

# RESSALVA

Atendendo solicitação do(a)  
autor(a), o texto completo desta tese  
será disponibilizado somente a partir  
de 10/02/2024.



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
*Campus de Botucatu*



# INVESTIGAÇÃO SOBRE O PAPEL DOS RNAs LONGOS NÃO CODIFICANTES NO CONTEXTO DO SARCOMA DE EWING

**RAFAEL LUIZ BUOGO COAN**

**Botucatu - SP  
2022**

Instituto de Biociências – Seção Técnica de Pós-Graduação  
Distrito de Rubião Júnior s.n CEP 18618-970 Botucatu SP Brasil  
[posgraduacao@ibb.unesp.br](mailto:posgraduacao@ibb.unesp.br)



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
*Campus de Botucatu*



UNIVERSIDADE ESTADUAL PAULISTA  
"Julio de Mesquita Filho"

INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

## INVESTIGAÇÃO SOBRE O PAPEL DOS RNAs LONGOS NÃO CODIFICANTES NO CONTEXTO DO SARCOMA DE EWING

**RAFAEL LUIZ BUOGO COAN**

**PROF. DR. NEY LEMKE**

Tese apresentada ao Instituto de Biociências, Câmpus de Botucatu, UNESP, para obtenção do título de Doutor no Programa de Pós-Graduação em Ciências Biológicas (Genética).

**Botucatu - SP  
2022**

Instituto de Biociências – Seção Técnica de Pós-Graduação  
Distrito de Rubião Júnior s.n CEP 18618-970 Botucatu SP Brasil  
[posgraduacao@ibb.unesp.br](mailto:posgraduacao@ibb.unesp.br)

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

Coan, Rafael Luiz Buogo.

Investigação sobre o papel dos RNAs longos não codificantes no contexto do sarcoma de Ewing / Rafael Luiz Buogo Coan. - Botucatu, 2022

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Ney Lemke

Capes: 20205007

1. Sarcoma de Ewing. 2. Sequenciamento de nucleotídeo. 3. RNA Longo não Codificante. 4. Tumores dos tecidos moles. 5. Ossos - Câncer.

Palavras-chave: HDAC; RNA-Seq; Sarcoma de Ewing; lncRNA.

# Agradecimentos

Agradeço ao Prof. Dr. Ney Lemke, meu orientador, por todas as orientações e suporte durante o período de doutorado, além das valiosas discussões sobre os mais variados assuntos, não só científicos. Também à Dra. Marialva Sinigaglia pelas orientações e direcionamentos do período.

Aos colegas de laboratório pelo apoio durante o período de doutorado, em especial a Luiz Bovolenta, pelas várias conversas técnicas e sobre a vida acadêmica.

A todo pessoal do Departamento de Biofísica, assim como à pós-graduação em Ciências Biológicas (Genética), Instituto de Biociências de Botucatu e UNESP pelo suporte institucional. Agradecimentos à CAPES e CNPQ (processo 142129/2018-6) pelo apoio institucional e financeiro. Esta pesquisa foi possível com auxílio do “*Intramural Research Program of the National Institute of Health and National Cancer Institute*” e os dados acessados pelo “*NIH database for Genotypes and Phenotypes (dbGaP)*” com os números de acesso phs000768.v1.p1 e phs000804.v1.p1.

A todos os pacientes que disponibilizaram informações pessoais e seu material genético para estudos como este.

Agradecimentos a todos os amigos que compartilharam a jornada da pós-graduação conosco, muitos no mesmo barco. Aos amigos Fábio Patroni e Bruno Fantinatti por compartilhar a paixão pela computação.

Agradeço à família, em especial a minha esposa Bárbara pelo companheirismo, paciência e suporte durante todos os anos que estamos juntos, principalmente nos tempos de pandemia. Especial agradecimento também a pequena Helena, que iluminou nosso caminho nesses últimos dois anos e transformou nossas vidas.

# Resumo

O sarcoma de Ewing (SE) é um câncer que tipicamente afeta ossos e tecidos moles, especialmente em crianças e adultos jovens. As células neoplásicas do SE apresentam uma translocação recorrente entre os genes *EWSR1* e *FLII*, localizados respectivamente nos cromossomos 11 e 22. Esta translocação resulta na proteína fusionada EWS-FLI1, que age como um fator de transcrição aberrante desregulando a expressão de ampla gama de genes. Por exemplo, via ligação em repetições GGAA presentes de genes codificadores de proteínas, EWS-FLI1 recruta histonas deacetilases (HDACs), enzimas que modulam epigeneticamente a transcrição de genes codificadores e não codificadores, como os RNAs longos não codificantes (lncRNA). Estes atuam sob diversas formas, por exemplo, inibindo complexos enzimáticos, interferindo diretamente no DNA ou com outras moléculas de RNA. Como característica em comum, lncRNAs e HDACs participam em complexos remodeladores de cromatina (e.g. NuRD), assim regulando a expressão gênica globalmente. Estudos experimentais e clínicos recentemente reportaram que inibidores de HDACs (HDACi) têm promissor potencial antineoplásico contra células de SE. No entanto, pouco é conhecido sobre a associação entre os moduladores epigenéticos e os lncRNA no contexto do SE. Este estudo emprega a biologia de sistemas para investigar as relações entre as principais formas da translocação EWS-FLI1 encontradas no SE e moléculas envolvidas em mecanismos conhecidos de modulação epigenética da expressão gênica, como no remodelamento de cromatina (HDACs) e regulação pós-transcricional (lncRNA). Objetivamos, além, estabelecer um *pipeline* computacional para identificar novos transcritos não codificantes no SE. Por meio de dados de RNA-Seq provenientes de tumores sequenciados de SE, depositados no dbGaP, calculamos a expressão gênica de codificadores e não codificadores de proteína. Após, construímos uma rede de coexpressão que revela as associações entre lncRNA conhecidos e HDACs. O enriquecimento de vias e o uso de banco de dados auxilia-nos na interpretação dos resultados nas diferentes etapas do estudo. Ademais, encontramos lncRNA candidatos a novos transcritos no contexto do SE. Por fim, nosso estudo estabelece conexões entre HDACs e lncRNA, além de encontrar novos transcritos associados ao SE.

**Palavras-chave:** Sarcoma de Ewing, lncRNA, HDAC, RNA-Seq

# Abstract

Ewing sarcoma (ES) is a cancer that usually affects bones and soft tissues, especially in children and young adults. Neoplastic SE cells have a recurrent translocation between EWSR1 and FLI1, genes that are located on chromosomes 11 and 22, respectively. Such translocation results in the fusion protein EWS-FLI1, which acts as an aberrant transcriptional factor deregulating genome-wide gene expression. For example, through binding with GGAA repetitions present in protein-coding genes, EWS-FLI1 recruits histone deacetylases (HDAC), enzymes that epigenetically modulate gene expression of both protein-coding and non-coding genes, such as long non-coding RNAs (lncRNA). These act in many ways, for example, by inhibiting enzymatic complexes, directly interfering with DNA or other RNA molecules. As a common characteristic, lncRNAs and HDACs participate in chromatin remodeling complexes (e.g., NuRD), thus regulating gene-wide expression. Experimental and clinical studies report that HDACs inhibitors (HDACi) have promising potential as antineoplastic drugs against ES cells. However, little is known about the relationship between epigenetic modulators and lncRNA in the context of ES. This study employs systems biology to investigate the association between the main forms of EWS-FLI1 translocation and the molecules involved in known epigenetic modulation of gene expression, such as chromatin remodeling (HDACs) and post-transcriptional regulation (lncRNA). Besides, our objective is to establish a computational pipeline to identify new non-coding transcripts on ES. With RNA-Seq data from sequenced patient tumors of dbGaP, we calculated the expression of coding and non-coding genes. We constructed a coexpression network that shows the associations between known lncRNA and HDACs. Pathway and gene ontology enrichment analysis, together with external databases, helped us interpret the results. In addition, we found new candidates for lncRNA on ES. Therefore, our study establishes connections between HDACs and lncRNA, besides finding new transcripts associated with ES.

## Lista de Ilustrações

- Figura 3-1.** Dendograma com a distância entre as amostras usadas no estudo. **a)** dendrograma inicial; **b)** dendrograma após a aplicação do ponto de corte.....28
- Figura 3-2.** Dendograma com as amostras (quadro superior) e suas características genotípicas/fenotípicas (quadro inferior). As diversas características estão codificadas com cores (tabela detalhada no anexo – Tabela 9-1). Sex: sexo feminino em azul e masculino em ciano; Age: diversas cores de acordo com a idade; Type: tipo de translocação – I em azul, II em vermelho e outras em amarelo; STAG2: alteração gênica em azul; CDKN2A: deleção homozigótica em azul e perda de exon em vermelho; TP53: mutação em azul; BRCA2: mutação em azul; Location: diversas cores de acordo com a localização dos tumores.....29
- Figura 3-3.** Gráficos para escolha do valor  $\beta$ . **a)** análise da topologia da rede para diversos valores de  $\beta$ . O painel da esquerda mostra o índice para o modelo livre de escala (eixo y) em função dos valores de  $\beta$  (eixo x), enquanto o painel da direita mostra a conectividade média (eixo y) em função de  $\beta$  (eixo x). **b)** graus de conectividade da rede. O painel da esquerda mostra o histograma dos graus, enquanto o painel da direita mostra um gráfico log-log deles. ....30
- Figura 4-1.** Gráficos com os dados brutos de Brohl et al 2014 (phs000768) e Crompton et al 2014 (phs000804). **a)** *boxplot* com os *counts* brutos, transformados em escala logarítmica para melhor visualização. **b)** gráfico RLE indicando a falta de normalização. ....35
- Figura 4-2.** Gráficos com os dados normalizados de Brohl et al 2014 (phs000768) e Crompton et al 2014 (phs000804). **a)** *boxplot* com os dados transformados por VST. **b)** gráfico RLE evidenciando os dados normalizados. ....36
- Figura 4-3.** Expressão das 11 HDACs de interesse em amostras de SE provenientes dos estudos Brohl et al 2014 (phs000768) e Crompton et al 2014 (phs000804). ....37
- Figura 4-4.** 20 lncRNA com expressão no SE e seus valores de correlação de Pearson com as 11 HDAC. A cor do ponto indica a força da correlação, enquanto os pontos triangulares indicam correlações com valor *p* significativo (menor que 0,05). ....38
- Figura 4-5.** Representação da região genômica da *HDAC7* e *HDAC10* indicando os lncRNA com correlação *in-cis* (quadros em vermelho). **a)** e **b)** *HDAC7* e seus respectivos lncRNA. **c)** *HDAC10* e seus respectivos lncRNA. ....39

<b>Figura 4-6.</b> Módulos resultantes da clusterização da rede de coexpressão. O quadro superior mostra o dendrograma com todos os genes presentes no estudo. O quadro inferior exhibe os módulos antes da fusão através dos <i>eigengenes</i> (“ <i>Dynamic Tree Cut</i> ”) e após esta (“ <i>Merged dynamic</i> ”).	40
<b>Figura 4-7.</b> Correlações entre as expressões dos módulos e as características genotípicas/fenotípicas disponíveis. Valores em vermelho indicam correlações positivas, enquanto os em azul correspondem às negativas. No quadro estão os valores das correlações e em parênteses os valores p.	42
<b>Figura 4-8.</b> Ilustração dos genes que são coexpressos diretamente a <i>HDAC2</i> em nossa rede final. Em laranja estão os genes codificadores de proteínas, em verde os lncRNA <i>antisense</i> e em azul os demais lncRNA. A espessura das arestas é proporcional ao peso das conexões.	45
<b>Figura 4-9.</b> Ilustração dos genes remanescentes com a <i>HDAC2</i> após clusterização em comunidades. Em laranja estão os genes codificadores de proteínas, em verde os lncRNA <i>antisense</i> e em azul os demais lncRNA. A espessura das arestas é proporcional ao peso das conexões.	46
<b>Figura 4-10.</b> Expressão de alguns lncRNA e mRNA coexpressos com <i>HDAC2</i> em dois conjuntos de dados indicados nos quadros. Os primeiros seis genes representam lncRNA, enquanto os demais são mRNA.	48
<b>Figura 4-11.</b> Localização e expressão do lncRNA <i>AC020900.1</i> . <b>a)</b> localização no cromossomo 5, <i>antisense</i> ao gene <i>CAST</i> . O lncRNA está destacado em vermelho. <b>b)</b> Expressão no LncExpDB. O quadro a esquerda mostra a expressão em tecidos normais, enquanto o quadro da direita exhibe os valores em linhagens celulares. <b>c)</b> expressão em linhagens celulares de câncer, incluindo SE.	49
<b>Figura 4-12.</b> Localização e expressão do lncRNA <i>AC138207.4</i> . <b>a)</b> localização no cromossomo 17, no intron do gene <i>RNF135</i> . O lncRNA está destacado em vermelho. <b>b)</b> Expressão no LncExpDB. O quadro a esquerda mostra a expressão em tecidos normais, enquanto o quadro da direita exhibe os valores em linhagens celulares. <b>c)</b> expressão em linhagens celulares de câncer, incluindo SE.	51
<b>Figura 4-13.</b> Resumo do <i>pipeline</i> computacional empregado para encontrar transcritos candidatos a novos lncRNA associados ao SE. O detalhamento dos passos tomados encontra-se na seção “Material e métodos”.	56
<b>Figura 4-14.</b> Ilustração da intersecção entre dois preditores de potencial codificante em transcritos recém montados. Somente os 459 transcritos comuns foram considerados.	57

