



**UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
FACULDADE DE MEDICINA**

**Naiara da Costa Cinegaglia**

**microRNoma do carcinoma de pulmão  
de células não pequenas**

Dissertação apresentada à  
Faculdade de Medicina,  
Universidade Estadual Paulista  
“Júlio de Mesquita Filho”, Câmpus  
de Botucatu, para obtenção do título  
de Mestre em Bases Gerais da  
Cirurgia.

Orientadora: Profa. Dra. Patricia Pintor dos Reis  
Coorientador: Prof. Dr. Igor Jurisica

**Botucatu  
2015**

Naiara da Costa Cinegaglia

microRNoma do carcinoma de pulmão de células não  
pequenas

Dissertação apresentada à  
Faculdade de Medicina,  
Universidade Estadual Paulista  
“Júlio de Mesquita Filho”, Câmpus  
de Botucatu, para a obtenção do  
título de Mestre em Bases Gerais da  
Cirurgia.

Orientadora: Profa. Dra. Patricia Pintor dos Reis  
Coorientador: Prof. Dr. Igor Jurisica

Botucatu  
2015

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP  
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Cinegaglia, Naiara da Costa.

MicroRNoma do carcinoma de pulmão de células não  
pequenas / Naiara da Costa Cinegaglia. - Botucatu, 2015

Dissertação (mestrado) - Universidade Estadual Paulista  
"Júlio de Mesquita Filho", Faculdade de Medicina de  
Botucatu

Orientador: Patricia Pintor dos Reis

Coorientador: Igor Jurisica

Capes: 20200005

1. Adenocarcinoma. 2. Pulmões - Câncer. 3. Marcadores  
biológicos. 4. Expressão gênica. 5. Reação em cadeia de  
polimerase.

Palavras-chave: adenocarcinoma pulmonar; biomarcadores;  
microRNAs.

À **Deus** pela Luz Divina que nos abençoa em todas as circunstâncias  
e ajuda-nos o ideal na estrada imensa da vida.  
Em especial à minha família,  
Aos meus pais **Carlos e Silvia**,  
Meus exemplos de força, coragem e determinação,  
À eles serei eternamente grata por toda dedicação, confiança, incentivo e  
confiança  
Aos meus queridos irmãos **Heler e Caio**,  
Pela confiança, força, companheirismo, respeito e apoio sempre.

**dedico este trabalho.**

## AGRADECIMENTOS

À Faculdade de Medicina, Universidade Estadual Júlio de Mesquita Filho, Campus de Botucatu, em especial ao Programa de Pós-Graduação Bases Gerais da Cirurgia e ao Laboratório Neogene pela oportunidade de desenvolver o mestrado e por todo apoio e conhecimento adquiridos durante o desenvolvimento do projeto.

À Professora Dra. Patrícia Reis, pelas oportunidades, dedicação, respeito, disponibilidade, compreensão e paciência frente as minhas limitações e, sobretudo, pela amizade, confiança e grande apoio em todos os momentos.

Ao Dr. Igor Jurisica, pela orientação, disponibilidade, competência, sugestões e conhecimento.

o Professor Dr. Rogério Antonio de Oliveira, pela disponibilidade e colaboração nas análises estatísticas.

Ao Dr. Fábio Severino, pelo apoio e colaboração nas análises de bioinformática.

À Professora Dra. Silvia Regina Rogatto, pela oportunidade e disponibilização do laboratório.

À Dra. Sandra Drigo Linde, pela disponibilidade, conhecimento, colaboração e amizade.

Aos meus companheiros de laboratório, Natalia Bertoni, Tainara Felix, Jovita Ariede, Maísa Pinheiro e Márcio Carvalho, pela amizade, apoio, disponibilidade, paciência, respeito e companheirismo.

Ao meu namorado José Eduardo, pelo apoio, paciência, amor, carinho, confiança e principalmente por todo incentivo, paciência e companheirismo nessa fase final.

À Renata Bueno, pela sincera amizade, conselhos, respeito e um carisma contagiante.

À querida Eliete Trombini, pela amizade, positividade, disponibilidade, compreensão, conselhos e pelo grande apoio e incentivo durante todo esse período.

À minha amiga Bianca Picado, pelo apoio, carinho, conselhos, sinceridade, companheirismo e por todos os bons momentos que vivenciamos e compartilhamos desde a nossa graduação.

*No pensar, clareza  
No sentir, cordialidade,  
No querer, prudência:  
Almejando-as  
Posso então esperar,  
Que eu corretamente  
Possa encontrar-me  
Nas trilhas da vida  
Diante de corações humanos  
No âmbito do dever.  
Pois clareza  
Provém da luz da alma,  
E cordialidade  
Contém o calor espiritual,  
Prudência  
Intensifica a força da vida.  
E tudo isso,  
Em confiança em Deus anseia,  
Em caminhos humanos conduz  
A passos bons e seguros na vida.*

