr9	-0.974947524902878
ENSMUSG00000087143.9	A830082K12Rik	RP23-471B6.2	antisense	chr13	-0.972835453799715

ENSMUSG00000085457.3	1110046J04Rik	RP23-262J21.8	antisense	chr13	-0.960641453619757
ENSMUSG00000115092.1	E130202H07Rik	RP24-380K14.3	lincRNA	chr14	-0.949163890562736
ENSMUSG00000109231.1	Gm45737	RP23-323L3.1	TEC	chr7	-0.895759912541721
ENSMUSG00000113585.1	Gm19605	RP23-403P3.1	processed_transcript	chr12	-0.869205896891782
ENSMUSG00000110980.1	Gm47204	RP24-175C20.18	lincRNA	chr9	-0.863684498921084
ENSMUSG00000043903.3	Zfp469	RP24-286N17.2	lincRNA	chr8	-0.856655052032435
ENSMUSG00000096954.1	Gdap10	Gdap10	lincRNA	chr12	-0.826515771195121
ENSMUSG00000095123.1	Gm21781	RP24-538B17.1	lincRNA	chr10	-0.808908707653211
ENSMUSG00000100826.6	Snhg14	Snhg14	processed_transcript	chr7	-0.778629955233931
ENSMUSG00000086328.1	2700033N17Rik	RP23-162G11.1	lincRNA	chr2	-0.737110914211992
ENSMUSG00000089804.2	Gm16136	AC114008.1	antisense	chr15	-0.730471430714749
ENSMUSG00000110631.1	Gm42047	RP24-286J21.7	lincRNA	chr8	-0.721486256995187
ENSMUSG00000101335.1	Gm28229	RP24-286H7.4	sense_intronic	chr9	-0.687962063988583
ENSMUSG00000092274.2	Neat1	Neat1	lincRNA	chr19	-0.685208734435995
ENSMUSG00000085705.1	Gm16046	CH36-125B10.3	processed_transcript	chr17	-0.675130943530663
ENSMUSG00000072893.12	4933439C10Rik	RP23-210M6.5	processed_transcript	chr11	-0.661953894160515
ENSMUSG00000086894.2	Gm15708	AC139636.2	sense_intronic	chr5	-0.654207685141136
ENSMUSG00000087396.7	4933407K13Rik	RP23-235P16.3	processed_transcript	chrX	-0.623260864500755
ENSMUSG00000097589.9	Dleu2	Dleu2	processed_transcript	chr14	-0.568468219253958
ENSMUSG00000060143.7	Gm10076	RP23-4P23.3	lincRNA	chr14	-0.557521771509531
ENSMUSG00000063235.16	Ptpmt1	Ptpmt1	processed_transcript	chr2	-0.537241231386999
ENSMUSG00000104960.1	Snhg8	Snhg8	processed_transcript	chr3	-0.529582819467578
ENSMUSG00000101970.6	1810026B05Rik	RP23-32A8.1	processed_transcript	chr7	-0.5095351822237
ENSMUSG00000086841.1	2410006H16Rik	RP23-234K24.8	processed_transcript	chr11	-0.493029897117551
ENSMUSG00000108621.1	Gm37494	RP23-402K24.2	lincRNA	chr7	-0.474727328327119
ENSMUSG00000089726.2	Mir17hg	Mir17hg	lincRNA	chr14	-0.433662378596144
ENSMUSG00000087174.7	5530601H04Rik	RP23-250A14.2	processed_transcript	chrX	-0.430944551019005
ENSMUSG00000111394.1	Gm49759	RP23-320D23.6	TEC	chr9	-0.430588011436268
ENSMUSG00000092203.7	1110038B12Rik	RP23-349B4.7	processed_transcript	chr17	-0.425145312398768
ENSMUSG00000092341.2	Malat1	Malat1	lincRNA	chr19	-0.424404282788372

