

UNIVERSIDADE FEDERAL DO PARANÁ
SARA CRISTINA LOBO ALVES

PERFIL DE EXPRESSÃO DE RNAs E POLIMORFISMO DE lncRNAs NO PÊNFIGO
FOLIÁCEO ENDÊMICO

CURITIBA
2019

SARA CRISTINA LOBO ALVES

PERFIL DE EXPRESSÃO DE RNAs E POLIMORFISMO lncRNAs NO PÊNFIGO
FOLIÁCEO ENDÊMICO

Tese de doutorado apresentada ao Programa de Pós-Graduação em Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial para obtenção do título de doutora em genética.

Orientadora: Prof^a. Dr^a. Maria Luiza Petzl-Erler
Coorientadores: Dr. Rodrigo Coutinho de Almeida e Prof Dr. Saleh Ibrahim.

CURITIBA
2019

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Giana Mara Seniski Silva – CRB/9 1406)

Alves, Sara Cristina Lobo
Perfil de expressão de RNAs e polimorfismo lncRNAs no pênfigo
foliáceo endêmico. / Sara Cristina Lobo Alves. – Curitiba, 2019.
111 p. : il.

Orientadora: Maria Luiza Petzl-Erler
Coorientadores: Rodrigo Coutinho de Almeida e Saleh Ibrahim

Tese (doutorado) - Universidade Federal do Paraná, Setor de Ciências
Biológicas. Programa de Pós-Graduação em Genética.

1. Pênfigo 2. Expressão gênica 3. Doenças autoimunes 4. RNA I.
Título II. Petzl-Erler, Maria Luiza. III. Almeida, Rodrigo Coutinho de IV.
Ibrahim, Saleh V. Universidade Federal do Paraná. Setor de Ciências
Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 571.973



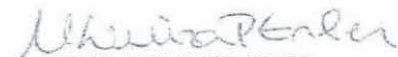
MINISTÉRIO DA EDUCAÇÃO
SETOR SETOR DE CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA -
40001016006P1

TERMO DE APROVAÇÃO

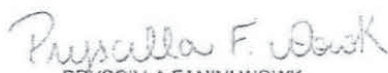
Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de SARA CRISTINA LOBO ALVES intitulada: Perfil de expressão de RNAs e polimorfismo de lncRNAs no pênfigo foliáceo endêmico, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

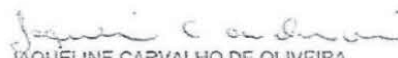
CURITIBA, 26 de Abril de 2019.


MARIA LUIZA PETZL ERLER
Presidente da Banca Examinadora


WILSON ARAÚJO DA SILVA JÚNIOR
Avaliador Externo (FMRPUSP)


PRISCILLA FANINI WOLK
Avaliador Externo (ICC-FIOCRUZ)


GABRIEL ADELMAN CIPOLLA
Avaliador Externo (UFPR)


JAQUELINE CARVALHO DE OLIVEIRA
Avaliador Interno (UFPR)


WANDERSON DUARTE DA ROCHA
Avaliador Externo (UFPR)

AGRADECIMENTOS

À minha orientadora, Prof^a Dr^a Maria Luiza Petzl-Erler pela oportunidade de desenvolver esse trabalho no Laboratório de Genética Molecular Humana – LGMH, pela orientação e contribuição para minha formação acadêmica, profissional e também pessoal.

Ao meu coorientador, Dr. Rodrigo Coutinho de Almeida pelo apoio e pelas ideias que ajudaram a compor esse trabalho.

Aos colegas e amigos do LGMH, pelo companherismo, boas conversas, suporte nos experimentos, compartilhamento de ideias e planos. A amizade de vocês foi essencial para que minha vida fosse mais feliz nesses vários anos de convivência.

Aos colegas do LIED em especial ao Prof Dr. Saleh Ibrahim, Prof Dr. Hauke Bush e à Dr^a Mareike White por me receberem e me apoiarem no meu período de estágio.

Aos meus pais, Luis e Romilda, por me ensinarem princípios e valores que fizeram e fazem diferença em tudo que faço na minha vida, inclusive a vida acadêmica. Agradeço por sempre me apoiarem e acreditarem em mim.

À minhas irmãs e irmão pela companhia durante a vida, pelo apoio durante o doutorado e por serem meus primeiros e melhores amigos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pela concessão da bolsa de doutorado, bolsa de doutorado sanduíche e apoio financeiro para o desenvolvimento desta pesquisa.

RESUMO

Nas últimas décadas, o uso de técnicas de sequenciamento em larga escala adicionou informações valiosas sobre a complexidade do genoma humano e de sequências não codificantes nele presentes. Por muito tempo uma porção considerável do genoma foi considerada DNA lixo por não codificar proteína, porém, estudos demonstraram que parte significativa do DNA é transcrita, mas não traduzida, e que muitos desses transcritos são importantes em processos essenciais às células. Entre eles estão os RNAs longos não codificadores (lncRNAs) que têm se destacado funcionalmente na regulação da expressão gênica. Compreender melhor a multiplicidade funcional do genoma contribui para o avanço no entendimento de doenças complexas nas quais fatores genéticos têm efeito. O pênfigo foliáceo (PF) é uma doença autoimune da pele na qual há a formação de autoanticorpos contra a desmogleína 1, uma proteína constituinte dos desmossomos responsáveis pela adesão entre queratinócitos. O PF se caracteriza pelo desprendimento da camada superior do epitélio, com formação de bolhas e erosões na pele. É uma doença em cuja patogênese atuam fatores ambientais e genéticos, e é endêmica em certas regiões geográficas do mundo, sendo o Brasil o país em que atinge sua maior incidência. Neste trabalho utilizamos duas abordagens principais com o objetivo de contribuir para uma melhor compreensão da patogênese do pênfigo foliáceo endêmico (PFE): uma, de análise de associação genética com polimorfismos de RNAs longos não codificadores (lncRNAs) e outra, de expressão diferencial em nível de RNA. Primeiramente investigamos a existência de associação entre 2.080 polimorfismos de um único nucleotídeo (SNPs) localizados em genes de lncRNAs e o PFE, a fim de descrevermos variantes de predisposição à doença. Encontramos um SNP associado, *rs7144332* localizado no lncRNA *AL110292.1*, e outros cinco SNPs sugestivos de associação, os quais podem alterar a função de lncRNAs e influenciar a susceptibilidade ao PFE. Na segunda abordagem, realizamos o sequenciamento de RNA (RNA-seq) de células T CD4⁺ de pacientes de PFE e de indivíduos sem a doença com o objetivo de descrever o perfil de mRNAs e lncRNAs com expressão alterada na condição de doença. Identificamos 324 genes diferencialmente expressos, entre esses, 32 genes de lncRNAs. Vários desses genes têm funções na resposta imune e inflamatória, e alguns deles já haviam sido encontrados diferencialmente expressos em PF em uma análise de microarray feita anteriormente. Por meio da técnica de PCR quantitativa (qPCR) validamos a diferença de expressão encontrada para os genes *LEF1-AS1*, *IL18RAP*, *GNLY* e *S1PR5* em células T CD4⁺. Também avaliamos a expressão dos mesmos genes em células B de pacientes e controles, sendo que *IL18RAP*, *GNLY* e *S1PR5* foram detectados diferencialmente expressos. Desta maneira, foi possível corroborar a participação de diversos genes codificadores e não codificadores de proteína no desencadeamento e/ou manutenção do PFE, além de apontar outros genes nunca antes implicados no pênfigo. Tais genes se destacam como candidatos interessantes para estudos posteriores, e podem tornar-se potenciais alvos terapêuticos em PF.

palavras-chave: pênfigo foliáceo, SNP, estudo de associação, lncRNA, RNA-seq, transcriptoma, expressão gênica, doença autoimune.

ABSTRACT

In recent decades, the use of large-scale sequencing techniques has added valuable insights into the complexity of the human genome and the non-coding sequences present in it. For a long time, a considerable portion of the genome was considered junk DNA because it did not encode protein. However, studies have shown that a significant part of DNA is transcribed but not translated, and many of these transcripts are important in processes essential to cells. Among them are the long non-coding RNAs (lncRNAs) that have been functionally prominent in the regulation of gene expression. Better understanding of the functional multiplicity of the genome contributes to advancing the understanding of complex diseases in which genetic factors have an effect. Pemphigus foliaceus (PF) is an autoimmune disease of the skin in which there is the formation of autoantibodies against desmoglein 1, a protein constituent of the desmosomes responsible for adhesion between keratinocytes. PF is characterized by the detachment of the upper epithelial layer, with formation of blisters and erosions on the skin. It is a disease whose pathogenesis affects environmental and genetic factors and is endemic in certain geographic regions of the world. Brazil is the country where it reaches its highest incidence. In this work we used two main approaches with the aim of contributing to a better understanding of the pathogenesis of endemic pemphigus foliaceus (EPF): genetic association analysis with polymorphisms of long non-coding RNAs (lncRNAs) and differential expression analysis at the RNA level. We first investigated the association between 2,080 single nucleotide polymorphisms (SNPs) located in lncRNAs genes and EPF, in order to describe variants of predisposition to the disease. We found one associated SNP, *rs7144332* located in the lncRNA *AL110292.1*, and other five SNPs suggestive of association, which may alter the function of lncRNAs and influence the susceptibility to EPF. In the second approach, we performed RNA sequencing (RNA-seq) of CD4⁺ T cells from EPF patients and individuals without the disease aiming to describe the profile of mRNAs and lncRNAs whose expression is altered in the disease condition. We identified 324 differentially expressed genes, among them, 32 lncRNA genes. Several of these genes have functions in the immune and inflammatory response and some of them had previously been found differentially expressed in EPF by microarray analysis. Using the quantitative PCR technique (qPCR) we validate the expression difference found for the *LEF1-AS1*, *IL18RAP*, *GNLY* and *S1PR5* genes in CD4⁺ T cells. We also evaluated the expression of the same genes in patients and controls B cells, and *IL18RAP*, *GNLY* and *S1PR5* were detected differentially expressed. In this way, it was possible to corroborate the participation of several protein-coding and non-coding genes of in the onset and/or maintenance of EPF, and to point out genes never previously implicated in pemphigus. Such genes stand out as interesting candidates for further studies, and may become potential therapeutic targets in PF.

Keywords: pemphigus foliaceus, SNP, association study, lncRNA, RNA-seq, transcriptome, gene expression, autoimmune disease.

Lista de abreviações e siglas

lncRNA	Long non-coding RNA/ RNA longo não codificador
RNA-seq	RNA sequencing/ Sequenciamento de RNA
qPCR	Polymerase chain reaction quantitative real time/ Reação em cadeia da polimerase quantitativa em tempo real
mRNA	Messenger RNA/ RNA mensageiro
SNP	Single nucleotide polymorphism/ Polimorfismo de nucleotídeo único
PF	Pemphigus foliaceus/ Pênfigo Foliáceo
PFE	Pênfigo Foliáceo Endêmico
PV	Pênfigo vulgar
FS	Fogo Selvagem
EGS	Extrato de glândula salivar
DSG1	Desmogleína 1
ELISA	Ensaio de Imunoabsorção enzimática
IL	Interleucina
PBMC	Peripheral blood mononuclear cells/ células mononucleares do sangue periférico
rRNA	Ribossomic RNA/ RNA ribossômico
tRNA	Transfer RNA/ RNA de transferência
ncRNA	Non-coding RNA/ RNA não codificador
miRNA	microRNA
AR	Artrite reumatoide
GWAS	Genome-wide association study/ Estudo de associação de genoma completo
EM	Esclerose múltipla
LES	Lupus eritematoso sistêmico
OR	Odds ratio

QC	Quality control/ Controle de qualidade
PCA	Principal component analysis/ Análise de Componente Principal
MFE	Minimum free-energy/ Mínima energia livre
DE	Diferencialmente expresso
DEGs	Differentially expressed genes/ Genes diferencialmente expressos
q value	Adjusted P/ Valor de P corrigido

SUMÁRIO

1. INTRODUÇÃO	13
2. REVISÃO BIBLIOGRÁFICA	15
2.1 Pênfigo foliáceo endêmico	15
2.1.1 Aspectos clínicos e epidemiologia	15
2.1.2 Etiologia	16
2.1.2.1 Fatores ambientais	16
2.1.2.2 Fatores genéticos	17
2.1.3 Patogênese	18
2.1.4 Diagnóstico e tratamento	20
2.2 RNAs longos não codificadores (lncRNAs)	21
2.2.1. Definição e principais características dos lncRNAs	21
2.2.2 Classificações dos lncRNAs	23
2.2.3 Mecanismos de ação e funções dos lncRNAs	26
2.2.4 lncRNAs no sistema imune	27
2.2.5 lncRNAs em doenças complexas	28
2.2.5.1 lncRNAs em doenças autoimunes	29
3. JUSTIFICATIVA, HIPÓTESE E OBJETIVOS	31
4. CAPÍTULO I	32
LONG NON-CODING RNA POLYMORPHISMS INFLUENCE SUSCEPTIBILITY TO ENDEMIC PEMPHIGUS FOLIACEUS	32
4.1 ABSTRACT	32
5. CAPÍTULO II	34
RNA-SEQ ANALYSIS REVEALS NEW INSIGHTS INTO ENDEMIC PEMPHIGUS FOLIACEUS	34
5.1 Abstract	34
5.2 Introduction	35
5.3 Material and methods	36
5.3.1 Sample collection and cohort	36
5.3.2 CD4 ⁺ T cell and CD19 ⁺ B cell isolation and RNA extraction	37
5.3.3 RNA sequencing	38
5.3.4 RNA sequencing data analysis	38
5.3.5 <i>In silico</i> functional analysis of differentially expressed genes	38

5.3.6 Validation of differential expression results for T cells and additional expression evaluation of B cells	39
5.4 Results	40
5.4.1 Genes differentially expressed in EPF	40
5.4.2 Validation in T cells and additional expression evaluation in B cells	42
5.3.3 <i>In silico</i> functional analysis of DEGs	45
5.4 Discussion	46
5.5 Conclusions and perspectives	50
5.6. References	50
6. Capítulo III	54
<i>Review Article</i> BESIDES PATHOLOGY: LONG NON-CODING RNA IN CELL AND TISSUE HOMEOSTASIS.	54
Abstract	54
6.1.Introduction	55
6.2 LncRNAs: expression patterns in tissues or cell types	56
6.2.2 Hematopoietic Cells	56
6.2.2.1 Erythrocytes	57
6.2.2.2 Other myeloid lineage cells.....	58
6.2.2.3 Lymphocytes	60
6.2.2.4 T lymphocytes.....	60
6.2.2.5 B lymphocytes.....	62
6.2.2.6 NK cells	63
6.2.3 Solid tissues	63
6.2.3.1 Epidermal tissue	64
6.2.3.2 Nervous tissue	65
6.2.3.4 Muscular tissue: cardiac, skeletal and smooth	67
6.2.3.5 Adipose and bone tissues	70
6.2.3.6 Hepatic tissue	72
6.2.3.7 Intestinal tissue	73
6.2.3.8 Gametogenesis and germ cells.....	74
6.2.4 Extracellular vesicles.....	75
6.3 Conclusions	76
6.4 References.....	76
7. DISCUSSÃO GERAL E CONSIDERAÇÕES FINAIS	89
REFERÊNCIAS	93

APÊNDICE	100
Apêndice A – Material suplementar Capítulo II.....	100

1. INTRODUÇÃO

O pênfigo foliáceo (PF) é uma doença autoimune, na qual há a formação de autoanticorpos contra a desmogleína 1, uma glicoproteína constituinte dos desmossomos da epiderme. A doença é endêmica em certas regiões do Brasil, onde é conhecida como fogo selvagem. Sua etiologia ainda não está bem elucidada, sabe-se, no entanto, que é uma doença complexa na qual atuam diferentes fatores, incluindo genéticos e ambientais. Assim como na maioria das doenças autoimunes o desenvolvimento do PF depende de vários genes que têm efeito baixo ou moderado sobre o fenótipo, sem que nenhum alelo seja suficiente ou necessário para que a doença ocorra.

Diversos polimorfismos presentes em genes codificadores de proteína já foram estudados quanto à suscetibilidade diferencial ao pênfigo foliáceo endêmico, porém a variação de genes não codificadores ainda não foi estudada de maneira sistematizada. Uma vez que crescem as evidências de que polimorfismos em regiões não codificadoras estão associados a várias patologias, estudá-los se torna um passo importante na compreensão de fatores genéticos implicados nas mesmas.

É também enriquecedor avaliar a dinâmica de expressão gênica pelo uso de tecnologias de sequenciamento de última geração, as quais têm revolucionado a pesquisa com as vantagens de alta produtividade, sensibilidade e velocidade. Uma vez que a expressão gênica é altamente variável e dependente de tempo, tipo celular e de estímulo, é especialmente importante conhecer o perfil do transcriptoma em condições de doença, pois auxilia na compreensão de quais genes influenciam e/ou têm sua expressão influenciada na patogênese do distúrbio em questão. Por sua capacidade de gerar proteínas, os RNAs mensageiros já são estudados há bastante tempo quanto a expressão em diferentes condições fisiológicas e patológicas. Porém, tem se tornado evidente que RNAs longo não codificadores (lncRNAs) também são importantes em processos biológicos e envolvidos em importantes funções celulares, especialmente na regulação da expressão gênica, e agem, inclusive, no desenvolvimento de células do sistema imune e no transcurso de respostas imunes.

O presente estudo, juntamente com diversos estudos já realizados pelo nosso grupo de pesquisa, tem o propósito de ajudar a compreender a patogênese do

pênfigo foliáceo, principalmente em seu componente genético e molecular. Estudamos o papel de polimorfismos em genes de lncRNAs na predisposição ao PF e também avaliamos diferenças na expressão de RNAs mensageiros e lncRNAs entre pacientes e indivíduos sem a doença. Com isso almejamos contribuir para a melhor compreensão do pênfigo foliáceo para que no futuro possa-se propiciar condições para a identificação de possíveis alvos terapêuticos.

2. REVISÃO BIBLIOGRÁFICA

2.1 Pênfigo foliáceo endêmico

2.1.1 Aspectos clínicos e epidemiologia

Pênfigo abrange um grupo de doenças autoimunes órgão-específicas da pele em que autoanticorpos contra desmogleínas, que são importantes glicoproteínas desmossomais, são detectados (JONES *et al.*, 1984). A ligação antígeno-anticorpo desencadeia a perda de adesão celular e com isso há a formação de bolhas intraepiteliais, processo que é chamado de acantólise (DIAZ *et al.*, 1989).

Os principais tipos de pênfigo são o pênfigo vulgar (PV) e o pênfigo foliáceo (PF). O PV se caracteriza pela produção de autoanticorpos contra a proteína desmogleína 3 e tanto epiderme quanto mucosas podem ser acometidas pela doença. O PF se caracteriza pela presença de autoanticorpos contra a desmogleína 1 e apenas a epiderme é acometida (UDEY e STANLEY, 1999). Os autoanticorpos anti-desmogleína são os principais em PV e PF; porém, outros autoanticorpos já foram detectados, principalmente contra outras proteínas de adesão celular da família das caderinas, como placoglobinas (KORMAN *et al.*, 1989), desmocolinas I e II (DMOCHOWSKI *et al.*, 1993), desmoplaquinas (KIM *et al.*, 2001), e também receptores colinérgicos (NGUYEN *et al.*, 1998).

Clinicamente, a lesão primária típica do PF é uma bolha superficial que facilmente se rompe, com sinal de Nikolsky positivo (camadas externas da pele se descolam das camadas internas quando pressionadas levemente) em pacientes com doença ativa. Os locais mais afetados incluem couro cabeludo e rosto, pescoço e tronco superior (HANS-FILHO *et al.*, 1999). Pacientes que apresentam a forma localizada possuem lesões, na maioria das vezes, limitadas às áreas seborréicas da pele. As vesículas e pequenas bolhas quando se rompem deixam áreas erosivas secundárias e crostas, podendo também ser eritematosas, violáceas ou pápulas hiperpigmentadas. A maioria dos pacientes permanece com essas lesões estabilizadas. Alguns, porém, podem apresentar disseminação das lesões para todo o tronco, abdômen e membros superiores e inferiores, evoluindo para a forma generalizada (CAMPBELL *et al.*, 2001).

Tanto PV quanto PF são doenças raras que ocorrem esporadicamente em todo o mundo. Porém, algumas regiões geográficas apresentam uma frequência significativamente maior do pênfigo foliáceo, nessas regiões o pênfigo foliáceo é

considerado endêmico. Brasil, Tunísia, Peru e Colômbia são alguns dos países que em o pênfigo foliáceo endêmico (PFE) ocorre. Além desses, recentemente foi reportado um foco endêmico nas Ilhas Canárias, na Espanha (GRAU-PEREZ *et al.*, 2018). Em todas essas regiões citadas não são observadas diferenças clínicas de manifestação tanto entre si quanto em relação à forma esporádica, porém, várias diferenças epidemiológicas ocorrem. Na Colômbia o PFE afeta principalmente homens acima de 40 anos envolvidos com a atividade mineradora (ROLBEDO *et al.*, 1988). Na Tunísia, a doença tem incidência de 3,9 casos por milhão de habitantes por ano e atinge principalmente mulheres de 25 a 34 anos de idade, sendo que a proporção sexual se considerados todos os tipos de pênfigo é de 4:1 (BASTUJIGARIN *et al.*, 1995; MASMOUDI *et al.*, 2018). No Brasil o pênfigo foliáceo é também conhecido como fogo selvagem (FS) e a incidência de casos é cerca de 25-35 casos por milhão por ano, com proporção sexual de 1:1 (DIAZ *et al.*, 1989). A distribuição dos focos endêmicos varia ao longo do tempo, uma vez que a doença decai com a urbanização, e novos focos aparecem em áreas de ocupação recente (CAMPBELL *et al.*, 2001). O fogo selvagem está presente nas áreas rurais, e a distribuição geográfica tende a seguir o curso de riachos e córregos. Um estudo recente na região endêmica do estado de São Paulo realizado com registros de 4 décadas (1965 a 2014) corroborou o fato que os casos de FS ocorrem em sua maioria perto de rios e áreas agrícolas (CELERE *et al.*, 2017). Até o momento, as principais regiões brasileiras afetadas incluem o centro-oeste e alguns estados do sudeste. No entanto, novas regiões endêmicas foram caracterizadas no norte do Brasil. Além disso, em algumas populações indígenas brasileiras a prevalência do FS é bastante alta; na reserva ameríndia de Limão Verde, localizada no estado do Mato Grosso do Sul, há cerca de 1.200 habitantes da tribo Terena e é um foco ativo da doença, com 3,4% de prevalência neste local (WARREN *et al.*, 2000). Entre os Xavantes, tribo indígena do estado de Mato Grosso a prevalência é de 1,4% (HANS-FILHO *et al.*, 1996).

2.1.2 Etiologia

2.1.2.1 Fatores ambientais

Pesquisadores têm hipotetizado fatores ambientais desencadeadores do PFE no Brasil, baseando-se principalmente na endemicidade da doença e no fato que as regiões endêmicas ocorrem em áreas geográficas próximas de regiões rurais ou

pouco urbanizadas. A exposição a picadas de insetos hematófagos é um dos principais fatores ambientais suspeitos de contribuir para o desenvolvimento da doença. Alguns estudos que apoiam essa hipótese mostram que regiões endêmicas do PFE se sobrepõe com a distribuição do mosquito *Simulium nigrimanum*, também conhecido como borrachudo (LOMBARDI *et al.*, 1992; EATON *et al.*, 1998). Um estudo recente realizado na área endêmica do estado de São Paulo analisou a distribuição de 3 espécies de mosquitos simulídeos e flebotomíneos em 60 municípios e relacionou com os níveis de anticorpos IgG contra extrato da glândula salivar (EGS) dos mosquitos *Simulium nigrimanum*, *Nyssomyia neivai* e *Aedes aegypti* (como controle) em pacientes, em seus parentes e também em vizinhos. Como resultado observaram que os espécimes de *S. nigrimanum* e *N. neivai* foram registrados em oito (13,3%) e 26 (43,3%) dos municípios analisados, respectivamente. Também reportaram que pacientes de PFE e seus familiares apresentaram níveis mais elevados de IgG contra os EGS de *S. nigrimanum* e *N. neivai*, mas não contra o EGS de *A. aegypti*, em comparação com os controles (VERNAL *et al.*, 2017).

Além disso, já foi visto que há uma alta prevalência não só de picadas de borrachudos, mas também de picadas de barbeiros e percevejos em pacientes de PFE (AOKI *et al.*, 2004). Há ainda uma coincidência na área endêmica do PFE, da leishmaniose cutânea e da doença de Chagas, as duas últimas sabidamente transmitidas por insetos vetores (SOUSA *et al.*, 2012). Uma das possibilidades de ação das picadas de insetos no desencadeamento da doença seria através do mimetismo molecular com peptídeos próprios do paciente de modo a provocar uma reação cruzada. Sobre isso um estudo demonstrou que anticorpos anti-desmogleína 1 de pacientes de PFE reagem de forma cruzada com antígenos salivares do mosquito-palha (*Lutzomyia longipalpis*), reconhecendo, mais precisamente, a proteína LJM11 (QIAN *et al.*, 2012).

2.1.2.2 Fatores genéticos

O componente genético implicado na susceptibilidade à doença evidencia-se pelo fato de que há uma agregação de casos entre indivíduos geneticamente relacionados. Ainda que se possa pensar que isso poderia ser decorrente do compartilhamento do mesmo ambiente, o que se observa é que o número de casos

entre indivíduos geneticamente relacionados é superior ao de indivíduos com apenas ligações sociais, demonstrando que existe uma predisposição genética a esta doença (HANS-FILHO *et al.*, 1999). A doença, porém, não pode ser explicada por um padrão de herança mendeliana, uma vez que as taxas de recorrência são inferiores às esperadas para genes únicos de penetrância completa, sendo, desta forma, compatíveis com um padrão de herança poligênica que atua em conjunto com fatores ambientais.

A fim de identificar os múltiplos genes que podem conferir susceptibilidade ao pênfigo foliáceo endêmico, diversos estudos de associação têm sido realizados. Esses estudos geralmente baseiam-se na escolha de genes candidatos, os quais em sua maioria desempenham papéis na regulação do sistema imune. Genes que codificam, por exemplo, antígenos-alvo, moléculas apresentadoras ou processadoras de antígeno, receptores de células B e T e citocinas, assim como moléculas reguladoras da apoptose (TRON *et al.*, 2005). O primeiro estudo de associação com o PFE detectou associação positiva com HLA-DR e HLA-DQ (PETZL-ERLER E SANTAMARIA, 1989), descrita com mais detalhes posteriormente (PAVONI *et al.*, 2003). Além dessas, foram observadas associações com polimorfismos de vários genes: *CD40L*, *CD40*, *BAFF* (MALHEIROS E PETZL-ERLER, 2009), *IL4*, *IL6* (PEREIRA *et al.*, 2004), *PDCD1* (BRAUN-PRADO E PETZL-ERLER, 2007), *CD28*, *CD86*, *CTLA4* (DALLA-COSTA *et al.*, 2010), *CIITA* (PIOVEZAN E PETZL-ERLER, 2013), genes *KIR* (AUGUSTO *et al.*, 2012; AUGUSTO *et al.*, 2015), *CD33*, *LAIR 1*, *LAIR 2* (CAMARGO *et al.*, 2016), *BAFF*, *APRIL*, *BCMA* e *TACI* (OLIVEIRA, 2015), *KLRG1* (CIPOLLA *et al.*, 2016).

Além de estudos associando polimorfismos de diversos genes à susceptibilidade ou proteção ao PF, a expressão gênica de pacientes de pênfigo também já foi avaliada. Utilizando a técnica de microarranjo, a expressão gênica global de mRNA em células T CD4⁺ foi quantificada. Foram analisados subgrupos de pacientes e indivíduos saudáveis quanto à expressão diferencial. Um total de 135 genes foram detectados diferencialmente expressos, sendo que grande parte deles possui funções relacionadas a adesão e migração de linfócitos, apoptose, proliferação celular, citotoxicidade e apresentação de antígenos (MALHEIROS *et al.*, 2014).

2.1.3 Patogênese

A desmogleína 1 (DSG1), caracterizada como principal antígeno do PF, é uma glicoproteína transmembrânica, integrante dos desmossomos e que tem uma participação importante na adesão celular da epiderme (WHEELER *et al.*, 1991). A detecção de autoanticorpos circulantes no soro é realizada principalmente por imunofluorescência indireta, e por imunofluorescência direta detectam-se os autoanticorpos depositados sobre a membrana dos queratinócitos da pele lesionada de pacientes (CAMPBELL *et al.*, 2001). Os anticorpos circulantes são principalmente IgG, das subclasses IgG1 e IgG4, sendo que o IgG4 é o tipo predominante entre os anticorpos depositados sobre os queratinócitos. Foi observado também que o estado de remissão da doença depende do decréscimo dos títulos de IgG4 (MARTEL e JOLY, 2001). Em 2003, Li *et al.* sugeriram um modelo para explicar a via imunopatogênica do PFE (figura 1). Segundo essa hipótese os autoanticorpos anti DSG1 em PFE são inicialmente formados contra o domínio EC5 da proteína DSG1, em indivíduos saudáveis da região endêmica; dentre esses, aqueles que forem geneticamente susceptíveis desenvolvem anticorpos contra outros epítomos da proteína (EC1 e EC2) desencadeando o processo patológico do PFE (LI *et al.*, 2003).

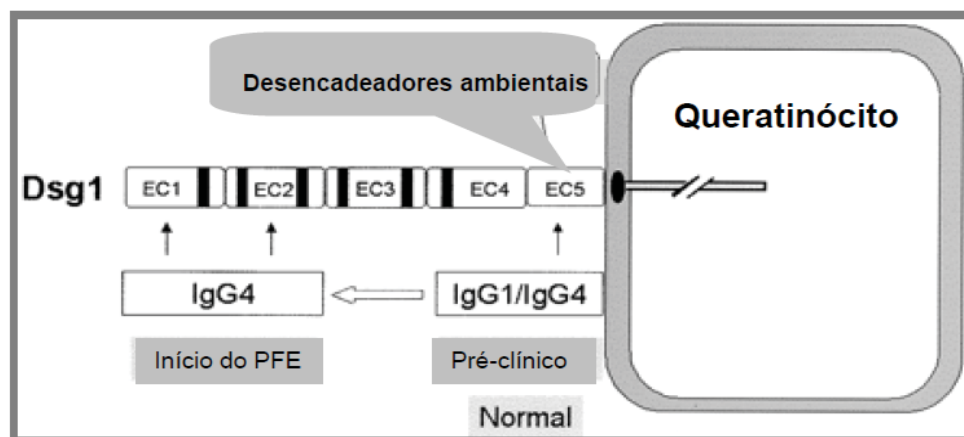


Figura 1: Modelo da via imunopatogênica do PFE: Inicialmente, fatores ambientais desencadeiam a produção de anticorpos IgG1 e IgG4 não-patogênicos contra o domínio EC5 da DSG1. Em indivíduos geneticamente predispostos, o espalhamento intramolecular de epítomos resulta na geração de anticorpos IgG4 contra os domínios EC1 e EC2, levando ao desenvolvimento do PFE (modificado de LI *et al.*, 2003).

Sendo o PF uma doença mediada por anticorpos é esperado que células T também participem da patogênese da doença, uma vez que a produção de anticorpos contra antígenos proteicos pelas células B geralmente requer a colaboração de células T CD4⁺. Através de um estudo realizado com pacientes da

tribo indígena Terena, foi demonstrado que linfócitos T de pacientes reconhecem a DSG1 recombinante como autoantígeno realizando proliferação antígeno específica (LIN *et al.*, 2000). A apresentação do antígeno é feita por moléculas HLA de classe II, cujas variantes genéticas já foram encontradas associadas ao pênfigo foliáceo endêmico, salientando a importância das células T no desenvolvimento da doença (PETZL-ERLER e SANTAMARIA, 1989; PAVONI *et al.*, 2003).

Uma investigação dos aspectos imunopatológicos do PF no Brasil determinou o nível de várias citocinas presentes no soro de pacientes e controles saudáveis. Foram quantificadas por ELISA as citocinas IL-2, -4, -5, -10, -12 e INF- γ em ambos os grupos. A concentração média de IL-2 e IL-4, IL-5 e INF- γ foi menor em pacientes quando comparada aos controles. Já as citocinas IL-10 e IL-12 estão em níveis mais elevados no soro dos pacientes com PF. IL-12 é produzida principalmente por linfócitos B, células de Langerhans e macrófagos e é responsável pela estimulação de citocinas do perfil Th1. Por outro lado, IL-10 inibe o perfil Th1. Segundo o estudo, o decréscimo dos níveis das citocinas IL-2 e INF- γ no PF, podem ser causados pela inibição do perfil Th1. A IL-10 encontra-se aumentada no PF, provavelmente, devido ao estímulo da proliferação de linfócitos B, sendo assim responsável pela ampla produção de autoanticorpos e inibição de Th1. Não foi encontrada associação entre diferentes níveis de citocinas e as formas localizada e generalizada de PF (ZEOTI *et al.*, 2000).

Outro estudo avaliou o perfil de citocinas e quimiocinas no pênfigo, em pacientes com FS, recém-diagnosticados e ainda sem intervenção terapêutica. Quando comparados a indivíduos saudáveis, tinham níveis reduzidos de IFN- γ e IL-2, concordante com o estudo citado acima, e também de IL-15 e CCL-11, e níveis mais elevados de IL-22 e CXCL-8. Além disso, o tratamento com drogas imunossupressoras aumentou a produção de IFN- γ , IL-2, CCL-5 e CCL-11, além de reduzir os níveis de IL-22 e CXCL-10 (TIMÓTEO *et al.*, 2017).

2.1.4 Diagnóstico e tratamento

O diagnóstico é baseado em sintomas clínicos e exames imuno- e histológico. Através do exame sorológico ocorre a detecção de autoanticorpos anti-DSG1 enquanto que o exame histológico permite a detecção da acantólise (CAMPBELL *et al.*, 2001). Uma das formas de diferenciação entre FS e PV é através da observação da ausência de lesões orais em FS. A imunofluorescência indireta (IFI) e o exame

sorológico de ELISA revelam autoanticorpos em pacientes com FS em níveis distintos que podem ser associados com a atividade da doença (CAMPBELL *et al.*, 2001; OLIVEIRA *et al.*, 2016). Autoanticorpos anti-DSG1 foram observados com prevalência de 84% em pacientes de PF sendo que pacientes em remissão diferiram daqueles pacientes com doença ativa. Os pacientes em remissão a longo prazo (mais de dois anos sem apresentar novas lesões) comportaram-se semelhantemente aos indivíduos controles endêmicos (OLIVEIRA *et al.*, 2016).

O pênfigo foliáceo é tratado durante toda a vida e o objetivo do tratamento utilizado é induzir e manter o estado de remissão. Este objetivo é geralmente alcançado pelo uso de agentes imunossupressores sendo que dose e período de administração dependem da resposta de cada paciente. O principal e mais bem estabelecido tratamento é o uso de glicocorticoides sistêmicos, como a prednisona. Uma vez que a terapia é prolongada, vários efeitos colaterais induzidos por esteroides podem ocorrer, alguns deles bastantes sérios a ponto de comprometer a adesão de pacientes ao tratamento. Sendo assim, pesquisadores continuam buscando novas modalidades que reduzam a necessidade de glicocorticoides.

Nas últimas décadas, vários outros tratamentos foram testados e alguns deles são usados como adjuvantes com eficácia variada. Alguns desses tratamentos são: azatioprina, ciclofosfamida, micofenolato mofetil, imunoglobulina intravenosa, rituximab. Segundo uma meta-análise do ano de 2014, nenhum dos tratamentos adjuvantes (combinados com glicocorticoides) foi mais eficaz do que os glicocorticoides sozinhos em termos de taxa de resposta completa. No entanto, a maioria dos tratamentos mostrou algum benefício como efeito poupador de esteroides, atraso no tempo de recaída, aceleração da cicatrização de lesões (ATZMONY *et al.*, 2014).

2.2 RNAs longos não codificadores (lncRNAs)

2.2.1. Definição e principais características dos lncRNAs

Por muito tempo o interesse e a pesquisa no genoma foram centrados nos genes codificadores de proteína e seus produtos. Porém, avanços em genômica e transcriptômica mostram que a fração do genoma transcrita em RNA é maior do que se pensava anteriormente, sendo que grande parte desses transcritos corresponde a RNAs não-codificadores (ncRNA) (HARROW *et al.*, 2012). Embora ncRNAs funcionais como os RNAs ribossomais (rRNAs) e os RNAs de transferência (tRNAs)

já venham sendo estudados por muitas décadas, nos últimos anos os RNAs longos não codificadores (lncRNAs) têm recebido atenção crescente por sua importância funcional (MORAN *et al.*, 2012; SALVIANO-SILVA *et al.*, 2019).

lncRNAs são definidos como transcritos maiores que 200 nucleotídeos (nt) que não codificam proteínas. Essa definição é bastante útil em aspectos técnicos de análise de transcriptomas por diferenciar os lncRNAs de mRNAs e RNAs pequenos não codificadores. Porém, há ainda bastante debate em torno dos aspectos que os definem como um grupo de ncRNAs. O limite de 200 nucleotídeos se estabeleceu mais por uma questão metodológica, uma vez que esse é geralmente o tamanho obtido nos procedimentos de purificação de RNA realizados por colunas de sílica (KAPRANOV *et al.*, 2007). Além disso, alguns transcritos classificados como lncRNAs já foram reportados como codificadores de peptídeos, inclusive peptídeos funcionais (JI *et al.*, 2015). À parte dessas questões conceituais, é cada vez mais evidente que lncRNAs são RNAs reguladores da expressão gênica, podendo atuar na regulação transcricional, pós-transcricional, traducional e pós-traducional (RINN e CHANG, 2012).

lncRNAs e mRNAs são estruturalmente semelhantes e compartilham grande parte do processo de biogênese. Porém, os lncRNAs apresentam algumas peculiaridades: são menos conservados evolutivamente, de modo geral são menos abundantes na célula e seu perfil de expressão é mais tecido específico que o dos mRNAs, o que sugere diferenças na regulação transcricional e pós-transcricional dos lncRNAs (CABILLI *et al.*, 2011; DERRIEN *et al.*, 2012; ULITSKY and BARTEL, 2013; HEZRONI *et al.*, 2015). Uma das razões do menor nível de expressão é provavelmente a ocorrência de uma transcrição menos eficiente de promotores na direção não-codificadora (SAMUDYATA *et al.*, 2018). Há indícios de que o processo de splicing de lncRNAs também é menos eficiente que o de mRNAs. Embora em ambos o splicing ocorra concomitante à transcrição, sequências internas sinalizadoras de splicing são mais fracas em lncRNAs que em mRNAs (TILGNER *et al.*, 2012; MELÉ *et al.*, 2017). Por outro lado, vários lncRNAs com função conhecida e importante como XIST, FIRRE e MIAT já demonstraram possuir um splicing eficiente, sendo este processo crucial para o desempenho de suas funções (MELÉ *et al.*, 2017).

Assim como os mRNAs, os lncRNAs são em sua maioria transcritos pela RNA polimerase II, o que contrasta com parte dos RNAs pequenos que podem ser

transcritos pela RNA polimerase III. Porém, ainda que mRNAs e lncRNAs sejam transcritos pela mesma RNA polimerase, um estudo demonstrou que o padrão de fosforilação CTD da polimerase II difere significativamente entre as que transcrevem lncRNAs e mRNAs, classificando a RNA polimerase II que transcreve lncRNAs como desregulada (SCHLACKOW *et al.*, 2017).

Outra similaridade entre lncRNAs e mRNAs é o processo de adição da cap 7-metil-guanosina na ponta 5' que ocorre em todos os RNAs nascentes transcritos pela RNA polimerase II, sejam eles codificadores ou não codificadores. Já o processo de poliadenilação parece ocorrer principalmente em mRNAs e menos frequentemente em lncRNAs. Alguns estudos já descreveram os lncRNAs como poliadenilados (CABILI *et al.*, 2011; MELÉ *et al.*, 2017). Porém, outros estudos os classificaram como proporcionalmente menos poliadenilados que os mRNA ou ainda, majoritariamente não-poliadenilados (DERRIEN *et al.*, 2012; SCHLACKOW *et al.*, 2017).

