

UNIVERSIDADE FEDERAL DO PARANÁ

CRISTIANE SATO MARA MULLER

IDENTIFICAÇÃO DE LncRNAs CONTENDO SÍTIOS PARA PUMILIO:
ANÁLISE *IN SILICO* DA EXPRESSÃO DIFERENCIAL, CO-EXPRESSÃO DE
SEUS ALVOS E IMPACTO NA SOBREVIDA EM DIFERENTES TIPOS DE
CÂNCER

CURITIBA
2021

CRISTIANE SATO MARA MULLER

IDENTIFICAÇÃO DE LncRNAs CONTENDO SÍTIOS PARA PUMILIO:
ANÁLISE *IN SÍLICO* DA EXPRESSÃO DIFERENCIAL, CO-EXPRESSÃO DE
SEUS ALVOS E IMPACTO NA SOBREVIVÊNCIA EM DIFERENTES TIPOS DE
CÂNCER

Dissertação apresentada ao Programa de Pós-Graduação em Genética do Departamento de Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, como parte das exigências para obtenção do título de Mestre em Genética.

Orientadora:
Prof.^a Dr.^a Daniela Fiori Gradia

Coorientadora:
Prof.^a Dr.^a Jaqueline Carvalho de Oliveira

CURITIBA
2021

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

Muller, Cristiane Sato Mara
Identificação de LncRNAs contendo sítios para PUMILIO : análise *in silico* da expressão diferencial, co-expressão de seus alvos e impacto na sobrevivência em diferentes tipos de câncer. / Cristiane Sato Mara Muller. – Curitiba, 2021.
107 p.: il.

Orientadora: Daniela Fiori Gradia.
Coorientadora: Jaqueline Carvalho de Oliveira.

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

1. RNA. 2. Câncer. 3. RNA longo não codificante. I. Título. II. Gradia, Daniela Fiori. III. Oliveira, Jaqueline Carvalho de. IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (22. ed.) 572.8



MINISTÉRIO DA EDUCAÇÃO
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 dissertação de Mestrado de CRISTIANE SATO MARA MULLER intitulada: *Identificação de lncRNAs contendo sítios de ligação a PUMILIO: análise in silico da expressão diferencial, co-expressão de seus alvos e impacto na sobrevivência em diferentes tipos de câncer.*, sob orientação da Profa. Dra. DANIELA FIORI GRADIA, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre 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, 01 de Julho de 2021.

Assinatura Eletrônica

06/07/2021 16:27:41.0

DANIELA FIORI GRADIA

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

06/07/2021 19:25:36.0

PATRICIA SHIGUNOV

Avaliador Externo (INSTITUTO CARLOS CHAGAS/FIOCRUZ)

Assinatura Eletrônica

06/07/2021 17:39:47.0

DANIELLE MALHEIROS FERREIRA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Setor de Ciências Biológicas, Centro Politécnico - CURITIBA - Paraná - Brasil
CEP 81531-980 - Tel: (41) 3361-1587 - E-mail: ppg-gen@ufpr.br

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.
Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 99570

Para autenticar este documento/assinatura, acesse <https://www.prpg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp>
e insira o código 99570

Dedico esse trabalho a minha família que com muito amor, compreensão e parceria possibilitaram que eu chegasse aonde cheguei e me incentivaram sempre a nunca desistir.

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me concedido a força e perseverança para não desistir.

Ao meu marido que sempre esteve ao meu lado e me incentivou a sempre continuar.

À minha filha pela compreensão sobre as longas manhãs e tardes que estive ausente. Obrigada por ser uma filha maravilhosa e independente.

Aos meus pais por sempre demonstrarem seu amor e orgulho, cujo apoio permitiu que eu chegasse até aqui.

À minha irmã, parceira em todos os momentos, minha melhor amiga.

Às minhas orientadoras, Daniela Fiori Gradia e Jaqueline Carvalho de Oliveira que sempre me fizeram sentir amiga, acolhida e orientada.

Às minhas colegas de laboratório por sempre me orientar quando precisei.

Às minhas amigas Luana e Branca que iniciaram esta jornada comigo e se tornaram minhas grandes amigas.

Ao Igor Giner que me auxiliou muitas vezes e, hoje faz parte deste trabalho.

À Carolina Mathias por me iniciar na bioinformática com toda paciência possível.

À Capes pelo apoio financeiro.

À Dr Enilze Ribeiro que está ao meu lado em um momento difícil.

E, novamente, um obrigada especial à minha orientadora Daniela Gradia que me apoia em todos os momentos e hoje é uma inspiração para mim.

RESUMO

Os RNAs longos não codificantes (lncRNA) representam uma família heterogênea de RNAs que atuam como reguladores de vários processos biológicos. Através de mecanismos distintos, controlam a expressão genética em múltiplos níveis e estão envolvidos em diversos cenários fisiológicos e patológicos. Os lncRNAs e as proteínas de ligação ao RNA (RBP) podem associar-se em complexos de ribonucleoproteínas (RNP), tanto no citoplasma como no núcleo. Neste estudo, concentramos nos lncRNAs que apresentam sítios de ligação para PUMILIO, uma família de RBPs envolvida na regulação pós-transcricional de centenas de RNAs. Comparamos os níveis de expressão dos lncRNAs em 9 tipos de câncer com seus respectivos tecidos não tumorais e buscamos genes co-expressos com estes lncRNAs usando a base de dados do TCGA (*The Cancer Atlas Genome*). Em seguida, analisamos o impacto da Sobrevida Global e Livre da Doença entre os lncRNA e seus genes alvos, selecionados na co-expressão, em cada tipo de câncer. Nossos resultados apontaram o lncRNA NORAD como sendo o mais relevante em câncer de mama, co-expresso significativamente (0,55) com o gene *RALGAP β* , um gene alvo da PUMILIO, relacionado com a estabilidade cromossômica durante a divisão celular. Além de co-expresso com NORAD, a expressão do gene *RALGAP β* também apresentou um p-valor significativo na Sobrevida Global (0,015) no câncer de mama. Esse estudo sugere pela primeira vez o potencial papel do lncRNA NORAD na instabilidade cromossômica através do eixo regulatório NORAD-PUMILIO-*RALGAP β* . No câncer de rim, os lncRNAs HCG27, MALAT1 e LINC01011, diferencialmente expressos no tumor em relação ao tecido normal, aparecem significativamente co-expressos (0,47), (0,25) e (0,18) respectivamente, com o gene *MSH5*, um gene alvo da PUMILIO envolvido no reparo do DNA. Além disso, as análises de Sobrevida Global para os lncRNAs MALAT1, LINC01011 e o gene *MSH5* apresentaram um valor-p significativo de 0,0099, 0,0073 e 0,00096, respectivamente. Isso sugere uma possível relação deste conjunto de lncRNAs e o gene *MSH5* no câncer de rim, podendo representar um eixo regulatório importante nos processos de reparo de DNA.

Palavras-chave: lncRNA, PUMILIO, câncer, NORAD, *RALGAP β* , HCG27, *MSH5*

ABSTRACT

Long non-coding RNAs (lncRNA) represent a heterogeneous family of RNAs that act as regulators of various biological processes. Through distinct mechanisms, they control gene expression at multiple levels and are involved in various physiological and pathological scenarios. lncRNAs and RNA-binding proteins (RBPs) can associate in ribonucleoprotein (RNP) complexes in both the cytoplasm and nucleus. In this study, we focused on lncRNAs that have binding sites for PUMILIO, a family of RBPs involved in post-transcriptional regulation of hundreds of RNAs. We compared the expression levels of lncRNAs in 9 cancer types with their respective non-tumor tissues and searched for genes co-expressed with these lncRNAs using the TCGA (The Cancer Atlas Genome) database. We then analyzed the impact of Overall and Disease Free Survival between the lncRNAs and their target genes, selected on co-expression, in each cancer type. Our results pointed out the lncRNA NORAD as the most relevant in breast cancer, co-expressed significantly (0.55) with the *RALGAP β* gene, a target gene of PUMILIO, related to chromosome stability during cell division. In addition to the co-expression analysis the *RALGAP β* gene also showed a significant p-value in Overall Survival (0.015) in breast cancer. This study suggests for the first time, the potential role of lncRNA NORAD in chromosomal instability through the NORAD-PUMILIO-*RALGAP β* regulatory axis. In kidney cancer, differentially expressed lncRNAs HCG27, MALAT1 and LINC01011, relative to normal tissue, appear significantly co-expressed (0.47), (0.25) and (0.18) respectively with the *MSH5* gene, a PUMILIO target gene involved in DNA repair. Furthermore, the Overall Survival analyses for the lncRNAs MALAT1, LINC01011 and the *MSH5* gene showed a significative p-value (0.0099), (0.0073) and (0.00096) respectively. This suggests a possible relationship of the set of lncRNAs and the *MSH5* gene in kidney cancer and may represent an important regulatory axis in DNA repair processes.

Keywords: lncRNA, PUMILIO, cancer, NORAD, *RALGAP β* , HCG27, *MSH5*

LISTA DE TABELAS

TABELA 1: TIPOS DE CÂNCER ANALISADOS E QUANTIDADE DE AMOSTRA.....	29
TABELA 2: RELAÇÃO DOS lncRNAs SELECIONADOS EM CADA TIPO DE CÂNCER.....	35
TABELA 3: LncRNAs DE COM SÍTIOS PARA PUMILIO E logFC > 0 EM PELO MENOS UM SUBTIPO DE BRCA.....	37
TABELA 4: LncRNAs DIFERENCIALMENTE EXPRESSOS (p-valor < 0,05) EM CADA TIPO DE CÂNCER E QUANTIDADES DE GENES CO-EXPRESSOS (p ajustado)*	38
TABELA 5: RELAÇÃO DE CO-EXPRESSÃO ENTRE lncRNAs COM SÍTIOS PARA PUMILIO E GENES ALVO DA PUMILIO NO BRCA (p < 0,00001).....	39
TABELA 6: LncRNAs E GENES CO-EXPRESSOS COM SOBREVIDA GLOBAL (p-valor) PARA OS CÂNCERES HNSC, KIRC, LIHC E PRAD.....	40
TABELA 7: LncRNAs E GENES CO-EXPRESSOS COM SOBREVIDA LIVRE DA DOENÇA (p-valor) PARA COAD, HNSC, LIHC, LUAD E PRAD.....	42

LISTA DE FIGURAS

FIGURA 1: Funções dos LncRNAs.....	15
FIGURA 2: Relação entre o lncRNA NORAD e a proteína PUMILIO.....	18
FIGURA 3: Classificação dos LncRNA.....	19
FIGURA 4: LncRNAs associados à carcinogênese.....	20
FIGURA 5: Domínio PUF de ligação da PUMILIO (PUM-HD).....	23
FIGURA 6: Mecanismo de ação da PUM.....	25
FIGURA 7: Comparação dos lncRNAs com sítios para PUMILIO e diferencialmente expressos nos cânceres avaliados	35
FIGURA 8: Análise da expressão diferencial do <i>MSH5</i> , logFC > 0,5 (p-valor < 0,05) nos tipos de câncer COAD, HNSC e KIRC.....	44
FIGURA 9: Análise da expressão diferencial do <i>MSH5</i> , logFC > 0,5 (p-valor < 0,05) nos tipos de câncer LIHC, LUAD e PRAD.....	45
FIGURA 10: <i>MSH5</i> e seus lncRNAs co-expressos nos tipos de câncer KIRC, COAD, LIHC, LUAD, PRAD e HNSC.....	45
FIGURA 11: Sobrevida global (p < 0,05) para o gene alvo <i>MSH5</i> em KIRC, LIHC e PRAD.....	46
FIGURA 12: Sobrevida livre da doença (p < 0,05) para o gene alvo <i>MSH5</i> em COAD e LIHC.....	46
FIGURA 13: Expressão diferencial do lncRNA NORAD em amostras do BRCA para subtipo Luminal A, Basal-like e tecido normal.....	47
FIGURA 14: Expressão diferencial do Gene <i>RALGAPβ</i> no BRCA, nos subtipos Basal-like, Luminal A e tecido normal.....	48
FIGURA 15: Comparação da expressão diferencial do lncRNA NORAD e o gene <i>RALGAPβ</i>	48
FIGURA 16: Análise de sobrevida global mostraram uma relação significativa do <i>RALGAPβ</i> no subtipo Luminal A.....	49
QUADRO 1: LncRNAs COM SÍTIOS PARA PUMILIO.....	34

Sumário

1. Introdução	12
2. Revisão Bibliográfica.....	14
2.1 RNAs longos não codificantes	14
2.1.1 Estrutura e função	14
2.1.2 Classificação	19
2.1.3 LncRNAs e o Câncer.....	20
2.2 Proteínas de Ligação ao RNA	22
2.3 PUMILIO	23
2.3.1 Estrutura.....	23
3. JUSTIFICATIVA	26
4.1 Objetivo Geral	27
4.2 Objetivos específicos	27
5. Materiais e Métodos	28
5.1 Seleção dos lncRNAs com sítios para PUMILIO	28
5.2 Análises de expressão diferencial para lncRNA	28
5.3 Análise de co-expressão.....	30
5.4 Análise de expressão diferencial para os alvos selecionados com sítios para PUMILIO.....	31
5.5 Análise de Sobrevida Global e Livre da Doença.....	31
5.6 Comparação dos lncRNA e Alvos co-expressos entre os 9 tipos de câncer	32
6. Resultados	32
6.1 Seleção dos lncRNA	32
6.2 Análise de expressão diferencial para lncRNA	35
6.3 Análise de co-expressão e expressão diferencial dos genes alvos	38
6.4 Análises de Sobrevida Global e Sobrevida Livre da Doença.....	40
6.5 Comparação dos lncRNAs e genes alvo de PUMILIO co-expressos entre os 9 tipos de câncer.....	44
7. Discussão	50
8. Conclusão	58
9. Referências Bibliográficas	59
10. TABELAS SUPLEMENTARES.....	70

1. Introdução

Os RNAs longos não codificantes (lncRNAs) representam uma família heterogênea de RNAs que foram definidos pela primeira vez por terem mais de 200 nucleotídeos na sua cadeia e pela ausência de qualquer quadro de leitura aberto (ORF) (Bhan et al., 2017), embora alguns deles possam produzir pequenos peptídeos funcionais (CABILI; TRAPNELL; GOFF; KOZIOL *et al.*, 2011).

O número exato de lncRNAs transcritos no genoma humano é uma questão em debate, mas a maioria das estimativas coloca o número em dezenas de milhares (Garzon et al., 2014). O catálogo dos lncRNAs, agora em constante expansão, tornou-se evidente a partir dos esforços para registrar as características funcionais do genoma humano, que mostraram que a grande maioria do genoma é transcrito (Moraes & Góes, 2016).

Os lncRNAs podem ser classificados em quatro grupos - *sense*, *antisense*, bidirecional e intergênico - de acordo com a posição genômica relativa do lncRNA a um gene codificador de proteínas (L. Wang et al., 2017). Estes RNAs funcionais têm papéis essenciais nos diferentes tipos celulares e durante a diferenciação celular podendo atuar de diferentes formas: como guias para modificadores de cromatina ao interagirem com proteínas associadas ao DNA; como reguladores transcricionais que afetam interações de polimerase de RNA e/ou interações de fatores de transcrição; como plataformas de interação, recrutando múltiplas proteínas e formando complexos ribonucleicos; e como iscas ao se associarem a outras moléculas, impedindo-as de se ligarem aos alvos originais (Zeng et al., 2018).

As proteínas de ligação ao RNA (RBPs) desempenham papéis vitais na regulação pós-transcricional atuando através da modulação do *splicing*, transporte, estabilidade, e tradução dos seus RNAs alvo (Gu et al., 2004).

Representando uma família conservada de proteínas de ligação ao RNA, PUMILIO atua como um regulador negativo da expressão gênica (Guan et al., 2018). Membros desta família de proteínas (proteínas PUF) regulam a expressão do mRNA através da ligação a regiões 3'UTR do seu mRNA alvo (L. Wang et al., 2017). Em humanos, PUMILIO 1 (PUM1) e PUMILIO 2 (PUM2) - representantes

desta família - se ligam a transcritos contendo a sequência de nucleotídeos UGUANAUA, referida como elemento de resposta à PUMILIO (PRE) (Smialek et al., 2020). Muitos alvos de PUM em mamíferos foram identificados usando abordagens de alto desempenho (Bohn et al., 2018), revelando várias funções destas proteínas na homeostase germinal e no controle do ciclo celular (Lin et al., 2018).

Em humanos, vários mRNAs envolvidos com doenças neurológicas, câncer e doenças cardiovasculares são alvos de PUM, sugerindo a influência da PUM na regulação genética destas doenças (Bohn JA. et al., 2017). Por exemplo, foram encontrados níveis elevados de PUM1 no adenocarcinoma pancreático, em comparação com os tecidos saudáveis adjacentes e, análises *in vivo* mostram que o silenciamento de PUM1 pode reduzir a proliferação celular, a capacidade de invasão, a metástase e promover a apoptose (Dai et al., 2019).

No sentido de controlar o ciclo celular, as proteínas PUMILIO aparecem associadas ao lncRNA NORAD (RNA longo não codificante ativado por danos no DNA), preservando a estabilidade do genoma. NORAD atua como uma plataforma de ligação molecular para as proteínas PUMILIO pela presença de vários sítios de ligação à proteína. Na ausência de NORAD, a hiperatividade da PUMILIO induz células cariotipicamente saudáveis a desenvolver aneuploidias durante a divisão celular (S. Lee et al., 2016). A elucidação da interação reguladora do NORAD e da proteína PUMILIO expandiu a compreensão dos lncRNAs e das suas funções.

Para além do NORAD, outros lncRNAs têm potenciais sítios de ligação para PUMILIO, mas pouco se sabe sobre os papéis biológicos e moleculares deles. Aqui pretendemos investigar, *in silico*, o conjunto de lncRNAs que apresentam múltiplos sítios de ligação para PUMILIO, analisando o seu nível de expressão no contexto de fenótipos normais ou em 9 tipos de câncer analisados no Projeto Pan Câncer (BLCA, COAD, ESCA, HNSC, KIRC, LIHC, LUAD, PRAD e BRCA), procurando genes co-expressos e a sua participação nos possíveis mecanismos de desregulação celular durante a tumorigênese mediada pela PUMILIO.

2. Revisão Bibliográfica

2.1 RNAs longos não codificantes

2.1.1 Estrutura e função

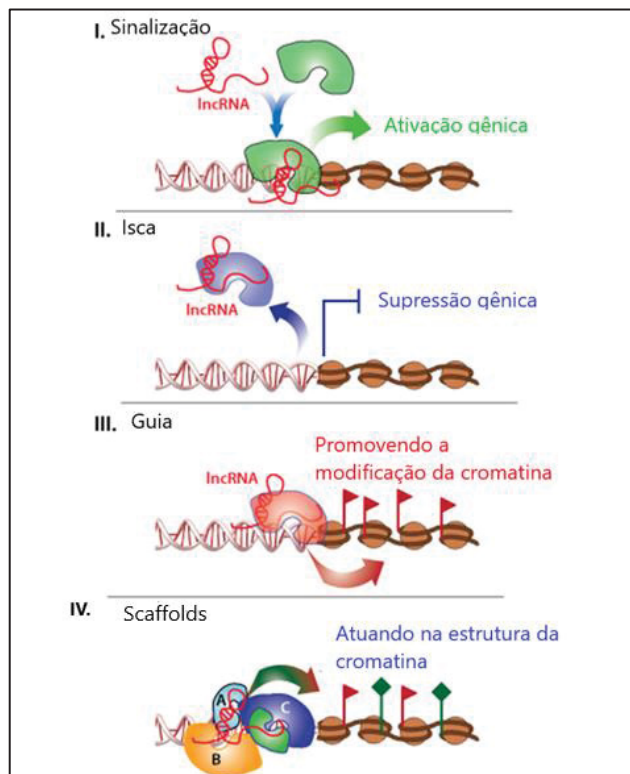
RNAs longos não codificantes (lncRNAs) são transcritos maiores que 200 nucleotídeos e em geral, não codificam proteínas (L. Sheng, M. Christopher, 2016) apesar de alguns serem capazes de produzir pequenos peptídeos funcionais (CABILI; TRAPNELL; GOFF; KOZIOL *et al.*, 2011). Sua biogênese é similar a dos RNAs codificantes de proteínas. São transcritos, em sua maioria, pela RNA polimerase II, podendo ser poliadenilados, capeados e sujeitos a processamento (Ma et al., 2013).

Há diversas fontes que tentam explicar o surgimento dos lncRNAs (Zhu et al., 2013). Eles podem se formar a partir de perturbações do quadro de leitura dos genes codificadores de proteínas; como rearranjos cromossômicos; por retrotransposição durante o processo de replicação; por repetições de sequências em *tandem*; e por fim, podem ser gerados pela inserção de um elemento transponível (Zhu et al., 2013).

A localização celular dos lncRNAs é variada, e provavelmente relacionada a sua função, podendo ser observados no núcleo, citoplasma, ou em um ou mais compartimentos das células (Mercer et al., 2008). No entanto, os padrões de localização de alguns lncRNAs são únicos, como o Gamafu, que está localizado exclusivamente em *speckles* nucleares (Carrero et al., 2006).

Sabemos, atualmente, que lncRNAs podem atuar de diferentes maneiras no ambiente celular: Atuando como guias; modificadores de estruturas da cromatina, por interação com proteínas associadas com DNA; como reguladores transcricionais, afetando interações da RNA polimerase e fatores de transcrição; recrutando múltiplas proteínas formando complexos ribonucleicos; e ligando-se a outras moléculas, impedindo que estas se liguem ao alvo original (FIGURA 1) (Morlando et al., 2015).

FIGURA 1: FUNÇÕES DOS LncRNAs NO NÚCLEO



I. lncRNA podem atuar como sinalizadores, participando da ativação gênica; II. como iscas, participando da supressão tumoral; III. podem atuar como guias e IV. *scaffolds* formando complexos ribonucleoproteicos agindo na estrutura da cromatina. Adaptado de K. C. Wang & Chang, 2012.

LncRNAs também têm um papel no controle do crescimento celular, principalmente através da regulação do ciclo celular e apoptose (Zhu et al., 2013). O lncRNA intergênico p21 (lincRNA-p21), ativado por p53, desempenha um papel importante na via de p53, levando à apoptose. A repressão transcricional por lincRNA-p21 é mediada através da ligação à ribonucleoproteína nuclear-K (hnRNP-K). Esta interação é necessária para direcionar o hnRNP-K para genes reprimidos (Huarte et al., 2011).

Também podem atuar no *imprinting* genômico (Kanduri, 2016) ou interferindo nas etapas de tradução e transcrição alterando a estabilidade genômica (Garzon et al., 2014). Uma grande proporção de genes de lncRNAs

está localizada em regiões de *imprinting*, como H19, XIST e MEG3 (X. Zhang et al., 2010).

No nível transcricional, os lncRNAs podem participar da transcrição direta, interagindo com o complexo transcricional ou com elementos de DNA, como promotores envolvidos na transcrição. Dessa forma podem estar envolvidos na modulação da estrutura de cromatina, recrutando enzimas modificadoras de cromatina, levando a expressão ou repressão de um grande número de genes (J. Chen et al., 2018). Por exemplo, o GAS5 (*Growth arrest-specific 5*) demonstrou competir com o Elemento de Resposta Glicocorticóide (GRE) no DNA, ocupando o domínio de ligação ao DNA do receptor glicocorticóide (GR) e impedindo o acesso do GR ao DNA alvo (Kino et al., 2010).

LncRNAs também são capazes de sequestrar ou alterar a localização celular de fatores de transcrição, competindo com o sítio de ligação ao DNA ou interferindo diretamente na atividade da RNA polimerase II (Morlando et al., 2015). Um estudo recente mostrou que a ativação dos retroelementos *Long Interspersed Element – 1* (LINE-1) leva a um aumento na expressão de *IFN β* e *ISGs* (*IFN-Stimulated Genes*) (Valadkhan & Gunawardane, 2020), porém o mecanismo pelo qual a replicação do LINE-1 leva a ativação da resposta IFN ainda não está clara.

Os lncRNAs também são capazes de atuar em alterações da expressão gênica interferindo no recrutamento ou na atividade de DNA metil-transferases modificando padrões de metilação de dinucleotídeos CpG (L. Wang et al., 2017). HOTAIR é um lncRNA localizado no locus HOXC no cromossomo 12. Ele funciona no recrutamento e ligação de complexos remodeladores de cromatina para o locus *HOXD* no cromossomo 2, resultando no direcionamento da PRC2 (*Polycomb Repressive Complex 2*), causando o silenciamento transcricional de uma região de cerca de 40 Kb de *HOXD* (Rinn et al., 2007). A expressão aumentada do HOTAIR tem mostrado induzir o redirecionamento de PRC2, levando a alteração da expressão gênica e está envolvido na ocorrência de metástase no câncer (Gupta et al., 2011).

A atuação dos lncRNAs pode ser em *cis*, interferindo na regulação da expressão de genes próximos ao seu local de síntese, ou em *trans*, regulando a expressão de genes localizados em outras regiões do genoma (Ferrè et al., 2016).

O *splicing* alternativo é um mecanismo importante para a diversidade genética atuando diretamente na regulação da expressão gênica. Os fatores de junção trabalham juntos para selecionar locais de *splicing* corretos em uma determinada condição (Pan et al., 2008). Os lncRNAs podem cooperar com ribonucleoproteínas nucleares heterogêneas (hnRNPs), como o hnRNP A1 (Kashima et al., 2007) e o hnRNPPC (Zarnack et al., 2013) para facilitar a sua ligação aos elementos silenciadores de *splicing* exônico (ESS) ou silenciadores de *splicing* intrônico (ISS) para *splicing* alternativo. O lncRNA MALAT1 pode regular o *splicing* alternativo de muitos genes, interagindo com as proteínas serina-arginina (SR) e influenciando sua localização subnuclear (Python.org, 2009). Além da sua influência no núcleo da célula, os lncRNAs também podem atuar no citoplasma e alterar a atividade e quantidade de proteínas e RNAs mensageiros (mRNA) (Smialek et al., 2020b). Eles podem atuar inibindo ou ativando a tradução, sequestrando proteínas que participam da regulação gênica (M. Zhang et al., 2017). Também podem atuar como competidores endógenos de microRNAs (miRNA) impedindo a repressão da tradução de mRNAs alvos por estes miRNAs (Y. C. Wang et al., 2019). Podem regular a montagem e função de corpos nucleares sem membrana, alterar a estabilidade e a tradução de mRNAs citoplasmáticos e interferir nas vias de sinalização (Statello et al., 2021).

No citoplasma os lncRNA provavelmente passam por processos de classificação específicos que atribuem diferentes lncRNAs a organelas específicas ou são distribuídos e se associam a diversas proteínas de ligação a RNA (RBPs) (Statello et al., 2021). Análises de transcriptoma mitocondriais humanos mostraram que lncRNAs exportados do núcleo podem ser classificados em mitocôndrias. O componente de RNA da endoribonuclease de processamento de RNA mitocondrial (RMRP) está associado ao RBPHur no núcleo e exportado para o citosol pela exportina 1. Assim que o RMRP chega à mitocôndria, é ligado e estabilizado pela sequência de RNA rica em G fator de ligação 1 (GRSF1), permitindo assim o seu acúmulo na matriz mitocondrial (Mercer et al., 2011).