**Rudolf Steiner**

## RESUMO

O câncer de pulmão é a causa mais frequente de óbito por câncer, com 1,6 milhões de óbitos mundialmente, a cada ano. Apesar dos avanços nas estratégias de tratamento, o prognóstico dos pacientes com carcinoma pulmonar ainda é pobre. O desenvolvimento de drogas com alvos moleculares têm levado a um aumento, embora modesto, na sobrevivência de pacientes com carcinomas pulmonares de células não pequenas (NSCLC), especialmente do subtipo histológico adenocarcinoma (AD). A identificação de vias moleculares como alvos terapêuticos é uma área promissora de investigação. Considerando que os miRNAs regulam a expressão gênica, os miRNAs têm o potencial de modular vias moleculares associadas à gênese e progressão do câncer. Investigamos alterações na expressão e ou mutações em sequências de miRNAs em AD pulmonar, utilizando plataformas de análise de sequenciamento de alto desempenho (Illumina) (N=17 AD e N=7 tecidos pulmonares histologicamente normais) e ensaios da PCR quantitativa em tempo real (RQ-PCR) (N=22 AD e N=9 N). A análise de miRNA-Seq identificou variações em nucleotídeos únicos (SNVs) bem como alterações na expressão de miRNAs em AD pulmonar. Três miRNAs (miR-328, miR-574 e miR-99a) foram validados, com expressão aumentada ( $p < 0,05$ ) em AD comparado com N. Adicionalmente, realizamos uma meta-análise de dados da literatura em NSCLC, com o objetivo principal de comparar os dados da literatura com os nossos dados de miRNA-Seq em AD. Os miRNAs identificados na meta-análise foram utilizados em análises de predição de genes-alvo potencialmente regulados por miRNAs. Redes de interação proteica e vias moleculares relevantes em AD pulmonar foram identificadas. mRNAs-alvo identificados nas redes de interação foram validados por RQ-PCR em 57 amostras de 38 pacientes com NSCLC (N=30 NSCLC e N=27 N). Os resultados mostraram que 3 genes-alvo de miRNAs (*PIK3R1*, *NOTCH1*, *KLF4*), estavam frequentemente desregulados, apresentando baixo nível de expressão nas amostras de NSCLC comparado com tecido normal. A expressão diminuída do gene *KLF4* foi estatisticamente correlacionada com o grau histológico avançado do tumor ( $p = 0,0371$ ). O gene *KLF4* estava diferencialmente expresso ( $p = 0,0338$ ) entre os subtipos histológicos AD e SCC. Os genes *KLF4* e *NOTCH1* apresentaram expressão diminuída nas amostras de AD quando comparado com tecido normal ( $p = 0,0022$  e  $p = 0,0041$ ). O gene *SATB2* estava com expressão aumentada em SCC comparado com tecido normal ( $p = 0,0216$ ). Estudos como esse podem levar à identificação de vias moleculares importantes na tumorigênese e progressão de NSCLC. O conhecimento dessas vias moleculares alteradas pode ser útil para o desenvolvimento futuro de estratégias de tratamento mais precisas para pacientes com AD pulmonar.

## ÍNDICE

<b>1. INTRODUÇÃO</b> .....	<b>1</b>
1.1. Câncer de pulmão .....	1
1.2. Genética dos carcinomas pulmonares .....	3
1.3. microRNAs (miRNAs) .....	4
1.4. miRNAs e câncer .....	6
1.5. miRNAs e alvos moleculares terapêuticos .....	7
<b>2. OBJETIVOS</b> .....	<b>9</b>
<b>3. MATERIAL E MÉTODOS</b> .....	<b>10</b>
3.1. Considerações sobre os aspectos éticos .....	10
3.2. Critérios de inclusão e exclusão dos pacientes .....	10
3.3. Parte I .....	10
3.3.1. Delineamento Experimental .....	10
3.3.2. Amostras .....	12
3.3.3. Extração do RNA .....	12
3.3.4. Sequenciamento de alto desempenho .....	13
3.3.5. Análises de sequenciamento .....	14
3.3.6. Validação da expressão de miRNAs .....	15
3.3.7. Análise estatística .....	16
3.4. Parte II .....	17
3.4.1. Meta-análise em NSCLC .....	17
3.4.2. Amostras utilizadas na análise da expressão gênica .....	20
3.4.3. Extração de RNA .....	20
3.4.4. Transcrição Reversa (RT) .....	21
3.4.5. Testes de eficiência dos iniciadores ( <i>primers</i> ) .....	21
3.4.6. Seleção dos genes endógenos .....	22
3.4.7. Análise dos dados de genes endógenos pela RT-QPCR .....	24
3.4.8. Análise de expressão gênica .....	24
3.4.9. Análise Estatística .....	24
<b>4. RESULTADOS</b> .....	<b>25</b>
4.1. Parte I .....	25
4.1.1. Sequenciamento de miRNAs (miRNA-Seq) .....	25
4.1.2. miRNAs com expressão alterada em adenocarcinoma .....	25



4.1.3. Resultados da análise nas amostras não pareadas: .....	25
4.1.4. Resultados da análise nas amostras pareadas: .....	25
4.1.5. Validação de alterações na expressão de miRNAs .....	28
4.1.6. miRNAs com variações em nucleotídeos únicos .....	30
4.2. Parte II.....	34
4.2.1. miRNAs desregulados em NSCLC: meta-análise de dados da literatura .....	34
4.2.2. Expressão gênica desregulada em NSCLC .....	37
<b>5. DISCUSSÃO .....</b>	<b>42</b>
5.1. Parte I.....	42
5.2. Parte-II .....	47
<b>6. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>50</b>
<b>ANEXOS .....</b>	<b>64</b>
<b>MANUSCRITO .....</b>	<b>70</b>

## 1. INTRODUÇÃO

### 1.1. Câncer de pulmão

O câncer de pulmão é um dos tipos de câncer mais comuns e a causa mais frequente de óbito por câncer, sendo responsável por 1,6 milhões de mortes mundialmente (GLOBOCAN, 2012) [1]. No Brasil, as estimativas de incidência para 2014/2015 apontam para a ocorrência de aproximadamente 27 mil casos novos de câncer de pulmão, sendo a maioria no sexo masculino (16.400 homens e 10.930 mulheres) (Instituto Nacional de Câncer, INCA, 2014) [2]. O Estado de São Paulo apresenta um dos maiores números de casos de carcinomas de traquéia, brônquios e pulmões, com incidência registrada, em 2010, de 22,94 casos/100.000 homens e 11,83 casos/100.000 mulheres [3].