2.2.2 Classificações dos lncRNAs

O repertório de lncRNA compreende milhares de transcritos, que possuem padrões e funções moleculares diversas. Dada essa heterogeneidade, a subdivisão em diferentes classes é uma tarefa difícil. Ainda assim, existem algumas classificações e caracterizações para os lncRNAs.

Uma das maneiras de classificar genes de lncRNAs em sugrupos considera a localização cromossômica desses em relação à de genes codificadores de proteínas. Desta forma, os lncRNAs são classificados em: lncRNAs intergênicos (lincRNAs), lncRNAs senso (intrônicos e de sobreposição), lncRNAs antisense (intrônicos ou transcritos antisense naturais) e lncRNAs bidirecionais (DERRIEN *et al.*, 2012). Os lincRNAs não se sobrepõem a genes codificadores de proteínas. Os lncRNAs senso se localizam na mesma fita e são transcritos no mesmo sentido que o gene codificador de proteína mais próximo, já os lncRNAs antisense são transcritos a partir da fita oposta a do gene codificador de proteína. Os lncRNAs bidirecionais estão localizados na fita oposta ao gene codificador de proteína, e tem o seu início de transcrição próximo ao início de transcrição do gene codificador, porém em sentido oposto (figura 2). No entanto, lncRNAs também podem pertencer a mais de uma dessas classes.

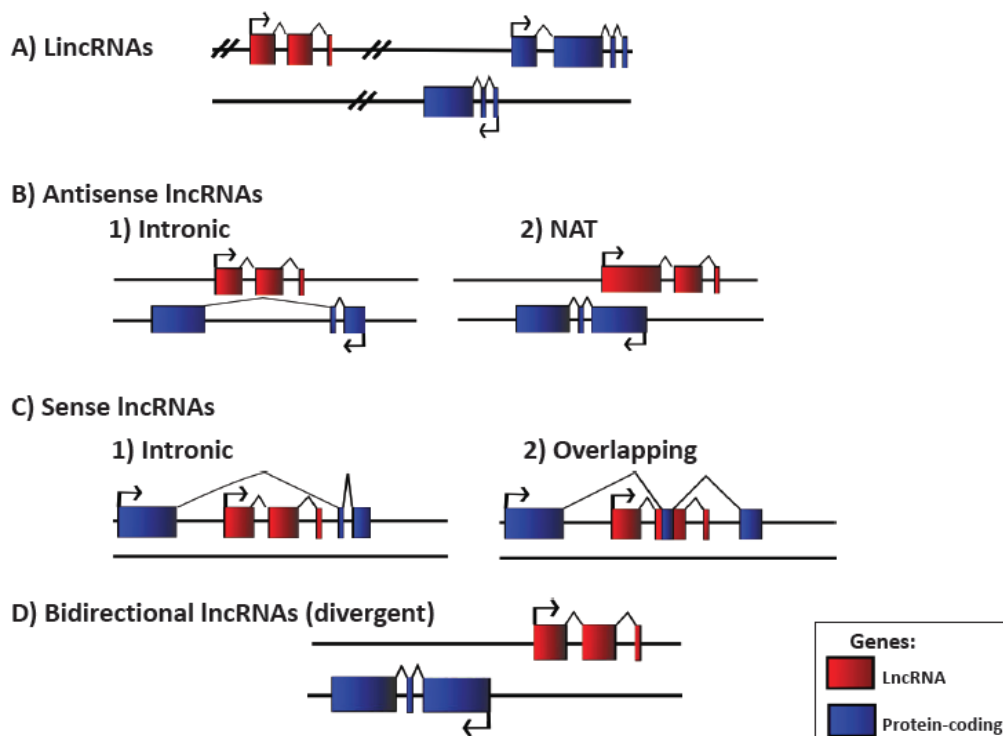


Figura 2 Classificação de lncRNAs em relação à posição do gene codificador de proteína mais próximo. (A) genes de RNA não-codificador intergênico longo (lincRNA) não se sobrepõem a genes codificadores de proteínas. (B) lncRNAs antisense são transcritos a partir da fita oposta do gene codificador de proteína com quem eles se sobrepõem (1) lncRNAs antisense intrônicos, quando o transcrito está completamente dentro dos limites do intron da fita oposta ou (2) transcritos antisense naturais (NAT) que se sobrepõem parcialmente ao gene codificador. (C) lncRNAs senso estão localizados na mesma fita e transcritos no mesmo sentido que o gene codificador de proteína. (D) lncRNAs bidirecionais ou divergentes estão localizados na fita antisense e seu local de início da transcrição (TSS) é perto do TSS do gene que codifica a proteína. Estes lncRNAs são transcritos no sentido inverso em relação ao gene que codifica a proteína. Legenda das cores: genes de codificação de proteínas (azul) e genes de lncRNA (vermelho). Fonte: HRDLICKOVA *et al.*, 2014a.

Os lncRNAs também podem ser classificados de acordo com a localização celular que ocupam, ou seja, eles podem ser citoplasmáticos, nucleares, ou ainda, ambos. A localização influencia suas interações e seus modos de ação. Com o objetivo de categorizar sistematicamente a localização subcelular de lncRNAs, Cabili e colaboradores (2015), utilizaram RNA-FISH de molécula única para quantificar e categorizar a localização de um conjunto representativo de 61 lncRNAs em três diferentes tipos de células. Na figura 3 é possível ver que existe uma variação gradual de expressão de lncRNAs desde os presentes exclusivamente em pontos específicos do núcleo, até os que são muito mais abundantes no citoplasma. Sendo que o grupo mais frequente, composto por 40% dos lncRNAs analisados pelos autores, são os lncRNAs com localização tanto nuclear quanto citoplasmática (CABILI *et al.*, 2015).

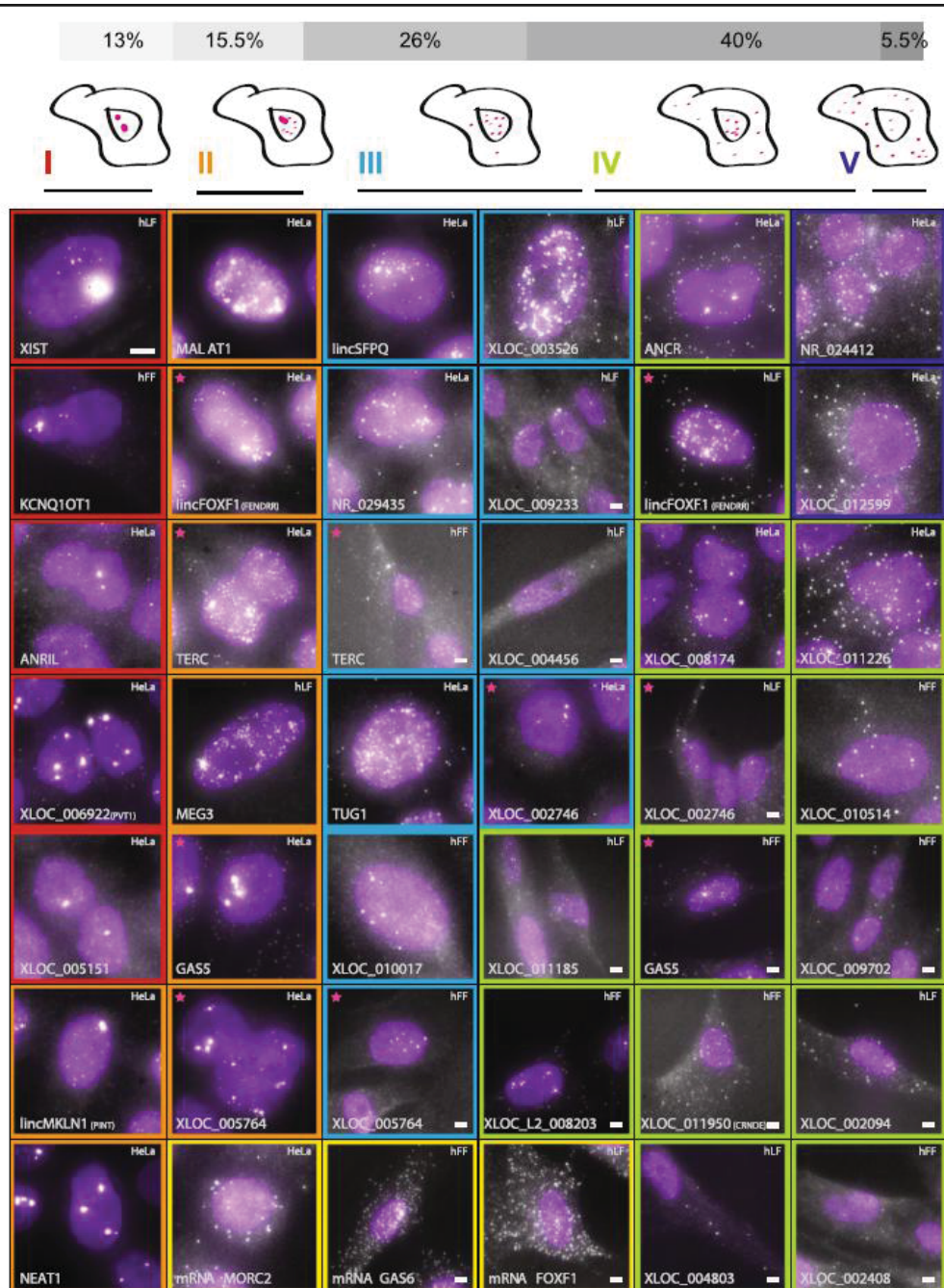


Figura 3. Variedade de padrões de localização celular dos lncRNAs. Micrografias de fluorescência de células expressando dezenas de lncRNAs para os quais existem sondas validadas. Os pares de tipo celular - lncRNA são classificados em localização celular tipos I a V (marcados pela cor da borda), sendo que: (I - vermelho) um ou dois grandes pontos nucleares; (II - laranja) tanto grandes pontos quanto moléculas únicas espalhadas pelo núcleo; (III - azul) predominantemente nuclear (sem pontos); (VI - verde) citoplasma e nuclear; e (V - roxo) predominantemente citoplásmica. Painel superior: fração de cada classificação para cada tipo em todo o conjunto de 70 pares de tipo celular - lncRNA válidos. Adaptado de Cabilli *et al.*, 2015.

Ainda, considerando o local de atuação, os lncRNAs se classificam, em geral, nos que atuam em *cis*, influenciando a expressão e/ou estado da cromatina de genes próximos e nos que atuam em *trans*, executando funções em genes mais distantes, no mesmo ou em outros compartimentos celulares (KOPP e MENDEL,

2018). Na próxima seção essas duas maneiras de ação serão mais bem descritas juntamente com o processo de regulação da expressão gênica por lncRNAs.

2.2.3 Mecanismos de ação e funções dos lncRNAs

As funções biológicas dos lncRNAs passaram a ser mais exploradas recentemente e em quase todos os sistemas biológicos estudados já foram identificados papéis críticos para os mesmos. lncRNAs têm importante atuação na regulação da expressão gênica através da interação com DNA ou RNA através de pareamento de bases de ácidos nucleicos ou através de estruturas conformacionais secundárias (MORAN *et al.*, 2012).

Os mecanismos moleculares de ação dos lncRNAs foram inicialmente organizados em quatro arquétipos por Wang e colaboradores (2011). (1) Arquétipo sinal - o lncRNA age como uma molécula sinalizadora, que em resposta a estímulos, pode ativar ou silenciar genes; (2) Arquétipo isca – lncRNAs que competem com outros RNAs ou proteínas por ligação a sequências ou estruturas e são considerados reguladores negativos; (3) Arquétipo guia – lncRNAs que se ligam a proteínas, formando complexos, e os transportam a sequências alvo específicas; (4) Arquétipo andaime - lncRNAs que se ligam a várias proteínas, aproximando-as e facilitando a interação das mesmas permitindo, por exemplo, a formação de complexos de ribonucleoproteínas. É importante ressaltar que muitos lncRNAs exercem mais de um dos mecanismos moleculares citados acima. Por exemplo, o lncRNA *HOTAIR* (RNA intergênico antisense HOX) que atua como arquétipo sinal, arquétipo guia e arquétipo andaime (revisado em HRDLICKOVA, 2014a).

Como já citado previamente os lncRNAs podem modular a transcrição de *loci* genômicos alvo em *cis* (genes vizinhos) ou em *trans* (genes distantes). Em uma revisão sobre o assunto, Kopp e Mendell (2018) destacaram três mecanismos de regulação em *cis* por lncRNAs.

Primeiro, o transcrito de lncRNA regula a expressão de genes vizinhos através da sua capacidade de recrutar fatores reguladores para o locus e/ou modular sua função. Segundo, o processo de transcrição e/ou splicing do lncRNA confere uma funcionalidade de regulação gênica que ocorre independentemente da sequência do transcrito de RNA, ou seja, a produção de um RNA de um locus de lncRNA é crítico para a regulação do gene local, mas essa funcionalidade do locus é independente do transcrito de lncRNA maduro ou sua sequência. Terceiro, a regulação depende

apenas de elementos de DNA dentro do promotor ou do gene de lncRNA e é completamente independente do lncRNA transcrito ou do seu processo de produção.

Os lncRNAs que atuam em *trans* também foram categorizados em três subgrupos. Primeiro, lncRNAs que regulam expressão de genes específicos em regiões distantes de seu sítio de transcrição, interagindo com promotores e acentuadores, ou proteínas ligadas a estas regiões, e influenciando o estado da cromatina e a atividade da RNA polimerase. Segundo, lncRNAs que influenciam a estrutura e organização nuclear para orquestrar a transcrição, o processamento de RNA e outras etapas na expressão gênica, bem como a regulação espacial sobre esses processos. Terceiro, lncRNAs que regulam a atividade ou abundância de proteínas e/ou outras moléculas de RNA com os quais interagem diretamente (KOPP E MENDEL *et al.*, 2018).

2.2.4 LncRNAs no sistema imune

A regulação da expressão de genes do sistema imunológico é fundamental para a capacidade de um organismo de gerar imunidade a patógenos e, ao mesmo tempo, evitar a autoimunidade a antígenos próprios. Estudos demonstram que os lncRNAs são expressos de maneira altamente célula-específica e modulam a diferenciação, a ativação e a função de células envolvidas na resposta imune inata e adaptativa (CHEN *et al.*, 2017).

Já foram descritos diversos lncRNAs envolvidos no processo de diferenciação de células tronco hematopoiéticas em células especializadas do sistema imune, como o lncRNA *Morbid* que é importante para diferenciação de células da linhagem mieloide e se encontra altamente expresso em neutrófilos, eosinófilos e monócitos (KOTZIN *et al.*, 2016).

Em células T CD4⁺ lncRNAs já foram reportados como moduladores da diferenciação das células T auxiliares, sendo, por exemplo, específicos de Th1, Th2 ou Th17 (RANZANI *et al.*, 2015; HUANG *et al.*, 2015). Em células Th0 que se diferenciarão em células Th1, o lncRNA *linc-MAF-4* reprime a expressão de MAF, um fator de transcrição de células Th2, para promover a diferenciação de células T em direção à linhagem de células Th1 (RANZANI *et al.*, 2015). O lncRNA *IFNG-AS1* também possui papéis em células do sistema imune; ele regula a expressão da citocina interferon gama em células Th1 (COLLIER *et al.*, 2014) e em células NK (STEIN *et al.*, 2019). Já o lncRNA *lincR-Ccr2-5'AS* regula positivamente a expressão

de genes de quimiocinas participantes na resposta Th2 (SPURLOCK *et al.*, 2015). Nas células Th17 o lncRNA *Rmrp* promove a interação da proteína DDX5 ao ROR γ t65, um fator de transcrição de células Th17 (HUANG *et al.*, 2015).

Além de participar dos múltiplos estágios do desenvolvimento de células do sistema imune, os lncRNAs participam também nas vias de resposta a patógenos. Na via de sinalização do NF- κ B, o lncRNA *NKILA* foi reportado como regulador negativo a nível pós-transcricional por interagir com o κ B (inibidor do NF κ B) (LIU *et al.*, 2015). O lncRNA *PACER* (do inglês *p50-associated COX-2 extragenic RNA*) age como uma isca na via de sinalização do NF- κ B. A expressão de *PACER* é induzida por estimulação com lipopolissacarídeo e promove a expressão do seu gene vizinho *PTGS2*, por sequestrar a subunidade p50 do NF- κ B, impedindo-a de inibir o promotor de *PTGS2*, gene que codifica COX-2 (ciclooxigenase 2), um conhecido mediador da resposta inflamatória (KRAWCZYK e EMERSON., 2014).

2.2.5 LncRNAs em doenças complexas

Uma vez que são importantes mediadores de diversos processos biológicos, a desregulação dos lncRNAs está intimamente ligada à patogênese das doenças humanas, como vários tipos de câncer, doenças inflamatórias e doenças autoimunes (revisado em CIPOLLA *et al.*, 2018 e DE OLIVEIRA *et al.*, 2019).

Em vários estudos foi descrita expressão alterada de lncRNAs em diferentes tipos de câncer. O lncRNA *MALAT1*, por exemplo, induz a expressão de genes associados a metástase (GUTSCHNER, 2013a). Em recentes estudos *in vitro* foi visto que células EBC-1 (células humanas de câncer de pulmão) podem ser inibidas pela interferência de oligonucleotídeos direcionados ao *MALAT1* (GUTSCHNER, 2013b). Outro exemplo do envolvimento de lncRNAs em câncer é o lncRNA *HOTAIR* encontrado superexpresso em tumores de mama e câncer colorretal (LU *et al.*, 2012; KOGO *et al.*, 2011). Já o *lincRNA-p21* funciona como um repressor de respostas transcricionais dependentes de p53, proteína codificada por um importante gene supressor tumoral (HUARTE *et al.*, 2010). Outras doenças podem também ser influenciadas por lncRNAs, o lncRNA *MIAT* (transcrito associado a infarto do miocárdio) tem sua expressão alterada por um SNP (polimorfismo de nucleotídeo único) localizado em sua sequência, o qual está associado com infarto do miocárdio (ISHII *et al.*, 2006).

Além disso, já foi demonstrado que 90% das variantes genéticas reportadas pelos estudos de associação de genoma total, GWAS (do inglês, *genome-wide associations studies*), estão localizadas em regiões não codificadoras, sugerindo um papel regulador dessas regiões e de seus produtos em doenças complexas (MAURANO *et al.*, 2012). Destas 90%, cerca de 10% das variantes estão mapeadas em genes de lncRNAs, o que sugere um papel importante para os lncRNAs na etiologia de doenças (RICAÑO-PONCE e WIJMENGA, 2013).

2.2.5.1 LncRNAs em doenças autoimunes

Nas doenças autoimunes, a desregulação de lncRNAs pode afetar a imunotolerância por influenciar a expressão de proteínas e RNAs envolvidos em vias de sinalização inflamatórias, vias de resposta humoral e celular, vias de morte celular, entre outras. A expressão e disponibilidade dos lncRNAs na célula pode também ser consequência da desregulação de vias de outros componentes importantes como citocinas e quimiocinas. O envolvimento de lncRNAs em diversas doenças autoimunes tem sido investigado, principalmente nas mais comuns e conhecidas, como nos exemplos que veremos nos próximos parágrafos.

Em pacientes de artrite reumatoide (AR) foi reportada expressão alterada de lncRNAs em células sanguíneas mononucleadas e em exossomos séricos através de análise por microarranjo. Os lncRNAs *HOTAIR*, *LUST*, *anti-NOS2A*, *MEG9*, *SNHG4*, *TUG1* e *NEAT1* estavam induzidos em pacientes de AR, enquanto os lncRNAs *mascrRNA*, *transcritos PR antisense*, *PRINS* e *HOXA3* estavam reprimidos (SONG *et al.*, 2015). O lncRNA *HOTAIR* foi reportado superexpresso também em soro de pacientes e, juntamente com o lncRNA *Cox2*, foi indicado como possível biomarcador para diagnóstico de AR (SHAKER *et al.*, 2019). Além disso, em monócitos de pacientes de AR a expressão de diversos lncRNAs é diferencialmente regulada pelas citocinas TNF- α e IL-6 (MULLER *et al.*, 2014).

Em psoríase, o lncRNA *PRINS* (do inglês, *psoriasis-associated RNA induced by stress*) tem expressão aumentada na epiderme não lesionada de pacientes quando comparada a epiderme lesionada dos mesmos e também a epiderme de indivíduos saudáveis. Foi sugerido que *PRINS* pode desempenhar um papel na susceptibilidade a psoríase em vez de atuar na manutenção das lesões psoriáticas (SONKOLY *et al.*, 2005). A análise global de expressão de lncRNAs comparando pele lesionada e não lesionada de pacientes de psoríase acrescentou ainda

centenas de outros lncRNAs que podem estar envolvidos na patogênese da psoríase (GUPTA *et al.*, 2015)

Na esclerose múltipla (EM), cerca de 10 lncRNAs foram encontrados reprimidos em PBMC (células mononucleares do sangue periférico) de pacientes quando comparado ao de controles. Dois desses, *NRON* e *TUG1*, foram replicados em amostra independente de pacientes e cotados como possíveis biomarcadores (FENOGLIO *et al.*, 2018). Recentemente, a análise integrada de lncRNAs, miRNAs e genes codificadores de proteínas evidenciou um novo lncRNA importante na esclerose múltipla, o lncRNA *ncRNA00201* é subexpresso em pacientes e tem como alvo diversos miRNAs, que por sua vez regulam genes envolvidos em funções biológicas afetadas pela EM (DOLCINO *et al.*, 2019).

Em lupus eritematoso sistêmico (LES), o lncRNA *NEAT1* tem expressão aumentada em monócitos de pacientes, sendo induzido por lipopolissacarídeo (LPS) pela ativação de p38. Além disso já foi demonstrado que o silenciamento de *NEAT1* reduz a expressão de um grupo de citocinas, entre elas, IL-6 e CXCL-10 (ZHANG *et al.*, 2016). Adicionalmente, os lncRNAs *IncDC*, *linc0597* e *GAS5* já foram reportados como subexpressos em pacientes de LES e cinco polimorfismos nesses genes foram investigados quanto a associação com a doença, porém não foram associados a menor ou maior predisposição à LES (LI *et al.*, 2017).

Em 2014, HRDLICKOVA *et al.* analisaram a expressão de lncRNAs cujos genes estão em *loci* associados a nove distúrbios imunes (doença de Crohn, doença celíaca, doença inflamatória intestinal, colite ulcerativa, artrite idiopática juvenil, cirrose biliar primária, psoríase, colangite esclerosante primária e artrite reumatoide) e constataram que esses lncRNAs estavam mais expressos em diferentes células do sistema imune, comparando com lncRNAs de outras partes do genoma. Além disso, a análise de coexpressão entre genes codificadores de proteínas e lncRNAs se mostrou promissora na aplicação de prever vias de sinalização que podem ser importantes na etiologia de distúrbios imunes (HRDLICKOVA *et al.*, 2014b).

3. JUSTIFICATIVA, HIPÓTESE E OBJETIVOS

Diversos SNPs em genes codificadores de proteína já foram associados à susceptibilidade diferencial ao pênfigo foliáceo endêmico, porém, polimorfismos em genes de lncRNAs ainda não haviam sido investigados em nenhuma das formas de pênfigo. Em diversas condições biológicas e patológicas, os lncRNAs e polimorfismos em seus genes têm sido reportados como sendo importantes. Estudá-los no pênfigo foliáceo endêmico contribui para compreender se eles são relevantes na patogênese da doença. A análise da expressão gênica em PFE abrangendo tanto mRNAs quanto lncRNAs, por sua vez, tem o potencial de gerar hipóteses e evidenciar fatores que afetam a predisposição ao PFE e/ou são afetados pelo transcurso da doença, ampliando nosso entendimento sobre o PFE.

Neste contexto, hipotetizamos que a expressão de RNAs, tanto os codificadores de proteínas (mRNAs) quanto os longos não codificadores (lncRNAs), influenciam ou são influenciados pela patogênese do pênfigo foliáceo endêmico; e que o polimorfismo de lncRNAs promove suscetibilidade diferencial a essa doença.

Assim sendo, o **objetivo geral** deste trabalho é identificar mRNAs e lncRNAs diferencialmente expressos em PFE além de explorar a participação dos lncRNAs pela análise de associação de polimorfismos presentes nos genes desses ncRNAs e o PFE.

Os **objetivos específicos** são:

- I - Verificar se SNPs localizados em genes de lncRNAs estão associados a susceptibilidade diferencial ao PFE.
- II - Inferir o possível efeito de SNPs (porventura associados) sobre a estrutura e função dos lncRNAs em que se localizam, através de análises *in silico*.
- III - Verificar se mRNAs e lncRNAs encontram-se diferencialmente expressos em células T CD4⁺ de pacientes de PFE comparados a indivíduos controle.
- IV - Validar os resultados encontrados no item III.
- V - Avaliar a expressão de genes diferencialmente expressos em células T CD4⁺ (de acordo com itens II e IV) em células B CD19⁺ de pacientes de PFE e indivíduos controles.

4. CAPÍTULO I

PRIMEIRO ARTIGO ORIGINAL – Estudo de associação caso-controle. Publicado no periódico científico British Journal of Dermatology em 17 de janeiro de 2019⁸. O conteúdo do artigo não é de acesso livre.

LONG NON-CODING RNA POLYMORPHISMS INFLUENCE SUSCEPTIBILITY TO ENDEMIC PEMPHIGUS FOLIACEUS

S.C. Lobo-Alves¹, D.G. Augusto^{1,2}, W.C.S. Magalhães^{3,4}, E. Tarazona-Santos⁴, M.F. Lima-Costa⁵, M.L. Barreto⁶, B.L. Horta⁷, R.C. de Almeida⁸, M.L. Petzl-Erler^{1*}

¹Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal do Paraná, Curitiba, Brazil; ²Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Ilhéus, Brazil; ³Núcleo de Ensino e Pesquisa - NEP, Instituto Mário Penna, Belo Horizonte, Minas Gerais, Brazil; ⁴Instituto de Ciências Biológicas, Departamento de Biologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; ⁵Instituto de Pesquisa Rene Rachou, Fundação Oswaldo Cruz, Belo Horizonte, MG, Brazil; ⁶Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, BA, Brazil; ⁷Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas, Pelotas, RS, Brazil; ⁸Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.

4.1 ABSTRACT

Background: Pemphigus foliaceus (PF) is an epidermal autoimmune disease, characterized by the presence of autoantibodies against the desmosomal protein desmoglein 1. Genetic and environmental factors contribute to PF, a complex disease that is endemic in Brazil and Colombia and neighbouring countries, and in Tunisia. Long non-coding RNAs (lncRNAs) may participate in gene regulation by interacting with DNA, proteins, and other RNAs. Dysregulation of lncRNAs has recently been recognized as an important co-player in the onset or progression of complex diseases. In addition, single-nucleotide polymorphisms (SNP) located in lncRNA genes have been associated with differential risk to cancer, autoimmunity and infection.

Objectives: Here, we aimed to investigate whether SNPs in lncRNA genes are associated with differential susceptibility to endemic PF.

Methods: We integrated data from the lncRNASNP database with genome-wide genotype data obtained for 229 patients and 6,681 controls. We tested the

association between endemic PF and 2,080 SNPs located in lncRNAs applying logistic regression.

Results: The most significantly associated SNP was *rs7144332* (OR = 1.63, $P = 2.8 \times 10^{-6}$), located in the lncRNA gene *AL110292.1*. Results for other five SNPs were suggestive of association ($P < 0.001$). *In silico* analysis indicated that five of the six SNPs impact transcription, three may influence lncRNA secondary structure, and three of them may alter microRNA-lncRNA interactions.

Conclusion: We showed, for the first time, that variation in lncRNA genes may influence pemphigus pathogenesis. Our findings highlight the importance of lncRNA variation in autoimmune and possibly other complex diseases and suggest polymorphisms for functional validation.

Keywords: Long non-coding RNA, pemphigus, fogo selvagem, genetic polymorphism, case-control association study, SNP, genetic susceptibility, autoimmunity.

&Lobo-Alves SC, Augusto DG, Magalhães WCS, Tarazona-Santos E, Lima-Costa MF, Barreto ML, Horta BL, de Almeida RC, Petzl-Erler ML. **Long non-coding RNA polymorphisms influence susceptibility to endemic pemphigus foliaceus.** Br J Dermatol. 2019 Jan 17. doi: 10.1111/bjd.17640.

5. CAPÍTULO II

RNA-SEQ ANALYSIS REVEALS NEW INSIGHTS INTO ENDEMIC PEMPHIGUS FOLIACEUS

Sara Cristina Lobo-Alves¹, Gabriel Adelman Cipolla¹, Rodrigo Coutinho de Almeida², Mareike Witte³, Hauke Busch³, Saleh Ibrahim³, Maria Luiza Petzl-Erler¹.

1. Laboratory of Human Molecular Genetics, Department of Genetics, Universidade Federal do Paraná, Curitiba, Paraná, Brazil.

2. Department of Biomedical Data Sciences, Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.

3. Lübeck Institute of Experimental Dermatology (LIED), University of Lübeck, Lübeck, Germany

5.1 Abstract

Pemphigus foliaceus (PF) is an epidermal autoimmune disease characterized by flaccid blisters and erosive lesions in the upper epidermal layers of the skin. The hallmark of this disease is the presence of pathogenic autoantibodies against desmoglein 1, an important cell-cell adhesion protein. In certain regions of Brazil, there is an endemic form of pemphigus foliaceus (EPF) also known as *fogo selvagem*. The pathogenesis of PF is not completely understood, yet, genetic and environmental factors generate and modulate the autoimmune response. The genetic factors already studied in pemphigus foliaceus include mainly protein-coding genes, and a more comprehensive approach is needed. In the present work, we employed transcriptome analysis to reveal insights into biological pathways and molecular mechanisms that regulate EPF pathogenesis. To this end, we sequenced messenger RNAs (mRNAs) and long non-coding RNAs (lncRNAs) from CD4⁺ T cell of four EPF patients not submitted to systemic treatment and five individuals without the disease. We identified 11 differentially expressed (DE) RNAs (q value < 0.1) and 313 genes suggestive of differential expression ($0.1 < q < 0.2$), including 32 lncRNA genes. Furthermore, we performed real time quantitative PCR (qPCR) to confirm the results for five genes in the same CD4⁺ T cell sample and evaluate their expression in an additional CD19⁺ B cell sample. For T cells four genes were confirmed differentially expressed, and for B cells, three genes were also detected DE. Moreover, we used previous results from an EPF microarray study to validate and corroborate the differential expression of eight protein-coding genes. The RNA-seq analysis provided insights into potential effects of both mRNA and lncRNA in pemphigus foliaceus. Gene set enrichment analysis helped us to clusterize the DE genes, and showed that most of them were involved in immune responses. Our findings reveal important previously unknown players in PF, like IL18RAP, S1PR5 and LEF1-AS1, and

highlight genes, that should be deeply investigated for clarify their involvement in the PF pathogenesis and their potential as novel therapeutic targets.

5.2 Introduction

Pemphigus foliaceus (PF) is a complex autoimmune disease, rare worldwide but endemic in certain countries, including Brazil, where PF reaches its highest incidence. PF is characterized by autoantibodies against the protein desmoglein 1 (DSG1), an important component of epidermal desmosomes (Jones *et al.*, 1984). The endemic form of the disease in Brazil is known as ‘fogo selvagem’ (FS) (meaning ‘wild fire’ in literal translation) due to the burning sensation caused by skin lesions.

The epidemiology of endemic PF (EPF) in Brazil indicates that particular environmental factors play a role in the etiology, acting together with predisposing genetic factors. The disease is associated with poor living conditions and exposure to ultraviolet radiation, among other factors. Specific suspect environmental triggers have been speculated to be involved in EPF pathogenesis, especially hematophagous insect bites (Lombardi *et al.*, 1992; Qian *et al.*, 2013). Molecular mimicry has been postulated as a potential initiating mechanism. In genetically susceptible individuals, a salivary protein or an infectious agent inoculated by hematophagous insect may initiate an antibody response that cross-reacts with DSG1. The genetic background of EPF is not entirely understood, but polymorphic variants of various genes are associated with EPF susceptibility, such as *HLA-DR*, *HLA-DQ* (Petzl-Erler & Santa Maria 1989, Pavoni *et al.*, 2003), *CD40L*, *CD40*, *BAFF* (Malheiros & Petzl-Erler, 2009) *IL4*, *IL6* (Pereira *et al.*, 2004), *PDCD1* (Braun-Prado & Petzl-Erler, 2007), *CD28*, *CD86*, *CTLA4* (Dalla-Costa *et al.*, 2010), *CIITA* (Piovezan & Petzl-Erler, 2013), *KIR* (Augusto *et al.*, 2012 Augusto *et al.*, 2015), *CD33*, *LAIR 1*, *LAIR 2* (Camargo *et al.*, 2016) and *KLRG1* (Cipolla *et al.*, 2016).

Gene expression profiling is the concerted analysis of the activity of thousands of genes, in order to understand their contribution to biological functions and conditions. Differential expression analysis comparing normal and pathological conditions produces gene expression signatures that are characteristic of the disease being studied and helps to elucidate its pathogenesis. In a previous study of EPF using the microarray technique, gene expression at the level of RNA was quantified in CD4⁺ T cells. Subgroups of EPF patients and healthy individuals were analyzed.

Comparisons of the expression profiles of untreated patients with the generalized form of the disease with the profiles obtained from healthy individuals, patients receiving immunosuppressive treatment, and patients with the localized form of the disease resulted in 135, 54 and 64 differentially expressed genes, respectively. Most of these genes were overexpressed, and their functions are related to lymphocyte adhesion and migration, apoptosis, cell proliferation, cytotoxicity and antigen presentation (Malheiros *et al.*, 2014).

The use of high-throughput sequencing technologies has some advantages over microarray approach, although both of them are very informative. One of the advantages of RNA sequencing (RNA-seq) is that it is exploratory and, as such, not limited to the quantitative assessment of previously described RNAs. Thus, RNA-seq has been instrumental in cataloging the diversity of novel transcript species including long non-coding RNAs (lncRNAs). lncRNAs are RNAs longer than 200 nucleotides that lack or have minimal protein coding potential (Harrow *et al.*, 2012). Initially, lncRNAs were considered 'transcriptional noise'; however, numerous reports in recent years have demonstrated that lncRNAs participate in diverse biological processes (Derrien *et al.*, 2012).

Here we report the first RNA-seq study in EPF, profiling mRNA and lncRNA to unravel differentially expressed genes between EPF patients and healthy controls. Using bioinformatics tools we aimed to interpret the complex relation of the various coding and long non-coding transcripts with the pathogenesis of EPF.

5.3 Material and methods

5.3.1 Sample collection and cohort

EPF patients were recruited at the 'Hospital Adventista do Pênfigo', in Campo Grande, Mato Grosso do Sul State; 'Hospital das Clínicas', USP, in Ribeirão Preto, São Paulo State; and 'Lar da Caridade Hospital do Fogo Selvagem', in Uberaba, Minas Gerais State. The control individuals were free of autoimmune diseases and mostly lived in the same endemic areas.

The sample comprised four subject groups: (I) patients with active disease and without treatment (n=4); (II) patients with active disease undergoing immunosuppressive therapy (n=6); (III) patients with inactive disease without treatment (n=6); and (IV) endemic controls (n=7) (table 1). The patients with active disease and without treatment and the endemic controls were chosen for RNA-seq;

the other patient groups were submitted to additional expression analysis by qPCR. The characterization of the individuals and whether the analysis was done through RNA sequencing and/or qPCR are summarized in table 1.

Patients and controls signed the informed consent in agreement with research participation and blood sample collection. Afterwards, 50 ml of venous blood were collected using tubes containing EDTA and stored at room temperature until cell isolation procedures. This study has been approved by the national ethics committee for research with human subjects (CONEP - Comissão Nacional de Ética em Pesquisa), under the protocol number 02727412.4.0000.0096, according to Brazilian Federal Laws.

Table 1. Sample characterization for RNA-seq and qPCR procedures

Individual Code	Gender	Age	Subgroup	CD4 ⁺ T cells		CD19 ⁺ B cells
				RNA seq	qPCR	qPCR
P325CEA	M	47	active disease w/o treatment	X	x	X
P362BCP	F	54		X	x	X
P391XQG	F	15		X	x	X
P414AIK	F	28		X	x	X
C691AXU	F	37	endemic control	X	x	X
C692LHF	M	66			x	X
C694UDW	F	22		X	x	X
C695NMV	F	40			x	
C696IDB	F	58		X	x	x
C697TEQ	M	40		X	x	x
C707RSY	F	41		X	x	x
P342GVN	F	56	active disease with treatment		x	x
P390BLW	F	32			x	x
P398HKU	F	49			x	x
P399LAF	M	33			x	x
P430VTH	F	15			x	x
P431RXF	M	56			x	x
P242MRS	F	62	inactive disease w/o treatment		x	x
P243ERS	M	70			x	x
P307LRS	F	31			x	x
P380MQJ	F	17			x	x
P422DSR	F	46			x	x
P508ASC	M	39			x	x

5.3.2 CD4⁺ T cell and CD19⁺ B cell isolation and RNA extraction

Peripheral blood mononuclear cells (PBMC) were isolated by ficoll-hypaque (Histopaque 1077; Sigma-Aldrich) density barrier centrifugation. The purification of CD4⁺ cells and CD19⁺ B cells was performed using a CD4 Multisort Kit and a CD19 Multisort kit (Miltenyi Biotec), respectively, and magnetic columns according to the manufacturer's guidelines. After CD4⁺ T cell isolation, the CD4 negative population was used for CD19⁺ B cell isolation. The degree of purity of the cell population was assessed by flow cytometry using CD3⁺+CD4⁺ doublestaining and CD19⁺ staining, respectively. The mean purity was of 97% for CD4⁺ T cells and of 88% for CD19⁺ cells.

RNA was extracted using *mirVana*TM *miRNA Isolation kit* (Ambion, Austin, TX, USA), following total RNA extraction protocol, with initial adaptations for trizol-lysed samples, following steps from the manufacturer's protocol.

5.3.3 RNA sequencing

CD4⁺ T cell RNA samples from patients with active disease without treatment (n=4) and from the control group (n=5) were individually sequenced in cooperation with Novogene Co., Ltd., (Hong Kong). The sequencing strategy started with ribosomal RNA removal (Ribo-ZeroTM Magnetic Kit); then, 250-300 bp insert strand specific libraries were prepared (NEB Next® UltraTM RNA Library Prep Kit). The Illumina Hi-seq platform was used to perform paired-end 150 bp sequencing.

5.3.4 RNA sequencing data analysis

Quality of RNA-seq reads was checked using FastQC version 0.11.5. Reads were pseudoaligned to the human transcriptome (Ensembl version 93, GRCh38) using both coding and long non-coding RNA genes. The quantification was achieved using Salmon version 0.11.0 with 30 bootstrap cycles. Principal component analysis was performed to help to assess the quality of RNA-seq data. Differential expression analysis was performed with the R package Sleuth version 0.29.0. Differentially expressed genes (DEGs) were determined based on likelihood ratio tests with Benjamini-Hochberg p-value correction. The established significance limit was adjusted *P* (q value) < 0.1. Besides, we considered genes with q value < 0.2 suggestive of differential expression. All statistical analyses were performed using R program version 3.4.4.

5.3.5 *In silico* functional analysis of differentially expressed genes

To identify functional categories of DEGs, enrichment analysis was performed using several databases. In order to cluster the whole set comprised by differentially expressed and suggestive of differential expression genes we used the Gene Set Enrichment Analysis (GSEA) with different pathway databases: Gene Ontology, Hallmark gene set, Consensus, Position Gene Sets and Transcription Factors. The gene lists were adopted from the Molecular Signatures Database (MSigDB). GSEA was performed using software gage version 2.28.2. We also used the integrative web-based software EnrichR (Kuleshov et al, 2016) that includes gene set libraries as an approach to rank terms of KEGG pathways, Panther and Human Gene Atlas. The STRING database (Szklarczyk *et al.*, 2017) was used to predict protein-protein interactions and explore the pathways related to the differentially expressed protein coding genes. The established significance limit was q value < 0.1.