Diversos lncRNA também são encontrados em condensados nucleares e são essenciais para a montagem e função destes (Banani et al., 2017). O transcrito 1 da montagem do condensado nuclear lncRNA (NEAT1) é a base da complexa organização e função dos condensados nucleares. O gene NEAT1

produz duas isoformas que compartilham uma extremidade 5' comum, mas têm extremidades 3' diferentes: NEAT1 curto, que tem uma cauda poli (A) produzida a partir de um sinal de poliadenilação a montante; e NEAT1 longo, que tem uma estável estrutura em hélice tripla que é clivado pela RNAase P. O NEAT1 longo é essencial na montagem do condensado nuclear (Hutchinson et al., 2007).

Estas moléculas apresentam uma grande versatilidade e complexidade de funções (Ferrè et al., 2016), e muitas delas apresentam um grande número de sítios ligantes de proteínas permitindo que atuem como plataformas para a formação de complexos ribonucleoproteicos e neste sentido, atuarem como reguladores da expressão gênica (Lu & Hall, 2011a). Um exemplo disso é o lncRNA NORAD, altamente conservado e abundante, com níveis de expressão de aproximadamente 500 a 1.000 cópias por célula. Sabemos que este lncRNA mantém a estabilidade genômica sequestrando proteínas PUMILIO impedindo que estas interfiram na estabilidade e tradução de mRNA alvos. A inativação do NORAD desencadeia uma dramática aneuploidia em linhagens celulares que antes estavam cariotipicamente normais. Na ausência do NORAD, a hiperatividade da PUMILIO reprime fatores mitóticos, de reparo do DNA e de replicação (FIGURA 2) (S. Lee et al., 2016).

FIGURA 2: RELAÇÃO ENTRE O lncRNA NORAD E A PROTEÍNA PUMILIO.

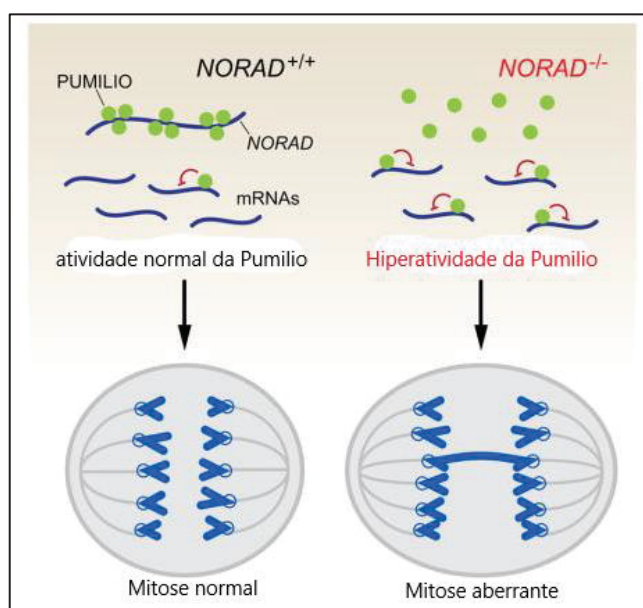


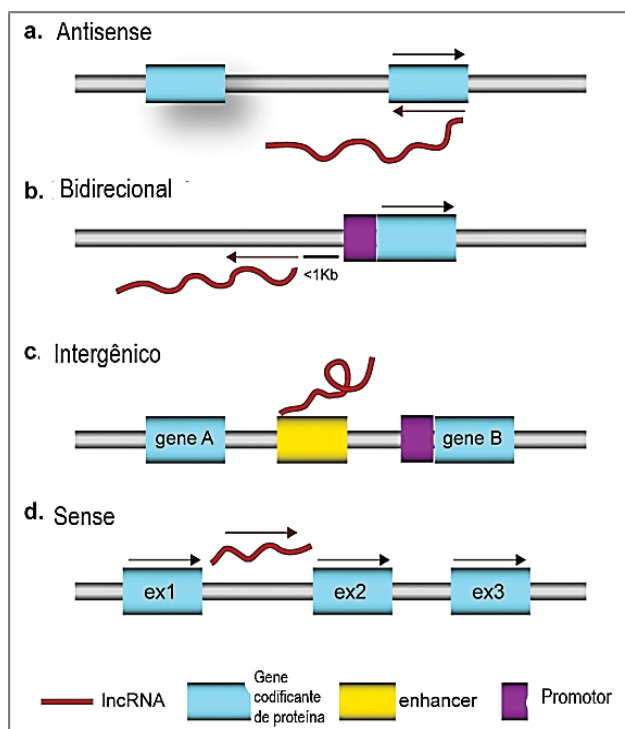
Figura esquemática mostrando a participação do lncRNAs NORAD na manutenção dos níveis normais da atividade de PUMILIO. Adaptado de S. Lee et al., 2016.

LncRNAs também têm participação na regulação da tradução. O lncRNA p21, por exemplo, serve como inibidor da tradução de junB e β -catenina de maneira dependente de HuR (Yoon et al., 2012).

2.1.2 Classificação

A classificação dos lncRNAs baseia-se na sua posição genômica em relação a um gene codificante de proteína (Ma et al., 2013). Como ilustrado na FIGURA 3, eles podem ser classificados como: a) lncRNA *antisense* – transcrito na fita oposta ao gene codificante de proteína; b) lncRNA bidirecional – lncRNA e o gene são expressos juntos, compartilhando a mesma região promotora, porém estão em fitas opostas; c) lncRNA intergênico (lincRNA) – localizado entre dois genes codificantes de proteína; d) lncRNA *sense* – transcrito da mesma cadeia do gene codificante de proteína (Jarroux et al., 2017).

FIGURA 3: CLASSIFICAÇÃO DOS LncRNA.



A posição genômica do lncRNA em relação a um gene codificador de proteína determina sua classificação em *antisense* (a), *bidirecional* (b), *intergênico* (c) ou *sense* (d). Adaptado de LANZAFAME ET AL., 2018

O sequenciamento completo do transcriptoma (RNA-seq) demonstrou que 81% dos lncRNAs não têm sua sequência de DNA conservada porém a sua região promotora é mais conservada do que a dos genes codificantes de proteínas sugerindo uma grande importância na regulação da expressão gênica (Charles Richard & Eichhorn, 2018).

2.1.3 LncRNAs e o Câncer

Diversos estudos mostram que os lncRNAs parecem apresentar um significado biológico em neoplasias, sendo associados com os processos de carcinogênese, progressão tumoral, metástase, proliferação celular, evasão da apoptose e o desenvolvimento de resistência a quimioterápicos (FIGURA 4) (Bhan et al., 2017)(DE OLIVEIRA; OLIVEIRA; MATHIAS; PEDROSO *et al.*, 2019).

FIGURA 4: LncRNAs ASSOCIADOS À CARCINOGENESE.

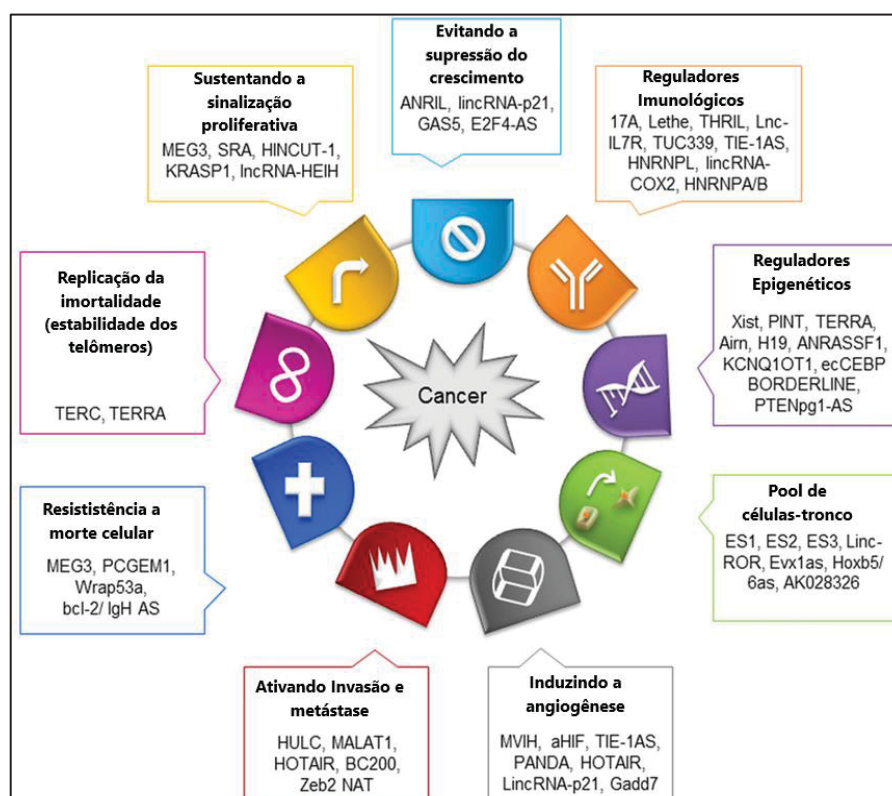


Imagem representativa do envolvimento de alguns lncRNAs nos *hallmarks* do câncer.
adaptado de Parasramka et al., 2016.

Entre os mais bem estudados, o lncRNA MALAT-1 (*Metastasis – Associated Lung Adenocarcinoma Transcript 1*), localizado no cromossomo 11 aparece relacionado ao câncer de pulmão (Schmidt et al., 2014) e também já foi observado super expresso em outros tipos de câncer, como o câncer de mama (Zeng et al., 2018) e câncer cervical (Y. Zhang et al., 2015). A super expressão do MALAT-1 em tumores de mama está correlacionada aos processos de metástase, proliferação celular e migração (Kim et al., 2019). O lncRNA SNHG15, localizado no cromossomo 7, foi relacionado com diversos cânceres incluindo câncer gástrico, hepatocelular, pancreático, colorretal, mama, tireoide, glioma, osteossarcoma, entre outros. Seu papel é crítico na manutenção da regulação da proliferação celular, migração e invasão de tumores. A sua super expressão está associada à menor sobrevida (Tong et al., 2019).

O lncRNA H19, localizado no cromossomo 11 aparece relacionado com o câncer colorretal, sua regulação positiva foi confirmada em um estudo comparando amostras de pacientes com câncer colorretal em diferentes estágios de metástase de linfonodos. O lncRNA também foi associado à quantidade de células tronco do câncer colorretal (Ren et al., 2018)

O lncRNA HCG11 (*HLA Complex Group 11*) apresenta expressão alterada em diferentes neoplasias como carcinoma hepatocelular, câncer de próstata, tumores de pulmão de células escamosas e melanomas cutâneos (Y. C. Wang et al., 2019). Para o câncer de mama, dados obtidos do *The Cancer Genoma Atlas (TCGA)*, mostram que a maior expressão de lncRNA HCG11 está correlacionada com um pior prognóstico e menor tempo de sobrevida, sendo observada uma diferença de expressão do lncRNA HCG11 entre os subtipos tumorais Luminal A e Triplo negativo (PEDROSO et al., 2018).

A proteína p53 é um supressor tumoral conhecido e frequentemente encontra-se mutado no câncer. p53 regula um grande conjunto de genes envolvidos no ciclo celular, apoptose e reparo do DNA e seu nível celular deve estar delicadamente equilibrado (A. Zhang et al., 2013). Estudos demonstram que um grande número de lncRNAs atuam como efetores de p53, participando de eventos a jusante de maneira dependente de p53. O nível da proteína p53 é baixo em condições fisiológicas normais, porém, após danos no DNA, a sua expressão é consideravelmente aumentada (Schmitt et al., 2017). O lncRNA

NORAD é induzido após dano ao DNA dependente de p53, apesar de não ser possível identificar um local óbvio de ligação à p53 nas proximidades do promotor do NORAD, estudos indicam a regulação indireta do NORAD pela p53 (Lee et al., 2016).

2.2 Proteínas de Ligação ao RNA

A regulação da expressão gênica muitas vezes é mediada pela ação de proteínas de ligação ao RNA (RBPs do inglês “*RNA binding proteins*”). Elas podem reconhecer e se ligar a moléculas de RNA específicas e determinar se uma molécula de RNA será degradada, levada à tradução ou estocada para uso posterior. Com efeito, as RBPs atuam como interruptores que ativam ou desativam a produção de proteínas (Goldstrohm et al., 2018).

As RBPs possuem domínios de ligação ao RNA formando complexos ribonucleoproteicos (RNPs). O complexo RNP pode ser transiente, e a RBP participa na função de maturação, localização ou decaimento do RNA. A dinâmica desses complexos depende de sua composição, alvos e cofatores (Gerstberger et al., 2014).

Estudos demonstram que RBPs individuais são capazes de coordenar a tradução ou repressão de mRNAs específicos, interagindo com motivos e elementos estruturais definidos em regiões não traduzidas no 5' UTRs ou 3' UTRs ou em regiões codificantes, direcionando programas de expressão gênica citoplasmática que são obrigados a responder a mudanças nas condições externas (King et al., 2014).

As técnicas para identificação de RBPs podem ser amplamente divididas em duas categorias: (a) RNA-centralizado, incluindo a captura de proteoma de ligação a RNA (RBPome) e métodos de afinidade de RNA para identificar proteínas que se ligam ao RNA celular; e (b) métodos centrados em proteínas, como RIP e CLIP, para estabelecer todos os RNAs ligados a uma única proteína (Faoro & Ataíde, 2014).

2.3 PUMILIO

2.3.1 Estrutura

As proteínas PUMILIO são membros da família PUF de proteínas de ligação a RNA (RBPs) que apresentam um domínio de reconhecimento de sequência específico e altamente conservado denominado Domínio de Homologia de PUMILIO (Pum-HD) (Lu & Hall, 2011a, Dai et al., 2019). Este domínio compreende oito repetições estruturais de um motivo de aproximadamente 36 aminoácidos formando três α -hélices (FIGURA 5) (S. Lee et al., 2016).

Foram primeiramente estudadas em *Drosophila*, onde percebeu-se um padrão repressor essencial no desenvolvimento embrionário agindo sobre o gene *hunchback* (hb) controlando a embriogênese (Goldstrohm et al., 2018).

Na altura da hélice 2 de cada repetição, na face interna do arco, estão presentes aminoácidos com características aromáticas e básicas que interagem com o RNA. A interação ocorre na extremidade 3'UTR de RNAs alvos com sequência consenso de 8 nucleotídeos (UGUANAUA) conhecida como Elemento de Resposta a PUMILIO (PRE, do inglês, *Pumilio Response Element*) (FIGURA 5) (Goldstrohm et al., 2018).

FIGURA 5: DOMÍNIO PUF DE LIGAÇÃO DA PUMILIO (PUM-HD)



O domínio PUF contém oito repetições seguidas que juntas formam um único domínio capaz de associar-se a uma sequência de RNA específica. Adaptado de X. Wang et al., 2002.

Apesar da região C-terminal ser bastante conservada entre as espécies, a porção N-terminal de sua cadeia não é podendo, no entanto, apresentar regiões regulatórias, incluindo motivos para glutamina e aspargina que facilitam a agregação de moléculas (Dai et al., 2019).

Os mamíferos possuem duas proteínas Pumilio, PUM1 e PUM2, ambas intimamente relacionadas com 91% de similaridade em relação ao domínio de ligação ao RNA (Bohn et al., 2018).

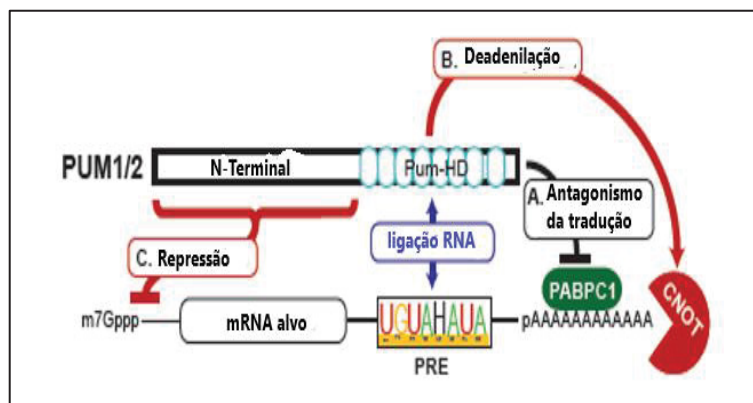
Proteínas Pumilio estão envolvidas em diversas funções na regulação da expressão gênica atuando em processos relacionados com diferenciação celular, fertilidade, funcionamento do sistema nervoso; e a sua disfunção está relacionada com doenças neurológicas e câncer (Goldstrohm et al., 2018).

Em *Drosophila*, Pumilio forma um complexo com a proteína Nanos (Nos), sendo ambos necessários para reprimir a diferenciação celular, indicando que a autorrenovação das células-tronco corresponde em parte à repressão do programa de diferenciação mediada pelo complexo Pum-Nos (Joly et al., 2013). No embrião, Nos e Pum atuam por dois mecanismos: inibição do início da tradução e recrutamento do complexo de deadenilação CCR4-NOT (Temme et al., 2010). O recrutamento do complexo CNOT é um mecanismo conservado de repressão da Pumilio de leveduras a seres humanos, interagindo diretamente com a subunidade enzimática deadenilase Pop2/Caf1/CNOT7/CNOT8 (Joly et al., 2013). Ainda em humanos, as proteínas PUMILIO foram associadas a processos de manutenção da pluripotência celular em modelos de diferenciação adipogênica (SHIGUNOV; SOTELO-SILVEIRA; KULIGOVSKI; DE AGUIAR *et al.*, 2012) e cardiomiogênica (SILVA; ROBERT; CABO; SPANGENBERG *et al.*, 2020).

A estabilidade e a tradução dos mRNAs são processos interligados. A tradução é promovida por duas modificações chave no mRNA, o cap5' e adição de cauda poli A (3'UTR), que são ligadas por eIF4E e proteínas de ligação de poli(A)(PABP), respectivamente. O cap 5' e a cauda poli A também protegem o RNA e sua remoção inicia a degradação do mRNA. Estudos sugerem que as proteínas PUMILIO reprimem o mRNA promovendo a deadenilação da cauda poli A através do complexo CNOT, mas também podem atuar de forma antagônica à proteína PABPC1 reprimindo o início da tradução, impedindo a

ligação dos fatores de início da tradução eIF4E e eIF5B (FIGURA 6) (Goldstrohm et al., 2018).

FIGURA 6 – MECANISMO DE AÇÃO DA PUMILIO



Ao associar-se a um mRNA alvo, as proteínas PUMILIO podem promover a perda de estabilidade deste RNA por dois mecanismos a) Deadenilação recrutando o complexo CNOT; b) Atuando como antagonista da proteína PABPC1, reprimindo a tradução. Adaptado de Goldstrohm et al., 2018.

3. Justificativa

Em humanos, vários mRNAs envolvidos com doenças neurológicas, câncer e doenças cardiovasculares são alvos de PUMILIO, sugerindo a sua influência na regulação genética destas doenças (Bohn JA. et al., 2017).

No sentido de controlar o ciclo celular, as proteínas PUMILIO aparecem associadas ao lncRNA NORAD (RNA longo não codificante ativado por danos no DNA), preservando a estabilidade do genoma. NORAD atua como uma plataforma de ligação molecular para as proteínas PUMILIO pela presença de vários sítios de ligação à proteína. A elucidação da interação reguladora do NORAD e da proteína PUMILIO expandiu a compreensão dos lncRNAs e das suas funções.

Sabendo da grande quantidade de mRNAs com sítios para PUMILIO e do número cada vez maior de lncRNAs descritos, acreditamos que muitos deles podem, assim como o NORAD, apresentar potenciais sítios de interação à PUMILIO, podendo atuar em um eixo regulatório promovendo o fenótipo do câncer.

Assim, a identificação destes lncRNAs bem como os genes alvos de PUMILIO, co-expressos pode levar a uma maior compreensão do papel destas moléculas no câncer.

4. Objetivos

4.1 Objetivo Geral

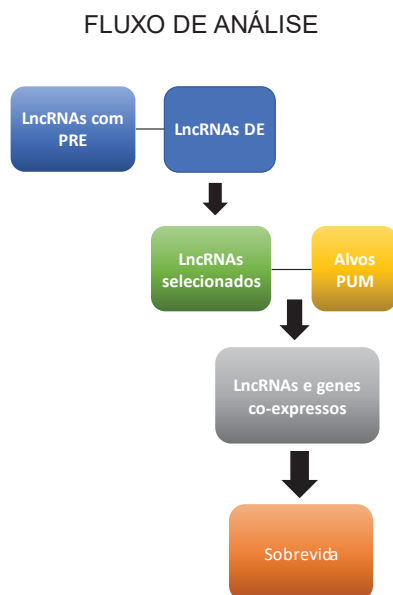
Investigar o conjunto de lncRNAs que apresentam múltiplos sítios de ligação para PUMILIO, analisando o seu nível de expressão no contexto de fenótipos normais ou de câncer em 9 tipos de câncer analisados no Projeto Pan-Câncer procurando por genes co-expressos que possivelmente estão participando de mecanismos de desregulação celular durante a tumorigênese mediada pela PUMILIO.

4.2 Objetivos específicos

- Selecionar RNA longos não codificantes com sítios para PUMILIO.
- Identificar entre estes lncRNAs, quais deles apresentam expressão diferencial positiva entre amostras tumorais e não tumorais em 9 tipos de câncer: Carcinoma urotelial de bexiga (BLCA), Adenocarcinoma do cólon (COAD), Carcinoma Esofágico (ESCA), Carcinoma das células escamosas da cabeça e do pescoço (HNSC), Carcinoma de células renais claras (KIRC), Carcinoma hepatocelular do fígado (LIHC), Adenocarcinoma pulmonar (LUAD), Adenocarcinoma da próstata (PRAD), Carcinoma invasivo da mama (BRCA).
- Verificar quais os genes alvo de PUMILIO apresentam-se co-expressos com os lncRNAs com expressão positiva alterada em tumores.
- Avaliar o impacto dos genes selecionados com valores de co-expressão mais elevados na sobrevida dos pacientes.

5. Materiais e Métodos

O fluxo de análise dos dados segue o esquema abaixo:



5.1 Seleção dos lncRNAs com sítios para PUMILIO

Para a seleção dos RNAs longos não codificantes (lncRNA) com sítios para PUMILIO foram utilizados os dados já publicados dos artigos Smialek et al., 2020 que usaram abordagens RIP-Seq e RNA-Seq identificando pools de mRNA regulados por PUM1 e PUM2 na linha celular TCam-2, um modelo de célula germinativa masculina humana e; Bohn et al., 2018 que, a partir de dados experimentais, utilizando análises RNA-Seq, testes estatísticos rigorosos e um corte de alteração de dobra derivado experimentalmente validaram um conjunto de RNAs com sítios de ligação para PUM (UGUANAUA) em células HEK293 humanas. A fim de selecionar somente os lncRNAs, utilizou-se a ferramenta *Biomart* (<https://www.ensembl.org/biomart/martview>) que permite, a partir do nome do gene, determinar sua classificação e *Ensemble ID*.

5.2 Análises de expressão diferencial para lncRNA

Os dados de RNA-seq dos 9 tipos de câncer analisados, foram extraídos utilizando o pacote *GDCRNATools* (Li et al., 2018) do banco de dados *The*

Cancer Genome Atlas (TCGA) seguindo os critérios éticos, legais e políticos do programa. Para cada tipo de câncer foram utilizados os *datasets Gene expression. RNA-Seq/HTSeq-FPKM*. As quantidades de amostras de cada tipo de tumor estão relacionadas na TABELA 1. A análise de expressão diferencial foi feita nos tipos tumorais, utilizando o mesmo pacote e a *pipeline* contida no pacote *limma* (Ritchie et al., 2015).

TABELA 1: TIPOS DE CÂNCER ANALISADOS E QUANTIDADE DE AMOSTRAS

Câncer	Número de Amostras	
	Tumor Primário	Tecido normal
Adenocarcinoma do cólon (COAD)	448	90
Carcinoma Urotelial da Bexiga (BLCA)	412	37
Carcinoma esofágico (ESCA)	180	64
Carcinoma das células da cabeça e do pescoço (HNSC)	527	82
Carcinoma de células renais claras (KIRC)	483	394
Carcinoma Hepatocelular do fígado (LIHC)	375	88
Adenocarcinoma da próstata (PRAD)	467	112
Carcinoma invasivo de mama (BRCA)	1091	113
Adenocarcinoma Pulmonar (LUAD)	564	259

Para cada tipo de câncer, foram confrontadas as tabelas de lncRNAs diferencialmente expressos (DE) no projeto Pan Câncer (Hoadley et al., 2018) ($p < 0,05$) e os lncRNAs com sítios para PUMILIO selecionados anteriormente. Foi utilizada a ferramenta *Interactiveen*, disponível por BMC *Bioinformatics* 16:169 (2015) (Heberle et al., 2015). Esta ferramenta permite a união de conjuntos em diagramas de *Venn* e a construção do diagrama possibilita a análise das uniões definidas, buscando elementos em comum. Ao confrontar o conjunto dos lncRNA DE em cada grupo de câncer (BLCA, COAD, ESCA, HNSC, KIRC, LIHC, LUAD e PRAD, BRCA) com os lncRNAs com sítios para PUMILIO a ferramenta mostra os lncRNA comuns. Como os tumores de mama podem ser subdivididos em subtipos com base em alterações moleculares distintas (Yeo & Guan, 2017) e, em estudos anteriores o lncRNA NORAD apresentou uma diferença de expressão entre os subtipos Luminal A e Basal-like (Mathias et al., 2021), o

BRCA foi analisado por sua classificação em subtipos moleculares: Luminal A (LA), Luminal B (LB), HER2- *enriched* e Basal-*like* (Sorlie et al., 2001).

Considerando o objetivo do projeto que é encontrar lncRNAs com expressão aumentada capazes ligar-se a PUMILIO, interferindo em sua capacidade de ligar-se aos sítios alvos, reprimindo-os, foi aplicado mais um filtro selecionando somente lncRNAs DE com $\log FC > 0$ nos cânceres BLCA, COAD, ESCA, HNSC, KIRC, LIHC, LUAD e PRAD. Para o BRCA foram selecionados os lncRNAs que apresentaram $\log FC > 0$ em pelo menos um subtipo. Após este filtro, a quantidade de lncRNAs para o BRCA ainda era muito grande e, por tratar-se do tipo de câncer com mais amostras analisadas no banco de dados TANRIC (<https://ibl.mdanderson.org/tanric/design/basic/main.html>), decidimos aplicar mais um filtro selecionando os lncRNAs que apresentaram expressão detectável em pelo menos 80% das amostras.