O carcinoma pulmonar é frequentemente diagnosticado em estádios avançados da doença, resultando em baixas taxas de sobrevida a longo prazo; a sobrevida de 5 anos varia entre 13 e 21% nos países desenvolvidos e entre 7 e 10% nos países em desenvolvimento (INCA, 2014) [2]. As taxas de mortalidade por carcinomas pulmonares superam as taxas de mortalidade por cânceres de próstata, estômago, esôfago, cólon e reto, combinadas [3].

Em 90% dos casos diagnosticados, o câncer de pulmão está associado ao uso do tabaco e/ou a exposição passiva à fumaça do cigarro, além de outros fatores de risco incluindo radiação e exposição à carcinógenos ocupacionais e ambientais [4-6]. Considerando que muitos dos pacientes novos com câncer de pulmão são ex-fumantes que pararam de fumar há mais de 10 anos, e que esta doença também acomete indivíduos não fumantes, é consenso geral que o câncer de pulmão é um grave problema de saúde pública, cujo controle, prevenção e o desenvolvimento de melhores estratégias de tratamento devem ser priorizados.

As decisões referentes ao tratamento de pacientes com câncer de pulmão são guiadas, em parte, pelos subtipos histológicos envolvidos. O câncer de pulmão é classificado em dois grupos histológicos principais: o carcinoma pulmonar de pequenas células (da sigla em inglês SCLC, *Small Cell Carcinoma*) e o carcinoma pulmonar de não pequenas células (da sigla em inglês NSCLC, *Non-Small Cell Lung Carcinoma*). O SCC ocorre em ~15% dos pacientes, caracterizando-se por tumores mais agressivos e associados a um prognóstico pobre com baixa sobrevida. O NSCLC é o tipo histológico mais comum (~85% dos casos) e compreende vários subtipos histológicos, sendo que os mais frequentes são o adenocarcinoma (aqui definido pela sigla em inglês AD, *Adenocarcinoma*), o carcinoma de células escamosas (aqui definido pela sigla em inglês SCC, *Squamous Cell Carcinoma*) e o carcinoma de grandes células (aqui definido pela sigla em inglês LCC, *Large Cell Carcinoma*) [7, 8]. Dentre estes, o

AD é o mais frequente (37% em homens, 47% em mulheres), seguido pelo SCC (32% homens, 25% mulheres) e o LCC (18% homens, 10% mulheres) [7, 9].

As principais opções de tratamento para pacientes diagnosticados com NSCLC compreendem a cirurgia, bem como estratégias de quimioterapia e ou radioterapia. Normalmente, a prática clínica identifica somente 20% dos casos de câncer de pulmão com estadiamento I, a fase mais precoce, quando a cirurgia resulta em cura em aproximadamente 70% dos casos [10]. Entretanto, a maioria dos pacientes com NSCLC são diagnosticados em estadiamento avançado da doença e tratados por quimioterapia de primeira linha (Paclitaxel + Carboplatina, Gencitabina + Carboplatina, Permetrexede + Carboplatina) ou terapia alvo, sendo que as drogas mais comumente utilizadas incluem Gefitinib ou Erlotinib (*EGFR* positivo) e Crizotinib (rearranjo *ALK* positivo) (NCCN Guidelines, 2014) [11].

Nos últimos anos, a descoberta de mutações no domínio tirosina-quinase (TK) do gene do receptor do fator de crescimento epidérmico (*EGFR*) resultou em avanços nas estratégias de tratamento de pacientes com NSCLC, com o desenvolvimento de inibidores TK do EGFR, erlotinibe e gefitinibe [12-14]. Mutações no *EGFR* ocorrem em ~10-40% dos casos de NSCLC, preferencialmente em AD, sendo raras em SCC [15, 16]. As mutações mais comumente encontradas no domínio TK do EGFR compreendem deleções nos éxons 19 e 21 (deleção *in-frame* em ~45%, nomeada delE746-A750 e mutação *missense* em ~40%, nomeada L858R, respectivamente), as quais conferem sensibilidade aos inibidores TK do EGFR; tais mutações ocorrem preferencialmente em pacientes com AD, não fumantes, na maioria das vezes do sexo feminino (revisado em [17]).

Outras mutações frequentes incluem mutações no gene supressor tumoral *TP53* em aproximadamente 50% dos tumores de pacientes com NSCLC (AD e SCC) [7]. Adicionalmente, a ativação do proto-oncogene *K-RAS*, um efetor *downstream* da via molecular do *EGFR*, ocorre em tumores de pacientes com historia de consumo excessivo de tabaco. Mutações em *K-RAS* ocorrem com uma frequência de 15-35% dos casos de AD e <5% dos casos de SCC (revisado em [18]). Mutações em *EGFR* e *K-RAS* são mutuamente exclusivas e a presença de mutações em *K-RAS* está associada à resistência ao tratamento com inibidores do *EGFR* [19, 20].

Estratégias de tratamento para pacientes com mutações no *K-RAS* ainda são limitadas. Estudos atuais têm investido nesse grupo de pacientes, com o desenvolvimento de bloqueadores da cascata de sinalização intracelular mediada pelo *K-RAS*. A via de sinalização *RAF–MEK–ERK*, a qual é ativada diretamente na via molecular do *K-RAS*, provou ser um alvo válido em modelos pré-clínicos e clínicos (revisado em [21]). Outro tratamento

direcionado tem como alvo o *MEK*, entretanto o uso de inibidores do oncogene *MEK* tem sido associado a efeitos de alta toxicidade (revisado em [21]).