5.3.6 Validation of differential expression results for T cells and additional expression evaluation of B cells

To validate our RNA-seq findings, we performed quantitative PCR (qPCR) on cDNA derived from CD4⁺ T cells of the PF patient subgroups and control group. We also extended the expression analysis for CD19⁺ B cells for the same individuals according to table 1. We chose five differentially expressed or suggestive of differential expression genes (*LEF1-AS1*, *LINC00921*, *GNLY*, *S1PR5*, and *IL18RAP*) for analysis by qPCR. The genes were chosen based on the most significant q values, biggest fold changes and/or their already known functions. Those already described as important in immune system or autoimmune diseases were prioritized.

The cDNA was synthesized from 1.5 µg of RNA in a final volume of 20 µL with random primers using the High capacity cDNA reverse transcription kit (ThermoFisher Scientific). Then, qPCR was performed by the DNA-binding dye Go taq® qPCR master mix (Promega) in an Ep Gradient Realplex Mastercycler (Eppendorf, Hamburg, Germany). Reactions were conducted in triplicates. The cycling conditions started with initial denaturation at 95°C, and then amplification for 40 cycles of 15 seconds at 95°C and 60 seconds at annealing/extension temperature optimized for each gene. Primer sequences and annealing temperatures are listed in table S1. The normalization factor was the β-actin (ACTB) gene. We analyzed the data using the $2^{-\Delta\Delta CT}$ method (Livak *et al.*, 2001). Statistical analyses were done by the non-

parametric Mann-Whitney test. P values of ≤ 0.05 were considered statistically significant.

5.3.7 Comparison with published microarray data.

To further validate our findings, we performed a comparison between our results and those of a previously published endemic pemphigus foliaceus microarray data (Malheiros *et al.*, 2014) to find the overlapping differentially expressed genes.

5.4 Results

5.4.1 Genes differentially expressed in EPF

To profile differentially expressed RNAs in EPF, we performed sequencing of total RNA of CD4⁺ T cell samples from four patients with active disease and five healthy controls. An average of 35 million paired-end reads per sample was obtained. The PCA of the RNA-seq data showed good discrimination between patients and controls (figure S1).

We identified 11 differentially expressed (DE) RNAs with a q value (adjusted P) < 0.1; of these, one, *AC007278.1*, is a lncRNA (table 2, figure 4). Moreover, given our small cohort size and the interpatient heterogeneity we also considered other 313 genes suggestive of differential expression with q value < 0.2; of these 31 are lncRNAs (Table S2).

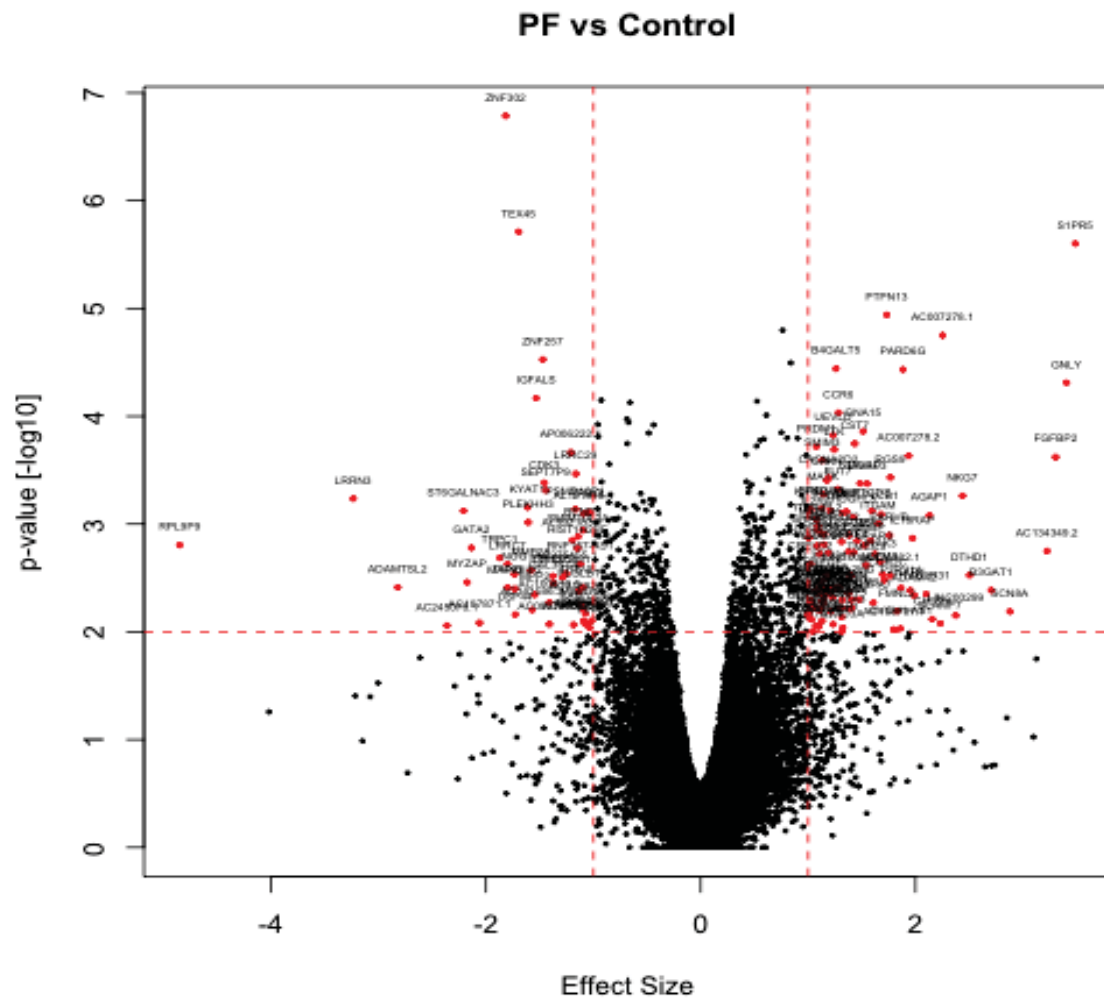


Figure 4. 'Volcano plot' of statistical significant DE genes against fold change between untreated active EPF patients and controls. Red dots represent up-regulated (right side) and downregulated genes (left side) genes, black dots represent non-differentially expressed genes.

Table 2. Differentially expressed genes in CD4⁺ T cells from EPF patients compared with controls (q value < 0.1)

Gene Symbol	q-value	Log fold change
<i>ZNF302</i>	0.0034	-1.82
<i>TEX45</i>	0.0171	-1.70
<i>S1PR5</i>	0.0171	3.49
<i>PTPN13</i>	0.0592	1.73
<i>SH2D2A</i>	0.0610	0.76
<i>AC007278.1</i>	0.0610	2.25
<i>ZNF257</i>	0.0760	-1.47
<i>LYAR</i>	0.0760	0.84

<i>B4GALT5</i>	0.0760	1.26
<i>PARD6G</i>	0.0760	1.88
<i>GNLY</i>	0.0916	3.41

q value: adjusted P. Negative fold change values mean underexpression and positive fold change values mean overexpression.

5.4.2 Validation in T cells and additional expression evaluation in B cells

Our qPCR results showed that four of the five tested RNAs (*LEF1-AS1*, *GNLY*, *S1PR5* and *IL18RAP*) were differentially expressed in pemphigus foliaceus CD4⁺ T cells in the same direction (fold change > 0, or < 0) as observed in our RNA-seq data (table 3 and figure 5). Furthermore, we replicated the results for genes *GNLY*, *S1PR5*, and *IL18RAP* in the B cell sample (table 3 and figure 6). In B cells, the *LEF1-AS1* was first analyzed for the same subjects used for RNA-seq; once expression levels did not differ significantly between patients with active disease and healthy controls, the other patient groups were not analyzed. Both RNA-seq and qPCR showed lower levels of lncRNA *LINC00921* in patients; however, the difference observed by qPCR was not statistically significant.

Table 3. Results from qPCR analysis for CD4⁺ T cells and B cells

Mann Whitney test		CD4 ⁺ T cells		CD19 ⁺ B cells	
Gene	Comparison	P value	Fold change	P value	Fold change
<i>LEF1-AS1</i>	Active w/o treat. x Controls	0.0121	0.53	0.6095	1.42
	Active with treat. x Controls	0.0012	0.24		
	Inactive w/o treat. X Controls	0.0047	0.54		
	Active w/o treat. x Inactive w/o treat	0.8286	0.99	not evaluated	
	Active with treat. x Inactive w/o treat.	0.0411	0.44		
	Active with treat x Active w/o treat	0.2571	0.44		
<i>GNLY</i>	Active w/o treat. x Controls	0.0121	10.99	0.0381	4.91
	Active with treat. x Controls	0.1014	2.29	0.9999	1.35
	Inactive w/o treat. x Controls	0.4394	2.18	0.7652	0.89
	Active w/o treat. x Inactive w/o treat.	0.3524	5.05	0.0727	5.53
	Active with treat. x Inactive w/o treat.	0.5714	3.51	0.6964	1.52
	Active w/o treat x Active with treat	0.1714	4.79	0.2381	3.64
<i>S1PR5</i>	Active w/o treat. x Controls	0.0061	3.94	0.0190	6.14

	Active with treat. x Controls	0.1817	2.01	0.7879	0.99
	Inactive w/o treat. x Controls	0.2331	2.15	0.1320	2.25
	Active w/o treat. x Inactive w/o treat	0.2571	1.83	0.0667	2.73
	Active with treat. x Inactive w/o treat	0.7879	0.93	0.0152	<u>0.44</u>
	Active w/o treat x Active with treat	0.7524	1.96	0.0095	6.16
<i>IL18RAP</i>	Active w/o treat. x Controls	0.0424	3.55	0.1714	2.33
	Active with treat. x Controls	0.6154	2.31	0.7879	0.99
	Inactive w/o treat. x Controls	0.7005	1.01	0.0022	0.31
	Active w/o treat. x Inactive w/o treat	0.0095	3.51	0.0095	7.45
	Active with treat. x Inactive w/o treat	0.2381	2.38	0.0260	3.19
	Active w/o treat x Active with treat	0.3524	1.54	0.3524	2.33
<i>LINC00921</i>	Active without treat. x Controls	0.8242	0.82	0.8286	0.95

5.4.3 Comparison of our data with published microarray results.

We compared our RNA-seq results with previously published EPF CD4⁺ T cells microarray results (Malheiros *et al.*, 2014), and eight protein-coding genes were found differentially expressed in both studies: *Interleukin 18 receptor accessory protein (IL18RAP)*, *integrin subunit alpha M (ITGAM)*, *granzyme A (GZMA)*, *natural killer cell granule protein 7 (NKG7)*, *cystatin F (CS7)*, *apolipoprotein B mRNA editing enzyme catalytic subunit 3G (APOBEC3G)*, *acetyl-CoA acyltransferase 2 (ACAA2)* and *fucosyltransferase 7 (FUT7)* (table 4). Among these, overexpression of *IL18RAP* was confirmed by our qPCR results; and overexpression of the *NKG7* gene was also detected in lesioned skin tissue from untreated EPF patients when compared with non-lesioned skin from the same patients by qPCR results from Malheiros *et al.* (2014).

Table 4. Genes differentially expressed both in the CD4⁺ microarray (p<0.01) and RNA-seq (q<0.2)

	Fold Change	
	Microarray	RNA-seq (log fold change)
<i>CST7</i>	1.8	1.4
<i>NKG7</i>	1.9	2.4
<i>APOBEC3G</i>	1.7	0.6
<i>ACAA2</i>	1.8	0.8

<i>GZMA</i>	1.5	1.7
<i>FUT7</i>	2.2	1.3
<i>ITGAM</i>	1.9	1.7
<i>IL18RAP</i>	1.7	2.0

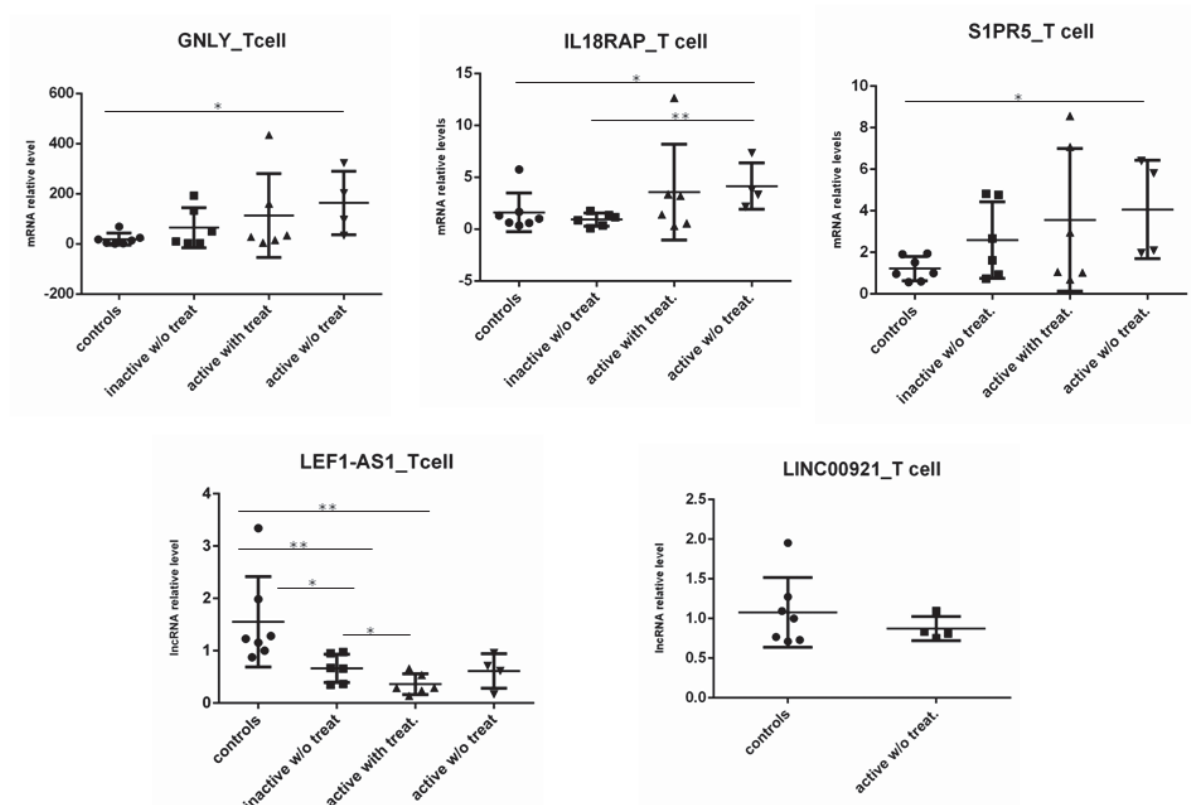


Figure 5: Dot plots showing the expression of *GNLY*, *S1PR5*, *IL18-RAP*, *LEF1-AS1* and *LINC00921* obtained by qPCRs performed for CD4⁺ T cells. P values ≤ 0.05 are considered statistically significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Horizontal lines mark the median values.

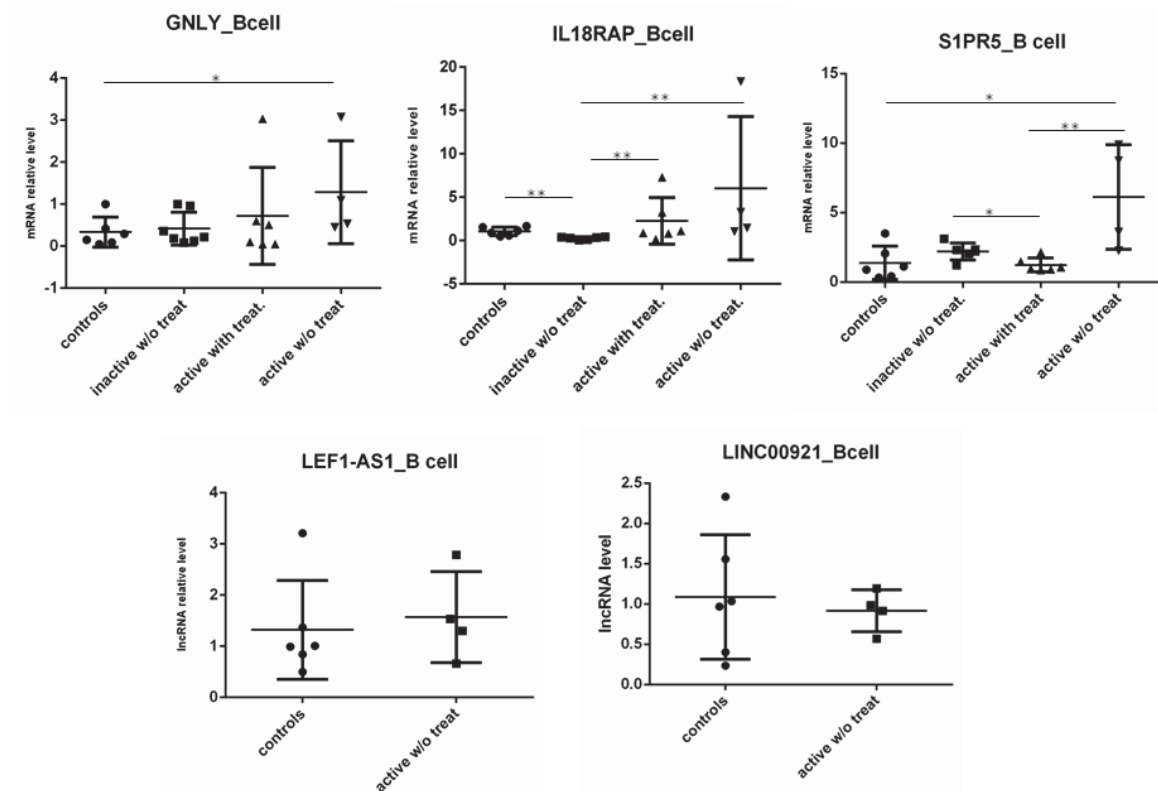


Figure 6: Dot plots showing the expression of *GNLY*, *S1PR5*, *IL18-RAP*, *LEF1-AS1* and *LINC00921* obtained by qPCRs performed for B cells. P values ≤ 0.05 are considered statistically significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Horizontal lines mark the median values.

5.3.3 *In silico* functional analysis of DEGs

Gene set enrichment analysis (GSEA) was performed to evaluate changes at the gene-set level (table S3). Gene Ontology (GO) analysis revealed that most DEGs were categorized into ‘immune response’, ‘inflammatory response’ and ‘cytokine-mediated signaling pathway’ as the three most significantly enriched terms among the 18 upregulated terms (q value < 0.1). None of the gene sets identified by GO terms was found downregulated. We also performed analyses for KEGG pathways with both EnrichR and STRING databases and identified the pathways ‘systemic lupus erythematosus’ and ‘TNF signaling pathway’ upregulated. In addition, we used EnrichR to report enriched terms for the Panther and Human Gene Atlas datasets for which “toll receptor signaling pathway” and “CD56⁺_NKCells” were found upregulated, respectively. The STRING database was used to construct a net of DEGs (figure 7). Our network has significantly more connections than expected ($P=2.37^{-13}$), which means that the proteins coded by the DEGs have more

interactions among themselves than expected for random protein sets of similar size, drawn from the genome. The enrichment indicates that the proteins are at least partially biologically connected, as a group.

Figure 7. Protein network for differentially expressed genes. Only connected nodes are shown.

5.4 Discussion

We compared the mRNA and lncRNA transcriptome of CD4⁺ T cells from pemphigus foliaceus patients and subjects without the disease. This approach allowed us to point out genes that are somehow involved in the pemphigus foliaceus pathogenesis.

We found 11 differentially expressed RNAs (q value < 0.1). Among these, *SH2D2A* and *GNLY* are notable because they have well described functions in immune responses and are likely to be important in autoimmune disease pathogenesis. The *SH2D2A* gene is overexpressed in EPF and encodes the T cell

specific adaptor protein (TSAd) that is involved in the control of signal transduction during T cell activation. TSAd interacts and modulates the activity of the Src family protein tyrosine kinase Lck (Hem *et al.*, 2015). Also, a polymorphism in the *SH2D2A* gene has previously been associated with multiple sclerosis and Sjögren syndrome (DAI *et al.*, 2001; Taylor *et al.*, 2017). The overexpression of *SH2D2A* in EPF can indicate the high activation status of the CD4⁺ T cells in the disease context. Possibly *SH2D2A* expression and function can affect or be affected by EPF pathogenesis.

The gene *GNLY* encodes granulysin, a protein usually present in cytotoxic granules of cytotoxic T lymphocytes and natural killer cells, where it has antimicrobial activity. In the present work, we show *GNLY* mRNA overexpression in CD4⁺ T cells of EPF patients. Diseases associated with *GNLY* include alopecia areata, in which *GNLY* serum levels are increased (ONO *et al.*, 2014); and psoriasis, in which granulysin expression in patients' skin increases with increasing clinical severity of the disease (ELGARHY *et al.*, 2015). We may suggest that increased levels of the *GNLY* protein correlate with the increased mRNA level reported for EPF in this study. As a proinflammatory factor, this protein is likely to contribute for the beginning and/or maintenance of inflammation and can alter the cytokine profile expression in EPF.

Other 313 genes were considered as suggestive ($0.1 < q \text{ value} < 0.2$) of differential expression in EPF. The GSEA helped us to systematically clusterize the DEG, based on prior knowledge. These genes perform functions in immune and inflammatory response pathways that are enriched in EPF patients, as well as pathways mediated by cytokines. Interestingly, the differentially expressed genes did not point to downregulated terms and pathways in any of the databases analyzed (GO, Consensus, Hallmark, and Transcription Factors).

We validated by qPCR the differential expression of one lncRNA gene (*LEF1-AS1*) and three protein-coding genes (*IL18RAP*, *GNLY*, *S1PR5*). The *interleukin 18 receptor accessory protein* (*IL18RAP*) gene was overexpressed in the RNA-seq analysis. Its overexpression was corroborated by the qPCR analysis. Moreover, the *IL18RAP* gene was overexpressed in EPF patients with generalized lesions when compared with patients with localized lesions (Malheiros *et al.*, 2014). The *IL18RAP* protein enhances IL18-binding by the IL-18 receptor and plays a role in signaling by the IL-18 cytokine. Polymorphisms in this gene are associated with Crohn's disease and inflammatory bowel disease, celiac disease, and leprosy (Festen *et al.*, 2011;

Yang *et al.*, 2015; Liu *et al.*, 2012). Furthermore, the lncRNAs *AC007278.1* and *AC007278.2* that overlap introns of *IL18RAP* were overexpressed in the present study (fold change 2.25 and 1.94, respectively). These lncRNAs are not well known. However, coexpression relationships between overlapping protein-coding and lncRNA genes represent a rich source of information, and it has been shown that functionally interacting genes tend to show similar expression profile (Lee *et al.*, 2004; Liao *et al.*, 2011). According to the GTEx Portal, sequence variants of *AC007278.2* influence the expression of *IL18RAP*, *IL18R1* (interleukin 18 receptor 1) and possibly also *IL18* (GTEx consortium - www.gtexportal.org). It is reasonable to expect that also the levels of these lncRNAs modify the IL-18 signaling pathway and have downstream effects in the immune response. Besides, we observed increased expression of the *IL18R1* mRNA in EPF. *IL18R1*, together with *IL18RAP* contribute to the IL-18 receptor complex formation and consequently to IL-18 mediated signal transduction. Collectively, the overexpression of *IL18RAP* in two independent samples, the overexpression of lncRNAs genes that overlap *IL18RAP*, and the increased *IL18R1* expression led us to hypothesize a crucial role for the IL-18 signaling pathways in the EPF pathogenesis.

The lncRNA gene *LEF1 antisense RNA 1 (LEF1-AS1)* is antisense to lymphoid enhancer binding factor 1 (*LEF1*) and was found underexpressed in PF patients. Several lncRNAs may regulate closely linked genes of which they are antisense transcripts (Lee, 2012; Villegas *et al.*, 2015). We did not observe differential expression for gene *LEF1* at the mRNA level; despite this, we hypothesize that the lncRNA *LEF1-AS1* may regulate nearby genes, as *LEF1*. The *LEF1* gene encodes a transcription factor that binds to a functionally important site in the T-cell receptor alpha enhancer, resulting in maximal enhancer activity. This transcription factor is also critical for sustaining regulatory T cell suppressive functions and preventing loss of self-tolerance (Xing *et al.*, 2019). Moreover, the lncRNA *LEF1-AS1* has been found deregulated in some cancer types (Congrains-Castillo *et al.*, 2019 Wang *et al.*, 2019; Liu *et al.*, 2019).

The gene sphingosine-1-phosphate receptor 5 (*S1PR5*) presented the highest fold change (3.49) among the overexpressed genes in the EPF RNA-seq analysis. This gene codes a receptor for the lysosphingolipid sphingosine 1-phosphate (S1P). S1P is a bioactive lysophospholipid that elicits diverse physiological effects on many cell types. As a chemoattractant, S1P guides T cell circulation among secondary

lymphoid organs, where T cells search for antigens. The receptor S1PR5 is a member of the G protein-coupled receptor family (S1PR1-S1PR5) that has been implicated in inflammation, immune cell trafficking, and T cell differentiation, among other cellular processes (Chi, 2010). S1PR5 is essential for NK cell migration (Drouillard *et al.*, 2018) and, together with other S1P receptors, has been explored as a potential therapeutic target for diseases as multiple sclerosis and systemic lupus erythematosus. Clinical and animal model studies have shown beneficial effects of S1P modulators in MS and SLE, which seem promising also for treatment of other autoimmune diseases (Faisser and Gold, 2018; Meadows *et al.*, 2018).

The expression level of the lncRNA *LINC00921* did not differ significantly between patient and control CD4⁺ T cells by the qPCR method. In the RNA-seq results, it was suggestively underexpressed in EPF. Therefore, likely because of the small sample size and the interindividual heterogeneity, it was not possible to conclude about the level of this lncRNA in the disease.

We focused mainly on the CD4⁺ T cell RNA profile, but we used qPCR to also investigate the CD19⁺ B cell expression for genes *GNLY*, *S1PR5*, *IL18RAP*, *LEF1-AS1*, and *LINC00921*. The *S1PR5* and *GNLY* expression levels were similar to those seen in T cells: both were overexpressed in untreated patients with active disease when compared with controls. As in CD4⁺ T cells, also in B cells *IL18RAP* was overexpressed, however, the difference between untreated patients with active disease and controls was not statistically significant. Still, *IL18RAP* was overexpressed in treated and untreated patients when compared with inactive EPF patients. The results for B cells support a possibly broader role for *GNLY*, *S1PR5* and *IL18RAP* in the EPF autoimmunity context, once they can have important functions in both T and B cells. On the other hand, the lncRNA *LEF1-AS1* was underexpressed in CD4⁺ T cells, but no difference was recognized in B cells.

In order to further validate our findings, we compared our RNA-seq results with results from a CD4⁺ T cell transcriptome study of EPF by microarray analysis (Malheiros *et al.*, 2014). Eight differentially expressed genes in the microarray were also differentially expressed or suggestive of differential expression in the RNA-seq results. Some of the genes have functions related to immune response: *IL18RAP* (discussed above), *ITGAM*, *GZMA*, *NKG7*, and *CS7*. Other genes have diverse cellular functions as *APOBEC3G*, *ACAA2*, and *FUT7*. It is worthy highlighting that among these genes, two (*NKG7* and *GZMA*) were validated by qPCR by Malheiros *et*

al. (2014), who also found *NKG7* overexpressed in lesioned skin of EPF patients (Malheiros *et al.*, 2014). Moreover, polymorphism of the *ITGAM* gene has been studied in pemphigus foliaceus, and the *rs12928810* was found associated (Bumiller-Bini *et al.*, 2018). The genes differentially expressed in both the microarray and RNA-seq studies are strong candidates to have a function in pemphigus pathogenesis and should be further analyzed in functional studies.

5.5 Conclusions and perspectives

Overall, this study identified various mRNAs and lncRNAs differentially expressed in CD4⁺ T cells of endemic pemphigus foliaceus patients. Many DE RNAs had never been studied in EPF or even in other complex diseases before and are novel potential players in disease development. Moreover, some of the DE mRNAs had previously been found differentially expressed in EPF in a microarray analysis, so that the present study strongly corroborated their importance in EPF pathogenesis.

Future studies can aim on deciphering if the DE mRNAs are differentially expressed also at the protein level and how these proteins are implicated in EPF condition. Functional assays should be performed to reveal the mechanisms through which DE lncRNAs act to regulate protein coding genes at both the transcriptional and translational level.

5.6. References

Augusto DG, Lobo-Alves SC, Melo MF, *et al.* Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. *PloS One* 2012; 7:7:e39991. doi: 10.1371/journal.pone.0039991

Augusto DG, O'Connor GM, Lobo-Alves SC, *et al.* Pemphigus is associated with KIR3DL2 expression levels and provides evidence that KIR3DL2 may bind HLA-A3 and A11 in vivo. *Eur J Immunol* 2015; 45:2052-60.

Braun-Prado K, Petzl-Erler ML. Programmed cell death 1 gene (PDCD1) polymorphism and pemphigus foliaceus (fogo selvagem) disease susceptibility. *Genetics Molecular Biology* 2007; 30: 314.

Bumiller-Bini V, Cipolla GA, de Almeida R C, Petzl-Erler ML, Augusto DG, Boldt A. Sparking Fire Under the Skin? Answers From the Association of Complement Genes With Pemphigus Foliaceus. *Frontiers in immunology*, 2018: 9, 695. doi:10.3389/fimmu.2018.00695

Camargo CM, Augusto DG, Petzl-Erler ML. Differential gene expression levels might explain association of LAIR2 polymorphisms with pemphigus. *Hum Genet* 2016; 135:233-44.

- Chi H. Sphingosine-1-phosphate and immune regulation: trafficking and beyond. *Trends in pharmacological sciences* 2010; 32: 16-24.
- Cipolla GA, Park JK, de Oliveira LA, *et al.* A 3'UTR polymorphism marks differential KLRG1 mRNA levels through disruption of a miR-584-5p binding site and associates with pemphigus foliaceus susceptibility. *Biochem Biophys Acta* 2016; 1859:1306-13.
- Congrains-Castillo A, Niemann FS, Santos Duarte AS, Olalla-Saad ST. LEF1-AS1, long non-coding RNA, inhibits proliferation in myeloid malignancy. *J Cell Mol Med.* 2019; 23:3021–3025. doi:10.1111/jcmm.14152
- Dai KZ, Harbo HF, Celius EG, Oturai A, Sorensen PS, Ryder LP, Datta P, Svejgaard A, Hillert J, Fredrikson S, Sandberg-Wollheim M, Laaksonen M, Myhr KM, Nyland H, Vartdal F, Spurkland A. The t cell regulator gene sh2d2a contributes to the genetic susceptibility of multiple sclerosis. *Genes Immun.* 2001; 2:263–268.
- Dalla-Costa R, Pincerati MR, Beltrame MH, Malheiros D, Petzl-Erler ML. Polymorphisms in the 2q33 and 3q21 chromosome regions including T cell coreceptor and ligand genes may influence susceptibility to pemphigus foliaceus. *Human Immunology* 2010; 71: 809
- Derrien T, Johnson R, Bussotti G, *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012; 22: 1775–1789. doi:10.1101/gr.132159.111
- Drouillard A, Mathieu A-L, Marçais A, Belot A, Viel PharmD S, Mingueneau M, *et al.* S1PR5 is essential for human natural killer cell migration toward sphingosine-1 phosphate. *J Allergy Clin Immunol.* 2018; 141:2265–8.
- Elgarhy LH, Shareef MM, Moustafa SM. Granulysin expression increases with increasing clinical severity of psoriasis. *Clin Exp Dermatol.* 2015; 40:361-6.
- Faissner S, Gold R. Oral Therapies for Multiple Sclerosis. *Cold Spring Harb Perspect Med.* 2019 2;9(1). pii: a032011. doi: 10.1101/cshperspect.a032011.
- Festen EA, Goyette P, Green T, *et al.* A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet.* 2011;7(1):e1001283. Published 2011 Jan 27. doi:10.1371/journal.pgen.1001283
- GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nature genetics*, 45(6), 580–585. 2013.
- Harrow J, Frankish A, Gonzalez JM, *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* 2012; 22: 1760-74.
- Hem CD, Sundvold-Gjerstad V, Granum S, *et al.* T cell specific adaptor protein (TSAd) promotes interaction of Nck with Lck and SLP-76 in T cells. *Cell Commun Signal.* 2015;13:31. doi:10.1186/s12964-015-0109-7
- Jones JC, Arn J, Stachelin LA, Goldman RD. Human antibodies against desmosomes: possible causative factors in pemphigus. *Proceedings of the National Academy of Sciences of the United States of America* 1984; 81:2781-5.

Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016; 44:W90–W97. doi:10.1093/nar/gkw377

Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P. Coexpression analysis of human genes across many microarray data sets. *Genome Res.* 2004; 14:1085–94. doi: 10.1101/gr.1910904

Lee JT. Epigenetic regulation by long noncodingRNAs. *Science* 2012; 338: 1435–1439

Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, Zhao G, Luo H, Bu D, Zhao H, Skogerboe G, Wu Z, Zhao Y. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res.* 2011; 39: 3864–78. doi: 10.1093/nar/gkq1348.

Liu H, Irwanto A, Tian H, et al. Identification of IL18RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases. *Am J Hum Genet.* 2012;91(5):935–941. doi:10.1016/j.ajhg.2012.09.010

Liu DC, Song LL, Liang Q, Hao L, Zhang ZG, Han CH. Long noncoding RNA LEF1-AS1 silencing suppresses the initiation and development of prostate cancer by acting as a molecular sponge of miR-330-5p via LEF1 repression. *J. Cell Physiol.* 2019 10.1002/jcp.27893

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta C_t$ method. *Methods* 2001; 25: 402-8.

Lombardi C, Borges PC, Chaul A, Sampaio SAP, Rivitti EA, Freidman H, et al. Environmental risk factors in endemic pemphigus foliaceus (Fogo selvagem). "The Cooperative Group on Fogo Selvagem Research". *The Journal of Investigative Dermatology* 1992; 98:847-50.

Malheiros D, Petzl-Erler ML. Individual and epistatic effects of genetic polymorphisms of B-cell co-stimulatory molecules on susceptibility to pemphigus foliaceus. *Genes Immunology* 2009; 10:547

Malheiros D, Panepucci RA, Roselino AM, Araújo AG, Zago MA, Petzl-Erler ML. Genome-wide gene expression profiling reveals unsuspected molecular alterations in pemphigus foliaceus. *Immunology.* 2014; 143:381–395. doi:10.1111/imm.12315

Meadows T.K. R., Steinberg, M. W., Clemons, B., Stokes, M. E., Opiteck, G. J., Peach, R., & Scott, F. L. Ozanimod (RPC1063), a selective S1PR1 and S1PR5 modulator, reduces chronic inflammation and alleviates kidney pathology in murine systemic lupus erythematosus. *PloS one* 2014, 13(4), e0193236. doi:10.1371/journal.pone.0193236

Ono S, Otsuka A, Yamamoto Y, Kataoka TR, Koyanagi I, Miyachi Y, Kabashima K. Serum granulysin as a possible key marker of the activity of alopecia areata. *J Dermatol Sci.* 2014; 73:74-9.

Pavoni DP, Roxo VMMS, Marquart Filho A, Petzl-Erler ML. Dissecting the associations of endemic pemphigus foliaceus (Fogo Selvagem) with HLA-DRB1 alleles and genotypes. *Genes and Immunity* 2003, 4: 110-6.

Pereira NF, Hansen JA, Lin MT, *et al.* Cytokine gene polymorphisms in endemic pemphigus foliaceus: a possible role for IL6 variants. *Cytokine* 2004; 28:233-41.

Petzl-Erler ML, Santamaria J. Are HLA class II genes controlling susceptibility and resistance to Brazilian pemphigus foliaceus (fogo selvagem)? *Tissue Antigens* 1989; 33: 408-414.

Piovezan BZ, Petzl-Erler ML. Both Qualitative and quantitative genetic variation of MHC Class II molecules may influence susceptibility to autoimmune diseases: the case of endemic pemphigus foliaceus. *Human Immunology* 2013;74:1134-40.

Qian Y, Jeong JS, Maldonado M, *et al.* Cutting Edge: Brazilian Pemphigus Foliaceus Anti-Desmoglein 1 autoantibodies cross-react with sandy fly salivary LJM11 antigen. *The journal of Immunology* 2013; 189:1535-39.

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017; 45:D362–D368.

Taylor KE, Wong Q, Levine DM, *et al.* Genome-Wide Association Analysis Reveals Genetic Heterogeneity of Sjögren's Syndrome According to Ancestry. *Arthritis Rheumatol.* 2017; 69:1294–1305. doi:10.1002/art.40040

Villegas VE, Zaphiropoulos PG. Neighboring gene regulation by antisense long non-coding RNAs. *Int J Mol Sci.* 2015; 16:3251–3266.

Xing S, Gai K, Li X, Shao P, Zeng Z, Zhao X, Zhao X, Chen X, Paradee WJ, Meyerholz DK, Peng W, Xue HH. Tcf1 and Lef1 are required for the immunosuppressive function of regulatory T cells. *J Exp Med.* 2019 216:847-866.

Wang A, Zhao C, Gao Y. *et al.* LEF1-AS1 contributes to proliferation and invasion through regulating miR-544a/ FOXP1 axis in lung cancer. *Invest New Drugs* 2019 <https://doi.org/10.1007/s10637-018-00721-z>.

Yang G, Zhang B, Huang W, Zhang N, Dong F, Jing L, Wang M *et al.* Systematic review and meta-analysis of the association between IL18RAP rs917997 and CCR3 rs6441961 polymorphisms with celiac disease risks. *Expert Rev Gastroenterol Hepatol.* 2015; 9:1327-38. doi: 10.1586/17474124.2015.1075880.

6. Capítulo III

Artigo de revisão publicado em compartilhamento de primeira autoria com Amanda Salviano Silva no periódico científico *Non-coding RNA* em março de 2018.

Salviano-Silva A, Lobo-Alves SC, Almeida RC, Malheiros D, Petzl-Erler ML. Besides Pathology: Long Non-Coding RNA in Cell and Tissue Homeostasis. *Noncoding RNA* 2018; **4**(1):3. doi: 10.3390/ncrna4010003.

Review Article BESIDES PATHOLOGY: LONG NON-CODING RNA IN CELL AND TISSUE HOMEOSTASIS.

Amanda Salviano-Silva^{1&}, Sara Cristina Lobo-Alves^{1&}, Rodrigo Coutinho de Almeida¹, Danielle Malheiros¹, Maria Luiza Petzl-Erler^{1,*}

¹Laboratory of Human Molecular Genetics, Department of Genetics, Universidade Federal do Paraná, Curitiba, Brazil.

[&]These joint-first authors contributed equally to this review.

*Corresponding author:

Prof. Dr. Maria Luiza Petzl-Erler
Departamento de Genética
Universidade Federal do Paraná
Caixa Postal 19071
81531-980 Curitiba, Brazil
(+55 41) 3361-1746; 3361-1724
perler@ufpr.br; mlperler@gmail.com

Abstract

A significant portion of mammalian genomes corresponds to genes that transcribe long non-protein-coding RNAs (lncRNAs). Along the last decade, the number of studies concerning the roles played by lncRNAs in different biological processes has increased considerably. That intense interest in lncRNA has produced a major shift in our understanding of gene and genome regulation and structure. It became apparent that lncRNAs regulate gene expression through diverse mechanisms. These RNAs function as transcriptional or post-transcriptional regulators through binding to histone-modifying complexes, to DNA, to transcription factors and other DNA binding proteins, to RNA polymerase II, to mRNA, or through modulation of microRNA function. Often the lncRNA transcription itself rather than the lncRNA product appears to be regulatory. In this review, we highlight studies identifying lncRNA in homeostasis of various cell and tissue types, or demonstrating their effects in expression of protein-coding or other non-coding RNA gene.