5.3 Análise de co-expressão

Para as análises de co-expressão foram consultados os dados disponíveis no TCGA e confrontados os lncRNAs selecionados com sítios para PUMILIO entre tecido tumoral e não tumoral de cada tipo de câncer (BLCA, COAD, ESCA, HNSC, KIRC, LIHC, LUAD, PRAD, BRCA) contra todo transcriptoma (60.483 genes) a fim de encontrar genes que estão co-expressos com estes lncRNAs. Foi utilizado o pacote RTN (Castro et al., 2015) para a reconstrução de redes regulatórias transcricionais (TRNs) e análise de regulons usando informação mútua (MI). Para gerar regulons sob condições estatísticas semelhantes para coortes diferentes entre os tipos de cânceres utilizamos a função de ajuste (p-value Cutoff) do pacote RTN. Para cada câncer o valor de p foi ajustado de acordo com o tamanho da amostra utilizando a fórmula:

$$\alpha_B \leftarrow \text{tni.alpha.adjust} (n_B = 100, n_A = 300, \alpha_A = 1e-5, \beta_A = 0.2)$$

onde n_A representa a maior coorte entre os tipos de câncer analisados e n_B as demais coortes. Foi testada a associação entre um determinado lncRNA e todos os alvos potenciais usando dados de transcriptoma. Primeiro foi calculado o valor absoluto de MI entre um lncRNA e todos os alvos potenciais (60.483 genes). O

valor absoluto de MI representa a quantidade que mede o quanto uma variável aleatória nos diz sobre outra. É representada por uma quantidade adimensional com unidades em bits e pode ser considerada como a redução de incerteza sobre uma variável aleatória, dado o conhecimento de outra, removendo associações não significativas por análises de 1000 permutações inferindo um p-valor ajustado. Neste estágio, cada alvo no TRN pode ser vinculado a vários lncRNAs. Como a regulação pode ocorrer por interações diretas (alvo – lncRNA) e indiretas (alvo – lncRNA – lncRNA) o pacote aplica um filtro (função *bootstrap*) para remover interações fracas em qualquer triplete formado por dois lncRNAs e um gene alvo comum, preservando o par lncRNA – gene alvo dominante.

Por fim, de todos os genes selecionados para cada lncRNA com sítio para PUMILIO, foram selecionados apenas aqueles que também são alvos de PUMILIO consultados nos artigos de Smialek et al., 2020 e Bohn et al., 2018. Foram escolhidos os pares lncRNA – gene alvo com valores de co-expressão > 0,15.

5.4 Análise de expressão diferencial para os alvos selecionados com sítios para PUMILIO

Para as análises de expressão diferencial de cada gene alvo da PUMILIO, consultamos o banco de dados disponibilizados pela plataforma GEPIA (<http://gepia2.cancer-pku.cn>). Decidimos avaliar apenas os alvos com valor de co-expressão > 0,15 interpretando uma correlação acima de 10% entre o lncRNA e o alvo. Cada alvo foi analisado quanto a sua expressão no tecido tumoral, comparado com o tecido normal, usando apenas dados do TCGA. Selecionamos os genes diferentemente expressos positivamente ($p > 0,05$) com valor logFC > 0.5.

5.5 Análise de Sobrevida Global e Livre da Doença.

Analizamos os lncRNA e seus genes co-expressos com os maiores valores de co-expressão. Selecionamos os top 20 genes co-expressos para cada tipo de câncer com valor significativo na Sobrevida Global e na Sobrevida Livre da Doença ($p < 0,05$).

Os dados clínicos foram obtidos no cBioPortal. O *dataset* utilizado foi o “Clinical Data” das coortes BLCA, COAD, ESCA, HNSC, KIRC, LIHC, LUAD, PRAD, BRCA (TCGA, PanCancer Atlas), disponível em <https://www.cbioportal.org>.

Os dados de expressão gênica foram obtidos no XenaBrowser, disponível em <https://xenabrowser.net>. O *dataset* utilizado foi o “gene expression RNAseq - HTSeq - FPKM”, das coortes “GDC TCGA BLCA, GDC TCGA COAD, GDC TCGA ESCA, GDC TCGA HNSC, GDC TCGA KIRC, GDC TCGA LIHC, GDC TCGA LUAD, GDC TCGA PRAD, GDC TCGA BRCA”.

A análise de regressão de riscos proporcional *Univariate Cox* foi realizada utilizando o pacote R “sobrevivência”, utilizando os pacotes “*survival* e *survminer*”. Um valor $p < 0,05$ foi considerado estatisticamente significativo. No caso de câncer de mama, a análise foi realizada por subtipo molecular: Luminal A, Luminal B, Her2-*enriched* e Basal-*like*.

5.6 Comparação dos lncRNA e Alvos co-expressos entre os 9 tipos de câncer

Para avaliar os lncRNAs e alvos co-expressos e diferencialmente expressos com Sobrevida Global e Livre da Doença significativos ($p < 0,05$) utilizamos a ferramenta *Interactiveen*, disponível por BMC *Bioinformatics* 16:169 (2015) (Heberle et al., 2015). Comparamos os lncRNAs e seus respectivos alvos co-expressos nos tipos de câncer BLCA, HNSC, ESCA, COAD, KIRC, LIHC, LUAD e PRAD, e selecionamos o gene alvo de PUM e seu respectivo lncRNA mais frequente entre os tipos de câncer com o maior valor de co-expressão. No BRCA selecionamos o lncRNA e seu alvo co-expresso comum nos quatros subtipos Luminal A, Luminal B, HER2-*enriched* e Basal-*like*.

6. Resultados

6.1 Seleção dos lncRNA

A fim de selecionar somente os lncRNAs, presentes nos trabalhos de Bohn et al., 2018 e Smialek et al., 2020, que relacionaram todos os genes com sítios de ligação para PUMILIO (UGUANAUA), foi utilizado a ferramenta *Biomart*

que permite, a partir do nome do gene, determinar sua classificação e identificação no banco de dados do *Ensembl ID* (<https://www.ensembl.org/index.html>). Entre os 8.248 RNAs identificados nos dois trabalhos, foram selecionados 49 lncRNAs (QUADRO 1).

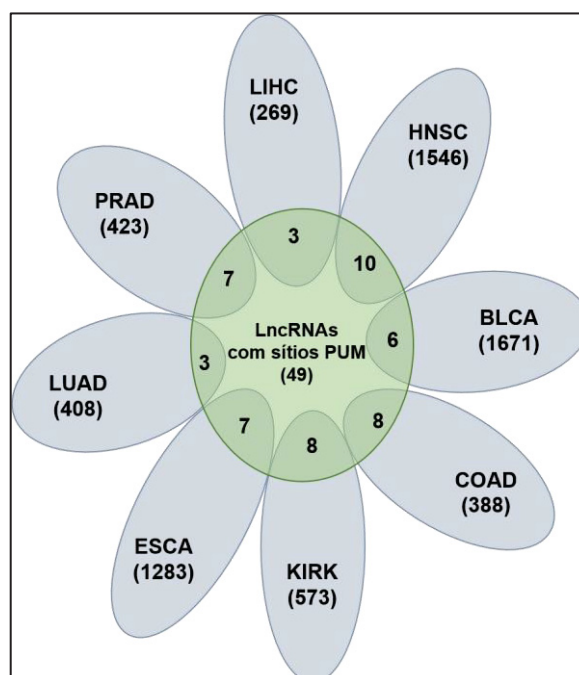
QUADRO 1: LncRNAs COM SÍTIOS PARA PUMILIO

<i>Ensembl ID</i>	<i>lncRNA</i>	<i>Nº sítios</i>
ENSG00000248360	LINC00504	71
ENSG00000248874	C5orf17	70
ENSG00000241316	SUCLG2-AS1	52
ENSG00000228223	HCG11	19
ENSG00000245573	BDNF-AS	18
ENSG00000260032	NORAD	17
ENSG00000249348	UGDH-AS1	11
ENSG00000251562	MALAT1	4
ENSG00000180861	LINC01559	4
ENSG00000186960	LINC01551	3
ENSG00000175746	C15orf54	3
ENSG00000280109	PLAC4	3
ENSG00000224281	SLC25A5-AS1	2
ENSG00000236333	TRHDE-AS1	2
ENSG00000257242	LINC01619	2
ENSG00000248801	C8orf34-AS1	2
ENSG00000176659	C20orf197	2
ENSG00000231889	TRAF3IP2-AS1	2
ENSG00000224687	RASAL2-AS1	2
ENSG00000198054	DSCR8	2
ENSG00000262117	BCAR4	2
ENSG00000246145	RR51-AS1	1
ENSG00000229153	EPHA1-AS1	1
ENSG00000255559	ZNF252P-AS1	1
ENSG00000196243	LINC00615	1
ENSG00000163915	IGF2BP2-AS1	1
ENSG00000280963	SERTAD4-AS1	1
ENSG00000237290	LINC01343	1
ENSG00000224165	DNAJC27-AS1	1
ENSG00000235919	ASH1L-AS1	1
ENSG00000267858	MZF1-AS1	1
ENSG00000234492	RPL34-AS1	1
ENSG00000203808	BVE5-AS1	1
ENSG00000205086	C2orf91	1
ENSG00000261485	PAN3-AS1	1
ENSG00000225880	LINC00115	1
ENSG00000254389	RHPN1-AS1	1
ENSG00000214188	ST7-OT4	1
ENSG00000235665	LINC00298	1
ENSG00000215424	MCM3AP-AS1	1
ENSG00000234235	BOK-AS1	1
ENSG00000099869	IGF2-AS	1
ENSG00000238148	HCG27	1
ENSG00000204860	FAM201A	1
ENSG00000267586	LINC00907	1
ENSG00000244041	LINC01011	1
ENSG00000234617	SNRK-AS1	1
ENSG00000253686	LINC01484	1
ENSG00000226397	C12orf77	1

6.2 Análise de expressão diferencial para lncRNA

Para cada tipo de câncer, confrontamos as tabelas de lncRNAs diferencialmente expressos (DE) no projeto Pan Câncer (Hoadley et al., 2018) e os 49 lncRNAs com sítios para PUMILIO selecionados anteriormente (FIGURA 7).

FIGURA 7 - COMPARAÇÃO DOS lncRNAs COM SÍTIOS PARA PUMILIO E DIFERENCIALMENTE EXPRESSOS NOS CÂNCERES



A região de intersecção ilustra os lncRNAs que são simultaneamente diferencialmente expressos (em parênteses - p-valor < 0,05) em cada tipo de câncer BLCA, COAD, KIRC, LIHC, ESCA, HNSC, LUAD e PRAD e que apresentam sítios de ligação à PUMILIO (em parênteses no centro).

Destes, foram selecionamos somente os lncRNAs com sítios para PUMILIO diferencialmente expressos que apresentam expressão aumentada no tumor em relação às amostras normais ($\log FC > 0$) (TABELA 2).

TABELA 2: RELAÇÃO DOS lncRNAs SELECIONADOS EM CADA TIPO DE CÂNCER

Câncer	lncRNAs	logFC
BLCA	ZNF252P-AS1	2,34
	LINC00115	1,16
	RHPN1-AS1	1,73

	LINC01484	1,02
COAD	FAM201A	0,72
	RHPN1-AS1	1,67
	PAN3-AS1	0,67
	MALAT1	1,12
	MCM3AP-AS1	0,64
ESCA	C15orf54	1,57
HNSC	C15orf54	1,08
	LINC00115	1,06
	LINC01619	1,25
	IGF2-AS	1,50
	LINC01343	1,81
	IGF2BP2-AS1	2,29
	LINC01484	1,15
KIRC	LINC01011	0,69
	MALAT1	0,94
	HCG27	2,43
LIHC	MCM3AP-AS1	1,03
	MZF1-AS1	0,69
	MALAT1	0,82
LUAD	RHPN1-AS1	1,42
	MCM3AP-AS1	0,73
PRAD	C8orf34-AS1	1,27
	FAM201A	0,98
	RHPN1-AS1	0,84
	LINC00115	1,19
	ZNF252P-AS1	1,03
	HCG27	0,77

Como os tumores de mama podem ser subdivididos em subtipos com base em alterações moleculares distintas, o BRCA foi analisado por subtipo: Luminal A (LA), Luminal B (LB) *HER2-enriched* e *Basal-like*. No total, para os quatro subtipos foram encontrados 34 lncRNAs com sítios para PUMILIO diferencialmente expressos com diferenças no logFC por subtipo.

A partir dos 34 lncRNAs diferencialmente expressos para os subtipos do BRCA, Luminal A, Luminal B, *HER2-enriched* e *Basal-like*, selecionamos os lncRNA com sítios para PUMILIO e logFC > 0 em pelo menos um subtipo do BRCA, diferencialmente expressos nas amostras de pacientes do banco de dados Pan Câncer, resultando em 28 lncRNAs (TABELA 3).

TABELA 3: LncRNAs DE COM SÍTIOS PARA PUMILIO E logFC > 0 EM PELO MENOS UM SUBTIPO DE BRCA

lncRNA	Her2- <i>enriched</i>	Luminal A	Luminal B	Basal-like
RHPN1-AS1	1,69	1,18	1,80	1,59
LINC00504	0,21	1,27	1,27	--
ZNF252P-AS1	0,46	0,30	0,49	0,41
LINC00115	0,46	0,31	0,39	0,67
ASH1L-AS1	0,46	0,41	0,53	0,73
LINC00907	0,38	0,18	0,27	0,29
LINC01343	0,40	0,19	0,28	0,32
LINC00615	0,40	0,19	0,28	0,30
LINC00298	0,40	0,18	0,28	0,31
MZF1-AS1	0,46	0,30	0,36	0,21
RRS1-AS1	0,41	0,17	0,29	0,29
RPL34-AS1	0,33	0,16	0,25	0,24
LINC01011	0,49	0,30	0,42	0,24
BCAR4	0,53	0,31	0,47	0,51
FAM201A	0,10	0,30	0,52	0,40
NORAD	0,05	0,25	0,43	--
SLC25A5-AS1	0,34	0,16	0,26	0,20
LINC01484	0,19	--	--	0,06
EPHA1-AS1	0,38	0,08	0,15	0,22
RASAL2-AS1	0,12	--	--	0,36
BVES-AS1	0,14	--	0,007	0,23
MALAT1	--	0,07	0,32	--
SUCLG2-AS1	0,05	--	0,014	--
IGF2-AS	0,17	0,06	0,04	0,12
SNRK-AS1	0,21	0,04	0,15	0,13
UGDH-AS1	--	0,04	0,10	--
DNAJC27-AS1	0,16	--	0,09	--
HCG11*	-0,87	-1,04	-1,21	-0,31

* O lncRNA HCG11 apresenta logFC < 0 no BRCA mas foi incluído por tratar-se de um gene parálogo e com alta identidade com o lncRNA NORAD.

Dos 28 lncRNAs DE com sítios para PUMILIO, 10 apresentaram expressão detectável em pelo menos 80% das amostras das 942 pacientes diagnosticadas com câncer de mama no banco de dados TANRIC (<https://www.tanric.org/>), sendo assim selecionados para a continuidade das análises: ASH1L-AS1,

RHPN1-AS1, SERTAD4-AS1, NORAD, LINC01011, LINC00504, LINC00431, MALAT1, LINC01963 e HCG11.

6.3 Análise de co-expressão e expressão diferencial dos genes alvos

Decidimos escolher os valores de co-expressão $> 0,15$ por representar uma relação acima de 10% entre a dupla lncRNA e o gene alvo. Os valores positivos de co-expressão entre os lncRNA e alvos para PUMILIO foram observados nos 9 tipos de câncer analisados (p ajustado). Selecionamos apenas os genes que apresentaram sítios para PUMILIO e $\log FC > 0,5$, ou seja, todos os genes alvos diferencialmente expressos positivamente na comparação de tecido tumoral com tecido normal em cada tipo de câncer. A TABELA 4 resume os resultados da co-expressão para cada tipo de câncer e os dados detalhados podem ser consultados nas Tabelas Suplementares 1 a 9.

TABELA 4: LncRNAs DIFERENCIALMENTE EXPRESSOS (p-valor $< 0,05$) EM CADA TIPO DE CÂNCER E QUANTIDADES DE GENES CO-EXPRESSOS (p ajustado)*

Câncer	lncRNA	Genes co-expressos
BLCA (p=0,017)*	LINC00115	1
	MCM3AP-AS1	28
COAD (p=0,012)*	PAN3-AS1	17
	MALAT1	1
ESCA (p=0,2)*	C15orf54	8
	IGF2-AS	20
	IGF2BP2-AS1	1
HNSC (p=0,005)*	LINC01484	2
	LINC01619	2
	LINC01343	4
	LINC00115	28
	LINC01011	47
KIRC (p=0,01)*	MALAT1	180
	HCG27	415
LIHC (p=0,025)*	MCM3AP-AS1	137
LUAD (p=0,003)*	MCM3A-AS1	47
PRAD (p=0,01)*	LINC00115	41
	C8orf34-AS1	1

RHPN1-AS1	1
ZNF252-AS1	3

*Em parênteses o p-valor calculado para cada coorte de câncer de acordo com a fórmula do RTN ($\alpha_B \leftarrow \text{tni.alpha.adjust}(n_B = 100, n_A = 300, \alpha_A = 1e-5, \beta_A = 0.2)$) nas análises de co-expressão.

Para o câncer de mama, as análises foram realizadas considerando cada subtipo separadamente (TABELA 5)

TABELA 5: RELAÇÃO DE CO-EXPRESSÃO ENTRE lncRNAs COM SÍTIOS PARA PUMILIO E GENES ALVO DA PUMILIO NO BRCA ($p < 0,00001$)

Subtipo BRCA	lncRNA	Gene Alvo PUM	Co-expressão
LumA	NORAD	<i>RALGAPB</i>	0,15
LumB	NORAD	<i>RALGAPB</i>	0,20
		<i>NCOA6</i>	0,17
		<i>ZNF7</i>	0,20
Her2-enriched	NORAD	<i>RALGAPB</i>	0,16
		<i>ELMO2</i>	0,16
Basal-like	ASHIL-AS1	<i>USP21</i>	0,15
	RHPN1-AS1	<i>ZNF7</i>	0,18
	MALAT*	<i>NKTR*</i>	0,21
	HCG11*	<i>ZNF322*</i>	0,55
		<i>MAPK14*</i>	0,31
		<i>KCTD20*</i>	0,30
		<i>PTPN11*</i>	0,30
		<i>ROCK1*</i>	0,27
		<i>PHF3*</i>	0,27
		<i>GPATCH8*</i>	0,25
		<i>CCP110*</i>	0,24
		<i>STAG1*</i>	0,24
		<i>ASHIL*</i>	0,22
		<i>EPC2*</i>	0,22
		<i>RALGAPB*</i>	0,55
		<i>KIF3B*</i>	0,31
		<i>RBM39*</i>	0,30
		<i>ZC3H11A*</i>	0,26
		<i>CREB1*</i>	0,23
		<i>RANBP2*</i>	0,22
		<i>NCOA3*</i>	0,21
		<i>NIPBL*</i>	0,21
	NORAD*		

*Os lncRNAs HCG11, NORAD e MALAT-1 apresentam uma expressão diferencial negativa (estão menos expressos) no subtipo Basal, assim selecionamos os genes alvos que também se apresentaram sub-expressos ($\log FC < 0$).

6.4 Análises de Sobrevida Global e Sobrevida Livre da Doença

Analizamos a Sobrevida Global e Sobrevida Livre da Doença para os top 20 pares lncRNAs e genes co-expressos com sítios para PUMILIO e diferencialmente expressos. Para os tipos de câncer BLCA e ESCA não foram encontrados valores significativos para Sobrevida Global ou Sobrevida Livre da Doença. Um valor $p < 0,05$ foi considerado estatisticamente significativo quando encontrado no gene alvo da PUMILIO. Nem todos os pares apresentaram Sobrevidas significativas nos dois contextos. As TABELAS 6 e 7 apresentam os resultados de Sobrevida Global e Sobrevida Livre da Doença, respectivamente, com resultados estatisticamente significativos.

TABELA 6: LncRNAs E GENES CO-EXPRESSOS COM SOBREVIDA GLOBAL (p-valor) PARA OS CÂNCERES HNSC, KIRC, LIHC E PRAD

CÂNCER	lncRNA	GENE CO-EXPRESSO
HNSC	IGF2-AS	<i>P4HA3</i> (p=0,028)
	IGF2BP2-AS1 (p=0,0058)	<i>IGF2BP2</i> (p=0,015)
KIRC	LINC01011 (p = 0,0073)	<i>CLK2</i> (p = 0,0013)
		<i>PPFIA4</i> (p = 0,029)
		<i>AGER</i> (p < 0,0001)
		<i>PRPF3</i> (p = 0,016)
		<i>MSH5</i> (p = 0,00096)
		<i>ANKZF1</i> (p = 0,0039)
		<i>WDR90</i> (p = 0,036)
		<i>CCNL2</i> (p < 0,0001)
		<i>ANKRD10</i> (p = 0,00079)
		<i>PARP6</i> (p = 0,0073)
		<i>KIAA0895L</i> (p = 0,027)
		<i>SIRT7</i> (p < 0,0001)
		<i>AMY2B</i> (p = 0,011)
		<i>SFSWAP</i> (p = 0,0078)
		<i>DMTF1</i> (p = 0,0037)
		<i>MAPK8IP3</i> (p < 0,0001)

	<i>RUSC1-AS1</i> (p = 0,00014) <i>ZNF276</i> (p = 0,00002) <i>INTS6L</i> (p = 0,0011)
MALAT1 (p = 0,0099)	<i>RUSC1-AS1</i> (p = 0,00014) <i>NKTR</i> (p = 0,0052) <i>TTLL3</i> (p < 0,0001) <i>LENG8</i> (p < 0,0001) <i>CSAD</i> (p < 0,0001) <i>LUC7L</i> (p < 0,0001) <i>AMY2B</i> (p = 0,011) <i>MSH5</i> (p = 0,00096) <i>MDM4</i> (p = 0,0074) <i>CCNL2</i> (p < 0,0001) <i>RSRP1</i> (p = 0,0019) <i>DMTF1</i> (p = 0,0037) <i>RPL32P3</i> (p = 0,00015) <i>MAPK8IP3</i> (p < 0,0001) <i>FNBP4</i> (p = 0,00043) <i>RBM33</i> (p = 0,0034) <i>PABPC1L</i> (p < 0,0001) <i>KIAA0895L</i> (p = 0,027) <i>ZNF841</i> (p = 0,0095) <i>ZNF276</i> (p = 0,0002)
HCG27	<i>MSH5</i> (p = 0,00096) <i>PRPF3</i> (p = 0,016) <i>AGER</i> (p < 0,0001) <i>RUSC1-AS</i> (p = 0,00014) <i>CCNL2</i> (p < 0,0001) <i>CLK2</i> (p = 0,00013) <i>KIAA0895L</i> (p = 0,027) <i>AMY2B</i> (p = 0,011) <i>ZNF276</i> (p = 0,011) <i>LENG8</i> (p < 0,0001) <i>DMTF1</i> (p = 0,0037) <i>MDM4</i> (p = 0,0074)

		<i>WDR90</i> (p = 0,036)
		<i>INTS6L</i> (p = 0,0011)
		<i>ANKZF1</i> (p = 0,0039)
		<i>PARP6</i> (p = 0,0073)
		<i>MAPK8IP3</i> (p < 0,0001)
		<i>CSAD</i> (p < 0.0001)
		<i>PPFIA4</i> (p = 0,029)
LIHC	MCM3AP-AS1 (p = 0,0099)	<i>PAXBP1</i> (p = 0,024)
		<i>U2SURP</i> (p = 0,00028)
		<i>DMTF1</i> (p = 0,091)
		<i>ILF3</i> (p = 0,0031)
		<i>FNBP4</i> (p = 0,0077)
		<i>PRPF39</i> (p = 0,00014)
		<i>ZNF207</i> (p = 0,00027)
		<i>SUGP2</i> (p = 0,0091)
		<i>CPSF6</i> (p = 0,0042)
		<i>TAF4</i> (p = 0,0032)
		<i>CCDC93</i> (p = 0,002)
		<i>ZNF37BP</i> (p = 0,0045)
		<i>MSH5</i> (p = 0,038)
		<i>ZNF337</i> (p = 0,022)
		<i>NAA25</i> (p = 0,0084)
		<i>DENND4B</i> (p = 0,025)
		<i>MAML1</i> (p = 0,024)
		<i>KCTD7</i> (p = 0,01)
PRAD	LINC00115	<i>MSH5</i> (p=0,034)

TABELA 7: LncRNAs E GENES CO-EXPRESSOS COM SOBREVIVÊNCIA LIVRE DA DOENÇA (p-valor) PARA COAD, HNSC, LIHC, LUAD E PRAD.

CÂNCER	lncRNA	GENE CO-EXPRESSO
COAD	MCM3AP-AS1	<i>FNBP4</i> (p = 0,013)
		<i>DDX39B</i> (p = 0,013)
		<i>MSH5</i> (p = 0,022)
		<i>ZNF37BP</i> (p = 0,021)

		<i>ZFC3H1</i> (p = 0,0092)
		<i>OGT</i> (p = 0,04)
		<i>PNISR</i> (p = 0,021)
		<i>LUC7L3</i> (p = 0,0022)
		<i>ZNF514</i> (p = 0,00044)
	PAN3-AS1	<i>STX16</i> (p = 0,011)
		<i>RBM39</i> (p = 0,035)
		<i>UPF3A</i> (p = 0,024)
HNSC	LINC01619	<i>CLEC2D</i> (p = 0,023)
	LINC01484	<i>PIM2</i> (p = 0,032)
LIHC	MCM3AP-AS1	<i>PAXBP1</i> (p = 0,024)
		<i>PRPF39</i> (p = 0,01)
		<i>ZNF207</i> (p = 0,0074)
		<i>SUGP2</i> (p = 0,016)
		<i>CCDC93</i> (p = 0,016)
		<i>MSH5</i> (p = 0,019)
		<i>NAA25</i> (p = 0,013)
		<i>MAML1</i> (p = 0,0076)
LUAD	MCM3AP-AS1	<i>DDX39B</i> (p = 0,034)
PRAD	LINC00115 (p = 0,025)	<i>CCNL2</i> (p = 0,00041)
		<i>RSRP1</i> (p = 0,0017)
		<i>KIAA0895L</i> (p = 0,0021)
		<i>LENG8</i> (p = 0,0012)
		<i>MSH5-SAPCD1</i> (p = 0,011)
		<i>ANKZF1</i> (p = 0,0031)
		<i>WDR90</i> (p < 0,0001)
		<i>ARGLU1</i> (p = 0,0081)
		<i>OGT</i> (p = 0,017)

Analizamos o BRCA pelos subtipos, adotando os mesmos critérios para os outros tipos de câncer, ou seja, um valor $p < 0,05$ foi considerado estatisticamente significativo quando encontrado no gene alvo da PUMILIO. No subtipo Luminal A, o gene *RALGAP β* , co-expresso com o lncRNA NORAD apresentou expressão associada significativamente com a Sobrevida Global das pacientes ($p=0,015$). Observamos também a relação significativa da expressão

do gene *NCOA3* ($p = 0,028$), também co-expresso com o lncRNA NORAD, com a sobrevida livre de doença, neste caso em pacientes do subtipo Basal-like.