Diferentemente, pacientes com SCC apresentam mutações em outros receptores tirosina-quinase, como mutação em *DDR2* e ampliações/mutações na família do *FGFR* (receptor de fator de desenvolvimento em fibroblasto), sendo que ambos são alvos terapêuticos para estes pacientes [22].

Outra via de sinalização importante que está ativada em câncer, é a via *PI3K/AKT/mTOR*, a qual apresenta papéis diversos em condições fisiológicas normais e em processos oncogênicos, incluindo proliferação celular, sobrevivência e diferenciação [23]. Estudos pré-clínicos já vêm utilizando inibidores da via *PI3K/AKT/mTOR* em pacientes com tumores sólidos avançados, incluindo o carcinoma pulmonar do subtipo histológico SCC [24]. Um dos genes que inibem a via de sinalização *PI3K/AKT/mTOR*, que por sua vez, é ativada por membros receptores tirosina quinase (ex. *EGFR* e *HER2*) é o gene *PTEN* (revisado em [17]).

## 1.2. Genética dos carcinomas pulmonares

Estudos têm sido realizados com o objetivo de identificar novas mutações ou outras alterações associadas à patogênese do carcinoma pulmonar. Um estudo de caracterização do perfil mutacional de SCC de pulmão realizado pelo projeto *The Cancer Genome Atlas* (TCGA)\* (<http://cancergenome.nih.gov/>) identificou 360 mutações em regiões exônicas, 323 segmentos com alterações no número de cópias do DNA e 165 rearranjos cromossômicos. Mutações no gene *TP53* foram observadas em 81% (144/178) amostras tumorais de CCE. Interessantemente, alterações relativamente raras em AD foram identificadas, tais como ampliações do gene *SOX2*, mutações em *NFE2L2* (fator nuclear eritróide 2, similar) e *KEAP1* (*kelch-like ECH-associated protein 1*, esse gene é um alvo molecular no processo biológico de stress oxidativo), modificações na via de sinalização molecular do *PI3K* (fosfatidil inositol quinase), amplificação em *FGFR1* e mutações em *DDR2* (receptor de tirosina-quinase) [25].

\*TCGA é um consórcio genômico internacional estabelecido para determinar as mutações, alterações genômicas, transcritômicas e epigenômicas mais frequentes em diferentes tipos de tumores humanos, incluindo carcinomas pulmonares dos subtipos histológicos principais; AD e SCC.

Outro estudo de sequenciamento de alto desempenho revelou 1.063 genes desregulados, sendo que os genes *CBX3* e *CRABP2* estavam altamente expressos em NSCLC comparado com tecido pulmonar histologicamente normal (sem evidências de alterações histológicas). O nível aumentado de expressão do gene *CBX3* e de sua proteína sugere que *CBX3* pode contribuir com o desenvolvimento tumoral [26].

Embora tais estudos tenham identificado mutações em NSCLC, ainda não há dados de integração de alterações genéticas e epigenéticas identificadas como condutoras da tumorigênese (*drivers*), os quais são importantes para o entendimento das vias moleculares alteradas na doença. Além disso, é consenso geral que aproximadamente 50% das alterações moleculares responsáveis pela tumorigênese do carcinoma pulmonar, especialmente o subtipo AD, ainda não são conhecidas. Considerando que as terapias com alvos moleculares têm beneficiado uma fração pequena dos pacientes com AD, ainda é necessária a identificação de novos alvos moleculares nesses carcinomas [27].

Nos últimos anos, têm-se evidenciado o papel de RNAs não codantes, especialmente os microRNAs (miRNAs) em diversas patologias, incluindo doenças crônico-degenerativas como o câncer.

### 1.3. microRNAs (miRNAs)

Estudos de biomarcadores e alvos terapêuticos têm evidenciado o papel biológico dos RNAs não-codantes, em particular os miRNAs, no desenvolvimento e progressão de neoplasias humanas. miRNAs são RNAs pequenos (~18-22 nucleotídeos), não codificadores de proteínas, transcritos a partir do DNA e com papel na regulação pós-transcricional da expressão gênica [28]. Os miRNAs são transcritos pela RNA polimerase II como um transcrito primário longo (pri-miRNA), caracterizado por conter uma estrutura em grampo (aqui definida como *hairpin*). O pri-miRNA é clivado no núcleo pela enzima RNase III (Drosha), a qual forma um complexo de ~650 KDa acoplado ao co-fator DGCR8 (*Di George syndrome critical region 8*). O DGCR8 contém dois domínios de ligação de RNA dupla fita (dsRNA). A atividade desse complexo em uma molécula de pri-miRNA resulta em um miRNA precursor (pré-miRNA) em forma de *hairpin* e de ~70 a 100 nucleotídeos de comprimento [29, 30]. Posteriormente, o pré-miRNA é exportado para o citoplasma pela atividade da proteína Exportina-5, onde a RNase III (Dicer) em conjunto com o complexo *transactivating response RNA-binding protein* (TRBP), cliva a região do *loop* (laço) originando um RNA de fita dupla (dsRNA) de ~22 nucleotídeos; essa molécula é denominada miRNA/miRNA\*. Em um mecanismo posterior, uma das fitas é selecionada

(principalmente baseado na estabilidade termodinâmica da molécula), como a fita guia funcional (miRNA), e a outra fita complementar (miRNA\*) é sujeita a degradação [30] ou ainda pode ser funcional e apresentar papel biológico importante. Recentes estudos revelaram que a fita complementar miRNA\* é ativamente selecionada por proteínas Ago e podem participar da montagem do RISC, resultando na atividade funcional contra o gene alvo (revisado em [31]).

