

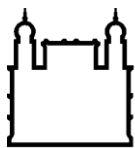
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Mestrado em Biologia Computacional e Sistemas

ANÁLISE DO PERFIL DE EXPRESSÃO DE MICRORNAs, RNAs QUE
INTERAGEM COM PIWI E RNAs NUCLEOLARES PEQUENOS EM
DIFERENTES SUBTIPOS DE CÂNCER DE TIROIDE

MAYLA ABRAHIM COSTA

Rio de Janeiro
Abril de 2017



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

MAYLA ABRAHIM COSTA

Análise do perfil de expressão de microRNAs, RNAs que interagem com PIWI e RNAs nucleolares pequenos em diferentes subtipos de câncer de tireoide

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Ciências

Orientador: Dr. Fabio Passetti

RIO DE JANEIRO

Abril de 2017

Abraham, Mayla .

Análise do perfil de expressão de microRNAs, RNAs que interagem com PIWI e RNAs nucleolares pequenos em diferentes subtipos de câncer de tireoide / Mayla Abraham. - Rio de Janeiro, 2017.

83 f.; il.

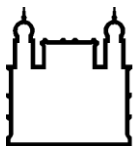
Dissertação (Mestrado) - Instituto Oswaldo Cruz, Pós-Graduação em Biologia Computacional e Sistemas, 2017.

Orientador: Fabio Passetti.

Bibliografia: f. 59-70

1. pequenos RNAs não codificadores. 2. bioinformática. 3. câncer de tireoide. 4. piRNA. 5. snoRNA. I. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da Biblioteca de Manguinhos/ICICT com os dados fornecidos pelo(a) autor(a).



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AUTOR: MAYLA ABRAHIM COSTA

**ANÁLISE DO PERFIL DE EXPRESSÃO DE MICRORNA, RNAs QUE INTERAGEM
COM PIWI E RNAs NUCLEOLARES PEQUENOS EM DIFERENTES SUBTIPOS
DE CÂNCER DE TIROIDE**

ORIENTADOR: Dr. Fabio Passetti

Aprovada em: 24/04/2017

EXAMINADORES:

Prof. Dr. Thiago Estevam Parente Martins - Presidente (IOC/FIOCRUZ)

Prof. Dr^a. Edna Teruko Kimura (USP)

Prof. Dr. Diogo Antonio Tschoeke (UFRJ)

Prof. Dr. Antonio Basilio de Miranda (IOC/FIOCRUZ)

Prof. Dr. Francisco Prosdocimi (UFRJ)

Rio de Janeiro, 24 de Abril de 2017



Ministério da Saúde

Fundação Oswaldo Cruz
Instituto Oswaldo Cruz

DECLARAÇÃO

Declaramos, para fins curriculares, que **Mayla Abraham Costa**, sob orientação do Dr. Fabio Passetti, foi aprovada em 24/04/2017, em sua defesa de dissertação de mestrado intitulada: **“Análise do perfil de expressão de microRNAs, RNAs que interagem com PIWI e RNAs nucleolares pequenos em diferentes subtipos de câncer de tireoide”**, área de concentração: Genômica Funcional, Evolução e Filogenômica. A banca examinadora foi constituída pelos Professores: Dr. Thiago Estevam Parente Martins - IOC/FIOCRUZ (presidente), Dr. Diogo Antonio Tschoeke - UFRJ/RJ e Dr^a. Edna Teruko Kimura - USP/SP e como suplentes: Dr. Antonio Basilio de Miranda - IOC/FIOCRUZ e Dr. Francisco Prosdocimi - UFRJ/RJ.

A Pós-graduação *Stricto sensu* em Biologia Computacional e Sistemas (Mestrado e Doutorado) está credenciada pela CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, de acordo com Portaria n° 1.146, de 12 de setembro de 2008, tendo validade no Brasil para todos os efeitos legais, e alcançando conceito 4 (quatro) na última avaliação trienal da CAPES.

Informamos ainda que, de acordo com as normas do Programa de Pós-graduação, **a liberação do Diploma e do Histórico Escolar está condicionada à entrega da versão definitiva da dissertação/tese em capa espiral (1 cópia), juntamente com o termo de autorização de divulgação da dissertação/tese on line e o CD-rom com a dissertação completa em PDF.**

Rio de Janeiro, 24 de abril de 2017.

Dr. Ernesto Raúl Caffarena
Matrícula SIAPE 13555863

Coordenador do Programa de Pós-Graduação em Biologia Computacional e Sistemas
Instituto Oswaldo Cruz/Fiocruz

Com muito carinho dedico este trabalho aos meus pais Marcia Abraham e Eduardo Costa pelo incondicional incentivo aos meus sonhos, por tudo que fizeram por mim até aqui e por sempre estarem ao meu lado em todos os momentos.

AGRADECIMENTOS

Agradeço primeiramente a Deus por me abençoar e me iluminar até aqui.

Aos meus pais Marcia Abraham e Eduardo Costa pela excelente criação e por tudo que sempre fizeram por mim. Por sempre estarem ao meu lado, me apoiando e incentivando, me dando ótimos exemplos, amizade, muito carinho e amor, que foram fundamentais na construção do meu caráter. Sem eles nada na minha vida seria possível.

As minhas companheiras de quatro patas, Lolly, Lilly e Lolla, fonte renovadora de energia.

Aos meus amigos que estiveram ao meu lado me ajudando, apoiando e incentivando direta ou indiretamente, contribuindo que eu pudesse crescer. À Thais Martins pelo companheirismo e amizade "*Cause, baby, you're a firework*" e ao Edson Machado pelo carinho, amizade e assistência.

Ao Dr. Fabio Passeti por acreditar que eu era capaz, recebendo-me sem conhecer meu potencial, que abriu as portas, como um pai que abre os braços para receber um filho. Por todas as oportunidades concedidas, aprendizado, incentivo, confiança depositada, por todo comprometimento profissional perante seus discentes, disposto a auxiliar e orientar de fato. Fonte inesgotável para a minha motivação, dedicação e equilíbrio na realização do projeto.

Aos meus colegas de laboratório em especial ao Dr. Rafael Tavares por sempre estar no lugar certo na hora certa. À Natasha por toda paciência, ensinamentos, sempre disposta a contribuir positivamente em todos os momentos permitindo os melhores resultados na confecção do projeto.

À pós-graduação em Biologia Computacional e Sistemas por todo o apoio, em especial à Rose pela disponibilidade e contribuição e ao Dr. Ernesto Caffarena que sempre esteve presente e disposto a ajudar, principalmente no momento de grande dificuldade conseguiu transforma-lo em uma oportunidade positiva.

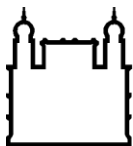
Ao Dr. Rodrigo Jardim por me mostrar os encantos da bioinformática e todos os ensinamentos durante os primeiros passos na iniciação científica perdurando até o mestrado, um exemplo como pessoa e pesquisador.

Aos doutores da banca, por aceitarem o convite, todos os conselhos e correções.

Ao CNPq pelo auxílio financeiro.

*“Look deep into nature, and then you will
understand everything better”*

Albert Einstein



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

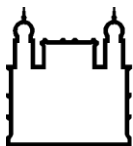
ANÁLISE DO PERFIL DE EXPRESSÃO DE MICRORNA, RNAs QUE INTERAGEM COM PIWI E RNAs NUCLEOLARES PEQUENOS EM DIFERENTES SUBTIPOS DE CÂNCER DE TIROIDE

RESUMO

DISSERTAÇÃO DE MESTRADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Mayla Abraham Costa

O câncer de tireoide é um relevante problema de saúde pública por ser o tumor maligno mais comum do sistema endócrino. Para a obtenção de melhores tratamentos, estratégias distintas vêm sendo desenvolvidas, dentre elas a identificação de marcadores moleculares. Ao longo dos anos, RNAs não codificadores (ncRNA) passaram a ser identificados com a finalidade de serem usados como potenciais marcadores moleculares e alvos terapêuticos. Os ncRNAs pequenos (sncRNA) desempenham papel importante na ocorrência do câncer e resposta ao tratamento, a exemplo dos microRNAs (miRNA), RNAs que interagem com PIWI (piRNA) e RNAs nucleolares pequenos (snoRNAs). A fim de avaliar a expressão de sncRNAs das famílias miRNA, piRNA e snoRNA em amostras dos diferentes subtipos de câncer de tireoide, foram analisados dados públicos de sequenciamento de alto desempenho de amostras de câncer de tireoide dos subtipos tumorais carcinoma papilífero (PTC) (49 pacientes), carcinoma papilífero, variante folicular, (PCF) (7 pacientes) e carcinoma papilífero, variante Células Altas, (PCC) (3 pacientes), totalizando 118 amostras pareadas de amostra tumorais e normais adjacentes ao tumor. Ao todo foram obtidos 46 sncRNA diferencialmente expressos, dos quais 40 são miRNAs, sendo 21 mais e 19 menos expressos em amostras tumorais do que em amostras normais adjacentes aos tumores. Foram identificados seis snoRNAs, dois mais e quatro menos expressos em amostras tumorais. Consideramos como constitutivamente expressos 34 sncRNAs, dentre os quais, destacando o piRNA hsa-piR-009294 constitutivamente expresso em todos os subtipos tumorais. Ao integrar os dados da expressão diferencial e de dispersão, foram identificados três miRNAs que apresentaram padrão de expressão similar nos subtipos tumorais PCC e PTC comparado ao padrão de expressão encontrado em amostras adjacentes aos tumores e amostras tumorais do subtipo tumoral PCF. Tais resultados mostram que as ferramentas usadas nesse trabalho são eficientes para identificar sncRNAs diferencialmente e constitutivamente expressos em amostras dos diferentes subtipos de câncer de tireoide. Os resultados encontrados no presente estudo corroboram dados obtidos por outros autores, bem como apresentam resultados inéditos, evidenciando uma alternativa viável na obtenção de novos potenciais marcadores moleculares.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

ANALYSIS OF EXPRESSION PROFILE THE MICRORNA, PIWI-INTERACTING RNA AND SMALL NUCLEOLAR RNA IN DIFFERENT SUBTYPES OF THE THYROID CANCER

ABSTRACT

MASTER DISSERTATION IN SYSTEMS AND COMPUTATIONAL BIOLOGY

Mayla Abraham Costa

Thyroid cancer is a public health problem and is considered the most common malignant tumor of the endocrine system. Aiming to obtain better treatments, different strategies have been developed, including the identification of molecular markers. Over the years, non-coding RNAs (ncRNAs) have been identified as potential molecular markers with the purpose of identifying novel therapeutic interventions. Studies show that small ncRNAs (sncRNAs), such as microRNAs (miRNAs), Piwi-interacting RNA (piRNA) and small nucleolar RNAs (snoRNAs) play important roles in cancer and response to treatment. In order to identify the miRNA, piRNA and snoRNAs constitutively and differentially expressed in samples of the different thyroid cancer subtypes. We analysed public data from high-throughput sequencing of thyroid cancer samples from the tumor subtypes carcinoma papillary (PTC) (49 patients), carcinoma papillary, follicular variant, (PCF) (7 patients), and carcinoma papillary, columnar cell variant, (PCC) (3 patients), accounting for 118 paired samples. Forty six differentially expressed sncRNAs were obtained, of which 40 were miRNAs. Among them, 21 were detected as upregulated in tumor samples and 19 were downregulated. A total of six snoRNAs have been detected in all tumor subtype, 2 more expressed in tumor samples and 4 snoRNAs more expressed in normal samples. We identified 34 constitutively expressed sncRNAs and the piRNA has-piR-009294 can be highlighted as detected in the 3 tumor subtypes. The integration of the differential expression and dispersion analysis revealed three miRNAs presenting similar expression pattern in tumor subtypes PCC and PTC when compared to the constitutive expression pattern in normal and tumor samples of the PCF subtype. These results show that it was possible to detect sncRNAs differentially and constitutively expressed in samples of the different thyroid cancer subtypes. Our results corroborate those obtained by others and present novel findings, evidencing a viable alternative to search for novel potential molecular markers.

SUMÁRIO

RESUMO.....	IX
ABSTRACT.....	X
1. INTRODUÇÃO.....	1
1.1. BIOLOGIA COMPUTACIONAL E O SEQUENCIAMENTO.....	1
1.1.1. Bases de dados públicas	2
1.1.2. Reprodutibilidade de resultados	3
1.1.2.1. Reuso de dados públicos	5
1.2. CÂNCER.....	5
1.2.1. O que é o câncer.....	6
1.2.2. Câncer no mundo, fatores e ocorrência	6
1.2.3. Fatores que levam ao câncer.....	6
1.3. CÂNCER DE TIROIDE	7
1.3.1. A Tireoide.....	7
1.3.2. Os diferentes tipos de câncer de tireoide	8
1.3.2.1. Câncer papilífero.....	8
1.3.2.1.1. Carcinoma papilífero, variante folicular	9
1.3.2.1.2. Carcinoma papilífero, variante Células Altas.....	9
1.3.2.2. Câncer folicular.....	10
1.3.2.3. Adenonoma.....	10
1.3.2.4. Câncer anaplásico	11
1.3.3. Genética do câncer de tireoide	11
1.3.4. Prognóstico e biomarcadores.....	11
1.3.4.1. BIOMARCADORES	12
1.4. RNAs NÃO CODIFICADORES.....	12
1.4.1. RNAs não codificadores pequenos	13
1.4.1.1. MicroRNA	13
1.4.1.1.1. Biogênese do microRNA	14
1.4.1.2. RNA que interage com PIWI.....	16
1.4.1.2.1. Biogênese do piRNA	17
1.4.1.3. RNAs nucleolares pequenos.....	18
1.4.1.3.1. SnoRNA C/D <i>box</i>	19
1.4.1.3.2. SnoRNA H/ACA <i>box</i>	19
1.4.1.3.3. Biogênese do snoRNA	20
1.5. JUSTIFICATIVA	21
2. OBJETIVOS	23
2.1. OBJETIVO GERAL	23
2.2. OBJETIVOS ESPECÍFICOS	23
3. MATERIAL E MÉTODOS.....	24
3.1. OBTENÇÃO DOS DADOS	24
3.1.1. Dados de sequenciamento de alto desempenho das amostras pareadas de tecido normal adjacente ao tumor e do tumor do mesmo paciente 24	
3.1.2. Obtenção dos dados de sequenciamento e anotação	25
3.2. PRÉ-PROCESSAMENTO	25

3.2.1.	Conversão e alinhamento contra o genoma humano.....	25
3.3.	MAPEAMENTO, IDENTIFICAÇÃO E CONTAGEM DOS SNCRNAS	26
3.4.	NORMALIZAÇÃO E IDENTIFICAÇÃO DA EXPRESSÃO DIFERENCIAL.....	26
3.5.	ANÁLISE DA DISPERSÃO	26
4.	RESULTADOS	28
4.1.	PRÉ-PROCESSAMENTO E ALINHAMENTO	28
4.2.	IDENTIFICAÇÃO E CONTAGEM DOS SNCRNAS	31
4.3.	COMPARAÇÃO ENTRE AS ABORDAGENS	32
4.4.	IDENTIFICAÇÃO DA EXPRESSÃO DIFERENCIAL	33
4.4.1.	Identificação da expressão diferencial de sncRNA	34
4.4.2.	Identificação da expressão diferencial de miRNA	36
4.4.3.	Identificação da expressão diferencial de piRNA	38
4.4.4.	Identificação da expressão diferencial de snoRNA	39
4.5.	ANÁLISE DE DISPERSÃO DOS DADOS DE EXPRESSÃO DE SNCRNAS	41
4.5.1.	MiRNA construtivamente expressos	42
4.5.2.	PiRNAs constitutivamente expressos.....	43
4.5.3.	SnoRNAs constitutivamente expressos	44
4.6.	INTEGRAÇÃO DAS EXPRESSÕES DIFERENCIAL E CONSTITUTIVA DE SNCRNAS	45
5.	DISCUSSÃO.....	46
6.	PERSPECTIVAS	56
7.	CONCLUSÕES.....	57
8.	REFERÊNCIAS BIBLIOGRÁFICAS	58
9.	ANEXOS.....	70
9.1.	ANEXO 1	70
9.2.	ANEXO 2	76
9.3.	ANEXO 3	79

LISTA DE FIGURAS

Figura 1.1: Custo em dólares do sequenciamento por genoma entre os anos de 2001 e 2015.	2
Figura 1.2: Anatomia da tiroide.....	7
Figura 1.3: Resumo dos principais tipos de câncer de tiroide derivados das células C e células foliculares e subtipos do câncer papilífero.....	8
Figura 1.4: Resumo da classificação dos RNAs não codificadores. De acordo com seu tamanho os RNAs não codificadores podem ser divididos em: RNAs não codificadores longos (lncRNAs) e RNAs não codificadores pequenos (sncRNAs)	13
Figura 1.5: Biogênese resumida do miRNA.....	16
Figura 1.6: Biogênese resumida do piRNA	18
Figura 1.7: Estrutura dos subtipos de snoRNA. A) <i>Box C/D</i> B) <i>Box H/ACA</i>	20
Figura 1.8: Biogênese do snoRNA.....	21
Figura 3.1: Fluxograma ilustrando os passos das análises realizadas nesse estudo.	27
Figura 4.1: Diagrama de Venn comparando a quantidade de miRNA diferencialmente expressos encontrados no presente estudo (vermelho) e a abordagem desenvolvida por Mancikova e colaboradores (2015) (azul).	33
Figura 4.2: Diagrama de Venn comparando a quantidade total de sncRNAs mais expressos em amostras tumorais ($FDR < 0,05$ e $\log FC \geq 1,5$) presentes nos três diferentes subtipos de câncer de tiroide.....	35
Figura 4.3: Diagrama de Venn comparando a quantidade total de sncRNAs menos expressos ($FDR < 0,05$ e $\log FC \leq -1,5$) presentes nos três diferentes subtipos de câncer de tiroide.....	36
Figura 4.4: Diagrama de Venn comparando a quantidade de miRNA mais expressos ($FDR < 0,05$ e $\log FC \geq 1,5$) presentes nos três diferentes subtipos de câncer de tiroide.....	37
Figura 4.5: Diagrama de Venn comparando a quantidade de miRNA menos expressos ($FDR < 0,05$ e $\log FC \leq -1,5$) presentes nos três diferentes subtipos de câncer de tiroide	38

Figura 4.6: Diagrama de Venn comparando a quantidade de snoRNA mais expressos (FDR < 0,05 e logFC ≥ 1,5) presentes nos três diferentes subtipos de câncer de tireoide.....	39
Figura 4.7: Diagrama de Venn comparando a quantidade de snoRNA menos expressos (FDR < 0,05 e logFC ≥ 1,5 ou ≤ -1,5) presentes nos três diferentes subtipos de câncer de tireoide.....	40
Figura 9.1: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero variante Células Altas.....	76
Figura 9.2: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero variante folicular	77
Figura 9.3: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero.....	78

LISTA DE TABELAS

Tabela 1.1: Resumo dos RNAs não codificadores pequenos analisados no presente estudo: miRNA, piRNA e snoRNA.....	13
Tabela 3.1: Subtipos de câncer de tireoide, número de pacientes com dados pareados de amostras normais adjacentes aos tumores e tumorais do mesmo paciente utilizadas no trabalho	24
Tabela 3.2: Metadados dos dados obtidos na base de dados The Cancer Genome Atlas	25
Tabela 4.1: Análise quantitativa dos <i>reads</i> sequenciados, dos <i>reads</i> mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero, variante Células Altas.....	28
Tabela 4.2: Análise quantitativa dos <i>reads</i> sequenciados, dos <i>reads</i> mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero, variante folicular	29
Tabela 4.3: Análise quantitativa dos <i>reads</i> sequenciados, dos <i>reads</i> mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero.....	30
Tabela 4.4: Quantidade de <i>reads</i> alinhados nas regiões dos RNAs não codificadores pequenos: miRNA, piRNA e snoRNA por subtipo tumoral PCC (Carcinoma papilífero variante Células Altas), PCF (Carcinoma papilífero variante folicular) e PTC (Carcinoma papilífero) através do qual foram verificados a menor (Min.), maior (Max.) e a média total dos sncRNAs em cada amostra tumoral (Med.).....	32
Tabela 4.5: Comparação dos miRNAs diferencialmente expressos encontrados no presente trabalho e pela abordagem de Mancikova e colaboradores (2015)	33
Tabela 4.6: Comparação quantitativa dos sncRNAs diferencialmente expressos utilizando os filtros FDR menor que 0,05 e logFC maior ou igual a 1,5 ou menor ou igual a -1,5	34

Tabela 4.7: SncRNAs mais expressos (FDR < 0,05 e logFC ≥ 1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	35
Tabela 4.8: SncRNAs menos expressos (FDR < 0,05 e logFC ≤ -1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	36
Tabela 4.9: MiRNAs mais expressos (FDR < 0,05 e logFC ≥ 1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	37
Tabela 4.10: MiRNAs menos expressos (FDR < 0,05 e logFC ≤ -1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	38
Tabela 4.11: SnoRNAs mais expressos (FDR < 0,05 e logFC ≥ 1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	40
Tabela 4.12: SnoRNAs menos expressos (FDR < 0,05 e logFC ≤ -1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.....	41
Tabela 4.13: Quantidade de sncRNAs com expressão uniforme nos subtipos tumorais Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas utilizando os filtros: desvio padrão inferiores 1 e logFC entre 1 e -1.....	42
Tabela 4.14: Quantidade dos sncRNAs constitutivamente expressos e menor desvio padrão nos subtipos tumorais Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas desvio padrão inferiores 0,5 e logFC entre 0,1 e -0,1.....	42
Tabela 4.15: MiRNAs constitutivamente expressos e com menor desvio padrão, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.....	43
Tabela 4.16: PiRNAs constitutivamente expressos, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.....	44

Tabela 4.17: SnoRNAs constitutivamente expressos e com menor desvio padrão, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.....	44
Tabela 4.18: Integração da expressão diferencial e constitutiva utilizando os sncRNAs diferencialmente expressos, logFC, FDR e desvio padrão (DP).	45
Tabela 9.1: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA, divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero.....	70

LISTA DE SIGLAS E ABREVIATURAS

10A	Adenina na 10 ^a posição
1U	Uracila na região 5'
Ago3	Argonaute 3
ATC	Câncer Anaplásico da Tireoide
Aub	Aubergine
DGE	Expressão de genes de dados digitais
DNA	Ácido desoxirribonucleico
FDR	Taxa de falsa descoberta
FTC	Câncer folicular da tireoide
lncRNA	Longo RNA não codificador
LogFC	Logaritmo de <i>Fold Change</i>
MAPK	Proteína quinase ativada por mitógeno9
miRNA	MicroRNA
mRNA	RNA mensageiro
MTC	Câncer medular da tireoide
NCI	Do inglês: <i>National Cancer Institute</i>
ncRNA	RNA não codificador
NGS	Sequenciamento de nova geração
NHGRI	Do inglês: <i>National Human Genome Research Institute</i>
NOS	Do inglês: <i>Not Otherwise Specified</i>
PCC	Carcinoma papilífero, variante células altas
PCF	Carcinoma papilífero, variante folicular
piRNA	RNA que interage com PIWI
Pré-miRNA	Precursor de miRNA
Pré-piRNA	Precursor de piRNA precursores
Pri-miRNA	miRNA primário
PTC	Câncer papilífero da tireoide
RISC	Complexo de Indução do Silenciamento do RNA
RNA	Ácido ribonucleico
rRNA	RNA ribossomal
RPKM	Do inglês: <i>Reads per kilobase per million mapped reads</i>

scaRNA	Pequeno RNA do corpo Cajal
siRNA	RNAs de interferência pequenos
sncRNA	RNA não codificador pequeno
SNORD	SnoRNA C/D box
snoRNA	RNA nucleolar pequeno
snoRNP	ribonucleoproteína nucleolar pequena
T3	Triiodotironina
T4	Tiroxina
TCGA	Do inglês: <i>The Cancer Genome Atlas</i>
TE	Elemento de transposição

1. INTRODUÇÃO

1.1. Biologia computacional e o sequenciamento

Durante o início da década de 1960, os computadores emergiram como ferramentas importantes para a biologia surgindo assim, a biologia computacional (HAGEN, 2000). Uma nova ciência que envolve diversas linhas de conhecimentos como a tecnologia da computação, matemática e biologia molecular, que combinadas podem responder perguntas fundamentais nas ciências da vida (HAGEN, 2000). O desenvolvimento dos sequenciadores automáticos de DNA exigiu recursos computacionais cada vez mais eficientes para analisar e armazenar a grande quantidade de dados gerados (BANSAL, 2005).

Descrito pela primeira vez em 1977, o sequenciamento de Sanger surgiu como uma técnica para determinar a sequência (ou: as sequências) de nucleotídeos de um genoma (SANGER *et al.*, 1977). A partir de 2005, a abordagem tradicional baseada em Sanger para sequenciamento de DNA sofreu mudanças inovadoras (MARDIS, 2013), assim o método de Sanger foi considerado como uma tecnologia de primeira geração e métodos mais recentes são referidos como sequenciamento de nova geração (NGS).

As tecnologias de NGS reduziram drasticamente o tempo e custo para sequenciar genomas grandes quando comparadas a tecnologia de primeira geração (HENSON *et al.*, 2014) (Figura 1). Nos últimos anos, o NGS tem sido utilizado para diversos estudos como: metagenômica, doenças complexas, incluindo doenças cardíacas e neurológicas, bem como, a detecção de mutações e rastreamento da expressão de diversos tipos de RNA em diferentes condições biológicas (BENNETT & FARAH, 2014; MORINI *et al.*, 2015; HU, *et al.*, 2016; LAVEZZO *et al.*, 2016). Assim, com o barateamento do custo de sequenciamento por nucleotídeo, uma quantidade crescente de dados oriundos dessa tecnologia tem sido depositados em bases de dados públicas.

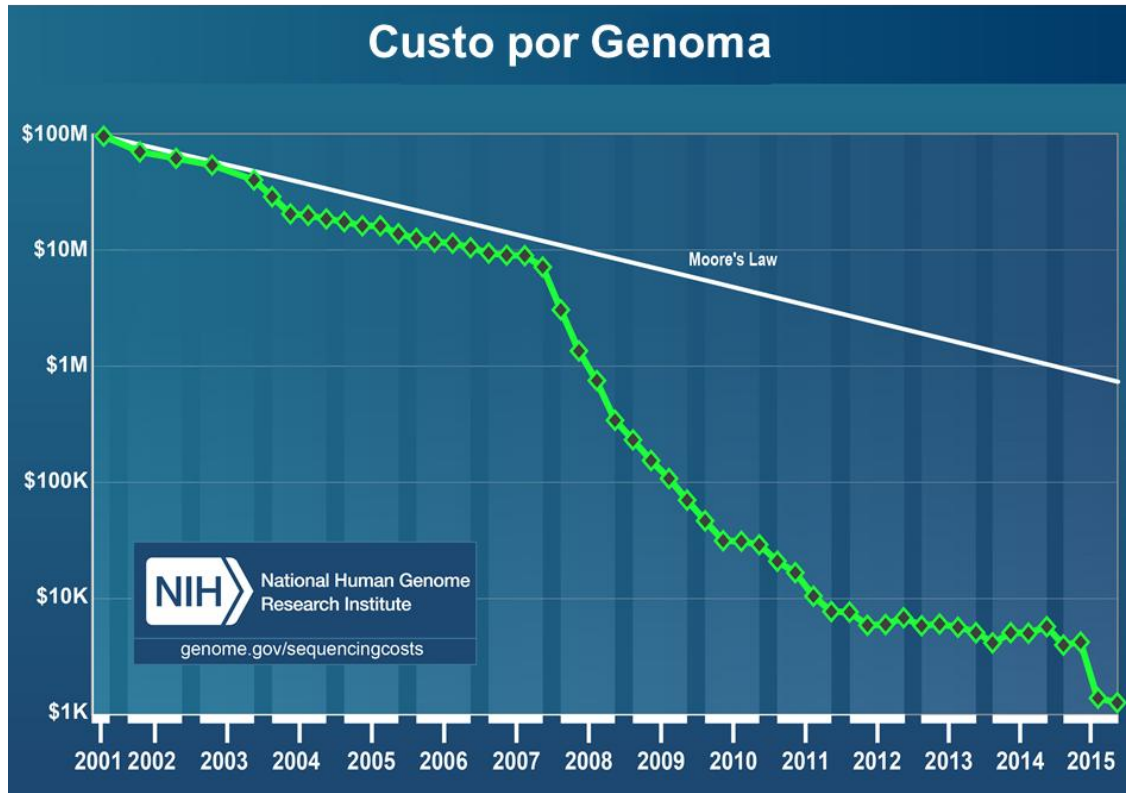


Figura 1.1: Custo em dólares do sequenciamento por genoma entre os anos de 2001 e 2015.
Fonte: Adaptado de genome.gov/sequencingcostsdata/

1.1.1. Bases de dados públicas

Uma parte da ciência adotou o reuso de dados públicos oriundos de métodos de produção de dados em larga escala, seja em genômica, proteômica ou análise de expressão gênica para o estudo de células, tecidos ou organismos inteiros. Assim sendo, há a necessidade crescente de bases de dados capazes de organizar e disponibilizar tais dados, bem como recursos computacionais cada vez mais eficientes para processar e analisar tanta informação (RICE *et al.*, 2000; PROSDOCIMI *et al.*, 2002). À medida que o volume e a variedade dos dados públicos se expandem, ferramentas que ajudam na pesquisa, visualização, organização e recuperação destes se tornam cada vez mais úteis (HINRICHS *et al.*, 2006).