<b>Figura 4-15.</b> <i>Heatmaps</i> exibindo os valores de expressão dos lncRNA candidatos. <b>a)</b> gráficos separados por estudo. <b>b)</b> gráfico com os estudos agrupados. Em geral, os lncRNA apresentam expressão consistente entre amostras.....	58
<b>Figura 4-16.</b> Valores de expressão dos lncRNA candidatos <i>MSTRG.36401</i> e <i>MSTRG.36401</i> em dois estudos. Nota-se expressão maior que lncRNA conhecidos como <i>EWSAT1</i> e <i>HOTAIR</i> . .....	59
<b>Figura 4-17.</b> Região do lncRNA <i>MSTRG.36401</i> mostrando seus dois transcritos (quadro em vermelho), além de outros transcritos na região (parte inferior). Nenhum exon é sobreposto na mesma orientação.....	59
<b>Figura 4-18.</b> Região do lncRNA <i>MSTRG.22951</i> mostrando seus dois transcritos (quadro em vermelho), além de outros transcritos na região (parte inferior). Nenhum exon é sobreposto na mesma orientação.....	60
<b>Figura 7-1.</b> Clusterização dos <i>eigengenes</i> dos módulos da rede de coexpressão. A linha vermelha indica o ponto de corte para fusão dos módulos. ....	81
<b>Figura 7-2.</b> Regressão linear entre os diferentes pares de HDAC-lncRNA encontrados em janelas de 100kb às HDACs. ....	82
<b>Figura 7-3.</b> Correlações entre as amostras dos dois estudos analisados. No eixo vertical, amostras de Brohl et al 2014 e no eixo horizontal amostras de Crompton et al 2014. Valores em azul denotam altas correlações positivas. ....	87

# Lista de Tabelas

<b>Tabela 3-1.</b> Resumo das amostras do estudo phs000768.v2.p1 escolhidas para o trabalho. Run: código no <i>sequence read archive</i> (SRA); Amostra: nome da amostra dada pelos autores do estudo; Tipo: tipo de translocação de EWS-FLI1 (“Outra” refere-se a translocações tipo III e demais); Origem: origem tecidual ou celular; Sexo: sexo do paciente.....	19
<b>Tabela 3-2.</b> Resumo das amostras do estudo phs000804.v1.p1 escolhidas para o trabalho. Run: código no <i>sequence read archive</i> (SRA); Amostra: nome da amostra dada pelos autores do estudo; Tipo: tipo de translocação de EWS-FLI1 (“Outra” refere-se a translocações tipo III e demais); Origem: origem tecidual ou celular; Sexo: sexo do paciente.....	20
<b>Tabela 3-3.</b> Principais referências usadas em nosso estudo. Arquivo: nome do arquivo; Conteúdo: breve descrição do tipo de conteúdo do arquivo. ....	21
<b>Tabela 3-4.</b> Arquivos para anotação customizada do enriquecimento de ontologias. ....	32
<b>Tabela 4-1.</b> Identificadores, correlação de Pearson e valor <i>p</i> dos pares HDAC-lncRNA com correlações significativas. Os tipos referem-se à anotação padrão.....	38
<b>Tabela 4-2.</b> Módulos da rede de coexpressão com as HDACs estudadas. Ensembl: identificador do Ensembl; HGNC: nome do gene; Mod_number: número do módulo em nossas análises; Mod_color: cor do módulo.....	41
<b>Tabela 4-3.</b> Módulos encontrados na rede de coexpressão, seu número total de genes e seu hub. O módulo “grey” é composto por genes não agrupados em módulos. Módulo: nome do módulo proveniente do pacote WGCNA; Número de genes: total de genes no módulo; Hub Gene: principal hub do módulo; Gene ID: identificador Ensembl do módulo; Hub Gene Type: biotipo de gene. ....	43
<b>Tabela 4-4.</b> Estatística descritiva das características dos lncRNA candidatos, em especial número de exons e tamanho dos transcritos. ....	57
<b>Tabela 7-1.</b> Características genótípicas e fenótípicas consideradas no estudo. Sex: sexo; Age: idade; STAG2: presença ou ausência de mutação em <i>STAG2</i> ; CDKN2A: deleção homozigótica ou perda de exon em <i>CDKN2A</i> ; TP53: mutação em <i>TP53</i> ; BRCA2: mutação em <i>BRCA2</i> ; Location: localização do tumor no paciente. ....	80
<b>Tabela 7-2.</b> Valores encontrados para diferentes $\beta$ . ....	81
<b>Tabela 7-3.</b> Tabelas suplementares, que devido ao tamanho, foram depositas em <a href="https://github.com/rlbc/suplementar">https://github.com/rlbc/suplementar</a> . O endereço eletrônico abaixo leva diretamente ao arquivo em formato <i>raw</i> , pronto para ser baixado. Arquivos em csv podem ser visualizados diretamente	

no GitHub (via <https://github.com/rlbc/suplementar>), enquanto os xlsx precisam vistos no computador do usuário.....83

**Tabela 7-4.** Lista de genes que são coexpressos diretamente com a HDAC2 em nossa rede final, após várias etapas de filtragem. ....84

**Tabela 7-5.** Interações entre os lncRNA *AC020900.1* (ENST00000502568.1 - ENSG00000249180.1 - AC020900.1-201 - 565) e *AC138207.4* (ENST00000580979.1 - ENSG00000264456.1 - AC138207.4-201 - 1661) e proteínas selecionadas, encontradas por meio do lncPro. “*Score*” refere-se à qualidade da interação predita. Os valores entre parênteses são, na sequência: identificador do transcrito, identificador do gene, nome do transcrito e seu tamanho.....86

## Lista de abreviaturas e siglas

BALF	<i>bronchoalveolar lavage fluid</i>
ESCA	<i>esophageal carcinoma</i>
FPKM	<i>fragments per kilobase of transcript per million fragments mapped</i>
HGNC	<i>HUGO gene nomenclature committee</i>
kb	kilo bases
miRNA	<i>micro RNA</i>
mRNA	<i>messenger RNA</i>
OV	<i>ovarian serous cystadenocarcinoma</i>
pb	pares de bases
PBMC	<i>peripheral blood mononuclear cell</i>
PPI	<i>protein protein interaction</i>
PRAD	<i>prostate adenocarcinoma</i>
RNA-Seq	<i>RNA sequencing</i>
ROC	<i>receiver operating characteristic</i>
SKCM	<i>skin cutaneous melanoma</i>
TCGA	<i>The Cancer Genome Atlas</i>
TMM	<i>trimmed mean of m values</i>
TPM	<i>transcripts per million</i>

# Sumário

<b>1</b>	<b><u>INTRODUÇÃO</u></b>	<b>4</b>
1.1	SARCOMA DE EWING	4
1.2	RNAS LONGOS NÃO CODIFICANTES	7
1.2.1	SARCOMA DE EWING E LNCRNA	9
1.3	HISTONA DEACETILASES	9
1.3.1	MODIFICADORES EPIGENÉTICOS E SUA RELAÇÃO COM SARCOMA DE EWING E LNCRNA	11
1.4	BIOLOGIA DE SISTEMAS	13
1.4.1	REDES DE COEXPRESSION	15
1.5	ESTRUTURA DA TESE	16
<b>2</b>	<b><u>OBJETIVOS</u></b>	<b>17</b>
2.1	OBJETIVO GERAL	17
2.2	OBJETIVOS ESPECÍFICOS	17
<b>3</b>	<b><u>MATERIAIS E MÉTODOS</u></b>	<b>18</b>
3.1	INFRAESTRUTURA DE BIOINFORMÁTICA	18
3.2	AMOSTRAS	18
3.3	REFERÊNCIAS	21
3.4	QUALIDADE E FILTRAGEM DOS DADOS BRUTOS	22
3.5	MONTAGEM DO TRANSCRIPTOMA	23
3.5.1	ALINHAMENTO	23
3.5.2	MONTAGEM	23
3.5.3	FILTRAGEM DOS TRANSCRITOS CANDIDATOS	24
3.5.4	POTENCIAL CODIFICANTE DOS TRANSCRITOS	24
3.6	QUANTIFICAÇÃO DA EXPRESSÃO	25
3.7	NORMALIZAÇÃO DA EXPRESSÃO	25
3.8	ANÁLISE <i>IN CIS</i> DE LNCRNA	26
3.9	ANÁLISE <i>IN TRANS</i> DE LNCRNA	27
3.9.1	CONSTRUÇÃO DA REDE DE COEXPRESSION (WGCNA)	27
3.9.2	ANÁLISE DA REDE DE COEXPRESSION	31
3.9.3	VISUALIZAÇÃO DA REDE DE COEXPRESSION	32
3.10	ENRIQUECIMENTO DE ONTOLOGIAS GÊNICAS	32

3.11	DEMAIS FONTES DE DADOS EXTERNAS .....	33
<b>4</b>	<b><u>RESULTADOS E DISCUSSÃO .....</u></b>	<b>35</b>
4.1	PADRÃO DE EXPRESSÃO NO SE.....	35
4.2	INVESTIGAÇÃO DE LNCRNAs OPERANDO <i>IN CIS</i> E <i>IN TRANS</i> A HDACs NO SE.....	37
4.2.1	CORRELAÇÕES <i>IN CIS</i> ENTRE LNCRNA E HDACs.....	37
4.2.2	CORRELAÇÕES <i>IN TRANS</i> ENTRE LNCRNA E HDACs.....	39
4.2.3	DISCUSSÃO .....	51
4.3	CANDIDATOS A NOVOS LNCRNAs NO SE .....	56
4.3.1	PIPELINE COMPUTACIONAL E APRESENTAÇÃO DE NOVOS TRANSCRITOS.....	56
4.3.2	DISCUSSÃO .....	60
<b>5</b>	<b><u>CONCLUSÃO.....</u></b>	<b>63</b>
<b>6</b>	<b><u>REFERÊNCIAS.....</u></b>	<b>64</b>
<b>7</b>	<b><u>ANEXOS.....</u></b>	<b>80</b>
7.1	MATERIAL E MÉTODOS.....	80
7.2	RESULTADOS .....	82

# 1 Introdução

## 1.1 Sarcoma de Ewing

O Sarcoma de Ewing (SE) é uma rara neoplasia maligna mesenquimal de células redondas e azuis que acomete especialmente crianças, adolescentes e adultos jovens (Grünewald et al. 2018). Compreende aproximadamente 1% das neoplasias em crianças até 14 anos e 2% em adolescentes de 15 a 19 anos (Siegel et al. 2021). No mundo, o pico de incidência é entre os 10 a 19 anos, com maior frequência em pacientes do sexo masculino e populações de ascendência europeia (Spector et al. 2021). O tumor pode atingir tecidos moles (cerca de 18% dos casos) e ósseos, nestes com particular incidência na pelve, seguida por demais ossos do tronco e extremidades (Leavey et al. 2021; Hesla, Papakonstantinou, and Tsagkozis 2021). As taxas de sobrevida de 5 anos variam conforme a extensão da doença, com a doença localizada permanecendo em torno de 70-80%. Pacientes com metástase ao diagnóstico (cerca de 30%) apresentam sobrevida menores que 30%, salvo os que exibem apenas metástase pulmonar (sobrevida ~50%) (Ladenstein et al. 2010; Grünewald et al. 2018). Ademais, um relatório recente não encontrou relação da sobrevida livre de eventos e localização do tumor primário (Leavey et al. 2021). Estudos realizados no Brasil encontraram números equiparáveis aos internacionais, com taxas levemente maiores de metástase ao diagnóstico (39%) (Brunetto et al. 2015).

Em termos moleculares, caracteristicamente apresentam fusão gênica envolvendo genes codificadores de proteína da família FET, e genes codificadores de proteína ligadoras de RNA da família ETS (fusões tipo FET-ETS), deste modo gerando uma proteína quimérica (Doyle 2014; Vilanova 2017; Grünewald et al. 2018).

O diagnóstico do SE requer análise de radiografias, tomografias ou ressonância magnética aliadas à análise histopatológica, possibilitando estimativa da extensão de destruição tecidual pela neoplasia. Marcadores moleculares, como CD99, também podem ser utilizados, apesar deste não ser específico ao SE. Técnicas de citogenética molecular, como a hibridação fluorescente *in situ* (*Fluorescent In Situ Hybridization* - FISH), e métodos baseados em transcrição reversa seguida de reação em cadeia da polimerase (*Reverse Transcription Polymerase Chain Reaction* - RT-PCR), possibilitam identificar a quebra cromossômica, de modo que são inicialmente empregados para concluir o diagnóstico da doença (Ozaki 2015; Hesla, Papakonstantinou, and Tsagkozis 2021).

O tratamento do SE depende da localização do tumor primário, gravidade da lesão e presença de metástases. Na maior parcela dos casos envolve excisão cirúrgica da lesão, seguida de cirurgia reconstrutiva, radioterapia e quimioterapia. A radioterapia é especialmente necessária em casos de comprometimento de margem cirúrgica e para tumores em sítios de difícil acesso. Já a quimioterapia é empregada na forma neoadjuvante, para redução de volume tumoral pré-cirurgia, ou adjuvante, para consolidação do tratamento cirúrgico ou tratamento de doença metastática (Ozaki 2015; Khanna, Pandey, and Bajpai 2017).

A origem do SE ainda é controversa, mas está relacionada às células da crista neural ou células-tronco mesenquimais (*Mesenchymal Stem Cell - MSC*), esta última hipótese a mais aceita na literatura recente (Tu et al. 2017).

A hipótese de origem na crista neural se baseia na expressão de propriedades imunofenotípicas compartilhados entre o SE e tais tecidos, incluindo expressão de enolase neurônio-específica (*Neuron Specific Enolase - NSE*) e proteína S-100 (Franchi et al. 2001). Admite-se que células-tronco humanas da crista neural (*human Neural Crest Stem Cells - hNCSC*) podem se diferenciar em células tronco mesenquimais derivadas da crista neural (*hNC-MS*C), adquirindo fenótipo mesenquimal e capacidade de diferenciação em adipócitos e osteócitos. Porém, a transdução da proteína quimérica EWS-FLI1 em culturas de hNC-MS C causou expressão de genes associados a hNCSC e repressão de genes associados a hNC-MS C, o que evidencia reativação do estado de hNCSC e indica que essas são possíveis candidatas a célula de origem da doença (von Levetzow et al. 2011).