Key-words: long non-coding RNA; homeostasis; physiological regulatory mechanisms; gene expression; gene regulation; transcriptome.

Short title: lncRNA in homeostasis.

6.1.Introduction

Until few years ago, non-coding RNAs (ncRNA) were an unexplored component of the transcriptome considered as transcriptional noise. The only well characterized ncRNAs were the rRNAs, tRNAs, and small RNAs involved in processing of primary transcripts, being the messenger RNAs (mRNA) thought as the most numerous and diversified class of RNAs. However, technological advances, including large-scale sequencing and bioinformatics tools, have revealed that ncRNAs are the major part of the total set of transcripts [1,2]. This finding aroused the interest in studying ncRNAs, and among them the long non-coding RNAs (lncRNAs).

LncRNAs are defined as a heterogeneous class of ncRNAs longer than 200 nucleotides [1]. These transcripts are located predominantly in the nucleus and secondarily in the cytoplasm, being mostly polyadenylated and transcribed by RNA polymerase II. LncRNAs genes may present tissue- and cell-specific expression, and differ from mRNAs genes for having longer and fewer exons [3]. According to genomic position relative to the nearest coding gene, lncRNAs genes can be categorized as intergenic (lincRNAs), bidirectional (divergent), sense and antisense (intronic or overlapping) [1] (figure 8A). According to regulatory mechanisms on expression of target genes, the lncRNAs can act as signals, decoys, scaffolds, guides, competing endogenous RNAs (ceRNAs) and others [4] (figure 8B).

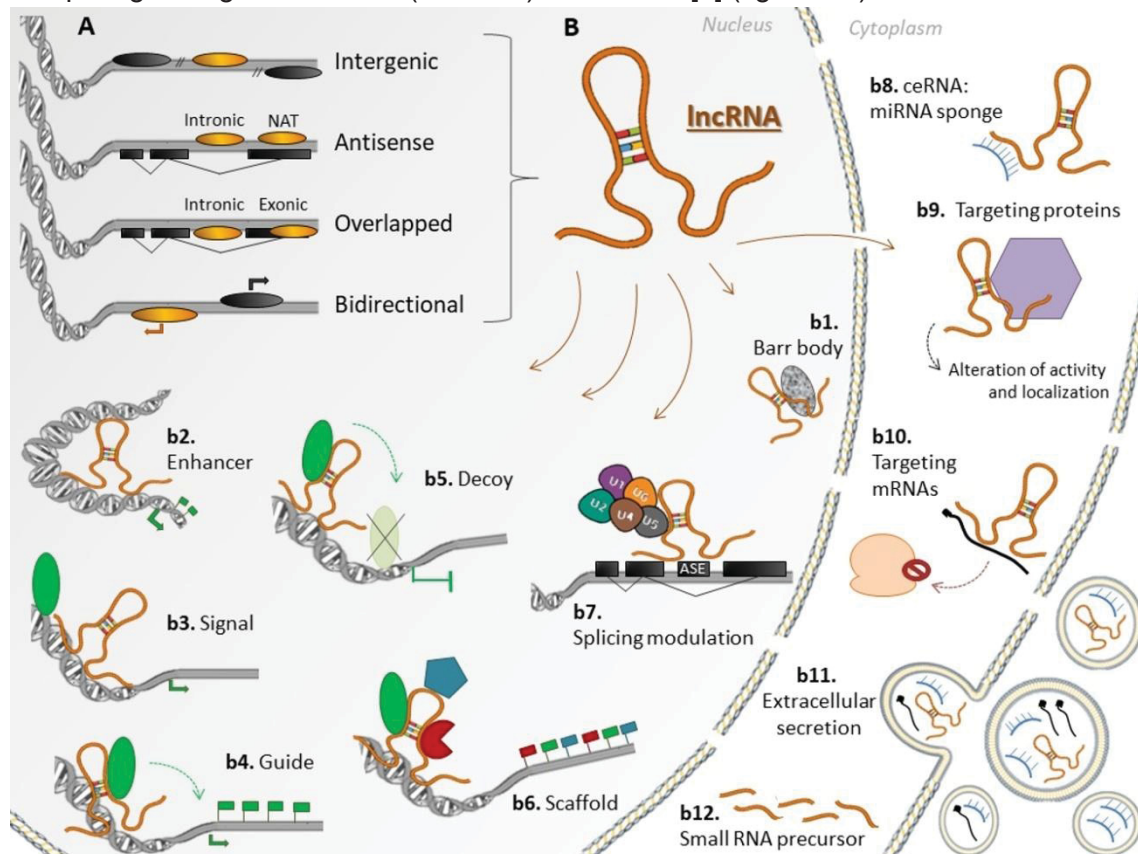


Figure 8. Genomic location relative to protein-coding genes, and regulatory mechanisms of lncRNAs in the nucleus, cytoplasm and extracellular compartments. **(A)** Nomenclature of lncRNAs genes (gold ellipses), according to their genomic location relative to the nearest coding gene (black ellipses) and/or to exons of coding genes (black rectangles). **(B)** lncRNAs regulatory mechanisms: (b1) lncRNA Xist, as component of Barr body in females; (b2) Acting as enhancers, inducing transcription *in cis* or *in trans*; (b3) as molecular signals, activating or silencing gene expression through signalling to regulatory pathways; (b4) Guiding proteins (in general, chromatin modifiers) to specific target sites; (b5) decoy to regulatory proteins, such as transcription factors and chromatin modifiers, blocking their binding to DNA; (b6) as scaffolds, binding different proteins and forming ribonucleoprotein (RNP) complexes, which also affect gene expression; (b7) modulating alternative splicing of primary transcripts; (b8) as competing endogenous RNA (ceRNA), serving as a sponge for miRNAs, blocking their effect; (b9) targeting proteins, forming molecular complexes which can block or induce functional effects, or

even alter their location in the cell; (b10) targeting mRNAs, inhibiting their translation in ribosomes. In addition, lncRNAs can be (b11) transferred to other cells by extracellular vesicles (EVs), where they can produce effects; (b12) be precursors of miRNAs and other regulatory small RNA. A lncRNA can act by multiple regulatory mechanisms, in both the nucleus and/or in the cytoplasm. The b11 itself is not exactly a regulatory feature, however, release of these functional lncRNAs through EVs is a way of regulating gene expression in other tissues. ASE – alternatively spliced exon.

Many studies have demonstrated involvement of lncRNAs in processes of prime relevance for homeostasis [5]. One of the first lncRNAs studied, even before the recognition of these class of transcripts, was the *X inactivation specific transcript (Xist)*, located in the X inactivation centre region and involved with X chromosome inactivation in mammalian females [6,7]. Overall, lncRNAs are known to be involved in gene expression regulation - at the transcriptional and post-transcriptional levels, by epigenetic or other mechanisms, such as interfering with the recruitment of RNA polymerase II or inducing chromatin remodelling. Furthermore, they participate in genomic imprinting; in nuclear and cytoplasmic trafficking; in protein localization and activity; in interaction with miRNAs, among other processes [reviewed in 8]. In addition, they can be further processed to small ncRNAs [9] or even encode functional micropeptides [10,11]. However, little is known about how these transcripts control gene expression.

lncRNAs are strictly regulated [12,13] and participate in or are products of many biological processes [14,15]. Mutations in the primary sequence of lncRNAs, as well as aberrant variations of their expression, have been associated with several disorders, pointing to their potential as disease biomarkers [16]. Therefore, the lncRNAs have been largely studied in tissues' homeostasis and pathology to understand their physiological effects and the consequences of their deregulation in complex diseases.

We performed an extensive search of the literature for articles presenting data about lncRNAs involved in homeostasis of different tissues and cell types and/or that have a tissue-specific expression pattern. We present the information by cell or tissue type throughout this review.

6.2 lncRNAs: expression patterns in tissues or cell types

lncRNAs are strictly regulated and present cell-specific expression, substantiating their crucial role in physiological mechanisms [1,3,17]. In the following, we summarize what is currently known about lncRNA expression along cell development and differentiation, and in specific pathways. A summary of that information is presented in table S1.

6.2.2 Hematopoietic Cells

Ontogenesis of blood cells from the hematopoietic stem cells (HSCs) occurs throughout the whole individual's life, and is highly controlled by transcription factors and non-coding RNA. Circulating blood, where most of these cells are found, is easy to acquire and to work with, being routinely used in molecular studies. Yet, some authors did analyse bone marrow and thymus, to understand early stages of hematopoiesis and development of the different cell lineages.

H19 is one of the most abundant and conserved mammalian noncoding transcripts and a key regulator of differentiation of various cell and tissue types, as will be seen also in other topics of this review. H19 is a transcript of the H19/IGF2 genomic imprinted cluster. H19 is transcribed from the maternally-inherited locus, and IGF2 (insulin-like growth factor-II) from the paternally-inherited locus. During murine hematopoiesis, the growth-restricting lincRNA

H19 was downregulated in HSCs before their proliferation, and upregulated in long-term HSCs. *H19* is localized downstream of *Igf2* in the *H19/Igf2* locus. Both genes are co-expressed and have an antagonistic effect on cell proliferation during hematopoiesis [18]. *H19* also inhibits HSCs activation and proliferation serving as a precursor of miR-675, a miRNA that targets the insulin-like growth factor 1 receptor (*Igf1r*) in placenta [19]. In another study, among the 323 novel lncRNAs identified in murine HSCs RNA sequencing, 159 were HSC-enriched or likely HSC-specific (lncHSCs), in comparison with the transcriptome of differentiated B cells and granulocytes. Among these 159 lncRNAs, the *lncHSC-1* was identified as involved in myeloid differentiation, and *lncHSC-2* as involved in HSC self-renewal and T cell differentiation. In addition, *lncHSC-2* is enriched with target sites for important hematopoietic-specific transcription factors, especially E2A [20].

In the following, we will comment well-established lncRNAs involved in ontogeny and homeostasis of circulating blood cells and their progenitors (fig. 11).

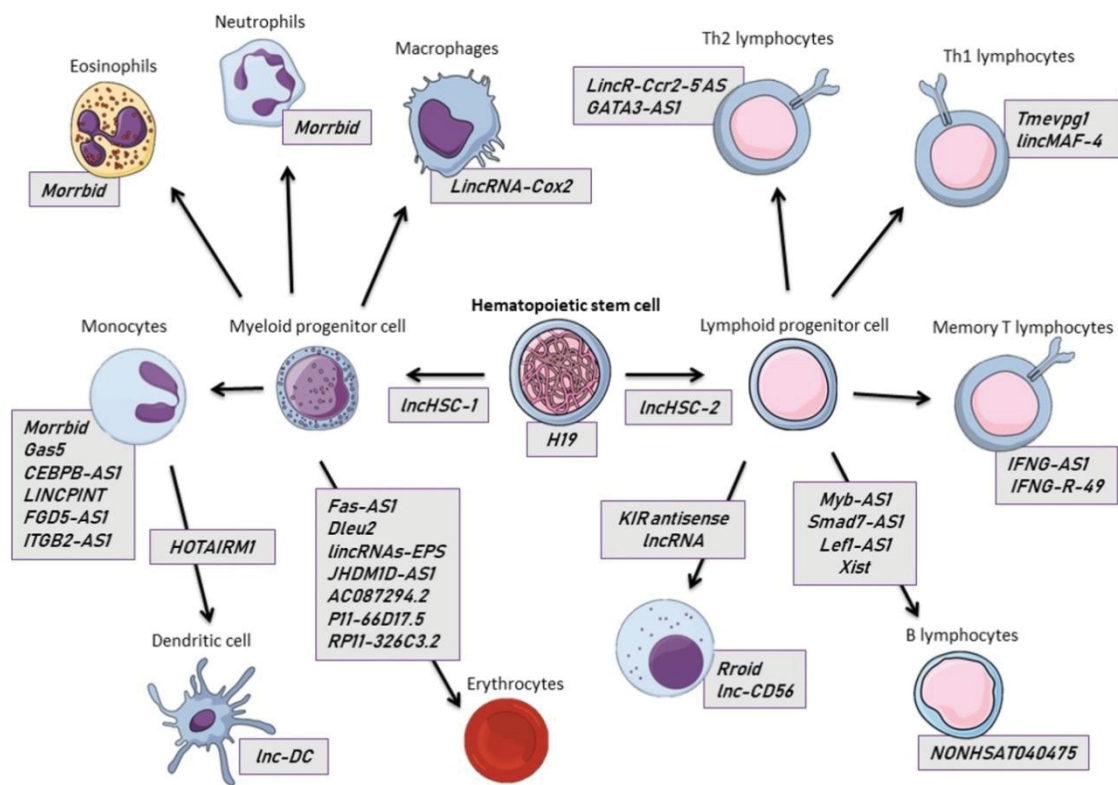


Figure 9. lncRNAs described in physiology of mature and progenitor hematopoietic cells, derived from myeloid (left) and lymphoid (right) differentiation from a hematopoietic stem cell (HSC), in which the lncRNA *H19* plays a central role. In the grey rectangles, are listed the lncRNAs specifically or differentially expressed in each cell type (rectangles at the side of cells), or lncRNAs involved in differentiation and maturation of these cells (upon the arrows).

6.2.2.1 Erythrocytes

During red blood cells development, various lncRNAs regulate erythroid gene expression. Although their specific functions are still not elucidated, some critical processes for the homeostasis of erythropoiesis involve pro- and anti-apoptotic pathways. *Fas-antisense 1* (*Fas-AS1*, also known as *Saf*), a lncRNA antisense to *Fas* cell surface death receptor coding gene, was upregulated during differentiation stage of erythropoiesis by the erythroid transcription factors GATA-1 and KLF1, and negatively regulated by NF- κ B. In erythroblasts

derived from CD34⁺ hematopoietic stem cells, the overexpression of *Fas-AS1* reduced the Fas surface levels and protected cells from death signals [21].

Depletion of 12 out of 13 lncRNAs selected as candidate modulators of erythropoiesis impaired erythrocyte maturation, suggesting that they play crucial roles during late stages of maturation [22]. One of these candidates RNAs is the *lincRNA-EPS*, which promotes erythroid differentiation by preventing apoptosis [23].

The lncRNAs *lincRNA-EC2*, *-E4* and *-EC9*, were found as required for erythrocyte maturation in the fetal liver, but they were absent in the adult bone marrow. The enhancer-derived *alncRNA-EC7* is required for activation of the neighbouring gene *SLC4A1*, which encodes BAND3, a crucial anion transporter of erythrocyte membrane. The lncRNA *alncRNA-EC7* has an evolutionary conserved enhancer site, and there is evidence of chromatin looping between this site and the *SLC4A1* promoter-proximal region, supporting a model whereby *alncRNA-EC7* modulates erythropoiesis. In erythroblasts, expression of *elncRNA-EC3* is positively correlated with expression of *KIF2A*, which encodes a microtubule-associated kinesin motor protein, required for mitotic progression. Moreover, the lncRNA *DLEU2*, encoded by *shlncRNA-EC6*, is expressed from distinct promoters in erythroblasts, compared to other cell types, and represses the neighbour gene *SPRYD7*. A deletion in *shlncRNA-EC6* is frequent in B cell malignancies, and *DLEU2* might have an important role in erythropoiesis [22].

The expression of several lncRNAs is restricted to specific stages of erythrocyte differentiation. Often, these lncRNAs are co-expressed with mRNAs and miRNAs in the same stages of cell differentiation. Among the lncRNAs with potential role during hematopoiesis, *JHDM1D-AS1*, *AC087294.2* and *RP11-66D17.5* were frequently associated with genes of erythrocyte mature network, and the *RP11-326C3.2*, localized between *NLRP6* and *ATHL1* coding-genes, is co-expressed with *ATHL1*, both increased during erythrocyte maturation. Moreover, as observed in *in silico* analysis, the lncRNA *TOPORS-AS1* could act as ceRNA for three microRNAs (miR-532-3p, miR-191-5p, and hsa-miR-652-3p) involved in inhibition of erythrocyte development [24]. In a similar previous study, it was found co-expression between a lncRNA (not identified by the authors) and the miR-4732, a key modulator miRNA during human erythropoiesis and abundant in erythrocytes within the erythroid-enriched miR-144/451 locus, a cluster upregulated by GATA1 during erythrocyte development. This ~250bp lncRNA overlaps with miR-4732, and is proximal to the miR-144/451 cluster. This cluster has been highly related to erythropoiesis, and all RNAs in this locus were highly expressed in mature erythrocytes, in comparison with RNAs from other genomic regions [25].

6.2.2.2 Other myeloid lineage cells

Beyond the erythrocytes, the myeloid lineage includes the neutrophils, eosinophils, monocytes and dendritic cells, among other important cells for the innate immunity.

The lifespan of short-lived myeloid cells (neutrophils, eosinophils and Ly6Chi classical monocytes) was found in mice as regulated by the *myeloid RNA regulator of Bim-induced death (Morrbid)*, in response to pro-survival cytokines. The nuclear lncRNA *Morrbid*, specifically expressed in mature short-lived myeloid cells, neighbours and cis-regulates the pro-apoptotic gene *Bcl2l11 (Bim)*, by promoting an H3K27me3 modification at Bim promoter through interaction with the polycomb repressive complex 2 (PRC2 – a histone modifying

complex), repressing the apoptosis in these cells. *Morrbid* deficiency resulted in a drastic reduction of these myeloid cells in blood and spleen, increasing susceptibility to bacterial infection and protecting from eosinophil-driven allergic lung inflammation. Moreover, *Morrbid* dysregulation in humans is associated with hypereosinophilic syndrome, a disorder characterized by severe eosinophilia. Collectively, the lncRNA *Morrbid*, in response to extracellular pro-survival signals, is crucial for the physiologic control of neutrophils, eosinophils and classical monocytes lifespan [26].

Various lncRNAs were highly expressed in monocytes, including the well-known lncRNAs *growth arrest specific 5 (GAS5)*, *CCAAT/enhancer binding protein beta anti-sense (CEBPB-AS1)*, *lincRNA P53 induced transcript (LINCPINT)*, *integrin beta 2 complement component 3 receptor 3/4 subunit anti-sense (ITGB2-AS1)* and *FYVE RhoGEF/PH domain containing 5 antisense (FGD5-AS1)* [27]. Several other annotated, but still uncharacterized lncRNAs, were also highly expressed in monocytes, as the *LINC00678*, *LINC00657*, *LINC00989* and *LINC01420*, which might have a potential role in monocyte regulation. Moreover, novel lncRNAs, such as *LINC01420*, whose gene is located between the *UBQLN2* and *SPIN3* coding genes, and with a moderated expression in other hematopoietic cell types, were also detected in monocytes, apart from the lncRNAs *TCONS_00282615* and *TCONS_00008494*, specific to these cells [27].

During the differentiation of monocytes to dendritic cells (DCs), important epigenetic modifications were found in the promoter region of *HOXA transcript antisense RNA myeloid-specific 1 (HOTAIRM1)*, a lncRNA gene located between *HOXA1* and *HOXA2* genes. *HOTAIRM1* was downregulated in differentiated DCs, and its silencing caused expression changes of several monocyte differentiation markers. Moreover, upregulation of miR-3960 could downregulate both *HOTAIRM1* and the DC differentiation repression gene *HOXA1*, inducing the differentiation into DCs. The results indicate competitive binding between *HOTAIRM1* and miR-3960, regulating DC differentiation [28].

The *lincRNA-COX2*, neighbour of the prostaglandin-endoperoxide synthase 2 (*Ptgs2/Cox2*) gene in mice, is highly induced in innate immune response [29,14] and was first described in DCs after toll-like receptor 4 (TLR4) stimulation with lipopolysaccharide (LPS) [14]. Some years later, it was demonstrated that in macrophages stimulated with Pam3CSK4 (a TLR2 ligand), numerous immunological genes were induced, including 62 lncRNAs, of which *lincRNA-COX2* was the most highly induced [29]. In the same study, the authors observed that silencing the *lincRNA-COX2* did not alter *COX2* expression, but induced 787 genes in unstimulated cells, including chemokines (such as *Ccl5*, *Cx3cl1*), chemokine receptors (such as *Ccr1*) and interferon-stimulated genes (such as *Irf7*). In stimulated cells, *lincRNA-COX2* silencing attenuated the expression of 713 genes, including *Tlr1*, *interleukin (IL-) 6* and *IL-23a*. It was also demonstrated *in vitro* that *lincRNA-Cox2* interacts with the heterogeneous ribonucleoprotein (hnRNP) A/B and A2/B1, and knockdown of both hnRNPs repressed *Ccl5* and *Irf7* [29]. In contrast, *lincRNA-Cox2* was seen to induce *Ccl5* expression in LPS-stimulated immortalized macrophage cell lines, by recruitment of the chromatin-remodelling complex SWI/SNF to the *Ccl5* promoter. Furthermore, *lincRNA-Cox2* is required in the innate response for the expression of NF- κ B-regulated late-primary inflammatory response genes [30].

A lncRNA exclusive to DCs is *lnc-DC*, highly expressed from the human *Wdnm1-like* pseudogene at the moment of DC differentiation, driven by PU.1 transcription factor [31,32],

being present in Lin⁻MHCII⁺CD11c⁺ conventional DCs, and absent or lowly expressed in CD14⁺ monocytes, B cells, and other hematopoietic cells. *Lnc-DC* binds to STAT3, a transcription factor that regulates DC differentiation, protecting STAT3 from inactivation by the tyrosine phosphatase SHP1. Therefore, through promoting the STAT3 signalling, the *Lnc-DC* leads to DC differentiation from monocytes to DCs, *in vitro* and *in vivo*, and to the DC capacity to stimulate the activation of T lymphocytes [31].

6.2.2.3 Lymphocytes

Lymphocytes are a class of leucocytes of prime importance for both the innate immunity (innate lymphoid cells – ILCs, where the natural killer cells are included) and acquired immunity (T and B cells).

Many novel lncRNAs were identified by global gene-expression profiling of human bone marrow and thymic progenitor cells spanning the earliest stages of B and T lymphoid lineages commitment. These lncRNA were predominantly intergenic, followed by anti-sense lncRNAs. Expression of a high proportion of the lncRNAs was positively correlated with that of neighbouring protein-coding genes, including genes involved in hematopoiesis and regulation of lymphocyte development and proliferation. In addition, the authors identified lncRNAs co-expressed with known protein-coding genes specific for different lymphocyte lineages, such as *PAX5*, *EBF1*, *CD3*, *BCL11B*, *NOTCH1*, *RAG1* and *RAG2*, among many others. The results provide a resource of candidate lncRNA to be further investigated in functional studies to clarify their role in lymphopoiesis [33].

Moreover, *NONHSAT040475* is a lncRNA widely expressed in human peripheral blood leukocytes. Its expression was compared between T cells (CD3⁺), B cells (CD19⁺) and natural killer (NK) cells (CD56⁺). B cells were the lymphocyte subset with highest expression of *NONHSAT040475*, in comparison with NK cells and granulocytes, which presented lowest expression among the analysed cell types [34].

The next topics will address the lncRNAs associated with specific lymphocyte types.

6.2.2.4 T lymphocytes

T lymphocytes (or T cells), as one of central cell type of the immune system, perform crucial roles in adaptive immune responses, which are under constant regulation to maintain body homeostasis. T lymphocytes are subdivided in partially overlapping subgroups based on their more prominent known functions. So, CD4⁺ T helper lymphocytes (Th1, Th2, Th17) are essential for activation and regulation of B lymphocytes and cytotoxic T lymphocytes (CTL or Tc); CD8⁺ T cytotoxic lymphocytes eliminate virus-infected cells, tumour cells or otherwise damaged cells; regulatory T lymphocytes (Treg) are essential for immunological peripheral tolerance; memory T lymphocytes, in their turn, provide a quick response upon re-exposure to their specific antigen. In the first genome-wide characterization of lncRNAs expressed in mammalian (human and mouse) CD8⁺ T lymphocytes, over a thousand lncRNAs were identified, many of which presented stage- or tissue-specific expression, mapped close to or overlapped shorter functional RNA or protein-coding genes with quite characterized roles in CD8⁺ T cells [35]. Most lncRNAs expressed in CD8⁺ T cells presented signatures of evolutionary conservation, and their secondary structures and the characteristics of the promoters were further evidences of functional relevance [35].

LncRNAs also affect transcription factors present in T cells, as the NFAT complex (nuclear factor of activated T cells). In resting cells, NFAT proteins are located in the cytoplasm, where they are heavily phosphorylated. Increase of intracellular Ca²⁺ concentration activates calcineurin that dephosphorylates the NFAT regulatory domain, ultimately resulting in nuclear accumulation of NFAT and transcription of NFAT target genes. Phosphorylated NFAT1 is present with the lincRNA *ncRNA repressor of the nuclear factor of activated T cells (NRON)* and the multidomain scaffold protein IQGAP1 in a cytoplasmic RNA-protein complex that also contains three known NFAT inhibitory kinases. This supports the hypothesis that lincRNAs may function as scaffolds for transcriptional regulators in large RNA-protein complexes [36]. LncRNAs can influence Th1/Th2 differentiation by affecting expression of critical cytokines. The mouse lincRNA *Tmevpg1* is required for control of Theiler's virus infection and, as its human orthologue *TMEVPG1*, contributes to interferon gamma expression. Expression of *TMEVPG1* and *IFNG* correlate positively. This lincRNA is preferentially expressed in Th1 cells, depending on transcription factors STAT4 and TBET/T-bet, which drive the Th1 differentiation program [37].

High-throughput RNA sequencing was used to investigate lncRNAs expression along the development and differentiation of T cell lineages. Forty-two subsets of thymocytes and peripheral T cells at multiple time points of differentiation were analysed and 1,524 genomic regions expressing lincRNAs were identified. The effect of *TMEVPG1* in Th1 lymphocyte differentiation (see above) was corroborated, but none of the numerous other lincRNAs had been previously described in T cells. Results revealed highly dynamic stage- and cell-specific expression patterns for lincRNAs during T cell differentiation. Key transcription factors, specific for the Th1 (T-bet and STAT4) and Th2 (GATA-3 and STAT6) lineages were largely responsible for the lineage-specific expression of T cell lincRNAs [38]. A more specific result was obtained for lincRNA *LincR-Ccr2-5'AS*, which together with transcription factor GATA-3, was a fundamental component of the chemokine signalling pathway in Th2 subset cells and important for migration of these cells [38]. More recently, other 13 subsets of T cells and B cells were analysed by RNA sequencing and over 500 previously unknown lincRNAs were identified, suggesting that lincRNAs might collaborate to determination of lymphocytes identity and to modulation of their functional plasticity [39]. Beyond that, the influence of a specific lincRNA *LincMAf-4* in T cells differentiation was partially clarified. *LincMAf-4* is a chromatin-associated lincRNA specific to the Th1 subset of helper T cells, whose expression was inversely correlated with expression of MAF, a Th2-associated transcription factor. *Linc-MAF-4* regulated *MAF* transcription through the recruitment of chromatin modifiers LSD1 and EZH2; downregulation of *linc-MAF-4* alter T cell differentiation toward the Th2 phenotype [40]. Furthermore, progenitor cells spanning the earliest stages of B lymphoid and T lymphoid specification were isolated from the bone marrow and thymus for RNA sequencing analysis. More than 3,000 genes encoding yet unexplored lncRNAs were described. Highly stage-specific lncRNA expression patterns characterized lymphoid commitment, that is lncRNAs mark the beginning of the divergence of B and T lineages [33]. In addition, more than 2,000 lncRNAs were identified in human T cell cultures by RNA sequencing, moreover a cluster of antisense lncRNA transcribed from the *RAD50* locus that regulate the expression of Th2 cytokines, IL-4, IL-5 and IL-13 was described [41]. Other Th2 specific lincRNA, *GATA3-AS1* was analysed in 27 human tissues and immune cell types where it was largely undetectable or had very limited expression supporting *GATA3-AS1* Th2-specific expression pattern. Indicating it might serve as a particular indicator of Th2 response [42]. Microarray analysis in CD4⁺ T cells at different stages also confirmed distinguishable lncRNAs

expression profiles and highlights lncRNAs importance in CD4⁺ T cell development and activation [43]. Specific lncRNA genes involved in early differentiation of Th1 and Th2 subsets were identified and a positive trend in expression of the lineage-specific lncRNA and the neighbouring protein-coding gene was observed [44].

LncRNAs profiles of naive T cells and two subsets of memory CD4⁺ T cells (central and effector) were obtained by RNA sequencing and a total of 57,000 novel lncRNAs were found in at least one cell subtype. Many lineage-specific lncRNAs were reported, most of them mapped in a non-random manner, positioned close to lineage-specific protein-coding genes with which they are co-expressed. Furthermore, it has been demonstrated that lncRNAs mapped near to the *IFNG* locus are necessary for *IFNG* prompt expression, since the expression inhibition of lncRNAs *IFNG-AS1* and *IFNG-R-49* by siRNAs abolishes *IFNG* expression by effector memory T lymphocytes [45].

Considering the huge amount of data generated, especially by high throughput sequencing technologies, and the recent progress in lncRNAs identification, it is possible to conclude that the role of lncRNAs in T cells development as well in immune responses becomes evident and concrete. It has become clear that lncRNAs are expressed in a highly lineage-specific manner and help to control the differentiation and function of these important immune cells. Although there is much to explore about T cells and lncRNAs, the importance of these non-coding RNAs is already a reality.

6.2.2.5 B lymphocytes

B lymphocytes are responsible for the production of immunoglobulins, also known as antibodies, which recognize foreign antigens, being thus essential for the so-called humoral immunity. Activated B cells can evolve to plasmocytes and memory cells. In a transcriptome analysis of 11 B cells subsets, several lncRNAs expressed along human B-cell development and correlated with other RNAs were identified. Among the lncRNAs expressed during the early stage of B-cell development, the transcripts *MYB-AS1*, *SMAD-AS1* and *LEF1-AS1* are antisense to genes of known early B-cells transcription factors (*MYB*, *SMAD1* and *LEF1*, respectively). Another transcript identified was the lncRNA *CTC-436K13.6*, located between the genes *EBF1* (which encodes the transcription factor early B-cell factor 1) and *CLINT1*. Several lncRNAs were also identified in proliferative stages and in the germinal centre. In germinal centre, the lncRNAs *LINC00487*, *LINC00877*, and *RP11-203B7.2* were found as highly connected genes (hub genes). Furthermore, the gene for lncRNA *RP11-132N15.3*, found as predominantly expressed in centroblasts, is localized upstream of the *BCL6* gene, the main germinal centre regulator [46].

In another transcriptome analysis of B cell subtypes of mice, 4,516 lncRNAs were identified as differentially expressed during B cell development and activation, of which 702 were predicted to be associated with enhancers (eRNAs) and 192 with promoter regions (pRNAs). Among them, a strong positive correlation between the eRNA *LNCGme02323* and the cluster lightpink4, which contains 31 protein-coding genes regulated in marginal zone B cells was observed. Moreover, 784 lncRNAs of pro-B cells and 717 lncRNAs of mature B cells overlapped or were at less than 1 kb downstream of binding sites for PAX5, a transcription factor required for B-cell development. Evidence of regulation by PAX5 was found for 109 of these lncRNA, indicating a role for these lncRNA in B cell homeostasis [47].

Finally, recently it was demonstrated that during female B cell development, the inactivated X chromosome (Xi) loses the heterochromatin modifications and the co-localization within the lncRNA *Xist*, which are required to compensate between sexes for dosage of X-linked gene products. The lack of *Xist* at the Xi occurs at the pro-B cell stage and persists in non-activated B cells, being restored to Xi during the activation of mature B cells. For restoring of *Xist* localization and chromatin modifications, the transcription factor YY1, involved in V(D)J recombination of B cells is required. While it remains unknown why the Xi loses heterochromatic modifications and *Xist* RNA during female B cell development, the authors suggested that these features might enable rapid immune-related X-linked genes reactivation in response to infections, and could also explain some of the immunological differences between the sexes [48].

6.2.2.6 NK cells

lncRNAs were also seen to be involved in NK cell function and differentiation. In a lncRNA profiling analysis of different populations of NK cells, numerous novel lncRNAs were identified, among them, the *lnc-CD56*, a lncRNA specific to NK cells. *lnc-CD56* positively correlated with the expression of the classical NK cells marker CD56, in primary and some differentiated human NK cells. Moreover, *lnc-CD56* is predicted to interact with the NK-transcription factors TBX21, IRF2, IKZF2, ELF4 and EOMES [49].

A novel killer cell immunoglobulin-like receptor (KIR) gene-body associated antisense transcript not expressed in mature NK cells was described in cells with stem cell properties. The promoter of this spliced lncRNA is at intron 2 of several *KIR* genes, and its sequence is complementary to the proximal promoter of the *KIR* genes and KIR-coding exons 1 and 2. Its promoter contains binding sites for MZF-1, a bi-functional transcription factor that activates transcription in cells of hematopoietic origin and represses transcription in most other cells. In support of a role of that lncRNA in the initial silencing of *KIR* genes, differentiation of human embryonic stem cells into CD34-expressing hematopoietic progenitor cells was accompanied by loss of this lncRNA [50].

Moreover, it was recently demonstrated that the *RNA-demarcated regulatory region of id2 (Rroid)* of mice is a lncRNA locus located ~220 kb from the *id2* gene, a transcriptional regulator that inhibits the expression of T and B cell-specific genes, and is essential for lineage commitment and maintenance of mature group 1 innate lymphoid cells (ILC1s). The lncRNA *Rroid* is highly and specifically expressed in ILC1s, where the NK cells are included. In ILC1s, the *Rroid* locus (not the lncRNA itself) directly interacts with the *id2* promoter and, in response to IL-15, promotes chromatin accessibility and deposition of the transcription factor STAT5 at the *id2* promoter, thus acting as a cis-regulatory element of *id2* gene and promoting its expression in ILC1s [51].

6.2.3 Solid tissues

Several lncRNAs were identified in healthy solid tissues and organs, exercising critical roles during their differentiation, metabolism and physiological functions (Fig. 12). These lncRNAs and the regulatory mechanisms that they play in specific solid tissues are reviewed in this topic.

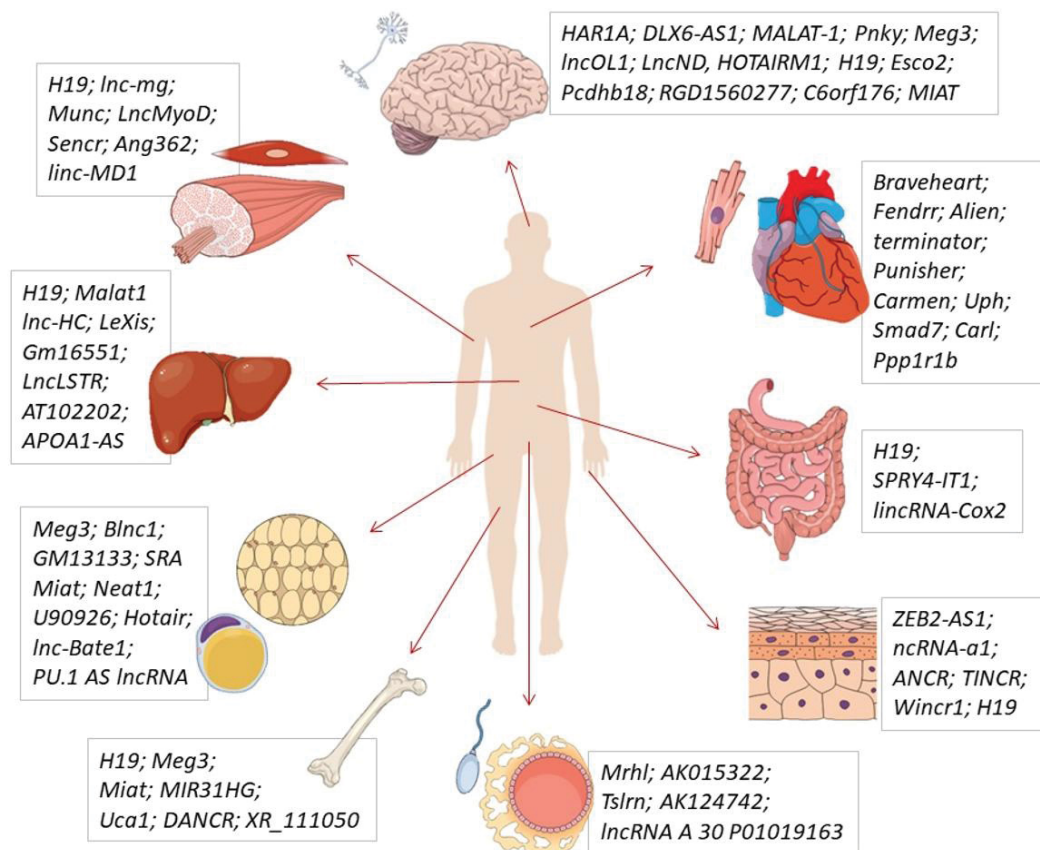


Figure 10. Main lncRNAs described as playing physiological roles (in function, homeostasis and/or differentiation) in the following tissues (clockwise direction): nervous, cardiac, intestinal, epidermal, germ, bone, adipose, hepatic and muscular (skeletal and smooth).

6.2.3.1 Epidermal tissue

Epidermal epithelial tissue is a stratified surface epithelium providing a barrier to the external environment. The epithelial cells are closely interconnected (juxtaposed). The epidermis is the outermost layer of the skin and has a high turnover rate due to the continuous shedding of the outer layer of cornified cells. The importance of lncRNAs in epithelial cells has been demonstrated by diverse studies, especially with regard to gene expression regulation [52,53]. One of the first examples is the natural antisense transcript (NAT) *Zeb2-AS1*, a lncRNA that is an important component of the epithelial-mesenchymal transition (EMT) [52]. Among the regulatory proteins required for EMT is the Zeb2 transcriptional repressor of E-cadherin. *Zeb2-AS1* acts inducing the expression of the *Zeb2/Sip1 (ZFHX1B)* gene by promoting the maintenance of a large intron in the 5'UTR of the *Zeb2* transcript. This region contains an internal ribosome entry site (IRES) necessary for *Zeb2* expression, which in turn participates of downregulation of E-cadherin mRNA and protein levels [52]. Furthermore, several lncRNAs have been identified as differentially expressed during keratinocyte (KCs) differentiation by non-coding RNA array analysis of epidermal equivalents grown *in vitro* [54], including *AF005081*, *UC003af*, *BC020554* and *AK022798*, upregulated at different times and differentiation stages. In a study of primary human keratinocytes, lncRNAs expression was measured by microarray before and after induction of differentiation. A total of 687 lncRNAs were detected, of which 104 were differentially expressed following differentiation [55]. The *activating long ncRNA 1 (ncRNA-a1)* was found upregulating *ECM1*, a neighbour protein coding gene, and both, the protein and the lncRNA, were induced after keratinocyte

differentiation [55]. Other examples are two lncRNAs that control the balance between the progenitor and differentiated compartment of human epidermis: *anti-differentiation non-coding RNA (ANCR)*, which is strongly suppressed in differentiated cells, and *terminal differentiation-induced ncRNA (TINCR)*, that is highly induced during epidermal differentiation [56,57,53]. *ANCR* is required to enforce the undifferentiated cell state within epidermis [56], whereas *TINCR* promotes human epidermal differentiation by a post-transcriptional mechanism, mediated by interactions with *staufen1* (STAU1) protein [57]. Moreover, *ANCR* and *TINCR* act as regulators of transcription factors MAF and MAFB, which are important for epidermal differentiation regulation [53]. Interestingly, a lncRNA-miRNA interaction has been reported as influencing keratinocyte differentiation. The lncRNA *H19* can function as competitive endogenous RNA for miR-130b-3p. This miRNA, targets the desmoglein 1 (DSG1) transcript, resulting in decreased DSG1 expression and action, thus, keratinocyte differentiation is inhibited. However, lncRNA *H19* adds a regulation point in this differentiation process by acting as miR-130b-3p sponge and preventing its activity, hence, *H19* can act promoting keratinocyte differentiation [58]. In addition, lncRNAs, including the aforementioned *H19*, can play a role also in regulation of dermal papilla (DP) cells function, influencing hair follicle development and regeneration. Comparison of lncRNA differentially expressed in early- and late-passage DP cells revealed numerous differentially expressed lncRNA. Three of these, *RP11-766N7.3*, *H19* and *HOTAIR*, may participate in hairy gene regulation by targeting the Wnt signalling pathway, which plays critical roles in skin development and homeostasis and is required for hair follicle initiation [59]. Recently, 111 lncRNAs were shown to be differentially expressed in response to activation of Wnt/ β -catenin signalling in mouse dermal fibroblasts. The Wnt signalling-induced lncRNA *Wincrl* is highly expressed in primary dermal fibroblasts and may affect complex cellular behaviours including collective cell migration and collagen processing and contraction. Also, *Wincrl* affects expression of several genes, such as *Col1a1* and *Mmp10* [60].