Projetos genômica melhoraram consideravelmente o estudo e as aplicações na saúde e outros campos da biotecnologia (LIU *et al.*, 2012). Foram criados repositórios específicos para dados de NGS voltados para a pesquisa e o desenvolvimento de estratégias preventivas, métodos de diagnóstico e terapias contra o câncer, como por exemplo, *The Cancer Genome Atlas* (TCGA) em 2005 e o

International Cancer Genome Consortium (ICGC) em 2008. (HUDSON *et al.*, 2010; TOMCZAK *et al.*, 2015).

O TCGA é um projeto público que tem como objetivo catalogar e descobrir grandes alterações genômicas relacionadas ao desenvolvimento e câncer em mais de 30 tumores humanos. Através de dados de genômica em larga escala e análises multidimensionais integradas, há o propósito de melhorar os métodos de diagnóstico, padrões de tratamento e, finalmente, a prevenção do câncer com a disponibilização pública dos conjuntos de dados (TOMCZAK *et al.*, 2015).

O ICGC tem objetivo de obter uma descrição abrangente e catalogar as alterações genômicas, transcriptômicas e epigenômicas em 50 diferentes tipos de tumores e subtipos que são de importância clínica e social em todo o mundo. Os dados possuem alta qualidade, são disponibilizados publicamente e possuem restrições mínimas, a fim de acelerar as pesquisas sobre as causas e o controle do câncer. Além de facilitar a comunicação entre os membros e fornecer um fórum para maximizar a eficiência entre os cientistas que trabalham para compreender, tratar e prevenir essas doenças (HUDSON *et al.*, 2010).

A fase piloto do projeto GTEx (do inglês *Genotype-Tissue Expression*) terminou em janeiro de 2013 e o conjunto de dados de alta qualidade foi constituído por dados de 190 indivíduos (LONSDALE *et al.*, 2013) com finalidade de propiciar análises genômicas extensas, correlação entre a variação genética humana, expressão gênica específica entre o tecido de indivíduos não-doentes e avaliar como a correlação entre as variações genéticas se correlacionam com o desenvolvimento do câncer (CARITHERS *et al.*, 2015). A partir de setembro de 2014 a base de dados GTEx criou um atlas abrangente com 4502 amostras de 215 doadores representando 55 tecidos sadios (ROSENBLOOM *et al.*, 2015).

1.1.2. Reprodutibilidade de resultados

Uma vez que diferentes abordagens estatísticas podem ser empregadas nas análises de dados de genômica e expressão gênica, a reprodutibilidade dos resultados pela comunidade científica é ponto fundamental para a confirmação destes e evolução dos métodos estatísticos e computacionais existentes. Assim, foi criado em 2001 o projeto Bioconductor (MORGAN, 2016), no qual são oferecidos uma série de programas de código aberto para a análise dos referidos dados.

A quantificação do perfil da expressão de genes em larga escala avançou consideravelmente ao longo dos anos. As primeiras abordagens levavam em consideração que a abundância de um transcrito particular é medida de forma eficaz pela intensidade de fluorescência. Com o passar dos anos e os avanços na tecnologia de sequenciamento, a abundância da expressão de genes de dados digitais (DGE) em amostras de NGS é observada como uma contagem capaz de identificar diferenças estatisticamente significativas na abundância de RNAs ou características arbitrárias entre diferentes estados biológicos (ROBINSON *et al.*, 2009; HARSHBARGER *et al.*, 2017). A DGE é uma tecnologia capaz de mensurar a expressão de genes por contagem de *tags* de sequência e sensível para medir a expressão em escala genômica, sem a necessidade do conhecimento prévio da sequência, demonstrando ser uma técnica que pode ser utilizada para contabilizar eficientemente os transcritos (ROBINSON & SMYTH, 2007).

Diversos testes estatísticos para avaliar a expressão diferencial foram propostos para análise de DGE. Inicialmente, utilizava-se o teste de *Poisson* para tais análises. Porém, este teste não é utilizável quando há variações biológicas e apresenta resultados falso-positivos. Assim, tal abordagem vem sendo descrito como método não eficaz para avaliar a variação na análise de expressão diferencial de DGE (ROBINSON & SMYTH, 2007; LANGMEAD *et al.*, 2010). Atualmente, uma estratégia para se obter resultados estatísticos relevantes no estudo de DGE é a utilização do teste binomial negativo. Esta metodologia pode ser usada para determinar se as diferenças observadas nas contagens de *tags* podem ser atribuídas ao acaso ou não (ROBINSON & SMYTH, 2007). O programa EdgeR foi escrito na linguagem R, faz parte do projeto Bioconductor e analisa de forma empírica a DGE utilizando o teste binomial negativo (ROBINSON *et al.*, 2009). O EdgeR foi concebido para encontrar as alterações entre dois ou mais grupos, quando pelo menos um deles possuir réplicas e detectar um pequeno número de genes diferencialmente expressos com baixa taxa de falsas descobertas (ROBINSON *et al.*, 2009; ZHOU *et al.*, 2014).

1.1.2.1. Reuso de dados públicos

O número de estudos baseados na reanálise de conjuntos de dados públicos vem crescendo ao longo dos anos, agregando valor à base de conhecimentos científicos a custos mais baixos. O reuso destes dados disponibilizados publicamente pode proporcionar a detecção de sinais que não puderam ser encontrados anteriormente, abordar questões diferentes daquelas apresentadas nos estudos originais e fornecer amplo material de teste para avaliação do desempenho de novos programas ou métodos estatísticos (RUNG & BRAZMA, 2012).

Diversos autores já utilizaram a abordagem de reuso de dados públicos para estudar com diferentes perspectivas o câncer de tireoide. Huang e colaboradores (2014) a partir de conjuntos de dados disponibilizados publicamente no TCGA analisaram a expressão de miRNA e mRNA de câncer papilífero da tireoide e revelaram os microRNAs que regulam especificamente a resposta imune nesse tipo de câncer, além de destacarem os microRNAs miR-221-3p, miR-222-3p, e miR-146b-5p no câncer de tireoide. Tokowy e colaboradores (2016) analisaram miRNA e o perfil de expressão de mRNA de 466 amostras de câncer de tireoide disponibilizados no TCGA. Martinez e colaboradores (2015) utilizando a amostras de 12 tipos de câncer identificaram três piRNAs com padrão de expressão diferente entre câncer da tireoide e os demais tipos tumorais. Visto isso, fica evidenciado o potencial de reutilizar dados de amostras de câncer disponíveis publicamente com propósito de desenvolver e/ou aprimorar os métodos de diagnóstico, tratamento e a prevenção do câncer e buscar RNAs não codificadores pequenos.

1.2. Câncer

Os tecidos sadios são compostos por células características de um determinado órgão. Porém, quando há desregulação de uma célula sadia que afete a diferenciação celular, crescimento, replicação e/ou morte celular, elas passam a ter a capacidade de células cancerosas, acumulando inúmeras alterações e anomalias genéticas (WILLIS, 2016). Neste contexto, a gama de mutações identificadas em tumores idênticos histopatologicamente pode ser diferente (CAIRNS *et al.*, 2011).

1.2.1. O que é o câncer

O câncer é uma doença definida como o crescimento anormal de uma população clonal de células com a capacidade de invadir e destruir tecidos (MAKOHON-MOORE & IACOBUZIO-DONAHUE, 2016). É um grande problema de saúde pública em diferentes partes do mundo (CHENG *et al.*, 2011) e tem sido considerado como a principal causa de mortalidade por doenças não transmissíveis em todo o mundo (GOLD-SMITH *et al.*, 2016).

1.2.2. Câncer no mundo, fatores e ocorrência

Segundo as estatísticas da Agência Mundial da Saúde, em 2012 existiram cerca de 14,1 milhões de casos de câncer no mundo com 8,2 milhões de mortes e estima-se que até 2025 terão 20 milhões de novos casos. Sendo assim, reconhecido como um dos problemas mais cruciais de saúde em todo o mundo devido ao grande aumento da sua incidência e mortalidades significativas (FERLAY *et al.*, 2015; CHEN *et al.*, 2016).

O aumento do número de casos de câncer pode estar relacionado a fontes de instabilidade genômica, tais como fatores exógenos e endógenos (MOYANO & STEFANI, 2015). Assim sendo, a ocorrência do câncer é de etiologia multifatorial podendo estar relacionada com o aumento da expectativa de vida da população mundial, adoção de hábitos pouco saudáveis, tais como tabagismo, sedentarismo, má alimentação, infecções sexualmente transmissíveis, dentre outros fatores (MIRANDA *et al.*, 2016; NIKIFOROV *et al.*, 2016). Ainda, a maioria dos casos podem ser resultantes de interações complexas entre fatores genéticos, epigenéticos e ambientais (ASHTON-PROLLA & SEUANEZ, 2016).

1.2.3. Fatores que levam ao câncer

O desenvolvimento do câncer é um processo de passos múltiplos. No qual o primeiro passo na carcinogênese é a ocorrência de mutações em genes críticos, que predisõem a célula afetada e suas descendentes à transformação neoplásica subsequente. O segundo passo é descrito como uma progressão do processo neoplásico, causando o início da instabilidade genética, taxa de crescimento mais elevada, mudanças nos processos bioquímicos e metabólicos e alterações morfológicas. Mutações adicionais resultam no surgimento de subpopulações

malignas heterogênicas com maior capacidade de sobrevivência (CERUTTI *et al.*, 2003; WILLIS, 2016). Tal processo pode atingir células de diferentes órgãos tais como pulmão, pâncreas, esôfago, intestino, tireoide, entre outros.

1.3. Câncer de tireoide

O câncer de tireoide é a neoplasia maligna endócrina mais comum, provoca o maior número de mortes entre qualquer outro câncer do sistema endócrino e é o segundo câncer mais frequente diagnosticado durante a gravidez (KHALED *et al.*, 2016). Fatores que podem influenciar especificamente o número de casos de câncer de tireoide incluem: obesidade, sexo (aproximadamente 77% dos casos ocorrem em mulheres) e uma variedade de agentes ambientais incluindo a radiação ionizante, ingestão de iodo na dieta e exposição a poluentes ambientais (KITAHARA & SOSA, 2016; NIETO & BOELAERT, 2016).

1.3.1. A Tireoide

A tireoide é uma importante glândula endócrina localizada na parte anterior do pescoço e composta por dois lóbulos em forma de asa e um istmo que os conecta (Figura 1.1). É a primeira glândula endócrina desenvolvida durante a vida embrionária, no 22º dia após concepção em seres humanos. A tireoide utiliza o iodo para secretar os hormônios triiodotironina (T3) e tiroxina (T4) que controlam a frequência cardíaca, pressão arterial, temperatura corporal e taxa metabólica basal (NGUYEN *et al.*, 2015; STOUPA *et al.*, 2016).

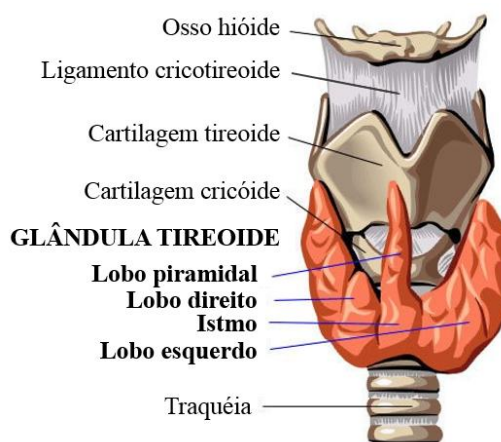


Figura 1.2: Anatomia da tireoide
Fonte: Adaptado de <http://fmcienfederal.com.ar/>

1.3.2. Os diferentes tipos de câncer de tireoide

O câncer de tireoide é a neoplasia maligna endócrina mais generalizada. Existem quatro principais variedades de câncer da tireoide: papilífero (PTC), folicular (FTC), anaplásico (ATC) e câncer medular da tireoide (MTC). PTC, FTC e ATC são originários a partir de células foliculares e apenas MTC se origina a partir de células parafoliculares (células C). Mais de 95% dos cânceres de tireoide derivam de células foliculares da tireoide, enquanto que, cerca de 5% são originários das células C (Figura 1.2) (HU *et al.*, 2016; KITAHARA & SOSA, 2016; PERDAS *et al.*, 2016).

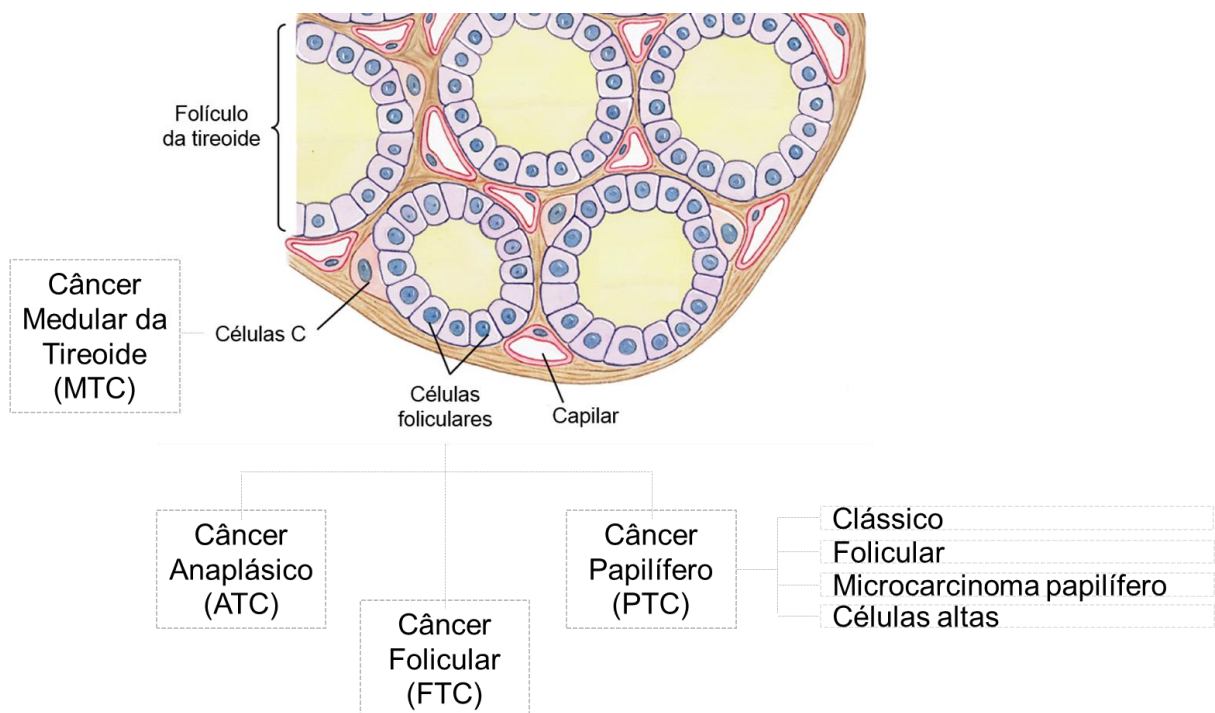


Figura 1.3: Resumo dos principais tipos de câncer de tireoide derivados das células C e células foliculares e subtipos do câncer papilífero.

Fonte: Adaptado de http://pt-br.aia1317.wikia.com/wiki/Arquivo:Tiroide_macro.jpg.

1.3.2.1. Câncer papilífero

Derivado de células foliculares, o câncer papilífero da tireoide (PTC) também conhecido como adenocarcinoma papilífero da tireoide (PAT) (*International Classification of Disease for Oncology*, 3ª edição) é o subtipo de câncer de tireoide mais comum (YIP & SOSA, 2016), ocorre mais frequentemente em mulheres e geralmente ocorrem na faixa etária entre 50 e 60 anos. PTC, tende a invadir os vasos sanguíneos e sofrer metástase mais comumente em linfonodos cervicais e, menos comumente, nos pulmões (CABANILLAS *et al.*, 2016; PARAMESWARAN *et*

al., 2016). Este subtipo de câncer pode progredir para carcinoma mal diferenciado ou perder totalmente a diferenciação e dar origem a ATC (HU *et al.*, 2016).

A exposição à radiação ionizante é um fator de risco para o desenvolvimento do PTC. Tal fato foi observado após o acidente no reator nuclear de Chernobyl, em 1986, quando houve um aumento acentuado na incidência de carcinomas papilíferos, afetando principalmente crianças. Semelhantemente os casos de câncer de tireoide aumentaram após as explosões das bombas atômicas em Hiroshima e Nagasaki em 1945 e em pessoas que recebem radioterapia externa para condições benignas ou malignas da cabeça e pescoço (WILLIAMS, 2015; FAGIN & WELLS, 2016).

PTC pode ser subdividido em: convencional, variante folicular, microcarcinoma papilífero, oncocística, Células Altas, esclerosante difusa, sólido, célula clara, morular cribiforme, macrofolicular, PTC com características proeminentes *hobnail*, PTC com facete como estroma, combinado papilífero e carcinoma medular e PTC de diferenciação ao carcinoma anaplásico (Figura 1.2) (LLOYD *et al.*, 2011).

1.3.2.1.1. Carcinoma papilífero, variante folicular

Carcinoma papilífero, variante folicular, (PCF) é o subtipo mais comum de PTC, constituindo entre 9% e 22,5% de todos os casos (YU *et al.*, 2013) e pode apresentar padrões metastáticos incomuns (NWAENZE *et al.*, 2015). São compostos de folículos de tamanhos variados, dentro destes folículos as células ocasionalmente são grandes e multinucleadas com coloração distintas geralmente com o meio mais escuro e pode ter aparência de chiclete (LLOYD *et al.*, 2011).

1.3.2.1.2. Carcinoma papilífero, variante Células Altas

Carcinoma papilífero, variante Células Altas (PCC) representa 7,5% dos casos de PTC e pode sofrer metástases para traqueia (LIVOLSI 2011). As células desse tipo tumoral possuem altura pelo menos três vezes maior do que a sua largura, têm abundante citoplasma eosinofílico e características nucleares semelhantes à PTC convencional. Os tumores tendem a ser maiores, mais volumosos e agressivos do que os PTC convencionais, de modo que a necrose e atividades mitóticas são mais comuns. A maioria dos pacientes com PCC são idosos (LLOYD *et al.*, 2011). As células altas devem representar 50% ou mais das células de carcinoma papilar para

fazer o diagnóstico de variante de células altas. O prognóstico para esta variante é menos favorável do que para o câncer papilífero convencional, frequentemente não é reconhecida e é patologicamente confundida com outras enfermidades (LIVOLSI 2011).

1.3.2.2. Câncer folicular

Câncer folicular da tireoide abrange 15% dos casos de câncer de tireoide e é um tipo de câncer de alto risco que têm tendência a sofrer metástase para locais distantes, em particular, aos pulmões e ossos (MACIEL *et al.*, 2005; CABANILLAS *et al.*, 2016). É derivado de células foliculares e tende a ser mais comum em países onde as pessoas não recebem iodo suficiente na sua dieta (HU *et al.*, 2016). A sobrevivência em casos FTC está associada com idade avançada no momento do diagnóstico, maior tamanho do tumor, invasão capsular, sexo masculino e presença de metástases (CABANILLAS *et al.*, 2015). FTC apresenta um desafio diagnóstico especial devido às semelhanças morfológicas e moleculares com o Carcinoma papilífero (PTC) (WEBER *et al.*, 2006) e assim como o PTC, FTC também pode progredir para carcinoma mal diferenciado ou podem perder totalmente a diferenciação e originar câncer anaplásico da tireoide (HU *et al.*, 2016).

1.3.2.3. Adenoma

Os adenomas são tumores benignos da tireoide, abrangem 2% dos casos, são sólidos, podem possuir forma redonda ou oval, com uma cápsula fibrosa circundante que é geralmente regular e fina, variam de tamanho entre 1 e 3 cm e podem ser observadas mudanças que incluem degeneração cística, hemorragia, ossificação, calcificação, fibrose e até necrose. Adenomas são cinco vezes mais frequentes do que os carcinomas foliculares e são considerados verdadeiros neoplasmas, uma vez que vários estudos confirmaram sua natureza clonal (MACIEL *et al.*, 2005; BOYANTON *et al.*, 2008; ESZLINGER *et al.*, 2008; NIKIFOROVA *et al.*, 2008; ROSSING *et al.*, 2012).

Adenoma possui semelhanças morfológicas e moleculares com FTC (WEBER *et al.*, 2006), a característica chave que os distingue é a invasão capsular e/ou vesicular, que não pode ser detectado por ultrassom ou citologia. Biomarcadores para distinguir adenomas de FTC antes da cirurgia são necessários (HUANG *et al.*,

2015) e considerados como uma forma de diagnóstico mais preciso (ROSSING *et al.*, 2012).

1.3.2.4. Câncer anaplásico

Câncer anaplásico de tireoide é uma forma rara de câncer de tireoide (menos de 1%) e têm mau prognóstico por causa do crescimento tumoral rápido (FAGIN & WELLS, 2016). Os pacientes muitas vezes desenvolvem rouquidão, disfagia e dispneia e ao serem examinados, a maioria dos pacientes com câncer de tireoide anaplásico têm uma massa palpável grande e firme na tireoide com ou sem adenopatia cervical (CABANILLAS *et al.*, 2016).

ATC pode sofrer metástase para órgãos mais distantes mais comumente no pulmão, seguido de ossos e cérebro, muitas vezes surge e pode coexistir com câncer diferenciado da tireoide. Os médicos devem suspeitar de transformação anaplásicas em pacientes com histórico de câncer diferenciado da tireoide de longa data (CABANILLAS *et al.*, 2016).

1.3.3. Genética do câncer de tireoide

Estudos que sequenciaram o DNA de amostras de câncer de tireoide revelaram que a via de sinalização Proteína Quinase Ativada por Mitógeno (MAPK) pode desenvolver papel importante no câncer de tireoide. Desta forma, a alta frequência de alterações genéticas na via MAPK afetando os principais genes *BRAF*, tais como *RAS* ou *RET*, que pode levar ao câncer de tireoide. Estas alterações estão presentes em cerca de 70% dos carcinomas papilífero da tireoide (OLER *et al.*, 2008; YAMASHITA *et al.*, 2013; CABANILLAS *et al.*, 2015; MANCIKOVA *et al.*, 2015). A via MAPK transmite sinais a partir da membrana plasmática para o núcleo e desempenha papel em vários processos celulares tais como: regulação da proliferação e sobrevivência celular, expressão gênica, mitose e apoptose (CABANILLAS *et al.*, 2016).

1.3.4. Prognóstico e biomarcadores

Apesar do câncer de tireoide geralmente possuir um prognóstico favorável, indivíduos afetados sofrem considerável morbidade, tanto da doença, bem como dos efeitos colaterais da terapia. A biópsia aspirativa com agulha fina e histologia são o

padrão ouro para o diagnóstico de câncer da tireoide, mas até 25% das biópsias realizadas na prática não são conclusivas (WITCZAK *et al.*, 2016). Diferentes estratégias para a obtenção de melhores tratamentos e diagnósticos para o câncer de tireoide vêm sendo desenvolvidas tal como a identificação de marcadores tumorais e uma melhor compreensão dos seus mecanismos moleculares na proliferação e invasão do tumor (LUO *et al.*, 2015).

1.3.4.1. Biomarcadores

Biomarcadores ou marcadores biológicos podem ser moléculas específicas utilizadas como indicadores fisiológicos ou patológicos de ecossistemas ou organismos podendo ser utilizadas em diagnósticos, na medição do progresso de doenças ou na eficácia de terapias específicas (VILELA *et al.*, 2012). O potencial dos RNAs não codificadores pequenos na identificação de diversos estados fisiológicos e patológicos foram demonstrados em diferentes estudos. Por exemplo, Lee e colaboradores (2013) identificaram microRNA como biomarcadores para câncer papilífero de tireoide. Em outro estudo, Mannoer e colaboradores (2014) relataram RNAs nucleolares pequenos como potenciais biomarcadores para tumor de pulmão. Kichukova e colaboradores (2015) estudaram o perfil de expressão de microRNAs como potenciais biomarcadores para o diagnóstico de transtornos neuropsiquiátricos. Krishnan e colaboradores (2016) descreveram RNAs que interage com PIWI como novos marcadores biológicos para prognósticos de câncer de mama.

1.4. RNAs não codificadores

Ao longo dos anos, RNAs não codificadores (ncRNA) vem sendo identificados como potenciais biomarcadores a fim de propiciar novas intervenções terapêuticas de câncer, podendo se tornar uma prática mais tangível em um futuro próximo (ASSUMPÇÃO, 2015). NcRNAs são RNAs que não transcrevem para proteína (KENTWELL *et al.*, 2014) e de acordo com seu tamanho podem ser classificados em: 1) longos RNAs não codificadores (lncRNA) com tamanho superior a 200 nucleotídeos 2) RNAs não codificadores pequenos (sncRNA) com tamanho inferior a 200 nucleotídeos (BHARTIYA & SCARIA, 2016; AIGNER, *et al.*, 2016).

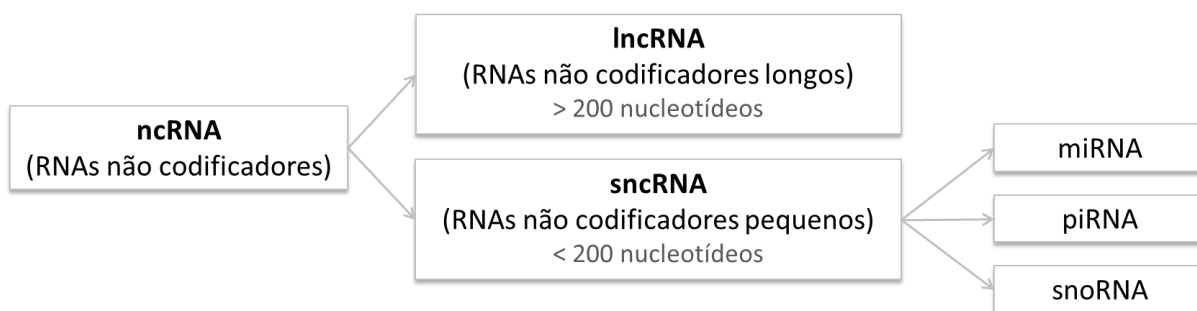


Figura 1.4: Resumo da classificação dos RNAs não codificadores. De acordo com seu tamanho os RNAs não codificadores podem ser divididos em: RNAs não codificadores longos (lncRNAs) com mais de 200 nucleotídeos; RNAs não codificadores pequenos (sncRNAs) com menos de 200 nucleotídeos. Os sncRNAs podem ser subdivididos de acordo com sua biogênese e origens genômicas e destacamos os subtipos miRNA, piRNA e snoRNA que foram utilizados no presente estudo.

1.4.1. RNAs não codificadores pequenos

NcRNAs pequenos são geralmente classificados de acordo com diferentes vias de biogênese e origens genômicas (RAMALHO-CARVALHO *et al.*, 2016). Dentre os sncRNAs, destacamos microRNA (miRNA), RNA que interage com PIWI¹ (piRNA), RNAs nucleolares pequenos (snoRNA) (ESTELLER, 2011) (Tabela 1.1).