A hipótese de origem nas células-tronco mesenquimais apresenta indícios sólidos. Células-tronco tumorais CD133+ foram isoladas de tumores primários de SE e cultivadas de modo equivalente à indução de diferenciação de MSCs primárias normais. As células nessas condições foram capazes de se diferenciarem em adipócitos, condrócitos e osteócitos, revelando que manutenção de propriedades de MSC mesmo na vigência de expressão da proteína de fusão EWS-FLI1 (Suvà et al. 2009). A expressão de EWS-FLI1 em células progenitoras mesenquimais (*Mesenchymal Progenitor Cells – MPCs*) provenientes da medula óssea de ratos C57BL/6 foram capazes de formar tumores em camundongos com síndrome imunodeficiência combinada severa (*Severe Combined Immunodeficiency - SCID*) mesmo após duas reinjeções. Os tumores exibiram aumento da expressão de genes reconhecidos como alvos da proteína de fusão, tais como *MYC*, *MMP-3*, *ID2* e *CCND1*, além da repressão de *CDKN1A* (Nicolò Riggi et al. 2005). Além disso, nota-se que o silenciamento de EWS-FLI1

em células A-673 provenientes do SE resgata os padrões de expressão gênica global e o fenótipo das MSCs, evidenciando uma das possíveis origens (Tirode et al. 2007).

O SE apresenta características moleculares bem definidas que influenciam diretamente em sua carcinogênese. De modo geral há a fusão FET-ETS que em cerca de 85% dos casos decorrem da translocação t(11,22)(q24,q12) e originam a proteína fusionada EWS-FLI1 (provenientes dos genes *EWSR1* e *FLI1*, respectivamente). Nesses casos tipicamente se observa junção entre o exon 7 de *EWSR1* e o exon 6 (tipo 1) ou exon 5 (tipo 2) de *FLI1*. Entretanto, outras 7 combinações também são relatadas: em 10% dos casos há a t(21,22)(q22,q12), que gera a fusão *EWSR1-ERG*, e os casos remanescentes correspondem a translocações incomuns, como *EWSR1-ETV1* e *EWSR1-FEV* (Brohl et al. 2014; Tirode et al. 2014; Sand, Szuhai, and Hogendoorn 2015). Mutações são incomuns no SE, porém alterações em *STAG2*, *TP53* e *CDKN2A* são detectadas em alguns casos. Em geral as alterações em *STAG2* e *TP53* estão relacionadas a um pior prognóstico da doença (Brohl et al. 2014; Grünewald et al. 2018).

A proteína EWS-FLI1 liga-se a repetições GGAA, especialmente em regiões de sequências potencializadoras (*enhancers*). A proteína quimérica torna-se um fator de transcrição aberrante, levando a modulação gênica errônea e alterando o perfil transcricional celular. Um dos mecanismos de ação da proteína de fusão é recrutar remodeladores de cromatina como p300, BAF e LSD1 que por sua vez causam modificações de histonas, propiciando ativação ou repressão gênica. Ainda, a presença em EWS de um pequeno domínio semelhante a prion é suficiente para a atividade anormal da proteína quimérica (Nicolò Riggi et al. 2014; Boulay et al. 2017).

A dualidade ativação/repressão ocasionada pela proteína quimérica do SE afeta tanto RNAs mensageiros (*messenger RNAs* - mRNAs) como RNAs não-codificantes (*non-coding RNAs* - ncRNAs). Dentre os ncRNAs alterados podemos citar os micro-RNAs (miRNA) e RNAs longos não codificantes (*long non-coding RNA* - lncRNA) (Hancock and Lessnick 2008; Nicolò Riggi et al. 2010; Marques Howarth et al. 2014; Karnuth et al. 2014). As alterações transcricionais decorrentes de EWS-FLI1 acarretam fenótipos característicos da progressão do câncer, incluindo aumento das taxas de proliferação, pluripotência, aumento de motilidade e migração celular e indução de angiogênese (Sand, Szuhai, and Hogendoorn 2015).

## 1.2 RNAs longos não codificantes

Os lncRNAs são definidos como moléculas de RNA com mais de 200 nucleotídeos e que não codificam proteínas. Integram com diversas moléculas, formando complexos RNA-proteína, RNA-DNA e RNA-RNA. A ligação com proteínas decorre da capacidade de formar estruturas secundárias. Quando ligados ao DNA, podem recrutar complexos remodeladores de cromatina, ativando ou suprimindo a expressão de genes. Ademais, interagem com mRNAs e atuam como “esponjas” de miRNAs. Assim, lncRNAs exercem regulação gênica pré e pós-transcricional (Guttman et al. 2009; Prensner and Chinnaiyan 2011).

Os lncRNA são expressos em determinados tecidos e situação celulares – i.e., apresentam expressão espaço-temporal específica – usualmente concomitante a genes codificadores de proteínas. Apresentam localização intracelular dependente de sua função: apenas no núcleo, no citoplasma ou em ambos (Mallory and Shkumatava 2015; H. Zhao et al. 2016). Alguns lncRNA podem ser excretados em vesículas extracelulares, sendo detectados em fluídos corporais como sangue e urina (Takahashi et al. 2014). Portanto, além de sua importância na regulação gênica, seu uso como marcadores moleculares começa a ser explorado. Por exemplo, recentemente o lncRNA *PCA3*, que é encontrado na urina, foi aprovado para uso diagnóstico no carcinoma de próstata (Schmitt and Chang 2016). Tais características tornaram os lncRNA promissores marcadores prognósticos e diagnósticos (Byron et al. 2016).

O potencial regulatório dos lncRNA ainda é o mais explorado, especialmente considerando que a maioria dos transcritos conhecidos ainda apresenta função desconhecida (Palazzo and Lee 2015). Os lncRNA exercem regulação gênica por meio de duas formas principais: 1) regulação *in cis*, quando o lncRNA atua em seu próprio locus ou 2) *in trans*, quando o lncRNA exerce sua função em um locus distante de sua origem (Quinn and Chang 2016; Kopp and Mendell 2018). Serão demonstrados alguns exemplos destes dois tipos de regulação a seguir.

A regulação *in cis* ocorre por alguns processos distintos. Os lncRNA podem comportar-se como RNAs *enhancer* (*enhancer RNA* – eRNA) atuando na mediação de loops cromossômicos. O lncRNA *LUNARI* age desta forma, mediando o recrutamento de proteínas auxiliares por meio da formação do arcabouço cromossômico na região do gene *IGF1R*, que por consequência é expresso. Os lncRNA atuam também em regiões de *imprinting*, e seu silenciamento epigenético leva a modulação gênica dependente da origem parental. O eixo que compreende o lncRNA *H19* e o gene codificador de proteínas *IGF2* exhibe tal mecanismo. *H19* é expresso maternalmente e silencia o gene *IGF2*; este último é expresso somente no alelo

paterno. A dosagem gênica é outro processo controlado por lncRNAs operando *in cis*, que tem como exemplo a inativação do cromossomo X por meio da expressão de *Xist*. A transcrição *antisense* de lncRNA é também um exemplo de regulação *in cis*. Genes codificadores de proteínas são silenciados quando os complexos de RNA Polimerase II (RNA Pol II) transcrevem em direções opostas, ocasionando colisão entre eles. Tal tipo de evento é descrito para o lncRNA *Air* e o gene codificador de proteína *Igf2r*. Por fim, hibridizações *sense-antisense* silenciam genes alvo, a exemplo dos lncRNA *ZEB2-AS* e *BACE1-AS*, que desta forma modulam a expressão de *ZEB2* e *BACE1*, respectivamente (Quinn and Chang 2016).

A regulação *in trans* transcorre pela ligação dos lncRNA com proteínas, DNA e outros RNAs. *HOTAIR*, um lncRNA transcrito na região do gene *HOXC*, interage diretamente com complexo PRC2 e por conseguinte silenciando o gene *HOXD*. *HOTAIR* e PRC2 são necessários para que a região de *HOXD* mantenha a marca epigenética repressiva correspondente a trimetilação da lisina 27 na histona H3 (H3K27me3). Adicionalmente, lncRNA influenciam a arquitetura cromossômica. O lncRNA *Firre* causa a ancoragem de seu locus de origem no cromossomo X com cinco loci distantes. O silenciamento de *Firre* acarretam a perda de contato entre estas regiões cromossômicas. Outro exemplo de regulação *in trans* dá-se pela ligação direta entre o lncRNA e uma proteína. PUM1 e PUM2 são proteínas responsáveis pela inibição e degradação de mRNA contendo as sequências UGUANAUA (*PUMILIO response element - PRE*) em suas regiões 3'-UTR. O lncRNA *NORAD* liga-se as sequências PRE presentes PUM1 e PUM2, inativando-as. Por fim, lncRNA capturam miRNAs, a exemplo do lncRNA *CDRIas*, que sequestra o miR-7 e impede o silenciamento de seus alvos (Kopp and Mendell 2018).

Há relação direta entre os lncRNA e diversos cânceres. *HOTAIR* liga-se ao complexo PRC2 e em conjunto com a subunidade EZH2 altera as marcações H3K27me3, ocasionando a repressão de genes supressores de metástases no câncer de mama, como *HOXD10*, *JAM2* e *PCDH10*. Já a supressão de *HOTAIR* está relacionada a diminuição da capacidade de invasão celular (Chisholm et al. 2012; Shen, Qi, and Du 2015). Um mecanismo semelhante de ligação à PRC2 está associado ao lncRNA *PCAT-1*, aumentado no câncer de próstata (Prensner and Chinnaiyan 2011). No carcinoma de células não-pequenas do pulmão, *MALAT1* é marcador prognóstico indicador de menor sobrevida, envolvido na formação de metástases por alterar a expressão de genes relacionados à metástase de modo semelhante ao *HOTAIR*, mas relacionado a PRC1 (Shen, Qi, and Du 2015). No câncer de bexiga, *MALAT1* ativa a via Wnt/ $\beta$ -catenina, reprimindo a expressão de caderina E, de modo a propiciar transição epitelial mesenquimal (*Epithelial Mesenquimal Transition - EMT*) (Ying et al. 2012). Já os lncRNA *PPP3CB*,

*MAP3K14*, e *DAPK1*, por sua vez, estão aumentados em casos de adenocarcinoma do ducto pancreático, e alteram genes relacionados a via MAPK, envolvida na progressão do câncer (Shen, Qi, and Du 2015).

Portanto, percebemos a que os lncRNA participam da patogênese dos cânceres, com implicações em sua progressão.

### 1.2.1 Sarcoma de Ewing e lncRNA

A caracterização de lncRNA no SE ainda está em seus estágios iniciais, porém alguns transcritos já indicam a importância desses transcritos (Barrett et al. 2021). O lncRNA *EWSATI* é expresso concomitantemente a EWS-FLI e auxilia no potencial oncogênico da célula de SE. Seu mecanismo de ação é relacionado a ligação com a proteína HNRNPK, formação de um complexo lncRNA-proteína e posterior regulação da interação da proteína com seus alvos (Marques Howarth et al. 2014). O mesmo lncRNA regula *MEG3* no osteosarcoma, contribuindo para migração, invasão e proliferação celular (Sun et al. 2016).

Já os lncRNA *HULC* e *SOX2OT* são superexpressos em linhagens de SE. O silenciamento de *HULC* diminui proliferação e clonogênese, enquanto o silenciamento de *SOX2OT* diminui proliferação, invasão e sobrevida. Também atuam como esponjas de miRNA para regular outros transcritos. Deste modo, mostrou-se que ambos contribuem para o fenótipo maligno em células de SE (Barrett et al. 2021). Por meio de uma estratégia computacional, recentemente encontrou-se sete lncRNA (*SNHG17*, *WAC-AS1*, *LINC00623*, *SSBP3-AS1* e *TDRG1*) que, com expressão elevada, são sugestivos de pior prognóstico em pacientes com SE (Z. Chen et al. 2021). Ademais, o lncRNA *FOXP4-AS1* é superexpresso em células de SE e seu silenciamento suprimiu proliferação, migração e invasão celular. Sua localização é predominantemente citoplasmática e é excretado por vesículas extracelulares (Xiong et al. 2021).

Isto demonstra a complexidade a ser ainda descoberta quanto as interações dos lncRNA com os demais atores celulares. Os estudos citados evidenciam a importância dos lncRNA no SE, ao mesmo tempo que demandam novas investigações acerca deste conjunto de transcritos na biologia deste sarcoma.

## 1.3 Histona deacetilases

As proteínas HDAC clássicas, dependentes de zinco e que serão abordadas neste trabalho, são divididas em classe I (HDAC1, HDAC2, HDAC3 e HDAC8), classe IIa (HDAC4, HDAC5,

HDAC7 e HDAC9), classe IIb (HDAC6 e HDAC10) e classe IV (HDAC11). Temos ainda as de classe III (SIRT 1-7) que compreendem as sirtuínas dependentes de NAD<sup>+</sup>. A exceção da classe IIb, que tipicamente são encontradas no citoplasma, e as de classe I, encontradas primariamente no núcleo, as demais HDACs estão presentes no núcleo ou citoplasma, conforme a sinalização celular que recebem (de Ruijter et al. 2003; B. Chen et al. 2015).