6.5 Comparação dos lncRNAs e genes alvo de PUMILIO co-expressos entre os 9 tipos de câncer

Entendendo que muitos mecanismos celulares são comuns em diferentes tumores, e a fim de identificar biomarcadores e possíveis alvos terapêuticos nos 9 tipos de cânceres analisados, comparamos os dados de expressão diferencial, co-expressão, sobrevida global e sobrevida livre da doença buscando lncRNAs e genes alvos de PUM comuns entre os tipos de câncer.

Nesta análise identificamos o gene *MSH5*, que aparece diferencialmente expresso com $\logFC > 0,5$ ($p < 0,05$) em COAD, HNSC, KIRC, LIHC, LUAD e PRAD (FIGURAS 8 e 9).

FIGURA 8: ANÁLISE DA EXPRESSÃO DIFERENCIAL DO *MSH5*, $\logFC > 0,5$ ($p < 0,05$) NOS TIPOS DE CâNCER COAD, HNSC E KIRC

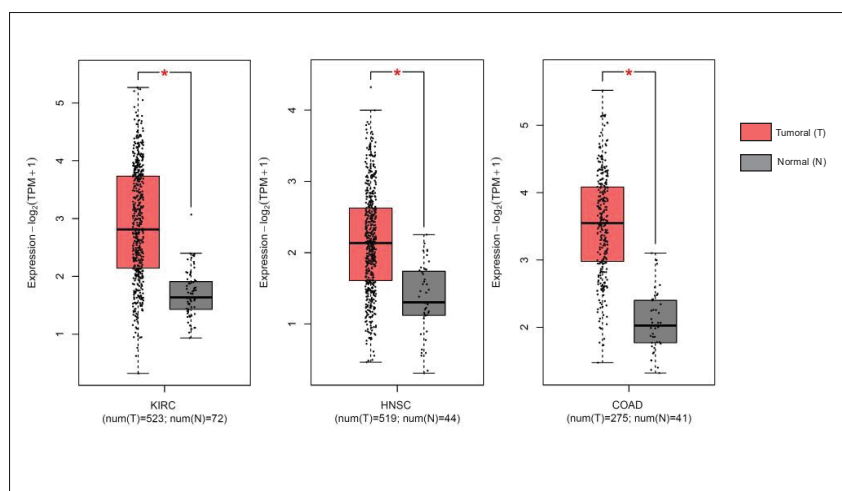
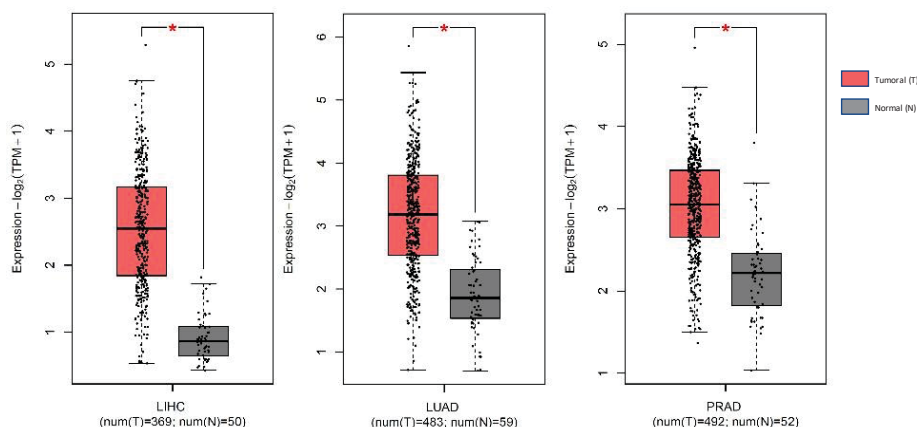
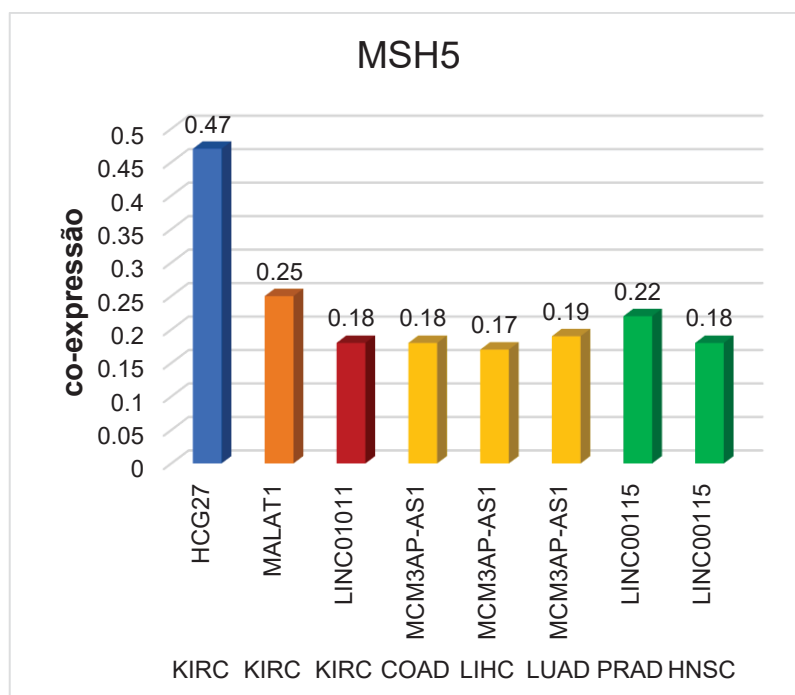


FIGURA 9: ANÁLISE DA EXPRESSÃO DIFERENCIAL DO *MSH5*, $\logFC > 0,5$ ($p < 0,05$) NOS TIPOS DE CâNCER LIHC, LUAD E PRAD



Avaliando os lncRNAs co-expressos com este gene, observamos que diferentes lncRNAs estavam correlacionados à sua expressão nos diferentes tipos de câncer. A maior correlação ocorreu no carcinoma renal de células claras do rim (KIRC) com o lncRNA HCG27, sendo observada também outras associações em KIRC, COAD, LIHC, LUAD, PRAD e HNSC com os lncRNAs MALAT1, LINC01011, MCM3AP-AS1 e LINC00115, ilustrado na FIGURA 10.

FIGURA 10: *MSH5* E SEUS lncRNAs CO-EXPRESSOS NOS TIPOS DE Câncer KIRC, COAD, LIHC, LUAD, PRAD E HNSC



A análise de sobrevida global de *MSH5*, mostrou uma associação da expressão significativa nos tipos de câncer KIRC, LIHC e PRAD (FIGURA 11); e sobrevida livre da doença significativa nos tipos de câncer COAD e LIHC (FIGURA 12). A linha vermelha representa na sobrevida global e sobrevida livre da doença dos pacientes relacionados a uma menor expressão diferencial do gene *MSH5* e a linha azul representa a sobrevida global dos pacientes relacionados com uma maior expressão diferencial do gene *MSH5*.

FIGURA 11: SOBREVIDA GLOBAL ($p < 0,05$) PARA O GENE ALVO *MSH5* EM KIRC, LIHC E PRAD

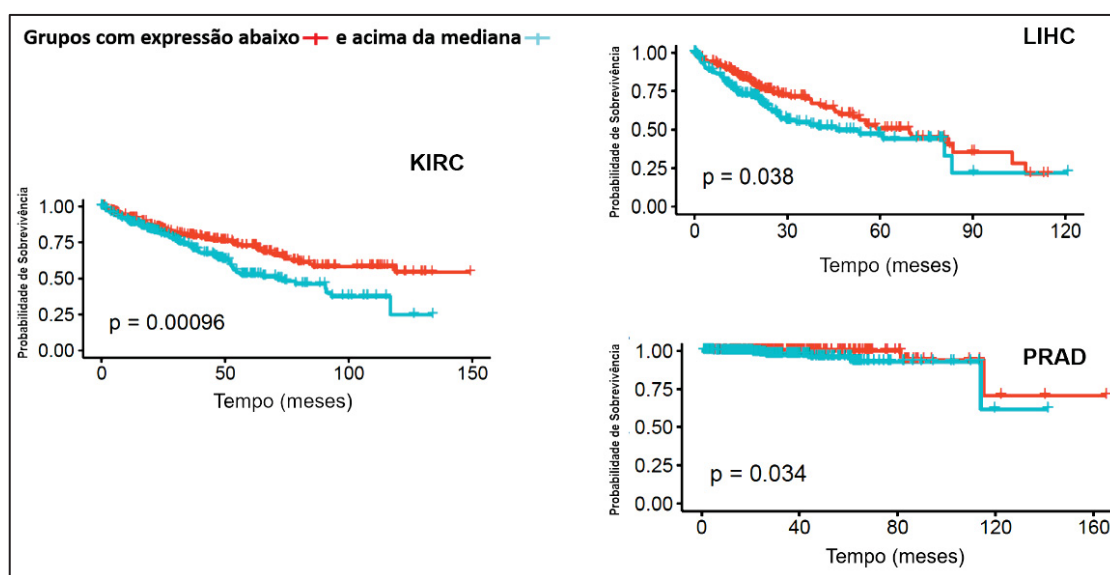
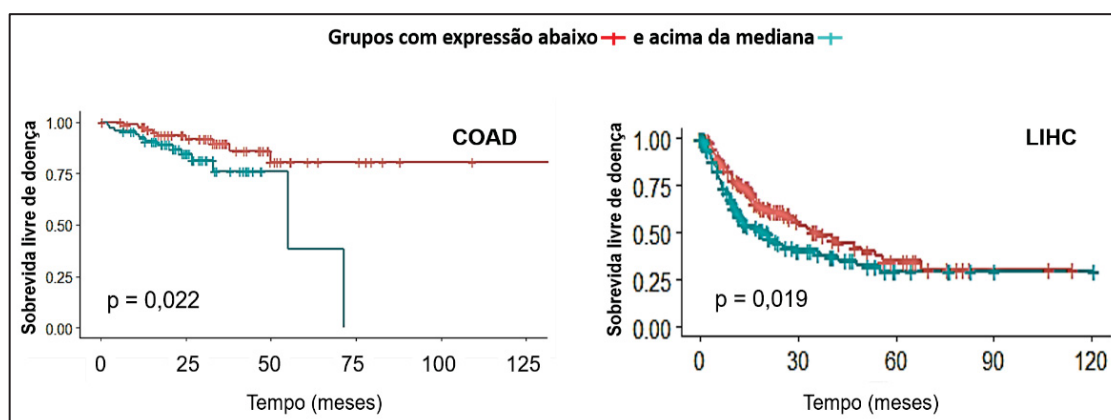


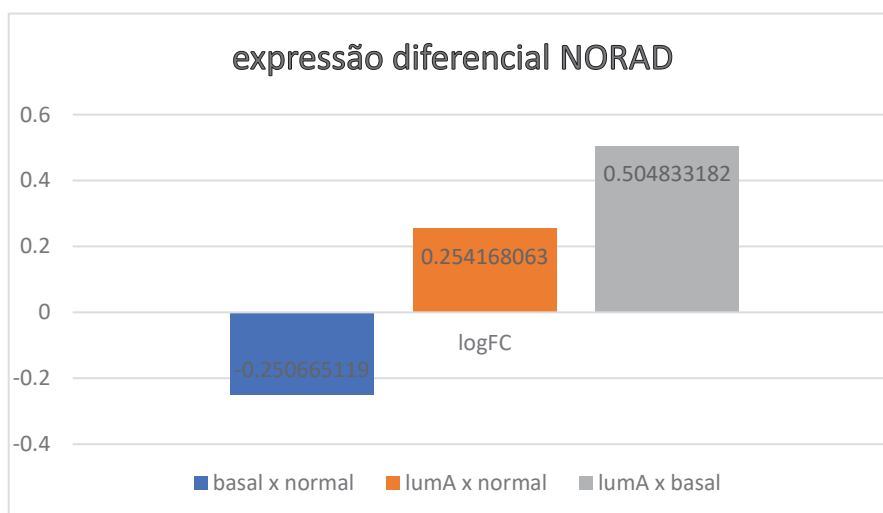
FIGURA 12: SOBREVIDA LIVRE DA DOENÇA ($p < 0,05$) PARA O GENE ALVO *MSH5* EM COAD E LIHC



No BRCA o lncRNA NORAD e o gene alvo de PUMILIO *RALGAP β* , tem valores significativos de co-expressão nos quatro subtipos, Luminal A, Luminal B, *Her2-enriched* e *Basal-like* e apresentando o maior valor em *Basal-like* (0,55).

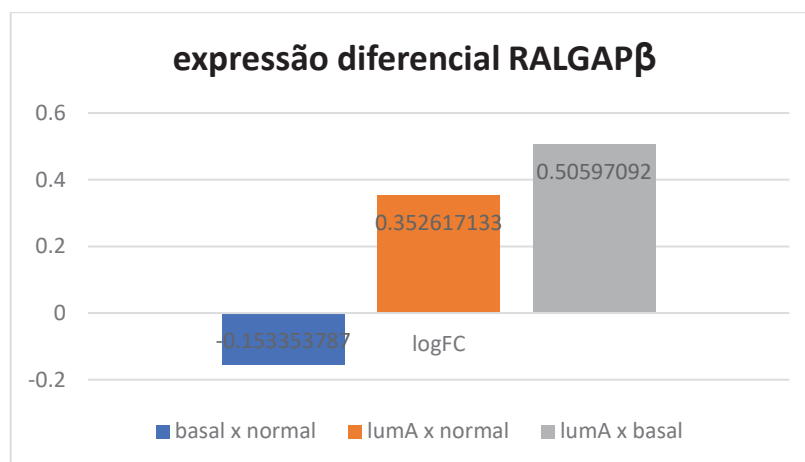
Fazendo-se a comparação entre o subtipo *Basal-like* e o tecido normal, observamos que o lncRNA NORAD aparece menos expresso em *Basal-like* ($\log FC = -0,25$). Já no subtipo LA este lncRNA aparece com a expressão aumentada em relação ao tecido não tumoral ($\log FC = 0,25$). Comparando-se os dois subtipos entre si, Luminal A e *Basal-like* o lncRNA NORAD aparece mais expresso em Luminal A ($\log FC = 0,50$) (FIGURA 13).

FIGURA 13: EXPRESSÃO DIFERENCIAL DO lncRNA NORAD EM AMOSTRAS DO BRCA PARA SUBTIPO LUMINAL A, BASAL *like* E TECIDO NORMAL



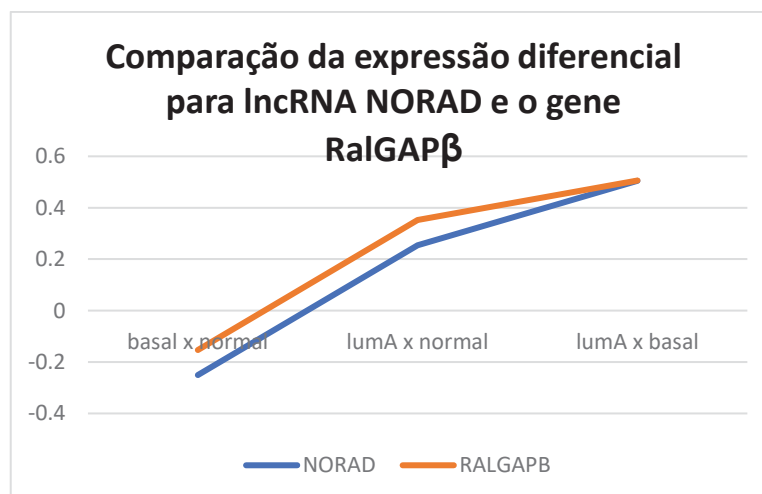
Como o lncRNA NORAD aparece co-expresso com o gene *RalGAP β* (co-expressão $\geq 0,5$) em nossas análises, verificamos a expressão diferencial deste gene nos dois subtipos do BRCA que apresentam diferença de expressão do lncRNA NORAD. O gene *RalGAP β* aparece menos expresso no subtipo *Basal-like*, quando comparado ao tecido normal ($\log FC = -0,15$) e expressão aumentada no subtipo Luminal A em comparação ao tecido normal ($\log FC = 0,35$). Na comparação entre os subtipos Luminal A e *Basal-like*, o gene *RalGAP β* aparece mais expresso no subtipo Luminal A ($\log FC = 0,50$) (FIGURA 14).

FIGURA 14: EXPRESSÃO DIFERENCIAL DO GENE *RALGAP β* NO BRCA, NOS SUBTIPOS BASAL *like*, LUMINAL A E TECIDO NORMAL



Comparando os padrões de expressão do lncRNA NORAD e o gene alvo *RALGAP β* observamos a correlação entre eles. (FIGURA 15).

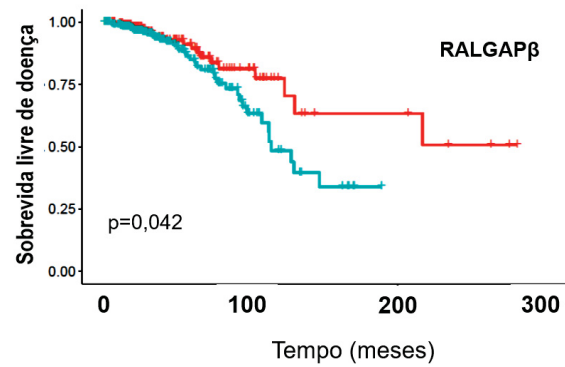
FIGURA 15: COMPARAÇÃO DA EXPRESSÃO DIFERENCIAL DO lncRNA NORAD E O GENE *RALGAP β*



Por fim, foi observada uma relação significativa da expressão do gene *RALGAP β* com a sobrevida global das pacientes com câncer de mama no subtipo Luminal A (FIGURA 16).

FIGURA 16: ANÁLISE DE SOBREVIDA GLOBAL SIGNIFICATIVA PARA O GENE *RALGAP β* NO SUBTIPO LUMINAL A

Grupos com expressão abaixo —+ e acima da mediana —+



7. Discussão

A expressão gênica é altamente regulada, desde a formação de transcritos até a geração de proteínas funcionais. Esta regulação é fundamental para criar células especializadas, expressando o que é necessário para um determinado momento para a célula, e para responder adequadamente aos estímulos ambientais (Lipshitz et al., 2017).

A regulação pós-transcricional da expressão gênica no nível do RNA, ocorre durante seus processos de processamento, transporte, maturação, modificação ou degradação. É uma das etapas regulatórias mais críticas para a célula, visto que 7% dos genes codificados para proteínas estão envolvidos na regulação pós-transcricional, e aproximadamente 25% do transcriptoma está envolvido em algum nível no metabolismo do RNA (Ramakrishnan & Janga, 2019).

A estabilidade da molécula de mRNA e o controle da tradução, são pontos essenciais na regulação da expressão gênica, sendo coordenados por outros tipos de RNAs (como miRNAs e lncRNAs) e proteínas de ligação ao RNA (RBPs), geralmente associados a sítios regulatórios presentes na região 3'-UTR dos mRNAs (Gerstberger et al., 2014; Nishtala et al., 2016).

As RBPs são proteínas que possuem domínios para se ligar ao RNA, e quando associados formam complexos de ribonucleoproteína (RNPs). As RBPs podem exercer uma função conjunta no complexo RNP, como ocorre em ribossomos, ou ainda o complexo pode ser transitório, onde a RBP desempenhará algum papel na maturação, localização ou decadência do RNA. A dinâmica desses complexos depende de sua composição, metas e cofatores (Gerstberger et al., 2014). As RBPs têm um ou mais domínios de ligação ao RNA e são classificadas em famílias de acordo com a anotação desses domínios, tais como domínios RRM, KH, CSD, Dedo de Zinco e PUF. A análise da estrutura das RBPs pode ajudar a entender sua função e como seus domínios interagem com o RNA (Glisovic et al., 2008).

Como já referido, os lncRNAs participam de diversos processos biológicos como reguladores do metabolismo celular, podendo estar associados com RBPs (Zhu et al., 2013).

As proteínas PUMILIO (PUM) são membros de uma família evolutivamente conservada de RBPs conhecida como PUF (PUmilio-Fem-3-binding factor)(WICKENS; BERNSTEIN; KIMBLE; PARKER, 2002). A família PUF é amplamente distribuída com membros encontrados em muitos eucariotos, incluindo organismos unicelulares, plantas e animais, porém não tendo sido observado até o momento em eubactérias e *archaea* (WICKENS; BERNSTEIN; KIMBLE; PARKER, 2002).

A característica de todas as proteínas Puf é a presença do domínio de ligação a RNA C-terminal conhecido como *Pumilio homology domain* (PUM-HD), formado por oito repetições imperfeitas em *tandem* de 36 aminoácidos cada. Associadas, estas estruturas formam uma molécula na forma de um arco (WANG; MCLACHLAN; ZAMORE; HALL, 2002; WICKENS; BERNSTEIN; KIMBLE; PARKER, 2002).

A PUMILIO atua na repressão da tradução de RNAs mensageiros com os quais interage através de domínios em sua estrutura capazes de reconhecer sequências nucleotídicas específicas na região 3'UTR do transcrito alvo. A repressão da tradução pode ocorrer por meio do recrutamento de fatores que irão simplesmente impedir a tradução ou ainda através da degradação do mRNA alvo pelo recrutamento de complexos de deadenilases (Spassov, 2002; Van Etten *et al.*, 2012).

A perda da fina regulação mediada por PUMILIO poderia oferecer condições ideais para o desenvolvimento de tumores, uma vez que as células cancerosas se beneficiariam da plasticidade proteômica gerada para suprir suas demandas de crescimento e metástase (David e Manley, 2010). Interessantemente, tanto a superexpressão como o silenciamento de PUM1/PUM2 em células eucarióticas resultou em aneuploidias e pontes anafásicas, indicando que a disponibilidade dessas proteínas deve ser mantida dentro de níveis específicos, com um intervalo estreito de variação, a fim de evitar consequências deletérias para a célula (Kopp e Mendell, 2018).

Foi verificado que o lncRNA NORAD interage com as proteínas PUMILIO (PUM1/PUM2), controlando sua disponibilidade no citoplasma. Como estas proteínas regulam uma ampla classe de transcritos, muitos deles envolvidos com a proteção do genoma e controle do ciclo celular, alterações nos níveis de NORAD poderiam afetar a expressão proteica dos alvos de PUMILIO, e por

consequência desequilibrar a orquestração de mecanismos importantes para a homeostase celular (Kedde *et al.*, 2010).

Partindo disso, decidimos investigar se os lncRNAs, com sítios de ligação para PUMILIO, que estivessem diferencialmente expressos em 9 tipos de câncer, em comparação com tecidos normais, poderiam ter sua expressão diretamente correlacionada com os alvos de PUMILIO, além de impacto na sobrevida dos pacientes.

Usando estratégias de análise de bioinformática, fomos capazes de identificar 6 lncRNAs diferencialmente expressos (HCG27, MALAT1, LINC01011, MCM3AP-AS1, LINC00115 e NORAD), correlacionados a alvos de PUMILIO e com impacto na sobrevida em 7 tipos tumorais (KIRC, COAD, LIHC, LUAD, PRAD, HNSC e BRCA).

Neste trabalho, exploramos particularmente os resultados observados para os eixos regulatórios HCG27-PUMILIO-MSH5 em carcinoma renal de células claras do rim (KIRC) e NORAD-PUMILIO-RALGABP em câncer de mama (BRCA).

O lncRNA HCG27 está localizado no cromossomo 6, tem 1 sítio para PUMILIO, está diferencialmente expresso em amostras tumorais em relação às normais ($\log_{FC} > 0,05$) e aparece significativamente co-expresso (0,47) com o gene alvo de PUMILIO *MSH5*.

Poucos dados existem sobre este lncRNA. Alguns estudo têm associado o HCG27 com o diabetes (Saleh *et al.*, 2020), com o AVC isquêmico (L. Zhang *et al.*, 2020) e com a osteoporose (Fei *et al.*, 2020), mas pouco se sabe na sua relação com o câncer.

O gene *MSH5* (*MutS homólogo 5*) é um membro da família MutS e está principalmente ligado ao mecanismo de reparo (MMR – MisMatch Repair). Entre todos os homólogos MSH, identificados em eucariotos, os heterodímeros *MSH4-MSH5* desempenham um papel importante no reparo de recombinação homóloga (HR) para quebras de fita dupla de DNA (DSBs) (Guo *et al.*, 2017). Deficiências no sistema de reparo (MMR) já foram relacionadas com o câncer, juntamente com a expressão negativa de genes que participam destas vias (Bei, 2017). Quando uma incorporação incorreta escapa da revisão pelas DNA polimerases, o reparo de DNA (MMR) fornece uma última linha de defesa. Consequentemente, a inativação de qualquer um dos genes MMR causa um

fenótipo que pode levar a célula à apoptose ou ao fenótipo câncer (Y. Zhang et al., 2005).

Em estudo com células da granulosa de pacientes com insuficiência ovariana prematura tratadas, observou-se a redução da expressão de MSH5, provocando um atraso no reparo do DNA e promovendo a apoptose induzida por etoposídeo (WANG; ZHANG; DANG; LI *et al.*, 2020).

Diversos trabalhos têm identificado mutações em MSH5 em contextos tumorais, como no câncer de mama masculino (BEN KRIDIS-REJEB; BEN AYED-GUERFALI; AMMOUS-BOUKHRIS; AYADI *et al.*, 2020), no câncer de pulmão (DOHERTY; SAKODA; LOOMIS; BARNETT *et al.*, 2013; KAZMA; BABRON; GABORIEAU; GÉNIN *et al.*, 2012) e carcinoma hepatocelular (LIU; ZHANG; JIA; TANG *et al.*, 2015).

A relação da expressão de *MSH5* com a progressão tumoral não está muito clara. Em estudo com 29 pacientes com carcinoma colorretal esporádico, não houve diferença nos níveis de expressão de *MSH5* em amostras tumorais comparadas com tecido normal (IOANA; ANGELESCU; BURADA; MIXICH *et al.*, 2010). Em um estudo mais recente comparando os níveis de transcritos entre pacientes de câncer colorretal de ascendência afro-americana e caucasiana americana também não mostrou diferenças significativas na expressão de *MSH5* entre amostras de pacientes e controles (PAREDES; ZABALETA; GARAI; JI *et al.*, 2020).

Já em estudo com 15 casos de leucemia mieloide crônica, Nowick *et al.*, observaram uma expressão significativamente maior de *MSH5* nas amostras dos pacientes em relação aos controles (NOWICKI; PAWLOWSKI; FISCHER; HESS *et al.*, 2003). Em glioma, dois estudos mostram que o aumento de expressão de MSH5 está relacionado com menor sobrevida (GRABOWICZ; WILCZYŃSKI; KAMIŃSKA; ROURA *et al.*, 2020; GUSEV; BHUVANESHWAR; SONG; ZENKLUSEN *et al.*, 2018).

Aqui, observamos que o *MSH5* está com uma expressão positiva ($\log_{FC} > 0,05$) nos tipos de câncer COAD, HNSC, KIRC, LIHC, LUAD e PRAD e que particularmente em KIRC, o aumento de expressão tem um impacto muito significativo na sobrevida global ($p = 0,00096$) dos pacientes. Sabemos que células tumorais têm uma maior frequência de replicação, estando sujeitas a maiores erros no DNA. Sugerimos então que a expressão positiva do *MSH5*

pode estar contribuindo na manutenção das células cancerosas, através do reparo MMR, contribuindo para a evasão da apoptose.