Tabela 1.1: Resumo dos RNAs não codificadores pequenos analisados no presente estudo: miRNA, piRNA e snoRNA.

sncRNA	Tamanho (nucleotídeos)	Função
miRNA	18 a 25	Regulação da expressão de genes e processos biológicos (MANCIKOVA <i>et al.</i> , 2015)
piRNA	26 a 31	Regulação gênica, apoptose, metilação do DNA e controle dos elementos de transposição (ESTELLER, 2011; TÓTH <i>et al.</i> , 2016)
snoRNA	60 a 300	Metilação de rRNA e pseudouridilação de nucleotídeos (STEPANOV <i>et al.</i> , 2015)

1.4.1.1. MicroRNA

MicroRNAs (miRNAs) são RNAs não codificadores pequenos descobertos pela primeira vez em 1993 no nematóide *Caenorhabditis elegans* (HAYES *et al.*, 2014). Os miRNAs possuem tamanho entre 18 e 25 nucleotídeos de comprimento (LUDVÍKOVÁ *et al.*, 2015), desempenham papéis importantes na regulação da

¹ As proteínas PIWI pertencem a família das proteínas Argonaute (TÓTH *et al.*, 2016) e quando ligadas aos piRNAs formam o complexo piRNA+PIWI podendo bloquear a transcrição de genes alvo (ASSUMPÇÃO *et al.*, 2015).

expressão de genes e processos biológicos, incluindo a diferenciação celular, desenvolvimento, padronização e metabolismo do ciclo celular (CHENG *et al.*, 2011) e são a maior classe de sncRNAs em células somáticas de animais e plantas (VILELA *et al.*, 2012). Os miRNAs têm importância na biologia do câncer através do controle da expressão de RNAs mensageiros (mRNA) alvos no crescimento do tumor, invasão, angiogênese e evasão imune (HAYES *et al.*, 2014). Tal qual sua desregulação demonstrou ser uma característica do câncer (MANCIKOVA *et al.*, 2015).

A nomenclatura dos miRNAs destinam-se a transmitir informações limitadas sobre as relações funcionais entre os miRNAs maduros e é atribuída utilizando um prefixo com 3 ou 4 caracteres para designar a qual espécie pertence. Por exemplo, para a espécie *Homo sapiens* é utilizado o prefixo “hsa”, seguido de identificadores numéricos sequenciais por exemplo os miRNAs hsa-miR-101 em humanos e mmu-miR-101 em camundongo. Para sequências parálogas cujos miRNA diferem em apenas uma ou duas posições, um caractere é usado como sufixo: hsa-miR-146a e hsa-miR-146b. As sequências maduras são designadas por “miR” e as sequências precursoras são chamadas de “mir”, como segue: hsa-miR-121a expresso a partir do precursor hsa-mir-121a. Os *loci* distintos da estrutura em formato de grampo de cabelo que dão origem a miRNAs maduros idênticos têm sufixos numéricos: dme-mir-281-1 e dme-mir-281-2 em *Drosophila melanogaster*. Quando dois miRNA maduros são originados do mesmo pré-miRNA, um sufixo 5p ou 3p é adicionado para indicar que o miRNA corresponde à fita da região 5' ou 3' do grampo, respectivamente. Plantas e vírus tem nomenclaturas ligeiramente diferentes. Let-7 e lin-4 são exceções para o sistema de nomenclatura, os nomes são mantidos por razões históricas e seus homólogos também recebem tais nomes (AMBROS *et al.*, 2003; GRIFFITHS-JONES *et al.*, 2006).

1.4.1.1.1. Biogênese do microRNA

A biogênese de miRNA (Figura 1.3) inicia-se no núcleo quando a RNA polimerase II sintetiza o gene de miRNA formando o miRNA primário (pri-miRNA) que possui uma longa sequência de nucleotídeos e contém estruturas em forma de grampo de cabelo. Na segunda etapa, o pri-miRNA entra em um complexo que consiste na enzima Drosha e um cofator essencial Pasha que irá transformar o pri-

miRNA em um precursor menor (pré-miRNA) sendo depois exportado por exportina-5/Ran-GTP para o citoplasma. O pré-miRNA é formado por uma sequência menor com 70 nucleotídeos de comprimento em formato de grampo de cabelo (LUDVÍKOVÁ *et al.*, 2015).

Ao nível citoplasmático, pré-miRNAs de fita dupla são clivados pelas enzimas Dicer e helicase e, subsequentemente, o duplex formado é desenovelado em duas cadeias de miRNA individuais. Uma das fitas do pré-miRNA é usado para produzir miRNA maduro, porém a segunda fita parece ser geralmente degradada. O miRNA maduro é incorporado ao complexo RISC (Complexo de Indução do Silenciamento do RNA) que irá mediar a interação entre o miRNA e o mRNA alvo. Dependendo do grau de pareamento de bases entre o miRNA e a região 3' UTR do mRNA o efeito final desta interação poderá ser a clivagem do mRNA ou a repressão da tradução (HAYES *et al.*, 2014; LUDVÍKOVÁ *et al.*, 2015).

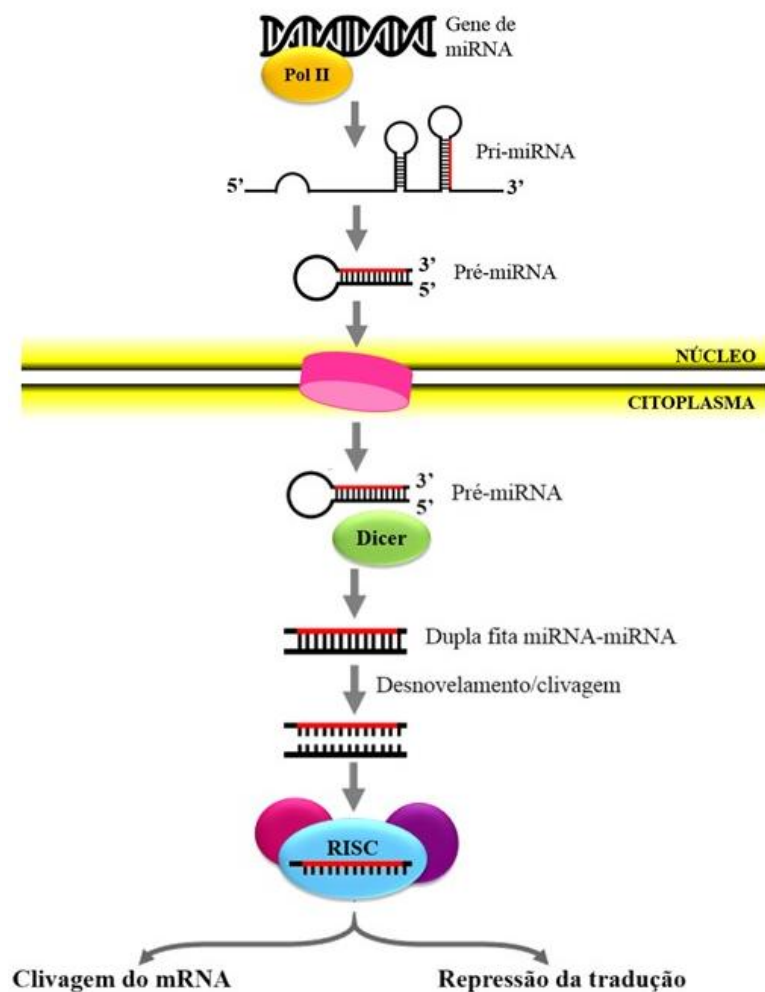


Figura 1.5: Biogênese resumida do miRNA: RNA polimerase II transcreve o gene de miRNA gerando o miRNA primário (pri-miRNA). O pri-miRNA é transformado em precursor menor (pré-miRNA) por um complexo proteico. O pré-miRNA é exportado do núcleo para o citoplasma onde proteínas irão se ligar e clivar o pré-miRNA em cadeia dupla fita, subsequentemente, o duplex formado é desenovelado e clivado em duas fitas de miRNA individuais. Uma das fitas de miRNA se liga ao complexo RISC e depende do grau de emparelhamento de bases entre o mRNA e o miRNA o efeito final desta interação poderá ser a clivagem do mRNA ou a repressão da tradução.

1.4.1.2. RNA que interage com PIWI

RNA que interage com PIWI (piRNA) é um dos sncRNA descoberto mais recentemente (ASSUMPÇÃO, 2015). Os piRNAs apresentam tamanho entre 26 a 31 nucleotídeos (MOYANO & STEFANI, 2015), seu papel em células germinativas de animais foi estabelecido e agora também está sendo classificado como regulador pós-transcricional na expressão gênica em células somáticas (KRISHNAN *et al.*, 2016). PiRNAs atuam em animais na regulação gênica, apoptose, metilação do DNA e controle dos elementos de transposição (TE) (ESTELLER, 2011; TÓTH *et al.*, 2016). Assim, piRNAs e proteínas PIWI desempenham um papel importante na

ocorrência de câncer, prognóstico e tratamento (ASSUMPÇÃO, 2015). Através da caracterização molecular dos piRNAs, foi descrito o enriquecimento de uracila na região 5' (1U) dos RNAs antisense e adenina na 10ª posição (10A) nos RNAs senso (AGUIAR *et al*, 2016). PiRNAs em cadeias opostas podem mostrar uma sobreposição de 10 nucleotídeos entre as extremidades 5' caracterizando o mecanismo *ping-pong* (AGUIAR *et al*, 2016).

PiRNAs são classificados em dois grupos com base na sua biogênese: piRNAs primários e secundários. A biogênese de piRNA primário é pouco conhecida, enquanto piRNAs secundários são provenientes de um mecanismo de amplificação *ping-pong* que permite a produção de muitos piRNAs no citoplasma (NOHNICK, 2015).

1.4.1.2.1. Biogênese do piRNA

A primeira fase da síntese do piRNA se baseia na teoria de transcrição das sequências de nucleotídeos por uma RNA polimerase II, formando os piRNAs precursores (pré-piRNA). Após a exportação para o citoplasma, os transcritos são processados em sequências menores e se ligam a proteína PIWI, para formar um complexo piRNA/PIWI. O complexo migra para o núcleo e emparelha a bases complementares do DNA, atingindo o seu gene alvo e mobilizando maquinarias de silenciamento para bloquear a transcrição do gene alvo. Desta forma, piRNAs atuam como reguladores da transcrição, principalmente em sequências dos elementos transponíveis (ASSUMPÇÃO, 2015).

No segundo mecanismo, conhecido como *ping-pong*, os piRNAs em vez de se associar com proteínas PIWI, se ligam com proteínas Ago3 (Argonaute3) ou Aub (Aubergine), dos quais, Ago3 possui afinidade a fita senso (10A) e Aub possui afinidade a fita antisense 1U (CZECH & HANNON, 2016). Desta forma, o complexo piRNA/AGO3 cliva uma sequência de RNA resultando em uma nova sequência de RNA que irá funcionar como um substrato para a formação de um novo piRNA que é capaz de se ligar a uma proteína Aub. Do mesmo modo, o complexo piRNA/Aub resultante irá clivar uma sequência de RNA complementar, resultando na produção de substratos de RNA adicional que formam novos complexos piRNA/AGO3. Quando o piRNA maduro é incorporado ao complexo RISC (piRNA/RISC), no

citoplasma, este complexo poderá atuar na proteção do genoma contra a invasão de elementos transponíveis dentre outras funções (ASSUMPÇÃO, 2015) (figura 1.4).

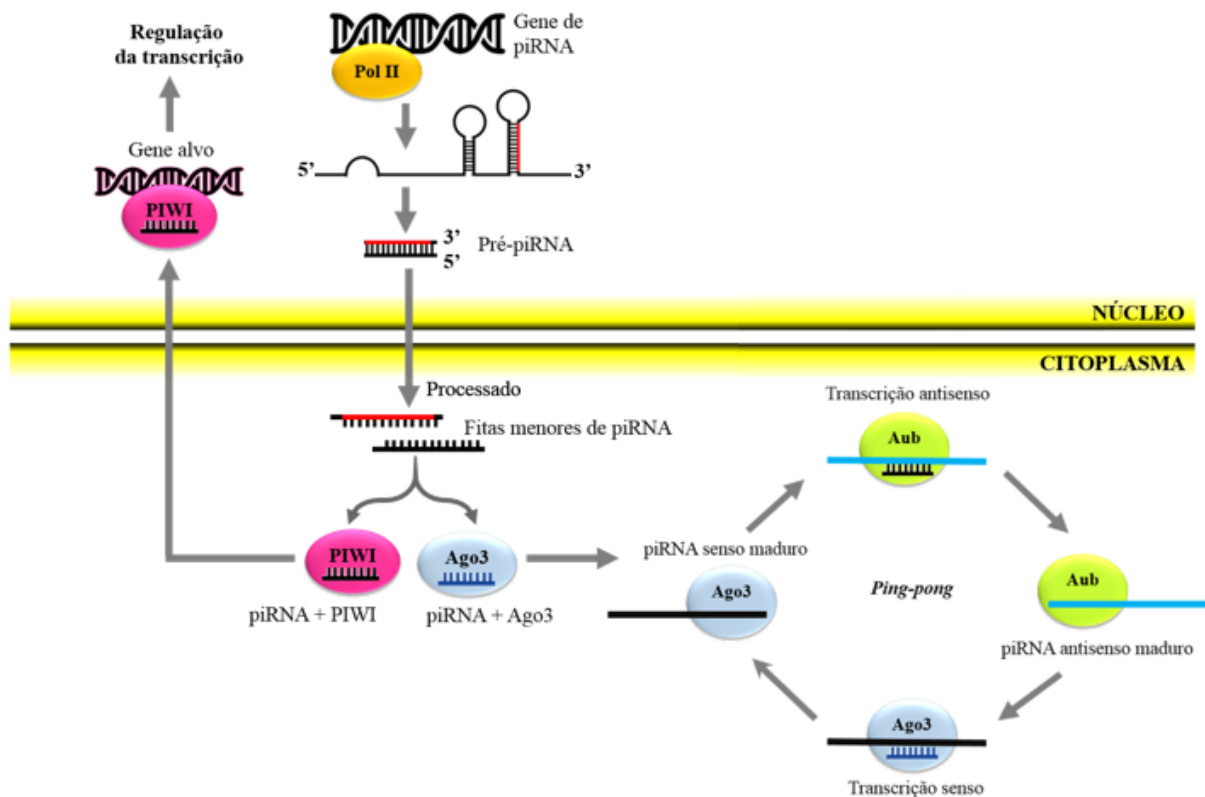


Figura 1.6: Biogênese resumida do piRNA: RNA polimerase II transcreve o gene de piRNA gerando o piRNA primário (pri-piRNA). O pri-piRNA é transformado em precursor menor (pré-piRNA) por um complexo proteico. O pré-piRNA é exportado do núcleo para o citoplasma onde proteínas se ligam e processam o pré-piRNA em duas fitas de piRNA individuais. Uma das fitas de piRNA se liga a proteína PIWI e esse complexo piRNA + PIWI pode ser transportado para o núcleo para a regulação da transcrição por um processo ainda mal descrito ou realizar o processo de amplificação *ping-pong*. No processo *ping-pong* a fita de piRNA antisense se liga a proteína Aub que irá servir de molde para a síntese da fita senso, piRNA senso se liga a proteína Ago3. O piRNA maduro é incorporado ao complexo RISC (piRISC), no citoplasma, para atuar na proteção do genoma contra a invasão de elementos transponíveis dentro outras funções.

1.4.1.3. RNAs nucleolares pequenos

RNAs nucleolares pequenos (snoRNA) possuem tamanho intermediário de 60 nucleotídeos até 300 nucleotídeos (ESTELLER, 2011) e apesar de poder alcançar tamanhos superiores a 200 nucleotídeos, é descrito na literatura como pertencente a classe dos sncRNAs (RAVO *et al.*, 2015). São responsáveis por modificações nos RNAs ribossomais (rRNA), porém há evidências que snoRNAs também estão envolvidos em outros processos, incluindo a regulação de *splicing* alternativo, tradução e estresse oxidativo (MAKAROVA *et al.*, 2013). Podem ser divididos em

duas famílias de acordo com sua estrutura: 1) snoRNA C/D *box*, associado a metilação de rRNA 2) snoRNA H/ACA *box*, responsável pela pseudouridilação de nucleotídeos (STEPANOV *et al.*, 2015; BHARTIYA & SCARIA, 2016). Os snoRNAs estão envolvidos na regulação da modificação pós-transcricional de RNAs ribossomais, podem afetar a condição fisiológica de células, tecidos e órgãos. Desta maneira, a alteração no nível de expressão dos snoRNA pode conduzir à (ao desenvolvimento de) várias doenças. Assim, estudos destinados a avaliar os níveis de expressão dos snoRNAs em células humanas e de mamíferos podem apoiar o desenvolvimento de abordagens para diagnósticos e geração de novos agentes terapêuticos para doenças tal qual o câncer (STEPANOV *et al.*, 2015).

1.4.1.3.1. SnoRNA C/D *box*

SnoRNA C/D *box* (SNORD) de célula eucariota possuem tipicamente o comprimento de 70 a 120 nucleotídeos e contém duas caixas conservadas: caixa de motivo C (sequência consenso: XUGAUGA, onde X equivale a A ou G) e caixa de motivo D (sequência de consenso: CUGA). São estreitamente próximas nas extremidades 5' e 3' terminais da molécula de RNA (Figura 1.5 - A) (KISHORE *et al.*, 2013; DUPUIS-SANDOVAL *et al.*, 2015; STEPANOV *et al.*, 2015). É descrito que SNORD pode reconhecer até dois substratos diferentes (DUPUIS-SANDOVAL *et al.*, 2015).

1.4.1.3.2. SnoRNA H/ACA *box*

SnoRNA H/ACA *box* é mais longo, quando comparado ao SNORD, medindo entre 120 a 140 nucleotídeos de comprimento e apresenta uma estrutura secundária característica que consiste em dois ganchos. Família H/ACA *box* possui em sua estrutura secundária uma região denominada "*hairpin-hinge-hairpin-tail*" que inclui dois domínios em forma de grampo de cabelo ligadas com a região de cadeia simples (dobradiça) e região 3'-terminal (cauda). Caixas de H e ACA estão localizadas em estreita proximidade em um dos grampos nas regiões de dobradiça e cauda, respectivamente (DUPUIS-SANDOVAL *et al.*, 2015; STEPANOV *et al.*, 2015).

A caixa H está na região que liga as estruturas em forma de grampo de cabelo e é formada por uma sequência consenso: ANANNA onde N pode ser qualquer nucleotídeo. A caixa ACA é altamente conservada e se encontra exatamente três

nucleotídeos a montante da extremidade 3', imediatamente após a segunda estrutura em forma de grampo de cabelo. A região de guia de segmentação no rRNA está localizada na protuberância entre as estruturas de grampo de cabelo, especificando pela complementaridade da posição exata a ser pseudouridilação no alvo. (Figura 1.5 - B) *Box H/ACA* snoRNAs pode reconhecer até dois substratos diferentes (DUPUIS-SANDOVAL *et al.*, 2015; STEPANOV *et al.*, 2015).

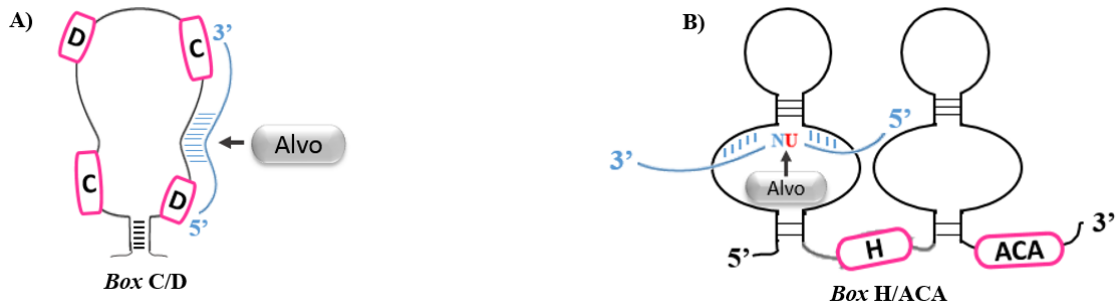


Figura 1.7: Estrutura dos subtipos de snoRNA. A) *Box C/D*: contém duas regiões conservadas: caixa de motivo C (sequência consenso: XUGAUGA, onde X = A ou G) e caixa de motivo D (sequência de consenso: CUGA). B) *Box H/ACA*: Com duas estruturas em forma de grampo de cabelo e caixas consenso: H (com região consenso: ANANNA onde N pode ser qualquer nucleotídeo) e caixa ACA.

1.4.1.3.3. Biogênese do snoRNA

SnoRNAs estão predominantemente localizados nas regiões intrônicas. Com o processo de transcrição, há a formação do pré-mRNA a partir do processo de *splicing*, que remove os íntrons do pré-mRNA e une os éxons formando os mRNAs. Subsequentemente há degradação exonucleotídica, montagem de complexos de ribonucleoproteínas nucleolares (snoRNP) e agrupamento dos snoRNAs maduros com o núcleo proteico. SnoRNPs podem ser exportados do núcleo para o nucléolo onde atuam no processamento do rRNA, ou se mantem no núcleo, onde são envolvidos no processo de *splicing* alternativo e suponha-se que possam existir funções adicionais ainda desconhecidas (ESTELLER, 2011) (Figura 1.6).

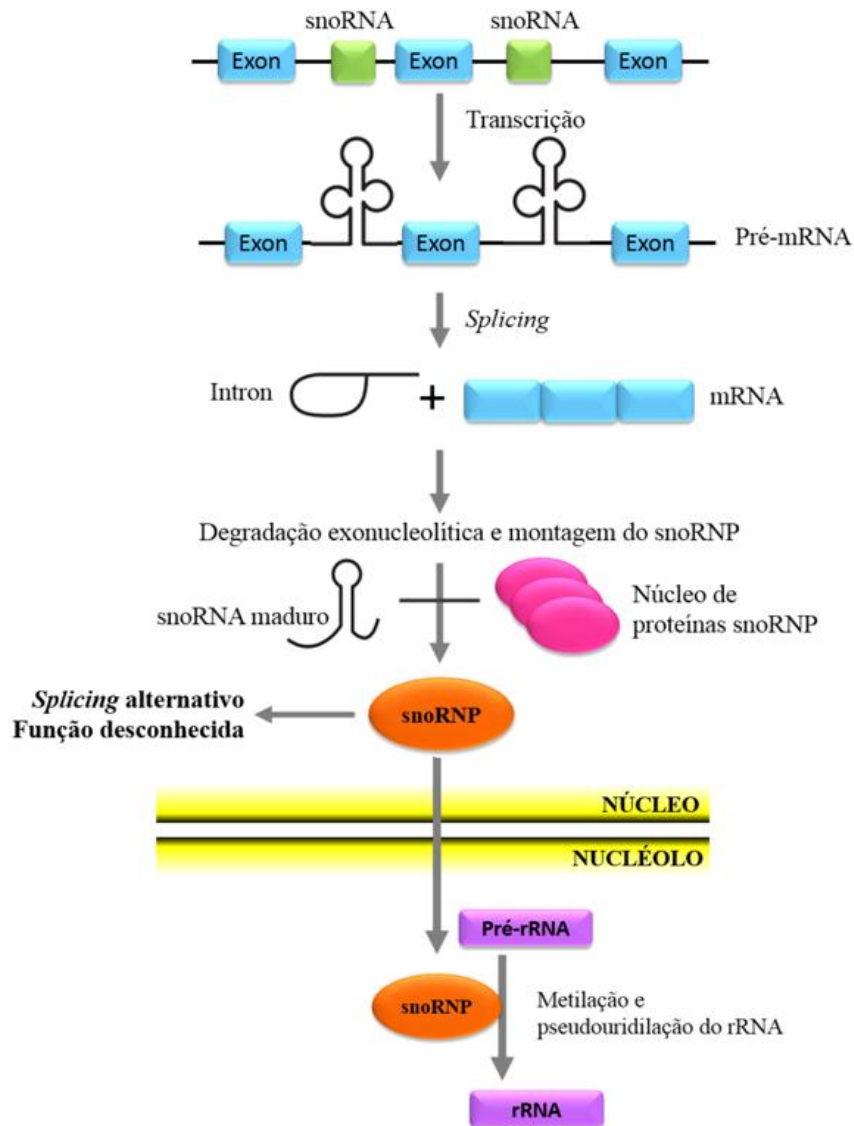


Figura 1.8: Biogênese do snoRNA: SnoRNAs estão predominantemente localizados nas regiões intrônicas. Com o processo de transcrição há a formação do pré-mRNA a partir do processo de *splicing*, desramificação, degradação exonucleotílica, montagem das pequenas ribonucleoproteínas nucleolares (snoRNP) e agrupamento dos snoRNAs maduros com o núcleo proteico os snoRNPs são exportados do núcleo para o nucléolo onde irão atuar no processamento do RNA ribossomal (rRNA), ou se mantem no núcleo, onde são envolvidos no *splicing* alternativo e funções adicionais ainda desconhecidas.

1.5. Justificativa

Com o surgimento de novas tecnologias como o sequenciamento de nova geração, a compreensão dos mecanismos moleculares de câncer aumentou ao longo dos anos (ASHTON-PROLLA & SEUANEZ, 2016). Porém, há uma infinidade de anormalidades genéticas dentro das células cancerosas que ainda são desconhecidas. Deste modo, é pertinente buscar e compreender de forma clara a

natureza das alterações genéticas e moleculares, bem como a sua relevância para o prognóstico do câncer.

Mancikova e colaboradores (2015) com finalidade de complementar o conhecimento existente sobre microRNAs desregulados no desenvolvimento de câncer de tireóide e possíveis novos biomarcadores utilizaram a abordagem de sequenciamento de Sanger em amostras tumorais de pacientes de dois hospitais na Espanha com câncer de tireoide. Porém, em virtude da grande quantidade de dados disponíveis publicamente em bases de dados com foco no sequenciamento de transcritos, Martinez e colaboradores (2015) fizeram uma nova análise de dados públicos disponibilizados na base de dados TCGA identificaram três diferentes piRNAs em câncer de tireoide e com potencial uso como biomarcadores para câncer de tireoide. Outros autores como Ab Mutalib e colaboradores (2016) e Stokowy e colaboradores (2016) conseguiram identificar miRNA por meio do reuso de dados de amostras de câncer de tireoide disponíveis publicamente na base de dados TCGA, evidenciando a relevância e importância da reutilização de dados disponíveis publicamente na busca por RNAs não codificadores pequenos em tecidos tumorais de tireoide.

RNAs não codificadores pequenos estão amplamente descritos em tecidos de eucariotos, assim como em diferentes tipos de câncer em humanos. Entretanto, a expressão de piRNA em tumores malignos e não malignos permanecem em grande parte inexplorados (MARTINEZ *et al.*, 2015). Portanto, a busca por diferentes padrões de expressão de miRNA, piRNA e snoRNA nos diferentes subtipos de câncer de tireoide, evidencia uma alternativa viável na obtenção de potenciais biomarcadores de prognóstico.

2. OBJETIVOS

2.1. Objetivo geral

Avaliar a expressão de miRNA, piRNA e snoRNA em amostras de diferentes subtipos de câncer de tireoide.

2.2. Objetivos específicos

2.2.1. Obter os dados de diferentes subtipos de câncer de tireoide disponíveis na base de dados TGCA.

2.2.2. Identificar os miRNAs, piRNAs e snoRNAs encontrados.

2.2.3. Analisar a expressão de miRNA, piRNA e snoRNA em amostras de câncer de tireoide

3. MATERIAL E MÉTODOS

3.1. Obtenção dos dados

3.1.1. Dados de sequenciamento de alto desempenho das amostras pareadas de tecido normal adjacente ao tumor e do tumor do mesmo paciente

Foram obtidos na base de dados The Cancer Genome Atlas (<https://gdc-portal.nci.nih.gov/>) (TCGA) (versão de 06/06/2016) os arquivos em formato BAM das amostras pareadas dos 59 pacientes reunindo 118 amostras (Figura 3.1, passo 1, Figura 3.2 e Anexo 1) de dados de tecidos normais adjacentes ao tumor e do tumor do mesmo paciente dos diferentes subtipos de câncer de tireoide (Tabela 2.1). Isto é importante para reduzir a heterogeneidade do perfil de expressão que ocorre em populações.