A modulação epigenética da transcrição dá-se, dentre outros processos, pela acetilação e deacetilação de lisinas presentes nas caudas das histonas componentes dos nucleossomas, que empacotam o DNA e previnem a expressão gênica. A interação entre histonas carregadas positivamente e o DNA, que tem carga negativa, interrompe o acesso da RNA Pol II e evita a transcrição dos genes encontrados na região. A acetilação dos resíduos de lisina nas caudas das histonas por meio da histona acetiltransferase (HAT), neutraliza a carga dessas, com consequente relaxamento dos nucleossomas e, portanto, expressão gênica. A recuperação da carga positiva dá-se pela deacetilação das caudas das histonas por meio da ação das histonas deacetilases (HDAC), que recuperam a estrutura compacta da cromatina, resultando em repressão gênica (Park and Kim 2020).

As várias classes de HDACs são expressas em todos os tecidos, com modulações diferentes conforme a enzima e o tecido analisado. Muitas participam de complexos proteicos para exercer sua função em sítios específicos no genoma. Por exemplo, as HDAC1 e HDAC2 são parte dos complexos CoREST, Sin3 A/B e NuRD, enquanto a HDAC11 é componente do complexo SMN. Portanto, demonstra-se que a interação das HDACs com os demais componentes celulares é fundamental para sua função (Milazzo et al. 2020; Banks et al. 2020; J. Zhang et al. 2020).

A atividade das HDACs é estritamente controlada, devido a sua importância como reguladores epigenéticos e controladores da expressão gênica. A desregulação das HDACs ocorre em diversos tipos de câncer, e relacionada a baixa sobrevida em alguns. Por exemplo, a *HDAC2* tem expressão aumentada em tumores pulmonares (adenocarcinoma e carcinoma de células escamosas), cervicais, uterinos e gástricos, com menor sobrevida reportada no adenocarcinoma pulmonar, mesotelioma e em sarcomas de tecido mole, quando analisadas amostras do TCGA (Milazzo et al. 2020).

Diversas drogas com atividade anti-histona deacetilase foram e estão sendo pesquisadas para combater seus efeitos deletérios nos cânceres. Os inibidores de HDACs (HDACi) podem agir de maneira específica em determinada HDAC ou de forma generalizada. Drogas como Vorinostat (SAHA) e Belinostat (PXD101) têm ação generalizada, enquanto Ácido Valprônico

(VPA) e Entinostat (MS-275) têm ação específica em HDACs de classe I (Milazzo et al. 2020). A baixa especificidade de alguns inibidores pode acarretar quadros de toxicidade nos pacientes, incluindo, mas não limitado a trombocitopenia, neutropenia, náusea e vômitos. Portanto, o estudo dos efeitos biológicos dos HDACi ainda requer atenção para formatação de terapias mais seguras no futuro (Eckschlager et al. 2017; Park and Kim 2020).

Os diversos HDACi atuam em diferentes vias, como indução de apoptose, parada do ciclo celular e autofagia, por exemplo. Também alteram a expressão de RNAs não codificantes. SAHA induz a expressão do miR-129-5p, miRNA necessário para estimular morte celular em células de câncer de tireoide. LncRNA também são encontrados diferencialmente expressos em linhagens de câncer hepatocelular e o lncRNA *Xist* é um biomarcador para o uso do inibidor Abexinostat no câncer de mama (Salvador et al. 2013; Eckschlager et al. 2017).

### 1.3.1 Modificadores epigenéticos e sua relação com sarcoma de Ewing e lncRNA

O SE é permeado por alterações epigenéticas decorrentes da ação da proteína fusionada EWS-FLI1 (Tomazou et al. 2015; Boulay et al. 2017; Nicolò Riggi et al. 2014; Boulay et al. 2018). Devido ao papel de tais alterações na carcinogênese, somado a sua reversibilidade, as enzimas controladoras de processos epigenéticos têm sido utilizadas como alvos terapêuticos. Dentre tais, os HDACi já são usados na rotina clínica com bons resultados. Além disto, novos inibidores estão em testes clínicos a fim de ampliar a gama de tratamentos disponíveis (C. Souza and Chatterji 2015; Yu et al. 2019). Porém, a rápida aparição de novas drogas traz a necessidade de maiores esclarecimentos quanto a seus mecanismos de ação e suas interações com os demais componentes celulares.

A disfunção na acetilação de histonas está presente em tumores sólidos e hematológicos (Marks et al. 2001). No SE a HAT p300 atua conjuntamente com EWS-FLI1 nos sítios GGAA que são alvos da proteína fusionada, causando acetilação dos mesmos e consequente ativação da transcrição gênica (Nicolò Riggi et al. 2014). A mesma importância dá-se as HDACs, que têm sua desregulação também associada a oncogênese. O recrutamento de HDACs ocorre em leucemias e linfoma não-Hodgkin's, por exemplo, levando ao quadro patogênico (Marks et al. 2001). Tais consequências evidenciam a importância do balanço da acetilação no desenvolvimento de diversos tumores.

Terapias mais eficazes são necessárias para o tratamento do SE, e alvejar os modificadores epigenéticos é especialmente promissor para este propósito. Inibidores de HDAC, como Vorinostat (SAHA) e Etinostat (SNDX-275), estão sendo avaliados em testes clínicos, com resultados preliminares favoráveis (Lawlor and Thiele 2012). Como exemplo, recentemente mostrou-se que o Etinostat induz apoptose nas células de SE em testes *in vitro* e diminui a carga tumoral e aumenta a sobrevida em testes *in vivo* (Ma et al. 2019). A droga atinge seletivamente células expressando EWS-FLI1 e EWS-ERG, o que confere uma importante vantagem para o tratamento. Similar seletividade para célula do SE é observada em outros HDACi como Vorinostat, NaB e MS-275 (Sonnemann et al. 2007; B. K. Souza et al. 2018). O uso de HDACi no tratamento do SE é uma estratégia promissora, portanto elucidar os mecanismos relacionados a HDACs na patogênese do SE é de suma relevância.

Em várias condições, HDACi têm direta conexão com a expressão e atividade de transcritos não-codificantes, incluindo, mas não limitando-se aos lncRNA. Por meio da superexpressão do transcrito *lncRNA-A330023F24Rik*, foi reportado que Vorinostat reduz arteriosclerose em camundongos. No geral, o mecanismo de ação para esta droga engloba a alteração de metilação de sítios promotores de lncRNA (Rafehi and El-Osta 2016; Ye et al. 2018). Diante dos milhares de lncRNA conhecidos, há relativamente pouca informação quanto a possibilidade de sua interferência com a atividade de HDACs, expondo a necessidade de investigações adicionais.

As HDACs são constitutivamente transcritas no SE. O *knockout* da HDAC1 e HDAC2, por meio de CRISPR/Cas9, diminuiu a proliferação de células de SE, assim como reduziu o crescimento tumoral em modelos *in vivo*. Altos níveis de HDACs classe I também são associados a menor taxa de sobrevida em pacientes (Schmidt et al. 2021). A inibição seletiva da HDAC6 com BML-281 induziu apoptose e reduziu a tumorigênese *in vitro*, assim como apresentou atividade antitumoral *in vivo* (García-Domínguez et al. 2021). Ademais, a HDAC2 liga-se diretamente a quimera EWS-FLI1, fato evidenciado por meio de experimentos com cromatografia líquida acoplada à espectrometria de massas (Selvanathan et al. 2015).

Não há dúvidas quanto à existência de interações entre lncRNA e modulares epigenéticos. Neste quesito, HDACs são conhecidas por causar ou sofrer influência de lncRNA. Mesmo assim estas relações necessitam de melhores elucidações a nível molecular, o que é crucial para avaliar o uso de HDACs como alvos terapêuticos no SE.

## 1.4 Biologia de Sistemas

A chegada do sequenciamento de nova geração propiciou estudos genômicos em larga escala devido a enorme quantidade de dados gerados. Entretanto, este cenário demanda novas formas de estudar a complexidade dos sistemas biológicos. A biologia de sistemas, que utiliza grafos ou redes para elucidar o funcionamento de sistemas associados a processos vitais torna-se chave neste momento (Vidal, Cusick, and Barabási 2011; Samyn et al. 2015; J. Yan et al. 2017).

Redes são estruturas matemáticas compostas por nodos, que representam objetos, e arestas, que indicam a ligação entre os objetos. Nas redes biológicas, os nodos representam os diversos componentes celulares (e.g., genes, transcritos ou proteínas) enquanto as arestas são as interações entre estes componentes (e.g., interações entre proteínas). As redes podem ser direcionais, quando as arestas indicam a direção entre os nodos (ex. redes metabólicas) ou não-direcionais, quando não há indicação desta direção, somente se os componentes interagem ou não (ex. redes proteína-proteína) (Barabasi and Oltvai 2004).

A posição de um nodo em uma rede e sua possível importância é representada pelas medidas de centralidade. Tais medidas em redes biológicas são empregadas para apontar os componentes biológicos mais relevantes ao sistema, como por exemplo, quais genes são indispensáveis aos processos celulares. Uma importante medida de centralidade é o grau de conectividade  $k(g)$  ou a quantidade de interações que um nodo  $g$  possui (Barabasi and Oltvai 2004; Acencio et al. 2013). Redes biológicas em geral configuram-se como “livre de escala”, ou seja, vários nodos têm baixo grau de conectividade enquanto poucos nodos têm alto grau de conectividade. Esta propriedade está relacionada a redundância nos sistemas biológicos; um processo celular importante apresenta diversas formas de ocorrer, e a falha de um componente com baixa conectividade é circunscrita por outro. Nodos com alta conectividade são denominados “hubs” (Jeong et al. 2000).

As células podem ser representadas por redes de interação entre seus vários componentes biológicos. Tais redes são denominadas interactomas e os dados utilizados como entrada podem ser compilados de banco de dados preexistentes, preditos computacionalmente ou mesmo coletados por experimentos de larga escala. Os dados podem gerar diversos tipos de redes: interações proteína-proteína, redes de regulação gênica, redes de coexpressão, redes metabólicas, interações hospedeiro-patógeno, entre outras (Vidal, Cusick, and Barabási 2011).

A atual onipresença de dados de expressão gênica, representados principalmente por experimentos de RNA-Seq, facilita a construção de redes de coexpressão. Este tipo específico

de rede não-direcional apresenta seus nodos como genes e suas arestas como a correlação entre a expressão gênica. A facilidade em sua construção, aliado a possibilidade da adição de outros tipos de informação à rede (ex. interações proteína-proteína), torna as redes de coexpressão uma valiosa ferramenta para seleção de candidatos a ensaios funcionais (Contreras-López et al. 2018). A análise de redes de coexpressão possibilita a separação dela em módulos que são correlacionados a processos intracelulares por meio do enriquecimento de vias ou de termos ontológicos. O estudo dos “hubs” presentes nestes módulos tende a revelar genes importantes ao funcionamento do sistema biológico (van Dam et al. 2018).

A análise das redes biológicas e suas propriedades auxilia na elucidação dos mais variados problemas. Conhecimento do mecanismo de ação de doenças, processos metabólicos desconhecidos, interações entre moléculas, biomarcadores e alvos de drogas são alguns dos usos para as redes. Regiões densas das redes, chamadas módulos topológicos, apresentam funções semelhantes e identificam componentes relacionados a um processo celular ou doença. A análise das relações presentes no interactoma pode gerar uma lista de genes relacionados a doenças, especialmente devido a relação genótipo-fenótipo presente em tais redes (Alvarez et al. 2016; W. Yan et al. 2016; Cowen et al. 2017; van Dam et al. 2018).

Um interactoma contendo genes, proteínas e fenótipos demonstrou bom valor preditivo na busca por receptores glicocorticoides (Bakker et al. 2017). A análise das medidas de topologia, por exemplo, através de métodos de inteligência artificial, é capaz de gerar sub-redes com novas vias moleculares a serem estudadas em laboratório, como em interações oncogênicas (Acencio et al. 2013). No carcinoma de nasofaringe, redes de interação mRNA-lncRNA identificaram lncRNA participantes na migração, proliferação, crescimento e apoptose celular, validados pela reação em cadeia em tempo real (qPCR) (Q. Yang and Deng 2015). O uso do sequenciamento de nova geração, associado a biologia de sistemas na pesquisa do câncer, já revelou genes relacionados ao câncer de pulmão, marcadores prognósticos e preditivos para quimioterapia no câncer de mama além de genes relacionados ao câncer de próstata (W. Yan et al. 2016). A priorização através do algoritmo VIPER em uma rede de regulação gênica (GRN), criada com o algoritmo ARACNe, identificou o fator de transcrição TRIM25 como molécula chave na metástase do carcinoma de mama (Walsh et al. 2017). Isto demonstra o uso dos interactomas na busca de marcadores moleculares e elucidação de vias de sinalização.

### 1.4.1 Redes de coexpressão

As redes de coexpressão são amplamente empregadas nas análises de dados de expressão. Nela, nodos representam genes e as arestas a correlação de suas expressões. Um *framework* popular para criar redes de coexpressão chamado WGCNA (*weighted gene co-expression network*) aloca pesos para as conexões entre os nodos da rede usando um valor  $\beta$  (chamado aqui de “limiar suave”) escolhido pelo usuário. A implementação do WGCNA é realizada por meio de um pacote para o ambiente R (Langfelder and Horvath 2008).

Brevemente, a primeira etapa consiste em definir uma medida de similaridade entre os genes. Neste contexto utiliza-se o valor absoluto da correlação de Pearson. A partir destes valores, a matriz de adjacência é construída elevando-os ao coeficiente  $\beta$ . O valor de  $\beta$  é escolhido para adequar-se a premissa de que redes biológicas têm um padrão “livre de escala”, e a rede resultante terá esta característica, preferencialmente. O próximo passo é encontrar módulos de genes altamente conectados. Para tal, emprega-se a medida topológica de sobreposição (*topological overlap measure*) que resulta na matriz TOM (*topological overlap matrix*), uma medida de similaridade entre os nodos. Os módulos são então concatenados com base em seus autovalores (*eigenvectors*) e relacionados com características genotípicas ou fenotípicas de interesse (Langfelder and Horvath 2008; B. Zhang and Horvath 2005).