Sobre a relação do lncRNA NORAD (lncRNA ativado por dano ao DNA) com os alvos da PUMILIO, o maior valor de co-expressão aparece com o gene alvo *RALGAP β* (co-expressão = 0,55) no subtipo Basal-*like* do BRCA, porém, a dupla aparece também nos demais subtipos: Luminal A, Luminal B e Her2-*enriched* sugerindo uma forte relação ao fenótipo BRCA.

O gene *NORAD* está localizado no cromossomo 20q11.23, na fita 5'→3' (*reverse strand*), contém 5.339 nucleotídeos, apenas 1 éxon, e está localizado preferencialmente no citoplasma celular (Tichon et al., 2016). É altamente conservado e abundante, com níveis de expressão de aproximadamente 500 a 1.000 cópias por célula. Sabemos que este lncRNA mantém a estabilidade genômica sequestrando proteínas PUMILIO impedindo que estas reprimam a estabilidade e tradução de mRNA alvos. Na ausência do NORAD, a hiperatividade da PUMILIO reprime fatores mitóticos, de reparo do DNA e de replicação (S. Lee et al., 2016).

Semelhante com alguns outros lncRNAs, como XIST e FIRRE (Hacisuleyman et al., 2016) a sequência do NORAD contém elementos repetidos – 12 unidades de repetições NORAD (NRUs) – reconhecíveis e similares à sequência, que provavelmente se originaram por duplicação em tandem no surgimento de mamíferos e ainda compartilham substancial homologia de sequência (Tichon et al., 2016). A maioria das NRUs contém um ou dois locais de ligação para os dois homólogos de PUMILIO em mamíferos: PUM1 e PUM2. O NORAD contém pelo menos 17 desses elementos de reconhecimento de PUM (PREs), que possuem a sequência de consenso UGURNAUA (R=A/G). A deleção do NORAD em linhagens celulares resulta em um aumento da repressão de genes portadores de PREs em suas regiões 3'UTR, e a superexpressão tem efeito inverso (S. Lee et al., 2016).

Recentemente foi mostrado que lncRNA NORAD promove a condensação e sequestro de proteínas PUM em corpos NP (NORAD-PUMILIO) (ELGUINDY; MENDELL, 2021). Embora NORAD não seja o único, ele é muito mais abundante do que outros transcritos contendo PRE. Isso torna o NORAD um arcabouço

forte e específico para a montagem destes condensados, permitindo que atue como um potente inibidor de PUM.

Os alvos de PUM regulados pelo NORAD são enriquecidos em genes envolvidos na divisão celular (Tichon et al., 2016), e a depleção deste lncRNA resulta em defeitos na mitose e no acúmulo de células com instabilidade no número de cromossomos (S. Lee et al., 2016).

Em nossas análises de co-expressão entre os lncRNAs com sítios PRE e os genes alvos de PUMILIO, encontramos uma correlação positiva mais alta entre o NORAD e o gene alvo *Ra/GAP β* no câncer de mama.

RaGAPs (proteínas ativadoras de Ral GTPase) são complexos de múltiplas subunidades compreendendo uma subunidade alfa catalítica (α) e subunidades regulatórias beta (β). Atuam como ativadores de GTPase para as pequenas GTPases RALA e RALB. A função e a estabilidade do complexo requerem a associação das subunidades α e β . Na célula, o *Ra/GAP β* encontra-se no núcleo e no complexo de Golgi durante a intérfase e sofre uma dramática redistribuição durante a mitose, migrando para o fuso mitótico e para a ponte intercelular na citocinese (Personnic et al., 2014).

Como o lncRNA NORAD é capaz de sequestrar proteínas PUMILIO, impedindo que estas reprimam mRNAs alvos, acreditamos que a modulação da expressão do longo interfere no sequestro da PUMILIO. Como o gene *Ra/GAP β* é um alvo da PUMILIO, o aumento ou diminuição da expressão do lncRNA NORAD, interfere na eficiência de interação com PUMILIO, devendo levar a um aumento ou diminuição da expressão do gene *Ra/GAP β* no BRCA.

Nossos dados mostram que em câncer de mama, NORAD e *RALGAP β* têm sua expressão correlacionada nos quatro subtipos. O valor de co-expressão mais importante (0,55) foi com o subtipo Basal-*like* de câncer de mama, tendo ambos significativa redução de expressão em relação às amostras não tumorais ($p < 0,05$). Este perfil tumoral está associado com o pior prognóstico em relação aos demais subtipos de BRCA (maior risco de metástase), possui alto grau histológico (menor diferenciação celular) e acomete mulheres mais jovens. Além disso, não responde bem aos tratamentos hormonais e imunoterápicos por não apresentar alvos terapêuticos definidos (Eroles et al., 2012). Um estudo em

carcinoma oral de células escamosas (GAO; LIU; YOSHIDA; SHI *et al.*, 2019) mostrou que o silenciamento de *RalGAP β* aumentou o nível de ativação de RalA promovendo a migração e invasão de células HSC-2 *in vitro*, perfil característico de células metastáticas. Em um estudo com células HeLa a depleção da proteína *RalGAP β* no fuso mitótico e na ponte intercelular causou um aumento considerável no número de células metafásicas com desalinhamento cromossômico aberrante no plano equatorial (PERSONNIC; LAKISIC; GOUIN; ROUSSEAU *et al.*, 2014).

Neste trabalho, não observamos valores significativos relacionados à expressão do NORAD e sobrevida global em pacientes do subtipo Basal-like. Este resultado difere do anteriormente publicado pelo nosso grupo, onde observamos uma associação entre o aumento da expressão do NORAD com diminuição na sobrevida ($p < 0,0023$). Esta diferença deve-se a uma modificação metodológica, visto que no estudo anterior (MATHIAS; PEDROSO; PABST; LIMA *et al.*, 2021), fora considerado um período de tempo menor (170 meses) na análise, em relação à este (280 meses). Após 170 meses, há uma grande quantidade de dados censurados, que podem interferir nos resultados. Como no trabalho atual as análises de sobrevida envolvendo todos os genes estudados, usou o tempo total de estudo disponível no TCGA, optamos por usar a mesma condição para o NORAD. Desta forma, acreditamos que com um número maior de casos com seguimento contínuo das pacientes, o impacto da expressão de NORAD possa ser observado e associado com a redução de sobrevida, podendo ser usado como informação preditiva da progressão da doença para este subtipo.

No subtipo Luminal A, NORAD e *RALGAP β* apresentaram expressão correlacionada sendo ambos significativamente mais expressos em tumores em relação aos tecidos normais ($p < 0,05$). Além disso, a análise de sobrevida global indicou que em amostras Luminal A, o aumento da expressão de *RALGAP β* está associada à uma menor sobrevida ($p = 0,015$). Em estudo com células HeLa, a superexpressão do *RalGAP β* interferiu na divisão celular, levando à binucleação, multinucleação e morte celular (PERSONNIC; LAKISIC; GOUIN; ROUSSEAU *et al.*, 2014).

Com base no artigo publicado que identificou o interatoma do NORAD (MUNSCHAUER; NGUYEN; SIROKMAN; HARTIGAN *et al.*, 2018), sabendo que as proteínas PUMILIO são capazes de regular centenas de outros mRNAs, flutuações na expressão desse lncRNA deve também interferir na montagem de outros complexos, alterando assim o metabolismo celular por meio de diferentes mecanismos. Por exemplo, em células de glioma, foi mostrado que a associação de NORAD e AKR1B1 ativa a via ERK, promovendo um fenótipo maligno (LUO; CHEN; HE; CAI *et al.*, 2020). De maneira oposta, o NORAD funciona como um tumor supressor pela ligação de FUBP1 e promove apoptose celular no câncer endometrial (HAN; WU; HU; CHEN *et al.*, 2020). NORAD também serve como uma plataforma para ligação de S100P, suprimindo a metástase em câncer de pulmão (TAN; YANG; SINGH; CHOU *et al.*, 2019).

Assim, nossos resultados observados nos subtipos Basal-*like* e Luminal A, reforçam a ideia de que o valor prognóstico do lncRNA NORAD pode ser diferente também em subtipos moleculares específicos no câncer de mama.

8. Conclusão

Neste trabalho selecionamos 49 lncRNAs com sítios para PUMILIO, diferencialmente expressos em 9 tipos de câncer e de genes, potencialmente regulados por PUMILIO, co-expressos com os lncRNAs, em cada tipo de tumor. Buscando encontrar candidatos a marcadores de prognóstico, através de análises de sobrevida, selecionamos os genes com expressão associada com valores mais significativos.

Nossos resultados sugerem que o lncRNA HCG27 tem a capacidade de sequestrar proteínas PUMILIO impedindo sua atividade repressora sobre o gene alvo *MSH5* (associado ao reparo de DNA). Aqui, observamos que o *MSH5* está com uma expressão aumentada em seis tipos de câncer: COAD, HNSC, KIRC, LIHC, LUAD e PRAD. Assim, propomos que aumento da expressão do HCG27 pode contribuir no fenótipo tumoral por resultar no aumento dos níveis da proteína MSH5, promovendo a evasão da apoptose e desta forma contribuindo na manutenção das células cancerosas pela manutenção da integridade do DNA. Particularmente em KIRC, observamos que o aumento de expressão de *MSH5* tem um impacto muito significativo na sobrevida global ($p = 0,00096$) dos pacientes.

Este trabalho também evidenciou a já conhecida relevância do lncRNA NORAD no contexto do câncer, por sua associação à fenótipos aneuplóides. De forma inédita, observamos que o gene alvo de PUMILIO *RALGAP β* , aparece significativamente relacionado ao NORAD, podendo em parte explicar os fenótipos associados à desregulação deste lncRNA, por tratar-se de um fator mitótico importante tanto na organização do fuso mitótico como na ponte intercelular na citocinese. Ainda, a expressão de *RALGAP β* está significativamente relacionada à sobrevida global das pacientes do subtipo molecular Luminal A.

Assim, nosso estudo sugere pela primeira vez um eixo regulatório lncRNA HCG27-PUMILIO-*MSH5* no câncer de rim e, um eixo regulatório lncRNA NORAD-PUMILIO-*RALGAP β* no câncer de mama, ambos com impacto na sobrevida dos pacientes.

9. Referências Bibliográficas

- Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. In *Nature Reviews Molecular Cell Biology* (Vol. 18, Issue 5). <https://doi.org/10.1038/nrm.2017.7>
- Bhan, A., Soleimani, M., & Mandal, S. S. (2017). Long noncoding RNA and cancer: A new paradigm. *Cancer Research*, 77(15), 3965–3981. <https://doi.org/10.1158/0008-5472.CAN-16-2634>
- Bei, Y. (2017). Genetic and genomic basis of the mismatch repair system involved in Lynch syndrome. *Physiology & Behavior*, 176(3), 139–148. <https://doi.org/10.1016/j.dnarep.2015.11.019>
- Ben Kridis-Rejeb, W.; Ben Ayed-Guerfali, D.; Ammous-Boukhris, N.; Ayadi, W. et al. Identification of novel candidate genes by exome sequencing in Tunisian familial male breast cancer patients. *Mol Biol Rep*, 47, n. 9, p. 6507-6516, Sep 2020.
- Bohn, J. A., Van Etten, J. L., Schagat, T. L., Bowman, B. M., McEachin, R. C., Freddolino, P. L., & Goldstrohm, A. C. (2018b). Identification of diverse target RNAs that are functionally regulated by human PUMILIO proteins. *Nucleic Acids Research*, 46(1), 362–386. <https://doi.org/10.1093/nar/gkx1120>
- Cabili, M. N.; Trapnell, C.; Goff, L.; Koziol, M. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*, 25, n. 18, p. 1915-1927, Sep 2011.
- Carrero, G., Hendzel, M. J., & De Vries, G. (2006). Modelling the compartmentalization of splicing factors. *Journal of Theoretical Biology*, 239(3), 298–312. <https://doi.org/10.1016/j.jtbi.2005.07.019>
- Castro, M. A. A., De Santiago, I., Campbell, T. M., Vaughn, C., Hickey, T. E., Ross, E., Tilley, W. D., Markowitz, F., Ponder, B. A. J., & Meyer, K. B. (2015). Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nature Genetics*, 48(1), 12–21. <https://doi.org/10.1038/ng.3458>
- Charles Richard, J. L., & Eichhorn, P. J. A. (2018). Platforms for Investigating LncRNA Functions. *SLAS Technology*, 23(6), 493–506. <https://doi.org/10.1177/2472630318780639>
- Chen, J., Wang, S., Jia, S., Ding, G., Jiang, G., & Cao, L. (2018). Integrated analysis of Long Non-Coding RNA and mRNA expression profile in pancreatic cancer derived exosomes treated dendritic cells by microarray analysis. *Journal of Cancer*, 9(1), 21–31. <https://doi.org/10.7150/jca.21749>
- Chen, Y., Bao, C., Zhang, X., Lin, X., Huang, H., & Wang, Z. (2019). Long non-coding RNA HCG11 modulates glioma progression through cooperating with miR-496/CPEB3 axis. *Cell Proliferation*. <https://doi.org/10.1111/cpr.12615>
- Cheong, C. G., & Tanaka Hall, T. M. (2006). Engineering RNA sequence specificity of PUMILIO repeats. *Proceedings of the National Academy of Sciences of the United States of America*, 103(37), 13635–13639. <https://doi.org/10.1073/pnas.0606294103>
- Chheda, N., & Gupta, M. K. (2014). RNA as a Permutation. March. <http://arxiv.org/abs/1403.5477>

- Cloonan, N., Forrest, A. R. R., Kolle, G., Gardiner, B. B. A., Faulkner, G. J., Brown, M. K., Taylor, D. F., Steptoe, A. L., Wani, S., Bethel, G., Robertson, A. J., Perkins, A. C., Bruce, S. J., Lee, C. C., Ranade, S. S., Peckham, H. E., Manning, J. M., McKernan, K. J., & Grimmond, S. M. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods*, 5(7), 613–619. <https://doi.org/10.1038/nmeth.1223>
- Collart, M. A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdisciplinary Reviews: RNA*, 7(4), 438–454. <https://doi.org/10.1002/wrna.1332>
- Dai, H., Shen, K., Yang, Y., Su, X., Luo, Y., Jiang, Y., Shuai, L., Zheng, P., Chen, Z., & Bie, P. (2019). PUM1 knockdown prevents tumor progression by activating the PERK/eIF2/ATF4 signaling pathway in pancreatic adenocarcinoma cells. *Cell Death and Disease*, 10(8). <https://doi.org/10.1038/s41419-019-1839-z>
- Demaria, C. T., & Brewer, G. (1996). AUF1 binding affinity to A+U-rich elements correlates with rapid mRNA degradation. *Journal of Biological Chemistry*, 271(21), 12179–12184. <https://doi.org/10.1074/jbc.271.21.12179>
- De Oliveira, J. C.; Oliveira, L. C.; Mathias, C.; Pedroso, G. A. et al. Long non-coding RNAs in cancer: Another layer of complexity. *J Gene Med*, 21, n. 1, p. e3065, Jan 2019.
- Doherty, J. A.; Sakoda, L. C.; Loomis, M. M.; Barnett, M. J. et al. DNA repair genotype and lung cancer risk in the beta-carotene and retinol efficacy trial. *Int J Mol Epidemiol Genet*, 4, n. 1, p. 11-34, 2013.
- Elguindy, M. M.; Mendell, J. T. NORAD-induced Pumilio phase separation is required for genome stability. *Nature*, 2021/06/09 2021.
- Faoro, C., & Ataide, S. F. (2014). Ribonomic approaches to study the RNA-binding proteome. *FEBS Letters*, 588(20), 3649–3664. <https://doi.org/10.1016/j.febslet.2014.07.039>
- Fei, Q., Li, X., Lin, J., Yu, L., & Yang, Y. (2020). Identification of aberrantly expressed long non-coding rnas and nearby targeted genes in male osteoporosis. *Clinical Interventions in Aging*, 15, 1779–1792. <https://doi.org/10.2147/CIA.S271689>
- Fishel, R. (2015). Mismatch repair. *Journal of Biological Chemistry*, 290(44), 26395–26403. <https://doi.org/10.1074/jbc.R115.660142>
- Ferrè, F., Colantoni, A., & Helmer-Citterich, M. (2016). Revealing protein-lncRNA interaction. *Briefings in Bioinformatics*, 17(1), 106–116. <https://doi.org/10.1093/bib/bbv031>
- Fletcher, M. N. C., Castro, M. A. A., Wang, X., De Santiago, I., O'Reilly, M., Chin, S. F., Rueda, O. M., Caldas, C., Ponder, B. A. J., Markowitz, F., & Meyer, K. B. (2013). Master regulators of FGFR2 signalling and breast cancer risk. *Nature Communications*, 4(May), 1–12. <https://doi.org/10.1038/ncomms3464>
- Gao, P.; LIU, S.; Yoshida, R.; Shi, C. Y. et al. Ral GTPase Activation by Downregulation of RalGAP Enhances Oral Squamous Cell Carcinoma Progression. *J Dent Res*, 98, n. 9, p. 1011-1019, 08 2019.

- Garzon, R., Volinia, S., Papaioannou, D., Nicolet, D., Kohlschmidt, J., Yan, P. S., Mrózek, K., Bucci, D., Carroll, A. J., Baer, M. R., Wetzler, M., Carter, T. H., Powell, B. L., Kolitz, J. E., Moore, J. O., Eisfeld, A. K., Blachly, J. S., Blum, W., Caligiuri, M. A., ... Bloomfield, C. D. (2014). Expression and prognostic impact of lncRNAs in acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 111(52), 18679–18684. <https://doi.org/10.1073/pnas.1422050112>
- Genome, C., & Analysis, A. P. (2015). 张坤 综述 王红 审校 【. *Chinese Journal of Lung Cancer*, 18(4), 219–223. <https://doi.org/10.3779/j.issn.1009-3419.2015.04.02>
- Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins. *Nature Reviews Genetics*, 15(12), 829–845. <https://doi.org/10.1038/nrg3813>
- Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters*, 582(14), 1977–1986. <https://doi.org/10.1016/j.febslet.2008.03.004>
- Goldstrohm, A. C., Hall, T. M. T., & Mckenney, K. M. (2018). Post-transcriptional Regulatory Functions of Mammalian PUMILIO Proteins. *Trends Genet.* <https://doi.org/10.1016/j.tig.2018.09.006>
- Grabowicz, I. E.; Wilczyński, B.; Kamińska, B.; Roura, A.-J. et al. Epigenetic regulation of differentially expressed genes between various glioma types. *bioRxiv*, p. 2020.2008.2029.272013, 2020.
- Gu, W., Deng, Y., Zenklusen, D., & Singer, R. H. (2004). A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization. *Genes and Development*, 18(12), 1452–1465. <https://doi.org/10.1101/gad.1189004>
- Guo, T., Zhao, S., Zhao, S., Chen, M., Li, G., Jiao, X., Wang, Z., Zhao, Y., Qin, Y., Gao, F., & Chen, Z. J. (2017). Mutations in MSH5 in primary ovarian insufficiency. *Human Molecular Genetics*, 26(8), 1452–1457. <https://doi.org/10.1093/hmg/ddx044>
- Guan, X., Chen, S., Liu, Y., Wang, L. li, Zhao, Y., & Zong, Z. H. (2018). PUM1 promotes ovarian cancer proliferation, migration and invasion. *Biochemical and Biophysical Research Communications*, 497(1). <https://doi.org/10.1016/j.bbrc.2018.02.078>
- Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, J., Tsai, M., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., Vijver, M. J. Van De, Sukumar, S., & Chang, H. Y. (2011). Promote Cancer Metastasis. *Program*, 464(7291), 1071–1076. <https://doi.org/10.1038/nature08975>
- Gusev, Y.; Bhuvaneshwar, K.; Song, L.; Zenklusen, J.-C. et al. The REMBRANDT study, a large collection of genomic data from brain cancer patients. *Scientific Data*, 5, n. 1, p. 180158, 2018/08/14 2018.
- Hacisuleyman, E., Shukla, C. J., Weiner, C. L., & Rinn, J. L. (2016). Function and evolution of local repeats in the Firre locus. *Nature Communications*, 7, 1–12. <https://doi.org/10.1038/ncomms11021>

- Han, T.; Wu, Y.; Hu, X.; Chen, Y. et al. NORAD orchestrates endometrial cancer progression by sequestering FUBP1 nuclear localization to promote cell apoptosis. *Cell Death Dis*, 11, n. 6, p. 473, Jun 18 2020.
- Heberle, H., Meirelles, V. G., da Silva, F. R., Telles, G. P., & Minghim, R. (2015). InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*, 16(1), 1–7. <https://doi.org/10.1186/s12859-015-0611-3>
- Hoadley, K. A., Yau, C., Hinoue, T., Wolf, D. M., Lazar, A. J., Drill, E., Shen, R., Taylor, A. M., Cherniack, A. D., Thorsson, V., Akbani, R., Bowlby, R., Wong, C. K., Wiznerowicz, M., Sanchez-Vega, F., Robertson, A. G., Schneider, B. G., Lawrence, M. S., Noushmehr, H., ... Laird, P. W. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell*, 173(2), 291-304.e6. <https://doi.org/10.1016/j.cell.2018.03.022>
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Magdalena, J., Kenzelmann-broz, D., Khalil, A. M., Zuk, O., Amit, I., Attardi, L. D., Regev, A., Lander, E. S., Jacks, T., & John, L. (2011). Nihms227441. 142(3), 409–419. <https://doi.org/10.1016/j.cell.2010.06.040.A>
- Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J. B., & Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics*, 8, 1–16. <https://doi.org/10.1186/1471-2164-8-39>
- Iijima, S. (1991). © 19 9 1 Nature Publishing Group 그라첼꺼. *Nature*, 354, 56–58.
- Ioana, M.; Angelescu, C.; Burada, F.; Mixich, F. et al. MMR gene expression pattern in sporadic colorectal cancer. *J Gastrointestin Liver Dis*, 19, n. 2, p. 155-159, Jun 2010.
- Jarroux, J., Morillon, A., & Pinskaya, M. (2017). Long Non Coding RNA Biology. In *Advances in experimental medicine and biology* (Vol. 1008). <https://doi.org/10.1007/978-981-10-5203-3>
- Joly, W., Chartier, A., Rojas-Rios, P., Busseau, I., & Simonelig, M. (2013). The CCR4 deadenylase acts with Nanos and PUMILIO in the fine-tuning of Mei-P26 expression to promote germline stem cell self-renewal. *Stem Cell Reports*, 1(5), 411–424. <https://doi.org/10.1016/j.stemcr.2013.09.007>
- Kanduri, C. (2016). Long noncoding RNAs: Lessons from genomic imprinting. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(1), 102–111. <https://doi.org/10.1016/j.bbagrm.2015.05.006>
- Kashima, T., Rao, N., David, C. J., & Manley, J. I. (2007). hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Human Molecular Genetics*, 16(24), 3149–3159. <https://doi.org/10.1093/hmg/ddm276>
- Kazma, R.; Babron, M. C.; Gaborieau, V.; Génin, E. et al. Lung cancer and DNA repair genes: multilevel association analysis from the International Lung Cancer Consortium. *Carcinogenesis*, 33, n. 5, p. 1059-1064, May 2012.
- Kim, J., Piao, H., Kim, B., Yao, F., Han, Z., Xiao, Z., Siverly, A. N., Lawhon, S. E., Ton, B. N., Zhou, Z., Gan, B., Nakagawa, S., Ellis, M. J., Liang, H., Hung,

- C., You, M. J., Sun, Y., & Ma, L. (2019). Long noncoding RNA MALAT1 suppresses breast cancer metastasis. 50(12), 1705–1715. <https://doi.org/10.1038/s41588-018-0252-3>.Long
- King, H. A., Cobbold, L. C., Pichon, X., Pöyry, T., Wilson, L. A., Booden, H., Jukes-Jones, R., Cain, K., Lilley, K. S., Bushell, M., & Willis, A. E. (2014). Remodelling of a polypyrimidine tract-binding protein complex during apoptosis activates cellular IRESs. *Cell Death and Differentiation*, 21(1), 161–171. <https://doi.org/10.1038/cdd.2013.135>
- Kino, M., Hur, D. E., Ichijo, T., Nader, N., & Chrousos, G. P. (2010). Kino T, Hurt D E, Ichijo T, et al. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor.[J]. *Science Signaling*, 2010, 3(107):ra8. *Science Signalling*, 3(107), 1–16. <https://doi.org/10.1126/scisignal.2000568>.Noncoding
- L. Sheng, M. Christopher, A. M. (2016). 乳鼠心肌提取 HHS Public Access. *Physiology & Behavior*, 176(1), 100–106. <https://doi.org/10.1016/j.gde.2016.03.011>
- Lee, J. T. (2010). The X as model for RNA's niche in epigenomic regulation. *Cold Spring Harbor Perspectives in Biology*, 2(9), 1–12. <https://doi.org/10.1101/cshperspect.a003749>
- Lee, S., Kopp, F., Chang, T.-C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., & Mendell, J. T. (2016). Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell*, 164 (1-2), 69–80. <https://doi.org/10.1016/j.cell.2015.12.017>
- Li, R., Qu, H., Wang, S., We, J., Zhang, L., Ma, R., Lu, J., Zhu, J., Zhong, W. De, & Jia, Z. (2018). GDCRNATools: An R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics*, 34(14), 2515–2517. <https://doi.org/10.1093/bioinformatics/bty124>
- Li, Y., Syed, J., & Sugiyama, H. (2016). RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chemical Biology*, 23(11), 1325–1333. <https://doi.org/10.1016/j.chembiol.2016.09.011>
- Li, Z., Hou, P., Fan, D., Dong, M., Ma, M., Li, H., Yao, R., Li, Y., Wang, G., Geng, P., Mihretab, A., Liu, D., Zhang, Y., Huang, B., & Lu, J. (2017). The degradation of EZH2 mediated by lncRNA ANCR attenuated the invasion and metastasis of breast cancer. *Cell Death and Differentiation*, 24(1), 59–71. <https://doi.org/10.1038/cdd.2016.95>
- Lin, K., Zhang, S., Shi, Q., Zhu, M., Gao, L., Xia, W., Geng, B., Zheng, Z., & Xu, E. Y. (2018). Essential requirement of mammalian PUMILIO family in embryonic development . *Molecular Biology of the Cell*, 29(24), 2922–2932. <https://doi.org/10.1091/mbc.e18-06-0369>
- Lipshitz, H. D., Claycomb, J. M., & Smibert, C. A. (2017). Post-transcriptional regulation of gene expression. *Methods*, 126, 1–2. <https://doi.org/10.1016/j.ymeth.2017.08.007>
- Liu, Y.; Zhang, X.; Jia, J.; Tang, L. et al. Correlation between polymorphisms in DNA mismatch repair genes and the risk of primary hepatocellular carcinoma