6.2.3.2 Nervous tissue

Many lncRNAs are specifically expressed in the brain. These transcripts have been considered essential pieces in regulation of a broad spectrum of neuronal processes, including brain development, establishment and maintenance of neural cell types, synapse formation and function, stress responses, and age-associated changes [61]. Some lncRNAs have been implicated in brain development, such as *HAR1A*, which is expressed specifically in Cajal-Retzius neurons in the developing human neocortex from 7 to 19 gestational weeks, a crucial period for cortical neuron specification and migration. *HAR1A* is co-expressed with *reelin*, that is of fundamental importance in specifying the six-layer structure of the human brain cortex [62]. Another example is the lncRNA *DLX6 antisense RNA 1 (DLX6-AS1, or Evf2)*, which influences the establishment of murine regional forebrain organization. This lncRNA acts in ventral forebrain development recruiting DLX and MECP2 transcription factors to DNA regulatory elements in the *Dlx5/6* intergenic region and controlling *Dlx5*, *Dlx6* and *Gad1* expression through trans and cis-acting mechanisms [63]. Moreover, *Pinky (Pnky)* is a lncRNA that influences murine neural stem cells (NSCs) differentiation. *Pnky* interacts with the PTBP1 protein and regulates neurogenesis from NSCs in the embryonic and postnatal brain [64]. In the Catarrhini branch of primates, the lncRNA for neuronal development (*LncND*) is highly expressed in neural progenitor cells and downregulated in neurons. It controls the expression of Notch signalling pathway, which may be necessary to maintain stemness of outer sub-ventricular zone radial glia-like cells (oRG), by function as a miRNA sponge for the microRNA mir-143-3p. The evolutionary emergence of *LncND* its co-

evolution with miRNA-143-3p may have contributed to the expansion of the cerebral cortex in primates [65].

Global RNA sequencing techniques has also been used to investigate lncRNAs expression in brain development and function. Neurons derived from induced pluripotent stem cells (iPSCs) have been analysed by RNA sequencing and emphatic changes in lncRNAs expression during transition from iPSCs to early differentiating neurons were detected. *HOTAIRM1* is among lncRNAs expressed early in neurogenesis, being important for *HOXA* genes regulation during myelopoiesis [66]. Besides, through a microarray approach 934 lncRNAs were found differently expressed when comparing human embryonic stem cells (hESCs) with neural progenitor cells (NPCs). Of these lncRNAs, four were subjected to knockdown that indicated that these lncRNAs are required for adequate neuronal differentiation [67]. Microarray assay was also used to explore differentiation of bone marrow mesenchymal stem cells (BMSCs) into neural cells. After differentiation induction 23 lncRNAs were upregulated or downregulated. Further, RT-qPCR was performed to validate *H19*, *Esco2*, *Pcdhb18* and *RGD1560277* differential expression, corroborating lncRNAs as key regulators of BMSCs differentiation [68]. A transcriptome and splicing database was generated for eight different brain cell types from mouse cerebral cortex: neurons, astrocytes, oligodendrocyte precursor cells, newly formed oligodendrocytes, myelinating oligodendrocytes, microglia, endothelial cells, and pericytes [69]. Expression of 811 lncRNAs was detected and some of them were among the highest expressed transcripts in the mouse brain, with 12 placing among the 7% most highly expressed genes in the RNA dataset. Moreover, some lncRNAs were expressed in cell type-specific or enriched manners [69]. RNA sequencing was also used to profile lncRNA expression in mouse auditory forebrain during postnatal development at time points before and after the onset of hearing [70]. This study revealed that the lncRNA expression profile was distinct between primary auditory cortex and medial geniculate body and between different postnatal time point, indicating spatial and temporal specificity during maturation of the auditory forebrain [70]. Recently, the human neocortex transcriptome was also analysed at single cell level [71]: hundreds of cells from human neocortex at different stages of development were individually sequenced, revealing that many lncRNAs are abundantly expressed in individual cells and are cell type-specific. For example, *LOC646329*, a lncRNA that regulates cell proliferation, is enriched in single radial glia cells but is detected at low abundance in bulk tissues [71]. Besides, 155 differentially expressed lncRNAs have been reported in a neuronal cell line after stimulation by brain-derived neurotrophic factor (BDNF), a critical neurotrophin for neural development. lncRNAs with fold changes above 2 included the already known *C6orf176* and *MIAT*, however many differentially expressed lncRNAs are not currently understood [72]. For oligodendrocytes, a dynamic expression profile at different stages of development was established. Among oligodendrocyte-restricted lncRNAs, the evolutionary conserved chromatin-associated *lncOL1* was functionally explored. Its overexpression promotes oligodendrocyte differentiation in developing brain whereas its knockdown inhibited expression of myelin genes, causing defects in CNS myelination [73].

lncRNA are also involved in the regulation of synaptic processes. The nuclear lncRNA *metastasis-associated lung adenocarcinoma transcript 1* (*MALAT1*), also known as *NEAT2* (*nuclear-enriched abundant transcript 2*) or NCRNA00047, is highly expressed in neurons and regulates synapse formation by modulating the expression of genes involved in synapse formation and/or maintenance [74]. Furthermore, the lncRNA *maternally expressed gene 3*

(*Meg3*) in mice is important for maintaining synaptic plasticity by modulating PTEN/PI3K/AKT signaling pathway. In addition, knockdown of *Meg3* in primary cortical neurons prevents the glycine-induced increase of the GluA1 subunit of AMPARs (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors) in the cell membrane. Thus, *Meg3* is important for modulating AMPARs expression, which mediate fast synaptic transmission in the central nervous system [75].

6.2.3.4 Muscular tissue: cardiac, skeletal and smooth

There are three types of muscle: smooth, skeletal and cardiac. Recently, lncRNAs have been reported as important regulators of myocyte gene expression, and/or have a role in contraction mechanisms, being involved in muscle development, differentiation and homeostasis. The cardiac muscle composes the heart, the first functional organ formed during embryo development. The multistep process involved in cardiac myocyte differentiation has been detailed elsewhere [reviewed in 76], and involves highly coordinated pathways of gene expression modulated by the products of protein-coding and non-coding genes [reviewed in 77].

The murine heart-associated lncRNA *Braveheart* (*Bvht* or *AK143260*) was one of the first lncRNAs associated with cardiac development. *Bvht* is required for cardiovascular lineage commitment of nascent mesoderm, through the activation of a core regulatory network of cardiovascular transcription factors, and possibly mediating also epigenetic regulation during cardiomyocyte differentiation, by interaction with SUZ12, a component of polycomb-repressive complex 2 (PRC2) [78]. In another study using a mouse model, the mesoderm-specific foetal-lethal non-coding developmental regulatory RNA (*Fendrr*) was found to be an activity modulator of two histone-modifying complexes. *Fendrr* binds to the PRC2 and trithorax group/mixed lineage leukemia (TrxG/MLL) complexes, regulating the epigenetic role of both through attenuation of the expression of transcription factors that are important in lateral mesoderm differentiation. Thereby, *Fendrr* is essential for normal heart and body wall development [79].

The cardiac-enriched lncRNA *Upperhand* (*Uph*; also known as *HAND2-AS1* or *DEIN*) shares a bidirectional promoter with the *Hand2* gene, which encodes a transcription factor involved in heart development, and is precisely controlled by an upstream super-enhancer locus. *Uph* transcription *in cis* is required for the permissive epigenetic signature, GATA4 binding, and transcription through the enhancer. Therefore, loss of *Uph* transcription abolishes *Hand2* expression and development of the heart [80].

Transcriptome analysis of human cardiovascular progenitor cells (CPCs) revealed 570 lncRNAs possibly involved in cardiac differentiation. Among them, the *cardiac mesoderm enhancer associated noncoding RNA* (*CARMEN*), was one of the most highly expressed. The *CARMEN* lncRNA is upstream of the cardiac mesoderm-specifying gene regulatory network, and interacts with the PRC2 components SUZ12 and EZH2. Knockdown of the mouse orthologue (*Carmen* or *AK087736*) inhibited cardiac specification and differentiation of CPCs. Furthermore, *CARMEN* also plays a role in maintaining cardiac identity of differentiated cardiomyocytes [81].

In a transcriptome analysis searching for ncRNAs differentially regulated during hESC differentiation to vascular endothelial cells, 75 novel lncRNAs whose sequence and functionality might be conserved across vertebrate species were identified. Among them,

ALIEN, *TERMINATOR* and *PUNISHER* were functionally characterized in human and other vertebrates' cardiovascular development. The lincRNA *ALIEN*, was expressed in vascular progenitors, and was related to cardiovascular development before vascular or cardiac specialization, being strongly correlated with genes involved in mesoderm and cardiovascular commitment, such as *GATA4*, *EOMES*, *MIXL1*, and *T (Brachyury)*. Two outcomes were observed after knockdown of *ALIEN*: (1) downregulation of 503 genes involved in blood vessel development and angiogenesis; (2) upregulation of 738 genes related to cell adhesion and extracellular matrix remodelling [82]. Moreover, the lincRNA *TERMINATOR*, was expressed in cardiovascular progenitors, and was positively correlated with genes related to cell cycle, DNA repair and chromatin assembly, and negatively correlated with genes involved in proliferation regulation and cell death. *TERMINATOR* was also associated with pluripotency regulators *POU5F1 (Oct4)*, *SOX2*, *ZIC5*, and *REX1*, and its knockdown resulted in downregulation of 185 genes involved in cell-cell interactions and chromatin remodelling related to pluripotent maintenance. Finally, the anti-sense lincRNA *PUNISHER*, expressed in mature endothelial cells, was positively correlated with definitive vascular development genes, such as *GATA2*, and negatively correlated with genes involved in cell-cycle regulation, DNA damage response and chromatin modification. In addition, *PUNISHER*'s knockdown resulted in decrease of histone H3 phosphorylation and expression of 802 genes involved in cell division, as well as increased expression of 831 genes involved in cell adhesion and extracellular interactions, altogether, compromising endothelial cell function [82].

In another study in ESCs, 4,145 multi-exonic lincRNAs (2,608 novel) with an enhancer activity were identified during differentiation to cardiomyocytes. One of them was the *SMAD7-lincRNA*, proximal to the *Smad7* gene, and highly expressed in cardiac fibroblasts. *SMAD7-lincRNA* knockdown reduced *Smad7* gene expression, demonstrating the relationship between enhancer-derived lincRNAs and the expression of proximal target genes in cardiac fibroblasts [81].

In a transcriptome analysis of left and right ventricles of neonatal mouse heart, 196 of the 23,949 identified lincRNAs presented dynamic regulation along the cardiac maturation process. In addition, 5% of the 2,262 lincRNAs located within 50 kb of protein coding genes, were correlated with their neighbouring genes' expression (90.4% positively correlated) [84]. Furthermore, the impact of the *Ppp1r1b-lincRNA* on regulation of its neighbouring partner gene *Tcap* was demonstrated. *Tcap* encodes the *Titin Cap* protein that anchors Titin filaments at the Z disk in cardiomyocyte sarcomere organization and myogenesis. Interestingly, the expression ratio of *Ppp1r1b-lincRNA/Tcap* differed between normal cardiac maturation and congenital heart defect phenotypes [84].

In heart metabolism, the first lincRNA identified under physiological conditions was the *cardiac apoptosis-related lincRNA (CARL or AK017121)*. As a sponge lincRNA highly expressed in the heart, CARL binds miR-539 that negatively regulates PHB2. Thereby CARL maintains the expression of PHB2, suppressing the effects of anoxia, such as mitochondrial fission and cardiomyocyte apoptosis [85].

In other muscular tissues, the *myogenesis-associated lincRNA (linc-mg)* was recently identified in murine skeletal muscle development and during skeletal differentiation of muscle stem cells (muSCs). The *linc-mg* is enriched in skeletal muscle, and its expression is gradually increased during myocyte differentiation. The authors demonstrated that *linc-mg*

overexpression promotes cell differentiation *in vitro*, and muscle hypertrophy *in vivo*. Accordingly, knockout of *Inc-mg* in mice promoted muscle atrophy and loss of muscular endurance. The *Inc-mg* was shown to function as ceRNA, acting as sponge for miR-125b, a negative regulator of insulin-like growth factor 2 (Igf2). Thus, *Inc-mg* overexpression downregulates miR-125b, leading to Igf2 protein increase and enhancing the skeletal myogenesis [86].

As already mentioned above, in the *H19/Igf2* imprinted genomic locus, is also localized *H19* [87], a trans-regulatory lncRNA also involved in skeletal muscle differentiation and regeneration. *H19* is highly expressed in the majority of embryonic and neonatal tissues, especially in skeletal muscles, and is induced by MyoD, the main regulator of muscle differentiation. The *H19* first exon gives rise to miR-675-3p and miR-675-5p, two conserved miRNAs equally upregulated in skeletal muscle differentiation, which downregulate the DNA replication initiation factor *Cdc6* and the anti-differentiation transcription factors *Smad1* and *Smad5*. Thus, the miRNAs miR-675-3p and miR-675-5p, transcribed within *H19*, repress *Cdc6*, *Smad1* and *Smad5*, mediating skeletal muscle differentiation and regeneration [88]. Furthermore, *H19* may have a role in controlling timing of muscle differentiation, as this lncRNA also acts as a sponge for miRNAs of the let-7 family, which are involved in muscle differentiation anticipation and in glucose metabolism in muscle [88].

Another myogenic lncRNA is the *MyoD upstream noncoding (MUNC or DRR(eRNA))*, located 5 kb upstream the *MyoD* transcription start site (TSS), and which 5' end overlaps with the cis-acting distal regulatory region (DRR) of *MyoD*. *MUNC* lncRNA induces myoblast differentiation by promoting MyoD association with the DRR enhancer and with *Myogenin* promoter, and also by increasing the expression of *MyoD*, *Myogenin*, and *Myh3*, and the expression of non MyoD-inducible genes [89]. In addition, next to the *MyoD* gene is *LncMyoD*, a lncRNA activated by MyoD during myogenesis and myoblast differentiation. *LncMyoD* binds to IGF2-mRNA-binding protein 2 (IMP2), inhibiting N-Ras and c-Myc mRNAs translation, and then, facilitating cell-cycle exit and promoting terminal differentiation [90].

In the regulation of muscle differentiation, the muscle-specific lncRNA *linc-MD1* acts as a ceRNA in human and murine myoblasts, presenting a decoy activity for two different miRNAs, miR-133 and miR-135. These miRNAs regulate the expression of the transcription factors MAML1 and MEF2C, respectively, which activate the expression of muscle-specific genes. In this way, *Linc-MD1* controls the timing of myoblast differentiation, and when it is overexpressed, the anticipation of the muscle differentiation program occurs [91].

Angiotensin II (AngII) is a hormone involved in vasoconstriction, cellular growth and other processes. Transcriptome analysis of rat vascular smooth muscle cells (VSMCs) revealed five previously annotated lncRNAs and 24 novel lncRNAs regulated by AngII [92]. Among them, *Inc-Ang362*, an AngII-regulated lncRNA, was identified as the host transcript for miR-221 and miR-222, miRNAs involved in cell proliferation. Knockdown of *Inc-Ang362* leads to decrease of VSMCs proliferation [91]. The smooth muscle and endothelial cell-enriched migration/differentiation-associated lncRNA (*SENCR*) was seen in loss of function studies as related to decrease of myocardin expression and to increase of pro-migratory genes expression, which indicates that *SENCR* is a repressor of smooth muscle cell migration and possibly stabilizes the smooth muscle cell contractile phenotype [93].

Finally, it has been discovered that some annotated lncRNAs also contain ORFs that encode functional peptides, as the case of the muscle-specific lncRNA-derived micropeptide called

dwarf open reading frame (DWORF). DWORF has 34 aminoacids, is localized in sarcoplasmatic reticulum and enhances the activity of the Ca^{2+} adenosine triphosphatase SERCA calcium pump, enhancing muscle contractility [10]. In this process, SERCA inhibitors are displaced, including the myoregulin (MLN) [10], another micropeptide encoded by a skeletal muscle-specific putative lncRNA, which has an opposite effect to DWORF, inhibiting SERCA calcium pump and reducing muscle performance [11].

6.2.3.5 Adipose and bone tissues

Adipose tissue is composed predominantly by adipocytes, cells specialized in storing energy as fat. The white fat (subcutaneous and visceral) is the largest energy reservoir in the body, while brown (muscle-derived) and beige (or brite) fat (white-derived) are critical for thermogenesis, energy homeostasis, and metabolism. Besides these primary functions, adipose tissue also acts as a thermal insulation and as a protection against mechanical shocks. The prevalence of obesity has led to a surge of interest in understanding the mechanisms underlying adipocyte development, the process of differentiation of the pre-adipocyte precursor cells into adipocytes. There are strong evidences that lncRNAs participate in adipogenesis regulation in physiological and pathological contexts.

Global expression patterns of lncRNAs have been explored in adipocytes, and 175 lncRNAs were identified as specifically regulated during adipogenesis, a substantial fraction of which are under the control of the same key transcription factors that activate and repress coding genes during adipocyte differentiation [94].

During adipogenesis, most of the identified lncRNAs are involved in and/or bind the peroxisome proliferator-activated receptors (PPARs), a nuclear hormone receptor superfamily which regulates metabolic functions and energy homeostasis. The lncRNA *steroid receptor RNA activator* (SRA) has been observed as an adipogenesis promoter, acting in preadipocyte differentiation by associating with PPAR γ , a master transcriptional regulator of adipogenesis [95]. *NEAT1* is another lncRNA upregulated in adipogenesis, regulating the PPAR γ 2 splicing [96]. *NEAT1* expression is enhanced by *miR-140*, a regulator of developmental pathways and stem cells differentiation, leading to adipogenic phenotype [97]. Finally, the lnc-*U90926* decreases 3T3-L1 adipocyte differentiation in mice, via inhibiting the transactivation of PPAR γ 2 or PPAR γ [98].

PU.1 is a transcription factor involved in myelo-lymphoid differentiation and expressed also in the adipocyte lineage. In mice, *PU.1* antisense lncRNA (*PU.1 AS*) was shown to promote adipogenesis through preventing *PU.1* mRNA translation via binding to *PU.1* mRNA to form a *PU.1* mRNA/*PU.1 AS* lncRNA duplex in preadipocytes [99].

The lncRNA *HOTAIR* was found expressed in gluteal, but not in abdominal adipose tissue. The ectopic *HOTAIR* expression in abdominal preadipocytes increased the proportion of differentiated cells and the expression of important adipogenic genes, such as PPAR γ and lipoprotein lipase (LPL, the main enzyme involved with lipid storage in adipocyte). These data indicate that *HOTAIR* expression may underline the heterogeneity of human fat distribution, and might be involved in key processes of adipocyte differentiation [100].

Besides that, it has been demonstrated that adipocyte differentiation in brown fat is also regulated by lncRNAs, as the *nuclear brown fat lncRNA 1* (*Blnc1*), that forms a ribonucleoprotein complex with transcription factor EBF2, promoting brown and beige adipocyte differentiation [101]. This process is enhanced by the heterogeneous nuclear

ribonucleoprotein U (hnRNPU or hnRNP-U) assembly, which physically interacts both to *Blnc1* and to EBF2, inducing the ribonucleoprotein complex Blnc1/EBF2 formation and transcriptional activation of thermogenic genes in brown adipocytes [102].

The *Inc-BATE1* was identified in a murine transcriptome analysis as a lncRNA required for development and maintenance of brown adipocytes. *Lnc-BATE1* sustains the activation of brown adipose related genes, represses white adipose related genes and also interacts with hnRNP-U, during brown adipogenesis [103]. Moreover, *GM13133* is a lncRNA enriched in murine brown adipocytes and induced by cold exposure and β 3-adrenergic agonist and cAMP stimulation. When upregulated, *GM13133* inhibits white adipocytes differentiation and may increase the expression of brown adipose related genes and the mitochondrial biogenesis, possibly leading to conversion of white to brown adipocytes [104]. In addition, 704 upregulated and 896 downregulated lncRNAs were identified in brown adipose tissue when compared with skeletal muscle, indicating, again, an important role for lncRNAs in adipocytes differentiation, and in transdifferentiation between brown adipose tissue and skeletal muscle, both of which have abundant mitochondria and energy consumption capacity [105].

Adipose tissues have been also studied in osteogenic processes, where lncRNAs are also required. During human adipose stem cells (hASCs) differentiation, the expression of the tumour suppressor lncRNA *MEG3* was increased during osteogenesis and decreased during adipogenesis. Accordingly, *MEG3* knockdown blocked the osteogenic differentiation, promoting the adipogenic differentiation of these hASCs. Moreover, the adipogenic and osteogenic balance promoted by *MEG3* during hASCs differentiation might be modulated by miR-140-5p, a miRNA negatively correlated with *MEG3* in both processes [106].

The osteogenic differentiation is suppressed by cytokine-induced inflammation. Downregulation of the lncRNA *myocardial infarction-associated transcript (MIAT)* leads to osteogenic differentiation of hASCs and reverses the negative effects of the inflammation. Treatment with tumor necrosis factor increases *MIAT* expression, inhibiting hASCs osteoinduction [107]. The lncRNA *MIR31HG* was not only found involved in osteogenic differentiation of hASCs, but also regulating the inflammation. The interaction between *MIR31HG* and NF- κ B pathways may attenuate osteogenic differentiation of hASCs, leading to an inflammation-mediated inhibition of bone formation [108].

As mentioned above, the lncRNA *H19* is a key regulator for differentiation of various cell and tissue types, and gives rise to miR-675, that is partially responsible for the pro-osteogenic effect of *H19* by downregulating the transforming growth factor- β 1 (TGF- β 1), inhibiting Smad3 phosphorylation, and also downregulating histone deacetylase HDAC 4/5, reducing recruitment of HDACs to the Runt-related transcription factor 2 (Runx2)-binding DNA sequence [109].

Bone morphogenetic protein 9 (BMP9, or GDF2) is one of the most osteoinductive factors. The well-coordinated expression of lncRNA *H19* may be essential in BMP9-induced osteogenic differentiation of mesenchymal stem cells (MSCs). *H19* is highly expressed at the early stage of BMP9 stimulation of MSCs, followed by a rapid decrease and gradual return to the basal level. Possibly *H19* functions as a key mediator of BMP9 signalling by modulating miRNAs that target Notch receptors and/or ligands [110].

Profiling of RNAs of human bone marrow mesenchymal stem cells (hBMMSCs, or hBMSCs) revealed 433 upregulated and 232 downregulated lncRNAs during osteogenic differentiation. The lncRNA *ENST00000502125.2* was further examined and the results pointed to its role as a negative regulator of osteogenic differentiation, since its downregulation promoted osteogenic differentiation of hBMMSCs [111]. Global profiling of RNAs during osteogenic differentiation of hBMSCs found 785 lncRNAs upregulated and 623 lncRNAs downregulated along with osteogenic differentiation. Among the differentially expressed lncRNA were *XR_111050*, *NR_024031*, *FR374455*, *FR401275*, *FR406817*, and *FR148647*. One of them, lncRNA *XR_111050*, could enhance the osteogenic differentiation potential of MSCs [112].

Furthermore, the lncRNA *urothelial cancer-associated 1 (UCA1)* was found as expressed in human normal chondrocytes and inducing chondrocytic differentiation in a murine chondrocyte precursor [113]. Still during osteogenic differentiation, the lncRNA *differentiation antagonizing non-protein coding RNA (DANCR)* was decreased in human bone marrow-derived mesenchymal stem cells (HBMSCs). When silenced, the proliferation of HBMSCs and osteogenic differentiation is significantly enhanced, and vice versa. This negative control of osteogenic differentiation and proliferation by *DANCR* is through inactivation of p38 MAPK pathway, and *DANCR* is suggested by the authors as a potential therapeutic target [114].

6.2.3.6 Hepatic tissue

The liver is the biggest internal organ of the body, and also an exocrine and endocrine gland. As the adipose tissue, the liver is responsible, among other important functions, for adipogenesis and lipid metabolism. The lncRNAs involved in liver homeostasis are shown in this topic.

During liver development, the lncRNA *H19* was found in a ChIP-seq analysis as the most highly expressed lncRNA. *H19* blocks the interaction between actin and the heterogeneous hnRNP-U, affecting the Wnt/ β -catenin signalling pathway and proliferation of foetal liver cells [115].

In liver metabolism, the NAT lncRNA *APOA1-AS*, negatively regulates apolipoprotein A1 (*APOA1*) gene expression by epigenetic mechanisms. It was demonstrated, both *in vitro* and *in vivo*, that knockdown of *APOA1-AS* alters certain histone methylation patterns through the recruitment of histone-modifying enzymes, increasing the expression of the closely linked *APOA1*, *APOA4* and *APOC3* genes [116]. The process of cholesterol efflux and inhibition of cholesterol biosynthesis in the liver is promoted by ligand activation of liver X receptors (LXRs), and mediated by the lncRNA *LeXis*. *LeXis* can be induced by pharmacologic LXR activation and Western diet feeding, and interacts with ribonucleoprotein RALY, required for cholesterologenic gene expression, affecting in hepatic and plasmatic cholesterol levels in mice [117].

The *liver-specific triglyceride regulator lncLSTR* regulates plasmatic triglycerides (TG) in mice. The knockdown of *lncLSTR* leads to downregulation of *Cyp8b1*, altering the ratio of the most abundant bile acids, muricholic acid and cholic acid. This alteration results in upregulation of apoC2 expression and LPL activity, increasing FXR activity (an important bile acid receptor) and plasmatic TG clearance. However, the increased expression of apoC2 in primary hepatocytes did not lead to the same effects of TG regulation, indicating the presence of other hepatic factors involved in plasmatic TG clearance [118].

A novel sense-overlapping lncRNA named *lnc-HC* was identified in rats, and homologues exist in mice and humans. It overlaps the predicted 3'UTR region of the coding gene *Slc25a15* (mitochondrial carnitine transporter). The *lnc-HC* is a negative regulator of cholesterol metabolism, and is upregulated in hepatocytes through C/ERBb when exposed to high cholesterol. *lnc-HC* physically interacts with hnRNPA2B1, forming the lnc-HC/hnRNPA2B1 complex, which targets to mRNA of *Cyp7a1* and *Abca1*, decreasing their protein levels and, consequently, downregulating bile acid biosynthesis and cholesterol export [119].

Among the 359 tissue-specific lncRNAs expressed in metabolic homeostasis, found in transcriptome analysis of murine major metabolic organs, the liver metabolically sensitive lncRNA *Gm16551* was identified. The findings support a model in which SREBP1c, the master transcription factor of lipogenesis, induces *Gm16551* to activate an lncRNA-mediated negative feedback loop to regulate SREBP1c activity in the liver. The *Gm16551* knockdown led to upregulation of key lipogenic genes and to high circulating triglycerides levels [120].

The benefits of green tea in cholesterol metabolism, due to the high levels of catechins, mainly the epigallocatechin-3-gallate (EGCG) are well-known, but the molecular mechanisms underlying this effect are incompletely understood. Microarray analysis of RNA expressed in cultured human hepatocytes (HepG2) revealed 15 genes related to cholesterol metabolism and 285 lncRNAs whose levels were altered by EGCG treatment. Among them, the lncRNA *AT102202* mapped within the gene of its potential target, 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR). HMGCR is the rate-regulating enzyme in the cholesterol biosynthetic pathway. EGCG treatment enhanced *AT102202* expression and decreased expression of HMGCR. Downregulation of HMGCR expression leads to LDL receptor upregulation and cholesterol uptake by hepatocytes [121].

In liver regeneration, it was found *in vitro* that the lncRNA *MALAT1* (or *NEAT2*), upregulated in this process by p53, stimulates cell cycle progression from G1 to S phase and inhibits apoptosis, leading to hepatocyte proliferation. *MALAT1* also activated the Wnt/ β -catenin pathway and promoted cyclinD1 expression. These data indicate the critical role of *MALAT1* in liver regeneration, and its potential therapeutic benefit in liver failure and transplantation [122].

6.2.3.7 Intestinal tissue

In intestinal mucosa, the epithelial cells form a single-cell layer that constitutes the largest and most important barrier protecting the intestinal lumen against intraluminal toxins, antigens and enteric flora. The epithelium maintains its selective barrier function through the formation of complex protein-protein networks that mechanically link adjacent cells and seal the intercellular space.

Among the few studies investigating the lncRNAs in intestinal tissue, it was demonstrated that the lncRNA *SPRY4-IT1* increases the stability and translation of the mRNA encoding tight junction (TJ) proteins claudin-1, claudin-3, occludin, and JAM-1. Accordingly, the silencing of *SPRY4-IT1* repressed the expression of these TJ proteins in intestinal epithelial barrier. The direct interaction between *SPRY4-IT1* and TJ mRNAs is enhanced by the RNA-binding protein HuR. Patients with some intestinal diseases presented lower levels of *SPRY4-IT1*, suggesting that decreased *SPRY4-IT1* and the resulting lower expression of TJ proteins contribute to the pathogenesis of gut barrier dysfunction in patients [123]. Also regulating homeostasis of the intestinal epithelial barrier, the lncRNA *H19* may act as a

ceRNA, competing with the miR-874, a microRNA which suppresses the expression of the aquaporin-3 gene (*AQP3*) by binding to the 3'UTR of its mRNA. The *AQP3* downregulation opens the TJ complex, leading to dysfunction of intestinal barrier. *H19* competes with miR-874, reducing its levels and maintaining normal levels of *AQP3*, therefore playing a protective role in intestinal barrier [124]. However, overexpression of *H19* and the consequent increase of miR-675 inhibited synthesis of the TJ protein ZO-1 and the adherens junction protein E-cadherin, leading to intestinal epithelial barrier dysfunction. The epithelial barrier normal function was restored by increase of the expression of HuR in these cells [125].

A genome-wide transcriptome analysis of murine intestinal epithelial cells showed that TNF- α stimulation up- and downregulated several lincRNAs genes. Among them, the *lincRNA-Cox2* was upregulated in an early response to TNF- α , through the NF- κ B signalling pathway activation. *LincRNA-Cox2* knockdown altered the gene expression profile, in response to TNF- α stimulation. It was found that *lincRNA-Cox2* regulates the transcription of the proinflammatory cytokine interleukin 12b gene (*Il12b*) through epigenetic mechanisms, indicating a role of this lincRNA in inflammatory responses of the intestinal epithelia [126].

6.2.3.8 Gametogenesis and germ cells

Gametogenesis is a highly regulated complex cellular differentiation process. The accurate development of oocytes and sperm is among the most fundamental processes for the success of sexual reproduction. Similar to other reproductive traits, many features of gametogenesis and especially of spermatogenesis are under strong selective pressure to change. However, some fundamental features of gametogenesis are evolutionary conserved. Moreover, among all tissues, the transcriptome of testis displays the highest diversity and specificity.

The lncRNA/messenger RNA gene pair, *AK124742* and *PSMD6* presents elevated expression in human cumulus cells, the cluster of cells that surrounds the oocyte, and was correlated with oocyte maturation, fertilization, embryo quality and clinical pregnancy outcome [127]. Furthermore, a miRNA-lncRNA analysis indicated 41 lncRNAs able to interact with oocyte microRNAs and potentially involved in regulation of folliculogenesis [128]. In a study that investigated the role of lncRNAs in mediating the effects of β -crystallin B2 (*CRYBB2*) in ovary development in mice, RNA microarray profiling showed 157 differentially expressed lncRNAs and 1,085 differentially expressed mRNAs between ovary tissues from wild type and *CRYBB2* knockout mice. In *CRYBB2* knockout mice, expression of lncRNA A-30-P01019163 was downregulated, as well as its likely downstream target gene *P2rx7*, affecting ovarian development [129].

Concerning mammalian spermatogenesis, analysis of the Gc1-Spg cell line, derived from mouse spermatogonial cells, demonstrated that the lncRNA *mrhl* (meiotic recombination hot spot locus) negatively regulates the Wnt signaling pathway (that is implicated in regulation of mammalian spermatogenesis) through Ddx5/p68 RNA helicase [130]. Furthermore, genes at 37 loci (referred to as GRPAM) also showed altered expression upon *mrhl* downregulation. The p68 helicase interacted with *mrhl* in chromatin at these loci, suggesting a role for p68 mediated *mrhl* RNA occupancy in regulating GRPAM expression [131]. Moreover, a functional study reported that the lncRNA *AK015322* promotes proliferation of the mouse spermatogonial stem cell line C18-4 *in vitro* by function as a decoy of miR-19b-3p. *AK015322* antagonizes the function of that microRNA thus indirectly upregulating its endogenous target

transcriptional factor Ets-variant 5 (ETV5) which is a pivotal gene for spermatogonial stem cell self-renewal [132].

LncRNA profiling in mature mouse cells led to identification of 20,907 known and 4,088 novel lncRNAs. Compared to round spermatids, 1,794 lncRNAs were upregulated and 165 downregulated, and 14,259 potential target coding genes of differently expressed lncRNAs were suggested [133].

Purified cells at key stages of murine male spermatogenesis were submitted to RNA sequencing. The dynamic expression pattern of lncRNAs during progression of spermatogenesis was reported and 1,630 lncRNAs differently expressed in spermatogonia versus pachytene and 1,534 lncRNAs differently expressed in pachytene versus spermatids were identified. To examine the possible role of one lncRNA, a mouse knockout of *Tslrn1* (a X-linked lncRNA) was generated. These mice had normal fertility but expressive reduction in sperm count [134].

6.2.4 Extracellular vesicles

Extracellular vesicles (EVs) play an important role in intercellular communication through the transference of their cargo between cells. They are sub-micron membrane vesicles released to the extracellular environment by all cell types. According to their biogenesis, membrane, size and composition, EVs can be classified into exosomes, apoptotic bodies, microvesicles, ectosomes, among others [reviewed in 135 and 136]. EVs have been found in body fluids, including blood and urine. They are composed of a lipid bilayer from the plasma membrane and contain a variety of proteins, nucleic acids, lipids, and metabolites from donor cells, transferring biomolecular information between tissues and mediating diverse biological functions. Accordingly, RNA transcripts contained in the EVs also may be transferred to the target cells and present functional impact [137, 138]. The number of studies concerning lncRNAs in EVs is steadily increasing. Most aim at finding potential disease markers or at understanding pathological mechanisms of human diseases. Yet some of the findings are useful also for comprehension of physiological processes.

Interestingly, lncRNAs with low expression in cells may be enriched in EVs, as *lincRNA-p21*, *HOTAIR*, and *ncRNA-CCND1* [139, 140]. In fact, in a study comparing normal cell lines with a prostate cancer cell line, the population of exosomal lncRNAs differ between both cell lines. Furthermore, some lncRNAs enriched in EVs present complementary motifs to some microRNAs, such as members of the let-7 family, among others, suggesting that these lncRNAs can act as sponges for EV microRNAs, and to distinct RNA-binding proteins, such as ELAVL1 (HuR), which may play a role in exosomal transport of lncRNAs [140]. Another functional study provided evidence that specific lncRNAs are preferentially packaged into exosomes, and some of these exosomal lncRNAs (as *Exo2*, *Exo4* and *RMRP*) can modulate protein functions and cell viability by direct interactions with L-lactate dehydrogenase B (LDHB), high-mobility group protein 17 (HMG-17), and CSF2RB, inside recipient [141].

In a total transcriptome profiling of plasma-derived EVs of five healthy women, thousands of different transcripts were found. Among these, up to 220 lncRNA were detected. The lncRNAs *RN7SL1*, *RPPH1*, *VTRNA2-1*, *RMRP*, *RNU11* and *MALAT1* were the most highly expressed [142]. To our knowledge, this is the first study characterizing EV-lncRNA exclusively in healthy individuals. Thus, the EV-lncRNA cargo remains largely unknown, and the need for more research to characterize EVs contents and biological effects becomes

evident. Apart from the relevance of such studies for learning more about extracellular lncRNAs (and other biomolecules) in physiological settings, there is great interest in developing EVs analysis for “liquid biopsies”, which are minimally invasive assays that will be of value in biology and medicine.

6.3 Conclusions

Although lncRNA were overlooked for a long time, now their importance in various biological processes became evident. However, their description and especially their effects and functions in homeostasis (and in disease) remain largely elusive. This is not surprising in view of the complex interactions among the multiple players involved in every biological process, and the limitations of the available biological materials, and methods. These limitations certainly will be overcome by improvement of systems biology approaches in wet laboratories as well as bioinformatics, and development of more powerful integrative analysis methods. In this review, we provided a broad overview of the information available in the literature regarding lncRNA identification and/or effects in specific mammalian - and especially human - cell and tissue types, aiming at understanding their effects in homeostasis. We limited our search to cells and tissues for which presently there is growing interest among researchers. Future studies will expand the knowledge about lncRNAs in cell biology and physiological processes, contributing to a deeper understanding of their function in the context of the multiple interactions between lncRNAs, other ncRNAs, mRNAs, proteins, and small molecules. Furthermore, information about lncRNAs in homeostasis provides a framework for future studies of their functions in health and disease and in development of novel disease markers and therapeutic targets.

Supplementary Material: Table S1. lncRNAs described in homeostasis of different cells and tissues.

Acknowledgments: We acknowledge the Graduate Program of Genetics of the Department of Genetics at the Federal University of Paraná, and the funding agencies CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

Conflicts of Interest: The authors declare no conflict of interest.

6.4 References

- [1] Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S.; Barnes, I.; Bignell, A.; Boychenko, V.; Hunt, T.; Kay, M.; Mukherjee, G.; Rajan, J.; Despacio-Reyes, G.; Saunders, G.; Steward, C.; Harte, R.; Lin, M.; Howald, C.; Tanzer, A.; Derrien, T.; Chrast, J.; Walters, N.; Balasubramanian, S.; Pei, B.; Tress, M.; Rodriguez, J.M.; Ezkurdia, I.; van Baren, J.; Brent, M.; Haussler, D.; Kellis, M.; Valencia, A.; Reymond, A.; Gerstein, M.; Guigó, R.; Hubbard, T.J. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **2012**, *22*, 1760–1774, DOI: 10.1101/gr.135350.111.
- [2] Iyer, M.K.; Niknafs, Y.S.; Malik, R.; Singhal, U.; Sahu, A.; Hosono, Y.; Barrette, T.R.; Prensner, J.R.; Evans, J.R.; Zhao, S.; Poliakov, A.; Cao, X.; Dhanasekaran, S.M.; Wu, Y.M.; Robinson, D.R.; Beer, D.G.; Feng, F.Y.; Iyer, H.K.; Chinnaiyan, A.M. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* **2015**, *47*(3), 199–208, DOI: 10.1038/ng.3192
- [3] 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 Res.* **2012**, *22*, 1775-1789, DOI: 10.1101/gr.132159.111.

[4] Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol Cell.* **2011**, *43*, 904-914; DOI: 10.1016/j.molcel.2011.08.018.

[5] Moran, V.A.; Perera R.J.; Khalil, A.M. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic Acids Res.* **2012**, *40*, 6391–6400; DOI: 10.1093/nar/gks296

[6] Brockdorff, N.; Ashworth, A.; Kay, G.F.; McCabe, V.M.; Norris, D.P.; Cooper, P.J.; Swift, S.; Rastan, S. The Product of the Mouse Xist Gene Is a 15 kb Inactive X-Specific Transcript Containing no Conserved ORF and Located in the Nucleus. *Cell* **1992**, *71*, 515-526.