O uso de WGCNA para criação de redes de coexpressão auxilia no estudo da função dos lncRNA, que em grande parte são desconhecidos. Aplica-se a técnica para correlacionar a expressão de lncRNA a mRNA conhecidos e determinar quais processos biológicos esses estão envolvidos. Esta estratégia denomina-se “*guilty-by-association*” e têm a premissa que transcritos coexpressos compartilham funções, reguladores ou processos celulares (Signal, Gloss, and Dinger 2016). Por exemplo, um estudo com diversos câncer, que empregou o método WGCNA, relacionou 236 lncRNA com o desenvolvimento desses cânceres, além de implicá-los em funções como estabilidade genômica, sinalização celular, adesão e motilidade celular, entre outros (S. Li et al. 2017). Também por meio de WGCNA, encontrou-se sete lncRNA envolvidos na modulação das células imunes no microambiente tumoral do melanoma, incluindo os novos candidatos *SNHG14*, *HCG18*, e *MAGI2-AS3*, previamente não implicados neste tipo de câncer (Cheng et al. 2020).

## 1.5 Estrutura da tese

Aborda-se nesta tese conhecidos relacionados ao sarcoma de Ewing, as HDACs e aos lncRNA, classe de transcritos que gera crescente interesse da comunidade. Dividimos os conhecimentos em cinco partes, sendo introdução, objetivos, material e métodos, resultados e discussão, além da conclusão. Portanto, apresenta-se uma visão geral de cada parte da tese nos parágrafos a seguir.

A primeira parte, de introdução, buscou abordar a base teórica de nossos estudos, com foco no SE, lncRNA e HDACs, além da visão geral de alguns métodos empregues. Segue-se a segunda parte, de objetivos, que esperamos que o leitor veja como concluídos ao final da leitura.

Na terceira parte, de material e métodos, elencamos as metodologias computacionais e programas que nos auxiliaram a cumprir nossos objetivos. Alguns métodos, como pré-processamento de amostras e quantificação da expressão são comuns à todas as análises, adaptadas apenas ao tipo de abordagem. A montagem do transcriptoma, por exemplo, foi exclusiva a criação de um *pipeline* para encontrar novos lncRNA.

A parte quatro, de resultados e discussão, mostra e contextualiza efetivamente nossos resultados. Dividimos os resultados em duas partes: a investigação de lncRNA associados a HDACs e a criação de um *pipeline* para encontrar novos lncRNA no SE. Na primeira parte, nosso racional foi aplicar métodos da biologia de sistemas para encontrar lncRNA coexpressos a HDACs. Na segunda parte, esperávamos que, por meio de uma estratégia computacional, encontrássemos lncRNA ausentes de repositórios públicos.

## 5 Conclusão

Neste trabalho, empregamos diversas técnicas computacionais para gerar conhecimento acerca do papel dos lncRNA no sarcoma de Ewing. Aplicamos métodos da biologia de sistemas e da biologia computacional e bioinformática para encontrar genes candidatos a estudos aprofundados, tanto *in vitro* quanto *in vivo*.

Na primeira parte dos resultados, investigamos a relação entre lncRNA e HDACs por meio de redes de coexpressão. Obtivemos resultados compatíveis com observações anteriores no SE, assim como algumas relações interessantes. Por exemplo, encontramos os pares *HDAC7-AC004466.1* e *HDAC10-CR559946.2* com correlação *in cis*, ensejando estudos adicionais desta dinâmica. As interações *in trans* revelaram o eixo *HDAC2, HSP90AA1* e *SETD1A* como pertencentes ao complexo HCF-1, e os lncRNA *AC020900.1* e *AC138207.4* associados a alguns desses transcritos codificantes. Entretanto, novos estudos são necessários para estabelecer com exatidão um eixo de causalidade das associações aqui descritas no sarcoma de Ewing.

Na segunda parte dos resultados, empregamos um *pipeline* computacional para encontrar, em amostras de pacientes, transcritos ausentes de banco de dados. Selecionamos o RNACentral para filtragem por ser um agregador de outras fontes, esperando assim razoável restringência. Com uma breve explanação sobre alguns desses transcritos, idealizamos que possam também ser motivo de estudos aprofundados no SE.

Ao fim de nosso estudo, temos uma visão geral de alguns mecanismos da EWS-FLI1. A quimera trabalha alterando os estados de cromatina e, portanto, a expressão gênica em diversas regiões. Sua ação dá-se com o auxílio das HDAC, que expõe regiões inacessíveis do genoma, liberando, dentro outros, os lncRNA. Em contrapartida, esses transcritos não codificantes também influenciam tais moduladores epigenéticos. As associações entre HDAC e lncRNA são encontradas no contexto da fosforilação oxidativa, e com candidatos a novos transcritos no genoma mitocondrial, sugere-se que processos metabólicos são importantes nesse cenário. Portanto, os lncRNA são implicados não somente como moléculas que ressurgem quando a cromatina é afetada, mas também participando ativamente neste desenvolvimento, pela associação direta com as HDAC.

Tal panorama, mesmo com lacunas, enseja estudos aprofundados das relações entre HDAC, lncRNA e EWS-FLI1, sejam eles computacionais ou funcionais, para melhor compreender todas as implicações genômicas da quimera EWS-FLI1.

## 6 Referências

- Abrams, Zachary B., Travis S. Johnson, Kun Huang, Philip R.O. Payne, and Kevin Coombes. 2019. “A Protocol to Evaluate RNA Sequencing Normalization Methods.” *BMC Bioinformatics* 20 (Suppl 24): 1–7. <https://doi.org/10.1186/s12859-019-3247-x>.
- Acencio, Marcio Luis, Luiz Augusto Bovolenta, Esther Camilo, and Ney Lemke. 2013. “Prediction of Oncogenic Interactions and Cancer-Related Signaling Networks Based on Network Topology.” *PLoS ONE* 8 (10). <https://doi.org/10.1371/journal.pone.0077521>.
- Alvarez, Mariano J, Yao Shen, Federico M Giorgi, Alexander Lachmann, B Belinda Ding, B. Hilda Ye, and Andrea Califano. 2016. “Functional Characterization of Somatic Mutations in Cancer Using Network-Based Inference of Protein Activity.” *Nature Genetics* 48 (8): 838–47. <https://doi.org/10.1038/ng.3593>.
- Andrews, Simon. 2016. “Fastqc.” <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Aynaud, Marie Ming, Olivier Mirabeau, Nadege Gruel, Sandrine Grossetête, Valentina Boeva, Simon Durand, Didier Surdez, et al. 2020. “Transcriptional Programs Define Intratumoral Heterogeneity of Ewing Sarcoma at Single-Cell Resolution.” *Cell Reports* 30 (6): 1767-1779.e6. <https://doi.org/10.1016/j.celrep.2020.01.049>.
- Bakker, Emyr, Kun Tian, Luciano Mutti, Constantinos Demonacos, Jean-Marc Schwartz, and Marija Krstic-Demonacos. 2017. “Insight into Glucocorticoid Receptor Signalling through Interactome Model Analysis.” *PLoS Computational Biology* 13 (11): e1005825. <https://doi.org/10.1371/journal.pcbi.1005825>.
- Banks, Charles A S, Ying Zhang, Sayem Miah, Yan Hao, Mark K. Adams, Zhihui Wen, Janet L. Thornton, Laurence Florens, and Michael P. Washburn. 2020. “Integrative Modeling of a Sin3/HDAC Complex Sub-Structure.” *Cell Reports* 31 (2): 107516. <https://doi.org/10.1016/j.celrep.2020.03.080>.
- Barabasi, Albert-Laszlo, and Zoltan N. Zoltán N Oltvai. 2004. “Network Biology: Understanding the Cell’s Functional Organization.” *Nature Reviews Genetics* 5 (2): 101–13. <https://doi.org/10.1038/nrg1272>.
- Barrett, Connor, Anuj Budhiraja, Vijay Parashar, and Mona Batish. 2021. “The Landscape of Regulatory Noncoding RNAs in Ewing’s Sarcoma.” *Biomedicine* 9 (8): 933. <https://doi.org/10.3390/biomedicine9080933>.
- Betancur, Juan G. 2016. “Pervasive LncRNA Binding by Epigenetic Modifying Complexes — The Challenges Ahead.” *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1859 (1): 93–101. <https://doi.org/10.1016/J.BBAGRM.2015.10.009>.

- Bolger, Anthony M, Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- Boulay, Gaylor, Gabriel J. Sandoval, Nicolo Riggi, Sowmya Iyer, Rémi Buisson, Beverly Naigles, Mary E. Awad, et al. 2017. “Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain.” *Cell* 171 (1): 163-178.e19. <https://doi.org/10.1016/j.cell.2017.07.036>.
- Boulay, Gaylor, Angela Volorio, Sowmya Iyer, Liliane C. Broye, Ivan Stamenkovic, Nicolo Riggi, and Miguel N. Rivera. 2018. “Epigenome Editing of Microsatellite Repeats Defines Tumor-Specific Enhancer Functions and Dependencies.” *Genes & Development* 32 (15–16): 1008–19. <https://doi.org/10.1101/gad.315192.118>.
- Brohl, Andrew S, David A Solomon, Wendy Chang, Jianjun Wang, Young Song, Sivasish Sindiri, Rajesh Patidar, et al. 2014. “The Genomic Landscape of the Ewing Sarcoma Family of Tumors Reveals Recurrent STAG2 Mutation.” *PLoS Genetics* 10 (7): e1004475. <https://doi.org/10.1371/journal.pgen.1004475>.
- Brunetto, Algemir L., Luis A. Castillo, Antonio S. Petrilli, Carla D. Macedo, Erica Boldrini, Cecilia Costa, Maria T. Almeida, et al. 2015. “Carboplatin in the Treatment of Ewing Sarcoma: Results of the First Brazilian Collaborative Study Group for Ewing Sarcoma Family Tumors-EWING1.” *Pediatric Blood and Cancer* 62 (10): 1747–53. <https://doi.org/10.1002/pbc.25562>.
- Byron, Sara A., Kendall R. van Keuren-Jensen, David M. Engelthaler, John D. Carpten, and David W. Craig. 2016. “Translating RNA Sequencing into Clinical Diagnostics: Opportunities and Challenges.” *Nature Reviews Genetics* 17 (5): 257–71. <https://doi.org/10.1038/nrg.2016.10>.
- Chen, Bing, Wenwen Zang, Juan Wang, Yajun Huang, Yanhua He, Lingling Yan, Jiajia Liu, and Weiping Zheng. 2015. “The Chemical Biology of Sirtuins.” *Chemical Society Reviews* 44 (15): 5246–64. <https://doi.org/10.1039/c4cs00373j>.
- Chen, Zhihui, Xinyu Wang, Guozhu Wang, Bin Xiao, Zhe Ma, Hongliang Huo, and Weiwei Li. 2021. “A Seven-LncRNA Signature for Predicting Ewing’s Sarcoma.” *PeerJ* 9 (June): e11599. <https://doi.org/10.7717/peerj.11599>.
- Cheng, Yaqi, Chengxiu Liu, Yurun Liu, Yaru Su, Shoubi Wang, Lin Jin, Qi Wan, et al. 2020. “Immune Microenvironment Related Competitive Endogenous RNA Network as

- Powerful Predictors for Melanoma Prognosis Based on WGCNA Analysis.” *Frontiers in Oncology* 10 (October): 577072. <https://doi.org/10.3389/fonc.2020.577072>.
- Chisholm, Karen M, Yue Wan, Rui Li, Kelli D Montgomery, Howard Y Chang, and Robert B West. 2012. “Detection of Long Non-Coding RNA in Archival Tissue: Correlation with Polycomb Protein Expression in Primary and Metastatic Breast Carcinoma.” *PLoS One* 7 (10): e47998. <https://doi.org/10.1371/journal.pone.0047998>.
- Contreras-López, Orlando, Tomás C. Moyano, Daniela C. Soto, and Rodrigo A. Gutiérrez. 2018. “Step-by-Step Construction of Gene Co-Expression Networks from High-Throughput Arabidopsis RNA Sequencing Data.” In *Root Development: Methods and Protocols*, edited by Daniela Ristova and Elke Barbez, 1761:275–301. Humana Press, New York, NY. [https://doi.org/10.1007/978-1-4939-7747-5\\_21](https://doi.org/10.1007/978-1-4939-7747-5_21).
- Cowen, Lenore, Trey Ideker, Benjamin J. Raphael, and Roded Sharan. 2017. “Network Propagation: A Universal Amplifier of Genetic Associations.” *Nature Reviews Genetics* 18 (9): 551–62. <https://doi.org/10.1038/nrg.2017.38>.
- Crompton, Brian D., Chip Stewart, Amaro Taylor-Weiner, Gabriela Alexe, Kyle C. Kurek, Monica L. Calicchio, Adam Kiezun, et al. 2014. “The Genomic Landscape of Pediatric Ewing Sarcoma.” *Cancer Discovery* 4 (11): 1326–41. <https://doi.org/10.1158/2159-8290.CD-13-1037>.
- Dam, Sipko van, Urmo Vösa, Adriaan van der Graaf, Lude Franke, and João Pedro de Magalhães. 2018. “Gene Co-Expression Analysis for Functional Classification and Gene-Disease Predictions.” *Briefings in Bioinformatics* 19 (4): 575–92. <https://doi.org/10.1093/bib/bbw139>.
- Dasgupta, Atreyi, Matteo Trucco, Nino Rainusso, Ronald J. Bernardi, Ryan Shuck, Lyazat Kurenbekova, David M. Loeb, and Jason T. Yustein. 2017. “Metabolic Modulation of Ewing Sarcoma Cells Inhibits Tumor Growth and Stem Cell Properties.” *Oncotarget* 8 (44): 77292–308. <https://doi.org/10.18632/oncotarget.20467>.
- Dillies, Marie-Agnès Agnès, Andrea Rau, Julie Aubert, Christelle Hennequet-Antier, Marine Jeanmougin, Nicolas Servant, Céline Keime, et al. 2013. “A Comprehensive Evaluation of Normalization Methods for Illumina High-Throughput RNA Sequencing Data Analysis.” *Briefings in Bioinformatics* 14 (6): 671–83. <https://doi.org/10.1093/bib/bbs046>.
- Doyle, Leona A. 2014. “Sarcoma Classification: An Update Based on the 2013 World Health Organization Classification of Tumors of Soft Tissue and Bone.” *Cancer* 120 (12): 1763–74. <https://doi.org/10.1002/cncr.28657>.