- for the Han population in northern China. *Scand J Gastroenterol*, 50, n. 11, p. 1404-1410, 2015.
- Lu, G., & Hall, T. M. T. (2011a). Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure*, 19(3), 361–367. <https://doi.org/10.1016/j.str.2010.12.019>
- Lu, G., & Hall, T. M. T. (2011b). Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure*, 19(3), 361–367. <https://doi.org/10.1016/j.str.2010.12.019>
- Luo, L.; Chen, C.; He, H.; Cai, M. et al. Silencing of Long Non-Coding RNA (LncRNA) Non-Coding RNA Activated by DNA Damage (NORAD) Inhibits Proliferation, Invasion, Migration, and Promotes Apoptosis of Glioma Cells via Downregulating the Expression of AKR1B1. *Med Sci Monit*, 26, p. e922659, Jul 21 2020.
- Ma, L., Bajic, V. B., & Zhang, Z. (2013). On the classification of long non-coding RNAs. *RNA Biology*, 10(6), 924–933. <https://doi.org/10.4161/rna.24604>
- Mathias, C., Pedroso, G. A., Pabst, F. R., De Lima, R. S., Kuroda, F., Cavalli, I. J., De Oliveira, J. C., De Souza Fonseca Ribeiro, E. M., & Gradia, D. F. (2021). So alike yet so different. Differential expression of the long non-coding RNAs NORAD and HCG11 in breast cancer subtypes. *Genetics and Molecular Biology*, 44(1), 1–10. <https://doi.org/10.1590/1678-4685-gmb-2020-0153>
- Mercer, T. R., Dinger, S. N. M. E., Crawford, J., Smith, M. A., Shearwood, A.-M. J., Haugen, E., Bracken, C. P., Rackham, O., Stamatoyannopoulos, J. A., Filipovska, A., & S. Mattick, J. (2011). The human mitochondrial transcriptome. *Cell*, 100(2), 130–134. <https://doi.org/10.1016/j.cell.2011.06.051>.
- Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F., & Mattick, J. S. (2008). Specific expression of long noncoding RNAs in the mouse brain. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2), 716–721. <https://doi.org/10.1073/pnas.0706729105>
- Moraes, F., & Góes, A. (2016). A decade of human genome project conclusion: Scientific diffusion about our genome knowledge. *Biochemistry and Molecular Biology Education*, 44(3), 215–223. <https://doi.org/10.1002/bmb.20952>
- Morlando, M., Ballarino, M., & Fatica, A. (2015). Long non-coding RNAs: New players in hematopoiesis and leukemia. *Frontiers in Medicine*, 2(APR), 1–5. <https://doi.org/10.3389/fmed.2015.00023>
- Munschauer, M.; Nguyen, C. T.; Sirokman, K.; Hartigan, C. R. et al. The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature*, 561, n. 7721, p. 132-136, Sep 2018.
- Nishtala, S., Neelamraju, Y., & Janga, S. C. (2016). Dissecting the expression relationships between RNA-binding proteins and their cognate targets in eukaryotic post-transcriptional regulatory networks. *Scientific Reports*, 6(May), 1–14. <https://doi.org/10.1038/srep25711>

- Nowicki, M. O.; Pawlowski, P.; Fischer, T.; Hess, G. et al. Chronic myelogenous leukemia molecular signature. *Oncogene*, 22, n. 25, p. 3952-3963, Jun 2003.
- Oliveira, J. C., Oliveira, L. C., Mathias, C., Pedroso, G. A., Lemos, D. S., Salviano, A. S., Jucoski, T. S., Lobo, S. C. A., Zambalde, E. P., Cipolla, G. A., & Gradia, D. F. (2018). Long non-coding RNAs in cancer: Another layer of complexity. *The Journal of Gene Medicine*, 21(1).
- Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, 40(12), 1413–1415. <https://doi.org/10.1038/ng.259>
- Parasramka, M. A., Maji, S., Matsuda, A., Yan, I. K., & Patel, T. (2016). Long non-coding RNAs as novel targets for therapy in hepatocellular carcinoma. *Pharmacology and Therapeutics*, 161(March), 67–78. <https://doi.org/10.1016/j.pharmthera.2016.03.004>.
- Paredes, J.; Zabaleta, J.; Garai, J.; Ji, P. et al. Immune-Related Gene Expression and Cytokine Secretion Is Reduced Among African American Colon Cancer Patients. *Front Oncol*, 10, p. 1498, 2020.
- Pedroso, G. A., Mathias, C., Pabst, F. R., Lima, R. S., Kurada, F., Cavalli, I. J., de Oliveira, J. C., Ribeiro, E. M. de S. F., & Gradia, D. F. (2019). So alike but so different. Differential expression of the long non coding RNAs NORAD and HCG11 in breast cancer subtypes. In Press.
- Personnic, N., Lakisic, G., Gouin, E., Rousseau, A., Gautreau, A., Cossart, P., & Bierne, H. (2014). A role for Ral GTPase-activating protein subunit β in mitotic regulation. *FEBS Journal*, 281(13), 2977–2989. <https://doi.org/10.1111/febs.12836>
- Python.org. (2009). Python FAQ. 39(6), 925–938. <https://doi.org/10.1016/j.molcel.2010.08.011>.The
- Ramakrishnan, A., & Janga, S. C. (2019). Human protein-RNA interaction network is highly stable across mammals. *BMC Genomics*, 20(Suppl 12), 1–14. <https://doi.org/10.1186/s12864-019-6330-9>
- Ren, J., Ding, L., Zhang, D., Shi, G., Xu, Q., Shen, S., Wang, Y., Wang, T., & Hou, Y. (2018). Carcinoma-associated fibroblasts promote the stemness and chemoresistance of colorectal cancer by transferring exosomal lncRNA H19. *Theranostics*, 8(14), 3932–3948. <https://doi.org/10.7150/thno.25541>
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., & Chang, H. Y. (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*, 129(7), 1311–1323. <https://doi.org/10.1016/j.cell.2007.05.022>
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47. <https://doi.org/10.1093/nar/gkv007>
- Rnas, N., Riken, T., Gene, M., Project, E., When, U., For, U., & Information, S. (2002). <Nature01266.Pdf>. 563–573.

- Saleh, A. A., Kasem, H. E., Zahran, E. S., & El-Hefnawy, S. M. (2020). Cell-free long non-coding RNAs (LY86-AS1 & HCG27_201 and GAS5) as biomarkers for pre-diabetes and type 2 DM in Egypt. *Biochemistry and Biophysics Reports*, 23(May), 100770. <https://doi.org/10.1016/j.bbrep.2020.100770>
- Schmidt, L. H., Görlich, D., Spieker, T., Rohde, C., Schuler, M., Mohr, M., Humberg, J., Sauer, T., Thoenissen, N. H., Hüge, A., Voss, R., Marra, A., Faldum, A., Müller-Tidow, C., Berdel, W. E., & Wiewrodt, R. (2014). Prognostic impact of Bcl-2 depends on tumor histology and expression of MALAT-1 lncRNA in non-small-cell lung cancer. *Journal of Thoracic Oncology*, 9(9), 1294–1304. <https://doi.org/10.1097/JTO.0000000000000243>
- Schmitt, A. M., Garcia, J. T., Hung, T., Flynn, R. A., Shen, Y., Payumo, A. Y., Broz, D. K., Guo, S., Chen, J. K., Attardi, L. D., Chang, H. Y., & Pharmaceuticals, I. (2017). HHS Public Access. 48(11), 1370–1376. <https://doi.org/10.1038/ng.3673>.
- Shigunov, P.; Sotelo-Silveira, J.; Kuligovski, C.; De Aguiar, A. M. et al. PUMILIO-2 is involved in the positive regulation of cellular proliferation in human adipose-derived stem cells. *Stem Cells Dev*, 21, n. 2, p. 217-227, Jan 2012.
- Silva, I. L. Z.; Robert, A. W.; Cabo, G. C.; Spangenberg, L. et al. Effects of PUMILIO1 and PUMILIO2 knockdown on cardiomyogenic differentiation of human embryonic stem cells culture. *PLoS One*, 15, n. 5, p. e0222373, 2020.
- Smialek, M. J., Ilaslan, E., Sajek, M. P., Swiercz, A., Janecki, D. M., Kusz-Zamelczyk, K., Wozniak, T., Kotecki, M., Handschuh, L., Figlerowicz, M., & Jaruzelska, J. (2020a). Characterization of RNP Networks of PUM1 and PUM2 Post-Transcriptional Regulators in TCam-2 Cells, a Human Male Germ Cell Model. *Cells*, 9(4), 984. <https://doi.org/10.3390/cells9040984>
- Smialek, M. J., Ilaslan, E., Sajek, M. P., Swiercz, A., Janecki, D. M., Kusz-Zamelczyk, K., Wozniak, T., Kotecki, M., Handschuh, L., Figlerowicz, M., & Jaruzelska, J. (2020b). Characterization of RNP Networks of PUM1 and PUM2 Post-Transcriptional Regulators in TCam-2 Cells, a Human Male Germ Cell Model. *Cells*, 9(4), 1–23. <https://doi.org/10.3390/cells9040984>
- Sonoda, J., & Wharton, R. P. (1999). Recruitment of nanos to hunchback mRNA by PUMILIO. *Genes and Development*, 13(20), 2704–2712. <https://doi.org/10.1101/gad.13.20.2704>
- Sorlie T; Perou CM; Tibshirani R; Geisler S; Johnsen H, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci.*, 98(19):10869- 74, 2001.
- Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, 22(2), 96–118. <https://doi.org/10.1038/s41580-020-00315-9>
- Tan, B.-S.; Yang, M.-C.; Singh, S.; Chou, Y.-C. et al. LncRNA NORAD is repressed by the YAP pathway and suppresses lung and breast cancer metastasis by sequestering S100P. *Oncogene*, 2019/04/09 2019.

- Tang, Z. et al. (2017) GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*, 10.1093/nar/gkx247.
- Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., & Wahle, E. (2010). Subunits of the *Drosophila* CCR4-NOT complex and their roles in mRNA deadenylation. *Rna*, 16(7), 1356–1370. <https://doi.org/10.1261/rna.2145110>
- Tichon, A., Gil, N., Lubelsky, Y., Solomon, T. H., Lemze, D., Itzkovitz, S., Stern-Ginossar, N., & Ulitsky, I. (2016). A conserved abundant cytoplasmic long noncoding RNA modulates repression by PUMILIO proteins in human cells. *Nature Communications*, 7, 1–10. <https://doi.org/10.1038/ncomms12209>
- Tong, J., Ma, X., Yu, H., & Yang, J. (2019). SNHG15: A promising cancer-related long noncoding RNA. *Cancer Management and Research*, 11, 5961–5969. <https://doi.org/10.2147/CMAR.S208054>
- Ulitsky, I., & Bartel, D. P. (2013). Ulitsky, Igor, and David P Bartel. 2013. “lincRNAs: Genomics, Evolution, and Mechanisms.” *Cell* 154 (1): 26–46. doi:10.1016/j.cell.2013.06.020.lincRNAs: genomics, evolution, and mechanisms. *Cell*, 154(1), 26–46. <https://doi.org/10.1016/j.cell.2013.06.020>
- Valadkhan, S., & Gunawardane, L. S. (2020). lncRNA-mediated regulation of the interferon response. January.
- Van Etten, J., Schagat, T. L., Hrit, J., Weidmann, C. A., Brumbaugh, J., Coon, J. J., & Goldstrohm, A. C. (2012). Human PUMILIO proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *Journal of Biological Chemistry*, 287(43), 36370–36383. <https://doi.org/10.1074/jbc.M112.373522>
- Wang, K. C., & Chang, H. Y. (2012). NIH Public Accessfile:///Users/ssheppa1/Desktop/Dropbox/ShaoLine Sheppard - Master 17/Resources/Papers/Plant Cell-2013-Jabnounge-tpc.113.116251.pdf. *Molecular Cell*, 43(6), 904–914. <https://doi.org/10.1016/j.molcel.2011.08.018>.
- Wang, L., Yang, F., Jia, L. T., & Yang, A. G. (2017). Missing Links in Epithelial-Mesenchymal Transition: Long Non-Coding RNAs Enter the Arena. *Cellular Physiology and Biochemistry*, 44(4), 1665–1680. <https://doi.org/10.1159/000485766>
- Wang, X., Mclachlan, J., Zamore, P. D., & Tanaka Hall, T. M. (2002). Forbes and proteins generally. *Cell*, 110, 501–512.
- Wang, X., Zamore, P. D., & Tanaka Hall, T. M. (2001). Crystal structure of a PUMILIO homology domain. *Molecular Cell*, 7(4), 855–865. [https://doi.org/10.1016/S1097-2765\(01\)00229-5](https://doi.org/10.1016/S1097-2765(01)00229-5)
- Wang, Y. C., He, W. Y., Dong, C. H., Pei, L., & Ma, Y. L. (2019). lncRNA HCG11 regulates cell progression by targeting miR-543 and regulating AKT/mTOR pathway in prostate cancer. *Cell Biology International*, 43(12), 1453–1462. <https://doi.org/10.1002/cbin.11194>
- Wang, X.; Zhang, X.; Dang, Y.; Li, D. et al. Long noncoding RNA HCP5 participates in premature ovarian insufficiency by transcriptionally regulating

- MSH5 and DNA damage repair via YB1. *Nucleic Acids Research*, 48, n. 8, p. 4480-4491, 2020.
- Weidmann, C. A., Qiu, C., Arvola, R. M., Lou, T. F., Killingsworth, J., Campbell, Z. T., Tanaka Hall, T. M., & Goldstrohm, A. C. (2016). *Drosophila nanos* acts as a molecular clamp that modulates the RNA-binding and repression activities of PUMILIO. *ELife*, 5(AUGUST), 1–28. <https://doi.org/10.7554/eLife.17096>
- Wickens, M.; Bernstein, D. S.; Kimble, J.; Parker, R. A PUF FAMILY portrait: 3'UTR regulation as a way of life. *Trends Genet*, 18, n. 3, p. 150-157, Mar 2002.
- Xu, Y., Zheng, Y., Liu, H., & Li, T. (2017). Modulation of IGF2BP1 by long non-coding RNA HCG11 suppresses apoptosis of hepatocellular carcinoma cells via MAPK signaling transduction. *International Journal of Oncology*, 51(3), 791–800. <https://doi.org/10.3892/ijo.2017.4066>
- Yeo, S. K., & Guan, J. L. (2017). Breast Cancer: Multiple Subtypes within a Tumor? *Trends in Cancer*, 3(11), 753–760. <https://doi.org/10.1016/j.trecan.2017.09.001>
- Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., Huarte, M., Zhan, M., Becker, K. G., & Gorospe, M. (2012). LincRNA-p21 Suppresses Target mRNA Translation. *Molecular Cell*, 47(4), 648–655. <https://doi.org/10.1016/j.molcel.2012.06.027>
- Zarnack, K., König, J., Tajnik, M., Martincorena, I., Eustermann, S., Stévant, I., Reyes, A., Anders, S., Luscombe, N. M., & Ule, J. (2013). Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell*, 152(3), 453–466. <https://doi.org/10.1016/j.cell.2012.12.023>
- Zeng, L., Cen, Y., & Chen, J. (2018). Long non-coding RNA MALAT-1 contributes to maintenance of stem cell-like phenotypes in breast cancer cells. *Oncology Letters*, 15(2), 2117–2122. <https://doi.org/10.3892/ol.2017.7557>
- Zhang, L., Liu, B., Han, J., Wang, T., & Han, L. (2020). Competing endogenous RNA network analysis for screening inflammation-related long non-coding RNAs for acute ischemic stroke. *Molecular Medicine Reports*, 22(4), 3081–3094. <https://doi.org/10.3892/mmr.2020.11415>
- Zhang, Y., Yuan, F., Presnell, S. R., Tian, K., Gao, Y., Tomkinson, A. E., Gu, L., & Li, G. M. (2005). Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell*, 122(5), 693–705.
- Zhang, A., Zhou, N., Huang, J., Liu, Q., Fukuda, K., Ma, D., Lu, Z., Bai, C., Watabe, K., & Mo, Y. Y. (2013). The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Research*, 23(3), 340–350. <https://doi.org/10.1038/cr.2012.164>
- Zhang, M., Chen, D., Xia, J., Han, W., Cui, X., Neuenkirchen, N., Hermes, G., Sestan, N., & Lin, H. (2017). Post-transcriptional regulation of mouse neurogenesis by PUMILIO proteins. *Genes and Development*, 31(13), 1354–1369. <https://doi.org/10.1101/gad.298752.117>

- Zhang, X., Rice, K., Wang, Y., Chen, W., Zhong, Y., Nakayama, Y., Zhou, Y., & Klibanski, A. (2010). Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: Isoform structure, expression, and functions. *Endocrinology*, 151(3), 939–947. <https://doi.org/10.1210/en.2009-0657>
- Zhang, Y., Wang, T., Huang, H. Q., Li, W., Cheng, X. L., & Yang, J. (2015). Human MALAT-1 long non-coding RNA is overexpressed in cervical cancer metastasis and promotes cell proliferation, invasion and migration. *Journal of B.U.ON.*, 20(6), 1497–1503.
- Zhu, J. J., Fu, H. J., Wu, Y. G., & Zheng, X. F. (2013). Function of lncRNAs and approaches to lncRNA-protein interactions. *Science China Life Sciences*, 56(10), 876–885. <https://doi.org/10.1007/s11427-013-4553-6>
- Zuckerman, B., Ron, M., Mikl, M., Segal, E., & Ulitsky, I. (2020). Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. *Molecular Cell*, 79(2), 251-267.e6. <https://doi.org/10.1016/j.molcel.2020.05.013>

10. TABELAS SUPLEMENTARES

TABELA SUPLEMENTAR 1: RESULTADO CO-EXPRESSÃO BLCA (p = 0,017)

BLCA			
lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
LINC00115	CCNL2	ENSG00000221978	0.156941527150894

TABELA SUPLEMENTAR 2: RESULTADO CO-EXPRESSÃO COAD (p = 0,012)

COAD			
IncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
MCM3AP-AS1	CCDC14	ENSG00000175455	0.188764086644963
	FNBP4	ENSG00000109920	0.181502293312037
	DDX39B	ENSG00000198563	0.17956690189663
	PAXBP1	ENSG00000159086	0.178872464265189
	MSH5	ENSG00000204410	0.177818846030882
	UBE2G2	ENSG00000184787	0.177716594619973
	ZNRD1ASP	ENSG00000204623	0.176734244876741
	NKTR	ENSG00000114857	0.174500501035574
	RUSC1-AS1	ENSG00000225855	0.172531586695897
	ADAT2	ENSG00000189007	0.172359064837888
	RPL32P3	ENSG00000251474	0.170204473104258
	TARBP1	ENSG00000059588	0.168249941580419
	ZNF37BP	ENSG00000234420	0.166301913457622
	SUGP2	ENSG00000064607	0.164731338689381
	ZFC3H1	ENSG00000133858	0.162356301353013
	CFAP44	ENSG00000206530	0.160816870066836
	OGT	ENSG00000147162	0.159821320482339
	PNISR	ENSG00000132424	0.159059466344467
	ZNF587	ENSG00000198466	0.158335611465845
	CREBZF	ENSG00000137504	0.155293757276541
	CTAGE7P	ENSG00000233122	0.155072304747112
	ZNF621	ENSG00000172888	0.154354667850882
	CCDC150	ENSG00000144395	0.154120453851786
	SLC25A27	ENSG00000153291	0.153463607296775
	RBM33	ENSG00000184863	0.15325940993752
	LUC7L3	ENSG00000108848	0.153061728251519
	TRIM52	ENSG00000183718	0.152695927767071
	ZNF514	ENSG00000144026	0.151228141031679
MALAT1	PLAC4	ENSG00000280109	0.162094485719424
PAN3-AS1	PAN3	ENSG00000152520	0.192541471947555
	ARGLU1	ENSG00000134884	0.190653587970003
	KLHL31	ENSG00000124743	0.184453924840646
	N4BP2L2	ENSG00000244754	0.1786638879419
	SUPT20H	ENSG00000102710	0.178182736233861
	STX16	ENSG00000124222	0.174684973015032
	RBM39	ENSG00000131051	0.173864188844221
	USPL1	ENSG00000132952	0.169926426851392
	UPF3A	ENSG00000169062	0.167721921057754
	ANKRD10	ENSG00000088448	0.161625874446672
	MTRF1	ENSG00000120662	0.161332163761455
	SLC25A27	ENSG00000153291	0.160849571590422
	CENPJ	ENSG00000151849	0.159565003717247

SETDB2	ENSG00000136169	0.155706736136761
RBM26	ENSG00000139746	0.151441864473122
CDC16	ENSG00000130177	0.150301761276136
NHLRC3	ENSG00000188811	0.150076009872017

TABELA SUPLEMENTAR 3: RESULTADO CO-EXPRESSÃO ESCA (p = 0,2)

ESCA			
lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
C15orf54	PCDH17	ENSG00000118946	0.212751300315791
	DCLK3	ENSG00000163673	0.189502431092283
	TM4SF18	ENSG00000163762	0.178754596449786
	CEP135	ENSG00000174799	0.174403285190894
	ESM1	ENSG00000164283	0.172887340183701
	TMEM26	ENSG00000196932	0.166399701579977
	PKD1L1	ENSG00000158683	0.154214045417706
	COL15A1	ENSG00000204291	0.153217805112918

TABELA SUPLEMENTAR 4: RESULTADO CO-EXPRESSÃO HNSC (p = 0,005)

HNSC			
lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
IGF2-AS	ITGA11	ENSG00000137809	0.209937491373607
	ASPN	ENSG00000106819	0.196027191413169
	MMP11	ENSG00000099953	0.186739648319119
	COL11A1	ENSG00000060718	0.186637355637647
	LRRC15	ENSG00000172061	0.183280248185452
	SGIP1	ENSG00000118473	0.176975018789425
	MAGEL2	ENSG00000254585	0.176294705614346
	COL3A1	ENSG00000168542	0.17547714188521
	NOX4	ENSG00000086991	0.175214009216353
	LUM	ENSG00000139329	0.172001269976558
	CTHRC1	ENSG00000164932	0.168423140313823
	SOX11	ENSG00000176887	0.162351795661029
	DACT3	ENSG00000197380	0.160541536676246
	COL24A1	ENSG00000171502	0.159418743678109
	POSTN	ENSG00000133110	0.157604506302713
	CHRD	ENSG00000090539	0.155891638826914
	P4HA3	ENSG00000149380	0.154861174363772
	EFEMP2	ENSG00000172638	0.154172274753986
	COL8A1	ENSG00000144810	0.152971632106235
	HEPH	ENSG00000089472	0.151402833057732
IGF2BP2-AS1	IGF2BP2	ENSG00000073792	0.170106324676006
LINC00115	CCNL2	ENSG00000221978	0.246202813893336
	KIAA0895L	ENSG00000196123	0.233783381585519
	PABPC1L	ENSG00000101104	0.227079280533609
	RSRP1	ENSG00000117616	0.225462327787625
	GABBR1	ENSG00000204681	0.216132641634285
	DDX39B	ENSG00000198563	0.212666531649381
	PRPF39	ENSG00000185246	0.209139259584281
	RUSC1-AS1	ENSG00000225855	0.203840085362902
	MSH5	ENSG00000204410	0.181660905874949
	TAF1C	ENSG00000103168	0.179581025693003
	CLK2	ENSG00000176444	0.176151535637646
	RBM39	ENSG00000131051	0.17469969204749
	CLK1	ENSG00000013441	0.174053077912732
	LENG8	ENSG00000167615	0.173921554396137
	CSAD	ENSG00000139631	0.173124771531303
	ZNF354B	ENSG00000178338	0.169105454043846
	CCNL1	ENSG00000163660	0.168690339705937
	VMP1	ENSG00000062716	0.168677725923137
	CCDC150	ENSG00000144395	0.166513362520498
	NFYC-AS1	ENSG00000272145	0.163456996727092
	E4F1	ENSG00000167967	0.162533379024985
	LUC7L3	ENSG00000108848	0.162198794701393

	CDK11A	ENSG00000008128	0.160928429627907
	AGER	ENSG00000204305	0.152727423586834
	SRSF11	ENSG00000116754	0.152528641136052
	ZNF252P-AS1	ENSG00000255559	0.152067923186907
	DMTF1	ENSG00000135164	0.152058715214973
	DDTL	ENSG00000099974	0.150230473026006
LINC01343	POU3F1	ENSG00000185668	0.289670861180825
	PPP2R2C	ENSG00000074211	0.16677189638267
	DSG1	ENSG00000134760	0.163838030166739
	GPR1	ENSG00000183671	0.150292480622749
LINC01484	PIM2	ENSG00000102096	0.163712691588748
	FCRL5	ENSG00000143297	0.161345684596262
LINC01619	BTG1	ENSG00000133639	0.183689165571676
	CLEC2D	ENSG00000069493	0.163114114854017

TABELA SUPLEMENTAR 5: RESULTADO CO-EXPRESSÃO KIRC (p = 0,01)

KIRC

lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
LINC01011	HCG27	ENSG00000206344	0.234567511495511
	CLK2	ENSG00000176444	0.220184765319562
	PPFIA4	ENSG00000143847	0.190976360470808
	AGER	ENSG00000204305	0.190623751466492
	PRPF3	ENSG00000117360	0.187176425737743
	MSH5	ENSG00000204410	0.184421399503268
	ANKZF1	ENSG00000163516	0.183218965851472
	WDR90	ENSG00000161996	0.178315582167007
	CLK1	ENSG00000013441	0.176457884315696
	RABL2A	ENSG00000144134	0.176301955296657
	CCNL2	ENSG00000221978	0.17540705163329
	CLK4	ENSG00000113240	0.175006386566427
	PHYKPL	ENSG00000175309	0.17221769773126
	ANKRD10	ENSG00000088448	0.171250591490764
	LUC7L3	ENSG00000108848	0.170305586914625
	DMAP1	ENSG00000178028	0.169514320861188
	PARP6	ENSG00000137817	0.169083218771916
	KIAA0895L	ENSG00000196123	0.168990857399068
	PDE6C	ENSG00000095464	0.1689813714035
	SIRT7	ENSG00000187531	0.168592965129578
	PRKRIP1	ENSG00000128563	0.168317578315379
	AMY2B	ENSG00000240038	0.167962916923241
	SFSWAP	ENSG00000061936	0.167529958368305
	DMTF1	ENSG00000135164	0.166742877205266
	DDX39B	ENSG00000198563	0.166338838710236
	LINC00115	ENSG00000225880	0.166275390051689
	MAPK8IP3	ENSG00000138834	0.165615290362929
	NUP85	ENSG00000125450	0.164367388973327
	RUSC1-AS1	ENSG00000225855	0.162680524024113
	ZNF276	ENSG00000158805	0.162415302883135
	TTC31	ENSG00000115282	0.162137639374202
	PCGF3	ENSG00000185619	0.161219508431581
	PPP4R1L	ENSG00000124224	0.160062529157458
	INTS6L	ENSG00000165359	0.158223076995904
	TAF1C	ENSG00000103168	0.157582653326205
	TRIM52	ENSG00000183718	0.156746863698199
	FAM228B	ENSG00000219626	0.156683332388237
	RAB24	ENSG00000169228	0.155503674530965
	FNBP4	ENSG00000109920	0.154818620517999
	CFAP44	ENSG00000206530	0.15326493602124
	GTPBP2	ENSG00000172432	0.152847225490246
	CCDC14	ENSG00000175455	0.15281485498332
	ZNF337	ENSG00000130684	0.152259009774979