Tabela 3.1: Subtipos de câncer de tireoide, número de pacientes com dados pareados de amostras normais adjacentes aos tumores e tumorais do mesmo paciente utilizadas no trabalho

Subtipo de câncer de tireoide	Número de pacientes
Carcinoma papilífero, variante Células Altas	3
Carcinoma papilífero, variante folicular	7
Carcinoma papilífero, NOS	49

Abreviação: NOS do inglês: *Not Otherwise Specified*

Tabela 3.2: Metadados dos dados obtidos na base de dados The Cancer Genome Atlas (<https://gdc-portal.nci.nih.gov/>) (TCGA) (versão de 06/06/2016)

		Carcinoma papilífero, variante Células Altas	Carcinoma papilífero variante folicular	Carcinoma papilífero	Normal
Sexo	Feminino	1	7	35	43
	Masculino	2	0	14	16
Etnia	Não hispânico ou latino	3	2	37	42
	Hispânico ou latino	0	0	4	4
	Não reportado	0	5	8	13
Raça	Preto ou américo-africano	0	2	3	5
	Branco	2	2	39	43
	Asiático	1	0	2	3
Idade	Não reportado	0	3	5	8
	Média	65	57	55	59
	Mínima-Máxima	39-78	33-81	24-94	24-94

3.1.2. Obtenção dos dados de sequenciamento e anotação

Os arquivos de anotação dos sncRNAs foram obtidos a partir das bases de dados: miRBase (versão 21: junho 2014) (KOZOMARA & GRIFFITHS-JONES, 2014) para obtenção dos dados de miRNA, piRNAbank (versão 1) (SAI LAKSHMI & AGRAWAL, 2008) para obtenção dos piRNAs e UCSC (hg18, *download*: 09/2016) (ROSENBLOOM *et al.*, 2015) para os dados de snoRNAs, ambos no formato GTF anotados na versão hg18 do genoma humano usando a ferramenta liftOver disponibilizada gratuitamente no endereço <https://genome-store.ucsc.edu/> (Figura 2.1, passo 1).

3.2. Pré-processamento

3.2.1. Conversão e alinhamento contra o genoma humano

Afim de proporcionar maior confiabilidade dos dados os arquivos no formato BAM obtidos na base de dados TCGA, foram convertidos para o formato fastq com programa Bam2Bed do pacote BedTools (versão 2.21.0-26-g31878f3) (QUINLAN & HALL, 2010). As sequências convertidas foram alinhadas com genoma humano de referência (versão GRCh38/hg38) utilizando o programa Novoalign (versão 3.02.13) (JAYA & MALAYSIA, 2017) (`novoalign -r None -m -d`

hg18.no_random.no_hap.novoalign.ndx -o SAM -c 8 -f \$fq > \$a_out 2> \$nlog") (Figura 3.1, passo 2). Para propiciar maior especificidade e acurácia aos nossos dados foi utilizado o parâmetro para retornar apenas os *reads* alinhados em um único local no genoma de referência (parâmetro -r None).

3.3. Mapeamento, identificação e contagem dos sncRNAs

O mapeamento através das coordenadas genômicas, identificação e contagem dos transcritos sequenciados foram realizadas utilizando o *software* BedTools (bedtools multicov -s -q 20 -bams arquivo_alinhado.novoalign_hg18.sort.bam.bam -bed arquivo_bed_arquivos_de_anotação.txt > arquivo_de_saida.cont), que requer o arquivo de anotação GTF dos sncRNA: miRNA, piRNAs e snoRNAs obtidos nas bases de dados miRBase, piRNABank e UCSC respectivamente (Figura 3.1, passo 3).

3.4. Normalização e identificação da expressão diferencial

A normalização utilizando método TMM e análise da expressão diferencial dos genes utilizando o teste binomial negativo foram realizadas com *script* em linguagem R utilizando o pacote EdgeR (versão 3.3) (ROBINSON *et al.*, 2009) do projeto Bioconductor (Figura 3.1, passo 4). Foram considerados como diferencialmente expressos os sncRNAs com taxa de falsa descoberta (FDR) menor do que 0,05 e logaritmo de *Fold Change* (logFC) maior ou igual a 1,5 ou menor ou igual a -1,5 (anexo 3).

3.5. Análise da dispersão

Foi utilizado *script* em R para análise da dispersão dos sncRNAs, utilizando adaptação dos padrões descritos por Eisenberg e Levanon (2013): 1) Expressão observada em todos os tecidos; 2) Baixa variação sobre tecidos: desvio padrão $|\log_2(\text{RPKM})| < 1$; 3) Nenhum valor de $|\logFC| \geq 1,5$ (Figura 3.1, passo 5).

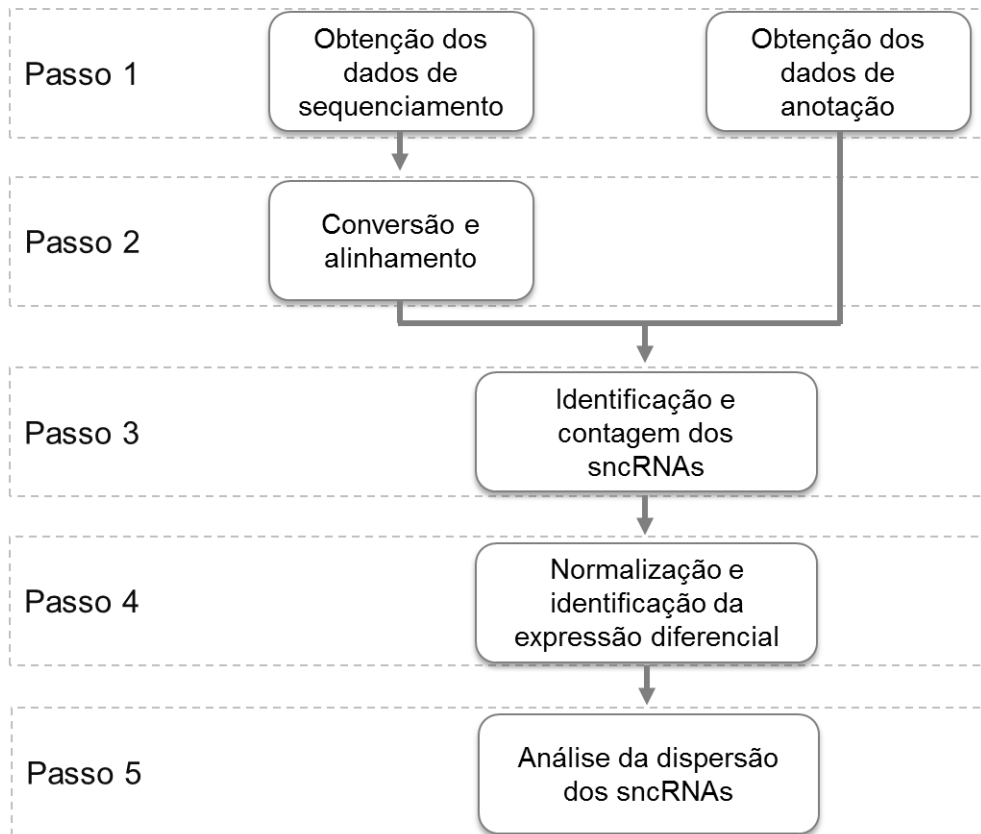


Figura 3.1: Fluxograma ilustrando os passos das análises realizadas nesse estudo.

4. RESULTADOS

4.1. Pré-processamento e alinhamento

Foram alinhados um total 1.478.569.907 *reads* sendo 684.502.827 mapeados unicamente na sequência do genoma humano (versão hg18), sendo obtida a menor e a maior quantidade de alinhamentos para o subtipo Carcinoma papilífero (PTC) com 1.345.897 e *reads* 19.568.904 *reads* respectivamente (Tabela 4.1-4.3).

No somatório total das amostras de cada subtipo tumoral de câncer de tireoide, foi observado a maior média de alinhamentos no subtipo tumoral Carcinoma papilífero, variante folicular (PCF), com 78.537.933 *reads* alinhados, seguido do subtipo Carcinoma papilífero com 5.824.819 *reads* alinhados e o subtipo Carcinoma papilífero, variante Células Altas (PCC), foram alinhados 3.738.457 *reads* contra o genoma humano (versão hg18) (Tabela 4.1-4.3).

No subtipo tumoral Carcinoma papilífero, variante Células Altas, foram alinhados contra o genoma humano dados de sequenciamento pareados de amostras normais adjacentes aos tumores e tumorais de três pacientes. A maior quantidade de *reads* sequenciados foi encontrada nas amostras normais adjacentes aos tumores, porém a maior porcentagem de mapeamentos únicos foi encontrada nas amostras tumorais. A menor quantidade de *reads* únicas foi de 2.730.095, o maior valor foi 7.448.845 e a média total de *reads* mapeados unicamente em todas as amostras de PCC foi de 4.356.855 (Tabela 4.1).

Tabela 4.1: Análise quantitativa dos *reads* sequenciados, dos *reads* mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero, variante Células Altas

Subtipo tumoral	Paciente	N/T	<i>Reads</i> sequenciados	<i>Reads</i> Mapeamentos unicamente	Porcentagem de mapeamentos únicos
Carcinoma papilífero, variante Células Altas	P1	T	6.228.484	2.730.095	43,8%
		N	9.955.682	3.556.622	35,7%
	P2	T	5.742.102	2.744.973	47,8%
		N	7.070.399	2.955.186	41,8%
	P3	T	11.721.483	6.705.410	57,2%
		N	13.252.939	7.448.845	56,2%

Abreviações: P: paciente. N: amostra normal adjacente ao tumor. T: amostra tumoral

Foram alinhados contra o genoma humano dados de sequenciamento pareados de amostras normais adjacentes aos tumores e tumorais de sete pacientes

para o subtipo tumoral Carcinoma papilífero, variante folicular. O menor valor de *reads* únicos foi de 2.607.968, o maior valor foi 12.465.766 e a média total de *reads* mapeados unicamente em todas as amostras de PCF foi de 5.609.852 (Tabela 4.2).

Tabela 4.2: Análise quantitativa dos *reads* sequenciados, dos *reads* mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero, variante folicular

Subtipo tumoral	Paciente	N/T	<i>Reads</i> sequenciados	<i>Reads</i> Mapeamentos unicamente	Porcentagem de mapeamentos únicos
Carcinoma papilífero, variante folicular	P1	T	10.812.846	5.311.259	49,1%
		N	13.664.916	7.011.116	51,3%
	P2	T	7.044.837	3.892.179	55,2%
		N	9.116.941	3.875.799	42,5%
	P3	T	9.969.207	5.434.585	54,5%
		N	12.446.961	6.388.154	51,3%
	P4	T	5.372.333	2.607.968	48,5%
		N	11.729.059	5.148.825	43,9%
	P5	T	10.573.295	4.863.709	46,0%
		N	9.487.283	4.958.491	52,3%
	P6	T	5.936.987	2.915.590	49,1%
		N	12.465.879	5.556.752	44,6%
	P7	T	14.525.447	8.107.740	55,8%
		N	22.379.947	12.465.766	55,7%

Abreviações: P: paciente. N: amostra normal adjacente ao tumor. T: amostra tumoral

Para o subtipo Carcinoma papilífero, foram alinhados contra o genoma humano dados de sequenciamento pareados de amostras normais adjacentes aos tumores e tumorais de 49 pacientes. Um total de 37 das 49 amostras normais adjacentes aos tumores deste subtipo apresentou maior quantidade de *reads* sequenciados e a maior porcentagem de *reads* mapeamentos unicamente em amostras tumorais. O menor valor de *reads* mapeados unicamente foi de 1.345.897, o maior valor foi 19.568.904 e a média total de *reads* mapeados unicamente de todas as amostras de PTC foi de 5.916.569 (Tabela 4.3).

Tabela 4.3: Análise quantitativa dos *reads* sequenciados, dos *reads* mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero

Subtipo tumoral	Paciente	N/T	<i>Reads</i> sequenciados	<i>Reads</i> Mapeamentos unicamente	Porcentagem de mapeamentos únicos
Carcinoma papilífero	P1	T	7.685.385	3.389.325	44,1%
		N	13.872.074	6.472.420	46,7%
	P2	T	13.794.854	7.608.232	55,2%
		N	15.490.152	8.679.443	56,0%
	P3	T	10.828.396	4.953.799	45,7%
		N	12.827.678	5.539.242	43,2%
	P4	T	18.453.353	8.025.402	43,5%
		N	13.990.146	5.393.198	38,5%
	P5	T	10.271.684	6.056.213	59,0%
		N	16.389.020	8.271.274	50,5%
	P6	T	15.866.970	8.224.925	51,8%
		N	13.806.131	6.751.283	48,9%
	P7	T	14.376.694	7.370.767	51,3%
		N	9.307.057	3.772.930	40,5%
	P8	T	8.700.934	4.119.228	47,3%
		N	18.372.629	9.427.488	51,3%
	P9	T	10.168.663	4.666.994	45,9%
		N	10.326.119	5.093.810	49,3%
	P10	T	10.173.899	4.114.192	40,4%
		N	17.346.856	6.344.384	36,6%
	P11	T	12.464.115	7.224.628	58,0%
		N	15.345.677	7.917.716	51,6%
	P12	T	11.698.811	5.069.186	43,3%
		N	16.183.667	7.225.000	44,6%
	P13	T	9.775.634	5.416.817	55,4%
		N	13.500.204	6.282.971	46,5%
	P14	T	6.710.905	3.531.896	52,6%
N		17.800.736	9.175.754	51,5%	
P15	T	3.917.956	2.148.697	54,8%	
	N	14.504.103	7.315.471	50,4%	
P16	T	12.401.037	5.813.256	46,9%	
	N	11.331.844	5.312.820	46,9%	
P17	T	17.715.008	6.926.101	39,1%	
	N	15.136.105	5.580.369	36,9%	
P18	T	9.435.198	4.015.887	42,6%	
	N	11.596.492	5.357.691	46,2%	
P19	T	9.026.082	4.501.649	49,9%	
	N	13.341.658	6.104.167	45,8%	
P20	T	37.917.362	19.568.904	51,6%	
	N	16.714.237	5.874.105	35,1%	
P21	T	11.093.060	4.828.862	43,5%	
	N	19.331.979	9.975.783	51,6%	
P22	T	8.854.982	4.253.419	48,0%	
	N	14.709.467	6.504.986	44,2%	
P23	T	14.021.552	5.097.783	36,4%	
	N	15.516.265	6.034.078	38,9%	
P24	T	11.130.475	4.253.684	38,2%	
	N	13.836.057	5.277.267	38,1%	
P25	T	7.965.654	4.452.615	55,9%	
	N	13.569.449	6.755.804	49,8%	
P26	T	11.955.272	5.465.385	45,7%	
	N	15.917.814	5.745.986	36,1%	
P27	T	12.440.597	5.247.096	42,2%	
	N	4.499.842	1.345.897	29,9%	

Continuação tabela 4.3: Análise quantitativa dos *reads* sequenciados, dos *reads* mapeamentos unicamente e porcentagem de mapeamentos únicos em cada amostra tumoral do subtipo de câncer de tireoide Carcinoma papilífero

Subtipo tumoral	Paciente	N/T	Reads sequenciados	Reads Mapeamentos unicamente	Porcentagem de mapeamentos únicos
Carcinoma papilífero	P28	T	8.965.323	4.970.150	55,4%
		N	6.909.314	3.272.493	47,4%
	P29	T	13.376.566	5.443.662	40,7%
		N	9.554.869	3.684.721	38,6%
	P30	T	10.670.834	6.076.091	56,9%
		N	16.712.013	7.779.622	46,6%
	P31	T	11.805.921	4.059.141	34,4%
		N	15.261.752	7.604.218	49,8%
	P32	T	7.365.020	4.054.816	55,1%
		N	12.623.873	5.407.055	42,8%
	P33	T	6.949.458	4.027.190	57,9%
		N	12.165.018	6.673.264	54,9%
	P34	T	12.287.210	4.192.316	34,1%
		N	15.509.007	5.987.721	38,6%
	P35	T	13.372.732	4.637.616	34,7%
		N	15.632.269	5.454.627	34,9%
	P36	T	17.117.387	5.769.280	33,7%
		N	15.660.773	5.161.596	33,0%
	P37	T	7.930.792	3.534.827	44,6%
		N	15.095.919	6.286.593	41,6%
	P38	T	14.484.541	7.398.216	51,1%
		N	8.082.437	4.726.199	58,5%
	P39	T	6.810.229	4.246.033	62,3%
		N	12.845.240	6.817.413	53,1%
	P40	T	13.314.603	7.049.725	52,9%
		N	14.703.265	6.043.411	41,1%
	P41	T	10.358.140	5.346.101	51,6%
		N	14.271.545	6.208.482	43,5%
	P42	T	11.886.472	7.140.568	60,1%
		N	14.157.863	5.893.344	41,6%
	P43	T	13.575.935	7.400.133	54,5%
		N	19.752.785	10.215.655	51,7%
	P44	T	9.612.773	5.021.163	52,2%
		N	14.041.287	6.598.857	47,0%
	P45	T	13.118.411	4.323.500	33,0%
		N	14.776.773	4.725.891	32,0%
P46	T	9.005.018	3.709.198	41,2%	
	N	11.812.463	4.658.283	39,4%	
P47	T	11.720.763	5.388.749	46,0%	
	N	14.443.222	6.482.955	44,9%	
P48	T	13.170.680	6.106.808	46,4%	
	N	17.956.667	8.600.739	47,9%	
P49	T	10.978.279	4.788.239	43,6%	
	N	15.819.449	6.980.823	44,1%	

Abreviações: P: paciente. N: amostra normal adjacente ao tumor. T: amostra tumoral

4.2. Identificação e contagem dos sncRNAs

A quantidade total obtida de *reads* alinhados nas regiões de miRNA, piRNA e snoRNA foi de 602.690.797 para todos os três subtipos tumorais Carcinoma

papilífero variante Células Altas, Carcinoma papilífero variante folicular e Carcinoma papilífero. Através do qual foi verificada a menor (787.470) e a maior (17.257.080) quantidade de *reads* alinhados a região de miRNA nas amostras do subtipo Carcinoma papilífero. Os *reads* alinhados nas regiões de piRNA em menor quantidade foram na amostra de Carcinoma papilífero com 11.323 *reads* e a maior quantidade de alinhamentos foi na amostra de Carcinoma papilífero variante folicular com 135.528 *reads*. Os snoRNAs foram alinhados em menor quantidade com 68.275 *reads* na amostra do Carcinoma papilífero variante Células Altas e em maior quantidade na amostra de Carcinoma papilífero com 674.355 *reads* alinhados (Tabela 4.4).

Tabela 4.4: Quantidade de *reads* alinhados nas regiões dos RNAs não codificadores pequenos: miRNA, piRNA e snoRNA por subtipo tumoral PCC (Carcinoma papilífero variante Células Altas), PCF (Carcinoma papilífero variante folicular) e PTC (Carcinoma papilífero) através do qual foram verificados a menor (Min.), maior (Max.) e a média total dos sncRNAs em cada amostra tumoral (Med.).

Subtipos	miRNA			piRNA			snoRNA		
	Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.
PCC	2.332.485	6.503.852	3.759.420,6	13.516	99.555	50.001,6	68.275	302.441	167.129,8
PCF	2.253.305	11.047.651	4.772.026,5	11.969	135.528	47.877,5	79.890	456.108	244.995,9
PTC	787.470	17.257.080	4.922.406,2	11.323	117.934	49.070,3	76.509	674.355	211.410

4.3. Comparação entre as abordagens

Uma forma de avaliar a precisão do nosso método de reuso de dados disponíveis na base de dados TCGA, foi comparar os nossos resultados de miRNA diferencialmente expressos encontrados com aqueles publicados por Mancikova e colaboradores (2015) que utilizaram amostras de carcinoma papilífero de tireoide obtidas de pacientes de dois hospitais da Espanha. No presente estudo, identificamos 40 miRNAs diferencialmente expressos, anotados na base de dados miRBase (versão 21) e para melhor comparação entre as abordagens foi utilizada a nomenclatura preexistente de tais miRNAs, totalizando 35 miRNAs diferencialmente expressos. Assim, quando comparados esses resultados, dos miRNAs diferencialmente expressos identificados pela nossa abordagem, 7 dos 10 miRNAs

descritos por Mancikova e colaboradores (2015) também foram identificados (Figura 4.1, Tabela 4.5). Desta forma, consideramos que a nossa abordagem é capaz de reproduzir resultados publicados previamente e poderia ser empregada para a busca de outras classes de RNAs não codificadores pequenos.

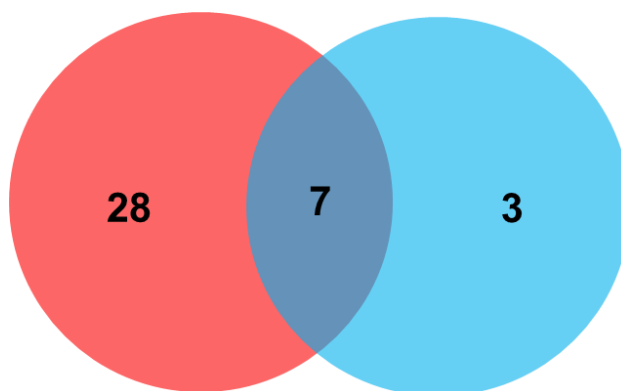


Figura 4.1: Diagrama de Venn comparando a quantidade de miRNA diferencialmente expressos encontrados no presente estudo (vermelho) e a abordagem desenvolvida por Mancikova e colaboradores (2015) (azul).

Tabela 4.5: Comparação dos miRNAs diferencialmente expressos encontrados no presente trabalho e pela abordagem de Mancikova e colaboradores (2015)

Fonte	Quantidade	miRNA
Presente estudo e Mancikova e colaboradores (2015)	7	hsa-miR-222, hsa-miR-21, hsa-miR-221, hsa-miR-204, hsa-miR-31, hsa-miR-146, hsa-miR-451
Mancikova e colaboradores (2015)	3	hsa-miR-486, hsa-miR-1179, hsa-miR-7
Presente estudo	28	hsa-miR-503, hsa-miR-134, hsa-miR-200, hsa-miR-508, hsa-miR-1978, hsa-miR-183, hsa-miR-127, hsa-miR-210, hsa-miR-1251, hsa-miR-92, hsa-miR-363, hsa-miR-150, hsa-miR-144, hsa-miR-375, hsa-miR-96, hsa-miR-100, hsa-miR-181a-2, hsa-miR-379, hsa-miR-205, hsa-miR-708, hsa-miR-153, hsa-miR-182, hsa-miR-34, hsa-miR-138, hsa-miR-1247, hsa-miR-20, hsa-miR-891, hsa-miR-199

4.4. Identificação da expressão diferencial

Foram considerados como diferencialmente expressos os sncRNAs que apresentaram taxa de falsa descoberta (FDR) menor que 0,05 e logaritmo de *Fold Change* (logFC) maior ou igual a 1,5 ou menor ou igual a -1,5 (Tabela 4.6) (Anexo 2, 3 e 4). Desta forma, quando analisados todos os subtipos tumorais, foram descritos como mais expressos as amostras que tiveram maior expressão em amostras tumorais e menos expressos quando a maior expressão é oriunda de amostras

normais adjacentes aos tumores. Com isso, foram obtidos um total de 46 RNAs não codificadores pequenos, dos quais seis são comuns para os três subtipos tumorais, sendo eles: hsa-miR-146b-3p, hsa-miR-222-5p, hsa-miR-146b-5p, hsa-miR-221-3p e hsa-miR-222-3p mais expressos em amostras tumorais e hsa-miR-204-5p mais expresso em amostras normais adjacentes aos tumores (Figura 4.2-11).

Tabela 4.6: Comparação quantitativa dos sncRNAs diferencialmente expressos utilizando os filtros FDR menor que 0,05 e logFC maior ou igual a 1,5 ou menor ou igual a -1,5

Subtipo de câncer de tireoide	Carcinoma papilífero, variante Células Altas	Carcinoma papilífero, variante folicular	Carcinoma papilífero
Sem filtro	223/1/101 (325)	237/2/94(333)	240/2/98(340)
FDR < 0,05	19/0/2(21)	104/1/33(137)	215/2/78(296)
logFC ≥ 1,5	19/0/2(21)	22/0/1(23)	21/0/3(24)

Legenda: miRNA/piRNA/snoRNA (total)

4.4.1. Identificação da expressão diferencial de sncRNA

Avaliamos os sncRNAs detectados exclusivamente em um subtipo. Identificamos nove sncRNAs presentes apenas em Carcinoma papilífero, sendo três mais expressos em amostras tumorais e seis mais expressos em amostras normais adjacentes aos tumores. No subtipo tumoral Carcinoma Células Altas encontramos quatro diferencialmente expressos, dos quais dois foram mais expressos em amostra tumoral e dois foram mais expressos em amostras normais adjacentes aos tumores. Por fim, no subtipo Carcinoma papilífero, variante folicular, foram observados 11 RNAs não codificadores pequenos detectados exclusivamente em um subtipo, sendo quatro mais expressos em amostras tumorais e sete mais expressos em amostras normais adjacentes aos tumores (Figuras 4.2 e 4.3 e Tabelas 4.7 e 4.8).

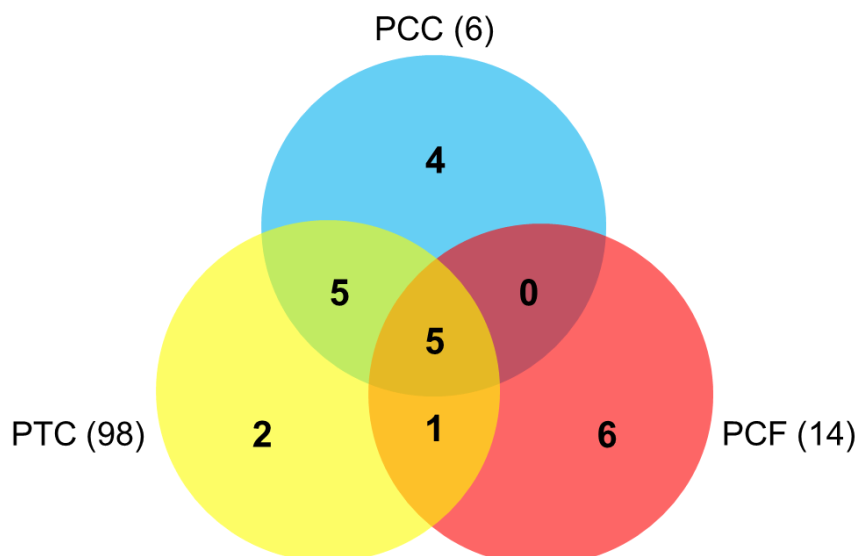


Figura 4.2: Diagrama de Venn comparando a quantidade total de sncRNAs mais expressos em amostras tumorais ($FDR < 0,05$ e $\log FC \geq 1,5$) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo.

Tabela 4.7: SncRNAs mais expressos ($FDR < 0,05$ e $\log FC \geq 1,5$) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.