- Dragomir, Mihnea, Baoqing Chen, and George A. Calin. 2018. “Exosomal LncRNAs as New Players in Cell-to-Cell Communication.” *Translational Cancer Research* 7 (Suppl 2): S243–52. <https://doi.org/10.21037/tcr.2017.10.46>.
- Eckschlager, Tomas, Johana Plch, Marie Stiborova, and Jan Hrabeta. 2017. “Histone Deacetylase Inhibitors as Anticancer Drugs.” *International Journal of Molecular Sciences* 18 (7): 1414. <https://doi.org/10.3390/ijms18071414>.
- Esparza-Moltó, Pau B., and José M. Cuezva. 2020. “Reprogramming Oxidative Phosphorylation in Cancer: A Role for Rna-Binding Proteins.” *Antioxidants and Redox Signaling* 33 (13): 927–45. <https://doi.org/10.1089/ars.2019.7988>.
- Ewels, Philip, Måns Magnusson, Sverker Lundin, and Max Källér. 2016. “MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report.” *Bioinformatics (Oxford, England)* 32 (19): 3047–48. <https://doi.org/10.1093/bioinformatics/btw354>.
- Fitzpatrick, Christopher, Maximiliano F. Bendek, Macarena Briones, Nicole Farfán, Valeria A. Silva, Gino Nardocci, Martín Montecino, et al. 2019. “Mitochondrial NcRNA Targeting Induces Cell Cycle Arrest and Tumor Growth Inhibition of MDA-MB-231 Breast Cancer Cells through Reduction of Key Cell Cycle Progression Factors.” *Cell Death & Disease* 2019 10:6 10 (6): 1–12. <https://doi.org/10.1038/s41419-019-1649-3>.
- Franchi, A., G. Pasquinelli, G. Cenacchi, C. della Rocca, C. Gambini, M. Bisceglia, G. N. Martinelli, and M. Santucci. 2001. “Immunohistochemical and Ultrastructural Investigation of Neural Differentiation in Ewing Sarcoma/PNET of Bone and Soft Tissues.” *Ultrastructural Pathology* 25 (3): 219–25. <https://doi.org/10.1080/01913120120194>.
- Gandolfo, Luke C., and Terence P. Speed. 2018. “RLE Plots: Visualizing Unwanted Variation in High Dimensional Data.” *PLoS ONE* 13 (2): e0191629. <https://doi.org/10.1371/journal.pone.0191629>.
- García-Domínguez, Daniel J., Nabil Hajji, Sara Sánchez-Molina, Elisabet Figuerola-Bou, Rocío M. de Pablos, Ana M. Espinosa-Oliva, Eduardo Andrés-León, et al. 2021. “Selective Inhibition of HDAC6 Regulates Expression of the Oncogenic Driver EWSR1-FLI1 through the EWSR1 Promoter in Ewing Sarcoma.” *Oncogene* 40 (39): 5843–53. <https://doi.org/10.1038/s41388-021-01974-4>.
- Gil, Noa, and Igor Ulitsky. 2019. “Regulation of Gene Expression by Cis-Acting Long Non-Coding RNAs.” *Nature Reviews Genetics*. <https://doi.org/10.1038/s41576-019-0184-5>.

- Giulietti, Matteo, Alessandra Righetti, Giovanni Principato, and Francesco Piva. 2018. “LncRNA Co-Expression Network Analysis Reveals Novel Biomarkers for Pancreatic Cancer.” *Carcinogenesis* 39 (8): 1016–25. <https://doi.org/10.1093/CARCIN/BGY069>.
- Grünewald, Thomas G. P., Florencia Cidre-Aranaz, Didier Surdez, Eleni M. Tomazou, Enrique de Álava, Heinrich Kovar, Poul H. Sorensen, Olivier Delattre, and Uta Dirksen. 2018. “Ewing Sarcoma.” *Nature Reviews. Disease Primers* 4 (1): 5. <https://doi.org/10.1038/s41572-018-0003-x>.
- Guttman, Mitchell, Ido Amit, Manuel Garber, Courtney French, Michael F Lin, David Feldser, Maite Huarte, et al. 2009. “Chromatin Signature Reveals over a Thousand Highly Conserved Large Non-Coding RNAs in Mammals.” *Nature* 458 (7235): 223–27. <https://doi.org/10.1038/nature07672>.
- Han, Siyu, Yanchun Liang, Ying Li, and Wei Du. 2016. “Long Noncoding RNA Identification: Comparing Machine Learning Based Tools for Long Noncoding Transcripts Discrimination.” *BioMed Research International* 2016: 14. <https://doi.org/10.1155/2016/8496165>.
- Hancock, Jeffrey D, and Stephen L Lessnick. 2008. “A Transcriptional Profiling Meta-Analysis Reveals a Core EWS-FLI Gene Expression Signature.” *Cell Cycle (Georgetown, Tex.)* 7 (2): 250–56. <https://doi.org/10.4161/cc.7.2.5229>.
- Harrow, Jennifer, Adam Frankish, Jose M Gonzalez, Electra Tapanari, Mark Diekhans, Felix Kokocinski, Bronwen L Aken, et al. 2012. “GENCODE: The Reference Human Genome Annotation for The ENCODE Project.” *Genome Research* 22 (9): 1760–74. <https://doi.org/10.1101/gr.135350.111>.
- Hesla, Asle Charles, Andri Papakonstantinou, and Panagiotis Tsagkozis. 2021. “Current Status of Management and Outcome for Patients with Ewing Sarcoma.” *Cancers* 13 (6): 1202. <https://doi.org/10.3390/cancers13061202>.
- Hinger, Scott A., Diana J. Cha, Jeffrey L. Franklin, James N. Higginbotham, Yongchao Dou, Jie Ping, Lihua Shu, et al. 2018. “Diverse Long RNAs Are Differentially Sorted into Extracellular Vesicles Secreted by Colorectal Cancer Cells.” *Cell Reports* 25 (3): 715-725.e4. <https://doi.org/10.1016/j.celrep.2018.09.054>.
- Işın, Mustafa, Ege Uysaler, Emre Özgür, Hikmet Köseoğlu, Öner Şanlı, Ömer B. Yücel, Uğur Gezer, and Nejat Dalay. 2015. “Exosomal LncRNA-P21 Levels May Help to Distinguish Prostate Cancer from Benign Disease.” *Frontiers in Genetics* 6 (MAY): 168. <https://doi.org/10.3389/fgene.2015.00168>.

- Jalali, Saakshi, Shruti Kapoor, Ambily Sivadas, Deeksha Bhartiya, and Vinod Scaria. 2015. “Computational Approaches towards Understanding Human Long Non-Coding RNA Biology.” *Bioinformatics* 31 (14): 2241–51. <https://doi.org/10.1093/bioinformatics/btv148>.
- Jeong, H, B Tombor, R Albert, Z N Oltvai, and A L Barabasi. 2000. “The Large-Scale Organization of Metabolic Networks.” *Nature* 407 (6804): 651–54. <https://doi.org/10.1038/35036627>.
- Kang, Yilin, Michael James Baker, Michael Liem, Jade Louber, Matthew McKenzie, Ishara Atukorala, Ching-Seng Ang, Shivakumar Keerthikumar, Suresh Mathivanan, and Diana Stojanovski. 2016. “Tim29 Is a Novel Subunit of the Human TIM22 Translocase and Is Involved in Complex Assembly and Stability.” *ELife* 5 (AUGUST). <https://doi.org/10.7554/eLife.17463>.
- Karnuth, Bianca, Nicolas Dedy, Tilmann Spieker, Elizabeth R Lawlor, Stefan Gattenlohner, Andreas Ranft, Uta Dirksen, Heribert Jurgens, and Andreas Brauninger. 2014. “Differentially Expressed MiRNAs in Ewing Sarcoma Compared to Mesenchymal Stem Cells: Low MiR-31 Expression with Effects on Proliferation and Invasion.” *PloS One* 9 (3): e93067. <https://doi.org/10.1371/journal.pone.0093067>.
- Khanna, Nehal, Avinash Pandey, and Jyoti Bajpai. 2017. “Metastatic Ewing’s Sarcoma: Revisiting the ‘Evidence on the Fence’.” *Indian Journal of Medical and Paediatric Oncology : Official Journal of Indian Society of Medical & Paediatric Oncology* 38 (2): 173–81. [https://doi.org/10.4103/ijmpo.ijmpo\\_24\\_17](https://doi.org/10.4103/ijmpo.ijmpo_24_17).
- Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. 2019. “Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype.” *Nature Biotechnology* 37 (8): 907–15. <https://doi.org/10.1038/s41587-019-0201-4>.
- Kopp, Florian, and Joshua T. Mendell. 2018. “Functional Classification and Experimental Dissection of Long Noncoding RNAs.” *Cell* 172 (3): 393–407. <https://doi.org/10.1016/j.cell.2018.01.011>.
- Ladenstein, Ruth, Ulrike Pötschger, Marie Cécile le Deley, Jeremy Whelan, Michael Paulussen, Odile Oberlin, Henk van den Berg, et al. 2010. “Primary Disseminated Multifocal Ewing Sarcoma: Results of the Euro-EWING 99 Trial.” *Journal of Clinical Oncology* 28 (20): 3284–91. <https://doi.org/10.1200/JCO.2009.22.9864>.

- Langfelder, Peter, and Steve Horvath. 2008. “WGCNA: An R Package for Weighted Correlation Network Analysis.” *BMC Bioinformatics* 9: 559. <https://doi.org/10.1186/1471-2105-9-559>.
- Lawlor, Elizabeth R, and Carol J Thiele. 2012. “Epigenetic Changes in Pediatric Solid Tumors: Promising New Targets.” *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research* 18 (10): 2768–79. <https://doi.org/10.1158/1078-0432.CCR-11-1921>.
- Lawrence, Michael, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T. Morgan, and Vincent J. Carey. 2013. “Software for Computing and Annotating Genomic Ranges.” *PLoS Computational Biology* 9 (8): e1003118. <https://doi.org/10.1371/journal.pcbi.1003118>.
- Leavey, Patrick J., Nadia N. Laack, Mark D. Krailo, Allen Buxton, R. Lor Randall, Steven G. DuBois, Damon R. Reed, et al. 2021. “Phase III Trial Adding Vincristine-Topotecan-Cyclophosphamide to the Initial Treatment of Patients With Nonmetastatic Ewing Sarcoma: A Children’s Oncology Group Report.” *Journal of Clinical Oncology* 39 (36): 4029–38. <https://doi.org/10.1200/jco.21.00358>.
- Levetzow, Cornelia von, Xiaohua Jiang, Ynnez Gwye, Gregor von Levetzow, Long Hung, Aaron Cooper, Jessie Hao Ru Hsu, and Elizabeth R. Lawlor. 2011. “Modeling Initiation of Ewing Sarcoma in Human Neural Crest Cells.” *PLoS ONE* 6 (4): 1–10. <https://doi.org/10.1371/journal.pone.0019305>.
- Li, Aimin, Junying Zhang, and Zhongyin Zhou. 2014. “PLEK: A Tool for Predicting Long Non-Coding RNAs and Messenger RNAs Based on an Improved k-Mer Scheme.” *BMC Bioinformatics* 15 (1): 311. <https://doi.org/10.1186/1471-2105-15-311>.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, Suqing, Bin Li, Yuanting Zheng, Menglong Li, Leming Shi, and Xuemei Pu. 2017. “Exploring Functions of Long Noncoding RNAs across Multiple Cancers through Co-Expression Network.” *Scientific Reports* 7 (1): 754. <https://doi.org/10.1038/s41598-017-00856-8>.
- Li, Zhao, Lin Liu, Shuai Jiang, Qianpeng Li, Changrui Feng, Qiang Du, Dong Zou, Jingfa Xiao, Zhang Zhang, and Lina Ma. 2021. “LncExpDB: An Expression Database of Human Long

- Non-Coding RNAs.” *Nucleic Acids Research* 49 (D1): D962–68. <https://doi.org/10.1093/NAR/GKAA850>.
- Love, Michael I, Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lu, Qiongshi, Sijin Ren, Ming Lu, Yong Zhang, Dahai Zhu, Xuegong Zhang, and Tingting Li. 2013. “Computational Prediction of Associations between Long Non-Coding RNAs and Proteins.” *BMC Genomics* 14 (1): 1–10. <https://doi.org/10.1186/1471-2164-14-651/FIGURES/2>.
- Luo, Haitao, Dechao Bu, Liang Sun, Shuangfang Fang, Zhiyong Liu, and Yi Zhao. 2017. “Identification and Function Annotation of Long Intervening Noncoding RNAs.” *Briefings in Bioinformatics* 18 (5): 789–97. <https://doi.org/10.1093/bib/bbw046>.
- Ma, Yan, Michael Baltezor, Lian Rajewski, Jennifer Crow, Glenson Samuel, Vincent S. Staggs, Katherine M. Chastain, Jeffrey A. Toretsky, Scott J. Weir, and Andrew K. Godwin. 2019. “Targeted Inhibition of Histone Deacetylase Leads to Suppression of Ewing Sarcoma Tumor Growth through an Unappreciated EWS-FLI1/HDAC3/HSP90 Signaling Axis.” *Journal of Molecular Medicine*, April, 1–16. <https://doi.org/10.1007/s00109-019-01782-0>.
- Mallory, Allison C, and Alena Shkumatava. 2015. “LncRNAs in Vertebrates: Advances and Challenges.” *Biochimie* 117 (October): 3–14. <https://doi.org/10.1016/j.biochi.2015.03.014>.
- Marks, Paul A., Richard A. Rifkind, Victoria M. Richon, Ronald Breslow, Thomas Miller, and William K. Kelly. 2001. “Histone Deacetylases and Cancer: Causes and Therapies.” *Nature Reviews Cancer* 1 (3): 194–202. <https://doi.org/10.1038/35106079>.
- Marques Howarth, Michelle, David Simpson, Siu P Ngok, Bethsaida Nieves, Ron Chen, Zurab Siprashvili, Dedeepya Vaka, et al. 2014. “Long Noncoding RNA EWSAT1-Mediated Gene Repression Facilitates Ewing Sarcoma Oncogenesis.” *The Journal of Clinical Investigation* 124 (12): 5275–90. <https://doi.org/10.1172/JCI72124>.
- Mattick, John S., and John L. Rinn. 2015. “Discovery and Annotation of Long Noncoding RNAs.” *Nature Structural & Molecular Biology* 22 (1): 5–7. <https://doi.org/10.1038/nsmb.2942>.