	THOC1	ENSG00000079134	0.152205919292658
	GIGYF1	ENSG000000146830	0.151765766143152
	RSRP1	ENSG000000117616	0.150474970148334
	VEGFA	ENSG000000112715	0.150405432962112
MALAT1	RUSC1-AS1	ENSG000000225855	0.282772680003312
	NKTR	ENSG000000114857	0.272972634953167
	DDX39B	ENSG000000198563	0.271440853208955
	TTLL3	ENSG000000214021	0.267970893564039
	LENG8	ENSG000000167615	0.267647332629511
	CSAD	ENSG000000139631	0.266257330415469
	L3MBTL1	ENSG000000185513	0.264068790790442
	C3orf35	ENSG000000198590	0.261851665280707
	LUC7L	ENSG000000007392	0.256140730672894
	AMY2B	ENSG000000240038	0.249634963143005
	MCM3AP-AS1	ENSG000000215424	0.249297433021498
	MSH5	ENSG000000204410	0.246633509173638
	MDM4	ENSG000000198625	0.245587563100406
	CCNL2	ENSG000000221978	0.244629538083812
	ADAM20P1	ENSG000000259158	0.243192600940982
	ATP2A1	ENSG000000196296	0.239667147948841
	RSRP1	ENSG000000117616	0.236352665788119
	DMTF1	ENSG000000135164	0.23592358883917
	WSB1	ENSG000000109046	0.235409685026463
	RPL32P3	ENSG000000251474	0.234442139411591
	MAPK8IP3	ENSG000000138834	0.233287727180601
	KCNAB3	ENSG000000170049	0.233067871371788
	TRIM52	ENSG000000183718	0.23239472360088
	TNRC6A	ENSG000000090905	0.231181167758365
	PNISR	ENSG000000132424	0.230811239835429
	FNBP4	ENSG000000109920	0.229669860852224
	LUC7L3	ENSG000000108848	0.229665205284162
	GOLGA6L9	ENSG000000197978	0.229381450469148
	ZNF26	ENSG000000198393	0.228458331638199
	RBM33	ENSG000000184863	0.228275486204707
	PABPC1L	ENSG000000101104	0.227129771368069
	DCAF4L1	ENSG000000182308	0.226084862609814
	KIAA0895L	ENSG000000196123	0.225366287192641
	LINC00115	ENSG000000225880	0.224561369377393
	ZNF841	ENSG000000197608	0.224361464917003
	ZNF276	ENSG000000158805	0.223690171374507
	CBWD3	ENSG000000196873	0.223276231625289
	NFYC-AS1	ENSG000000272145	0.223112495766738
	PDE6C	ENSG000000095464	0.222810525572191
	CCNL1	ENSG000000163660	0.221966838011184
	CAPRIN2	ENSG000000110888	0.221159903014933
	ZNF337	ENSG000000130684	0.220025818100917

FAM111A	ENSG00000166801	0.219546019719202
SLC25A27	ENSG00000153291	0.219122598911962
ZNF700	ENSG00000196757	0.21872326536234
RRN3P1	ENSG00000248124	0.216560017515539
ZNF37BP	ENSG00000234420	0.215349913210014
ARGLU1	ENSG00000134884	0.215301497470596
RBM39	ENSG00000131051	0.215118258694124
CTAGE7P	ENSG00000233122	0.214795674373117
AGER	ENSG00000204305	0.214534311564417
FAM186A	ENSG00000185958	0.214112969578417
ANKRD36	ENSG00000135976	0.212843980768386
TTC14	ENSG00000163728	0.211807365148253
CFAP44	ENSG00000206530	0.211069676703339
UCP3	ENSG00000175564	0.209847462362193
THOC1	ENSG00000079134	0.209512365263917
GIGYF1	ENSG00000146830	0.208387765684356
MAN2C1	ENSG00000140400	0.208032796837725
ABCA10	ENSG00000154263	0.207192893057575
PCGF3	ENSG00000185619	0.207003847445225
ZNF354B	ENSG00000178338	0.206839718272814
ZNF514	ENSG00000144026	0.205263296671224
ZNRD1ASP	ENSG00000204623	0.205165027340773
OGT	ENSG00000147162	0.204289315456733
CCDC14	ENSG00000175455	0.203387880878715
ANKRD10	ENSG00000088448	0.201283170159246
CLK1	ENSG00000013441	0.201070505018215
ZNF345	ENSG00000251247	0.200966674189667
LRRC69	ENSG00000214954	0.20069327975039
N4BP2L2	ENSG00000244754	0.200083088519852
CREBZF	ENSG00000137504	0.198940517863158
PRPF3	ENSG00000117360	0.198927075217344
TAF1C	ENSG00000103168	0.197305977498298
RBM5	ENSG00000003756	0.196928827074076
MAK	ENSG00000111837	0.196445440460158
NFATC2IP	ENSG00000176953	0.1948085413414
ANKRD36C	ENSG00000174501	0.194327120261506
ZNF169	ENSG00000175787	0.19354388958128
TIA1	ENSG00000116001	0.193113982487144
OFD1	ENSG00000046651	0.193105754008459
PRPF39	ENSG00000185246	0.193022599548842
DDX17	ENSG00000100201	0.192773247400454
LRRC39	ENSG00000122477	0.19263349593849
ZNF587	ENSG00000198466	0.192479434915236
HCG27	ENSG00000206344	0.191976129319483
KLC1	ENSG00000126214	0.191900047947478
SREK1	ENSG00000153914	0.191853900189884

NPPA	ENSG00000175206	0.19184615693464
GABBR1	ENSG00000204681	0.189898933873561
CLK2	ENSG00000176444	0.189531852205894
PAXBP1	ENSG00000159086	0.189505229669379
INTS6L	ENSG00000165359	0.189389769333429
SFSWAP	ENSG00000061936	0.18858880316274
ADAT2	ENSG00000189007	0.188235356962631
GSAP	ENSG00000186088	0.187563354525842
PARP6	ENSG00000137817	0.187069710846118
ZNF23	ENSG00000167377	0.1856956275558
ZGRF1	ENSG00000138658	0.183819954742325
MSANTD2	ENSG00000120458	0.18359037785262
ZFC3H1	ENSG00000133858	0.182173769510874
ZNF333	ENSG00000160961	0.181788130906036
CEP152	ENSG00000103995	0.181605018444637
DDX55	ENSG00000111364	0.180957738950002
PIDD1	ENSG00000177595	0.180884968260876
CELF6	ENSG00000140488	0.180557064924084
TRA2A	ENSG00000164548	0.179690123265919
MZF1-AS1	ENSG00000267858	0.179319696367007
ARPC4-TTLL3	ENSG00000250151	0.179278860343369
GOLGA6L10	ENSG00000278662	0.179172119482636
MYH3	ENSG00000109063	0.178384990569504
ZNF529	ENSG00000186020	0.17781094115349
RBM25	ENSG00000119707	0.176953180205267
RUFY3	ENSG00000018189	0.176181807232726
TTN	ENSG00000155657	0.176163099148974
PPP4R1L	ENSG00000124224	0.175135281248483
CRAMP1	ENSG00000007545	0.17370831645185
TBRG1	ENSG00000154144	0.172820985889826
VIP	ENSG00000146469	0.172567821635523
ANAPC4	ENSG00000053900	0.171932533572406
QRICH2	ENSG00000129646	0.171610461016664
RIMKLB	ENSG00000166532	0.171558220940558
ANKRD36B	ENSG00000196912	0.170949429066458
STX16	ENSG00000124222	0.170506370017855
PATL2	ENSG00000229474	0.169836019988681
ZNF117	ENSG00000152926	0.168833572014576
CC2D2B	ENSG00000188649	0.168733535427416
SRSF11	ENSG00000116754	0.168003859734073
FPGT-TNNI3K	ENSG00000259030	0.167522815221499
CDK11A	ENSG00000008128	0.167518711951752
C5orf58	ENSG00000234511	0.167131957591208
ZNF182	ENSG00000147118	0.166429199020329
TRMT13	ENSG00000122435	0.166349233594316
AKAP8L	ENSG00000011243	0.165654645557604

LUC7L2	ENSG00000146963	0.165642821249383
SLC26A5	ENSG00000170615	0.165441511873589
CLK4	ENSG00000113240	0.164815204686914
CBWD6	ENSG00000215126	0.164809966015127
GOLGA6L4	ENSG00000184206	0.164723203632021
PNN	ENSG00000100941	0.164499873760432
SLC4A5	ENSG00000188687	0.164083964315993
PABPN1	ENSG00000100836	0.163401894470134
METTL17	ENSG00000165792	0.1627983167831
CCNT2	ENSG00000082258	0.162596152548406
ZNF449	ENSG00000173275	0.162442191004434
AGBL2	ENSG00000165923	0.162398686106893
GABRE	ENSG00000102287	0.162311980010931
VMP1	ENSG00000062716	0.162229221385258
SPDYE6	ENSG00000260097	0.161564245582893
SETD4	ENSG00000185917	0.160427849274152
DIP2A	ENSG00000160305	0.160409174904481
PHYKPL	ENSG00000175309	0.160304712791906
ZRANB2	ENSG00000132485	0.160283500807913
SYT2	ENSG00000143858	0.160231755744147
ZNF252P-AS1	ENSG00000255559	0.160188868936051
RABL2A	ENSG00000144134	0.16001148070263
ERMN	ENSG00000136541	0.159291698294981
UBN2	ENSG00000157741	0.158905537753309
TVP23C	ENSG00000175106	0.15877995130546
MAP3K12	ENSG00000139625	0.158710074566037
ZNF160	ENSG00000170949	0.158686044901352
NADSYN1	ENSG00000172890	0.157714251050799
NXF5	ENSG00000126952	0.157116448483693
RPAIN	ENSG00000129197	0.156432921038405
EFCAB13	ENSG00000178852	0.15615659223938
MYSM1	ENSG00000162601	0.155912363753838
CLHC1	ENSG00000162994	0.155611205628519
ZNF266	ENSG00000174652	0.155529440494397
GABRR2	ENSG00000111886	0.155269801849435
SCML1	ENSG00000047634	0.155059306459588
STARD9	ENSG00000159433	0.154734508235725
WASH2P	ENSG00000146556	0.153545902369435
ZNF682	ENSG00000197124	0.152927830094934
SRSF5	ENSG00000100650	0.152502588849237
ADCY10	ENSG00000143199	0.152281762400331
ANKHD1	ENSG00000131503	0.152109555398336
CCDC66	ENSG00000180376	0.151390584768212
DNA2	ENSG00000138346	0.150994865021279
PKD2L2	ENSG00000078795	0.150738598262663
PPFIA4	ENSG00000143847	0.150164316516959

HCG27	MSH5	ENSG00000204410	0.469682618201323
	PRPF3	ENSG00000117360	0.44378574153153
	AGER	ENSG00000204305	0.443006744776533
	RUSC1-AS1	ENSG00000225855	0.438799015671911
	CLK4	ENSG00000113240	0.433857858362074
	TRIM52	ENSG00000183718	0.43250452812785
	CCNL2	ENSG00000221978	0.414903853823867
	RRN3P1	ENSG00000248124	0.412397774507704
	CLK2	ENSG00000176444	0.41191093407398
	PHYKPL	ENSG00000175309	0.411374061218494
	KIAA0895L	ENSG00000196123	0.410128887772259
	AMY2B	ENSG00000240038	0.408351229111443
	DDX39B	ENSG00000198563	0.4078364453137
	CLK1	ENSG00000013441	0.406426321405153
	ZNF276	ENSG00000158805	0.405664885009879
	LENG8	ENSG00000167615	0.38979087438479
	DMTF1	ENSG00000135164	0.389751398231011
	MDM4	ENSG00000198625	0.383210582893274
	WDR90	ENSG00000161996	0.380342305090543
	INTS6L	ENSG00000165359	0.378694362906799
	ANKZF1	ENSG00000163516	0.376833868303986
	PARP6	ENSG00000137817	0.375402591276276
	MAPK8IP3	ENSG00000138834	0.374438312532153
	RABL2A	ENSG00000144134	0.369876242737863
	CSAD	ENSG00000139631	0.368684540320214
	PDE6C	ENSG00000095464	0.360641480464769
	PPFIA4	ENSG00000143847	0.359755271759912
	SFSWAP	ENSG00000061936	0.356548593384659
	RSRP1	ENSG00000117616	0.356136792166311
	ZNF354B	ENSG00000178338	0.355933313010074
	LUC7L3	ENSG00000108848	0.349793024558869
	RPL32P3	ENSG00000251474	0.348011865149515
	FNBP4	ENSG00000109920	0.345085300299788
	TTLL3	ENSG00000214021	0.344495358813673
	PATL2	ENSG00000229474	0.342174003395947
	ATP2A1	ENSG00000196296	0.341822429888544
	ZNF337	ENSG00000130684	0.341604980384019
	CCDC14	ENSG00000175455	0.341256742610813
	LINC00115	ENSG00000225880	0.341025099395676
	TTC31	ENSG00000115282	0.340195908016417
	L3MBTL1	ENSG00000185513	0.338013384049159
	FAM186A	ENSG00000185958	0.335570706026565
	GIGYF1	ENSG00000146830	0.335121789678199
	FAM111A	ENSG00000166801	0.33372467372198
	THOC1	ENSG00000079134	0.332749180882594
	CLEC2D	ENSG00000069493	0.332243827566089

PPP4R1L	ENSG00000124224	0.330907747697916
STX1B	ENSG00000099365	0.327423708461175
PCGF3	ENSG00000185619	0.327252207780055
ADAM20P1	ENSG00000259158	0.326452966933365
RBM33	ENSG00000184863	0.322976335646522
GSAP	ENSG00000186088	0.322192679103163
ANKRD10	ENSG00000088448	0.321707769730367
LINC01011	ENSG00000244041	0.321446483606418
CFAP44	ENSG00000206530	0.321171613726505
CBWD3	ENSG00000196873	0.321110790412793
N4BP2L1	ENSG00000139597	0.320648566483753
RAB24	ENSG00000169228	0.319897714395431
LUC7L	ENSG00000007392	0.317695342195814
PABPC1L	ENSG00000101104	0.317571730126807
NUP85	ENSG00000125450	0.315865821083773
ARGLU1	ENSG00000134884	0.315782806705869
TAF1C	ENSG00000103168	0.314771513485467
ZNF26	ENSG00000198393	0.312697425948262
AGBL2	ENSG00000165923	0.311481994742524
SIRT7	ENSG00000187531	0.310685366443851
WSB1	ENSG00000109046	0.310348495240044
ZNF37BP	ENSG00000234420	0.309859377658169
TIA1	ENSG00000116001	0.309458772482746
ZNF700	ENSG00000196757	0.308582267414893
MAN2C1	ENSG00000140400	0.308239562882223
CTAGE7P	ENSG00000233122	0.308163669983636
ZNRD1ASP	ENSG00000204623	0.307864096587296
RBM39	ENSG00000131051	0.307155602419428
N4BP2L2	ENSG00000244754	0.304292239340295
MYH3	ENSG00000109063	0.303708633020071
ANKRD36	ENSG00000135976	0.30244193650642
CCDC73	ENSG00000186714	0.301159008324609
RPAIN	ENSG00000129197	0.300918031085298
KCNAB3	ENSG00000170049	0.300110931501699
ADCY10	ENSG00000143199	0.297396621563162
NFATC2IP	ENSG00000176953	0.296264069503707
ZGRF1	ENSG00000138658	0.296119660814791
OFD1	ENSG00000046651	0.294985358997892
ZNF345	ENSG00000251247	0.294410442262471
DDX55	ENSG00000111364	0.293867435874827
SH3BP2	ENSG00000087266	0.293659121303642
PNISR	ENSG00000132424	0.292801465687039
ZNF841	ENSG00000197608	0.291980724485583
PAXBPI	ENSG00000159086	0.288171428079592
PIDD1	ENSG00000177595	0.287817086993087
NKTR	ENSG00000114857	0.287704374855512

DENND4B	ENSG00000198837	0.287561903371126
TNFRSF14	ENSG00000157873	0.287449608605925
ATAD5	ENSG00000176208	0.28721696009574
UCP3	ENSG00000175564	0.286549688281668
NASP	ENSG00000132780	0.283064883538849
CREBZF	ENSG00000137504	0.282168524687278
DNA2	ENSG00000138346	0.281601995275445
NFYC-AS1	ENSG00000272145	0.281133702170016
ADAT2	ENSG00000189007	0.279192397464162
C3orf35	ENSG00000198590	0.278890159808194
DMAP1	ENSG00000178028	0.27783823618694
GTPBP2	ENSG00000172432	0.277714260583574
SEC16B	ENSG00000120341	0.276395378617052
XAF1	ENSG00000132530	0.275006531116714
VIP	ENSG00000146469	0.272538144320519
ANKHD1	ENSG00000131503	0.27251655820975
ZNF514	ENSG00000144026	0.272331960118372
PZP	ENSG00000126838	0.271861987870332
C5orf58	ENSG00000234511	0.271830275724673
MCM3AP-AS1	ENSG00000215424	0.271666051556617
S100PBP	ENSG00000116497	0.270311745990794
ZNF169	ENSG00000175787	0.267362485155717
MAP3K12	ENSG00000139625	0.267046724638998
TTC14	ENSG00000163728	0.265174185622583
NADSYN1	ENSG00000172890	0.26470296110962
MSANTD2	ENSG00000120458	0.263038982145215
CDH26	ENSG00000124215	0.262796549782473
ZNF333	ENSG00000160961	0.262525346697153
STARD9	ENSG00000159433	0.262437013375029
MAU2	ENSG00000129933	0.261999803076422
TTN	ENSG00000155657	0.261737647098338
MSTN	ENSG00000138379	0.261071594723428
ANKRD36C	ENSG00000174501	0.258544306436502
VEGFA	ENSG00000112715	0.25830632653687
STX16	ENSG00000124222	0.258084814289468
PRKRIP1	ENSG00000128563	0.258031012457438
CAPN15	ENSG00000103326	0.257969380430998
ZNF160	ENSG00000170949	0.257663837541664
SRSF11	ENSG00000116754	0.256484493391593
GABRE	ENSG00000102287	0.255716759722439
LACTB2-AS1	ENSG00000246366	0.255635263909749
CC2D2B	ENSG00000188649	0.255630349671403
TRIP10	ENSG00000125733	0.255190503952168
GOLGA6L9	ENSG00000197978	0.255131809382063
CTAGE9	ENSG00000236761	0.254402343737278
ZNF7	ENSG00000147789	0.253963503596689

MALAT1	ENSG00000251562	0.252736287663478
ZNF182	ENSG00000147118	0.25255109870807
TNK2	ENSG00000061938	0.251970878483753
ZFC3H1	ENSG00000133858	0.251581306513179
CDK11A	ENSG00000008128	0.250762460493971
TRMT13	ENSG00000122435	0.249163438862904
DDX17	ENSG00000100201	0.248510349047098
VMP1	ENSG00000062716	0.247798932222384
TENM1	ENSG00000009694	0.247658732900494
PCBP1-AS1	ENSG00000179818	0.247243607354119
AZIN2	ENSG00000142920	0.245645853659764
ADORA2A	ENSG00000128271	0.245168230527969
MS4A14	ENSG00000166928	0.244890006290039
CCNL1	ENSG00000163660	0.243473709988561
PPP1R13L	ENSG00000104881	0.24338832367968
NPPA	ENSG00000175206	0.242856033520481
SLC25A27	ENSG00000153291	0.241950473517888
SPATA6L	ENSG00000106686	0.241629728077076
POLR2F	ENSG00000100142	0.240759051134706
KLHL31	ENSG00000124743	0.239377560638841
RFX8	ENSG00000196460	0.23842465623542
ANAPC4	ENSG00000053900	0.238022074906812
C20orf197	ENSG00000176659	0.237748066335563
SLC25A37	ENSG00000147454	0.237062085494853
DIP2A	ENSG00000160305	0.236181266457925
IFI44	ENSG00000137965	0.235890451472248
KLF8	ENSG00000102349	0.235723686614878
FPGT-TNNI3K	ENSG00000259030	0.235591359081334
NEK3	ENSG00000136098	0.235586821025149
ZNF682	ENSG00000197124	0.235543359683024
DZIP1L	ENSG00000158163	0.235365811234327
TBC1D26	ENSG00000214946	0.234069308833903
ERCC3	ENSG00000163161	0.233821378850058
LUC7L2	ENSG00000146963	0.233634201184637
NXF1	ENSG00000162231	0.232925351591902
ZRANB2	ENSG00000132485	0.232759935375699
ANKRA2	ENSG00000164331	0.232620059992187
MSH4	ENSG00000057468	0.231864535029412
ARHGEF33	ENSG00000214694	0.229630289146585
FAM76B	ENSG00000077458	0.228340001238378
PHKA2	ENSG00000044446	0.227792433993483
ADAMTS10	ENSG00000142303	0.227755379481415
PPP1R12C	ENSG00000125503	0.227635835728473
AKAP8L	ENSG00000011243	0.226917358511135
CYP2R1	ENSG00000186104	0.22683429525723
TTC32	ENSG00000183891	0.22643921013349

ZNF449	ENSG00000173275	0.224331138538939
CCDC191	ENSG00000163617	0.224149194418273
PRPF39	ENSG00000185246	0.22400825461131
ANKRD36B	ENSG00000196912	0.223756453342382
ZNF302	ENSG00000089335	0.223002758682921
THUMPD2	ENSG00000138050	0.222973589317717
OGT	ENSG00000147162	0.222610393456992
RALGDS	ENSG00000160271	0.222400060729565
E4F1	ENSG00000167967	0.222246221710651
ARPC4-TTL3	ENSG00000250151	0.220584013076514
LRR36	ENSG00000159708	0.220458020862422
MMP25	ENSG00000008516	0.220395072495104
MICA	ENSG00000204520	0.220313792581239
KLC1	ENSG00000126214	0.220171512020929
ZCCHC8	ENSG00000033030	0.219768373944153
PPEF2	ENSG00000156194	0.219166691676701
EXOSC9	ENSG00000123737	0.219130550119442
WASH2P	ENSG00000146556	0.218946032399114
POGZ	ENSG00000143442	0.218436494671897
ZNF248	ENSG00000198105	0.218264896635119
RHEBL1	ENSG00000167550	0.218165214560028
CHFR	ENSG00000072609	0.217208545036425
ZNF266	ENSG00000174652	0.216940364725758
DENND3	ENSG00000105339	0.215722875045352
UNC13D	ENSG00000092929	0.215667767928993
ORC2	ENSG00000115942	0.215624624134236
ABCA10	ENSG00000154263	0.21452489224283
NXF5	ENSG00000126952	0.213708675262134
KDM4C	ENSG00000107077	0.21365443221287
PGS1	ENSG00000087157	0.213439598653779
ZNF558	ENSG00000167785	0.213287791015048
FAM228B	ENSG00000219626	0.213263156278068
MAK	ENSG00000111837	0.21308788926532
GPCPD1	ENSG00000125772	0.212655723953724
OGFR	ENSG00000060491	0.212502828195204
CARD8	ENSG00000105483	0.211228198571471
PLXNA3	ENSG00000130827	0.210584840373864
DCAF4L1	ENSG00000182308	0.210513128852776
GABRR2	ENSG00000111886	0.210073630812149
SYNGAP1	ENSG00000197283	0.209973050968681
SLFN13	ENSG00000154760	0.209861420070145
CCNT2	ENSG00000082258	0.209707117447876
ZNF23	ENSG00000167377	0.208656772575198
GABBR1	ENSG00000204681	0.208035510544423
ZC3H12D	ENSG00000178199	0.20799273074168
SEC14L3	ENSG00000100012	0.207424918027661

SREK1	ENSG00000153914	0.207014450738194
FBXO43	ENSG00000156509	0.206968250905252
ACTR5	ENSG00000101442	0.204698074501941
CBWD6	ENSG00000215126	0.204468235398133
SLC24A4	ENSG00000140090	0.204425858801826
AP5Z1	ENSG00000242802	0.203751416087215
IL18RAP	ENSG00000115607	0.203483240515543
METTL17	ENSG00000165792	0.203331469778913
TNRC6A	ENSG00000090905	0.203141091248868
CFLAR	ENSG00000003402	0.203058076222807
CELF6	ENSG00000140488	0.20275726034174
ZNF519	ENSG00000175322	0.202689281129086
C1orf52	ENSG00000162642	0.20258016314672
ZC3H8	ENSG00000144161	0.202495365543049
MSH5-SAPCD1	ENSG00000255152	0.202437237148453
CENPJ	ENSG00000151849	0.201920904028036
SPDYE6	ENSG00000260097	0.201656333144585
SAP30BP	ENSG00000161526	0.201307842177741
CAPRIN2	ENSG00000110888	0.200320378141221
ANKRD49	ENSG00000168876	0.20024128058108
RBM5	ENSG00000003756	0.200216403156156
CEP152	ENSG00000103995	0.199941775651999
VENTX	ENSG00000151650	0.199913694389222
KANTR	ENSG00000232593	0.199152284831062
GOLGA6L10	ENSG00000278662	0.198310816083834
PRPF38B	ENSG00000134186	0.198155757552411
TIAL1	ENSG00000151923	0.1963920899256
SETD4	ENSG00000185917	0.19638232423889
LRRC39	ENSG00000122477	0.196331269633468
SRSF2	ENSG00000161547	0.195484339222115
ALPK2	ENSG00000198796	0.194115561460069
PRDM15	ENSG00000141956	0.193995159307689
MZF1-AS1	ENSG00000267858	0.193982995816445
BTAF1	ENSG00000095564	0.193857840014262
NFKB2	ENSG00000077150	0.193723552749844
ZNF252P-AS1	ENSG00000255559	0.193071054823455
EDRF1	ENSG00000107938	0.193045257294469
KMT2B	ENSG00000272333	0.192810226187281
SUPT20H	ENSG00000102710	0.192105011125933
WRAP73	ENSG00000116213	0.19204593879274
ASIC4	ENSG00000072182	0.190489342091248
KIN	ENSG00000151657	0.190471355600859
AGBL3	ENSG00000146856	0.190149494447656
TRAF5	ENSG00000082512	0.190072126285779
PAN3	ENSG00000152520	0.190009206641332
ZMIZ2	ENSG00000122515	0.189477578399749