ENSMUSG00000053332.13	Gas5	Gas5	processed_transcript	chr1	-0.410583827398523
ENSMUSG00000106847.1	Peg13	Peg13	sense_intronic	chr15	-0.409893488589839
ENSMUSG00000098318.7	Lockd	Lockd	lincRNA	chr6	-0.406659332517355
ENSMUSG00000074578.13	Zfas1	Zfas1	processed_transcript	chr2	-0.312121653676142
ENSMUSG00000108414.1	Snhg1	Snhg1	processed_transcript	chr19	-0.253022398067805
ENSMUSG00000085385.7	Snhg17	Snhg17	processed_transcript	chr2	-0.218288453563759
ENSMUSG00000085396.7	Firre	Firre	processed_transcript	chrX	-0.206072341623136

10.3 Apêndice 3 – Lista completa dos lncRNAs hiperregulados após do ensaio de adesão

ID_Ensembl	Nome_Ensembl	Nome_Gencode	Tipo	Chr	Log(2)FC
ENSMUSG00000098959.1	Gm27230	RP24-282D16.2	processed_transcript	chr9	5.9655969914345
ENSMUSG00000103358.1	Gm37593	RP24-490B17.2	TEC	chr17	4.09199945251259
ENSMUSG00000085971.4	Gm15411	AC158782.7	processed_transcript	chr5	4.05899934931405
ENSMUSG00000084796.1	Mir142hg	Mir142hg	lincRNA	chr11	3.66200316824931
ENSMUSG00000063109.13	Dgkeos	Dgkeos	antisense	chr11	3.51854929686628
ENSMUSG00000113908.1	Gm10811	RP23-38L5.8	TEC	chr13	2.97052406437679
ENSMUSG00000104811.1	Gm33093	RP23-307D8.9	lincRNA	chr5	2.86206663521232
ENSMUSG00000109002.1	Gm38405	RP23-428N8.3	antisense	chr7	2.84112571329894
ENSMUSG00000113049.1	Gm48544	RP24-574N19.4	lincRNA	chr13	2.78821575342399
ENSMUSG00000106757.1	Gm43482	RP23-189N6.3	TEC	chr5	2.78217860226
ENSMUSG00000056145.5	AI504432	RP23-469H5.1	lincRNA	chr3	2.65901943848516
ENSMUSG00000106652.1	Gm43358	RP23-307D8.3	lincRNA	chr5	2.40530669133993
ENSMUSG00000097289.1	2010300F17Rik	2010300F17Rik	lincRNA	chr11	2.39614833032021
ENSMUSG00000107215.1	Gm43197	RP23-253G12.9	TEC	chr6	2.09266301920841
ENSMUSG00000097354.7	2310001H17Rik	RP24-530N5.3	lincRNA	chr6	2.09088404542494

ENSMUSG00000097636.7	Mirt1	Mirt1	lincRNA	chr19	2.04140688099855
ENSMUSG00000105826.1	Gm42141	RP23-41J13.2	TEC	chr5	1.81272324579733
ENSMUSG00000104344.1	Gm38077	RP23-444I6.3	TEC	chr9	1.71047036179709
ENSMUSG00000106734.3	Gm20559	RP24-328P2.5	processed_transcript	chr6	1.66521286674522
ENSMUSG00000097209.2	AU022754	RP23-64J7.5	antisense	chr15	1.58538357111041
ENSMUSG00000104982.1	Gm32554	RP23-307D8.7	lincRNA	chr5	1.49626155944636
ENSMUSG00000104554.1	Gm4610	RP23-71G16.6	antisense	chr3	1.49610664262705
ENSMUSG00000091422.8	Gm6455	RP24-113C21.1	lincRNA	chr5	1.43295718332153
ENSMUSG00000097252.1	Gm6634	Gm6634	lincRNA	chr3	1.25001479170818
ENSMUSG00000106874.1	Gm20186	RP23-148L8.3	lincRNA	chr6	1.0944472559947
ENSMUSG00000108179.1	Gm44274	RP23-200M5.8	TEC	chr6	0.992354759732477