- Meng, Linyan, Richard E. Person, and Arthur L. Beaudet. 2012. “Ube3a-ATS Is an Atypical RNA Polymerase II Transcript That Represses the Paternal Expression of Ube3a.” *Human Molecular Genetics* 21 (13): 3001–12. <https://doi.org/10.1093/hmg/dds130>.
- Milazzo, Giorgio, Daniele Mercatelli, Giulia di Muzio, Luca Triboli, Piergiuseppe de Rosa, Giovanni Perini, and Federico M. Giorgi. 2020. “Histone Deacetylases (HDACs): Evolution, Specificity, Role in Transcriptional Complexes, and Pharmacological Actionability.” *Genes* 11 (5). <https://doi.org/10.3390/GENES11050556>.
- Moazzam-Jazi, Maryam, Hossein Lanjanian, Samaneh Maleknia, Mehdi Hedayati, and Maryam S. Daneshpour. 2021. “Interplay between SARS-CoV-2 and Human Long Non-coding RNAs.” *Journal of Cellular and Molecular Medicine* 25 (12): 5823. <https://doi.org/10.1111/JCMM.16596>.
- Nieto, M. Angela. 2013. “Epithelial Plasticity: A Common Theme in Embryonic and Cancer Cells.” *Science (New York, N.Y.)* 342 (6159): 1234850. <https://doi.org/10.1126/science.1234850>.
- Nust, Daniel, Vanessa Sochat, Ben Marwick, Stephen J. Eglén, Tim Head, Tony Hirst, and Benjamin D. Evans. 2020. “Ten Simple Rules for Writing Dockerfiles for Reproducible Data Science.” *PLoS Computational Biology* 16 (11): 1–24. <https://doi.org/10.1371/journal.pcbi.1008316>.
- Ottina, Eleonora, Veera Panova, Laura Doglio, Anastasiya Kazachenka, Georgina Cornish, Joanna Kirkpatrick, Jan Attig, et al. 2021. “E3 Ubiquitin Ligase HECTD2 Mediates Melanoma Progression and Immune Evasion.” *Oncogene* 40 (37): 5567–78. <https://doi.org/10.1038/s41388-021-01885-4>.
- Ozaki, Toshifumi. 2015. “Diagnosis and Treatment of Ewing Sarcoma of the Bone: A Review Article.” *Journal of Orthopaedic Science*. <https://doi.org/10.1007/s00776-014-0687-z>.
- Palazzo, Alexander F., and Eliza S. Lee. 2015. “Non-Coding RNA: What Is Functional and What Is Junk?” *Frontiers in Genetics* 6 (JAN): 2. <https://doi.org/10.3389/fgene.2015.00002>.
- Park, Suk-Youl, and Jeong-Sun Kim. 2020. “A Short Guide to Histone Deacetylases Including Recent Progress on Class II Enzymes.” *Experimental & Molecular Medicine* 52 (2): 204–12. <https://doi.org/10.1038/s12276-020-0382-4>.
- Patro, Rob, Geet Duggal, Michael I Love, Rafael A Irizarry, and Carl Kingsford. 2017. “Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression.” *Nature Methods* 14 (4): 417–19. <https://doi.org/10.1038/nmeth.4197>.

- Pertea, Geo, and Mihaela Pertea. 2020. “GFF Utilities: GffRead and GffCompare.” *F1000Research* 9 (April): 304. <https://doi.org/10.12688/f1000research.23297.2>.
- Pertea, Mihaela, Daehwan Kim, Geo M Pertea, Jeffrey T Leek, and Steven L Salzberg. 2016. “Transcript-Level Expression Analysis of RNA-Seq Experiments with HISAT, StringTie and Ballgown.” *Nat Protocols* 11 (9): 1650–67. <https://doi.org/10.1038/nprot.2016.095>.
- Pertea, Mihaela, Geo M. Pertea, Corina M. Antonescu, Tsung-Cheng Chang, Joshua T. Mendell, and Steven L. Salzberg. 2015. “StringTie Enables Improved Reconstruction of a Transcriptome from RNA-Seq Reads.” *Nature Biotechnology* 33 (3): 290–95. <https://doi.org/10.1038/nbt.3122>.
- Prensner, John R, and Arul M Chinnaiyan. 2011. “The Emergence of LncRNAs in Cancer Biology,” no. October: 391–408. <https://doi.org/10.1158/2159-8290.CD-11-0209>.
- Quinn, Jeffrey J., and Howard Y. Chang. 2016. “Unique Features of Long Non-Coding RNA Biogenesis and Function.” *Nature Reviews Genetics* 17 (1): 47–62. <https://doi.org/10.1038/nrg.2015.10>.
- Rafehi, Haloom, and Assam El-Osta. 2016. “HDAC Inhibition in Vascular Endothelial Cells Regulates the Expression of NcRNAs.” *Non-Coding RNA* 2 (2): 4. <https://doi.org/10.3390/ncrna2020004>.
- Riggi, Nicolò, Luisa Cironi, Paolo Provero, Mario-Luca Suvà, Konstantinos Kaloulis, Carlos Garcia-Echeverria, Francesco Hoffmann, Andreas Trumpp, and Ivan Stamenkovic. 2005. “Development of Ewing’s Sarcoma from Primary Bone Marrow–Derived Mesenchymal Progenitor Cells.” *Cancer Research* 65 (24): 11459–68. <https://doi.org/10.1158/0008-5472.CAN-05-1696>.
- Riggi, Nicolò, Birgit Knoechel, Shawn M. Gillespie, Esther Rheinbay, Gaylor Boulay, Mario L. Suvà, Nikki E. Rossetti, et al. 2014. “EWS-FLI1 Utilizes Divergent Chromatin Remodeling Mechanisms to Directly Activate or Repress Enhancer Elements in Ewing Sarcoma.” *Cancer Cell* 26 (5): 668–81. <https://doi.org/10.1016/j.ccell.2014.10.004>.
- Riggi, Nicolo, Mario-Luca Suva, Claudio De Vito, Paolo Provero, Jean-Christophe Stehle, Karine Baumer, Luisa Cironi, et al. 2010. “EWS-FLI-1 Modulates MiRNA145 and SOX2 Expression to Initiate Mesenchymal Stem Cell Reprogramming toward Ewing Sarcoma Cancer Stem Cells.” *Genes & Development* 24 (9): 916–32. <https://doi.org/10.1101/gad.1899710>.
- Ruijter, Annemieke J M de, Albert H van Gennip, Huib N. Caron, Stephan Kemp, and André B P van Kuilenburg. 2003. “Histone Deacetylases (HDACs): Characterization of the

- Classical HDAC Family.” *The Biochemical Journal* 370 (Pt 3): 737–49. <https://doi.org/10.1042/BJ20021321>.
- Saleembhasha, A., and Seema Mishra. 2018. “Novel Molecules LncRNAs, TRFs and CircRNAs Deciphered from next-Generation Sequencing/RNA Sequencing: Computational Databases and Tools.” *Briefings in Functional Genomics* 17 (1): 15–25. <https://doi.org/10.1093/bfpg/elx013>.
- Salvador, Marion A., Julien Wicinski, Olivier Cabaud, Yves Toiron, Pascal Finetti, Emmanuelle Josselin, H el ene Leli evre, et al. 2013. “The Histone Deacetylase Inhibitor Abexinostat Induces Cancer Stem Cells Differentiation in Breast Cancer with Low Xist Expression.” *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research* 19 (23): 6520–31. <https://doi.org/10.1158/1078-0432.CCR-13-0877>.
- Samyn, Bart, Bart Samyn, Kjell Sergeant, Kjell Sergeant, Pedro Castanheira, Pedro Castanheira, Carlos Faro, Carlos Faro, Jozef Van Beeumen, and Jozef Van Beeumen. 2015. “Pathway and Network Analysis of Cancer Genomes.” *Nature Methods* 2 (3): 1–6. <https://doi.org/10.1038/NMETH>.
- Sand, Laurens G L, Karoly Szuhai, and Pancras C W Hogendoorn. 2015. “Sequencing Overview of Ewing Sarcoma: A Journey across Genomic, Epigenomic and Transcriptomic Landscapes.” *International Journal of Molecular Sciences* 16 (7): 16176–215. <https://doi.org/10.3390/ijms160716176>.
- Sannino, Giuseppina, Aruna Marchetto, Thomas Kirchner, and Thomas G.P. Gr unewald. 2017. “Epithelial-to-Mesenchymal and Mesenchymal-to-Epithelial Transition in Mesenchymal Tumors: A Paradox in Sarcomas?” *Cancer Research* 77 (17): 4556–61. <https://doi.org/10.1158/0008-5472.CAN-17-0032>.
- Schmidt, Oxana, Nadja Nehls, Carolin Prexler, Kristina von Heyking, Tanja Groll, Katharina Pardon, Heathcliff D. Garcia, et al. 2021. “Class I Histone Deacetylases (HDAC) Critically Contribute to Ewing Sarcoma Pathogenesis.” *Journal of Experimental & Clinical Cancer Research : CR* 40 (1): 322. <https://doi.org/10.1186/s13046-021-02125-z>.
- Schmitt, Adam M, and Howard Y Chang. 2016. “Long Noncoding RNAs in Cancer Pathways.” *Cancer Cell*. Elsevier Inc. <https://doi.org/10.1016/j.ccell.2016.03.010>.
- Selvanathan, Saravana P., Garrett T. Graham, Hayriye v. Erkizan, Uta Dirksen, Thanemozhi G. Natarajan, Aleksandra Dakic, Songtao Yu, et al. 2015. “Oncogenic Fusion Protein EWS-FLI1 Is a Network Hub That Regulates Alternative Splicing.” *Proceedings of the*

- National Academy of Sciences* 112 (11): E1307–16.  
<https://doi.org/10.1073/pnas.1500536112>.
- Shen, Xiao-han, Peng Qi, and Xiang Du. 2015. “Long Non-Coding RNAs in Cancer Invasion and Metastasis.” *Mod Pathol* 28 (1): 4–13. <https://doi.org/10.1038/modpathol.2014.75>.
- Sica, Valentina, José Manuel Bravo-San Pedro, Gautier Stoll, and Guido Kroemer. 2020. “Oxidative Phosphorylation as a Potential Therapeutic Target for Cancer Therapy.” *International Journal of Cancer* 146 (1): 10–17. <https://doi.org/10.1002/ijc.32616>.
- Siddiqui, Hasan, Julia Selich-Anderson, Joshua Felgenhauer, James Fitch, Vijay Nadella, Cenny Taslim, Laura Tomino, et al. 2018. “HOTAIR Primes the Ewing Sarcoma Family of Tumors for Tumorigenesis via Epigenetic Dysregulation Involving LSD1.” *BioRxiv*, January, 244558. <https://doi.org/10.1101/244558>.
- Siegel, Rebecca L., Kimberly D. Miller, Hannah E. Fuchs, and Ahmedin Jemal. 2021. “Cancer Statistics, 2021.” *CA: A Cancer Journal for Clinicians* 71 (1): 7–33. <https://doi.org/10.3322/caac.21654>.
- Signal, Bethany, Brian S. Gloss, and Marcel E. Dinger. 2016. “Computational Approaches for Functional Prediction and Characterisation of Long Noncoding RNAs.” *Trends in Genetics : TIG* 32 (10): 620–37. <https://doi.org/10.1016/j.tig.2016.08.004>.
- Soneson, Charlotte, Michael I. Love, and Mark D. Robinson. 2015. “Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences.” *F1000Research* 4 (December): 1521. <https://doi.org/10.12688/f1000research.7563.2>.
- Sonnemann, Jürgen, Linn Dreyer, Maite Hartwig, Chithra D. Palani, Le Thi Thu Hong, Ulrike Klier, Barbara Bröker, Uwe Völker, and James F. Beck. 2007. “Histone Deacetylase Inhibitors Induce Cell Death and Enhance the Apoptosis-Inducing Activity of TRAIL in Ewing’s Sarcoma Cells.” *Journal of Cancer Research and Clinical Oncology* 133 (11): 847–58. <https://doi.org/10.1007/s00432-007-0227-8>.
- Souza, Bárbara Kunzler, Patrícia Luciana da Costa Lopez, Pâmela Rossi Menegotto, Igor Araujo Vieira, Nathalia Kersting, Ana Lúcia Abujamra, André T. Brunetto, et al. 2018. “Targeting Histone Deacetylase Activity to Arrest Cell Growth and Promote Neural Differentiation in Ewing Sarcoma.” *Molecular Neurobiology* 55 (9): 7242–58. <https://doi.org/10.1007/s12035-018-0874-6>.
- Souza, Cristabelle, and Biswa Chatterji. 2015. “HDAC Inhibitors as Novel Anti-Cancer Therapeutics.” *Recent Patents on Anti-Cancer Drug Discovery* 10 (2): 145–62. <https://doi.org/10.2174/1574892810666150317144511>.