Subtipos	Quantidade	sncRNAs
PTC, PCF e PCC	5	hsa-miR-146b-3p, hsa-miR-222-5p, hsa-miR-146b-5p, hsa-miR-221-3p, hsa-miR-222-3p
PTC e PCF	1	hsa-miR-34a-5p
PTC e PCC	5	hsa-miR-31-3p, hsa-miR-21-5p, hsa-miR-375, hsa-miR-31-5p, hsa-miR-181a-2-3p
PCC e PCF	0	-
PTC	2	hsa-miR-508-3p, hsa-miR-503-5p
PCF	6	hsa-miR-891a-5p, hsa-miR-96-5p, hsa-miR-1251-5p, HBII-82, hsa-miR-182, hsa-miR-183-5p
PCC	4	hsa-miR-92b-3p, hsa-miR-200b-5p, hsa-mir-1978, ACA31

Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas

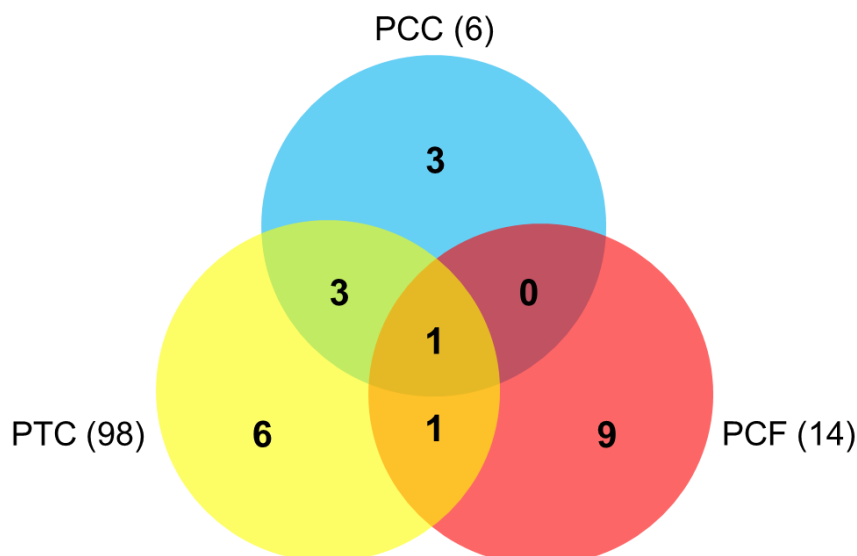


Figura 4.3: Diagrama de Venn comparando a quantidade total de sncRNAs menos expressos ($FDR < 0,05$ e $\log FC \leq -1,5$) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo.

Tabela 4.8: SncRNAs menos expressos ($FDR < 0,05$ e $\log FC \leq -1,5$) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.

Subtipos	Quantidade	sncRNAs
PTC, PCF e PCC	1	hsa-miR-204-5p
PTC e PCF	1	hsa-miR-199b-5p
PTC e PCC	3	hsa-miR-20b-5p, hsa-miR-144-5p, hsa-miR-451a
PCF e PCC	0	-
PTC	6	HBII-382, hsa-miR-1247-3p, mgU2-25/61, hsa-miR-144-3p, hsa-miR-363-3p, HBII-289
PCF	9	hsa-miR-205-5p, hsa-miR-134-5p, hsa-miR-708-5p, hsa-miR-379-5p, hsa-miR-150-5p, hsa-miR-127-3p, hsa-miR-127-5p, hsa-miR-210-3p, hsa-miR-100-5p
PCC	3	hsa-miR-153-5p, ACA3-2, hsa-miR-138-5p

Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas

4.4.2. Identificação da expressão diferencial de miRNA

Comparando os 40 miRNAs diferencialmente expressos, 21 deles apresentaram maior expressão em amostras tumorais e 19 foram identificados como mais expressos em amostras normais adjacentes aos tumores. Foram observados seis miRNAs pertencentes a todos os subtipos de câncer de tireoide (hsa-miR-146b-3p, hsa-miR-222-5p, hsa-miR-146b-5p, hsa-miR-221-3p, hsa-miR-222-3p e hsa-miR-204-5p) sendo cinco deles mais expressos em amostras tumorais. Presente apenas no subtipo tumoral, foram identificados cinco miRNAs em Carcinoma papilífero, dos

quais três foram mais expressos em amostras tumorais e dois mais expressos em amostras normais adjacentes aos tumores. Dos 14 miRNAs presentes apenas em PCF, cinco apresentaram maior expressão em amostras tumorais e sete como mais expressos em amostras normais adjacentes aos tumores. Já a análise do subtipo PCC mostrou cinco miRNAs, sendo um mais expresso em amostras normais adjacentes aos tumores e dois mais expressos em amostras tumorais (Figuras 4.4 e 4.5 e Tabelas 4.9 e 4.10).

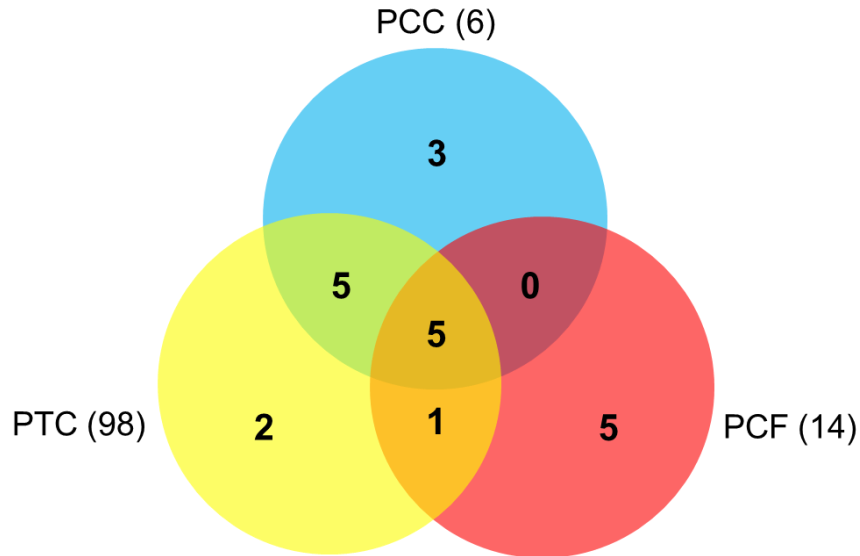


Figura 4.4: Diagrama de Venn comparando a quantidade de miRNA mais expressos (FDR < 0,05 e logFC ≥ 1,5) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo

Tabela 4.9: MiRNAs mais expressos (FDR < 0,05 e logFC ≥ 1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.

Subtipos	Quantidade	miRNAs
PTC, PCF e PCC	5	hsa-miR-146b-3p, hsa-miR-222-5p, hsa-miR-146b-5p, hsa-miR-221-3p, hsa-miR-222-3p
PTC e PCF	1	hsa-miR-34a-5p
PTC e PCC	5	hsa-miR-31-3p, hsa-miR-21-5p, hsa-miR-375, hsa-miR-31-5p, hsa-miR-181a-2-3p
PCC e PCF	0	-
PTC	2	hsa-miR-508-3p, hsa-miR-503-5p
PCF	5	hsa-miR-891a-5p, hsa-miR-96-5p, hsa-miR-1251-5p, hsa-miR-182, hsa-miR-183-5p
PCC	3	hsa-miR-92b-3p, hsa-miR-200b-5p, hsa-mir-1978

Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas

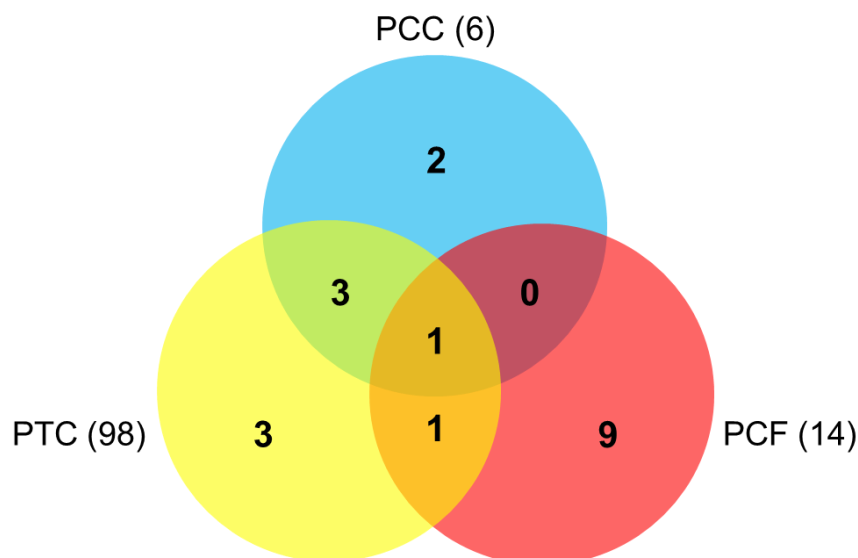


Figura 4.5: Diagrama de Venn comparando a quantidade de miRNA menos expressos (FDR < 0,05 e logFC ≤ -1,5) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo

Tabela 4.10: MiRNAs menos expressos (FDR < 0,05 e logFC ≤ -1,5) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.

Subtipos	Quantidade	miRNAs
PTC, PCF e PCC	1	hsa-miR-204-5p
PTC e PCF	1	hsa-miR-199b-5p
PTC e PCC	3	hsa-miR-20b-5p, hsa-miR-144-5p, hsa-miR-451a
PCF e PCC	0	-
PTC	3	hsa-miR-1247-3p, hsa-miR-144-3p, hsa-miR-363-3p
PCF	9	hsa-miR-205-5p, hsa-miR-134-5p, hsa-miR-708-5p, hsa-miR-379-5p, hsa-miR-150-5p, hsa-miR-127-3p, hsa-miR-127-5p, hsa-miR-210-3p, hsa-miR-100-5p
PCC	2	hsa-miR-153-5p, hsa-miR-138-5p

Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas

4.4.3. Identificação da expressão diferencial de piRNA

Após a análise da expressão diferencial das amostras dos três subtipos tumorais não foram encontrados piRNAs diferencialmente expressos empregando os filtros de significância estatística empregados neste estudo.

4.4.4. Identificação da expressão diferencial de snoRNA

Foram constatados seis snoRNAs com expressão diferencial, sendo dois mais expressos em amostras tumorais (HBII-82 e ACA31) e quatro mais expressos em amostras normais adjacentes aos tumores (mgU2-25/61 HBII-382, HBII-289, ACA3-2). Encontramos três snoRNAs presentes mais expressos em amostras normais adjacentes aos tumores apenas no subtipo Carcinoma papilífero, dois snoRNAs detectados exclusivamente no subtipo tumoral Carcinoma papilífero variante Células Altas, dos quais um foi mais expresso em amostras tumorais e um mais expresso em amostras normais adjacentes aos tumores. Por fim, foi encontrado um snoRNAs presente apenas no subtipo Carcinoma papilífero variante Células Altas, sendo mais expresso em amostra tumoral (Figuras 4.6 e 4.7 e Tabelas 4.11 e 4.12).

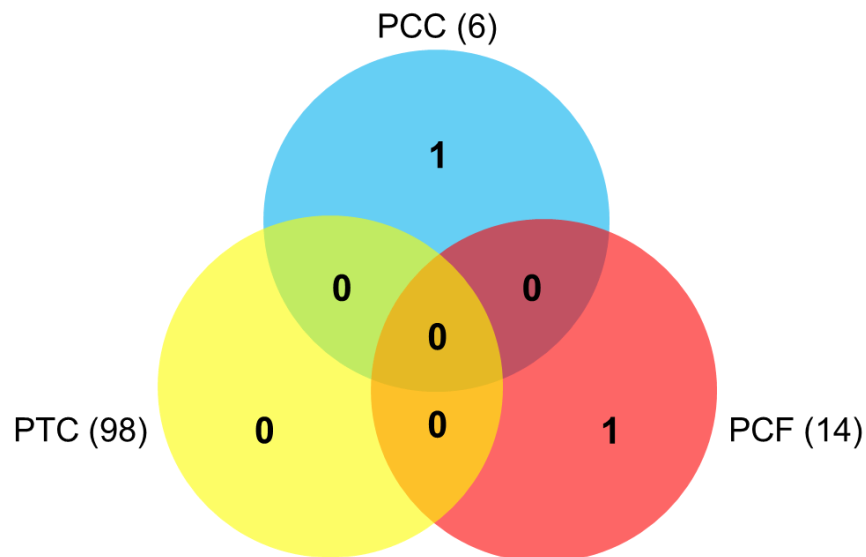


Figura 4.6: Diagrama de Venn comparando a quantidade de snoRNA mais expressos (FDR < 0,05 e logFC ≥ 1,5) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular. PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo.

Tabela 4.11: SnoRNAs mais expressos ($FDR < 0,05$ e $\log FC \geq 1,5$) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados. Coluna 1: subtipos tumorais. Coluna 2: a quantidade de sncRNA encontrado. Coluna 3: nomenclatura dos sncRNAs encontrados.

Subtipos	Quantidade	snoRNAs
PTC, PCF e PCC	0	-
PTC e PCF	0	-
PTC e PCC	0	-
PCC e PCF	0	-
PTC	0	-
PCF	1	HBII-82 (SNORD111)
PCC	1	ACA31 (SNORA31)

Abreviações: PCF: Carcinoma papilífero variante folicular

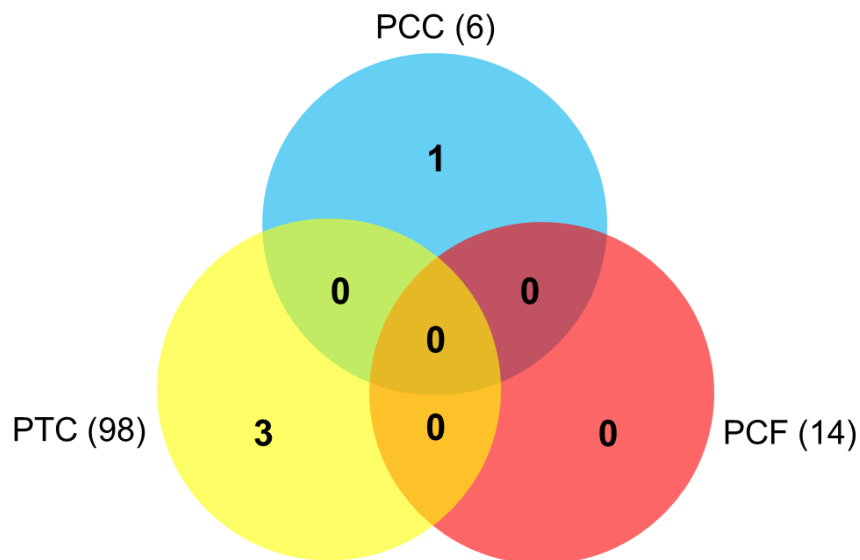


Figura 4.7: Diagrama de Venn comparando a quantidade de snoRNA menos expressos ($FDR < 0,05$ e $\log FC \geq 1,5$ ou $\leq -1,5$) presentes nos três diferentes subtipos de câncer de tireoide. Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular. PCC: Carcinoma papilífero variante Células Altas. Entre parênteses o número de amostra de cada subtipo.

Tabela 4.12: SnoRNAs menos expressos ($FDR < 0,05$ e $\log FC \leq -1,5$) detectados exclusivamente em um subtipo, compartilhada entre os subtipos ou presentes em todos os subtipos de câncer de tireoide pesquisados.

Subtipos	Quantidade	snoRNAs
PTC, PCF e PCC	0	-
PTC e PCF	0	-
PTC e PCC	0	-
PCF e PCC	0	-
PTC	3	mgU2-25/61 HBII-382 (SCARNA2), HBII-289 (SNORD89)
PCF	0	-
PCC	1	ACA3-2 (SNORD45)

Abreviações: PTC: Carcinoma papilífero; PCF: Carcinoma papilífero variante folicular; PCC: Carcinoma papilífero variante Células Altas

4.5. Análise de dispersão dos dados de expressão de sncRNAs

A análise dos sncRNAs constitutivamente expressos foi realizada entre as amostras tumorais e as amostras normais adjacentes ao tumor em cada subtipo de câncer de tireoide (PCC, PCF e PTC), a fim de identificar os sncRNAs que apresentaram menor variação da expressão. Foram considerados sncRNAs com expressão uniforme cujos valores de desvio padrão foram inferiores 1 e $\log FC$ entre 1 e -1 e considerados como os sncRNAs constitutivamente expressos e menor desvio padrão os valores de desvio padrão inferiores 0,5 e $\log FC$ entre 0,1 e -0,1. Assim, foram identificados 33 sncRNAs considerados como constitutivamente expressos. No subtipo Carcinoma papilar, variante Células Altas, foram encontrados 16 sncRNAs constitutivamente expressos com menor desvio padrão, 11 no subtipo tumoral Carcinoma papilar, variante folicular, e outros 7 no subtipo Carcinoma folicular (Tabela 4.13-3.16).

Tabela 4.13: Quantidade de sncRNAs com expressão uniforme nos subtipos tumorais Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas utilizando os filtros: desvio padrão inferiores 1 e logFC entre 1 e -1.

Subtipo tumoral	sncRNA			Total de sncRNAs
	miRNA	piRNA	snoRNA	
Carcinoma papilífero, variante Células Altas	142	1	80	222
Carcinoma papilífero, variante folicular	158	1	66	225
Adenocarcinoma papilar	166	1	76	243

Tabela 4.14: Quantidade dos sncRNAs constitutivamente expressos e menor desvio padrão nos subtipos tumorais Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas desvio padrão inferiores 0,5 e logFC entre 0,1 e -0,1.

Subtipo tumoral	sncRNA			Total de sncRNAs
	miRNA	piRNA	snoRNA	
Carcinoma papilífero, variante Células Altas	12	0	4	16
Carcinoma papilífero, variante folicular	10	0	1	10
Adenocarcinoma papilar	5	0	2	7

4.5.1. MiRNA constitutivamente expressos

Vinte e seis miRNAs constitutivamente expressos e com menor desvio padrão foram identificados, dos quais 12 foram identificados no subtipo tumoral Carcinoma papilífero, variante Células Altas, 10 no subtipo Carcinoma papilífero, variante folicular, cinco no subtipo Carcinoma papilífero e um em comum nos subtipos Carcinoma papilífero, variante Células Altas e Carcinoma papilífero (Tabela 4.15).

Tabela 4.15: MiRNAs constitutivamente expressos e com menor desvio padrão, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.

	Carcinoma papilífero variante, Células Altas		Carcinoma papilífero, variante folicular		Carcinoma papilífero	
	DP	logFC	DP	logFC	DP	logFC
hsa-let-7a-3	0,32	-0,02	0,58	-0,19	0,71	0,18
hsa-let-7c	0,28	-0,07	NA	NA	0,73	-0,23
hsa-let-7d	0,77	0,98	0,33	0,06	0,44	0,38
hsa-miR-1180-3p	0,46	0,02	0,52	0,15	0,71	-0,36
hsa-miR-135b-5p	NA	NA	0,45	-0,09	0,8	0,54
hsa-miR-15b-3p	0,27	0,06	0,66	-0,42	0,64	0,05
hsa-miR-17-5p	0,30	0,11	0,43	-0,12	0,50	0,06
hsa-miR-191-5p	0,28	0,04	0,45	0,17	0,49	0,21
hsa-miR-192-5p	0,57	0,53	0,64	-0,55	0,43	< 0,01
hsa-miR-200c-5p	0,66	0,43	0,33	< -0,01	0,81	-0,14
hsa-miR-24-3p	0,30	0,08	0,58	0,13	0,60	0,37
hsa-miR-29c-5p	0,54	-0,45	0,8	0,46	0,45	-0,05
hsa-miR-30a-3p	0,90	-0,95	0,46	-0,10	0,79	-0,50
hsa-miR-30e-3p	0,53	-0,47	0,48	0,03	0,62	-0,31
hsa-miR-335-5p	0,14	< -0,01	0,71	0,05	0,90	-0,16
hsa-miR-339-3p	0,88	-0,12	0,44	-0,08	0,50	-0,37
hsa-miR-361-5p	0,55	0,71	0,45	0,03	0,40	0,25
hsa-miR-374a-3p	0,32	< -0,01	0,71	0,12	0,87	0,08
hsa-miR-4677-3p	NA	NA	0,47	0,08	NA	NA
hsa-miR-501-3p	0,32	-0,05	NA	NA	0,52	0,09
hsa-miR-505-3p	0,46	0,10	0,26	-0,17	0,45	-0,07
hsa-miR-532-3p	0,28	-0,07	0,80	< -0,01	0,51	-0,28
hsa-miR-589-5p	0,77	0,46	0,44	0,059	0,52	0,11
hsa-miR-93-3p	NA	NA	0,67	0,13	0,47	0,08
hsa-miR-93-5p	0,54	0,42	0,49	0,08	0,53	0,08
hsa-miR-98-5p	0,26	0,01	0,44	-0,27	0,56	-0,02

Legenda: Em vermelho os miRNAs constitutivamente expressos e com menor desvio padrão. NA: dados não detectados após a análise da dispersão utilizando os filtros adaptados de Eisenberg e Levanon (2013).

4.5.2. PiRNAs constitutivamente expressos

Para identificar os piRNAs os filtros utilizados foram menos restritivos e considerados como constitutivamente expressos os piRNAs com desvio padrão menor que 1 e logFC entre 1 e -1. Com isso foi identificado o hsa-piR-009294 constitutivamente expresso em todos os subtipos tumorais: Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero (Tabela 4.16).

Tabela 4.16: PiRNAs constitutivamente expressos, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.

sncRNA	Carcinoma papilífero, variante Células Altas		Carcinoma papilífero, variante folicular		Carcinoma papilífero	
	DP	logFC	DP	logFC	DP	logFC
hsa-piR-009294	0,69	0,04	0,80	0,50	0,97	0,38

4.5.3. SnoRNAs constitutivamente expressos

Foram encontrados um total de sete snoRNAs constitutivamente expressos e com menor desvio padrão. Para o subtipo tumoral Carcinoma papilífero, variante Células Altas, foram encontrados quatro snoRNAs em amostras normais adjacentes aos tumores e tumorais. No Carcinoma papilífero, variante folicular foi encontrado um snoRNAs HBII-296A constitutivamente expressos. No subtipo Carcinoma papilífero foram identificados dois snoRNAs constitutivamente expressos (Tabela 4.17).

Tabela 4.17: SnoRNAs constitutivamente expressos e com menor desvio padrão, valor do desvio padrão (DP) e logFC dos subtipos tumorais Carcinoma papilífero variante, Células Altas, Carcinoma papilífero, variante folicular e Carcinoma papilífero.

	Carcinoma papilífero, variante Células Altas		Carcinoma papilífero, variante folicular		Carcinoma papilífero	
	DP	logFC	DP	logFC	DP	logFC
ACA36B (SNORA36B)	0,39	< -0,01	0,58	0,18	0,54	-0,11
HBI-43 (SNORD17)	NA	NA	0,62	0,16	0,49	0,01
HBII-239 (SNORD71)	0,36	0,08	0,50	0,47	0,60	-0,16
HBII-296A	0,57	-0,57	0,45	0,05	0,75	-0,15
U106 (SNORD91A)	0,41	0,07	0,64	0,63	0,56	0,30
U38A (SNORD38A)	0,48	0,06	0,75	0,30	0,78	-0,48
U43 (SNORD43)	0,38	0,24	0,70	-0,21	0,50	0,07

Legenda: Em vermelho os snoRNAs constitutivamente expressos e com menor desvio padrão. NA: dados não detectados após a análise da dispersão utilizando os filtros adaptados de Eisenberg e Levanon (2013).

4.6. Integração das expressões diferencial e constitutiva de sncRNAs

Ao integrar os dados das expressões diferencial e constitutiva de sncRNAs, foram identificados aqueles que apresentaram padrão de expressão similar. Assim, buscamos pelo padrão de expressão similar de tal forma que fossem identificados sncRNAs mais expressos diferencialmente em amostras normais adjacentes aos tumores e tumorais em todos subtipos analisados. Encontramos que Carcinoma papilífero, variante Células Altas e Carcinoma papilífero possuem padrão de expressão similar entre si e diferentes quando comparados com o padrão de expressão constitutiva em amostras normais adjacentes aos tumores e tumorais do terceiro subtipo tumoral (Carcinoma papilífero, variante folicular). Para esta análise, pode-se destacar os miRNAs que apresentaram padrão de expressão similar com valores de FDR e logFC estatisticamente significativos e considerados como diferencialmente expressos em dois subtipos e desvio padrão baixo considerado como constitutivamente expressos: hsa-miR-181a-2-3p, hsa-miR-20b-5p e hsa-miR-21-5p (Tabela 4.18). Destacamos ainda os miRNAs que apresentaram padrão de expressão similar com valores de FDR e logFC estatisticamente significativos e considerados como diferencialmente expressos no subtipo Carcinoma papilífero, variante folicular, e desvio padrão baixo considerado como constitutivamente expressos nos outros dois subtipos: hsa-miR-100-5p, hsa-miR-1251-5p e hsa-miR-183-5p.

Tabela 4.18: Integração da expressão diferencial e constitutiva utilizando os sncRNAs diferencialmente expressos, logFC, FDR e desvio padrão (DP).

sncRNA	Carcinoma papilífero, variante Células Altas			Carcinoma papilífero, variante folicular			Carcinoma papilífero		
	logFC	FDR	DP	logFC	FDR	DP	logFC	FDR	DP
hsa-miR-100-5p	-0,90	0,36	0,69	-1,79	0,02	1,72	-0,72	< 0,01	0,89
hsa-miR-1251-5p	-0,22	0,87	0,63	1,72	< 0,01	3,68	0,58	< 0,01	1,05
hsa-miR-181a-2-3p	2,31	< 0,01	1,52	0,76	0,28	0,96	1,74	< 0,01	1,23
hsa-miR-183-5p	1,39	0,08	1,22	2,17	< 0,01	1,93	0,91	< 0,01	0,99
hsa-miR-20b-5p	-2,21	< 0,01	1,71	-0,38	0,72	0,99	-1,64	< 0,01	2,06
hsa-miR-21-5p	2,37	< 0,01	2,28	0,13	0,87	0,67	2,04	< 0,01	2,04

Legenda: Vermelho: Valores dentro dos padrões de FDR e logFC demonstrando os sncRNAs diferencialmente expressos. Negrito: sncRNAs que apresentaram padrão de expressão similar

5. DISCUSSÃO

Os RNAs não codificadores pequenos (sncRNAs) referem-se à uma classe de RNAs sem a função de codificar para proteínas. Foram descritos desempenhando papéis fundamentais no desenvolvimento, proliferação, diferenciação e apoptose, podem atuar como supressores tumorais e estão associados ao desenvolvimento de doenças (LI *et al.*, 2014; YANG *et al.*, 2016). Diferentes autores citaram o papel de sncRNAs em diversos tipos de câncer. Aguiar e colaboradores (2010) descreveram o potencial de miRNA na carcinogênese oral. Cheng e colaboradores (2011) relataram que o piR-651 pode estar envolvido no desenvolvimento de câncer gástrico e outros tipos de tumorais, destacando-o como um potencial biomarcador para o diagnóstico de câncer. Martens-Uzunova e colaboradores (2012) mencionaram o papel dos miRNAs e snoRNAs no câncer de próstata. Huang e colaboradores (2013) e Hashim e colaboradores (2014) citaram piRNAs que desempenham papel no câncer da mama e atuam como marcadores tumorais. Kishikawa e colaboradores (2015) descreveram SNORD33, SNORD66 e SNORD76 em câncer de pulmão e diferentes miRNAs e piRNAs em câncer de gástrico. Wó Jcicka e colaboradores (2016) destacaram o alto potencial dos miRNAs como novas ferramentas terapêuticas e diagnósticas em carcinomas da tireoide. Portanto, evidencia-se a importância de se estudar tais sncRNAs e seu papel como potenciais biomarcadores para diferentes tipos tumorais.

Algumas bases de dados que disponibilizam publicamente sncRNAs são miRBase, piRNABank e UCSC. O miRBase (<https://mirbase.org/>) é o repositório de sequências de microRNA, suas anotações e coordenadas genômicas, foi criado em 2002 e atualmente contém 24.521 *loci* de microRNA de 206 espécies, que podem produzir 30.424 miRNAs maduros e 1.761 *loci* de miRNAs de 38 espécies anotados com alta confiabilidade (KOZOMARA & GRIFFITHS-JONES, 2014). O piRNABank (<https://pirnabank.ibab.ac.in/>) armazena sequências empiricamente dos piRNAs relatados em humanos, ratos e camundongos, anotações e outras informações relacionadas, atualmente, a base de dados possui sequências únicas de piRNA, que estão mapeando para *loci* únicos ou múltiplos no genoma correspondente totalizando 23.439 de humanos, 39.986 de camundongo e 38.549 de ratos (SAILAKSHMI & AGRAWAL, 2008). As sequências humanas são nomeadas de hsa_piR_000001 a hsa_piR_023439. Similarmente, prefixos de “mmu” e “rno” foram

usados para nomear os piRNAs de camundongo e rato, respectivamente (SAI LAKSHMI & AGRAWAL, 2008). O UCSC (<https://genome.ucsc.edu>) foi lançado em 2001 e possui anotação de snoRNAs de humanos e camundongos (HINRICHS *et al.*, 2006).