Após completar os passos de processamento, o miRNA maduro de fita simples é incorporado ao complexo de silenciamento induzido por RNA (aqui definido pela sigla em inglês RISC, *RNA-induced silencing complex*). O complexo RISC contém uma família de proteínas denominada Argonata (Ago 1 e Ago 2) e uma proteína de ligação ao RNA (TRBP, do inglês *Trans-activator RNA binding protein*) [30]. Como resultado desse processo, a fita funcional do miRNA maduro é capaz de regular a expressão gênica em nível pós-transcricional; essa atividade ocorre pela ligação, por pareamento complementar de bases, da extremidade 5' do miRNA (também denominada de sequência *seed*) à extremidade 3' não traduzida (3' UTR) do mRNA, levando, na maioria dos casos, à inibição da tradução do mRNA por homologia incompleta ou à degradação do mRNA por homologia completa [32].

Outras vias não canônicas ou alternativas, de processamento de miRNAs foram descritas em organismos de mamífero, drosófila e nematóide [33-35]. Mirtrons consistem em um tipo de miRNA processado por um mecanismo não canônico, onde moléculas de pré-miRNA são derivadas de um produto secundário da maquinaria de processamento alternativo de RNA (*splicing*) e processadas pela enzima Ldbr (*lariat-de-branching enzyme*), independente da clivagem pela enzima Drosha [34, 36-38].

Interessantemente, miRNAs também podem se ligar a ribonucleoproteínas na sequência *seed* de modo independente de RISC, e conseqüentemente, interferir na ligação miRNA:mRNA resultando na ativação da tradução. Adicionalmente, foi demonstrado que os miRNAs também podem regular a expressão gênica em nível transcricional, através da ligação direta com o DNA [30].

A biogênese e processamento de miRNAs por meio de vias moleculares canônicas e não canônicas demonstra a flexibilidade dos mecanismos de biogênese e processamento de RNAs e evidenciam como múltiplas funções podem ser realizadas por transcritos de RNA em nível celular.

Os miRNAs constituem, portanto, moléculas com papel relevante na regulação da expressão gênica e estão envolvidos em processos biológicos importantes, tais como o desenvolvimento embrionário, diferenciação, apoptose, proliferação celular [39-41] e em

mecanismos de oncogênese [42-44]. No câncer, os miRNAs atuam como moléculas reguladoras agindo de forma similar a oncogenes ou a genes supressores tumorais, sendo denominados *oncomiRs* [28, 45]. Atualmente, foram identificados 2.588 miRNAs maduros no genoma humano (Banco de dados miRBase, última atualização em Outubro de 2014, <http://www.mirbase.org/cgi-bin/browse.pl?org=hsa>), sendo que miRNAs são encontrados em localizações genômicas tais como regiões intrônicas de transcritos não codificadores, regiões intrônicas de transcritos codificadores (tais miRNAs compartilham o promotor gênico), regiões exônicas de transcritos não codificadores ou regiões exônicas de transcritos codificadores, além de miRNAs que constituem íntrons (miRtrons) [46, 47].

#### 1.4. miRNAs e câncer

Vários estudos já demonstraram a expressão desregulada de miRNAs associada à patogênese do câncer. Alterações na expressão de miRNAs podem ocorrer devido a diferentes mecanismos, tais como: anormalidades cromossômicas (como ampliações, deleções ou translocação), mutações e polimorfismos (SNPs) na sequência da molécula de miRNA, desregulação transcricional, defeitos na maquinaria da biogênese de miRNAs ou modificações epigenéticas (como alteração da metilação e remodelação da cromatina) (revisado em [48]).

Uma das primeiras evidências da associação dos miRNAs com a tumorigênese humana foi a identificação de uma família de miRNAs (miR-15a-16-1) mapeada em 13q14, uma região de deleção característica de células tumorais de leucemia linfocítica crônica (da sigla em inglês CLL, *chronic lymphocytic leukemia*). Esse achado associou essa deleção genômica com a perda da expressão de miR-15a-16-1 e conseqüentemente a perda do controle da expressão do gene *BCL2* (*B-cell CLL/Lymphoma 2*), o qual apresenta papel biológico anti-apoptótico em algumas células, tais como os linfócitos [49]. Esse e outros estudos evidenciaram que os miRNAs estão frequentemente localizados em regiões de ganhos e perdas genômicas relacionadas ao câncer (revisado em [30]).

Interessantemente, um estudo de genoma total identificou alterações no número de cópias do DNA associadas à desregulação da expressão de 41 miRNAs em diferentes tipos de câncer incluindo carcinomas de ovário, mama e melanoma [50]. Outro estudo mostrou que a presença de uma variação homozigota no mir-196a2 estava associada a baixa sobrevivência de pacientes com NSCLC; tais polimorfismos podem resultar em uma ligação enfraquecida do miRNA ao seu mRNA alvo ou alterações nos mRNAs-alvo [51].

Estudos de expressão global de miRNAs têm sido utilizados para a identificação de miRNAs como biomarcadores prognósticos em carcinomas pulmonares. Exemplos incluem a

identificação de expressão diminuída do miR-100 associada ao aumento da expressão do gene *PLK1* (*polo-like kinase 1*) em NSCLC, indicando que o miR-100 pode representar um biomarcador de prognóstico desfavorável; adicionalmente esse estudo demonstrou que a restituição da expressão do miR-100 suprimiu a tumorigenicidade das células, possivelmente devido ao aumento da apoptose associada a baixa expressão de *PLK1* e sugeriu que estratégias utilizando antagonistas de miRNAs (por exemplo, um oligonucleotídeo anti-miR-100) podem ter papel terapêutico [52].