- Spector, Logan G., Aubrey K. Hubbard, Brandon J. Diessner, Mitchell J. Machiela, Beau R. Webber, and Joshua D. Schiffman. 2021. “Comparative International Incidence of Ewing Sarcoma 1988 to 2012.” *International Journal of Cancer* 149 (5): 1054–66. <https://doi.org/10.1002/ijc.33674>.
- Sun, Lishan, Cheng Yang, Juan Xu, Yanhong Feng, Ligu Wang, and Tao Cui. 2016. “Long Noncoding RNA EWSAT1 Promotes Osteosarcoma Cell Growth and Metastasis Through Suppression of MEG3 Expression.” *DNA and Cell Biology* 35 (12): 812–18. <https://doi.org/10.1089/dna.2016.3467>.
- Suvà, Mario-Luca, Nicolò Riggi, Jean-christophe Stehle, Karine Baumer, Stéphane Tercier, Jean-marc Joseph, Domizio Suvà, et al. 2009. “Identification of Cancer Stem Cells in Ewing’s Sarcoma.” *Cancer Research* 69 (5): 1776–81. <https://doi.org/10.1158/0008-5472.CAN-08-2242>.
- Takahashi, Kenji, Irene K Yan, Takayuki Kogure, Hiroaki Haga, and Tushar Patel. 2014. “Extracellular Vesicle-Mediated Transfer of Long Non-Coding RNA ROR Modulates Chemosensitivity in Human Hepatocellular Cancer.” *FEBS Open Bio* 4: 458–67. <https://doi.org/10.1016/j.fob.2014.04.007>.
- Tam, Shirley, Ming-Sound Tsao, and John D. McPherson. 2015. “Optimization of MiRNA-Seq Data Preprocessing.” *Briefings in Bioinformatics* 16 (6): 950–63. <https://doi.org/10.1093/bib/bbv019>.
- The RNAcentral Consortium. 2017. “RNAcentral: A Comprehensive Database of Non-Coding RNA Sequences.” *Nucleic Acids Research* 45 (D1): D128–34. <https://doi.org/10.1093/nar/gkw1008>.
- Tirole, Franck, Karine Laud-Duval, Alexandre Prieur, Bruno Delorme, Pierre Charbord, and Olivier Delattre. 2007. “Mesenchymal Stem Cell Features of Ewing Tumors.” *Cancer Cell* 11 (5): 421–29. <https://doi.org/10.1016/j.ccr.2007.02.027>.
- Tirole, Franck, Didier Surdez, Xiaotu Ma, Matthew Parker, Marie Cécile le Deley, Armita Bahrami, Zhaojie Zhang, et al. 2014. “Genomic Landscape of Ewing Sarcoma Defines an Aggressive Subtype with Co-Association of STAG2 and TP53 Mutations.” *Cancer Discovery* 4 (11): 1342–53. <https://doi.org/10.1158/2159-8290.CD-14-0622>.
- Todoerti, Katia, Domenica Ronchetti, Vanessa Favasuli, Francesco Maura, Fortunato Morabito, Niccolò Bolli, Elisa Taiana, and Antonino Neri. 2021. “DIS3 Mutations in Multiple Myeloma Impact the Transcriptional Signature and Clinical Outcome.” *Haematologica*, May. <https://doi.org/10.3324/haematol.2021.278342>.

- Tomazou, Eleni M., Nathan C. Sheffield, Christian Schmidl, Michael Schuster, Andreas Schönegger, Paul Datlinger, Stefan Kubicek, Christoph Bock, and Heinrich Kovar. 2015. “Epigenome Mapping Reveals Distinct Modes of Gene Regulation and Widespread Enhancer Reprogramming by the Oncogenic Fusion Protein EWS-FLI1.” *Cell Reports* 10 (7): 1082–95. <https://doi.org/10.1016/j.celrep.2015.01.042>.
- Tu, Jian, Zijun Huo, Julian Gingold, Ruiying Zhao, Jingnan Shen, and Dung-Fang Lee. 2017. “The Histogenesis of Ewing Sarcoma.” *Cancer Reports and Reviews* 1 (2): 417-424.e5. <https://doi.org/10.15761/CRR.1000111>.
- Uszczyńska-Ratajczak, Barbara, Julien Lagarde, Adam Frankish, Roderic Guigó, and Rory Johnson. 2018. “Towards a Complete Map of the Human Long Non-Coding RNA Transcriptome.” *Nature Reviews. Genetics* 19 (9): 535–48. <https://doi.org/10.1038/s41576-018-0017-y>.
- Vidal, Marc, Michael Cusick, and Albert-László Barabási. 2011. “Interactome Networks and Human Disease.” *Cell* 144: 986–98. <https://doi.org/doi:10.1016/j.cell.2011.02.016>.
- Vilanova, Joan C. 2017. “WHO Classification of Soft Tissue Tumors.” In *Imaging of Soft Tissue Tumors*, edited by Filip M. Vanhoenacker, Paul M. Parizel, and Jan L. Gielen. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-46679-8>.
- Walsh, Logan A, Mariano J Alvarez, Erich Y Sabio, Marsha Reyngold, Vladimir Makarov, Suranjit Mukherjee, Ken Wing Lee, et al. 2017. “An Integrated Systems Biology Approach Identifies TRIM25 as a Key Determinant of Breast Cancer Metastasis.” *Cell Reports* 20 (7): 1623–40. <https://doi.org/10.1016/j.celrep.2017.07.052>.
- Wang, Lei, Sara J. Felts, Virginia P. van Keulen, Adam D. Scheid, Matthew S. Block, Svetomir N. Markovic, Larry R. Pease, and Yuji Zhang. 2018. “Integrative Genome-Wide Analysis of Long Noncoding RNAs in Diverse Immune Cell Types of Melanoma Patients.” *Cancer Research* 78 (15): 4411–23. <https://doi.org/10.1158/0008-5472.CAN-18-0529>.
- Wilson, Greg, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Nederbragt, and Tracy K Teal. 2017. “Good Enough Practices in Scientific Computing.” *PLoS Computational Biology* 13 (6): e1005510. <https://doi.org/10.1371/journal.pcbi.1005510>.
- Wu, Zhuomin, Xiaoxia Liu, Li Liu, Houliang Deng, Jingjing Zhang, Qian Xu, Bohong Cen, and Aimin Ji. 2014. “Regulation of LncRNA Expression.” *Cellular & Molecular Biology Letters* 19 (4): 561–75. <https://doi.org/10.2478/s11658-014-0212-6>.
- Wucher, Valentin, Fabrice Legeai, Benoît Hédan, Guillaume Rizk, Lætitia Lagoutte, Tosso Leeb, Vidhya Jagannathan, et al. 2017. “FEELnc: A Tool for Long Non-Coding RNA

- Annotation and Its Application to the Dog Transcriptome.” *Nucleic Acids Research* 45 (8): e57. <https://doi.org/10.1093/nar/gkw1306>.
- Wysocka, Joanna, Michael P. Myers, Carol D. Laherty, Robert N. Eisenman, and Winship Herr. 2003. “Human Sin3 Deacetylase and Trithorax-Related Set1/Ash2 Histone H3-K4 Methyltransferase Are Tethered Together Selectively by the Cell-Proliferation Factor HCF-1.” *Genes & Development* 17 (7): 896–911. <https://doi.org/10.1101/GAD.252103>.
- Xing, Xiaohua, Yao Huang, Sen Wang, Minhui Chi, Yongyi Zeng, Lihong Chen, Ling Li, et al. 2015. “Dataset for the Quantitative Proteomics Analysis of the Primary Hepatocellular Carcinoma with Single and Multiple Lesions.” *Data in Brief* 5 (December): 226–40. <https://doi.org/10.1016/j.dib.2015.08.036>.
- Xiong, Jiachao, Liang Wu, Lu Huang, Chunyang Wu, Zhiming Liu, Wenqiang Deng, Shengbiao Ma, Zhenhai Zhou, Honggui Yu, and Kai Cao. 2021. “LncRNA FOXP4-AS1 Promotes Progression of Ewing Sarcoma and Is Associated With Immune Infiltrates.” *Frontiers in Oncology* 11 (October): 718876. <https://doi.org/10.3389/fonc.2021.718876>.
- Yan, Jingwen, Shannon L Risacher, Li Shen, and Andrew J Saykin. 2017. “Network Approaches to Systems Biology Analysis of Complex Disease: Integrative Methods for Multi-Omics Data.” *Briefings in Bioinformatics* 16 (2): 85–97. <https://doi.org/10.1093/bib/bbx066>.
- Yan, Wenying, Wenjin Xue, Jiajia Chen, and Guang Hu. 2016. “Biological Networks for Cancer Candidate Biomarkers Discovery.” *Cancer Informatics* 15: 1–7. <https://doi.org/10.4137/CIN.S39458>.
- Yang, Qing-qing, and Yan-fei Deng. 2015. “Genome-Wide Analysis of Long Non-Coding RNA in Primary Nasopharyngeal Carcinoma by Microarray.” *Histopathology* 66 (7): 1022–30. <https://doi.org/10.1111/his.12616>.
- Yang, Yang, Lei Chen, Jin Gu, Hanshuo Zhang, Jiawei Yuan, Qiuyu Lian, Guishuai Lv, et al. 2017. “Recurrently Deregulated LncRNAs in Hepatocellular Carcinoma.” *Nature Communications* 8 (February): 14421. <https://doi.org/10.1038/ncomms14421>.
- Ye, Yicong, Xiliang Zhao, Yiyun Lu, Bo Long, and Shuyang Zhang. 2018. “Varinostat Alters Gene Expression Profiles in Aortic Tissues from ApoE<sup>-/-</sup> Mice.” *Human Gene Therapy Clinical Development* 29 (4): 214–25. <https://doi.org/10.1089/humc.2018.141>.
- Ying, Liang, Qi Chen, Yawei Wang, Zhihua Zhou, Yiran Huang, and Feng Qiu. 2012. “Upregulated MALAT-1 Contributes to Bladder Cancer Cell Migration by Inducing

- Epithelial-to-Mesenchymal Transition.” *Molecular BioSystems* 8 (9): 2289–94. <https://doi.org/10.1039/c2mb25070e>.
- Yu, Jie, Tianli Xie, Zhe Wang, Xuwen Wang, Su Zeng, Yu Kang, and Tingjun Hou. 2019. “DNA Methyltransferases: Emerging Targets for the Discovery of Inhibitors as Potent Anticancer Drugs.” *Drug Discovery Today*, September. <https://doi.org/10.1016/J.DRUDIS.2019.08.006>.
- Zhang, Bin, and Steve Horvath. 2005. “A General Framework for Weighted Gene Co-Expression Network Analysis.” *Statistical Applications in Genetics and Molecular Biology* 4 (1): Article17. <https://doi.org/10.2202/1544-6115.1128>.
- Zhang, Jianfu, Jianzhao Peng, Yiran Huang, Ling Meng, Qingrong Li, Feng Xiong, and Xiaoyu Li. 2020. “Identification of Histone Deacetylase (HDAC)-Associated Proteins with DNA-Programmed Affinity Labeling.” *Angewandte Chemie (International Ed. in English)* 59 (40): 17525–32. <https://doi.org/10.1002/anie.202001205>.
- Zhao, Hongxing, Maoshan Chen, Sara Bergström Lind, and Ulf Pettersson. 2016. “Distinct Temporal Changes in Host Cell LncRNA Expression during the Course of an Adenovirus Infection.” *Virology* 492 (May): 242–50. <https://doi.org/10.1016/j.virol.2016.02.017>.
- Zhao, Yijing, Lemeng Sun, Rachel R. Wang, Ji Fan Hu, and Jiuwei Cui. 2018. “The Effects of Mitochondria-Associated Long Noncoding RNAs in Cancer Mitochondria: New Players in an Old Arena.” *Critical Reviews in Oncology/Hematology* 131 (April): 76–82. <https://doi.org/10.1016/j.critrevonc.2018.08.005>.
- Zhao, Yingdong, Ming-Chung Li, Mariam M. Konaté, Li Chen, Biswajit Das, Chris Karlovich, P. Mickey Williams, Yvonne A. Evrard, James H. Doroshov, and Lisa M. McShane. 2021. “TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-Seq Data from the NCI Patient-Derived Models Repository.” *Journal of Translational Medicine* 19 (1): 1–15. <https://doi.org/10.1186/s12967-021-02936-w>.
- Zou, Ren-Chao, Zhi-Tian Shi, Shu-Feng Xiao, Yang Ke, Hao-Ran Tang, Tian-Gen Wu, Zhi-Tang Guo, Fan Ni, Sanqi An, and Lin Wang. 2019. “Co-Expression Analysis and CeRNA Network Reveal Eight Novel Potential LncRNA Biomarkers in Hepatocellular Carcinoma.” *PeerJ* 7: e8101. <https://doi.org/10.7717/peerj.8101>.