Mais de 200 formas de câncer já foram descritas e cada tipo pode ser caracterizado por diferentes perfis moleculares demandando estratégias terapêuticas exclusivas. O TCGA (do inglês: *The Cancer Genome Atlas*) é um projeto público, coordenado pelo NCI (do inglês: *National Cancer Institute*), NIH (do inglês: *National Institutes of Health*) e NHGRI (do inglês: *National Human Genome Research Institute*). Tal projeto tem como objetivo catalogar e descobrir as principais alterações genômicas causadoras de câncer para criar um atlas abrangente de perfis genômicos de câncer. Até o momento, já foram analisados mais de 30 tumores humanos provenientes do sequenciamento genômico e análises multidimensionais integradas. A genômica do câncer está fornecendo novas informações sobre o desenvolvimento do câncer e do seu comportamento, bem como novos conhecimentos sobre as alterações genéticas e vias moleculares. Portanto, disponibilizar tais dados publicamente ajuda a melhorar os métodos de diagnósticos, padrões de tratamento e possível cura para o câncer (TOMCZAK *et al.*, 2015; RIESCO-EIZAGUIRRE & SANTISTEBAN, 2016).

A expressão diferencial é avaliada aplicando testes estatísticos à contagem dos *reads* nas diferentes amostras que são normalizados pelo número total de *reads* mapeados do genoma (BORIA *et al.*, 2012). Tal abordagem vem sendo amplamente utilizada para estudos de possíveis biomarcadores para câncer de tireoide (CHEN *et al.*, 2008; CORBETTA *et al.*, 2010; BUDA *et al.*, 2012; GRAHAM *et al.*, 2015; STOKOWY *et al.*, 2016), bem como para câncer de mama (HUANG *et al.*, 2013), câncer pancreático (MÜLLER *et al.*, 2015), mesotelioma (AMATYA *et al.*, 2016), adenocarcinoma endometrial (JURCEVIC *et al.*, 2016) e câncer de pulmão (SUN *et al.*, 2016).

Uma estratégia para a avaliação do perfil de expressão de sncRNAs é o estudo *in vitro* e os nossos resultados *in silico* corroborando com alguns estudos publicados na literatura. Mancikova e colaboradores (2015) analisou experimentalmente a expressão de miRNA em tumores de tireoide do tipo adenoma, folicular e papilífero. Segundo a nossa comparação, houve concordância no padrão

de expressão de sete miRNAs presentes no subtipo papilífero (Tabela 4.5 e Figura 4.1). Ainda que ambos tenham utilizado metodologias distintas, foi evidenciada a eficácia das abordagens, demonstrando assim, que o presente estudo feito *in silico* pode servir de base para futuros estudos funcionais uma vez que, análises experimentais e *in silico* sugerem que a expressão miRNA anormal pode participar de processos biológicos fundamentais para a invasão e disseminação do câncer podendo contribuir na agressividade de PTC (Geraldo *et al.*, 2016). Os nossos resultados também corroboram com o estudo de Yu e colaboradores (2012) que utilizaram amostras de carcinoma papilífero e identificou dez miRNAs. Segundo nossas análises, encontramos quatro miRNAs (miR-100, miR-127, miR-144 e miR-222) com padrão de expressão similar ao encontrado pelos autores. Os nossos achados também estão de acordo com o estudo de Rossing e colaboradores (2012), no qual os autores encontram a expressão de miR-199b-5p e miR-144 em câncer de tireoide dos subtipos Carcinoma papilífero e carcinoma folicular. Similarmente, nossos dados corroboramos com os dados de Yoruker e colaboradores (2016), que utilizaram amostras de pacientes para identificar sete miRNAs diferencialmente expressos em tumores de tireoide que podem desempenhar papel no desenvolvimento de carcinoma papilífero da tireoide. Destes, encontramos cinco (miR-21, miR-31, miR-146b, miR-221, miR-222) mais expressos em amostras tumorais.

Outra estratégia para a análise do perfil de expressão é usar amostras pareadas de tecido tumoral e tecido normal adjacente ao tumor do mesmo paciente. Apesar da dificuldade de se conseguir dados públicos de amostras de câncer de tireoide pareadas, nota-se que a abordagem *in silico* para reuso de dados disponíveis publicamente corresponde aos resultados obtidos que usaram técnicas precisas de quantificação do perfil de expressão de genes. Swierniak e colaboradores (2013) quantificaram a expressão diferencial entre amostras pareadas de pacientes do subtipo tumoral carcinoma papilífero de pacientes. Os referidos autores encontraram que os miRNAs mais desregulados foram miR-146b-5p, miR-221-3p, miR-7-3p, miR-551b-3p, miR-486-3p e miR-144-3p. Corroboramos estes resultados ao encontrar como diferencialmente expressos os miRNAs miR-146b-5p, miR-221-3p e miR-144-3p. Estão de acordo também com os resultados de Suresh e colaboradores (2015). Estes autores analisaram amostras pareadas e identificaram seis miRNAs (miR-21,

miR-31, miR-146b, miR-221, miR-222, miR-3613) mais expressos em amostras tumorais e dois (miR-138, miR-98) mais expressos em amostras normais adjacentes aos tumores. Usando a nossa abordagem, encontramos seis miRNAs diferencialmente expressos nas respectivas amostras, miR-98 foi constitutivamente expresso e o miR-3613 não foi encontrado neste estudo. Zhang e colaboradores (2013) analisaram o perfil de miRNAs e detectaram com sucesso a expressão diferencial entre câncer papilífero da tireoide e tecido normal adjacentes aos tumores dos quais destacou que miR-146b como mais expresso em amostras tumorais, sendo este também encontrado no presente estudo.

Os miRNAs se tornaram componentes moleculares chave na identificação de células em estados normais e patológicos. Sua desregulação em células cancerosas foi relatada pela primeira vez em 2002 e vem sendo amplamente estudada proporcionando novas oportunidades diagnósticas e terapêuticas (HAYES *et al.*, 2014). Dos 40 miRNAs diferencialmente expressos encontrados neste estudo destacam-se os hsa-miR-146, hsa-miR-221 e hsa-miR-222 que foram mais expressos em amostras tumorais dos três subtipos Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas e hsa-miR-204 menos expresso em tais amostras tumorais. Presentes em apenas um subtipo foram considerados como mais diferencialmente expressos os miRNAs: hsa-miR-146-3p no Carcinoma papilífero variante Células Altas, hsa-miR-182 no Carcinoma papilífero variante folicular e hsa-miR-508-3p no Carcinoma papilífero.

Os microRNAs hsa-miR-221 e hsa-miR-222 estão localizados no cromossomo X, são altamente conservados em vertebrados, estão envolvidos na regulação do ciclo celular e apoptose e são conhecidos por suas desregulações em várias malignidades (CHEN *et al.*, 2008; GAROFALO *et al.*, 2012; ZHANG *et al.*, 2016). Estudos funcionais mostraram que esses miRNAs modificam negativamente muitos genes supressores de tumores, incluindo p27, p57, DDIT4, PTEN e TIMP3 (WANG & LI, 2012). Diversos autores já identificaram hsa-miR-221 e hsa-miR-222 como desregulados em diferentes tipos tumorais, dentre os quais câncer de tireoide e sugerem que estas moléculas desempenham papel chave, tanto nos primeiros passos do processo de transformação maligna na carcinogênese da tireoide quanto nas formas avançadas da doença, ou seja, eles podem atuar como reguladores no câncer de tireoide (RICARTE FILHO & KIMURA, 2006; MANCIKOVA *et al.*, 2015).

Vriens e colaboradores (2012) demonstraram que o bloqueio do miR-221 e miR-222 foi capaz de inibir o crescimento de uma linhagem celular de carcinoma papilífero da tireoide e o crescimento das mesmas células foi estimulado pela super-expressão do miR-221 e miR-222. Lee e colaboradores (2013) citaram que tais miRNAs podem ser utilizados para classificar os tipos de tumores da tireoide e diferenciar tumores malignos de benignos. Tais resultados estão em conformidade com o presente estudo que achou também hsa-miR-221 e hsa-miR-222 mais expresso em amostras tumorais.

O miR-146b desempenha um papel importante na progressão tumoral (YORUKER *et al.*, 2016). Já foi relatado a presença de miR-146 é diferencialmente expresso nas seguintes condições: pós-operatório em pacientes com tumores de tireoide maiores que dois centímetros (YORUKER *et al.*, 2016), em amostras tumorais em PTC (SCHNEIDER *et al.*, 2015). Tal miRNA é associado com o risco de recorrência de PTC e pode promover a migração e invasão celular. Outros dados sugerem também que miR-146b é superexpresso em muitos tipos de câncer, como melanoma, pulmonar, colorretal e tireoide (LIMA *et al.*, 2016; YORUKER *et al.*, 2016). A sua desregulação está correlacionada com a agressividade de PTC (DETTMER *et al.*, 2013), estes poderiam ser utilizados como biomarcadores para o câncer de tireoide (SCHNEIDER *et al.*, 2015) e sua inibição terapêutica representa uma maneira promissora de aumentar a eficácia da terapia radioativa (LUDVÍKOVÁ *et al.*, 2015).

O miR-204 é mais expresso em tecidos tumorais de amostras de câncer papilífero da tireoide (DETTMER *et al.*, 2013), foi proposto como supressor tumoral, (WÓ JCICKA *et al.*, 2016), está associado à presença da mutação BRAF e descrito prejudicando vários processos celulares, como aqueles que mantêm a fisiologia epitelial, apoptose e ciclo celular (MANCIKOVA *et al.*, 2015; AB MUTALIB *et al.*, 2016). Este miRNA é conhecido como um supressor de tumores e é regulado negativamente em vários tipos de câncer, incluindo carcinomas renais e tiroidianos (AB MUTALIB *et al.*, 2016).

MiR-182 possui papel potencial como oncogene em PTC e é um alvo terapêutico putativo neste tipo tumoral (WÓ JCICKA *et al.*, 2016). Outrossim, está associado com o tamanho do tumor, metástase dos linfonodos, estágio e prognóstico em câncer colorretal (WANG *et al.*, 2015). Zhu e colaboradores (2014) indicaram a correlação de miR-182 em PTC, descrevendo que este miRNA promove

a proliferação e invasão celular por meio da supressão direta e sua inibição pode ser utilizada como uma nova estratégia terapêutica contra PTC e outros subtipos de câncer de tireoide, bem como em outros tipos de câncer.

Estudos revelaram que a alta expressão de miR-508 tem papel importante na progressão de Carcinoma epidermóide esofágico e em glioma (LIN *et al.*, 2014; BAO *et al.*, 2016) e a menor expressão de miR-508 pode ser associada com melhor sobrevida de pacientes com câncer pancreático e maior expressão em amostras tumorais de câncer de pâncreas (ALI *et al.*, 2015). Zhai e colaboradores (2012) observaram que este miRNA inibiu acentuadamente a proliferação de células cancerosas, sugerindo o seu papel como potencial supressor tumoral em carcinoma renal. O miR-508 também está associado a regulação da resistência a múltiplos tratamentos no câncer gástrico. As amostras de câncer gástrico com menor expressão de miR-508-5p tende a ser resistente à quimioterapia (SHANG *et al.*, 2013, 2015). Isto posto, os níveis de miR-508-5p podem contribuir para a reversão da resistência a múltiplos fármacos em práticas clínicas futuras (SHANG *et al.*, 2015).

Entretanto, existem outras classes de sncRNAs que foram pouco exploradas na literatura de tumores de tireoide. SnoRNAs são RNAs não codificadores pequenos conhecidos por orientar a modificação pós-transcricional de outros RNAs não codificadores de proteínas, tais como rRNA e RNAs nucleares pequenos (JORJANI *et al.*, 2016). Os snoRNAs foram bem caracterizados e estão associados a certos tipos tumorais apontando assim, seu potencial como novo biomarcador e alvo terapêutico para o câncer. (MANNOOR *et al.*, 2012; THORENOOR & SLABY, 2014). No presente estudo utilizando amostras de câncer de tireoide foram encontrados seis diferentes snoRNAs, dos quais HBII-82 e ACA31 foram mais expressos em amostras tumorais e os snoRNAs ACA3-2, HBII-289, mgU2-25 e HBII-382 foram mais expressos em amostras normais adjacentes aos tumores quando comparados à sua expressão em amostras tumorais.

De acordo com a nova nomenclatura adotada em 2006 pelo Comitê Internacional HUGO Gene Nomenclature os nomes dos snoRNAs da família C/D são representados como SNORDn, onde "n" é um número e os nomes da família H/ACA são apresentados como SNORAn. No entanto, a forma a qual os snoRNAs foram nomeados antes da introdução da nova nomenclatura permanece amplamente

utilizada (THORENOOR & SLABY, 2014). ACA3-2 também conhecido como SNORA45, SNORA45B e SNORA3B pertence à família H/ACA box foi clonado por Gu e colaboradores (2005) a partir de células sanguíneas. HBII-289 também conhecido como SNORD89 é um RNA nucleolar pequeno e pertence a família C/D box (DUPUIS-SANDOVAL *et al.*, 2015). MgU2-25/HBII-382 são moléculas relacionadas, denominadas SCARNA2. Os pequenos RNAs do corpo de Cajal (scaRNAs) podem orientar tanto a metilação quanto a pseudouridilação dos RNAs spliceosômicos transcritos com a RNA Polimerase II (THORENOOR & SLABY, 2014). ACA31 também conhecido como SNORA31 (Ronchetti *et al.*, 2013), é predito para funcionar na modificação dos rRNAs 18S e 28S (Kiss *et al.*, 2004) e foi descrito por Ronchetti e colaboradores (2013) com função de codificar uma proteína crítica envolvida no controle e reversão de tumor. HBII-82 ou SNORD111 foi identificado em neurônios e células linfoblastoides e pertence a família C/D box (JORJANI *et al.*, 2016).

Os piRNAs é a classe de RNAs não codificadores pequenos mais nova, estão envolvidos no desenvolvimento da linha germinal, no silenciamento de elementos de transposição, regulação gênica, apoptose e metilação do DNA (CHENG *et al.*, 2011), além de serem menos estudados em câncer quando comparados a outros sncRNAs. Contudo, tais moléculas vêm sendo relatadas como importantes na ocorrência de diversos tipos de câncer, prognóstico e tratamento (HUANG *et al.*, 2014; ASSUMPÇÃO *et al.*, 2015; KRISHNAN *et al.*, 2016). Encontramos a partir da análise de dispersão em amostras de câncer de tireoide o piRNA hsa-piR-009294. Tal piRNA já foi identificado por Krishnan e colaboradores (2015) como um RNA não codificador pequeno diferencialmente expresso, com maior expressão em amostras tumorais de câncer de mama. Martinez e colaboradores (2015) utilizando 6.260 amostras (508 não malignas e 5.752 tumores) processadas pela base de dados TCGA e sncRNAs provenientes da base de dados miRBase (versão 20) e The Functional RNA Database (versão 3.0) identificaram o FR069557 como o piRNA mais diferencialmente expresso entre as amostras tumorais de tireoide e os demais tipos de câncer. Entretanto, pesquisas subsequentes tal piRNA teve a anotação alterada para o miRNA hsa-miR-138-1. A base de dados miRBase em sua última versão (versão 21) disponibiliza um subconjunto de dados de alta confiança, com base no padrão dos reads mapeados e diversos critérios, dos quais o hsa-miR-138-1

está anotado com alta confiabilidade. O conjunto de dados de microRNA de alta confiança está disponível ao lado da coleção completa de microRNA (http://www.mirbase.org/cgi-bin/miRNA_summary.pl?org=hsa) (KOZOMARA & GRIFFITHS-JONES, 2014).

A expressão aberrante de miRNAs em tumores da tireoide vem sendo relatada e tais moléculas emergiram como uma potencial ferramenta de diagnóstico, capaz de distinguir diferentes subtipos de câncer de tireoide e servir como marcadores para monitorar a recorrência do carcinoma papilífero da tireoide (PTC) (GERALDO & KIMURA, 2015). Este subtipo tumoral de câncer de tireoide é a malignidade endócrina mais comum (CHRUŚCIK & LAM, 2015) e pode ser dividido em subtipos (figura 1.1). A partir da base de dados TCGA, obtivemos dados dos subtipos tumorais Carcinoma papilífero, variante Células Altas (PCC), Carcinoma papilífero, variante folicular (PCF) e Carcinoma papilífero, NOS. NOS (do inglês: *Not Otherwise Specified*) é um termo que pode ser empregado com finalidade de: 1) termo morfológico não modificado; 2) termo morfológico com um adjetivo que não aparece em outra parte; 3) termo morfológico com sentido geral (*International Classification of Disease for Oncology*, 3ª edição). Em nossas análises, foram identificados os sncRNAs hsa-miR-181a-2-3p, hsa-miR-20b-5p e hsa-miR-21-5p, com padrões de expressão semelhantes entre dois dos subtipos (Carcinoma papilífero, variante Células Altas e Carcinoma papilífero, NOS) e discordantes com terceiro subtipo analisado, Carcinoma papilífero, variante folicular. Assim, pode-se especular que em alguns casos apesar do diagnóstico patológico de carcinoma papilífero, este possui um padrão de expressão de sncRNAs muito similar com Carcinoma papilífero, variante Células Altas, enquanto Carcinoma papilífero, variante folicular possui padrão de expressão diferenciado quando comparado aos outros dois subtipos analisados.

O miR-21 é superexpresso em muitos tipos tumorais, incluindo câncer de mama, glioblastoma, carcinoma hepatocelular, câncer de pulmão, estômago, colorretal e próstata (HAYES *et al.*, 2014). Especificamente, em câncer de tireoide miR-21 é diferencialmente expresso em amostras tumorais, principalmente em formas mais agressivas de PTC, podendo ser considerado como um evento crítico na sua patogênese, ademais, ajudar a distinguir as variantes de tumor papilífero (NETWORK *et al.*, 2014; MANCIKOVA *et al.*, 2015; YORUKER *et al.*, 2016). Apesar de estudos *in vivo* mostrarem a alta expressão de miR-21, resultando no aumento

de crescimento tumoral (YANG *et al.*, 2016), Yoruker e colaboradores (2016) citaram que dependendo do subtipo tumoral, miR-21 pode ser mais expresso ou menos expressos em amostras tumorais quando comparadas com amostras normais adjacentes aos tumores. O miR-21 é um dos miRNAs mais relacionados com câncer e sua regulação poderia aumentar o crescimento tumoral, metástases e invasão, além de estarem associados a redução da sensibilidade à quimioterapia por seus vários alvos (ZHOU *et al.*, 2014).

O miR-20b mostrou expressão aberrante em vários tipos de tumores, sua super-regulação é associada com pior prognóstico e pode afetar a viabilidade celular e apoptose (XUE *et al.*, 2015). Hong e colaboradores (2016) mostram que miR-20b funciona como um supressor tumoral *in vitro* e *in vivo* em pacientes com PTC. Wang e colaboradores identificaram que miR-21 promove a proliferação, migração, invasão e tumorigenicidade em pacientes com câncer esofágico (Wang *et al.*, 2016). Ao e colaboradores (2016) relataram que o miR-20b favorece a sobrevivência de células de câncer de mama e sua alta expressão pode promover metástases cerebrais. Xue e colaboradores verificaram que os níveis de expressão de tal microRNA aumentam nos tecidos de carcinoma hepatocelular em comparação com o tecido normal. Cong e colaboradores (2015) descreveram a associação do miR-20b à invasão e progressão de PTC. A regulação ascendente de miR-20b em PTC inibe a viabilidade celular, migração e invasão em células, sugerindo que tal miRNA pode desempenhar um papel importante na iniciação, progressão e metástase do câncer papilífero da tireoide e pode fornecer um potencial alvo terapêutico para PTC (Hong *et al.*, 2016).

MiR-181a-2 é membro da família miR-181 (RADY *et al.*, 2017) e evidências indicam que esse microRNA tem expressão nos tecidos tumorais, sugerindo um papel potencialmente importante no desenvolvimento e/ou progressão de tumores (LIU *et al.*, 2014). Diferentes autores identificaram tal miRNA em distintas amostras tumorais como em: pacientes com câncer de mama em comparação com indivíduos saudáveis (RADY *et al.*, 2017), alta expressão em glioblastoma (AGUIAR *et al.*, 2010), na progressão do câncer de ovário (PARIKH *et al.*, 2014) e desregulação em PTC (DETTMER *et al.*, 2013) podendo estar relacionado a reincidência em pacientes com carcinoma papilífero, variante folicular (WÓ JCICKA *et al.*, 2016). Wó Jcicka e colaboradores (2016) identificaram que miR-181a-5p tem expressão

significativamente elevada no plasma de pacientes com câncer de tireoide em comparação com o indivíduo controle e pacientes com outros tipos de tumorais, incluindo câncer de mama, pulmão, cólon e melanoma. Assim sendo, o miR-181a-5p tem sido descrito na patogênese, desenvolvimento, progressão, metástase, prognóstico e resposta terapêutica à quimioterapia e radioterapia em carcinoma de tireoide (SONG *et al.*, 2017). A função miR-181b pode ser única, dependendo do tipo de tumor e contexto celular. Compreendendo os genes alvo e as redes reguladoras de miR-181b pode-se usar a abordagem de biologia de sistemas, para determinar o potencial de miR-181b como biomarcador e novo alvo terapêutico (LIU *et al.*, 2014).

Assim, conclui-se pelos dados aqui apresentados e corroborados pela literatura, que piRNA e snoRNA são expressos em tecidos normais adjacentes aos tumores e em tumores de tireoide. Encontramos uma gama de sncRNAs capazes de serem testados como potenciais biomarcadores para os subtipos analisados, bem como a possibilidade de classifica-los em termos moleculares.

Em termos de bioinformática, este estudo propiciou o amadurecimento de uma metodologia que permite a identificação de miRNA, piRNA e snoRNA em dados de sequenciamento de alto desempenho oriundos do projeto TCGA. Os dados apresentados nesta dissertação mostram o potencial desta metodologia e dos parâmetros utilizados para a seleção de casos de corroborados por dados da literatura.

6. PERSPECTIVAS

Realizar análises *in vitro* com os RNAs não codificadores pequenos diferencialmente e constitutivamente expressos encontrados no presente estudo.

Pretendemos durante o projeto de pesquisa a ser executado durante o doutorado, ampliar o estudo realizado no mestrado com tumor de um órgão humano para tumores de 23 órgãos humanos diferentes.

7. CONCLUSÕES

- Foi possível identificar RNAs não codificadores pequenos dos tipos miRNA, piRNA e snoRNA diferencialmente e constitutivamente expressos em amostras dos diferentes subtipos de câncer de tireoide Carcinoma papilífero, Carcinoma papilífero variante folicular e Carcinoma papilífero variante Células Altas através de dados públicos disponibilizados na base de dados TCGA usando uma abordagem *in silico* desenvolvida pelo nosso grupo de pesquisa.

- Os resultados encontrados no presente estudo corroboram com dados obtidos por outros autores em câncer de tireoide. Apesar da utilização de metodologias distintas, a nossa abordagem se mostra como uma alternativa viável na obtenção de potenciais marcadores prognósticos utilizados no melhoramento do diagnóstico molecular.

- Foram obtidos um total de 46 RNAs não codificadores pequenos, dos quais 6 são comuns para os 3 subtipos tumorais. Hsa-miR-146b-3p, hsa-miR-222-5p, hsa-miR-146b-5p, hsa-miR-221-3p e hsa-miR-222-3p foram identificados maior aumento de expressão nos tecidos tumorais em comparação com os tecidos controles adjacentes aos tumores e hsa-miR-204-5p mais expressos em amostras normais adjacentes aos tumores. Foram considerados como mais diferencialmente expressos os miRNAs hsa-miR-146-3p no Carcinoma papilífero variante Células Altas, hsa-miR-182 no Carcinoma papilífero variante folicular e hsa-miR-508-3p no Carcinoma papilífero.

- Foram identificados 34 sncRNAs constitutivamente expressos o hsa-piR-009294 foi encontrado e considerado como constitutivamente expresso em todos os subtipos tumorais analisados.

- Ao integrar os dados da expressão diferencial e desvio padrão foram identificados sncRNAs que apresentaram padrão de expressão similar nos subtipos tumorais Carcinoma papilífero, variante Células Altas e Carcinoma papilífero quando comparado com o padrão de expressão constitutiva do subtipo tumoral Carcinoma papilífero, variante folicular. Dentre os quais destacam-se: hsa-miR-181a-2-3p, hsa-miR-20b-5p e hsa-miR-21-5p.

- Tais resultados demonstram a importância de estudar sncRNAs na biologia do câncer e seu uso como potenciais biomarcadores.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- Ab Mutalib NS, Othman SN, Mohamad Yusof A, Abdullah Suhaimi SN, Muhammad R, Jamal R. Integrated microRNA, gene expression and transcription factors signature in papillary thyroid cancer with lymph node metastasis. *PeerJ*. 2016;4:e2119.
- Aguiar B, Francisco C, Nonaka W, Freitas RDA, Souza LB De, Pinto LP. MicroRNAs – Biogênese , funções e seu papel potencial na carcinogênese oral. *Odontol Clin-Cient*. 2010;9(2):105–9.
- Aguiar ERGR, Olmo RP, Marques JT. Virus-derived small RNAs: molecular footprints of host-pathogen interactions. *Wiley Interdiscip Rev RNA*. 2016;7.
- Aigner A, Buesen R, Gant T, Gooderham N, Greim H, Hackermüller J, et al. Advancing the use of noncoding RNA in regulatory toxicology: Report of an ECETOC workshop. *Regul Toxicol Pharmacol*. 2016;1–13.
- Ali S, Dubaybo H, Brand RE, Sarkar FH. exists as a Biomarker for Pancreatic Cancer. *J Cancer Sci Ther*. 2016;7(11):336–46.
- Amatya VJ, Mawas AS, Kushitani K, El-Din MMM, Takeshima Y. Differential microRNA expression profiling of mesothelioma and expression analysis of miR-1 and miR-214 in mesothelioma. *Int J Oncol*. 2016;48(4):1599–607.
- Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. *RNA*. 2003;9(3):277–9.
- Ander BP, Barger N, Stamova B, Sharp FR, Schumann CM. Atypical miRNA expression in temporal cortex associated with dysregulation of immune, cell cycle, and other pathways in autism spectrum disorders. *Mol Autism*. *Molecular Autism*; 2015;6:37.
- Anwar SL, Albat C, Krech T, Hasemeier B, Schipper E, Schweitzer N, et al. Concordant hypermethylation of intergenic microRNA genes in human hepatocellular carcinoma as new diagnostic and prognostic marker. *Int J Cancer*. 2013;133(3):660–70.
- Ao X, Nie P, Wu B, Xu W, Zhang T, Wang S, et al. Decreased expression of microRNA-17 and microRNA-20b promotes breast cancer resistance to taxol therapy by upregulation of NCOA3. *Cell Death Dis*.; 2016;7(11):e2463.
- Arijit Sen, MD, Aasma Nalwa, MD, Sandeep, MDRM, Deepali Jain, MD, and Venkateswaran K. Iyer, MD. Cytomorphology of columnar cell variant of papillary carcinoma thyroid: A case report and review of the literature. *Cytojournal*. 2014 out; 11: 27.
- Ashton-Prolla P, Seuanez HN. The Brazilian Hereditary Cancer Network: historical aspects and challenges for clinical cancer genetics in the public health care system in Brazil. *Genet Mol Biol*. 2016;165:163–5.