AP4B1	ENSG00000134262	0.189356818915593
MYSM1	ENSG00000162601	0.189282211834203
LRR69	ENSG00000214954	0.188313849224158
CABP7	ENSG00000100314	0.188012890570578
TRA2A	ENSG00000164548	0.187928834963113
WDR74	ENSG00000133316	0.187864440173701
ZDHHC17	ENSG00000186908	0.187806190529233
ZNF528	ENSG00000167555	0.187525192831569
GON4L	ENSG00000116580	0.187208088547618
CRAMP1	ENSG00000007545	0.187152143210586
USP21	ENSG00000143258	0.187113392485618
FBRS	ENSG00000156860	0.186929907259064
RBM26	ENSG00000139746	0.186917116310896
UNK	ENSG00000132478	0.1850486981041
SRSF4	ENSG00000116350	0.184968113897335
HNRNPA1L2	ENSG00000139675	0.184848276702961
ZNF529	ENSG00000186020	0.184840276899951
PRSS27	ENSG00000172382	0.184422433804435
CCDC168	ENSG00000175820	0.184294270932835
NSMCE4A	ENSG00000107672	0.184000163254881
MYH15	ENSG00000144821	0.183022627874413
CCNE2	ENSG00000175305	0.182978477011021
ZNF587	ENSG00000198466	0.182945622070378
ANKHD1-EIF4EBP3	ENSG00000254996	0.18266012966367
PTBP2	ENSG00000117569	0.182341247198946
PKD1	ENSG00000008710	0.18139058599599
ZNF117	ENSG00000152926	0.181249179962916
RBM25	ENSG00000119707	0.180946809002503
CHRM5	ENSG00000184984	0.180799073899926
TRAF3IP3	ENSG00000009790	0.180613422357853
RNF44	ENSG00000146083	0.180424443707681
BTN2A2	ENSG00000124508	0.180221587818229
DONSON	ENSG00000159147	0.179808201867229
CFAP74	ENSG00000142609	0.179553330115714
GUSBP1	ENSG00000183666	0.178193881874495
SUPT7L	ENSG00000119760	0.178171892410639
DBF4B	ENSG00000161692	0.177632306489963
ZNF500	ENSG00000103199	0.177009202494131
PHF11	ENSG00000136147	0.176689715193922
USP6	ENSG00000129204	0.175867980016917
WDR75	ENSG00000115368	0.175769377890675
SZT2	ENSG00000198198	0.175512246375749
ZNF343	ENSG00000088876	0.175463691021159
BCL2L2-PABPN1	ENSG00000258643	0.175174497727047
NXNL2	ENSG00000130045	0.175153665280499
AGAP3	ENSG00000133612	0.174928687509111

KNTC1	ENSG00000184445	0.174858923257944
ERMN	ENSG00000136541	0.174800365349955
PABPN1	ENSG00000100836	0.174746070025853
FRA10AC1	ENSG00000148690	0.174161913306993
ZNF397	ENSG00000186812	0.17240443486451
PHF21A	ENSG00000135365	0.172207190434355
SBF1	ENSG00000100241	0.171994891112347
ZNF491	ENSG00000177599	0.171732841707165
ADAM11	ENSG00000073670	0.171277527888052
ATM	ENSG00000149311	0.170819043242177
MAST3	ENSG00000099308	0.170756433109761
CCDC82	ENSG00000149231	0.170426788107523
RWDD3	ENSG00000122481	0.170161686396452
TARBP1	ENSG00000059588	0.169622528192964
DSCAML1	ENSG00000177103	0.169293631472084
FAM193A	ENSG00000125386	0.168558959959797
RPL21P44	ENSG00000229585	0.168397948292752
ALPK1	ENSG00000073331	0.168357476598432
TRAPPC2	ENSG00000196459	0.168037728784886
DPY19L2P2	ENSG00000170629	0.167672199680691
RGS12	ENSG00000159788	0.16762950700889
PMS1	ENSG00000064933	0.167182781747868
CLHC1	ENSG00000162994	0.166871664873163
ATG12	ENSG00000145782	0.166493677524637
SAMD3	ENSG00000164483	0.166431674460975
ORMDL1	ENSG00000128699	0.166127863979412
OBSCN	ENSG00000154358	0.165652284368492
CRX	ENSG00000105392	0.165058926578397
FUS	ENSG00000089280	0.16469249523278
KAT8	ENSG00000103510	0.164498441478154
STAT4	ENSG00000138378	0.164376942673718
EXOSC10	ENSG00000171824	0.163020530212385
HPS4	ENSG00000100099	0.162867085667331
MTCL1	ENSG00000168502	0.162781017747668
LRIG2	ENSG00000198799	0.162523353632046
UPF3B	ENSG00000125351	0.162452648829279
UHRF2	ENSG00000147854	0.162236261078283
TET3	ENSG00000187605	0.16218086132502
HAUS3	ENSG00000214367	0.162128524620052
DGKD	ENSG00000077044	0.161979793933632
U2SURP	ENSG00000163714	0.161943085104866
SRSF5	ENSG00000100650	0.16173150026244
UBN2	ENSG00000157741	0.16156107268078
GOLGA6L4	ENSG00000184206	0.160466721262284
DDX56	ENSG00000136271	0.160160820718723
ILKAP	ENSG00000132323	0.160075761380465

ZNF76	ENSG00000065029	0.159526045972893
ZNF195	ENSG00000005801	0.159427043340516
ZSCAN30	ENSG00000186814	0.159408069897682
CNTRL	ENSG00000119397	0.158882101763075
CTC1	ENSG00000178971	0.158881794255447
DGKA	ENSG00000065357	0.15870600224084
FLT3LG	ENSG00000090554	0.158659165557022
KIF21B	ENSG00000116852	0.158025705017254
CEP290	ENSG00000198707	0.157983398837085
DCUN1D4	ENSG00000109184	0.157892040391911
USPL1	ENSG00000132952	0.157704437489166
VCPKMT	ENSG00000100483	0.157692235261583
KIAA0753	ENSG00000198920	0.15701504373597
RAPGEFL1	ENSG00000108352	0.156876320225958
MORN1	ENSG00000116151	0.156876184762449
FGR	ENSG00000000938	0.156682521440117
QRICH2	ENSG00000129646	0.156676430907573
XRCC2	ENSG00000196584	0.156157683441563
PTPN18	ENSG00000072135	0.156042733643853
MTF2	ENSG00000143033	0.155687457010504
MPZ	ENSG00000158887	0.155379532258754
TMEM259	ENSG00000182087	0.155251666172189
PLCXD1	ENSG00000182378	0.155105526944498
ARL17A	ENSG00000185829	0.154192811308577
MCF2L	ENSG00000126217	0.153718016342502
YPEL3	ENSG00000090238	0.153676348043122
PYHIN1	ENSG00000163564	0.153412230036651
RAPGEF6	ENSG00000158987	0.153207318114862
CAPN10	ENSG00000142330	0.153154321143651
INTS3	ENSG00000143624	0.153149558365675
KCNE3	ENSG00000175538	0.152881875008796
CELSR3	ENSG00000008300	0.152843356833769
SCML1	ENSG00000047634	0.152665056473207
SCLT1	ENSG00000151466	0.152655452346594
SPEF2	ENSG00000152582	0.15265508679984
RUFY3	ENSG00000018189	0.152263932782473
MYO3A	ENSG00000095777	0.152201097471025
HSPB11	ENSG00000081870	0.152149516496357
PLAGL1	ENSG00000118495	0.151471732577948
PMS2CL	ENSG00000187953	0.151286379206117
UIMC1	ENSG00000087206	0.151236185569968
CSNK1E	ENSG00000213923	0.151001352438813
RFX2	ENSG00000087903	0.150855187069698
FBXW7	ENSG00000109670	0.150617534201315
ZNF75D	ENSG00000186376	0.150448461426815
HSD17B7	ENSG00000132196	0.15018505759932

	COMMD3	ENSG00000148444	0.150182363055487
--	--------	-----------------	-------------------

TABELA SUPLEMENTAR 6: RESULTADO CO-EXPRESSÃO LIHC (p = 0,025)

LIHC			
lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
MCM3AP-AS1	PAXBP1	ENSG00000159086	0.188124795634632
	U2SURP	ENSG00000163714	0.183685967566862
	DIP2A	ENSG00000160305	0.182976291822738
	DMTF1	ENSG00000135164	0.177488059357598
	ILF3	ENSG00000129351	0.176590755132376
	FNBP4	ENSG00000109920	0.176329159676362
	PRPF39	ENSG00000185246	0.17578705669209
	ZNF207	ENSG00000010244	0.175470330577311
	MAU2	ENSG00000129933	0.173455749576296
	SUGP2	ENSG00000064607	0.172982477817105
	CPSF6	ENSG00000111605	0.17189427399611
	TET3	ENSG00000187605	0.171757517731733
	TAF4	ENSG00000130699	0.171157069599406
	CCDC93	ENSG00000125633	0.170877354270834
		ENSG00000214135	0.169945020536889
	ZNF37BP	ENSG00000234420	0.168469162771516
	MSH5	ENSG00000204410	0.168436084600045
	ZNF337	ENSG00000130684	0.167170956402828
	NAA25	ENSG00000111300	0.166859749587075
	DENND4B	ENSG00000198837	0.166777858320868
	MAML1	ENSG00000161021	0.166558232699677
	KCTD7	ENSG00000243335	0.165863762618941
	DBF4B	ENSG00000161692	0.165784819443391
	CBFA2T2	ENSG00000078699	0.16539535943182
	DDX39B	ENSG00000198563	0.16535644643963
	SOCS7	ENSG00000274211	0.164762244102317
	CEP295	ENSG00000166004	0.164618470934474
	ATXN7L3	ENSG00000087152	0.164251944892494
	KNTC1	ENSG00000184445	0.163628422107322
	TPR	ENSG00000047410	0.16361186164194
	ZNF587	ENSG00000198466	0.163366431066776
	ATAD5	ENSG00000176208	0.163225163028358
	UBAP2	ENSG00000137073	0.163027203615308
	DDX55	ENSG00000111364	0.162699648990907
	INCENP	ENSG00000149503	0.162685957475464
	PHF12	ENSG00000109118	0.162387808027165
	RBM12B	ENSG00000183808	0.161919959967356
	ZNF8	ENSG00000278129	0.161302374011323
	HECTD2	ENSG00000165338	0.161260775118713
	NKTR	ENSG00000114857	0.161117243238128
	MSANTD2	ENSG00000120458	0.160747646723089
	CEP290	ENSG00000198707	0.160325667969068
	ADNP	ENSG00000101126	0.159618963524593

EP400	ENSG00000183495	0.159486768529444
CKAP5	ENSG00000175216	0.159331851959939
NFYA	ENSG00000001167	0.159212875352432
KIF20B	ENSG00000138182	0.159201831835598
RNMT	ENSG00000101654	0.15915149544422
RBM25	ENSG00000119707	0.159144489242266
PRRC2A	ENSG00000204469	0.159047312670179
RBM39	ENSG00000131051	0.159002767793128
ZBTB40	ENSG00000184677	0.158844746109666
ANLN	ENSG00000011426	0.15853013312789
KPNB1	ENSG00000108424	0.158434122343292
DNAJC9	ENSG00000213551	0.158400054494195
MSH2	ENSG00000095002	0.158368898268229
ASXL1	ENSG00000171456	0.158346749276447
RNF44	ENSG00000146083	0.158269685458812
FAM219B	ENSG00000178761	0.15798768205723
NR2C2	ENSG00000177463	0.157969857859252
LENG8	ENSG00000167615	0.157796688849851
ZNF605	ENSG00000196458	0.157721636610258
YEATS2	ENSG00000163872	0.157619245875235
CLK2	ENSG00000176444	0.157477876649206
SETD5	ENSG00000168137	0.157167309080015
TOPBP1	ENSG00000163781	0.157162402758635
LUC7L3	ENSG00000108848	0.157039542756735
TIA1	ENSG00000116001	0.156969211922565
CCAR1	ENSG00000060339	0.15694898451237
KANSL1	ENSG00000120071	0.156868141558725
CCDC14	ENSG00000175455	0.156747272006466
MED12	ENSG00000184634	0.156705672991552
ZFC3H1	ENSG00000133858	0.156504760084201
XRCC2	ENSG00000196584	0.156260626375362
DNA2	ENSG00000138346	0.156246165871813
CEP192	ENSG00000101639	0.156146959748312
ZNF621	ENSG00000172888	0.155969466944673
ZGRF1	ENSG00000138658	0.155966410342563
BUB1B	ENSG00000156970	0.155931900272834
ZNF169	ENSG00000175787	0.155823056203628
SART3	ENSG00000075856	0.155680812551013
GPATCH8	ENSG00000186566	0.155656990270195
MSL1	ENSG00000188895	0.155539789182603
PAXIP1	ENSG00000157212	0.155195063312046
DGKD	ENSG00000077044	0.154933612198201
SENP1	ENSG00000079387	0.154868222900707
RBM12	ENSG00000244462	0.154828939474059
ZNF248	ENSG00000198105	0.154751326822119
NCOA6	ENSG00000198646	0.154731441552104

L3MBTL1	ENSG00000185513	0.154663983631897
RBM33	ENSG00000184863	0.15458468926774
DHX35	ENSG00000101452	0.154514434970138
RBM5	ENSG00000003756	0.154509674327744
ZNF142	ENSG00000115568	0.154504083419215
POGZ	ENSG00000143442	0.154457754102406
PRPF4B	ENSG00000112739	0.154273158886524
HNRNPA2B1	ENSG00000122566	0.154199424120532
KIF23	ENSG00000137807	0.154182831684319
NUP205	ENSG00000155561	0.154130459798562
CENPF	ENSG00000117724	0.154024411577195
ANKRD52	ENSG00000139645	0.153842575193807
TSC1	ENSG00000165699	0.153776053280507
ORC2	ENSG00000115942	0.153576338891391
TBC1D12	ENSG00000108239	0.153440974852241
RALGAPB	ENSG00000170471	0.153418014289863
DIDO1	ENSG00000101191	0.153240615117401
CENPE	ENSG00000138778	0.153163060976456
PMS1	ENSG00000064933	0.15312462094401
SF3B1	ENSG00000115524	0.153115907695107
ZNF37A	ENSG00000075407	0.153065011019694
KANSL2	ENSG00000139620	0.152940308209241
GPSM2	ENSG00000121957	0.152658312361122
GON4L	ENSG00000116580	0.152488877809557
STIL	ENSG00000123473	0.152480785241286
KMT2D	ENSG00000167548	0.152386385184232
SREK1	ENSG00000153914	0.152286287225549
XPO1	ENSG00000082898	0.152211630583265
PRDM10	ENSG00000170325	0.152153159292742
QSER1	ENSG00000060749	0.151910738191941
UBAP2L	ENSG00000143569	0.151540627255764
FOXM1	ENSG00000111206	0.151319136293824
LSM14B	ENSG00000149657	0.151220867163236
ZSCAN25	ENSG00000197037	0.151140614549845
CRLF3	ENSG00000176390	0.151035813671324
STX16	ENSG00000124222	0.151024189273049
KMT2B	ENSG00000272333	0.150916790861007
SRSF11	ENSG00000116754	0.150790245168156
CNOT6	ENSG00000113300	0.150720989823936
MAPKBP1	ENSG00000137802	0.150424831484245
RBM28	ENSG00000106344	0.150368288832291
SFSWAP	ENSG00000061936	0.150293696696659
CAD	ENSG00000084774	0.150211788932218
DCAF16	ENSG00000163257	0.150130173190545
NSD1	ENSG00000165671	0.150117538619697
ABCC10	ENSG00000124574	0.150109061124854

ZNF26	ENSG00000198393	0.150077668778153
DHX9	ENSG00000135829	0.1500098071855

TABELA SUPLEMENTAR 7: RESULTADO CO-EXPRESSÃO LUAD (p = 0,003)

LUAD			
lncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
MCM3AP-AS1	DIP2A	ENSG00000160305	0.209361737764881
	PAXBP1	ENSG00000159086	0.206054090235826
	PRDM15	ENSG00000141956	0.197624271885805
	RUSC1-AS1	ENSG00000225855	0.196176440120321
	ZNRD1ASP	ENSG00000204623	0.191923950954522
	MSH5	ENSG00000204410	0.190648974324805
	LUC7L3	ENSG00000108848	0.184905630626936
	RBM33	ENSG00000184863	0.184887226198263
	DDX39B	ENSG00000198563	0.179681162071099
	RBM25	ENSG00000119707	0.178665155315322
	TARBP1	ENSG00000059588	0.176123564694017
	CTAGE7P	ENSG00000233122	0.175377280365787
	DMTF1	ENSG00000135164	0.174557146200256
	RPL32P3	ENSG00000251474	0.173317440723576
	FNBP4	ENSG00000109920	0.169605191169103
	MYSM1	ENSG00000162601	0.169186647527297
	UBN2	ENSG00000157741	0.168959492920484
	NKTR	ENSG00000114857	0.168305374712232
	NFYC-AS1	ENSG00000272145	0.16824697208262
	RBM39	ENSG00000131051	0.167957855378946
	PCNT	ENSG00000160299	0.167877775702427
	CCDC14	ENSG00000175455	0.167789807821302
	PRPF39	ENSG00000185246	0.1642621470327
	SRSF11	ENSG00000116754	0.16358500218809
	CCNL2	ENSG00000221978	0.163482976020799
	CREBZF	ENSG00000137504	0.162729801142655
	PABPC1L	ENSG00000101104	0.162170526254401
	MDM4	ENSG00000198625	0.16192999603933
	GON4L	ENSG00000116580	0.161045435386705
	PNN	ENSG00000100941	0.160341541573521
	ZNF37BP	ENSG00000234420	0.158617827599451
	SS18L1	ENSG00000184402	0.158290138396448
	L3MBTL1	ENSG00000185513	0.156894483330721
	ZNF337	ENSG00000130684	0.156618230421769
	CLK2	ENSG00000176444	0.156322293435989
	S100PBP	ENSG00000116497	0.15600201429163
	ZNF169	ENSG00000175787	0.155851478235526
	SUGP2	ENSG00000064607	0.155789074750746
	NFATC2IP	ENSG00000176953	0.154845437673502
	ZNF514	ENSG00000144026	0.154212127091893
	PRPF3	ENSG00000117360	0.154113585680267
	EDRF1	ENSG00000107938	0.153836948349606
	POGZ	ENSG00000143442	0.15357591314927

ZC3H8	ENSG00000144161	0.153565203196492
TNRC6A	ENSG00000090905	0.153329633036071
PCGF3	ENSG00000185619	0.15113172435516
BRWD1	ENSG00000185658	0.150797725751078

TABELA SUPLEMENTAR 8: RESULTADO CO-EXPRESSÃO PRAD (p = 0,01)

PRAD			
IncRNA	Gene Alvo de PUM	Ensembl ID	co-expressão
LINC00115	CCNL2	ENSG00000221978	0.23740146879768
	MSH5	ENSG00000204410	0.219291772666359
	SLC25A27	ENSG00000153291	0.210510877119948
	RSRP1	ENSG00000117616	0.207707923394162
	CSAD	ENSG00000139631	0.200071114316384
	DDX39B	ENSG00000198563	0.19839816575215
	KIAA0895L	ENSG00000196123	0.195220163220195
	RUSC1-AS1	ENSG00000225855	0.193882841446249
	LENG8	ENSG00000167615	0.182441933192242
	NFYC-AS1	ENSG00000272145	0.181641363693323
	PABPN1	ENSG00000100836	0.180325116273024
	CLK2	ENSG00000176444	0.178090511900437
	RBM39	ENSG00000131051	0.178088217856548
	MSH5-SAPCD1	ENSG00000255152	0.175549457854274
	PCGF3	ENSG00000185619	0.174574809672295
	ZNF26	ENSG00000198393	0.173937549800184
	ANKZF1	ENSG00000163516	0.172560689982254
	ZNF76	ENSG00000065029	0.172381400903885
	FAM227A	ENSG00000184949	0.168698357004205
	ANKRD10	ENSG00000088448	0.16740387967031
	WDR90	ENSG00000161996	0.166455250792663
	LEAP2	ENSG00000164406	0.165778320972252
	RBM5	ENSG00000003756	0.165271963551798
	ARGLU1	ENSG00000134884	0.163799095994021
	SPATA24	ENSG00000170469	0.163097398688081
	SH2B1	ENSG00000178188	0.163086057583627
	OGT	ENSG00000147162	0.162424139451446
	MZF1-AS1	ENSG00000267858	0.161662563835797
	CLK1	ENSG00000013441	0.160669549436228
	CTAGE7P	ENSG00000233122	0.160477529112637
	PNN	ENSG00000100941	0.16038376842221
	TRIM52	ENSG00000183718	0.160032585451442
	PHYKPL	ENSG00000175309	0.159597944506123
	LUC7L3	ENSG00000108848	0.159326688829422
	LUC7L	ENSG00000007392	0.15623543854741
	MAPK8IP3	ENSG00000138834	0.155741390222736
	PRPF39	ENSG00000185246	0.155443613944322
	ZNF559-ZNF177	ENSG00000270011	0.153663906694728
	L3MBTL1	ENSG00000185513	0.153213815615804
	PRPF3	ENSG00000117360	0.153169322218172
	MSANTD2	ENSG00000120458	0.150650640955479
C8orf34-AS1	C8orf34	ENSG00000165084	0.294033253180102
RHPN1-AS1	ARHGAP39	ENSG00000147799	0.156617478961215

ZNF252P-AS1	CSAD	ENSG00000139631	0.166564971519545
	LEAP2	ENSG00000164406	0.15175336634443
	RSRP1	ENSG00000117616	0.151631239324304

TABELA SUPLEMENTAR 9: GENES ALVOS DE PUM
DIFERENCIAMENTE EXPRESSOS $\log FC > 0$

Câncer	lncRNA co-expresso	Gene Alvo
BLCA	LINC00115	CCNL2
COAD	MCM3AP-AS1	CCDC14 FNBP4 DDX39B PAXBP1 MSH5 NKTR RUSC1-AS1 ADAT2 TARBP1 ZNF37BP SUGP2 ZFC3H1 OGT PNISR CREBZF ZNF621 CCDC150 SLC25A27 RBM33 LUC7L3 ZNF514
	PAN3-AS1	PAN3 SUPT20H STX16 RBM39 USPL1 UPF3A ANKRD10 MTRF1 SLC25A27 CENPJ RBM26 CDC16 NHLRC3
ESCA	C15orf54	PCDH17 CEP135 ESM1 TMEM26
HNSC	IGF2-AS	ITGA11 ASPN MMP11 COL11A1

		LRRC15 SGIP1 COL3A1 NOX4 LUM CTHRC1 POSTN P4HA3 EFEMP2 COL8A1 HEPH
	IGF2BP2-AS1	IGF2BP2
	LINC01484	PIM2
	LINC01619	BTG1 CLEC2D
	LINC01343	POU3F1
	LINC00115	CCNL2 KIAA0895L PABPC1L RSRP1 GABBR1 DDX39B PRPF39 RUSC1-AS1 MSH5 TAF1C CLK2 RBM39 CLK1 LENG8 CSAD CCNL1 VMP1 CCDC150 E4F1 CDK11A AGER SRSF11 DMTF1
KIRC	LINC01011	HCG27 CLK2 PPFIA4 AGER PRPF3 MSH5 ANKZF1

	WDR90 CLK1 CCNL2 CLK4 PHYKPL ANKRD10 PARP6 KIAA0895L SIRT7 AMY2B SFSWAP DMTF1 MAPK8IP3 NUP85 RUSC1-AS1 ZNF276 TTC31 PCGF3 PPP4R1L INTS6L TRIM52 RAB24 FNBP4 GTPBP2 CCDC14 ZNF337 THOC1 RSRP1 VEGFA
MALAT1	RUSC1-AS1 NKTR TTLL3 LENG8 CSAD LUC7L AMY2B MSH5 MDM4 CCNL2 RSRP1 DMTF1 RPL32P3 MAPK8IP3 TRIM52 FNBP4 RBM33

PABPC1L
KIAA0895L
ZNF841
ZNF276
CCNL1
ZNF337
RRN3P1
ARGLU1
AGER
ANKRD36
THOC1
MAN2C1
PCGF3
ZNF354B
CCDC14
ANKRD10
CLK1
PRPF3
TIA1
OFD1
HCG27
CLK2
INTS6L
SFSWAP
GSAP
PARP6
ZGRF1
ZFC3H1
DDX55
PIDD1
RUFY3
PPP4R1L
RIMKLB
ANKRD36B
STX16
PATL2
CDK11A
CLK4
GOLGA6L4
GABRE
VMP1
DIP2A
PHYKPL
MAP3K12
ZNF160
NADSYN1

	STARD9
	DNA2
	PPFIA4
HCG27	MSH5
	PRPF3
	AGER
	RUSC1-AS1
	CLK4
	TRIM52
	CCNL2
	RRN3P1
	CLK2
	PHYKPL
	KIAA0895L
	AMY2B
	CLK1
	ZNF276
	LENG8
	DMTF1
	MDM4
	WDR90
	INTS6L
	ANKZF1
	PARP6
	MAPK8IP3
	CSAD
	PPFIA4
	SFSWAP
	RSRP1
	ZNF354B
	RPL32P3
	FNBP4
	TTLL3
	PATL2
	ZNF337
	CCDC14
	TTC31
	THOC1
	CLEC2D
	PPP4R1L
	STX1B
	PCGF3
	RBM33
	GSAP
	ANKRD10
	N4BP2L1

RAB24
LUC7L
PABPC1L
NUP85
ARGLU1
SIRT7
TIA1
MAN2C1
ANKRD36
ZGRF1
OFD1
DDX55
SH3BP2
ZNF841
PIDD1
NKTR
DENND4B
TNFRSF14
DNA2
GTPBP2
SEC16B
XAF1
S100PBP
MAP3K12
NADSYN1
STARD9
VEGFA
STX16
CAPN15
ZNF160
GABRE
TRIP10
CTAGE9
MALAT1
TNK2
ZFC3H1
CDK11A
VMP1
TENM1
PCBP1-AS1
ADORA2A
MS4A14
CCNL1
PPP1R13L
RFX8
SLC25A37

DIP2A
IFI44
KLF8
FAM76B
PHKA2
ADAMTS10
RALGDS
LRRC36
MMP25
ZCCHC8
ZNF248
RHEBL1
CHFR
DENND3
UNC13D
OGFR
CARD8
PLXNA3
SLFN13
ZC3H12D
AP5Z1
IL18RAP
SAP30BP
ALPK2
NFKB2
KMT2B
AGBL3
TRAF5
ZMIZ2
WDR74
GON4L
FBR5
CCNE2
PKD1
TRAF3IP3
RNF44
BTN2A2
DONSON
NXNL2
AGAP3
KNTC1
PHF21A
SBF1
ATM
MAST3
TARBP1

		DSCAML1 DPY19L2P2 RGS12 ATG12 SAMD3 OBSCN STAT4 HPS4 MTCL1 TET3 DGKD GOLGA6L4 DDX56 DGKA FLT3LG KIF21B RAPGEFL1 FGR PTPN18 MCF2L PYHIN1 KCNE3 RUFY3 MYO3A PLAGL1 CSNK1E RFX2 HSD17B7
LIHC	MCM3AP-AS1	PAXBP1 U2SURP DMTF1 ILF3 FNBP4 PRPF39 ZNF207 SUGP2 CPSF6 TAF4 CCDC93 ZNF37BP MSH5 ZNF337 NAA25 DENND4B MAML1 KCTD7

LUAD	MCM3AP-AS1	MSH5 DDX39B TARBP1 CCNL2 CREBZF PABPC1L GON4L CLK2 PRPF3 ZC3H8 PCGF3
PRAD	LINC00115	CCNL2 MSH5 SLC25A27 RSRP1 CSAD KIAA0895L LENG8 MSH5-SAPCD1 ANKZF1 ANKRD10 WDR90 ARGLU1 OGT