[7] Brown, C.J.; Hendrich, B.D.; Rupert, J.L.; Lafrenière, R.G.; Xing, Y.; Lawrence, J.; Willard, H.F. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **1992**, *71*, 527-542.

[8] Quinn J.J.; Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet.* **2016**, *17*, 47–62; DOI: 10.1038/nrg.2015.10.

[9] Jalali, S.; Jayaraj, G.G.; Scaria, V. Integrative transcriptome analysis suggest processing of a subset of long non-coding RNAs to small RNAs. *Biol Direct* **2012**, *7*:25, DOI: 10.1186/1745-6150-7-25.

[10] Nelson, B.R.; Makarewich, C.A.; Anderson, D.M.; Winders, B.R.; Troupes, C.D.; Wu, F.; Reese, A.L.; McAnally, J.R.; Chen, X.; Kavalali, E.T.; Cannon, S.C.; Houser, S.R.; Bassel-Duby, R.; Olson, E.N. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* **2016**, *351*, 271-275; DOI: 10.1126/science.aad4076.

[11] Anderson, D.M.; Anderson, K.M.; Chang, C.L.; Makarewich, C.A.; Nelson, B.R.; McAnally, J.R.; Kasaragod, P.; Shelton, J.M.; Liou, J.; Bassel-Duby, R.; Olson, E.N. A Micropeptide Encoded by a Putative long noncoding RNA Regulates Muscle Performance. *Cell* **2015**, *160*, 595–606, DOI: 10.1016/j.cell.2015.01.009.

[12] Dinger, M.E.; Amaral, P.P.; Mercer, T.R.; Pang, K.C.; Bruce, S.J.; Gardiner, B.B.; Askarian-Amiri, M.E.; Ru, K.; Soldà, G.; Simons, C.; Sunkin, S.M.; Crowe, M.L.; Grimmond, S.M.; Perkins, A.C.; Mattick, J.S. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **2008**, *18*, 1433–1445, DOI: 10.1101/gr.078378.108.

[13] Mercer, T.R.; Dinger, M.E.; Sunkin, S.M.; Mehler, M.F.; Mattick, J.S. Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci U S A.* **2008**, *105*, 716–721; DOI: 10.1073/pnas.0706729105

[14] Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; Cabili, M.N.; Jaenisch, R.; Mikkelsen, T.S.; Jacks, T.; Hacohen, N.; Bernstein, B.E.; Kellis, M.; Regev, A.; Rinn, J.L.; Lander, E.S. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223-227, DOI: 10.1038/nature07672.

[15] Wilusz, J.E.; Sunwoo, H.; Spector, D.L. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* **2009**, *23*, 1494–1504; DOI: 10.1101/gad.1800909.

- [16] Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; Shafer, A.; Neri, F.; Lee, K.; Kuttyavin, T.; Stehling-Sun, S.; Johnson, A.K.; Canfield, T.K.; Giste, E.; Diegel, M.; Bates, D.; et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. **2012**, *337*, 1190–1195. DOI: 10.1126/science.1222794.
- [17] Hrdlickova, B.; Kumar, V.; Kanduri, K.; Zhernakova, D.V.; Tripathi, S.; Karjalainen, J.; Lund, R.J.; Li, Y.; Ullah, U.; Modderman, R.; Abdulahad, W.; Lähdesmäki, H.; Franke, L.; Lahesmaa, R.; Wijmenga, C.; Withoff, S. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. *Genome Med.* **2014**, *6*(10):88, DOI: 10.1186/s13073-014-0088-0
- [18] Venkatraman, A.; He, X.C.; Thorvaldsen, J.L.; Sugimura, R.; Perry, J.M.; Tao, F.; Zhao, M.; Christenson, M.K.; Sanchez, R.; Yu, J.Y.; Peng, L.; et al. Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature*. **2013**, *500*, 345–349; DOI: 10.1038/nature12303.
- [19] Keniry, A.; Oxley, D.; Monnier, P.; Kyba, M.; Dandolo, L.; Smits, G.; Reik, W. The H19 lincRNA is a developmental reservoir of miR-675 which suppresses growth and Igf1r. *Nat Cell Biol.* **2012**, *14*(7), 659–665, DOI: 10.1038/ncb2521.
- [20] Luo, M.; Jeong, M.; Sun, D.; Park, H.J.; Rodriguez, B.A.; Xia, Z.; Yang, L.; Zhang, X.; Sheng, K.; Darlington, G.J.; Li, W.; Goodell, M.A. Long non-coding RNAs control hematopoietic stem cell function. *Cell Stem Cell.* **2015**, *16*, 426–438; DOI: 10.1016/j.stem.2015.02.002.
- [21] Villamizar, O.; Chambers, C.B.; Mo, Y.Y.; Torry, D.S.; Hofstrand, R.; Riberdy, J.M.; Persons, D.A.; Wilber, A. Fas-antisense long noncoding RNA is differentially expressed during maturation of human erythrocytes and confers resistance to Fas-mediated cell death. *Blood Cells Mol Dis.* **2016**, *58*, 57-66. DOI: 10.1016/j.bcmd.2016.03.002.
- [22] Alvarez-Dominguez, J.R.; Hu, W.; Yuan, B.; Shi, J.; Park, S.S.; Gromatzky, A.A.; van Oudenaarden, A.; Lodish, H.F. Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation. *Blood* **2014**, *123*, 570-581, DOI: 10.1182/blood-2013-10-530683.
- [23] Hu, W.; Yuan, B.; Flygare, J.; Lodish, H.F. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev.* **2011**, *25*, 2573–2578, DOI: 10.1101/gad.178780.111.
- [24] Ding, N.; Xi, J.; Li, Y.; Xie, X.; Shi, J.; Zhang, Z.; Li, Y.; Fang, F.; Wang, S.; Yue, W.; Pei, X.; Fang, X. Global transcriptome analysis for identification of interactions between coding and noncoding RNAs during human erythroid differentiation. *Front. Med.* **2016**, *10*(3), 297-310, DOI: 10.1007/s11684-016-0452-0.
- [25] Doss, J.F.; Corcoran, D.L.; Jima, D.D.; Telen, M.J.; Dave, S.S.; Chi, J.T. A comprehensive joint analysis of the long and short RNA transcriptomes of human erythrocytes. *BMC Genomics* **2015**, *16*:952, DOI: 10.1186/s12864-015-2156-2.
- [26] Kotzin, J.J.; Spencer, S.P.; McCright, S.J.; Kumar, D.B.; Collet, M.A.; Mowel, W.K.; Elliott, E.N.; Uyar, A.; Makiya, M.A.; Dunagin, M.C.; Harman, C.C.D.; Virtue, A.T.; Zhu, S.; Bailis, W.; Stein, J.; Hughes, C.; Raj, A.; Wherry, E.J.; Goff, L.A.; Klion, A.D.; Rinn, J.L.; Williams, A.; Flavell, R.A.; Henao-Mejia, J. The long non-coding RNA Morbid regulates Bim and short-lived myeloid cell lifespan. *Nature* **2016**, *537*, 239–243, DOI: 10.1038/nature19346.
- [27] Mirsafian, H.; Manda, S.S.; Mitchell C.J.; Sreenivasamurthy, S.-L.; Ripen A.M.; Mohamad S.B.; Merican A.F.; Pandey, A. Long non-coding RNA expression in human primary monocytes. *Genomics.* **2016**, *108*, 37-45; DOI: 10.1016/j.ygeno.2016.01.002.

- [28] Xin, J.; Li, J.; Feng, Y.; Wang, L.; Zhang, Y.; Yang, R. Downregulation of long noncoding RNA HOTAIRM1 promotes monocyte/dendritic cell differentiation through competitively binding to endogenous mir-3960. *Onco Targets Ther.* **2017**, *10*, 1307–1315; DOI: 10.2147/OTT.S124201.
- [29] Carpenter, S.; Aiello, D.; Atianand, M.K.; Ricci, E.P.; Gandhi, P.; Hall, L.L.; Byron, M.; Monks, B.; Henry-Bezy, M.; Lawrence, J.B.; O'Neill, L.A.; Moore, M.J.; Caffrey, D.R.; Fitzgerald, K.A. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **2013**, *341*, 789–792, DOI: 10.1126/science.1240925.
- [30] Hu, G.; Gong, A.Y.; Wang, Y.; Ma, S.; Chen, X.; Chen, J.; Su, C.J.; Shibata, A.; Strauss-Soukup, J.K.; Drescher, K.M.; Chen, X.M. LincRNA-Cox2 promotes late inflammatory gene transcription in macrophages through modulating SWI/SNF-mediated chromatin remodeling. *J. Immunol.* **2016**, *196*, 2799–2808, DOI: 10.4049/jimmunol.1502146.
- [31] Wang, P.; Xue, Y.; Han, Y.; Lin, L.; Wu, C.; Xu, S.; Jiang, Z.; Xu, J.; Liu, Q.; Cao, X. The STAT3-Binding Long Noncoding RNA Inc-DC Controls Human Dendritic Cell Differentiation. *Science*. **2014**, *344*, 310; DOI: 10.1126/science.1251456.
- [32] Dijkstra, J.M.; Ballingall, K.T. Non-human Inc-DC orthologs encode Wdm1-like protein. *F1000Res.* **2014**, *3*:160, DOI: 10.12688/f1000research.4711.2.
- [33] Casero, D.; Sandoval, S.; Seet, C.S.; Scholes, J.; Zhu, Y.; Ha, V.L.; Luong, A.; Parekh, C.; Crooks, G.M. Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. *Nat Immunol.* **2015**, *16*(12), 1282–1291, DOI: 10.1038/ni.3299.
- [34] Wang, F.Y.; Luo, L.; Zhao, S.S.; Wei, H.P.; Wang, L.; Gu, Z.Y.; Guan, L.X.; Yang, N.; Gao, Z.; Wang, L.L.; Gao, C.J. Expression Profile of lncRNA NONHSAT040475 in Peripheral Blood Leukocytes of Healthy Persons. *Zhongguo Shi Yan Xue Ye Xue Za Zhi.* **2016**, *24*, 1196–1200; DOI: 10.7534/j.issn.1009-2137.2016.04.043.
- [35] Pang, K.C.; Dinger, M.E.; Mercer, T.R.; Malquori, L.; Grimmond, S.M.; Chen, W.; Mattick, J.S. Genome-Wide Identification of Long Noncoding RNAs in CD8⁺ T Cells. *J. Immunol.* **2009**, *182*, 7738–7748, DOI: 10.4049/jimmunol.0900603.
- [36] Sharma, S.; Findlaya, G.M.; Bandukwalaa, H.S.; Oberdoerffera, S.; Bausta, B.; Lib, Z.; Schmidt, V.; Hogana, P.G.; Sacksb, D.B.; Raoa, A. Dephosphorylation of the nuclear factor of activated T cells (NFAT) transcription factor is regulated by an RNA-protein scaffold complex. *Proc Natl Acad Sci U S A.* **2011**, *108*, 11381–11386; DOI: 10.1073/pnas.1019711108.
- [37] Collier, S.P.; Collins, P.L.; Williams, C.L.; Boothby, M.R.; Aune, T.M. Cutting Edge: Influence of Tmevpg1, a Long Intergenic Noncoding RNA, on the Expression of Ifng by Th1 Cells. *J Immunol.* **2012**, *189*, 2084–2088, DOI: 10.4049/jimmunol.1200774.
- [38] Hu, G.; Tang, Q.; Sharma, S.; Yu, F.; Escobar, T.M.; Muljo, S.A.; Zhu, J.; Zhao, K. Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nature Immunology* **2013**, *14*, 1190–1198, DOI: 10.1038/ni.2712.
- [39] Bonnal, R.J.P.; Ranzani, V.; Arrigoni, A.; Curti, S.; Panzeri, I.; Gruarin, P.; Abrignani, S.; Rossetti, G.; Pagani, M. De novo transcriptome profiling of highly purified human lymphocytes primary cells. *Scientific Data* **2015**, *2*. DOI: 10.1038.2015
- [40] Ranzani, V.; Rossetti, G.; Panzeri, I.; Arrigoni, A.; Bonnal, R.J.P.; Curti, S.; Gruarin, P.; Provasi, E.; Sugliano, E.; Marconi, M.; De Francesco, R.; Geginat, J.; Bodega, B.; Abrignani, S.; Pagani, M. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat Immunol.* **2015**, *16*, 318–325; DOI: 10.1038/ni.3093.

- [41] Spurlock, C.F. 3rd; Tossberg, J.T.; Guo, Y.; Collier, S.P.; Crooke, P.H.; Aune, T.M. Expression and functions of long noncoding RNAs during human T helper cell differentiation. *Nat commun.* **2015**, *6*, 6932, DOI: 10.1038/ncomms7932.
- [42] Zhang, H.; Nestor, C.E.; Zhao, S.; Lentini, A.; Bohle, B.; Benson, M.; Wang, H. Profiling of human CD41 T-cell subsets identifies the TH2-specific noncoding RNA GATA3-AS1. *J Allergy Clin Immunol.* **2013**, *132*, 1005-1008; DOI: 10.1016/j.jaci.2013.05.033.
- [43] Xia, F.; Dong, F.; Yang, Y.; Huang, A.; Chen, Si.; Sun, D.; Xiong, S.; Zhang, J. Dynamic Transcription of Long Non-Coding RNA Genes during CD4⁺ T Cell Development and Activation. *PLoS ONE.* **2014**, *9*, e101588; DOI:10.1371/journal.pone.0101588.
- [44] Kanduri, K.; Tripathi, S.; Larjo, A.; Mannerström, H.; Ullah, U.; Lund, R.; Hawkins, R.D.; Ren, B.; Lähdesmäki, H.; Laheesmaa, R. Identification of global regulators of T-helper cell lineage specification. *Genome Med.* **2015**, *7*, 122, DOI: 10.1186/s13073-015-0237-0.
- [45] Spurlock, C.F. 3rd; Shaginurova, G.; Tossberg, J.T.; Hester, J.D.; Chapman, N.; Guo, Y.; Crooke, P.S. 3rd; Aune, T.M. Profiles of Long Noncoding RNAs in Human Naive and Memory T Cells. *J immunol.* **2017**, *199*, 547-558; DOI: 10.4049/jimmunol.1700232.
- [46] Petri, A.; Dybkær, K.; Bøgsted, M.; Thruø C.A.; Hagedorn, P.H.; Schmitz, A. Bødker, J.S.; Johnsen, H.E.; Kauppinen, S. Long Noncoding RNA Expression during Human B-Cell Development. *PLoS ONE.* **2015**, *10*, e0138236; DOI:10.1371/journal.pone.0138236.
- [47] Brazão, T.F.; Johnson, J.S.; Müller, J.; Heger, A.; Ponting, C.P.; Tybulewicz, V.L.J. Long non-coding RNAs in B cell development and activation. *Blood* **2016**, *128*(7), e10-e19, DOI: 10.1182/blood-2015-11-680843.
- [48] Syrett, C.M.; Sindhava, V.; Hodawadekar, S.; Myles, A.; Liang, G.; Zhang, Y.; Nandi, S.; Cancro, M.; Atchison, M.; Anguera, M.C. Loss of Xist RNA from the inactive X during B cell development is restored in a dynamic YY1-dependent two-step process in activated B cells. *PLoS Genet.* **2017**, *13*, e1007050; DOI: 10.1371/journal.pgen.1007050.
- [49] Zhang, R.; Ni, F.; Fu, B.; Wu, Y.; Sun, R.; Tian, Z.; Wei, H. A long noncoding RNA positively regulates CD56 in human natural killer cells. *Oncotarget.* **2016**, *7*, 72546-72558; DOI: 10.18632/oncotarget.12466.
- [50] Wright, P.W.; Huehn, A.; Cichocki F.; Li, H.; Sharma, N.; Dang, H.; Lenvik, T.R.; Woll, P.; Kaufman, D.; Miller, J.S.; Anderson, S.K. Identification of a KIR antisense lncRNA expressed by progenitor cells. *Genes Immun.* **2013**, *14*, 427-433; DOI: 10.1038/gene.2013.36.
- [51] Mowle, W.K.; McCright, S.J.; Kotzin, J.J.; Collet, M.A.; Uyar, A.; Chen, X.; DeLaney, A.; Spencer S.P.; Virtue, A.T.; Yang, E.; et al. Group 1 innate lymphoid cell lineage identity is determined by a cis-regulatory element marked by a long non-coding RNA. *Immunity.* **2017**, *47*, 435-449; DOI: 10.1016/j.immuni.2017.08.012.
- [52] Beltran, M.; Puig, I.; Peña, C.; García, J.M.; Álvarez, A.B.; Peña, R.; Bonilla, F.; Herreros, A.G. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes dev.* **2008**, *22*, 756-769, DOI: 10.1101/gad.455708.
- [53] Lopez-Pajares, V.; Qu, K.; Zhang, J.; Webster, D.E.; Barajas, B.C.; Siprashvili, Z.; Zarnegar, B.J.; Boxer, L.D.; Rios, E.J.; Tao, S.; Kretz, M.; Khavari, P.A. A LncRNA-MAF:MAFB Transcription Factor Network Regulates Epidermal Differentiation, *Dev Cell.* **2015**, *32*, 693-706; DOI: 10.1016/j.devcel.2015.01.028.
- [54] Mazar J.; Sinha, S.; Dinger, M.E.; Mattick, J.S.; Perera, R.J. Protein-coding and non-coding gene expression analysis in differentiating human keratinocytes using a three-

dimensional epidermal equivalent. *Mol Genet Genomics*. **2010**, 284,1–9. DOI: 10.1007/s00438-010-0543-6.

[55] Ørom, U.A.; Derrien, T.; Guigo, R.; Shiekhattar, R. Long Noncoding RNAs as Enhancers of Gene Expression. *Cold Spring Harb Symp Quant Biol*. **2010**, 75, 325-331; DOI: 10.1101/sqb.2010.75.058.

[56] Kretz, M.; Webster, D.E.; Flockhart, R.J.; Lee, C.S.; Zehnder, A.; Lopez-Pajares, V.; Qu, K.; Zheng, G.X.Y.; Chow, J.; Kim, G.E.; Rinn, J.L.; Chang, H.Y.; Siprashvili, Z.; Khavari, P.A. Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev*. **2012**, 26, 338–343, DOI: 10.1101/gad.182121.111.

[57] Kretz, M.; Siprashvili, Z.; Chu, C.; Webster, D.E.; Zehnder, A.; Qu, K.; Lee, C.S.; Flockhart, R.J.; Groff, A.F.; Chow, J.; Johnston, D.; Kim, G.E.; Spitale, R.C.; Flynn, R.A.; Zheng, G.X.Y.; Aiyer, S.; Raj, A.; Rinn, J.L.; Chang, H.Y.; Khavari, P.A. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **2013**, 493, 231-235, DOI: 10.1038/nature11661.

[58] Li, C.X.; Li, H.G.; Huang, L.T.; Kong, Y.W.; Chen, F.Y.; Liang, J.Y.; Yu, H.; Yao, Z.R. H19 lncRNA regulates keratinocyte differentiation by targeting miR-130b-3p. *Cell Death Dis*. **2017**, 8, doi:10.1038/cddis.2017.516.

[59] Lin, C.M.; Liu, Y.; Huang, K.; Chen, X.; Cai, B.; Li, H.; Yuan, Y.; Zhang, H.; Li, Y. Long noncoding RNA expression in dermal papilla cells contributes to hairy gene regulation. *Biochem Biophys Res Commun*. **2014**, 453, 508–514, DOI: 10.1016/j.bbrc.2014.09.119.

[60] Mullin, N.K.; Mallipeddi, N.V.; Hamburg-Shields E.; Ibarra, B.; Khalil, A.M.; Atit, R.P. Wnt/ β -catenin Signaling Pathway Regulates Specific lncRNAs That Impact Dermal Fibroblasts and Skin Fibrosis. *Front. Genet*. **2017**, 8, 183; DOI: 10.3389/fgene.2017.00183.

[61] Qureshi, I. A.; Mehler, M.F. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci*. **2012**, 13, 528–541; DOI: 10.1038/nrn3234.

[62] Pollard, K.S.; Salama, S.R.; Lambert, N.; Lambot, M.A.; Coppens, S.; Pedersen, J.S.; Katzman, S.; King, B.; Onodera, C.; Siepel, A.; Kern, A.D.; Dehay, C.; Igel, H.; Ares, M. Jr.; Vanderhaeghen, P.; Haussler, D. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*. **2006**, 443, 167-72; DOI: 10.1038/nature05113.

[63] Bond, A.M.; VanGompel, M.J.W.; Sametsky, E.A.; Clark, M.F.; Savage, J.C.; Disterhoft, J.F.; Kohtz, J.D. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat Neurosci*. **2009**, 12, 1020-1027, DOI: 10.1038/nn.2371.

[64] Ramos, A.D.; Andersen, R.E.; Liu, S.J.; Nowakowski, T.J.; Hong, S.J.; Gertz, C.; Salinas, R.D.; Zarabi, H.; Kriegstein, A.; Lim, D.A. The long noncoding RNA *Pnky* regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell*. **2015**, 16, 439–447; DOI: 10.1016/j.stem.2015.02.007.

[65] Rani, N.; Nowakowski, T.J.; Zhou H.; Godshalk, S.E.; Lisi V.; Kriegstein A.R.; Kosik K.S. A Primate lncRNA Mediates Notch Signaling During Neuronal Development by Sequestering miRNA. *Neuron*. **2016**, 90, 1174–1188; DOI: 10.1016/j.neuron.2016.05.005.

[66] Lin, M.; Pedrosa, E.; Shah, A.; Hrabovsky, A.; Maqbool S.; Zheng, D.; Lachman, H.M. RNA-Seq of Human Neurons Derived from iPS Cells Reveals Candidate Long Non-Coding RNAs Involved in Neurogenesis and Neuropsychiatric Disorders. *PLoS ONE*. **2011**, 6, e23356; DOI:10.1371/journal.pone.0023356.

- [67] Ng, S.Y.; Johnson, R.; Stanton, L.W. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J.* **2012**, *31*, 522–533; DOI: 10.1038/emboj.2011.459.
- [68] Wu, A.M.; Ni, W.F.; Huang, Z.Y.; Li, Q.L.; Wu, J.B.; Xu H.Z.; Yin, L.H. Analysis of differentially expressed lncRNAs in differentiation of bone marrow stem cells into neural cells. *J. Neurol Sci.* **2015**, *35*, 160–167; DOI: 10.1016/j.jns.2015.03.011.
- [69] Zhang, Y.; Chen, K.; Sloan, S.A.; Bennett, M.L.; Scholze, A.R.; O'Keefe, S.; Phatnani H.P.; Guarnieri, P.; Caneda, C.; Ruderisch, N.; Deng, S.; Liddelow, S.A.; Zhang, C.; Daneman, R.; Maniatis, T.; Barres, B.A.; Wu, J.Q. An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex. *J Neurosci.* **2014**, *34*, 11929–11947; DOI: 10.1523/JNEUROSCI.1860-14.2014.
- [70] Guo, Y.; Zhang, P.; Sheng, Q.; Zhao, S.; Hackett, T.A. lncRNA expression in the auditory forebrain during postnatal development. *Gene* **2016**, *593*, 201–216, DOI: 10.1016/j.gene.2016.08.027.
- [71] Liu, S.J.; Nowakowski, T.J.; Pollen A.A.; Lui J.H.; Horlbeck, M.A.; Attenello, F.J.; He, D.; Weissman, J.S.; Kriegstein, A.R.; Diaz, A.A.; Lim, D.A. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* **2016**, *17*, 67; DOI: 10.1186/s13059-016-0932-1.
- [72] Aliperti, V.; Donizetti, A. Long Non-coding RNA in Neurons: New Players in Early Response to BDNF Stimulation. *Front. Mol. Neurosci.* **2016**, *9*:15, DOI: 10.3389/fnmol.2016.00015.
- [73] He, D.; Wang, J.; Lu, Y.; Deng, Y.; Zhao, C.; Xu, L.; Chen, Y.; Hu, Y.C.; Zhou, W.; Lu, Q.R. lncRNA Functional Networks in Oligodendrocytes Reveal Stage-Specific Myelination Control by an lncOL1/Suz12 Complex in the CNS. *Neuron.* **2017**, *93*(2), 362–378, DOI: 10.1016/j.neuron.2016.11.044.
- [74] Bernard, D.; Prasanth, K.V.; Tripathi, V.; Colasse, S.; Nakamura, T.; Xuan, Z.; Zhang, M.Q.; Sedel, F.; Jourdain, L.; Couplier, F.; Triller, A.; Spector, D.L.; Bessis, A. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J.* **2010**, *29*, 3082–3093, DOI: 10.1038/emboj.2010.199.
- [75] Tan M.C.; Widagdo, J.; Chau, Y.Q.; Zhu, T.; Wong, J.J.; Cheung, A.; Anggono, V. The Activity-Induced Long Non-Coding RNA Meg3 Modulates AMPA Receptor Surface Expression in Primary Cortical Neurons. *Front. Cell. Neurosci.* **2017**, *11*, 124. DOI: 10.3389/fncel.2017.00124.
- [76] Bruneau, B.G. Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harb Perspect Biol.* **2013**, *5*(3), DOI: 10.1101/cshperspect.a008292.
- [77] Ounzain, S.; Burdet, F.; Ibberson, M.; Pedrazzini, T. Discovery and functional characterization of cardiovascular long noncoding RNAs. *J Mol Cell Cardiol.* **2015**, *89*, 17–26; DOI: 10.1016/j.jmcc.2015.09.013.
- [78] Klattenhoff, C.A.; Scheuermann, J.C.; Surface, L.E.; Bradley, R.K.; Fields, P.A.; Steinhauser, M.L.; Ding, H.; Butty, V.L.; Torrey, L.; Haas, S.; Abo, R.; Tabebordbar, M.; Lee, R.T.; Burge, C.B.; Boyer, L.A. Braveheart, a Long Noncoding RNA Required for Cardiovascular Lineage Commitment. *Cell* **2013**, *152*, 570–583, DOI: 10.1016/j.cell.2013.01.003.

- [79] Grote, P.; Wittler, L.; Hendrix, D.; Koch, F.; Währisch, S.; 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. *Dev Cell*. **2013**, *24*(2), 206–214, DOI: 10.1016/j.devcel.2012.12.012.
- [80] Anderson, K.M.; Anderson D.M.; McAnally, J.R.; Shelton, J.M.; Bassel-Duby, R.; Olson, E.N. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **2016**, *539*, 433–436, DOI: 10.1038/nature20128.
- [81] Ounzain, S.; Micheletti, R.; Arnan, C.; Plaisance, I.; Cecchi, D.; Schroen, B.; Reverter, F.; Alexanian, M.; Gonzales, C.; Ng, S.Y. et al. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *J Mol Cell Cardiol*. **2015**, *89*, 98–112; DOI: 10.1016/j.yjmcc.2015.09.016.
- [82] Kurian, L.; Aguirre, A.; Sancho-Martinez, I.; Benner, C.; Hishida, T.; Nguyen, T.B.; Reddy, P.; Nivet, E.; Krause, M.N.; Nelles, D.A.; Rodriguez Esteban, C.; Campistol, J.M.; Yeo, G.W.; Belmonte, J.C.I. Identification of Novel Long Noncoding RNAs Underlying Vertebrate Cardiovascular Development. *Circulation* **2015**, *131*(14), 1278–1290, DOI: 10.1161/CIRCULATIONAHA.114.013303.
- [83] Ounzain, S.; Pezzuto I.; Micheletti, R.; Burdet, F.; Sheta, R.; Nemir, M.; Gonzales, C.; Sarre, A.; Alexanian, M.; Blow, M.J.; et al. Functional importance of cardiac enhancer-associated noncoding RNAs in heart development and disease. *J. Mol Cell Cardiol*. **2014**, *76*, 55–70; DOI: 10.1016/j.yjmcc.2014.08.009.
- [84] Touma, M.; Kang, X.; Zhao, Y.; Cass, A.A.; Gao, F.; Biniwale, R.; Coppola, G.; Xiao, X.; Reemtsen, B.; Wang, Y. Decoding the Long Noncoding RNA During Cardiac Maturation: A Roadmap for Functional Discovery. *Circ Cardiovasc Genet*. **2016**, *9*, 395–407; DOI: 10.1161/CIRCGENETICS.115.001363.cardiomyocyte apoptosis
- [85] Wang, K.; Bo, Long.; Zhou, L.-Y.; Fang, Liu.; Zhou, Q.Y.; Liu, C.-Y.; Fan Y.Y.; Li, P.F. CARL lncRNA inhibits anoxia-induced mitochondrial fission and apoptosis in cardiomyocytes by impairing miR-539-dependent PHB2 downregulation. *Nat. Commun*. **2014**, *5*, 3596; DOI: 10.1038/ncomms4596.
- [86] Zhu, M.; Liu, J.; Xiao, J.; Yang, L.; Cai, M.; Shen, H.; Chen, X.; Ma, Y.; Hu, S.; Wang, Z.; Hong, A.; Li, Y.; Sun, Y.; Wang, X. Lnc-mg is a long non-coding RNA that promotes myogenesis. *Nat Commun*. **2017**, *8*, 14718; DOI: 10.1038/ncomms14718.
- [87] Kallen, A.N.; Zhou, X.B.; Xu, J.; Qiao, C.; Ma, J.; Yan, L.; Lu, L.; Liu, C.; Yi, J.S.; Zhang, H.; Min, W.; Bennett, A.M.; Gregory, R.I.; Ding, Y.; Huang, Y. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell*. **2013**, *52*(1), 101–112, DOI: 10.1016/j.molcel.2013.08.027.
- [88] Dey, B.K.; Pfeifer, K.; Dutta, A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev*. **2014**, *28*(5), 491–501, DOI: 10.1101/gad.234419.113.
- [89] Mueller, A.C.; Cicheviwicz, M.A.; Dey, B.K.; Layer, R.; Reon, B.J.; Gaganj.R.; Dutta, A. MUNC, a long noncoding RNA that facilitates the function of MyoD in skeletal myogenesis. *Mol Cell Biol*. **2015**, *35*, 498–513; DOI: 10.1128/MCB.01079-14.
- [90] Gong, C.; Li, Z.; Ramanujan, K.; Clay, I.; Zhang, Y.; Lemire-Brachat, S.; Glass, D.J. A long non-coding RNA, LncMyoD, regulates skeletal muscle differentiation by blocking IMP2-mediated mRNA translation. *Dev Cell*. **2015**, *34*(2), 181–191, DOI: 10.1016/j.devcel.2015.05.009.

- [91] Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell* **2011**, *147*, 358–369, DOI: 10.1016/j.cell.2011.09.028.
- [92] Leung, A.; Trac, C.; Jin, W.; Lanting, L.; Akbany, A.; Sætrom, P.; Schones, D.E.; Natarajan, R. Novel Long Noncoding RNAs Are Regulated by Angiotensin II in Vascular Smooth Muscle Cells. *Circ Res.* **2013**, *113*(3), 266–78, DOI: 10.1161/CIRCRESAHA.112.300849.
- [93] Bell, R.D.; Long, X.; Lin, M.; Bergmann, J.H.; Nanda, V.; Cowan, S.L.; Zhou, Q.; Han, Y.; Spector, D.L.; Zheng, D.; Miano, J.M. Identification and Initial Functional Characterization of a Human Vascular Cell–Enriched Long Noncoding RNA. *Arterioscler Thromb Vasc Biol.* **2014**, *34*, 1249–1259, DOI: 10.1161/ATVBAHA.114.303240.
- [94] Sun, L.; Goff, L.A.; Trapnell, C.; Alexander, R.; Lo, K.A.; Hacisuleyman, E.; Sauvageau, M.; Tazon-Vega, B.; Kelley, D.R.; Hendrickson, D.G.; Yuan, B.; Kellis, M.; Lodish, H.F.; Rinn, J.L. Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci USA.* **2013**, *110*, 3387–3392; DOI: 10.1073/pnas.1222643110.
- [95] Xu, B.; Gerin, I.; Miao, H.; Vu-Phan, D.; Johnson, C.N.; Xu, R.; Chen, X.W.; Cawthorn, W.P.; MacDougald, O.A.; Koenig, R.J. Multiple Roles for the Non-Coding RNA SRA in Regulation of Adipogenesis and Insulin Sensitivity. *PLoS ONE.* **2010**, *5*, e14199; DOI: 10.1371/journal.pone.0014199.
- [96] Cooper, D.R.; Carter, G.; Li, P.; Patel, R.; Watson, J.E.; Patel, N.A. Long Non-Coding RNA NEAT1 Associates with SRp40 to Temporally Regulate PPAR γ 2 Splicing during Adipogenesis in 3T3-L1 Cells. *Genes* **2014**, *5*, 1050–1063, DOI: 10.3390/genes5041050.
- [97] Gernapudi, R.; Wolfson, B.; Zhang, Y.; Yao, Y.; Yang, P.; Asahara, H.; Zhou, Q. MicroRNA 140 promotes expression of long noncoding RNA NEAT1 in adipogenesis. *Mol Cell Biol.* **2016**, *36*, 30–38, DOI: 10.1128/MCB.00702-15.
- [98] Chen, J.; Liu, Y.; Lu, S.; Yin, L.; Zong, C.; Cui, S.; Qin, D.; Yang, Y.; Guan, Q.; Li, X.; Wang, X. The role and possible mechanism of lncRNA U90926 in modulating 3T3-L1 preadipocyte differentiation. *Int J Obes.* **2016**, *41*(2), 299–308, DOI: 10.1038/ijo.2016.189.
- [99] Pang, W.J.; Lin, L-G.; Xiong, Y.; Wei, N.; Wang, Y.; Shen, Q-W.; Yang, G-S. Knockdown of PU.1 AS lncRNA Inhibits Adipogenesis Through Enhancing PU.1 mRNA Translation. *J Cell Biochem.* **2013**, *114*, 2500–2512; DOI: 10.1002/jcb.24595.
- [100] Divoux, A.; Karastergiou, K.; Xie, H.; Weimin, G.; Perera, R.J.; Fried, S.K.; Smith, S.R. Identification of a novel lncRNA in gluteal adipose tissue and evidence for its positive effect on preadipocyte differentiation. *Obesity (Silver Spring)* **2014**, *22*, 1781–1785, DOI: 10.1002/oby.20793.
- [101] Zhao, X-Y.; Li, S.; Wang, G-X.; Yu, Q.; Lin, J.D. A Long Noncoding RNA Transcriptional Regulatory Circuit Drives Thermogenic Adipocyte Differentiation. *Mol Cell.* **2014**, *55*, 372–382; DOI: 10.1016/j.molcel.2014.06.004.
- [102] Lin, Mi.; Zhao, X-Y.; Li, S.; Yang, G.; Lin, J.D. Conserved function of the long noncoding RNA Blnc1 in brown adipocyte differentiation. *Mol Metab.* **2017**, *6*, 101–110; DOI: 10.1016/j.molmet.2016.10.010.
- [103] Alvarez-Dominguez, J.R.; Bai, Z.; Xu, D.; Yuan, B.; Lo, K.A.; Yoon, M.J. De Novo reconstruction of adipose tissue transcriptomes reveals long non-coding RNA regulators of

brown adipocyte development. *Cell Metab.* **2015**, 21, 764–776, DOI: 10.1016/j.cmet.2015.04.003.

[104] You, L.; Zhou, Y.; Cui, X.; Wang, X.; Sun, Y.; Gao, Y.; Wang, X.; Wen, J.; Xie, K.; Tang, R.; Ji, C.; Guo, X. GM13133 is a negative regulator in mouse white adipocytes differentiation and drives the characteristics of brown adipocytes. *J Cell Physiol.* **2017**, 233, 313–324; DOI: 10.1002/jcp.25878.

[105] Zhang, J.; Cui, X.; Shen, Y.; Pang, L.; Zhang, A.; Fu, Z.; Chen, J.; Guo, X.; Gan, W.; Ji, C. Distinct expression profiles of LncRNAs between brown adipose tissue and skeletal muscle. *Biochem Biophys Res Commun.* **2014**, 443, 1028–1034; DOI: 10.1016/j.bbrc.2013.12.092.

[106] Li, Z.; Jin, C.; Chen, S.; Zheng, Y.; Huang, Y.; Jia, L.; Ge, W.; Zhou, Y. Long non-coding RNA MEG3 inhibits adipogenesis and promotes osteogenesis of human adipose-derived mesenchymal stem cells via miR-140-5p. *Mol Cell Biochem.* **2017**, 433, 51, DOI: 10.1007/s11010-017-3015-z.

[107] Jin, C.; Jia, L.; Huang, Y.; Zheng, Y.; Du, N.; Liu, Y.; Zhou, Y. Inhibition of lncRNA MIR31HG Promotes Osteogenic Differentiation of Human Adipose-Derived Stem Cells. *Stem Cells* **2016**, 34, 2707–2720, DOI: 10.1002/stem.2439.

[108] Jin, C.; Zheng, Y.; Huang, Y.; Liu, Y.; Jia, L.; Zhou, Y. Long non-coding RNA MIAT knockdown promotes osteogenic differentiation of human adipose-derived stem cells. *Cell Biol Int.* **2016**, 9999, 1–9, DOI: 10.1002/cbin.10697.

[109] Huang, Y.; Zheng, Y.; Jia, L.; Li, W. Long Noncoding RNA H19 Promotes Osteoblast Differentiation via TGF- β 1/Smad3/HDAC Signaling Pathway by Deriving miR-675. *Stem Cells.* **2015**, 33(12), 3481–3492. DOI: 10.1002/stem.2225.

[110] Liao, J.; Yu, X.; Hu, X.; Fan, J.; Wang, J.; Zhang, Z.; Zhao, C.; Zeng, Z.; Shu, Y.; Zhang, R.; et al. LncRNA H19 mediates BMP9-induced osteogenic differentiation mesenchymal stem cells (MSCs) through Notch signaling. *Oncotarget.* **2017**, 8, 53581–53601; DOI: 10.18632/oncotarget.18655.

[111] Qiu, X.; Jia, B.; Sun, X.; Hu, W.; Chu, H.; Xu, S.; Zhao, J. The Critical Role of Long Noncoding RNA in Osteogenic Differentiation of Human Bone Marrow Mesenchymal Stem Cells. *Biomed Res Int.* **2017**, DOI: 10.1155/2017/5045827.

[112] Zhang, W.; Dong, R.; Diao, S.; Du, J.; Fan, Z.; Wang, F. Differential long noncoding RNA/mRNA expression profiling and functional network analysis during osteogenic differentiation of human bone marrow mesenchymal stem cells. *Stem Cell Res Ther.* **2017**, 8(1), 30. DOI: 10.1186/s13287-017-0485-6.

[113] Ishikawa, T.; Nishida, T.; Ono, M.; Takarada, T.; Ha, N.; Kurihara, S.; Furumatsu, T.; Murase, Y.; Takigawa, M.; Ohashi, T.; Kamioka, H.; Kubota, S. Physiological role of urothelial cancer-associated 1 long noncoding RNA in human skeletogenic cell differentiation. *J Cell Physiol.* **2017**, DOI: 10.1002/jcp.26285.

[114] Zhang, J.; Tao, Z.; Wang, Y. Long non-coding RNA DANCER regulates the proliferation and osteogenic differentiation of human bone-derived marrow mesenchymal stem cells via the p38 MAPK pathway. *Int J Mol Med.* **2018**, 41(1), 213–219. DOI: 10.3892/ijmm.2017.3215.

[115] Wang, S.; Wu, X.; Liu, Y.; Yuan, J.; Yang, F.; Huang, F.; Meng, Q.; Zhou, C.; Liu, F.; Ma, J.; Sun, S.; Zheng, J.; Wang, F. Long noncoding RNA H19 inhibits the proliferation of fetal liver cells and the Wnt signaling pathway. *FEBS Letters.* **2016**, 590, 559–570; DOI: 10.1002/1873-3468.12078.