- Assumpção CB, Calcagno DQ, Araújo TMT, Santos SEB dos, Santos ÂKCR dos, Riggins GJ, et al. The role of piRNA and its potential clinical implications in cancer. *Epigenomics*. 2015;7(6):975–84.
- Bansal AK. Bioinformatics in microbial biotechnology--a mini review. *Microb Cell Fact*. 2005 Jun;4(ii):1–11.
- Bao G, Wang N, Li R, Xu G, Liu P, He B. MiR-508-5p Inhibits the Progression of Glioma by Targeting Glycoprotein Non-metastatic Melanoma B. *Neurochem Res* ; 2016;41(7):1684–90.
- Bennett NC, Farah CS. Next-generation sequencing in clinical oncology: Next steps towards clinical validation. *Cancers*. 2014;6(4):2296–312.
- Bhartiya D, Scaria V. Genomic variations in non-coding RNAs: Structure, function and regulation. *Genomics.*; 2016;107(2–3):59–68.
- Boria I, Boatti L, Pesole G, Mignone F. NGS-Trex: Next Generation Sequencing Transcriptome profile explorer. *BMC Bioinformatics*; 2012;14:1–8.
- Boyanton BL, Ascp MT, Luna RA, Mb C, Fasciano LR. DNA Pyrosequencing – Based Identification of Pathogenic *Candida* Species by Using the Internal Transcribed Spacer 2 Region. *Arch Pathol Lab Med*. 2008;132:667–74.
- Buda I, Bachar G, Gilam A, Modai S, Strenov Y, Feinmesser R, et al. Differential expression of microRNAs in aggressive and non-aggressive papillary thyroid carcinomas. *Head Neck Oncol*. 2012;4(2):1–12.
- Cabanillas ME, Dadu R, Hu MI, Lu C, Gunn GB, Grubbs EG, et al. Thyroid Gland Malignancies. *Hematol Oncol Clin North Am*. 2015;29(6):1123–43.
- Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet*. 2016;6736(16):1–13.
- Cairns R, Harris I, Mak T. Regulation of cancer cell metabolism. *Nat Rev Cancer*. 2011;11(2):85–95.
- Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank*. 2015;13(5):311–9.
- Cerutti JM, Riggins GJ, de Souza SJ. What can digital transcript profiling reveal about human cancers? *Brazilian J Med Biol Res*. 2003;36(8):975–85.
- Chen M, Wu J, Luo Q, Mo S, Lyu Y, Wei Y, et al. The Anticancer Properties of *Herba Epimedii* and Its Main Bioactive Components icariin and Icariside II. *Nutrients*. 2016;8(9):563.
- Chen Y-T, Kitabayashi N, Zhou XK, Fahey III TJ, Scognamiglio T. MicroRNA analysis as a potential diagnostic tool for papillary thyroid carcinoma. *Mod Pathol*. 2008;21(9):1139–46.

- Cheng J, Guo JM, Xiao BX, Miao Y, Jiang Z, Zhou H, et al. PiRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clin Chim Acta*. 2011;412(17–18):1621–5.
- Chruścik A, Lam AK yin. Clinical pathological impacts of microRNAs in papillary thyroid carcinoma: A crucial review. *Exp Mol Pathol*. 2015;99(3):393–8.
- Cong D, He M, Chen S, Liu X, Liu X, Sun H. Expression profiles of pivotal microRNAs and targets in thyroid papillary carcinoma: An analysis of The Cancer Genome Atlas. *Onco Targets Ther*. 2015;8:2271–7.
- Corbetta S, Vaira V, Guarnieri V, Scillitani A, Eller-Vainicher C, Ferrero S, et al. Differential expression of microRNAs in human parathyroid carcinomas compared with normal parathyroid tissue. *Endocr Relat Cancer*. 2010;17(1):135–46.
- Czech B, Hannon GJ. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem Sci* 2016;41(4):324–37.
- Dettmer M, Perren A, Moch H, Komminoth P, Nikiforov YE, Nikiforova MN. Comprehensive MicroRNA expression profiling identifies novel markers in follicular variant of papillary thyroid carcinoma. *Thyroid*. 2013;23(11):1383–9.
- Dluzen DF, Noren Hooten N, Zhang Y, Kim Y, Glover FE, Tajuddin SM, et al. Racial differences in microRNA and gene expression in hypertensive women. *Sci Rep* 2016;6:35815.
- Dupuis-Sandoval F, Poirier M, Scott MS. The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip Rev RNA*. 2015;6(4):381–97.
- Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet*; 2013;29(10):569–74.
- Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet*.; 2011;12(12):861–74.
- Eszlinger M, Krohn K, Hauptmann S, Dralle H, Giordano TJ, Paschke R. Perspectives for Improved and More Accurate Classification of Thyroid Epithelial Tumors. 2016;93:3286–94.
- Fagin JA, Wells SA. Biologic and Clinical Perspectives on Thyroid Cancer. *N Engl J Med*. 2016;375(11):1054–67.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–86.
- Gao L, Ma J, Mannoor K, Guarnera MA, Shetty A, Zhan M, et al. Genome-wide small nucleolar RNA expression analysis of lung cancer by next-generation deep sequencing. *Int J Cancer*. 2015;136(6):E623–9.

- Garofalo M, Quintavalle C, Romano G, Croce CM, Condorelli G. miR221/222 in cancer: their role in tumor progression and response to therapy. *Curr Mol Med.* 2012;12(1):27–33.
- Geraldo M V., Kimura ET. Integrated analysis of thyroid cancer public datasets reveals role of post-transcriptional regulation on tumor progression by targeting of immune system mediators. *PLoS One.* 2015;10(11):1–21.
- Geraldo MV, Nakaya HI, Kimura ET. Down-regulation of 14q32-encoded miRNAs and tumor suppressor role for miR-654-3p in papillary thyroid cancer. 2016;1–11.
- Gold-Smith F, Fernandez A, Bishop K. Mangiferin and Cancer: Mechanisms of Action. *Nutrients.* 2016;8(7):1–25.
- Graham ME, Hart RD, Douglas S, Makki FM, Pinto D, Butler AL, et al. Serum microRNA profiling to distinguish papillary thyroid cancer from benign thyroid masses. *J Otolaryngol Head Neck Surg. Journal of Otolaryngology - Head & Neck Surgery;* 2015;44(1):33.
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 2006;34(Database issue):D140-4.
- Gu A Di, Zhou H, Yu CH, Qu LH. A novel experimental approach for systematic identification of box H/ACA snoRNAs from eukaryotes. *Nucleic Acids Res.* 2005;33(22):20–5.
- Hagen JB. The origins of bioinformatics. *Nat Rev Genet.* 2000;1(3):231–6.
- Harshbarger J, Kratz A, Carninci P. DEIVA: a web application for interactive visual analysis of differential gene expression profiles. *BMC Genomics.* *BMC Genomics;* 2017;18(1):47. Available from: <http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-3396-5>
- Hashim A, Rizzo F, Marchese G, Ravo M, Tarallo R, Nassa G, et al. RNA sequencing identifies specific PIWI-interacting small non-coding RNA expression patterns in breast cancer. *Oncotarget* 2014;5(20):9901–10.
- Hayes J, Peruzzi PP, Lawler S. MicroRNAs in cancer: Biomarkers, functions and therapy. *Trends Mol Med.;* 2014;20(8):460–9.
- Henson J, Tischler G, Ning Z. Europe PMC Funders Group Next-generation sequencing and large genome assemblies. *Pharmacogenomics.* 2014;13(8):901–15.
- Hinrichs A, Karolchik D, Baertsch R, Barber G, Bejerano G, Clawson H. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* 2006;34:D590–8.
- Hu Y, Wang H, Chen E, Xu Z, Chen B, Lu G. Candidate microRNAs as biomarkers of thyroid carcinoma: a systematic review, meta-analysis, and experimental validation. *Cancer Med.* 2016 Jun;2602–14.

- Huang F-J, Zhou X-Y, Ye L, Fei X-C, Wang S, Wang W, et al. Follicular thyroid carcinoma but not adenoma recruits tumor-associated macrophages by releasing CCL15. *BMC Cancer*; 2015;16(1):98.
- Huang G, Hu H, Xue X, Shen S, Gao E, Guo G, et al. Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. *Clin Transl Oncol*. 2013;15(7):563–86/j.cca.2011.05.015.
- Huang Y, Bai JY, Ren HT. piRNA biogenesis and its functions. *Russ J Bioorganic Chem*. 2014;40(3).
- Hudson TJ, Anderson W, Artz A, Barker AD, Bell C, Bernabé RR, et al. International network of cancer genome projects. *Nature* 2010;464(7291):993–8.
- Jaya P, Malaysia S. Novoalign & NovoalignCS. 2017: 01-41.
- Jorjani H, Kehr S, Jedlinski DJ, Gumienny R, Hertel J, Stadler PF, et al. An updated human snoRNAome. *Nucleic Acids Res*. 2016 Jun 20;44(11):5068–82.
- Jurcevic S, Klinga-Levan K, Olsson B, Ejeskär K. Verification of microRNA expression in human endometrial adenocarcinoma. *BMC Cancer*; 2016;16(1):261.
- Kentwell J, Gundara JS, Sidhu SB. Noncoding RNAs in Endocrine Malignancy. *Oncologist*. 2014;19(5):483–91.
- Khaled H, Al Lahloubi N, Rashad N. A review on thyroid cancer during pregnancy: Multitasking is required. *J Adv Res*. 2016;7(4):565–70.
- Kichukova TM, Popov NT, Ivanov HY, Vachev TI. Circulating microRNAs as a Novel Class of Potential Diagnostic Biomarkers in Neuropsychiatric Disorders. *Folia Med (Plovdiv)*. 2015;57(3–4):159–72.
- Kishikawa T, Otsuka M, Ohno M, Yoshikawa T, Takata A, Koike K. Circulating RNAs as new biomarkers for detecting pancreatic cancer. *World J Gastroenterol*. 2015;21(28):8527–40.
- Kishore S, Gruber AR, Jedlinski DJ, Syed PTC, Jorjani H, Zavolan M. Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol*.; 2013;14(5):1–15.
- Kiss AM, Jádý BE, Bertrand E. Human Box H / ACA Pseudouridylation Guide RNA Machinery Human Box H / ACA Pseudouridylation Guide RNA Machinery †. *Mol Cell Biol*. 2004;24(13):5797–807.
- Kitahara CM, Sosa JA. The changing incidence of thyroid cancer. *Nat Rev Endocrinol*.; 2016;1–8.
- Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42(D1):68–73.

- Krishnan P, Ghosh S, Graham K, Mackey JR, Kovalchuk O, Damaraju S, et al. Piwi-interacting RNAs and PIWI genes as novel prognostic markers for Breast cancer. *Oncotarget*. 2016;5(25):37944–56.
- Langmead B, Hansen KD, Leek JT. Cloud-scale RNA-sequencing differential expression analysis with Myrna. *Genome Biol*. 2010;11(8):R83.
- Lavezzo E, Barzon L, Toppo S, Palù G. Third generation sequencing technologies applied to diagnostic microbiology: benefits and challenges in applications and data analysis. *Expert Rev Mol Diagn*. 2016 Jul;7159:1–13.
- Lee JC, Zhao JT, Clifton-Bligh RJ, Gill A, Gundara JS, Ip JC, et al. MicroRNA-222 and MicroRNA-146b are tissue and circulating biomarkers of recurrent papillary thyroid cancer. *Cancer* 2013 Dec 15;119(24):4358–65.
- Li P-F, Chen S-C, Xia T, Jiang X-M, Shao Y-F, Xiao B-X, et al. Non-coding RNAs and gastric cancer. *World J Gastroenterol*. 2014;20(18):5411–9.
- Lima CR, Geraldo M V, Fuziwara CS, Kimura ET, Santos MF. MiRNA-146b-5p upregulates migration and invasion of different Papillary Thyroid Carcinoma cells. *BMC Cancer*; 2016;16(1):108.
- Lin C, Liu A, Zhu J, Zhang X, Wu G, Ren P, et al. {miR-508} sustains phosphoinositide signalling and promotes aggressive phenotype of oesophageal squamous cell carcinoma. *Nat Commun*; 2014;5:4620.
- Liu J, Shi W, Wu C, Ju J, Jiang J. miR-181b as a key regulator of the oncogenic process and its clinical implications in cancer (Review). *Biomed reports*. 2014;2(1):7–11.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of Next-Generation Sequencing Systems. *J Biomed Biotechnol*. 2012;2012:1–11.
- LiVolsi V a. Papillary thyroid carcinoma: an update. *Mod Pathol* [Internet]. Nature Publishing Group; 2011;24 Suppl 2(S2):S1–9. Available from: <http://dx.doi.org/10.1038/modpathol.2010.129>
- Lloyd R V, Buehler D, Khanafshar E. Papillary Thyroid Carcinoma Variants. *Head Neck Pathol*. 2011;5(1):51–6.
- Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580–5.
- Ludvíková M, Kalfeřt D, Kholová I. Pathobiology of MicroRNAs and Their Emerging Role in Thyroid Fine-Needle Aspiration. *Acta Cytol*. 2015;59(6):435–444.
- Luo Y, Li X, Dong J, Sun W. microRNA-137 is downregulated in thyroid cancer and inhibits proliferation and invasion by targeting EGFR. *Tumor Biol*. 2015;(155):1–7.
- Maciel RMB, Kimura ET, Cerutti JM. Patogênese dos tumores diferenciados da tireóide (papilífero e folicular). *Arq Bras Endocrinol Metabol*. 2005;49(5):691–700.

- Makarova J, Ivanova SM, Tonevitsky a G, Grigoriev a I. New functions of small nucleolar RNAs. *Biochemistry*. 2013;78(6):638–50.
- Makohon-Moore A, Iacobuzio-Donahue CA. Pancreatic cancer biology and genetics from an evolutionary perspective. *Nat Rev Cancer*.; 2016;16(9):553–65.
- Mancikova V, Castelblanco E, Pineiro-Yanez E, Perales-Paton J, de Cubas A, Inglada-Perez L, et al. MicroRNA deep-sequencing reveals master regulators of follicular and papillary thyroid tumors. *Mod Pathol*.; 2015;28(6):1–10.
- Mannoor K, Liao J, Jiang F. Small nucleolar RNAs in cancer. 2012;100(2):130–4.
- Mannoor K, Shen J, Liao J, Liu Z, Jiang F. Small nucleolar RNA signatures of lung tumor-initiating cells. *Mol Cancer*. 2014;13(1):104.
- Mardis ER. Next-Generation Sequencing Platforms. *Annu Rev Anal Chem*. 2013 Mar;6:287–303.
- Martens-Uzunova ES, Jalava SE, Dits NF, van Leenders GJLH, Møller S, Trapman J, et al. Diagnostic and prognostic signatures from the small non-coding RNA transcriptome in prostate cancer. *Oncogene*. 2012;31(8):978–91.
- Martinez VD, Vucic EA, Thu KL, Hubaux R, Enfield KSS, Pikor LA, et al. Unique somatic and malignant expression patterns implicate PIWI-interacting RNAs in cancer-type specific biology. *Sci Rep*.; 2015 Abr;5:10423.
- Martinez-Sanchez A, Murphy CL. miR-1247 functions by targeting cartilage transcription factor SOX9. *J Biol Chem*. 2013;288(43):30802–14.
- Miranda B, Vidal SA, Mello MJG de, Lima JT de O, Rêgo JC, Pantaleão MC, et al. Cancer patients, emergencies service and provision of palliative care. *Rev Assoc Med Bras*. 2016;62(3):207–11.
- Morgan M. Bioconductor Annual Report, June 25, 2016. 2016. p. 1–9.
- Morini E, Sangiuolo F, Caporossi D, Novelli G, Amati F. Application of Next Generation Sequencing for personalized medicine for sudden cardiac death. *Front Genet*. 2015 Mar;6:55.
- Moyano M, Stefani G. piRNA involvement in genome stability and human cancer. *J Hematol Oncol*.2015;8:1–10.
- Müller S, Raulefs S, Bruns P, Afonso-Grunz F, Plötner A, Thermann R, et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdrRNAs and a piRNA in pancreatic cancer. *Mol Cancer*; 2015;14(1):94.
- The members of The Cancer Genome Atlas Research Network . Integrated Genomic Characterization of Papillary Thyroid Carcinoma. *Cell*. 2014;159(3):676–90.

- Nguyen QT, Lee EJ, Huang MG, Park YI, Khullar A, Plodkowski RA. Diagnosis and treatment of patients with thyroid cancer. *Am Heal drug benefits*. 2015;8(1):30–40.
- Nieto H, Boelaert K. Serum TSH in thyroid cancer : does it matter ? 2016 *Set*:1–33.
- Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LDR, et al. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. *JAMA Oncol*. 2016;2(8):1023–9.
- Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. *J Clin Endocrinol Metab*. 2008;93(5):1600–8.
- Nohnick WB. Non-coding RNA and the Reproductive System. 2015;51–77.
- Nwaeze O, Obidike S, Mullen D, Aftab F. Follicular variant papillary thyroid carcinoma with a twist. *Int J Surg Case Rep*. Elsevier Ltd; 2015;8:107–10.
- Oler G, Camacho CP, Hojajj FC, Michaluart P, Riggins GJ, Cerutti JM. Gene expression profiling of papillary thyroid carcinoma identifies transcripts correlated with BRAF mutational status and lymph node metastasis. *Clin Cancer Res*. 2008;14(15):4735–42.
- Organization WH. International Classification of Diseases for Oncology Third Edition. 3rd editio. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM, et al., editors. Switzerland: World Health Organization; 2013. 1-252 p.
- Paik WH, Song BJ, Kim HW, Kim HR, Hwang JH. MicroRNA-200c as a Prognostic Biomarker for Pancreatic Cancer. *Korean J Gastroenterol*. 2015;66(4):215–20.
- Parameswaran R, Shulin Hu J, Min En N, Tan W, Yuan N. Patterns of metastasis in follicular thyroid carcinoma and the difference between early and delayed presentation. *Ann R Coll Surg Engl*. 2016;1–4.
- Parikh A, Lee C, Joseph P, Marchini S, Baccharini A, Kolev V, et al. microRNA-181a has a critical role in ovarian cancer progression through the regulation of the epithelial-mesenchymal transition. *Nat Commun*. 2014;5:2977.
- Perdas E, Stawski R, Nowak D, Zubrzycka M. The Role of miRNA in Papillary Thyroid Cancer in the Context of miRNA Let-7 Family. *Int J Mol Sci*. 2016;17(6):1–7.
- Prosdocimi F, Cerqueira CG, Binneck E, Silva AF, Reis AN, Junqueira ACM, et al. *Bioinformática: Manual do Usuário*. Biotecnol Ciência e Desenvolvimentos. 2002;29:12–25.
- Rady M, Watzl C, Claus M, Khorshid O, Mahran L, Abou-Aisha K. Altered expression of miR-181a and miR-146a does not change the expression of surface NCRs in human NK cells. *Sci Rep*. 2017;7(August 2016):41381.

- Ramalho-Carvalho J, Fromm B, Henrique R, Jerónimo C. Deciphering the function of non-coding RNAs in prostate cancer. *Cancer Metastasis Rev. Cancer and Metastasis Reviews*; 2016 Maio;1–28.
- Ravo M, Cordella A, Rinaldi A, Bruno G, Alexandrova E, Saggese P, et al. Small non-coding RNA deregulation in endometrial carcinogenesis. *Oncotarget*. 2015;6(7):4677–91.
- Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet*. 2000;16(1):276–7.
- Ricarte Filho JCM, Kimura ET. MicroRNAs: nova classe de reguladores gênicos envolvidos na função endócrina e câncer. *Arq Bras Endocrinol Metabol*. 2006;50(6):1102–7.
- Riesco-Eizaguirre G, Santisteban P. ENDOCRINE TUMOURS: Advances in the molecular pathogenesis of thyroid cancer: lessons from the cancer genome. *Eur J Endocrinol*. 2016;175(5):R203–17.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009;26(1):139–40.
- Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*. 2007;23(21):2881–7.
- Ronchetti D, Mosca L, Cutrona G, Tuana G, Gentile M, Fabris S, et al. Small nucleolar RNAs as new biomarkers in chronic lymphocytic leukemia. *BMC Med Genomics*. *BMC Medical Genomics*; 2013;6(1):27.
- Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2015 update Jan. *Nucleic Acids Res*. 2015;43(D1):D670–81.
- Rossing M, Borup R, Henao R, Winther O, Vikesaa J, Niazi O, et al. Down-regulation of microRNAs controlling tumourigenic factors in follicular thyroid carcinoma. *J Mol Endocrinol*. 2012;48(1):11–23.
- Rung J, Brazma A. Reuse of public genome-wide gene expression data. *Nat Rev Genet*. Nature Publishing Group; 2012;(December):1–11.
- Sai lakshmi S, Agrawal S. piRNABank: A web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res*. 2008;36(Supl. 1):173–7.
- Sanger F, Nicklen S, Coulson a R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74(12):5463–7.
- Scaravilli M, Porkka KP, Brofeldt A, Annala M, Tammela TLJ, Jenster GW, et al. MiR-1247-5p is overexpressed in castration resistant prostate cancer and targets MYCBP2. *Prostate*. 2015;75(8):798–805.

- Schneider EB, Umbricht CB, Witwer KW, Zeiger MA. MicroRNA Expression and Association with Clinicopathologic Features in Papillary Thyroid Cancer: A Systematic Review. *Molecules*. 2015;25(12):1322–9.
- Shang Y, Feng B, Zhou L, Ren G, Zhang Z, Fan X, et al. The miR27b-CCNG1-P53-miR-508-5p axis regulates multidrug resistance of gastric cancer. *Oncotarget*. 2015;5(1):538–49.
- Siprashvili Z, Webster DE, Johnston D, Shenoy RM, Ungewickell AJ, Bhaduri A, et al. The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat Genet*. 2015;48(1):53–8.
- Song T, Liang Y, Cao Z, Du W, Li Y. Computational Analysis of Specific MicroRNA Biomarkers for Noninvasive Early Cancer Detection. 2017;2017.
- Stepanov GA, Filippova JA, Komissarov AB, Kuligina E V., Richter VA, Semenov D V., et al. Regulatory role of Small nucleolar RNAs in human diseases. *Biomed Res Int*. Hindawi Publishing Corporation; 2015;2015:1–10.
- Stokowy T, Gawel D, Wojtas B. Differences in miRNA and mRNA Profile of Papillary Thyroid Cancer Variants. *Int J Endocrinol*. Hindawi Publishing Corporation; 2016;2016:1–10.
- Stoupa A, Kariyawasam D, Carré A, Polak M. Update of Thyroid Developmental Genes. *Endocrinol Metab Clin North Am*. 2016;1–12.
- Sun L, Chen Y, Che G. Increased Plasma miRNA-30a as a Biomarker for Non-Small Cell Lung Cancer. 2016;647–55.
- Suresh R, Sethi S, Ali S, Giorgadze T, Sarkar FH. Differential Expression of MicroRNAs in Papillary Thyroid Carcinoma and Their Role in Racial Disparity. *J Cancer Sci Ther*. 2015;7(5):145–54.
- Swierniak M, Wojcicka A, Czetwertynska M, Stachlewska E, Maciag M, Wiechno W, et al. In-depth characterization of the MicroRNA transcriptome in normal thyroid and papillary thyroid carcinoma. *J Clin Endocrinol Metab*. 2013;98(8):1401–9.
- Thorenoor N, Slaby O. Small nucleolar RNAs functioning and potential roles in cancer. *Tumor Biology*. 2014. p. 41–53.
- Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Wspolczesna Onkol*. 2015;1A:A68–77.
- Tóth KF, Pezic D, Stuwe E, Webster A. The piRNA Pathway Guards the Germline Genome Against Transposable Elements. 2016. 1-36 p.
- Valleron W, Laprevotte E, Gautier E-F, Quelen C, Demur C, Delabesse E, et al. Specific small nucleolar RNA expression profiles in acute leukemia. *Leukemia*. 2012;26(9):2052–60.

- Vilela APP, Aguiar ERGR, Ferreira F V, Ribeiro LS, Olmo RP, Marques JT. Small Non-Coding RNAs as Biomarkers. *Recent Pat Biomark.* 2012;2(2):119–30.
- Vriens MR, Weng J, Suh I, Huynh N, Guerrero MA, Shen WT, et al. MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer. *Cancer.* 2012 Jul 1;118(13):3426–32.
- Wang B, Yang J, Xiao B, Li J, Liu D, Yang Y, et al. MicroRNA-20b (miR-20b) Promotes the Proliferation, Migration, Invasion, and Tumorigenicity in Esophageal Cancer Cells via the Regulation of Phosphatase and Tensin Homologue Expression. *PLoS One.* 2016;11(10):e0164105.
- Wang J, Li L-C. Small RNA and its application in andrology and urology. *Transl Androl Urol.* 2012;1(1):33–43.
- Wang J, Song Y-X, Ma B, Wang J-J, Sun J-X, Chen X-W, et al. Regulatory Roles of Non-Coding RNAs in Colorectal Cancer. *Int J Mol Sci.* 2015;16(8):19886–919.
- Weber F, Teresi RE, Broelsch CE, Frilling A, Eng C. A limited set of human MicroRNA Is deregulated in follicular thyroid carcinoma. *J Clin Endocrinol Metab.* 2006;91(9):3584–91.
- Williams D. Thyroid Growth and Cancer. *Eur Thyroid J.* 2015;4(3):164–73.
- Willis R. Targeted Cancer Therapy: Vital Oncogenes and a New Molecular Genetic Paradigm for Cancer Initiation Progression and Treatment. *Int J Mol Sci.* 2016;17(9):1–23.
- Witczak J, Taylor P, Chai J, Amphlett B, Soukias JM, Das G, et al. Predicting malignancy in thyroid nodules: feasibility of a predictive model integrating clinical, biochemical, and ultrasound characteristics. *Thyroid Res. Thyroid Research;* 2016;9:1–7.
- Wó Jcicka A, Kolanowska M, Jaz ˙dz ˙ewski K. MicroRNA in diagnostics and therapy of thyroid cancer. *Eur J Endocrinol.* 2016;174:89–98.
- Xue T-M, Tao L, Zhang M, Zhang J, Liu X, Chen G-F, et al. Clinicopathological Significance of MicroRNA-20b Expression in Hepatocellular Carcinoma and Regulation of HIF-1 α and VEGF Effect on Cell Biological Behaviour. *Dis Markers.* 2015;2015:325176.
- Yamashita AS, Geraldo MV, Fuziwara CS, Kulcsar MAV, Friguglietti CUM, da Costa RB, et al. Notch pathway is activated by MAPK signaling and influences papillary thyroid cancer proliferation. *Transl Oncol. Neoplasia Press Inc.;* 2013;6(2):197–205.
- Yang D, Sun L, Li Z, Gao P. Noncoding RNAs in Regulation of Cancer Metabolic Reprogramming Dongdong. 2016;927:317–36.
- Yip L, Sosa JA. Molecular-Directed Treatment of Differentiated Thyroid Cancer. *JAMA Surg.* 2016;15213(7):663–70.

- Yoruker EE, Terzioglu D, Teksoz S, Uslu FE, Gezer U, Dalay N. MicroRNA expression profiles in papillary thyroid carcinoma, benign thyroid nodules and healthy controls. *J Cancer*. 2016;7(7):803–9.
- Yu S, Liu Y, Wang J, Guo Z, Zhang Q, Yu F, et al. Circulating microRNA profiles as potential biomarkers for diagnosis of papillary thyroid carcinoma. *J Clin Endocrinol Metab*. 2012;97(6):2084–92.
- Yu X-M, Schneider DF, Levenson G, Chen H, Sippel RS. Follicular variant of papillary thyroid carcinoma is a unique clinical entity: a population-based study of 10,740 cases. *Thyroid*. 2013;23(10):1263–8.
- Zhang J, Liu Y, Liu Z, Wang XM, Yin DT, Zheng LL, et al. Differential expression profiling and functional analysis of microRNAs through stage I-III papillary thyroid carcinoma. *Int J Med Sci*. 2013;10(5):585–92.
- Zhang R, Hardin H, Chen J, Guo Z, Lloyd R V. Non-Coding RNAs in Thyroid Cancer. *Endocr Pathol*. 2016;27(1):12–20.
- Zhao F, Lv J, Gan H, Li Y, Wang R, Zhang H, et al. MiRNA profile of osteosarcoma with CD117 and stro-1 expression: MiR-1247 functions as an onco-miRNA by targeting MAP3K9. *Int J Clin Exp Pathol*. 2015;8(2):1451–8.
- Zhou X, Lindsay H, Robinson MD. Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res*. 2014;42(11):1–10.
- Zhou X, Wang X, Huang Z, Wang J, Zhu W, Shu Y, et al. Prognostic value of miR-21 in various cancers: An updating meta-analysis. *PLoS One*. 2014;9(7).
- Zhu H, Fang J, Zhang J, Zhao Z, Liu L, Wang J, et al. MiR-182 targets CHL1 and controls tumor growth and invasion in papillary thyroid carcinoma. *Biochem Biophys Res Commun*. Elsevier Inc.; 2014;450(1):857–62.

9. ANEXOS

9.1. ANEXO 1

Tabela 9.1: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero.