Estudos em mesotelioma pleural maligno (MPM) e em NSCLC mostraram a expressão diminuída do miR-let-7a e a expressão aumentada de seu mRNA-alvo *EphA2* (*ephrin type-A receptor 2*), o qual codifica um receptor tirosina quinase [53, 54]. Adicionalmente, observou-se que *EphA1*, um ligante específico do receptor *EphA2*, inibe a proliferação e migração celular [55, 56]. Estudos experimentais *in vitro* demonstraram a ligação do miR-let-7a ao mRNA-alvo *EphA2*. O miR-let-7a foi encapsulado a uma nanopartícula lipossomal e a proteína EphA1 foi conjugada à superfície da nanopartícula. O complexo mostrou-se eficaz quando testado em células de NSCLC, inibindo a proliferação, migração e crescimento tumoral, servindo como potencial alvo terapêutico [57].

### 1.5. miRNAs e alvos moleculares terapêuticos

Os avanços no entendimento da biologia do câncer têm conduzido ao desenvolvimento de terapias-alvo específicas, com o objetivo de bloquear vias moleculares desreguladas nas células tumorais. Esses tratamentos visam atuar seletivamente nas células tumorais contendo alterações moleculares específicas e assim reduzindo a citotoxicidade associada à quimioterapia convencional.

Recentemente, foi mostrada a associação entre a via de sinalização regulada pelo *EGFR* e a expressão alterada de miRNAs *in vivo*, *in vitro* e em amostras clínicas [58]. Neste estudo, a expressão aumentada do miR-21 foi associada a resistência ao tratamento com EGFR-TKI (gefitinib). Observou-se que o miR-21 pode promover a proliferação celular e induzir resistência ao EGFR-TKI através da inibição da expressão do PTEN (*phosphatase and tensin homolog*) e PDCD4 (*programmed cell death 4*) e ativação da via de sinalização PI3K/AKT. Adicionalmente, foi mostrado que o miR-21 circulante pode ser um indicador prognóstico robusto para pacientes com NSCLC resistentes ao EGFR-TKI [58].

Outro estudo mostrou que a alta expressão do miR-31 contribui com a tumorigênese pulmonar modulando a expressão dos oncogenes *K-RAS* e *EGFR* [59]. Foi demonstrado que os receptores *IGF-1R* (*insulin-like growth factor 1 receptor*), *TGFBR1* (*transforming growth*



*factor, beta receptor 1*) e *EGFR*, são alvos moleculares do miR-133a. Notavelmente, pacientes com baixa expressão do miR-133a apresentaram melhor sobrevida comparado aos pacientes com alta expressão desse miRNA [60].

Terapias baseadas em miRNAs têm sido um importante foco de pesquisas com a estratégia de restaurar ou inibir a função dos miRNAs usando miRNAs mimics ou anti-miRs (oligonucleotídeos miRNAs de dupla-fita) e antagomirs. Estudos *in vitro* e *in vivo* tem mostrado a eficácia do uso de miRNAs mimics contra o desenvolvimento tumoral (revisado em [61]). Atualmente, essa terapia está sendo utilizada em estudos de primeira-fase em pacientes com câncer de fígado primário ou metastático, com o objetivo de restaurar a perda da função supressora do miR-34, usando o miRNA mimic sintético MRX34, o qual consiste em um RNA de fita dupla entregue por lipossomos. A terapia usando MRX34 regula diretamente aproximadamente 24 oncogenes, envolvidos no ciclo celular e proliferação, anti-apoptose, metástase, quimioresistência e transcrição oncogênica [62].

Considerando a importância dos miRNAs nos processos associados à tumorigênese humana, a caracterização do perfil global de alterações na expressão e mutações em miRNAs em NSCLC deverá contribuir para o desenvolvimento de biomarcadores com potencial aplicação clínica. Biomarcadores moleculares podem ser úteis no desenvolvimento de estratégias para detecção precoce de doença, diagnóstico, prognóstico e terapêuticas mais precisas para o tratamento de pacientes com NSCLC.

## **2. OBJETIVOS**

1. Determinar o perfil global de mutações e alterações na expressão de miRNAs em adenocarcinoma pulmonar;
2. Correlacionar as alterações com dados clínicos e histopatológicos dos pacientes;
3. Identificar genes-alvo e vias moleculares moduladas por miRNAs como potenciais alvos terapêuticos em adenocarcinoma pulmonar.

As seções Material e Métodos, Resultados e Discussão estão divididas em Parte I e Parte II, as quais serão utilizadas para a elaboração de dois manuscritos independentes.

### **3. MATERIAL E MÉTODOS**

#### **3.1. Considerações sobre os aspectos éticos**

Este projeto de pesquisa foi aprovado pelos Comitês de Ética em Pesquisa (CEP) da Faculdade de Medicina de Botucatu (FMB) (Protocolo CEP 4319/2012), do Hospital AC Camargo (Protocolo CEP 1573/11) e do Hospital de Câncer de Barretos (Protocolo CEP 75907). O Termo de Consentimento Livre e Esclarecido (TCLE) foi aplicado a todos os pacientes elegíveis ao estudo.