- [116] Halley, P.; Kadakkuzha, B.M.; Faghihi, M.A.; Magistri, M.; Zeier, Z.; Khorkova, O.; Coito, C.; Hsiao, J.; Lawrence, M.; Wahlestedt, C. Regulation of apolipoprotein gene cluster by a long non coding RNA. *Cell Rep*. **2014**, *6*, 222-230, DOI: 10.1016/j.celrep.2013.12.015.
- [117] Sallam, T.; Jones, M.C.; Gilliland, T.; Zhang, L.; Wu X.; Eskin A.; Sandhu J.; Casero, D.; Vallim T.A.; Hong, C.; Katz, M.; Lee, R.; Whitelegge, J.; Tontonoz.; P. Feedback modulation of cholesterol metabolism by LeXis, a lipid-responsive non-coding RNA. *Nature*. **2016**, *534*, 124–128; DOI: 10.1038/nature17674.
- [118] Li, P.; Ruan, X.; Yang, L.; Kiesewetter, K.; Zhao, Y.; Luo, H.; Chen, Y.; Gucek, M.; Zhu, J.; Cao, H. A liver-enriched long non coding RNA, lncLSTR, regulates systemic lipid metabolism in mice. *Cell Metab*. **2015**, *21*, 455-467, DOI: 10.1016/j.cmet.2015.02.004.
- [119] Lan, X.; Yan, J.; Ren, J.; Zhong, B.; Li, J.; Li, Y.; Liu, L.; Yi, J.; Sun, Q.; Yang, X.; Sun, J.; Meng, L.; Zhu, W.; Holmdahl, R.; Li, D.; Lu S. A Novel Long Noncoding RNALnc-HC Binds hnRNPA2B1 to Regulate Expressions of Cyp7a1 and Abca1 in Hepatocytic Cholesterol Metabolism. *Hepatology* **2016**, *64*, 58-72, DOI: 10.1002/hep.28391.
- [120] Yang, L.; Li, P.; Yang, W.; Ruan, X.; Kiesewetter, K.; Zhu J.; Cao, H. Integrative Transcriptome Analyses of Metabolic Responses in Mice Define Pivotal LncRNA Metabolic Regulators. *Cell Metab*. **2016**, *24*, 627–639; DOI: 10.1016/j.cmet.2016.08.019.
- [121] Liu, G.; Zheng, X.; Xu, Y.; Lu, J.; Chen, J.; Huang, X. Long Non-coding RNAs Expression Profile in HepG2 Cells Reveals the Potential Role of Long Non-coding RNAs in the Cholesterol Metabolism. *Chin Med J*. **2015**, *128*, 91-97; DOI: 10.4103/0366-6999.147824.
- [122] Li, C.; Chang, L.; Chen, Z.; Liu, Z.; Wang, Y.; Ye, Q. The role of lncRNA MALAT1 in the regulation of hepatocyte proliferation during liver regeneration. *Int J Mol Med* **2017**, *39*, 347-356, DOI: 10.3892/ijmm.2017.2854.
- [123] Xiao, L.; Rao, J.N.; Cao, S.; Liu, L.; Chung, H.K.; Zhang, Y.; Zhang, J.; Liu, Y.; Gorospe, M.; Wang, J.Y. Long Noncoding RNA SPRY4-IT1 Regulates Intestinal Epithelial Barrier Function by Modulating the Expression Levels of Tight Junction Proteins. *Mol Biol Cell*. **2016**, *27*, 617-626; DOI: 10.1091/mbc.E15-10-0703.
- [124] Su, Z.; Zhi, X.; Zhang, Q.; Yang, I.; Xu H.; Xu, Z. LncRNA H19 functions as a competing endogenous RNA to regulate AQP3 expression by sponging miR-874 in the intestinal barrier. *FEBS Letters*. **2016**, *590*, 1354–1364; DOI: 10.1002/1873-3468.12171.
- [125] Zou, T.; Jaladanki, S.K.; Liu, L.; Xiao, L.; Chung, H.K.; Wang, J.Y.; Xu, Y.; Gorospe, M.; Wang, J.Y. H19 Long Noncoding RNA Regulates Intestinal Epithelial Barrier Function via MicroRNA 675 by Interacting with RNA-Binding Protein HuR. *Mol Cell Biol*. **2016**, *36*, 1332–1341; DOI: 10.1128/MCB.01030-15.
- [126] Tong, Q.; Gong, A.Y.; Zhang, X.; Lin, C.; Ma, S.; Chen, J.; Hu, G.; Chen, X-M. LincRNA-Cox2 modulates TNF- α -induced transcription of IL12b gene in intestinal epithelial cells through regulation of Mi-2/NuRD-mediated epigenetic histone modifications. *FASEB J*. **2016**, *30*, 1187-1197; DOI: 10.1096/fj.15-279166.
- [127] Li, J.; Cao, Y.; Xu, X.; Xiang, H.; Zhang, Z.; Chen, B.; Hao, Y.; Wei, Z.; Zhou, P.; Chen, D. Increased New lncRNA–mRNA Gene Pair Levels in Human Cumulus Cells Correlate With Oocyte Maturation and Embryo Development. *Reprod Sciences* **2015**, *22*, 1008-1014, DOI: 10.1177/1933719115570911.

- [128] Battaglia, R.; Vento, M.E.; Borzì, P.; Ragusa, M.; Barbagallo, D.; Arena, D.; Purrello, M.; Di Pietro, C. Non-coding RNAs in the Ovarian Follicle. *Front. Genet.* **2017**, *8*:57, DOI: 10.3389/fgene.2017.00057.
- [129] Gao, Q.; Ren, Q.; Chen, M.; Niu, Z.; Tao, H.; Jia, Y.; Zhang, J.; Li, W. Long non-coding RNAs regulate effects of β -crystallin B2 on mouse ovary development. *Mol Med Rep.* **2016**, *14*, 4223-4231, DOI: 10.3892/mmr.2016.5761.
- [130] Arun, G.; Akhade, V.S.; Donakonda, S.; Rao, M.R.S. mrhl RNA, a Long Noncoding RNA, Negatively Regulates Wnt Signaling through Its Protein Partner Ddx5/p68 in Mouse Spermatogonial Cells. *Mol Cell Biol.* **2012**, *32*, 3140–3152, DOI: 10.1128/MCB.00006-12.
- [131] Akhade, V.S.; Dighe, S.N.; Kataruka, S.; Rao, M.R.S. Mechanism of Wnt signaling induced down regulation of mrhl long non-coding RNA in mouse spermatogonial cells. *Nucleic Acids Research* **2016**, *44*, 387–401, DOI: 10.1093/nar/gkv1023.
- [132] Hu, K.; Zhang, J.; Liang, M. LncRNA AK015322 promotes proliferation of spermatogonial stem cell C18-4 by acting as a decoy for microRNA-19b-3p. *In Vitro Cell Dev Biol Anim.* **2016**, *53*(3), 277-284, DOI: 10.1007/s11626-016-0102-5.
- [133] Zhang, X.; Gao, F.; Fu, J.; Zhang, P.; Wang, Y.; Zeng, X. Systematic identification and characterization of long non-coding RNAs in mouse mature sperm. *PLoS ONE.* **2017**, *12*, e0173402; DOI: 10.1371/journal.pone.0173402.
- [134] Wichman, L.; Saigopal, S.; Breindel, C.; Valerio, D.M.; McCarrey, J.R.; Hodges, C.A.; Khalil, A.M. Dynamic expression of long non-coding RNAs reveals their potential roles in spermatogenesis and fertility. *Biol. Reprod.* **2017**, *97*(2), 313-323. DOI: 10.1093/biolre/iox084.
- [135] Hristov, M.; Erl, W.; Linder, S.; Weber, P.C. Apoptotic bodies from endothelial cells enhance the number and initiate the differentiation of human endothelial progenitor cells in vitro. *Blood* **2004**, *104*, 2761–2766, DOI: 10.1182/blood-2003-10-3614.
- [136] Yáñez-Mó, M.; Siljander, P.R.; Andreu, Z.; Zavec, A.B.; Borràs, F.E.; Buzas, E.I.; Buzas, K.; Casal, E.; Cappello, F.; Carvalho, J.; et al. Biological properties of extracellular vesicles and their physiological functions. *J Extracell Vesicles.* **2015**, *4*, 27066; DOI: 10.3402/jev.v4.27066.
- [137] Van der Grein, S.G.; Nolte-'t Hoen, E. N. “Small talk” in the innate immune system via RNA-containing extracellular vesicles. *Front. Immunol.* **2014**, *5*, 542; DOI: 10.3389/fimmu.2014.00542
- [138] Vader, P.; Breakefield, X.O.; Wood, M.J.A. Extracellular vesicles: emerging targets for cancer therapy. *Trends Mol Med.* **2014**, *20*, 385-393. DOI: 10.1016/j.molmed.2014.03.002.
- [139] Gezer, U.; Ozgur, E.; Cetinkaya, M.; Isin, M.; Dalay, N. Long non-coding RNAs with low expression levels in cells are enriched in secreted exosomes. *Cell Biol Int.* **2014**, *38*(9), 1076–9, DOI: 10.1002/cbin.10301.
- [140] Ahadi, A.; Brennan, S.; Kennedy, P.J.; Hutvagner, G.; Tran, N. Long non-coding RNAs harboring miRNA seed regions are enriched in prostate cancer exosomes. *Scientific Reports* **2016**, *6*, DOI: 10.1038/srep24922.
- [141] Hewson, C.; Capraro, D.; Burdach, J.; Whitaker, N.; Morris, K.V. Extracellular vesicle associated long non-coding RNAs functionally enhance cell viability. *Non-coding RNA Res.* **2016**, *1*, 3-11, DOI: 10.1016/j.ncrna.2016.06.001.

[142] Amorim, M.G.; Valieris, R.; Drummond, R.D.; Pizzi, M.P.; Freitas, V.M.; Sinigaglia-Coimbra, R.; Calin, G.A.; Pasqualini, R.; Arap, W.; Silva, I.T.; Dias-Neto, E.; Nunes, D.N. A total transcriptome profiling method for plasma-derived extracellular vesicles: applications for liquid biopsies. *Sci Rep.* **2017**; 7, DOI: 10.1038/s41598-017-14264-5.

7. DISCUSSÃO GERAL E CONSIDERAÇÕES FINAIS

No decorrer desta tese, utilizamos duas abordagens para analisar aspectos genéticos implicados no pênfigo foliáceo endêmico. No primeiro capítulo empregamos estudo de associação genética para identificar SNPs localizados em genes de lncRNAs que conferem risco diferencial ao pênfigo foliáceo endêmico. Já o segundo capítulo foi composto por um estudo de expressão gênica global que compreendeu tanto mRNAs quanto lncRNAs. Além de apresentar os resultados para essas duas abordagens em formato de artigo original, também apresentamos no capítulo III um artigo de revisão de literatura sobre a função de lncRNAs em células e tecidos humanos na condição de homeostase. Essa revisão foi útil para maior compreensão dos lncRNAs, suas funções e modos de ação, e visou disponibilizar conhecimento de forma abrangente e organizada a respeito desses ncRNAs, bastante investigados na biologia molecular atual. Nesta seção discutiremos os principais resultados obtidos nos capítulos I e II da tese.

Os resultados do estudo de associação apontaram um SNP fortemente associado ao PFE, o *rs7144332*, localizado no gene de lncRNA *AL110292.1*, e outros 5 SNPs sugestivos à associação. Esses SNPs estão localizados em genes de lncRNAs pouco conhecidos cuja função ainda não foi caracterizada. Existem diferentes maneiras pela qual um SNP pode afetar lncRNAs e suas funções; neste estudo algumas possibilidades foram consideradas e investigadas por ferramentas *in silico*. Primeiramente hipotetizamos que o SNP pode afetar a estrutura secundária de lncRNAs e, por meio de análises de predição identificamos que três dos SNPs associados podem alterar a conformação dos lncRNAs em que se localizam. Outra hipótese é que a variação na sequência primária do lncRNA crie ou desfaça sítios de interação com outras moléculas, como miRNAs ou proteínas. Nas análises *in silico* foi sugerido que três SNPs analisados possivelmente modificam sequências nas quais se ligam miRNAs. As análises de anotação funcional sugerem ainda que cinco desses SNPs encontram-se em elementos reguladores, possivelmente impactando o processo de transcrição; esse impacto poderia ocorrer na transcrição do próprio lncRNA ou de genes próximos. Diante desses resultados a importância da variação genética em genes de lncRNAs no PFE foi enfatizada. Dentre todos os resultados destacamos o SNP *rs6942557*, localizado nas duas isoformas do lncRNA *LINC0011776*, pois em todas as análises *in silico* obtivemos resultados interessantes

para o mesmo: esse SNP pode alterar a estrutura secundária do lncRNA em que está localizado, está presente em elemento regulador que provavelmente afeta a expressão gênica (escore regulome DB = 1b) e pode alterar sítios de ligação a miRNAs. Além disso, não há descrição de outros SNPs em desequilíbrio de ligação com o *rs6942557*, o que nos leva a propor que seja uma variante causal que pode ter um efeito direto na suscetibilidade ao PFE. Portanto, *rs6942557* é um forte candidato a validação funcional.

No segundo capítulo, descrevemos o perfil de expressão de mRNAs e lncRNAs de células TCD4⁺ de pacientes de PFE e indivíduos controles e comparados entre esses grupos. Identificamos 11 RNAs diferencialmente expressos entre pacientes e controles ($q < 0,1$) e 313 sugestivos à diferença de expressão ($0,1 < q < 0,2$).

Análises de enriquecimento de grupos gênicos, realizadas com os genes codificadores de proteína, evidenciaram que grande parte dos DEGs estão relacionados com resposta imune e resposta inflamatória, o que veio ao encontro do tipo celular analisado e do que se conhece sobre PFE, que é uma doença autoimune em que há presença de linfócitos B e T autorreativos com produção de autoanticorpos. Interessantemente, termos biológicos relacionados com a imunidade inata também foram vistos enriquecidos nas células T CD4⁺ de pacientes de pênfigo, como “via de sinalização dos receptores toll” e “células NK CD56⁺”.

Entre os DEGs relacionados a resposta imune destacamos *IL18RAP*, que codifica a proteína acessória do receptor de interleucina 18. *IL18RAP* foi encontrado superexpresso em pacientes no RNA-seq e validado por qPCR, além de já ter sido reportado superexpresso em pacientes de PFE com lesões generalizadas comparado com pacientes com lesões localizadas (MALHEIROS *et al.*, 2014). Na mesma região genômica de *IL18RAP*, sobrepostos a introns deste gene, estão localizados os genes de dois lncRNAs, *AC007278.1* e *AC007278.2*, que também estão superexpressos em pacientes de PFE. Embora não se conheçam suas funções, possivelmente esses lncRNAs regulam e/ou compartilham funções com genes próximos. Adicionalmente, o gene *IL18R1*, que também participa do complexo receptor de IL-18 tem expressão aumentada em PFE. Considerando todos esses fatores em conjunto nós destacamos *IL18RAP* como um dos nossos principais resultados e sugerimos que a via de sinalização da citocina IL-18 possui um papel importante na patogênese do PFE. É interessante notar que nem o receptor de IL-18, nem a via de sinalização de IL-18, tinham antes sido apontados como relevantes

em pênfigo, tanto vulgar como foliáceo, e, a partir desse achado, essa possibilidade surge e pode ser melhor investigada em estudos futuros.

No total de genes diferencialmente expressos e sugestivos a diferença de expressão, 32 corresponderam a lncRNAs. Entre os lncRNAs destacamos os AC007278.1 e AC007278.2, superexpressos em PFE e ainda inexplorados quanto a função. Porém, ambos são bastante interessantes no contexto geral de DEGs conforme discutido no parágrafo anterior. Ainda, destacamos o lncRNA *LEF1-AS1*, cuja subexpressão foi validada por qPCR, é antisenso ao gene *LEF1*, que por sua vez é um fator de transcrição importante em células T que contribui para a ativação de células T e já foi reportado crítico para as funções de células T reguladoras (XING *et al.*, 2019). Embora em nossos resultados nós não tenhamos observado expressão diferencial de *LEF1* é possível que *LEF1-AS1* regule a expressão de *LEF1*, uma vez que tem sido corroborado na literatura que muitos lncRNAs regulam a expressão de genes próximos aos quais são antisenso (LEE, 2012; VILLEGAS *et al.*, 2015).

Com o objetivo de explorar ainda mais nossos resultados, nós os comparamos àqueles do de uma análise de transcriptoma de PFE com microarranjo (MALHEIROS *et al.*, 2014). A expressão de oito genes é concordante nas duas abordagens. Entre eles, alguns já seriam candidatos de estudos no contexto da doença pelas suas funções relacionadas a resposta imune: *IL18RAP*, *ITGAM*, *GZMA*, *NKG7* e *CS7*. Outros são genes com funções celulares distintas: *APOBEC3G*, *ACAA2* e *FUT7*. A concordância entre estudos independentes torna esses genes fortemente corroborados como diferencialmente expressos em PFE e prováveis de contribuir de alguma forma com o início ou progressão da doença.

Por meio deste estudo identificamos vários mRNAs e lncRNAs diferencialmente expressos em células T CD4⁺ de pacientes com pênfigo foliáceo endêmico. Estudos posteriores poderão ajudar a decifrar o padrão de expressão dos mRNAs em nível de proteína e elucidar de que maneira estão implicadas em PFE. Para os lncRNAs diferencialmente expressos pode-se realizar estudos funcionais visando a identificação de interação com outras moléculas, como DNA, outros RNAs e proteínas e dos mecanismos de regulação desempenhados pelos lncRNAs.

Diante dos resultados apresentados nos capítulos dessa tese concluímos que nossos objetivos iniciais foram alcançados, uma vez que descrevemos SNPs de lncRNAs associados ao PFE e também detectamos vários genes, tanto

codificadores de proteína como de lncRNAs, diferencialmente expressos na condição de doença. A análise de transcriptoma contribuiu para melhor caracterização do PFE e de sua patogênese, sendo também responsável por sugerir genes e vias moleculares que possivelmente executam papel crucial no desencadeamento ou desenvolvimento da doença. Uma vez propostos neste trabalho, os genes com SNPs associados e os genes diferencialmente expressos podem ser estudados mais profundamente e surgir como possíveis alvos terapêuticos de PFE.

REFERÊNCIAS

- AOKI V., MILLIKAN R.C., RIVITTI E.A., HANS-FILHO G., EATON D.P., WARREN S.J., LI N., HILARIO-VARGAS J., HOFFMANN R.G., DIAZ L.A. Environmental risk factors in endemic pemphigus foliaceus (fogo selvagem). **J Investig Dermatol Symp Proc.** v.9, p.34–40, 2004.
- AUGUSTO D.G., LOBO-ALVES S.C., MELO M.F., *et al.* Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. **PloS One**, v.7, 7:e39991, 2012
- AUGUSTO D.G., O'CONNOR G.M., LOBO-ALVES S.C., *et al.* Pemphigus is associated with KIR3DL2 expression levels and provides evidence that KIR3DL2 may bind HLA-A3 and A11 in vivo. **Eur J Immunol** 2015, v.45, p.2052-2060, 2015.
- ATZMONY L., HODAK E., GDALEVICH M., ROSENBAUM O., MIMOUNI D. Treatment of pemphigus vulgaris and pemphigus foliaceus: a systematic review and meta-analysis. **Am J Clin Dermatol.** v.15, p.503–515, 2014.
- BASTUJI-GARIN S., SOUISSI R., BLUM L., TURKI H., NOUIRA R., JOMAA B., ZAHAF A., BEN OSMAN A., MOKHTAR I., FAZAA B. Comparative epidemiology of pemphigus in Tunisia and France: unusual incidence of pemphigus foliaceus in young Tunisian women. **The Journal of Investigative Dermatology**, v. 104, n. 2, p. 302-305, 1995.
- BRAUN-PRADO K, PETZL-ERLER ML. Programmed cell death 1 gene (PDCD1) polymorphism and pemphigus foliaceus (fogo selvagem) disease susceptibility. **Genetics Molecular Biology**, v.30, p.314, 2007.
- CABILI, M. N., TRAPNELL, C., GOFF, L., KOZIOL, M., TAZON-VEGA, B., REGEV, A., & RINN, J. L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. **Genes & development**, v.25, p.1915–1927, 2011.
- CABILI, M. N., DUNAGIN, M. C., MCCLANAHAN, P. D., BIAESCH, A., PADOVAN-MERHAR, O., REGEV, A., ... RAJ, A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. **Genome biology**, v.16, p. 20, 2015.
- CAMARGO C.M., AUGUSTO D.G., PETZL-ERLER M.L. Differential gene expression levels might explain association of LAIR2 polymorphisms with pemphigus. **Hum Genet.** v.135, p.233-244, 2016.
- CAMPBELL I., REIS V., AOKI V., CUNHA P., HANS FILHO G., ALVES G., ROCHA R., FRIEDMAN H. Pênfigo foliáceo endêmico/fogo selvagem. **Anais Brasileiros de Dermatologia**, v.76, p.13-33, 2001.
- CELERE B. S., VERNAL S., LA SERRA L., FRANCO BROCHADO M. J., MOSCHINI L. E., ROSELINO A. M., SEGURA-MUÑOZ, S. I. Spatial Distribution of Pemphigus Occurrence over Five Decades in Southeastern Brazil. **The American journal of tropical medicine and hygiene**, v.97, p.1737–1745, 2017.
- CHEN Y. G., SATPATHY A. T., CHANG, H. Y. (2017). Gene regulation in the immune system by long noncoding RNAs. **Nature Immunology**, v.18,p. 962–972, 2017.
- CIPOLLA G.A., PARK J.K., DE OLIVEIRA L.A., *et al.* A 3'UTR polymorphism marks differential KLRG1 mRNA levels through disruption of a miR-584-5p binding site and

associates with pemphigus foliaceus susceptibility. **Biochem Biophys Acta**, v.1859, p.1306-1313, 2016.

CIPOLLA, G. A., DE OLIVEIRA, J. C., SALVIANO-SILVA, A., LOBO-ALVES, S. C., LEMOS, D. S., OLIVEIRA, L. C., ... GRADIA, D. F. (2018). Long Non-Coding RNAs in Multifactorial Diseases: Another Layer of Complexity. **Non-coding RNA**, v. 4, 1p.3, 2018.

COLLIER S.P., HENDERSON M.A., TOSSBERG J.T., AUNE T.M. Regulation of the Th1 genomic locus from *Ilfn*g through *Tmevpg1* by T-bet. **J. Immunol.** v 193, p.3959–3965, 2014.

DALLA-COSTA R., PINCERATI M.R., BELTRAME M.H., MALHEIROS D., PETZL-ERLER M.L. Polymorphisms in the 2q33 and 3q21 chromosome regions including T cell coreceptor and ligand genes may influence susceptibility to pemphigus foliaceus. **Human Immunology**, v.71, p.809, 2010.

DE OLIVEIRA J.C., OLIVEIRA L.C., MATHIAS C, PEDROSO GA, LEMOS DS, SALVIANO-SILVA A, JUCOSKI TS, LOBO-ALVES SC, ZAMBALDE EP, CIPOLLA GA, GRADIA DF. Long non-coding RNAs in cancer: Another layer of complexity. **J Gene Med.** V.21, 2019.

DERRIEN T, JOHNSON R, BUSSOTTI G, *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. **Genome Res.**, v. 22, p.1775-1789, 2012.

DIAZ L.A., SAMPAIO S.A.P., RIVITTI E.A., MARTINS C.R., CUNHA P.R., LOMBARDI C., ALMEIDA F.A., CASTRO R.M., MACCA M.L., LAVRADO C., HANS FILHO G., BORGES P., CHAUL A., MINELLI L., EMPINOTTI J.C., FRIEDMAN H., CAMPBELL I., LABBIB R.S., ANHALT G.J. Endemic pemphigus foliaceus (fogo selvagem): II. Current and historical epidemiologic studies. **The Journal of Investigative Dermatology**, v.92, p.4-12, 1989.

DMOCHOWSKI M., HASHIMOTO T., GARROD D.R., NISHIKAWA T. Desmocollins I and II are recognized by certain sera from patients with various types of Pemphigus particularly Brazilian Pemphigus foliaceus. **The Journal of Investigative Dermatology**, v.100, p.380–384, 1993.

DOLCINO M., TINAZZI E., PUC CETTI A., LUNARDI C. In Systemic Sclerosis, a Unique Long Non Coding RNA Regulates Genes and Pathways Involved in the Three Main Features of the Disease (Vasculopathy, Fibrosis and Autoimmunity) and in Carcinogenesis. **Journal of Clinical Medicine**, v.8, p.320, 2019.

EATON D.P., HANS-FILHO G., SANTOS V.D., AOKI V., FRIEDMAN H., RIVITTI E.A., SAMPAIO S.A., GOTTLIEB M.S., GIUDICE G.J., LOPEZ A., CUPP E.W. Comparison of black fly species (Diptera: Simuliidae) on an Amerindian reservation with a high prevalence of fogo selvagem to neighboring disease-free sites in the State of Mato Grosso do Sul, Brazil. **The Cooperative Group on Fogo Selvagem Research. Journal of Medical Entomology**, v.35, p.120-131, 1998.

FENOGLIO C., OLDONI E., SERPENTE M., DE RIZ M.A., ARCARO M., D'ANCA M., PIETROBONI A.M., CALVI A., LECCHI E., GORIS A., MALLANTS K., DUBOIS B., COMI C., CANTELLO R., SCARPINI E., GALIMBERTI D. LncRNAs expression profile in peripheral blood mononuclear cells from multiple sclerosis patients. **J Neuroimmunol.** v. 324, p.129-135, 2018.

GRAU-PÉREZ M., GÓMEZ-DUASO A-J., SÁNCHEZ-GARCÍA F., DE LA ROSA-DEL REY P., GONZÁLEZ-MARTÍN J.M., CARRETERO-HERNÁNDEZ G. Epidemiological and genetic data suggest the presence of an endemic form of pemphigus foliaceus in Gran Canaria, Spain. **British Journal of Dermatology**. V.180, p.945-946, 2018.

GUPTA, R., AHN, R., LAI, K., MULLINS, E., DEBBANEH, M., DIMON, M., ARRON, S., ... LIAO, W. Landscape of Long Noncoding RNAs in Psoriatic and Healthy Skin. **The Journal of investigative dermatology**, v. 136, p. 603-609, 2015.

GUTSCHNER T, HÄMMERLE M, EISSMANN M, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. **Cancer Res**. 2012; v.73, p.1180–1189, 2012.

GUTSCHNER T., HÄMMERLE M., DIEDERICH S. *MALAT1* -- a paradigm for long noncoding RNA function in cancer. **Journal of Molecular Medicine (Berlin)**. v.91 p.791–801, 2013.

HANS-FILHO G., DOS SANTOS V., KATAYAMA J. H., AOKI V., RIVITTI E. A., SAMPAIO S. A., FRIEDMAN H., MORAES J. R., MORAES M. E., EATON D. P., LOPEZ A. L., HOFFMAN R. G., FAIRLEY J. A., GIUDICE G. J., e DIAZ L. A. Na active focus of high prevalence of fogo selvagem on an Amerindian reservation in Brazil. Cooperative Group on Fogo Selvagem Research. **J Invest Dermatol**, Baltimore, v. 107, p. 68-75, 1996.

HANS-FILHO G., AOKI V., ROVITTI E., EATON D.P., LIN M.S., DIAZ L. Endemic pemphigus foliaceus (fogo selvagem) -1998. **Clinics in Dermatology**, v.17, p.225-235, 1999.

HARROW J., FRANKISH A., GONZALEZ J.M., et al. GENCODE: the reference human genome annotation for The ENCODE Project. **Genome Res**. v.22, p.1760-1774, 2012.

HEZRONI H., KOPPSTEIN D., SCHWARTZ M.G., AVRUTIN A., BARTEL D.P., ULITSKY I. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. **Cell Rep**. v. 11, p.1110–1122. 2015.

HRDLICKOVA B., DE ALMEIDA R.C., BOREK Z., WITHOFF S. Genetic variation in the non-coding genome: Involvement of micro-RNAs and long non-coding RNAs in disease. **Biochim. Biophys. Acta**. v. 1842, p.1910–1922. 2014a.

HRDLICKOVA B., KUMAR V., KANDURI K., ZHERNAKOVA D. V., TRIPATHI S., KARJALAINEN J., ... WITHOFF S. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. **Genome medicine**, v.6, p.88.2014b.

HUANG, W., THOMAS, B., FLYNN, R. A., GAVZY, S. J., WU, L., KIM, S. V., ... LITTMAN, D. R. DDX5 and its associated lncRNA Rmrp modulate TH17 cell effector functions. **Nature**, v. 528, p.517–522, 2015.

HUARTE M., GUTTMAN, M., FELDSEY, D., GARBER, M., KOZIOL, M. J., KENZELMANN-BROZ, D.,...RINN J. L. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. **Cell**, v.142, p. 409–419, 2010.

ISHII N., OZAKI K., SATO H., MIZUNO H., SAITO S, TAKAHASHI A., MIYAMOTO Y.,

IKEGAWA S., KAMATANI N., HORI M., et al. Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. **J Hum Genet.** V.51, p.1087–1099, 2006.

JONES J.C., ARNN J., STACHELIN L.A., GOLDMAN R.D. Human antibodies against desmosomes: possible causative factors in pemphigus. **Proceedings of the National Academy of Sciences of the United States of America**, v.81, p.2781-2785, 1984.

KAPRANOV P., CHENG J., DIKE S., NIX D.A., DUTTAGUPTA R., WILLINGHAM A.T., STADLER P.F., HERTEL J., HACKERMULLER J., HOFACKER I.L., BELL I., CHEUNG E., DRENKOW J., DUMAIS E., PATEL S., HELT G., GANESH M., GHOSH S., PICCOLBONI A., SEMENTCHENKO V., TAMMANA H., GINGERAS T.R. RNA maps reveal new RNA classes and a possible function for pervasive transcription. **Science**, v.316, p.1484–1488, 2007.

KIM S.C., CHUNG Y.L., KIM J., CHO N.J., AMAGAI M. Pemphigus vulgaris with autoantibodies to desmoplakin. **The British Journal of Dermatology**, v.145, p.838–840, 2001.

KOGO R, SHIMAMURA T, MIMORI K, KAWAHARA K, IMOTO S, SUDO T, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. **Cancer Res.** v.71, p.6320–6326, 2011.

KOPP, F., MENDELL, J. T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. **Cell**, v. 172, p.393–407, 2018.

KORMAN N.J., EYRE R.W., KLAUS-KOVTUN V., STANLEY J.R. Demonstration of an adhering-junction molecule (plakoglobin) in the autoantigens of Pemphigus foliaceus and Pemphigus vulgaris. **The New England Journal of Medicine**, v.321, p. 631–635, 1989.

KOTZIN, J.J.; SPENCER, S.P.; MCCRIGHT, S.J.; KUMAR, D.B.; COLLET, M.A.; MOWEL, W.K.; ELLIOTT, E.N.; UYAR, A.; MAKIYA, M.A.; DUNAGIN, M.C.; HARMAN, C.C.D.; VIRTUE, A.T.; ZHU, S.; BAILIS, W.; STEIN, J.; HUGHES, C.; RAJ, A.; WHERRY, E.J.; GOFF, L.A.; KLION, A.D.; RINN, J.L.; WILLIAMS, A.; FLAVELL, R.A.; HENAO-MEJIA, J. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. **Nature**, v.537, p.239–243, 2016.

KRAWCZYK M., EMERSON, B.M. p50-associated COX-2 extragenic RNA (PACER) activates COX-2 gene expression by occluding repressive NF-κB complexes. **eLife** v.3, e01776, 2014.

LEE J.T. Epigenetic regulation by long noncodingRNAs. **Science**, v. 338, p.1435–1439, 2012.

LI J., WU G.C., ZHANG T.P., YANG X.K., CHEN S.S., LI L.J., XU Z.S., LV T.T., LENG R.X., PAN H.F., YE D.Q., Association of long noncoding RNAs expression levels and their gene polymorphisms with systemic lupus erythematosus. **Sci. Rep.** v.7, 2017.

LI N., AOKI V., HANS-FILHO G., RIVITTI E A., DIAZ L. A. The role of intramolecular epitope spreading in the pathogenesis of endemic pemphigus foliaceus (Fogo Selvagem). **The Journal of Experimental Medicine**, v. 197, p. 1501-1510, 2003.

LIN M.S., FU C.L., AOKI V., HANS-FILHO G., RIVITTI E.A., MORAES J.R., MORAES M.E., LAZARO A.M., GIUDICE G.J., STASTNY P., DIAZ L.A. Desmoglein-1-specific T

lymphocytes from patients with endemic pemphigus foliaceus (fogo selvagem). **The Journal of clinical investigation**, v.105, p.207-213, 2000.

LIU B., SUN L., LIU Q., GONG C., YAO Y., LV X., LIN L., YAO H., SU F., LI D., ZENG M., SONG E. A cytoplasmic NF- κ B interacting long noncoding RNA blocks I κ B phosphorylation and suppresses breast cancer metastasis. **Cancer Cell**, v. 27, p.370–381, 2015.

LOMBARDI C., BORGES P.C., CHAUL A., SAMPAIO S.A.P., RIVITTI E.A., FREIDMAN H., et al. Environmental risk factors in endemic pemphigus foliaceus (Fogo selvagem). "The Cooperative Group on Fogo Selvagem Research". **The Journal of Investigative Dermatology**, v. 98, p.847-50, 1992.

LU L., ZHU G., ZHANG C., DENG Q., KATSAROS D., MAYNE S.T., et al. Association of large noncoding RNA HOTAIR expression and its downstream intergenic CpG island methylation with survival in breast cancer. **Breast Cancer Res Treat.** v.136, p.875–883, 2012.

MALHEIROS D, PETZL-ERLER ML. Individual and epistatic effects of genetic polymorphisms of B-cell co-stimulatory molecules on susceptibility to pemphigus foliaceus. **Genes Immunology**; v.10, p547, 2009.

MALHEIROS D., PANEPUCCI R.A., ROSELINO A.M., ARAÚJO A.G., ZAGO M.A., PETZL-ERLER M.L. Genome-wide gene expression profiling reveals unsuspected molecular alterations in pemphigus foliaceus. **Immunology**, v.143, p.381–395, 2014.

MARTEL M., JOLY P. Pemphigus: autoimmune diseases of keratinocyte's adhesion molecules. **Clinics in Dermatology**, v.19, p.662-674, 2001.

MASMOUDI H., ABIDA O., MASMOUDI A., TURKI H. Update on immunogenetics of Tunisian endemic pemphigus foliaceus. **J Leukoc Biol.** p. 105:257-265, 2019.

MAURANO, M. T., HUMBERT, R., RYNES, E., THURMAN, R. E., HAUGEN, E., WANG, H., ... STAMATOYANNOPOULOS, J. A. (2012). Systematic localization of common disease-associated variation in regulatory DNA. **Science**, v.337, p.1190–1195. 2012.

MELÉ M., MATTIOLI K., MALLARD W., SCHECHNER D.M., GERHARDINGER C., RINN J.L. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. **Genome Res.** V. 27, p.27–37, 2017.

MORAN V.A., PERERA R.J., KHALIL, A.M. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. **Nucleic Acids Res.**, v.40, p.6391–6400; 2012.

MULLER N., DORING F., KLAPPER M., et al. Interleukin-6 and Tumour Necrosis Factor- α differentially regulate lincRNA transcripts in cells of the innate immune system in vivo in human subjects with rheumatoid arthritis. **Cytokine**, vol. 68,p. 65–68, 2014.

NGUYEN V.T., LEE T.X., NDOYE A., SHULTZ L.D., PITTELKOW M.R., DAHL M.V., LYNCH P.J., GRANDO S.A. The pathophysiological significance of non-desmoglein targets of Pemphigus autoimmunity. Pemphigus vulgaris and foliaceus patients develop antibodies against keratinocyte cholinergic receptors. **Archives of Dermatology**, v.134, p.971–980. 1998.

OLIVEIRA L. Coestimuladores de linfócitos B e pênfigo foliáceo: polimorfismos associados e expressão gênica diferencial. Curitiba, 2015. 149p. **Tese** (Doutorado em Genética) Departamento de Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná. Orientadora: Maria Luiza Petzl-Erler.

OLIVEIRA L.A., MARQUART-FILHO A., TREVILATO G., TIMOTEO R.P., MUKAI M., ROSELINO A.M., et al. Anti-desmoglein 1 and 3 autoantibody levels in endemic pemphigus foliaceus and pemphigus vulgaris from Brazil. **Clin Lab**. V.62, p.1209–1216, 2016.

PAVONI D.P., ROXO V.M.M.S., MARQUART FILHO A., PETZL-ERLER M.L. Dissecting the associations of endemic pemphigus foliaceus (Fogo Selvagem) with HLA-DRB1 alleles and genotypes. **Genes and Immunity**, 4: 110-6, 2003.

PEREIRA NF, HANSEN JA, LIN MT, *et al*. Cytokine gene polymorphisms in endemic pemphigus foliaceus: a possible role for IL6 variants. **Cytokine**, v. 28, p.233-241, 2004.

PETZL-ERLER M.L., SANTAMARIA J. Are HLA class II genes controlling susceptibility and resistance to Brazilian pemphigus foliaceus (fogo selvagem)? **Tissue Antigens**, v.33, p. 408-414, 1989.

PIOVEZAN B.Z., PETZL-ERLER M.L. Both Qualitative and quantitative genetic variation of MHC Class II molecules may influence susceptibility to autoimmune diseases: the case of endemic pemphigus foliaceus. **Human Immunology**, v.74, p.1134-1140, 2013.

QIAN Y., JEONG J.S., MALDONADO M, VALENZUELA J.G., *et al*. Cutting Edge: Brazilian Pemphigus Foliaceus Anti-Desmoglein 1 autoantibodies cross-react with sandy fly salivary LJM11 antigen. **The Journal of Immunology**; v.189, p.1535-1539, 2013.

ROBLEDO M.A.; STEMA PRADA, C.; JARAMILLO, D.; LEON, W. South American pemphigus foliaceus: Study of an epidemic in El Bagre and Nechi, Colombia 1982 to 1986. **British Journal of Dermatology**, v. 118, n. 6, p. 737-744, 1988.

RANZANI, V.; ROSSETTI, G.; PANZERI, I.; ARRIGONI, A.; BONNAL, R.J.P.; CURTI, S.; GRUARIN, P.; PROVASI, E.; SUGLIANO, E.; MARCONI, M.; DE FRANCESCO, R.; GEGINAT, J.; BODEGA, B.; ABRIGNANI, S.; PAGANI, M. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. **Nat Immunol**. V.16, p.318-325, 2015.

RICAÑO-PONCE I., WIJMENGA C. Mapping of immune-mediated disease genes. **Annual review of genomics and human genetics**, v.14, p.325–353, 2013.

RINN J.L., CHANG H.Y. Genome regulation by long noncoding RNAs. **Annu Rev Biochem**, v.81, p.145–166, 2012.

SALVIANO-SILVA A., LOBO-ALVES S.C., ALMEIDA R.C., MALHEIROS D., PETZL-ERLER M.L. Besides Pathology: Long Non-Coding RNA in Cell and Tissue Homeostasis. **Noncoding RNA**, v.4 p.3.

SAMUDYATA, CASTELO-BRANCO G., BONETTI A. Birth, coming of age and death: The intriguing life of long noncoding RNAs. **Semin Cell Dev Biol**. v. 79, p.143-152, 2018.

SCHLACKOW M., NOJIMA T., GOMES T., DHIR A., CARMO-FONSECA M., PROUDFOOT N.J. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. **Mol Cell**. V.65, p.25–38, 2017.

SHAKER O.G., MAHMOUD R.H., ABDELALHEEM O.O., AHMED T.I., FOUAD N.A., HUSSEIN H.A., NASSR M.H., ZAKI O.M., ABDELGHAFAR N.K., HEFZY E.M. Expression Profile of Long Noncoding RNAs, Inc-Cox2, and HOTAIR in Rheumatoid Arthritis Patients. **J Interferon Cytokine Res.** v.9, p.174-180. 2019.

SONG J., KIM D., HAN J., KIM Y., LEE M., JIN E.J. PBMC and exosome-derived HOTAIR is a critical regulator and potent marker for rheumatoid arthritis. **Clin Exp Med.** v.15, p.121–126, 2015.

SONKOLY E., BATA-CSORGO Z., PIVARCSI A., POLYANKA H., KENDERESSY-SZABO A., MOLNAR G., SZENTPALI K., BARI L., MEGYERI K., MANDI Y., DOBOZY A., KEMENY L., SZELL M. Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. **J Biol Chem.** V.280, p.24159–24167, 2005.