	Paciente	N/T	Amostra (srRNA-BAM ID)	TCGA- Paciente ID	ID amostra -Tumor
Carcinoma papilífero, variante Células Altas	P1	T	4ba277d8-ff62-47fb- a08d-17476ffbb787	TCGA- BJ-A28X	TCGA- BJ-A28X- 01
		N	b36f917b-3628-476d- a388-4df0b1d93da6		
	P2	T	6109aba2-9378-41b7- 97aa-28fdc33f96fb	TCGA- ET- A3DW	TCGA- ET- A3DW-01
		N	1a1881b7-d26d-4669- ab65-2b7a1bd6eda7		
	P3	T	33db5d5f-613e-4cbe- 8696-6abee8f27317	TCGA- BJ-A290	TCGA- BJ-A290- 01
		N	0c3c1ae7-e856-4446- 93d3-75c0c2ed28ad		
P1	T	32d79ac3-b093-4ba0- b141-3783298a3cfd	TCGA- KS-A41I	TCGA- KS-A41I- 01	
	N	bee97b09-9d73-4350- 81ec-1f500d56356c			
P2	T	0d1da331-db1f-4cf0- aff1-6e66b74fd2e2	TCGA- EM- A1YC	TCGA- EM- A1YC-01	
	N	c8473ec0-4c6f-44c9- b3a1-7a6f9d199ca1			
P3	T	b55076da-0863-4adf- adc0-ae7ce6dbdb0c	TCGA- DO- A1JZ	TCGA- DO- A1JZ-01	
	N	e64d1ba4-ecc2-46de- b162-da6f7d1bd287			
P4	T	1fb02162-e2a1-418c- b659-3a8b6356c15e	TCGA- KS-A41L	TCGA- KS-A41L- 01	
	N	3b782701-2c84-4379- b3ef-922c472d624c			
P5	T	0d4f39d5-14ce-4b7a- b645-1bf8abe62c49	TCGA- EM- A1CW	TCGA- EM- A1CW-01	
	N	2649e98b-21a9-4d56- 9688-f72380fac0ef			
P6	T	58c254cb-17d0-4f79- acbc-42bd2399b802	TCGA- EM- A3ST	TCGA- EM- A3ST-01	
	N	bed4478d-446c-4699- 84fb-4431355f75c5			

Continuação tabela: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero

Carcinoma papilífero variante folicular	P7	T	a471aad4-3fee-460c-8eac-2ccb299bb6e3	TCGA-ET-A2N5	TCGA-ET-A2N5-01
		N	f3664459-44de-4486-a89c-220a35a7cc52		
	P1	T	ca896aac-defa-47bf-90be-5539d9d66650	TCGA-ET-A2MX	TCGA-ET-A2MX-01
		N	3107e77c-6d34-4d6f-90f0-a4bbc9d7017a		
	P2	T	e8fd6ec2-011f-40ef-a3af-a6b6e17f4914	TCGA-BJ-A2N8	TCGA-BJ-A2N8-01
		N	d911e617-8cf1-479a-bd22-6e5abe7f830e		
	P3	T	37ebbca7-7f2d-412f-ac53-f0981112fe9a	TCGA-BJ-A3PU	TCGA-BJ-A3PU-01
		N	4ef60551-48ac-4ca5-84d4-b9bc11a0840c		
	P4	T	4481b820-6c4a-4770-80df-bd22c8c22ef8	TCGA-EL-A3MY	TCGA-EL-A3MY-01
		N	cd1f62ba-de10-4a09-b2aa-bd593e626ee4		
P5	T	92e47377-f513-4a6b-8159-fb6e4776eec0	TCGA-EM-A1CU	TCGA-EM-A1CU-01	
	N	15e44cc5-b72b-4d32-b574-e85d4360025d			
Carcinoma papilífero	P6	T	4c3e2109-1ae4-4024-9edd-e83749c6bccc	TCGA-EL-A3ZP	TCGA-EL-A3ZP-01
		N	85217e81-a1df-4e7d-a10e-592dbef7d5e7		
	P7	T	f2420c29-a62b-4f9e-af50-b4a0906b2610	TCGA-BJ-A2NA	TCGA-BJ-A2NA-01
		N	9c9d050c-0158-49b0-b350-d78cf2a6a319		
	P8	T	869bff71-f61a-4b53-bc85-70ccacece144	TCGA-BJ-A28R	TCGA-BJ-A28R-01
		N	b7058800-04d2-4e15-a01c-c4a2b73decc6		
	P9	T	e84d7507-7111-4402-94ea-10f9dd2470f4	TCGA-KS-A41J	TCGA-KS-A41J-01
		N	dc45e31c-8088-401d-8bb9-64cf13e70c70		
	P10	T	677288e7-ca0e-4f6a-a534-9c343d205251	TCGA-EL-A3T2	TCGA-EL-A3T2-01
		N	54750463-0383-4f85-a545-ee3ab29a70d5		

Continuação tabela: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero

Carcinoma papilífero	P11	T	2fbb897b-420b-4c34-aa59-ec217ad8c8de	TCGA-GE-A2C6	TCGA-GE-A2C6-01
		N	b7188a87-1368-4d71-adf0-d2ac7518de72		
	P12	T	a148fe16-b5be-4541-8c92-8438404c5832	TCGA-EL-A3ZL	TCGA-EL-A3ZL-01
		N	4966830e-3583-4612-8989-8fe72f5907ec		
	P13	T	7f99c998-5045-4f35-ba27-0242ad7afa02	TCGA-EL-A3ZH	TCGA-EL-A3ZH-01
		N	3c5a626f-4f44-4de0-8925-13f89d738335		
	P14	T	f3960ba6-c1cc-4ce8-a0ca-c6a8869cc455	TCGA-BJ-A2N7	TCGA-BJ-A2N7-01
		N	7bf8f55d-b9a5-417a-885e-2dc1f50e8409		
	P15	T	20fff6bf-5929-4c4c-844a-131f98f10325	TCGA-EM-A1CT	TCGA-EM-A1CT-01
		N	314d9298-5a27-429c-8e0d-063847fca824		
	P16	T	7e8ff428-5397-4855-9d52-38660a3f1402	TCGA-EL-A3ZT	TCGA-EL-A3ZT-01
		N	744682bd-c9c1-4861-8f50-3ef6a4f093b4		
	P17	T	4246e2c1-8cc7-4ca7-8630-a96888a379ca	TCGA-EL-A3TA	TCGA-EL-A3TA-01
		N	a4ff2bb9-b4d7-4ccc-a58a-5afaca9cdd4d		
	P18	T	3043ad32-a8e1-4d1e-b6d7-1eddf1e9823	TCGA-BJ-A3PR	TCGA-BJ-A3PR-01
		N	c26f0cc2-245e-49f4-89e7-853fb43e3795		
	P19	T	efffb9cf-789d-4641-8c48-6b7b24ef768c	TCGA-EL-A3N2	TCGA-EL-A3N2-01
		N	3971c5cc-610f-489e-a874-7b681e2c6310		
	P20	T	49b132dc-c069-4552-b574-e8eb8824c2f2	TCGA-EL-A3MX	TCGA-EL-A3MX-01
		N	002ee189-d4ba-409f-a2a0-cd5df639eca2		
	P21	T	3f501809-6a79-40ff-a9c3-f9c0dcc8c22f	TCGA-EM-A1CS	TCGA-EM-A1CS-01
N		8991409e-49e7-4ec3-8fcc-17091353fb96			

Continuação tabela: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero

Carcinoma papilífero	P22	T	15508fbe-ddf1-4c42-bfed-16068edce433	TCGA-EL-A3N3	TCGA-EL-A3N3-01
		N	cbfcc9c8-c134-4d54-a813-c66197e2f08a		
	P23	T	f3ffa0fc-45b2-42d8-ae1e-212f71c22795	TCGA-EL-A3T3	TCGA-EL-A3T3-01
		N	e8b907b7-c1c1-414b-82f1-c2f9562d69e1		
	P24	T	6860e703-dfef-481e-af81-6ff8adebca7c	TCGA-EL-A3T7	TCGA-EL-A3T7-01
		N	21d6e3e8-f1af-4caa-b207-e803ce228537		
	P25	T	27ee3861-eed1-48cc-9756-97c70ee5ccb3	TCGA-BJ-A2N9	TCGA-BJ-A2N9-01
		N	121f58c8-3113-4766-b412-ffd6a9fa1122		
	P26	T	792e5c4a-ef86-4183-a74d-ab0e4a287577	TCGA-EL-A3H7	TCGA-EL-A3H7-01
		N	874cb9d7-166d-4db7-b20b-21e0248c6cc8		
	P27	T	dd6b43d5-8007-434c-b8a5-46b1b7b6a7bf	TCGA-EL-A3GZ	TCGA-EL-A3GZ-01
		N	5363f479-0399-4c36-827d-2843b9368bcd		
	P28	T	7b94c2b5-8d66-4143-9faf-58e0fc97df28	TCGA-EL-A3ZS	TCGA-EL-A3ZS-01
		N	09b18290-7b32-4994-93ec-2b0e2ccdb03a		
	P29	T	d4dcabf4-705e-4422-9fb0-61e9be41aa68	TCGA-ET-A3DP	TCGA-ET-A3DP-01
		N	a2d7829a-f991-4b1b-8253-7cc8aef6f187		
P30	T	2c33e2bb-9590-480e-bfbb-92c171745db2	TCGA-EL-A3ZM	TCGA-EL-A3ZM-01	
	N	74f6c794-b225-4d9c-b2a6-40cd83d64e10			
P31	T	d8499ef0-43db-42c3-80d2-0ff317ea5e10	TCGA-H2-A3RI	TCGA-H2-A3RI-01	
	N	68aa16f6-a3af-4cc4-941b-ad0f6451df40			
P32	T	e00b7f52-c0b7-40d4-a7b2-77666e3243cf	TCGA-EL-A3ZK	TCGA-EL-A3ZK-01	
	N	e0ba5366-027f-4ed7-b1eb-8616c0c26e92			

Continuação tabela: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero

Carcinoma papilífero	P33	T	c374ae59-441d-44a2-aca8-0973765da6e6	TCGA-H2-A2K9	TCGA-H2-A2K9-01
		N	ebd1eaba-8352-4352-b679-2b6582fd6bbb		
	P34	T	ea76f113-ffb2-43e0-893b-d7afe976a5cd	TCGA-EL-A3T8	TCGA-EL-A3T8-01
		N	4e407151-a509-47c2-a1b6-199206c6cdb0		
	P35	T	31e64c16-0e6a-4792-b2b3-1c11def75bd8	TCGA-EL-A3TB	TCGA-EL-A3TB-01
		N	d12de476-71d3-45d2-ba94-d1d57f551328		
	P36	T	c9314ad1-7db3-4b3c-a62b-53a913994ff3	TCGA-EL-A3T0	TCGA-EL-A3T0-01
		N	ce370eef-b3dd-42dd-b5e0-5dd21280cf68		
	P37	T	ca10d6e7-2bf4-4af4-8eb8-9df67b2312fa	TCGA-EL-A3H2	TCGA-EL-A3H2-01
		N	b1d72607-0548-4180-94bc-e200759a9b6a		
	P38	T	865deb32-d45e-41f9-a30f-a8cf4a0ebfeb	TCGA-BJ-A28W	TCGA-BJ-A28W-01
		N	929a4b40-5849-45e1-9d89-f0c9f0119713		
	P39	T	6989eea3-1655-47fd-a7e1-69d99355888b	TCGA-E8-A2JQ	TCGA-E8-A2JQ-01
		N	5525ad18-22ea-4678-847a-c328be669230		
	P40	T	e7a8c272-8857-4e6f-b977-6e1fe9630f71	TCGA-FY-A3TY	TCGA-FY-A3TY-01
		N	4de425d9-c703-4272-ac98-a8353206c305		
P41	T	91d115e3-d55c-4c4c-b6c7-ade991a92cd7	TCGA-EL-A3ZR	TCGA-EL-A3ZR-01	
	N	cb7ab1eb-b5a8-4c2f-853d-c34697222746			
P42	T	d3e98adc-e4db-4782-94ca-14256abe06b8	TCGA-EL-A3ZO	TCGA-EL-A3ZO-01	
	N	7d1b9277-99bc-472e-ae09-e364ba799da0			
P43	T	88ef0eb7-20f7-4831-97bd-6af885f6e8cc	TCGA-EM-A1CV	TCGA-EM-A1CV-01	
	N	482324f9-bf3a-4027-a780-631708a64643			

Continuação tabela: Amostras pareadas (T: tumoral e N: normal adjacente ao tumor) de cada paciente utilizadas no presente estudo disponibilizadas no TCGA (<https://cancergenome.nih.gov/>), divididas nos subtipos tumorais Carcinoma papilífero variante Células Altas, Carcinoma papilífero variante folicular, Carcinoma papilífero

Carcinoma papilífero	P44	T	ad9ea32c-07f6-47e5-a745-247eabe6c2b6	TCGA-EL-A3T6	TCGA-EL-A3T6-01
		N	eadbc640-edc0-44e3-8127-ccc08077dc17		
	P45	T	f6525d5b-6424-40da-9bdc-545477ab76bd	TCGA-EL-A3H1	TCGA-EL-A3H1-01
		N	112c19db-b055-402f-8ff9-6f4f560bf9bb		
	P46	T	ec9fd10c-f564-47cc-bce6-a25b862c74d1	TCGA-EL-A3T1	TCGA-EL-A3T1-01
		N	46a9bc4a-2a9a-480b-a680-a2df00d3f7b0		
	P47	T	c218632e-33e1-4cc8-bb4f-0df7ffe2eed5	TCGA-EL-A3ZQ	TCGA-EL-A3ZQ-01
		N	a65e8fbd-4b6b-4358-89a0-851c1676ac40		
	P48	T	1b775de8-99d3-47e4-b9cd-a79865bd7ec7	TCGA-EL-A3MW	TCGA-EL-A3MW-01
		N	c083c5f5-fe67-4f64-8ce2-56825fc3cbfb		
P49	T	c703e014-15ad-43a0-b3d8-e42b9134891d	TCGA-EL-A3ZG	TCGA-EL-A3ZG-01	
	N	2c460e9e-3936-4d4c-a941-b19e06a98aee			

9.2. Anexo 2

Figura 9.1: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero variante Células Altas. Foram considerados como diferencialmente expressos os sncRNAs com FDR menor que 0,05 e logFC maior ou igual a 1,5 e menor ou igual a -1,5. Pontos vermelhos representam sncRNAs com valor de FDR menor do que 0,05, linhas azuis representam o logFC.

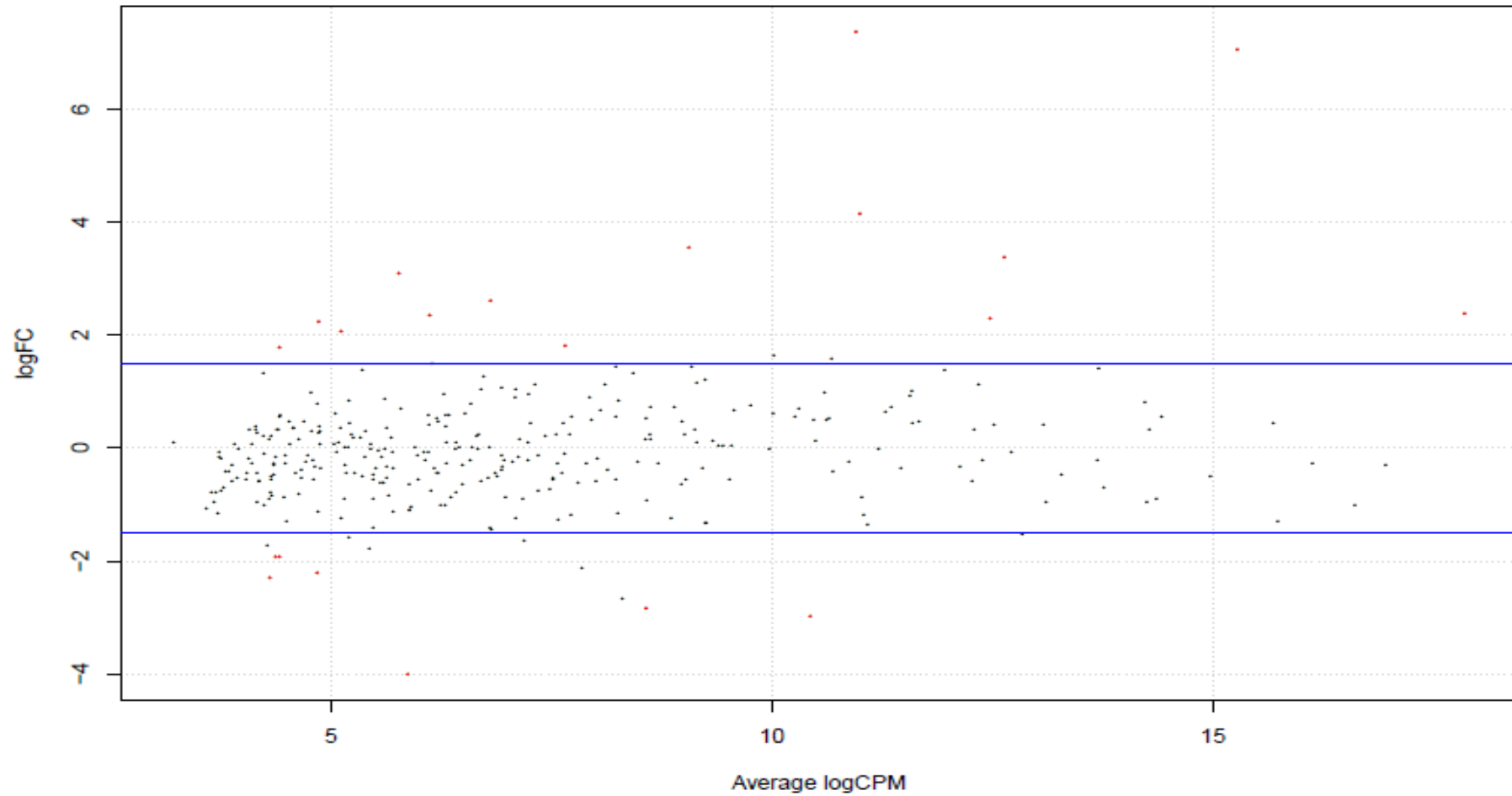


Figura 9.2: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero variante folicular. Foram considerados como diferencialmente expressos os sncRNAs com FDR menor que 0,05 e logFC maior ou igual a 1,5 e menor ou igual a -1,5. Pontos vermelhos representam sncRNAs com valor de FDR menor do que 0,05, linhas azuis representam o logFC.

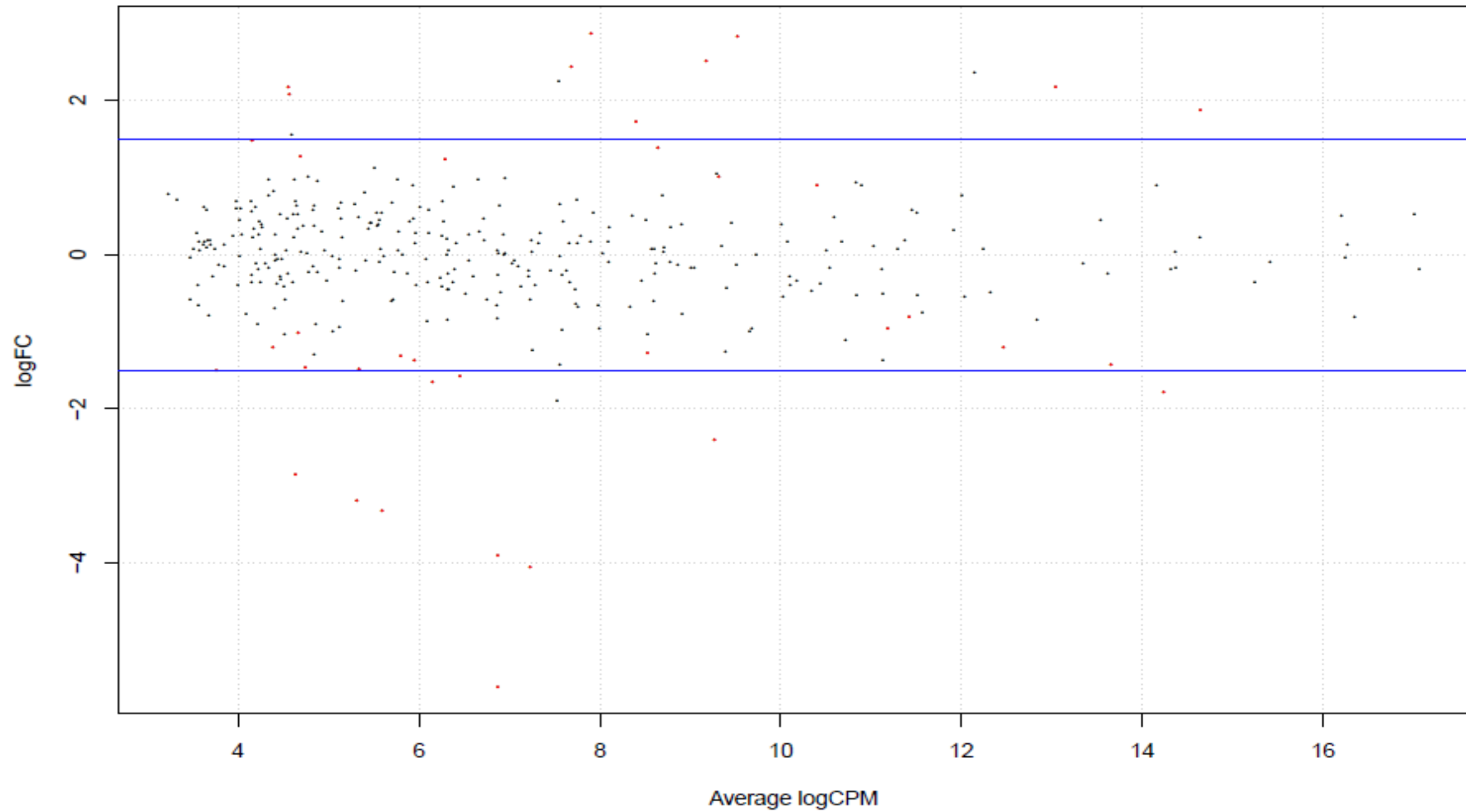
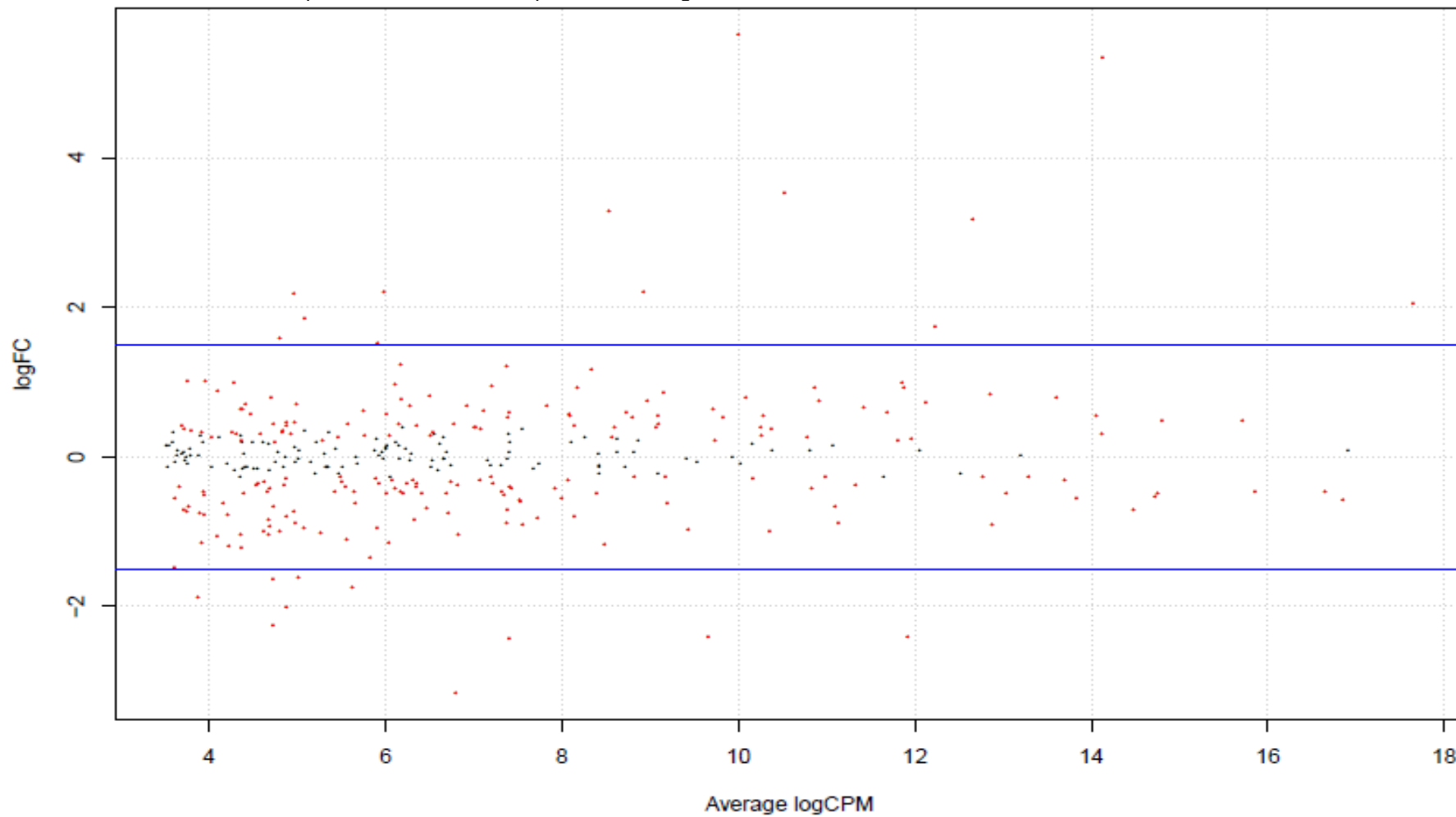


Figura 9.3: Gráfico demonstrando os sncRNAs encontrados no subtipo Carcinoma papilífero. Foram considerados como diferencialmente expressos os sncRNAs com FDR menor que 0,05 e logFC maior ou igual a 1,5 e menor ou igual a -1,5. Pontos vermelhos representam sncRNAs com valor de FDR menor do que 0,05, linhas azuis representam o logFC.



9.3. ANEXO 3

Script em linguagem R utilizando o pacote EdgeR (versão 3.3) (ROBINSON *et al.*, 2009) do projeto Bioconductor desenvolvido por Natasha Jorge para normalizar e analisar a expressão diferencial dos genes.

```
design.table = read.table("tabela_de_design_NT.txt",header=T)
archs = as.vector(design.table[,1])
cond1 = as.vector(design.table[,2])
cond2 = as.vector(design.table[,3])

#Getting counts. table
counts.table = read.table("tabela_de_contagem.tsv",header=T)
counts.table[is.na(counts.table)]<-0
rownames(counts.table)<-counts.table$gene
counts.table$gene <- NULL
counts.table = counts.table[rowSums(counts.table > 50) >= 49,] #Alterar para o
tamanho da menor amostra
counts.table = counts.table[rowSums(counts.table == 0) <=2,]

#Getting group condition;
group1 = vector()
group2 = vector()
count.order = as.vector(colnames(counts.table))
for (d in 1:length(count.order)){
  for (e in 1:length(archs)){
    if (count.order[d] == archs[e]){
      group1[d] = cond1[e]
      group2[d] = cond2[e]
    }
  }
}
count.order = as.factor(count.order)
group1 = as.factor(group1)
group2 = as.factor(group2)

#Start differential expression analysis
library(edgeR)
y=DGEList(counts=counts.table)
y = calcNormFactors(y)
data.frame(Sample=colnames(y),group1,group2)
design = model.matrix(~group1+group2)
rownames(design) = colnames(y)
y = estimateDisp(y, design)
fit = glmFit(y, design)

fit$design
lrt = glmLRT(fit, coef=50) #Alterar para a comparação coluna tumoral
```

```
de.top = topTags(lrt,n=length(lrt))
de.cpm = cpm(y)[rownames(de.top),]

#Writting outputs;
de.top$table$gene = rownames(de.top)
de.cpm = as.data.frame(de.cpm)
de.cpm$gene = rownames(de.cpm)
de.all = merge(de.cpm,de.top,by="gene",all=T)
de.all2 = de.all[order(de.all$PValue, de.all$FDR),]
write.table(de.all2,file="tabela_de_saida_NT.csv")
```