#### **3.2. Critérios de inclusão e exclusão dos pacientes**

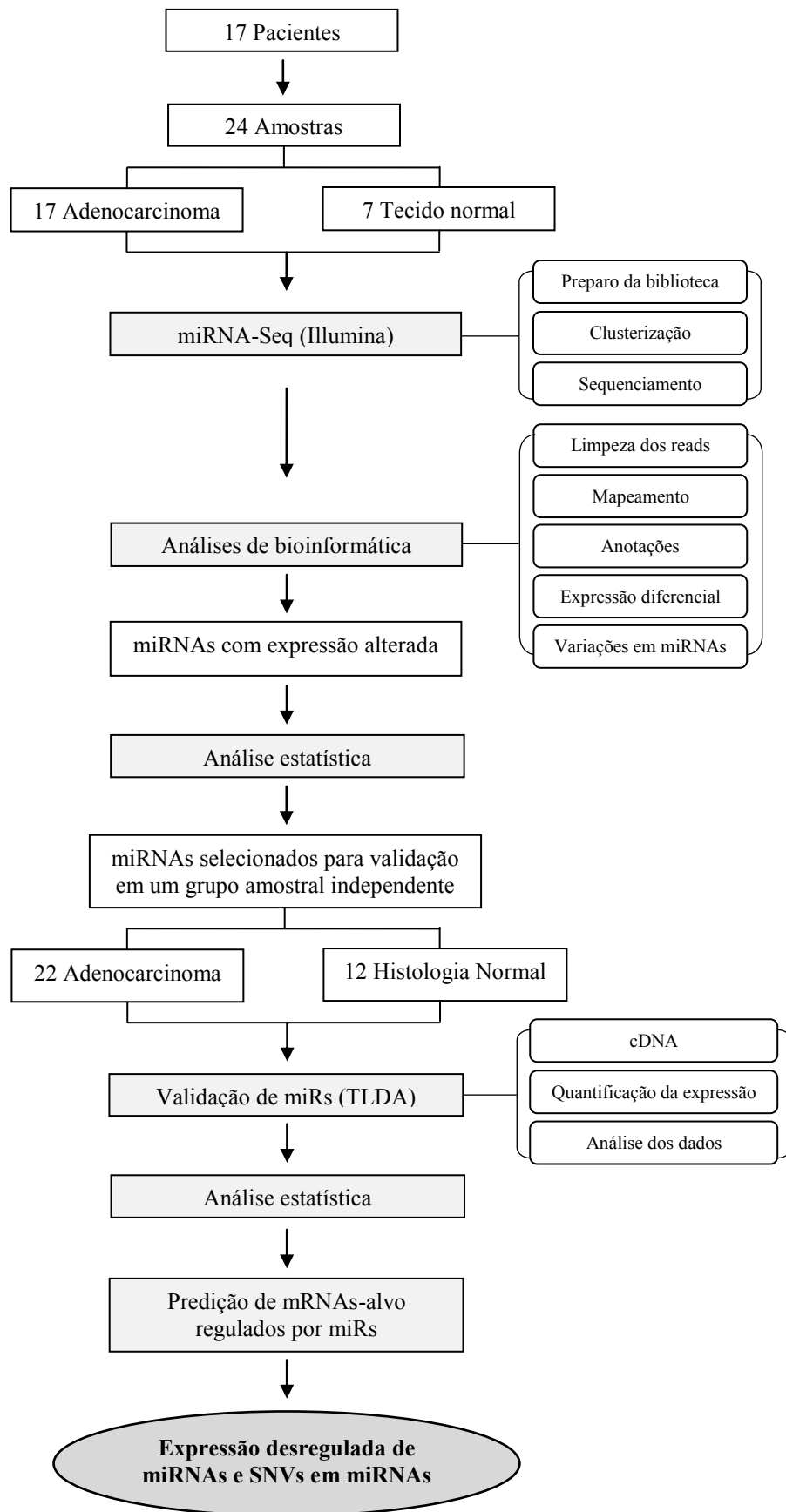
Os critérios de inclusão foram os seguintes: indivíduos maiores de 18 anos e capazes de entender o estudo e fornecer o consentimento informado, diagnóstico de NSCLC do subtipo histológico adenocarcinoma, pacientes submetidos à biópsia para fins de diagnóstico, ou cirurgia para fins de tratamento primário da doença, não sendo tratados previamente com quimioterapia e ou radioterapia, com tumores de qualquer localização anatômica nos pulmões e de qualquer estadiamento e grau histológico.

Os critérios de exclusão foram: pacientes menores de idade ou incapazes de entender o estudo e fornecer o consentimento informado, pacientes com outros tumores ou carcinoma pulmonar de outros subtipos histológicos.

#### **3.3. Parte I**

##### **3.3.1. Delineamento Experimental**

Esse projeto propôs a identificação de um painel de alterações na expressão e ou mutações em sequências de miRNAs maduros (miRNAs funcionais) em adenocarcinoma pulmonar (AD), por meio da utilização de plataformas de análise de sequenciamento de alto desempenho e de detecção pela reação em cadeia da polimerase (PCR) quantitativa em tempo real (aqui definida como RT-QPCR). O perfil desregulado de expressão de miRNAs foi identificado em um grupo amostral (grupo teste) seguido de validação em um grupo amostral independente (grupo de validação). Os grupos teste e validação estão definidos conforme descrito abaixo. A **Figura 1** abaixo apresenta o delineamento experimental, incluindo as amostras e as metodologias aplicadas.