SOUSA J. X., JR, DIAZ L. A., EATON D. P., HANS-FILHO G., FREITAS E. L., DELGADO, L., ... Cooperative Group on Fogo Selvagem Research . Profile of Trypanosoma cruzi reactivity in a population at high risk for endemic pemphigus foliaceus (Fogo selvagem). **The American journal of tropical medicine and hygiene**, v. 87, p.675–680. 2012.

SPURLOCK CF 3RD, TOSSBERG JT, GUO Y, COLLIER SP, CROOKE PS 3RD, AUNE TM. Expression and functions of long noncoding RNAs during human T helper cell differentiation. **Nat Commun.** V. 6, p.6932, 2015.

STEIN N, BERHANI O, SCHMIEDEL D, et al. IFNG-AS1 Enhances Interferon Gamma Production in Human Natural Killer Cells. **iScience.** 2019;11:466–473, 2019
doi:10.1016/j.isci.2018.12.034

TRON F., GILBERT D., MOUQUET H., JOLY P., DROUOT L., MAKNI S., MASMOUDI H., CHARRON D., ZITOUNI M., LOISEAU P., AYED M.B. Genetic factors in pemphigus. **Journal of Autoimmunity**, v.24, p.319-328, 2005.

TILGNER H., KNOWLES D.G., JOHNSON R., DAVIS C.A., CHAKRABORTTY S., DJEBALI S., *et al.* Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. **Genome Res.** v.22, p.1616–1625, 2012.

TIMÓTEO R.P., SILVA M.V., DA SILVA D.A.A., et al. Cytokine and Chemokines Alterations in the Endemic Form of Pemphigus Foliaceus (Fogo Selvagem). **Front Immunol.** v. 8, p.978, 2017.

UDEY M.C., STANLEY J.R. Pemphigus – diseases of antidesmosomal autoimmunity. **JAMA**, v.282, p.572–576, 1999.

ULITSKY I., BARTEL D. P. lincRNAs: genomics, evolution, and mechanisms. **Cell**, v. 154, p. 26–46, 2013.

VERNAL S., PEPINELLI M., CASANOVA C., GOULART T.M., KIM O., DE PAULA N.A., PINTO M.C., SÁ-NUNES A., ROSELINO A.M. Insights into the epidemiological link between biting flies and pemphigus foliaceus in southeastern Brazil. **Acta Tropica**, v.176, p.455-462, 2017.

VILLEGAS V.E., ZAPHIROPOULOS P.G. Neighboring gene regulation by antisense long non-coding RNAs. **International Journal of Molecular Science**. v. 16, p.3251–3266, 2015.

XING S., GAI K., LI X., SHAO P., ZENG Z., ZHAO X., ZHAO X., CHEN X., PARADEE W.J., MEYERHOLZ D.K., PENG W., XUE H.H. Tcf1 and Lef1 are required for the immunosuppressive function of regulatory T cells. **The Journal of Experimental Medicine** v.216, p847-866,2019.

WARREN S.J.P., LIN M.S., GIUDICE G.J., HOFFMANN R.G., HANS-FILHO G., AOKI V., RIVITTI E.A., SANTOS V., DIAZ L.A. The prevalence of antibodies against desmoglein 1 in endemic pemphigus foliaceus in Brazil. **The New England Journal of Medicine**, v.343, p.23-30, 2000.

WHEELER G.N., PARKER A.E., THOMAS C.L., ATALIOTIS P., POYNTER D., ARNEMANN J., RUTMAN A.F., PIDSLEY S.C., WATT F.M., REES D.A., BUXTON R.S., MAGEE A.I. Desmosomal glycoprotein DG1, a component of intercellular desmosome junctions, is related to the cadherin family of cell adhesion molecules. **Proceedings of the National Academy of Sciences of the United States of America**, v.88, 0.4796-4800, 1991.

ZHANG F., WU L., QIAN J., QU B, XIA S., LA T., WU Y., MA J., ZENG J., GUO Q., CUI Y., YANG W., HUANG J., ZHU W., YAO Y., SHEN N., TANG Y. Identification of the long noncoding RNA NEAT1 as a novel inflammatory regulator acting through MAPK pathway in human lupus. **J. Autoimmun.** v.75, p. 96–104, 2016.

ZEOTI D. M., FIGUEIREDO J. F., CHIOSSI M. P., e ROSELINO A. M. Serum cytokines in patients with Brazilian pemphigus foliaceus (fogo selvagem). **Braz J Med Biol Res**, v.33, n. 9, p. 1065-1068, 2000.

APÊNDICE

Apêndice A – Material suplementar Capítulo II

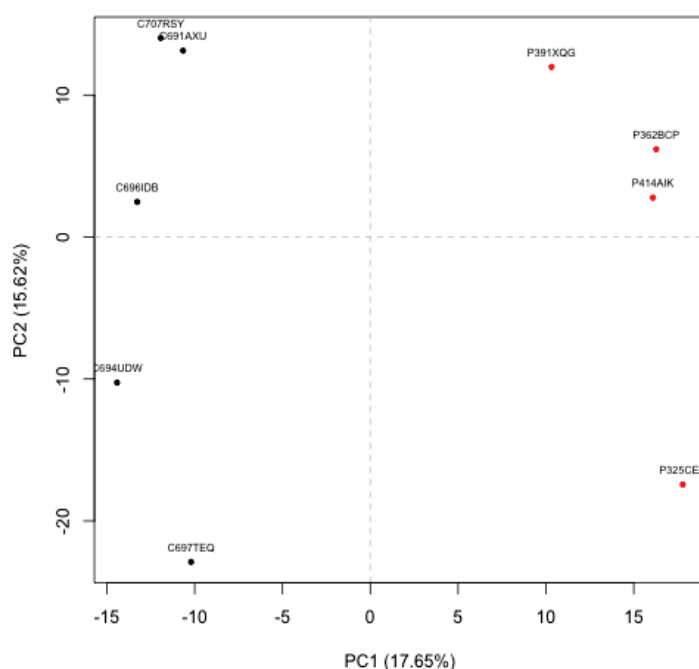


Figure S1. Principal component analysis (PCA) plot showing a map of the distances between samples according to gene expression (FPKM) levels.

Table S1. qPCR primer and conditions information

primer	sequence (5' - 3')	location	amplicon size	annealing temperature
LEF1-AS1_Foward	GAAGGCCAAGGGTTCTCTTC	exon 1		
LEF1-AS1_Reverse	TTTCAGGGTTTCTCCATTGC	exon2	116bp	64.5°C
GNLY_Foward	TAGAGTTACCCAGGGCCTCG	exon 4		
GNLY_Reverse	GAGGGAGACTGGAGAGTGGGA	exon 5	210	65.5°C
IL18RAP_Foward	CAGATATTCTGGATCCTGTCGAG	exon 7		
IL18RAP_Reverse	TGCTTTGCAGCTAATAGTTAAAGG	exon 8	74bp	62°C
LINC00921_Foward	GAGACTACTGTCTTCAGAGGGC	exon 2-3		
LINC00921_Reverse	GGTGTCCACTGTGGTTCCAG	exon 3	71bp	66.5°C
S1PR5_Foward	GACCAGGGCGCAGACCTTG	exon 1-2		
S1PR5_Reverse	GCCGGTGTAGTTGTAATGCAGG	exon 2	101bp	67°C
ACTB_Foward	ACATCCGCAAAGACCTGTACG	exon 7		
ACTB_Reverse	TTGCTGATCCA ATCTGCTGG	exon 8	213 bp	Normalization gene

Table S2. Differentially expressed and suggestive to differential expression in CD4⁺ T cells from EPF patients

Ensembl ID	p-value	q-value	Symbol	Entrez ID	fold change	NC
ENSG00000089335	1.63652E-07	0.00337	ZNF302	55900	-1.817	0
ENSG00000198723	1.94501E-06	0.01715	TEX45	374877	-1.696	0
ENSG00000180739	2.49888E-06	0.01715	S1PR5	53637	3.489	0
ENSG00000163629	1.15026E-05	0.05920	PTPN13	5783	1.732	0
ENSG00000027869	1.59124E-05	0.06100	SH2D2A	9047	0.763	0
ENSG00000234389	1.77792E-05	0.06100	AC007278.1		2.254	1
ENSG00000197134	2.97061E-05	0.07603	ZNF257	113835	-1.471	0
ENSG00000145220	3.19945E-05	0.07603	LYAR	55646	0.839	0
ENSG00000158470	3.61825E-05	0.07603	B4GALT5	9334	1.259	0
ENSG00000178184	3.69347E-05	0.07603	PAR6D6G	84552	1.883	0
ENSG00000115523	4.89448E-05	0.09159	GNLY	10578	3.406	0
ENSG00000099769	6.79013E-05	0.10263	IGFALS	3483	-1.532	0
ENSG00000261167	7.1076E-05	0.10263	AC107027.3		-0.927	1

ENSG00000203485	7.47882E-05	0.10263	INF2	64423	-0.659	0
ENSG00000117500	7.25706E-05	0.10263	TMED5	50999	0.525	0
ENSG00000223547	0.0001190	0.11475	ZNF844	284391	-0.958	0
ENSG00000237550	0.0001543	0.11475	RPL9P9		-0.955	0
ENSG00000184678	0.0001062	0.11475	HIST2H2BE	8349	-0.687	0
ENSG00000177732	0.0001141	0.11475	SOX12	6666	-0.672	0
ENSG00000268861	0.0001433	0.11475	AC008878.3	23370	-0.484	0
ENSG00000132823	0.0001202	0.11475	OSER1	51526	-0.439	0
ENSG00000169032	0.0001672	0.11475	MAP2K1	5604	0.553	0
ENSG00000148175	0.0001627	0.11475	STOM	2040	0.588	0
ENSG00000020633	0.0000977	0.11475	RUNX3	864	0.612	0
ENSG00000223639	0.0001421	0.11475	CLIC1	1192	0.751	0
ENSG00000042753	0.0001590	0.11475	AP2S1	1175	0.805	0
ENSG00000026652	0.0001597	0.11475	AGPAT4	56895	0.901	0
ENSG00000151116	0.0001497	0.11475	UEVLD	55293	1.233	0
ENSG00000112486	0.0000926	0.11475	CCR6	1235	1.284	0
ENSG00000060558	0.0001384	0.11475	GNA15	2769	1.512	0
ENSG00000203814	0.0001795	0.11549	HIST2H2BF	440689	-0.685	0
ENSG00000077984	0.0001795	0.11549	CST7	8530	1.435	0
ENSG00000172493	0.0001889	0.11601	AFF1	4299	0.521	0
ENSG00000057657	0.0001916	0.11601	PRDM1	639	1.078	0
ENSG00000062524	0.0002045	0.12025	LTK	4058	1.243	0
ENSG00000228463	0.0002167	0.12287	AP006222.1		-1.210	0
ENSG00000151883	0.0002276	0.12287	PARP8	79668	0.422	0
ENSG00000104490	0.0002308	0.12287	NCALD	83988	0.983	0
ENSG00000236525	0.0002334	0.12287	AC007278.2		1.937	1
ENSG00000137441	0.0002388	0.12287	FGFBP2	83888	3.306	0
ENSG00000256235	0.0002569	0.12898	SMIM3	85027	1.133	0
ENSG00000215302	0.0002783	0.13640	AC127502.1		-0.851	0
ENSG00000183735	0.0003110	0.14890	TBK1	29110	0.822	0
ENSG00000250506	0.0004142	0.15270	CDK3	1018	-1.457	0
ENSG00000125122	0.0003416	0.15270	LRRC29	26231	-1.165	0
ENSG00000213420	0.0004064	0.15270	GPC2	221914	-0.923	0
ENSG00000177189	0.0004050	0.15270	RPS6KA3	6197	0.425	0
ENSG00000163162	0.0004303	0.15270	RNF149	284996	0.472	0
ENSG00000114166	0.0003861	0.15270	KAT2B	8850	0.629	0
ENSG00000011566	0.0004243	0.15270	MAP4K3	8491	0.766	0
ENSG00000112029	0.0003503	0.15270	FBXO5	26271	0.774	0
ENSG00000164111	0.0003583	0.15270	ANXA5	308	0.788	0
ENSG00000015475	0.0003599	0.15270	BID	637	0.853	0
ENSG00000128656	0.0003927	0.15270	CHN1	1123	1.175	0
ENSG00000007402	0.0003710	0.15270	CACNA2D2	9254	1.191	0
ENSG00000189233	0.0004200	0.15270	NUGGC	389643	1.485	0
ENSG00000007516	0.0004222	0.15270	BAIAP3	8938	1.550	0
ENSG00000108370	0.0003698	0.15270	RGS9	8787	1.766	0
ENSG00000173114	0.0005806	0.15950	LRRN3	54674	-3.236	0

ENSG00000184005	0.0007544	0.15950	ST6GALNAC3	256435	-2.209	0
ENSG00000171097	0.0006989	0.15950	KYAT1	883	-1.612	0
ENSG00000120555	0.0004861	0.15950	SEPT7P9		-1.446	0
ENSG00000228663	0.0007307	0.15950	PSMD10P1		-1.167	0
ENSG00000229299	0.0008019	0.15950	AL121845.1		-1.090	1
ENSG00000261594	0.0007746	0.15950	TPBGL	1.01E+08	-1.039	0
ENSG00000232021	0.0007207	0.15950	LEF1-AS1	641518	-0.960	1
ENSG00000273148	0.0006326	0.15950	AL035563.1		-0.856	1
ENSG00000273802	0.0005353	0.15950	HIST1H2BG	8339	-0.842	0
ENSG00000223705	0.0008966	0.15950	NSUN5P1	155400	-0.815	0
ENSG00000235194	0.0008974	0.15950	PPP1R3E	90673	-0.784	0
ENSG00000244486	0.0006609	0.15950	SCARF2	91179	-0.776	0
ENSG00000137502	0.0008818	0.15950	RAB30	27314	-0.770	0
ENSG00000092096	0.0006557	0.15950	SLC22A17	51310	-0.767	0
ENSG00000221923	0.0007130	0.15950	ZNF880	400713	-0.709	0
ENSG00000198832	0.0006762	0.15950	SELENOM	140606	-0.618	0
ENSG00000203497	0.0007217	0.15950	PDCD4-AS1		-0.590	1
ENSG00000138381	0.0005207	0.15950	ASNSD1	54529	-0.560	0
ENSG00000090975	0.0006176	0.15950	PITPNM2	57605	-0.536	0
ENSG00000148925	0.0007429	0.15950	BTBD10	84280	-0.510	0
ENSG00000197903	0.0005912	0.15950	HIST1H2BK	85236	-0.469	0
ENSG00000263264	0.0006598	0.15950	AC119396.1		-0.456	0
ENSG00000185950	0.0006043	0.15950	IRS2	8660	-0.438	0
ENSG00000276141	0.0005944	0.15950	WHAMMP3		-0.438	0
ENSG00000156232	0.0007926	0.15950	WHAMM	123720	-0.392	0
ENSG00000067167	0.0007206	0.15950	TRAM1	23471	0.369	0
ENSG00000185624	0.0008631	0.15950	P4HB	5034	0.411	0
ENSG00000114850	0.0008739	0.15950	SSR3	6747	0.413	0
ENSG00000073921	0.0005095	0.15950	PICALM	8301	0.444	0
ENSG00000110876	0.0004601	0.15950	SELPLG	6404	0.451	0
ENSG00000136560	0.0008271	0.15950	TANK	10010	0.454	0
ENSG00000223482	0.0008504	0.15950	NUTM2A-AS1	728190	0.462	1
ENSG00000169155	0.0008190	0.15950	ZBTB43	23099	0.486	0
ENSG00000150977	0.0009066	0.15950	RILPL2	196383	0.506	0
ENSG00000086062	0.0006790	0.15950	B4GALT1	2683	0.563	0
ENSG00000168404	0.0005161	0.15950	MLKL	197259	0.570	0
ENSG00000239713	0.0008823	0.15950	APOBEC3G	60489	0.573	0
ENSG00000111885	0.0009044	0.15950	MAN1A1	4121	0.595	0
ENSG00000150961	0.0008627	0.15950	SEC24D	9871	0.643	0
ENSG00000034677	0.0007898	0.15950	RNF19A	25897	0.662	0
ENSG00000104312	0.0006624	0.15950	RIPK2	8767	0.837	0
ENSG00000062598	0.0008493	0.15950	ELMO2	63916	0.851	0
ENSG00000167508	0.0005402	0.15950	MVD	4597	0.939	0
ENSG00000103313	0.0007450	0.15950	MEFV	4210	0.984	0
ENSG00000196814	0.0006766	0.15950	MVB12B	89853	0.984	0
ENSG00000276734	0.0007366	0.15950	KIF5C	3800	1.025	0

ENSG00000134460	0.0009016	0.15950	IL2RA	3559	1.078	0
ENSG00000235106	0.0007150	0.15950	BRD3OS	266655	1.130	0
ENSG00000007264	0.0005291	0.15950	MATK	4145	1.140	0
ENSG00000090013	0.0007498	0.15950	BLVRB	645	1.206	0
ENSG00000180549	0.0004743	0.15950	FUT7	2529	1.279	0
ENSG00000115604	0.0007880	0.15950	IL18R1	8809	1.319	0
ENSG00000156127	0.0007676	0.15950	BATF	10538	1.355	0
ENSG00000166503	0.0008648	0.15950	HDGFL3	50810	1.423	0
ENSG00000107317	0.0007559	0.15950	PTGDS	5730	1.594	0
ENSG00000182621	0.0008123	0.15950	PLCB1	23236	1.679	0
ENSG00000157985	0.0008329	0.15950	AGAP1	116987	2.130	0
ENSG00000105374	0.0005465	0.15950	NKG7	4818	2.438	0
ENSG00000241587	0.0009243	0.16124	RN7SL482P		-0.860	1
ENSG00000068137	0.0009651	0.16555	PLEKHH3	79990	-1.606	0
ENSG00000112294	0.0009604	0.16555	ALDH5A1	7915	-0.976	0
ENSG00000137145	0.0009741	0.16572	DENND4C	55667	0.436	0
ENSG00000169896	0.0009906	0.16714	ITGAM	3684	1.655	0
ENSG00000178947	0.0013034	0.16769	SMIM10L2A	399668	-1.141	0
ENSG00000162373	0.0011430	0.16769	BEND5	79656	-1.097	0
ENSG00000229676	0.0012743	0.16769	ZNF492	57615	-0.972	0
ENSG00000167664	0.0012849	0.16769	TMIGD2	126259	-0.949	0
ENSG00000134461	0.0011764	0.16769	ANKRD16	54522	-0.889	0
ENSG00000253626	0.0012687	0.16769	EIF5AL1	143244	-0.881	0
ENSG00000140545	0.0012301	0.16769	MFGE8	4240	-0.766	0
ENSG00000127334	0.0011551	0.16769	DYRK2	8445	-0.719	0
ENSG00000180185	0.0010092	0.16769	FAHD1	81889	-0.694	0
ENSG00000237440	0.0012999	0.16769	ZNF737	1E+08	-0.608	0
ENSG00000122085	0.0011612	0.16769	MTERF4	130916	-0.512	0
ENSG00000138670	0.0011548	0.16769	RASGEF1B	153020	-0.503	0
ENSG00000184260	0.0012887	0.16769	HIST2H2AC	8338	-0.498	0
ENSG00000115896	0.0012048	0.16769	PLCL1	5334	-0.409	0
ENSG00000174652	0.0011143	0.16769	ZNF266	10781	-0.358	0
ENSG00000081237	0.0011821	0.16769	PTPRC	5788	0.425	0
ENSG00000166562	0.0011664	0.16769	SEC11C	90701	0.452	0
ENSG00000143870	0.0012435	0.16769	PDIA6	10130	0.458	0
ENSG00000125844	0.0012509	0.16769	RRBP1	6238	0.459	0
ENSG00000096063	0.0010674	0.16769	SRPK1	6732	0.504	0
ENSG00000168439	0.0011867	0.16769	STIP1	10963	0.520	0
ENSG00000110046	0.0011047	0.16769	ATG2A	23130	0.524	0
ENSG00000114209	0.0011965	0.16769	PDCD10	11235	0.538	0
ENSG00000028137	0.0010701	0.16769	TNFRSF1B	7133	0.572	0
ENSG00000140563	0.0011508	0.16769	MCTP2	55784	0.589	0
ENSG00000103494	0.0012255	0.16769	RPGRIP1L	23322	0.733	0
ENSG00000198018	0.0011946	0.16769	ENTPD7	57089	0.777	0
ENSG00000168280	0.0010440	0.16769	KIF5C	3800	0.825	0
ENSG00000048028	0.0011713	0.16769	USP28	57646	0.825	0

ENSG00000100031	0.0012492	0.16769	GGT1	2678	0.893	0
ENSG00000229491	0.0010304	0.16769	AC136489.1	1.05E+08	0.933	1
ENSG00000105186	0.0010266	0.16769	ANKRD27	84079	0.940	0
ENSG00000185697	0.0011384	0.16769	MYBL1	4603	1.087	0
ENSG00000205730	0.0010447	0.16769	ITPRIPL2	162073	1.092	0
ENSG00000169981	0.0012219	0.16769	ZNF35	7584	1.122	0
ENSG00000169252	0.0012545	0.16769	ADRB2	154	1.151	0
ENSG00000166548	0.0012755	0.16769	TK2	7084	1.381	0
ENSG00000188677	0.0012733	0.16769	PARVB	29780	1.758	0
ENSG00000278996	0.0013237	0.16827	FP671120.1		-0.801	1
ENSG00000101901	0.0013242	0.16827	ALG13	79868	-0.526	0
ENSG00000185669	0.0013387	0.16906	SNAI3	333929	1.005	0
ENSG00000169118	0.0013510	0.16930	CSNK1G1	53944	0.527	0
ENSG00000115607	0.0013570	0.16930	IL18RAP	8807	1.972	0
ENSG00000273748	0.0014163	0.17390	AL592183.1		-1.196	0
ENSG00000106133	0.0014228	0.17390	NSUN5P2	260294	-0.839	0
ENSG00000245532	0.0014361	0.17390	NEAT1	283131	0.442	1
ENSG00000271503	0.0014327	0.17390	CCL5	6352	1.457	0
ENSG00000169554	0.0014210	0.17390	ZEB2	9839	1.582	0
ENSG00000174080	0.0015148	0.17408	CTSF	8722	-0.992	0
ENSG00000166681	0.0014633	0.17408	BEX3	27018	-0.876	0
ENSG00000124635	0.0015194	0.17408	HIST1H2BJ	8970	-0.646	0
ENSG00000120837	0.0014905	0.17408	NFYB	4801	-0.505	0
ENSG00000180596	0.0014597	0.17408	HIST1H2BC	8347	-0.460	0
ENSG00000116489	0.0015222	0.17408	CAPZA1	829	0.387	0
ENSG00000054523	0.0015022	0.17408	KIF1B	23095	0.577	0
ENSG00000163110	0.0014769	0.17408	PDLIM5	10611	0.582	0
ENSG00000131149	0.0015024	0.17408	GSE1	23199	0.742	0
ENSG00000145107	0.0014557	0.17408	TM4SF19	116211	1.310	0
ENSG00000273673	0.0015630	0.17609	RPL9P9		-4.852	0
ENSG00000124243	0.0015692	0.17609	BCAS4	55653	-0.670	0
ENSG00000252835	0.0015761	0.17609	SCARNA21	677763	-0.325	1
ENSG00000153563	0.0015518	0.17609	CD8A	925	1.149	0
ENSG00000107554	0.0015826	0.17609	DNMBP	23268	1.525	0
ENSG00000230923	0.0015932	0.17632	LINC00309	150992	1.087	1
ENSG00000025293	0.0016019	0.17634	PHF20	51230	0.436	0
ENSG00000179348	0.0016618	0.18004	GATA2	2624	-2.136	0
ENSG00000186918	0.0016513	0.18004	ZNF395	55893	-0.499	0
ENSG00000178074	0.0016562	0.18004	C2orf69	205327	0.445	0
ENSG00000277224	0.0016730	0.18031	HIST1H2BF	8343	-1.147	0
ENSG00000196866	0.0017402	0.18049	HIST1H2AD	3013	-0.764	0
ENSG00000282540	0.0017751	0.18049	WHAMMP2		-0.653	0
ENSG00000123297	0.0018260	0.18049	TSFM	10102	-0.588	0
ENSG00000284691	0.0017779	0.18049	AC073111.5		-0.562	0
ENSG00000158373	0.0017576	0.18049	HIST1H2BD	3017	-0.560	0
ENSG00000180787	0.0018319	0.18049	ZFP3	124961	-0.536	0

ENSG00000132475	0.0018313	0.18049	H3F3B	3021	-0.374	0
ENSG00000165416	0.0018006	0.18049	SUGT1	10910	0.398	0
ENSG00000102572	0.0018068	0.18049	STK24	8428	0.399	0
ENSG00000175348	0.0018082	0.18049	TMEM9B	56674	0.442	0
ENSG00000272489	0.0016895	0.18049	AL132656.3		0.564	1
ENSG00000141068	0.0017557	0.18049	KSR1	8844	0.712	0
ENSG00000118402	0.0017119	0.18049	ELOVL4	6785	0.831	0
ENSG00000173852	0.0018135	0.18049	DPY19L1	23333	0.923	0
ENSG00000115183	0.0017259	0.18049	TANC1	85461	0.946	0
ENSG00000131459	0.0018063	0.18049	GFPT2	9945	1.378	0
ENSG00000180644	0.0018325	0.18049	PRF1	5551	1.423	0
ENSG00000256981	0.0017822	0.18049	AC134349.2		3.226	0
ENSG00000268279	0.0018518	0.18152	AC090004.1		1.190	0
ENSG00000168884	0.0018877	0.18248	TNIP2	79155	0.470	0
ENSG00000135821	0.0018771	0.18248	GLUL	2752	0.715	0
ENSG00000005020	0.0018902	0.18248	SKAP2	8935	1.107	0
ENSG00000171517	0.0018970	0.18248	LPAR3	23566	1.623	0
ENSG00000169738	0.0019276	0.18456	DCXR	51181	0.683	0
ENSG00000145882	0.0019470	0.18528	PCYOX1L	78991	-0.438	0
ENSG00000166949	0.0019532	0.18528	SMAD3	4088	0.537	0
ENSG00000144935	0.0020551	0.18692	TRPC1	7220	-1.872	0
ENSG00000072110	0.0021520	0.18692	ACTN1	87	-0.984	0
ENSG00000251623	0.0020592	0.18692	AC139491.5		-0.840	0
ENSG00000129535	0.0021205	0.18692	NRL	4901	-0.704	0
ENSG00000168298	0.0021471	0.18692	HIST1H1E	3008	-0.563	0
ENSG00000274641	0.0021091	0.18692	HIST1H2BO	8348	-0.530	0
ENSG00000131351	0.0020990	0.18692	HAUS8	93323	-0.490	0
ENSG00000255455	0.0020569	0.18692	AP003486.1	1.04E+08	-0.468	1
ENSG00000168970	0.0019807	0.18692	JMJD7-PLA2G4B	8681	-0.446	0
ENSG00000128699	0.0020480	0.18692	ORMDL1	94101	-0.409	0
ENSG00000132773	0.0020336	0.18692	TOE1	114034	-0.388	0
ENSG00000122417	0.0020934	0.18692	ODF2L	57489	0.362	0
ENSG00000135597	0.0020980	0.18692	REPS1	85021	0.394	0
ENSG00000113716	0.0021440	0.18692	HMGXB3	22993	0.429	0
ENSG00000131446	0.0020917	0.18692	MGAT1	4245	0.525	0
ENSG00000122678	0.0020648	0.18692	POLM	27434	0.549	0
ENSG00000123130	0.0021167	0.18692	ACOT9	23597	0.574	0
ENSG00000157557	0.0021417	0.18692	ETS2	2114	0.595	0
ENSG00000117010	0.0020612	0.18692	ZNF684	127396	0.854	0
ENSG00000243836	0.0020802	0.18692	WDR86-AS1	1E+08	0.928	1
ENSG00000120451	0.0021964	0.18704	SNX19	399979	-0.407	0
ENSG00000274290	0.0022171	0.18704	HIST1H2BE	8344	-0.392	0
ENSG00000198830	0.0022262	0.18704	HMG2	3151	-0.323	0
ENSG00000124783	0.0022057	0.18704	SSR1	6745	0.378	0
ENSG00000234883	0.0021679	0.18704	MIR155HG	114614	0.747	1
ENSG00000167315	0.0021962	0.18704	ACAA2	10449	0.771	0

ENSG00000258512	0.0022002	0.18704	LINC00239	145200	0.792	1
ENSG00000090376	0.0022255	0.18704	IRAK3	11213	1.675	0
ENSG00000282130	0.0022571	0.18735	NSMCE3	56160	-0.322	0
ENSG00000070882	0.0022560	0.18735	OSBPL3	26031	0.447	0
ENSG00000105717	0.0022509	0.18735	PBX4	80714	0.892	0
ENSG00000033122	0.0023499	0.18753	LRRC7	57554	-1.800	0
ENSG00000267128	0.0023501	0.18753	RNF157-AS1	1.01E+08	-1.121	1
ENSG00000281005	0.0024386	0.18753	LINC00921	283876	-0.870	1
ENSG00000075131	0.0023922	0.18753	TIPIN	54962	-0.631	0
ENSG00000230724	0.0023005	0.18753	LINC01001		-0.577	1
ENSG00000181274	0.0023766	0.18753	FRAT2	23401	-0.560	0
ENSG00000198040	0.0024477	0.18753	ZNF84	7637	-0.557	0
ENSG00000250889	0.0024250	0.18753	LINC01336	1.04E+08	-0.535	1
ENSG00000183513	0.0024494	0.18753	COA5	493753	-0.505	0
ENSG00000116863	0.0023111	0.18753	ADPRHL2	54936	-0.416	0
ENSG00000230124	0.0024191	0.18753	ACBD6	84320	-0.410	0
ENSG00000101654	0.0024296	0.18753	RNMT	8731	-0.319	0
ENSG00000120063	0.0023307	0.18753	GNA13	10672	0.312	0
ENSG00000174437	0.0023355	0.18753	ATP2A2	488	0.317	0
ENSG00000106948	0.0024506	0.18753	AKNA	80709	0.535	0
ENSG00000085511	0.0023189	0.18753	MAP3K4	4216	0.567	0
ENSG00000148248	0.0022995	0.18753	SURF4	6836	0.575	0
ENSG00000110660	0.0023621	0.18753	SLC35F2	54733	0.734	0
ENSG00000263083	0.0023324	0.18753	TAS2R64P		0.996	0
ENSG00000184164	0.0023536	0.18753	CRELD2	79174	1.023	0
ENSG00000057704	0.0024130	0.18753	TMCC3	57458	1.543	0
ENSG00000165915	0.0025071	0.18778	SLC39A13	91252	-0.645	0
ENSG00000168612	0.0025514	0.18778	ZSWIM1	90204	-0.516	0
ENSG00000274272	0.0025391	0.18778	AC069281.2		-0.399	1
ENSG00000265491	0.0024749	0.18778	RNF115	27246	0.345	0
ENSG00000074266	0.0025702	0.18778	EED	8726	0.377	0
ENSG00000136813	0.0025310	0.18778	ECPAS	23392	0.406	0
ENSG00000242247	0.0024910	0.18778	ARFGAP3	26286	0.440	0
ENSG00000141551	0.0025586	0.18778	CSNK1D	1453	0.465	0
ENSG00000072778	0.0025333	0.18778	ACADVL	37	0.505	0
ENSG00000146731	0.0025724	0.18778	CCT6A	908	0.532	0
ENSG00000011478	0.0024934	0.18778	QPCTL	54814	0.597	0
ENSG00000103222	0.0025634	0.18778	ABCC1	4363	0.831	0
ENSG00000173930	0.0025447	0.18778	SLCO4C1	353189	0.880	0
ENSG00000172663	0.0025843	0.18798	TMEM134	80194	-0.490	0
ENSG00000165355	0.0026440	0.19097	FBXO33	254170	-0.388	0
ENSG00000165312	0.0026395	0.19097	OTUD1	220213	-0.363	0
ENSG00000159256	0.0026717	0.19230	MORC3	23515	-0.355	0
ENSG00000264519	0.0026985	0.19355	RN7SL596P		0.897	1
ENSG00000271447	0.0027095	0.19366	MMP28	79148	-1.580	0
ENSG00000165943	0.0027362	0.19423	MOAP1	64112	-0.347	0

ENSG00000274093	0.0027362	0.19423	AC009032.1		0.904	1
ENSG00000167740	0.0027496	0.19450	CYB5D2	124936	-0.685	0
ENSG00000185838	0.0027743	0.19558	GNB1L	54584	-0.910	0
ENSG00000275221	0.0027906	0.19605	HIST1H2AK	8330	-0.479	0
ENSG00000283707	0.0028827	0.19645	AC275455.1		-1.254	0
ENSG00000102409	0.0028813	0.19645	BEX4	56271	-0.565	0
ENSG00000276966	0.0028528	0.19645	HIST1H4E	8367	-0.504	0
ENSG00000131381	0.0028332	0.19645	RBSN	64145	-0.367	0
ENSG00000181896	0.0028684	0.19645	ZNF101	94039	-0.364	0
ENSG00000131504	0.0028110	0.19645	DIAPH1	1729	0.327	0
ENSG00000118503	0.0028615	0.19645	TNFAIP3	7128	0.371	0
ENSG00000196440	0.0028806	0.19645	ARMCX4	1E+08	0.379	0
ENSG00000138376	0.0028515	0.19645	BARD1	580	0.791	0
ENSG00000220412	0.0028917	0.19645	AL356234.1		1.171	0
ENSG00000183691	0.0029149	0.19673	NOG	9241	-1.735	0
ENSG00000205181	0.0029138	0.19673	LINC00654	149837	-0.814	1
ENSG00000166387	0.0030410	0.19689	PPFIBP2	8495	-1.377	0
ENSG00000168876	0.0030165	0.19689	ANKRD49	54851	-0.443	0
ENSG00000069966	0.0029913	0.19689	GNB5	10681	-0.390	0
ENSG00000127980	0.0029827	0.19689	PEX1	5189	-0.388	0
ENSG00000125827	0.0030069	0.19689	TMX4	56255	0.487	0
ENSG00000197157	0.0029635	0.19689	SND1	27044	0.505	0
ENSG00000047365	0.0029403	0.19689	ARAP2	116984	0.535	0
ENSG00000172216	0.0030389	0.19689	CEBPB	1051	0.610	0
ENSG00000277452	0.0030511	0.19689	RN7SL473P		0.744	1
ENSG00000222041	0.0030430	0.19689	CYTOR	112597	0.890	1
ENSG00000256128	0.0030111	0.19689	LINC00944		1.396	1
ENSG00000145649	0.0029313	0.19689	GZMA	3001	1.694	0
ENSG00000285210	0.0030090	0.19689	AL136382.1		1.764	1
ENSG00000197057	0.0029876	0.19689	DTHD1	401124	2.504	0
ENSG00000102218	0.0030769	0.19755	RP2	6102	0.400	0
ENSG00000174946	0.0030805	0.19755	GPR171	29909	0.905	0
ENSG00000256625	0.0030981	0.19771	AC092747.2		-0.865	0
ENSG00000131148	0.0031023	0.19771	EMC8	10328	-0.399	0
ENSG00000257613	0.0031236	0.19845	LINC01481		-1.287	1

NC: 0 = protein coding gene/ 1= noncoding gene

Table S3. Results of gene set enrichment analysis Results (q value< 0.1)

Database	Term/ Pathway	q.value	set. size
Gene Ontology	immune response	0.000067	211
	inflammatory response	0.000067	250
	cytokine-mediated signaling pathway	0.000068	195
	external side of plasma membrane	0.000232	141

	neutrophil degranulation	0.000232	404
	focal adhesion	0.002643	349
	neutrophil chemotaxis	0.019314	34
	positive regulation of GTPase activity	0.025147	253
	MAPK cascade	0.028165	204
	innate immune response	0.044018	290
	cellular response to lipopolysaccharide	0.071555	91
	chemotaxis	0.078604	77
	cell adhesion	0.080870	261
	cell surface receptor signaling pathway	0.080870	161
	SH3 domain binding	0.080870	100
	vascular endothelial growth factor receptor signaling pathway	0.083274	59
	positive regulation of gene expression	0.093413	209
	actin binding	0.095751	210
Consensus	Cytokine-cytokine receptor interaction	0.012808	140
	Axon guidance	0.070456	197
	Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling	0.072542	27
	Hematopoietic cell lineage	0.072542	58
Hallmark	HALLMARK_INFLAMMATORY_RESPONSE	0.000002	156
	HALLMARK_ALLOGRAFT_REJECTION	0.000002	167
	HALLMARK_TNFA_SIGNALING_VIA_NFKB	0.000007	181
	HALLMARK_IL2_STAT5_SIGNALING	0.000015	173
	HALLMARK_INTERFERON_GAMMA_RESPONSE	0.000046	180
	HALLMARK_COMPLEMENT	0.000262	158
	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	0.000446	116
	HALLMARK_APICAL_JUNCTION	0.001144	142
	HALLMARK_MTORC1_SIGNALING	0.001144	191
	HALLMARK_COAGULATION	0.003090	79
	HALLMARK_ESTROGEN_RESPONSE_LATE	0.003386	136
	HALLMARK_KRAS_SIGNALING_UP	0.005405	138
	HALLMARK_UV_RESPONSE_DN	0.008798	117
	HALLMARK_IL6_JAK_STAT3_SIGNALING	0.015578	65
	HALLMARK_CHOLESTEROL_HOMEOSTASIS	0.017141	69
	HALLMARK_APOPTOSIS	0.019266	137
	HALLMARK_XENOBIOTIC_METABOLISM	0.025292	149
	HALLMARK_MITOTIC_SPINDLE	0.032129	178
	HALLMARK_ESTROGEN_RESPONSE_EARLY	0.048596	149
	HALLMARK_HYPOXIA	0.049914	152
	HALLMARK_ANGIOGENESIS	0.059442	21
	HALLMARK_GLYCOLYSIS	0.060683	161
	HALLMARK_ANDROGEN_RESPONSE	0.067492	88
	HALLMARK_PI3K_AKT_MTOR_SIGNALING	0.067492	96
	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	0.087219	108

Transcription Factors	JUN	0.000096	167
	FOS	0.003362	67
	ATF2.FOS.FOSB.FOSL1.FOSL2.JUN.JUNB.JUND	0.003612	104
	CEBPA	0.005055	103
	RELA	0.005055	227
	IRF8	0.005055	45
	TBX21	0.005055	33
	STAT1	0.007111	193
	NR3C1	0.007111	268
	NFKB1.NFKB2.REL.RELA.RELB	0.009305	139
	HIF1A	0.011545	188
	REL	0.012129	52
	STAT5B	0.012129	56
	ATF3	0.014746	32
	STAT4	0.014746	15
	CBFB	0.016229	12
	NFKB1	0.019780	83
	CEBPB	0.020352	176
	JUNB	0.033348	34
	MYB	0.043939	68
	AHR	0.043939	33
	ATF2	0.044156	39
	NFKB1.RELA	0.045800	80
	WT1	0.045800	51
	NFE2L2	0.045800	109
	FOXP3	0.081165	38
	POU2F2	0.081165	15
	ATF6	0.081165	16
	ETS2	0.081165	57
	IRF3	0.084149	21
	NFATC2	0.084149	42
	POU2F1	0.084149	38
	HSF1	0.084149	116
	PAX5	0.084149	51
	STAT5A.STAT5B	0.084149	43
	CREM	0.084528	37
	STAT3	0.084528	445
	GATA3	0.087320	45
	STAT5A	0.092316	65
	PPARG	0.095112	91
	TFAP4	0.095112	18
	YBX1	0.095112	26
	EBF1	0.095112	37
	IRF1	0.095112	182
	TP63	0.095112	144

ATF1	0.095112	23
KLF4	0.095112	63
CUX1	0.095112	77
CREB1	0.099604	247