10.4 Apêndice 4 – Lista completa dos lncRNAs hiporregulados após do ensaio de adesão

ID_Ensembl	Nome_Ensembl	Nome_Gencode	Tipo	Chr	Log(2)FC
ENSMUSG00000115762.1	Gm34907	RP23-296J16.3	lincRNA	chr14	-5.89257169219668
ENSMUSG00000097336.8	Fendrr	Fendrr	processed_transcript	chr8	-4.28041421183493
ENSMUSG00000086163.1	Gm14206	RP23-217I3.12	antisense	chr11	-2.91898491759021
ENSMUSG00000100094.1	1810008I18Rik	RP23-181K4.2	lincRNA	chr7	-2.50524442052799
ENSMUSG00000087143.9	A830082K12Rik	RP23-471B6.2	antisense	chr13	-2.28834716027512
ENSMUSG00000051832.7	E230016K23Rik	RP23-320E6.5	processed_transcript	chr11	-1.87340794765576
ENSMUSG00000112627.1	4933412E12Rik	RP23-413G8.2	bidirectional	chr10	-1.44061656832222

ANEXOS

11. ANEXOS

11.1 Anexo 1- Certificação da comissão de ética em experimentação animal



UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
COMISSÃO DE ÉTICA NO USO DE ANIMAIS





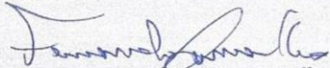
CERTIFICADO

Certificamos que o Protocolo intitulado "*Efeito de mutações no gene AIRE (Síndrome APSI) induzidas por CRISPR-Ca9 na conformação da proteína, no transcriptoma de células mTEC e na sua interação com tímócitos*", registrado com o número **003/2017-1**, sob a responsabilidade do **Prof. Dr. Geraldo Aleixo da Silva Passos Junior**, envolvendo a produção, manutenção ou utilização de animais pertencentes ao *filo Chordata, subfilo Vertebrata* (exceto humanos) para fins de pesquisa científica, encontra-se de acordo com os preceitos da Lei nº 11.794 de 8 de outubro de 2008, do Decreto nº 6.899 de 15 de julho de 2009 e com as normas editadas pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA), e foi **APROVADO** pela Comissão de Ética no Uso de Animais da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo em reunião de 26 de junho de 2017.

Este Protocolo prevê a utilização de 100 camundongos C57BL/6 fêmeas pesando 20g e oriundos do Serviço de Biotério da Prefeitura do Campus de Ribeirão Preto da Universidade de São Paulo. Vigência da autorização: 26/06/2017 a 30/07/2022.

We certify that the Protocol n° 003/2017-1, entitled "*Effect of AIRE gene mutations (APSI syndrome) induced by CRISPR-Cas9 in protein conformation, transcriptome of mTEC cells and its interaction with thymocytes*", is in accordance with the Ethical Principles in Animal Research adopted by the National Council for the Control of Animal Experimentation (CONCEA) and was approved by the Local Animal Ethical Committee from Ribeirão Preto Medical School of the University of São Paulo in 06/26/2017. This protocol involves the production, maintenance or use of animals from *phylum Chordata, subphylum Vertebrata* (except humans) for research purposes, and includes the use of 100 female C57BL/6 mice weighing 20g from the Central Animal House of Ribeirão Preto Medical School, University of São Paulo. This certificate is valid until 07/30/2022.

Ribeirão Preto, 26 de junho de 2017


Prof. Dr. Fernando Silva Ramalho
 Presidente da CEUA – FMRP – USP

Faculdade de Medicina de Ribeirão Preto - USP - Av. Bandeirantes, 3900 - Ribeirão Preto - SP - Brasil -14049-900
(16) 3315-3301 / 3275 / 3258