**Figura 1.** Delineamento experimental e amostras

### 3.3.2. Amostras

Grupo teste: Amostras de tecido tumoral (AD) e de tecido pulmonar adjacente não neoplásico e sem evidência de qualquer alteração histológica (aqui definido como histologicamente normal, N) foram obtidas do banco de tumores do Hospital AC Camargo, São Paulo, SP. Foram coletadas, prospectivamente, 24 amostras de 17 pacientes (17 AD e 7 tecidos pulmonares histologicamente normais; 5/7 amostras eram pareadas, ou seja, provenientes dos mesmos pacientes). As amostras cirúrgicas foram coletadas e congeladas imediatamente a  $-80^{\circ}\text{C}$ , segundo procedimento operacional padrão vigente no Biobanco do Hospital AC Camargo.

Grupo de validação: As amostras do grupo de validação foram obtidas retrospectivamente, no período de 2000-2012, sendo provenientes do Depto. de Patologia do Hospital de Câncer de Barretos e do Hospital das Clínicas (HC) da FMB. Esse grupo foi composto por 22 amostras de AD (N=20 de Barretos e N=2 do HCB), fixadas em formalina e embebida em parafina (FFPE). No grupo de validação, as amostras controle (N=12) foram tecidos pulmonares histologicamente normais coletados de margens cirúrgicas. Os tecidos normais utilizados como controle incluíram amostras de ambos os centros (Barretos e HCB). Para as amostras coletadas retrospectivamente, foi solicitada a dispensa de aplicação do Termo de Consentimento Livre e Esclarecido, pois a maioria dos pacientes foi a óbito pela doença ou não pode ser localizado.

Os dados clínicos, histopatológicos, sobrevida e resposta ao tratamento dos pacientes compondo o grupo teste e o grupo de validação estão apresentados nas **Anexos 1 e 2**, respectivamente.

### 3.3.3. Extração do RNA

O RNA das amostras de tecidos coletados a partir de cirurgias (provenientes do Biobanco do Hospital AC Camargo) foi extraído utilizando protocolo de extração do kit *RNeasy Mini Kit* (Qiagen, Hilden, Germany). O RNA foi obtido do banco de macromoléculas do mesmo hospital. Foi obtido, em média, 1 $\mu\text{g}$  de RNA total de cada amostra.

A extração de RNA das amostras FFPE (grupo de validação), foi realizada no laboratório NeoGene, Unidade de Pesquisa Experimental (UNIPLEX) da FMB. Para isolamento do RNA de amostras FFPE utilizamos o kit *RecoverAll Total Nucleic Acid Isolation* (Ambion/Life Technologies, Carlsbad, CA, USA), o qual é otimizado para extração de RNA de amostras FFPE.

Todas as amostras de RNA foram quantificadas utilizando o *NanoDrop 8000* (Thermo Fisher Scientific, Waltham, MA, USA) e a qualidade foi determinada utilizando *Bionalyzer 2100* (Agilent Technologies, Santa Clara, CA, USA).

### 3.3.4. Sequenciamento de alto desempenho

A análise de sequenciamento foi realizada no grupo teste (N=24 amostras, 17 AD e 7 N). 1 µg de RNA total foi utilizado para a construção de bibliotecas, purificação, amplificação de miRNAs e posterior sequenciamento de miRNAs em plataforma de alto desempenho. O sequenciamento foi realizado utilizando a plataforma MiSeq (Illumina, San Diego, CA, USA), no Laboratório de Biotecnologia Animal, Escola Superior de Agricultura Luiz de Queiroz (ESALQ), USP, Piracicaba, SP, em colaboração com o Dr. Luiz Lehmann Coutinho.

O método de sequenciamento utilizado consiste na clonagem, *in vitro*, dos fragmentos de RNA em uma plataforma sólida, processo denominado PCR de fase sólida. A plataforma MiSeq integra clusterização de fragmentos, sequenciamento e análise de dados em um único instrumento e tem capacidade de produzir seqüências nucleotídicas de 50 pb (leitura única ou *single read*) e até 7,5 gigabases por corrida. O sequenciamento ocorre em uma lâmina de fluxo ou superfície de clonagem (*flow cell*), a qual apresenta uma canaleta única (*lane*), possibilitando sequenciar até 12 amostras simultaneamente.

O protocolo de sequenciamento seguiu as recomendações do fabricante, o qual consiste em três etapas principais: preparo das bibliotecas, clusterização de fragmentos e sequenciamento. A construção das bibliotecas utilizou o kit *TruSeq Small RNA Sample Preparation*, conforme descrito abaixo.

*Etapas 1: Construção das bibliotecas.* As amostras de RNA foram diluídas conforme instruções do protocolo (a partir de 1 µg de RNA total), seguido da ligação das moléculas de RNA aos adaptadores. Os adaptadores apresentam um sítio de ligação ao iniciador (*primer*) de sequenciamento ou *index*, utilizado para identificar as amostras contidas no *pool* de amostras de RNA, e um sítio de ligação aos oligonucleotídeos iniciadores (*primers*) presentes na *flow cell* (superfície de clonagem) utilizados na amplificação dos fragmentos pela PCR. A transcrição reversa foi realizada para formar a segunda fita de cDNA, baseada na ligação dos pequenos RNAs com os adaptadores 5' e 3'. A biblioteca de cDNA foi obtida através da amplificação pela PCR, a qual permite enriquecer os fragmentos com os adaptadores de forma seletiva. Nesse processo são utilizados o *Master Mix*, *primers* de amplificação pela PCR e *primers index-específicos*, sendo que a amplificação ocorreu de acordo com as seguintes

condições: 98°C por 30 segundos, 98°C por 10 segundos; 60 °C, 30 segundos; 72°C, 15 segundos seguido de 72°C, 10 minutos, completando um total de 11 ciclos.

A validação da biblioteca de cDNA foi realizada no *Bionalyzer*; essa etapa é importante como controle de qualidade do cDNA. Posteriormente, as amostras a serem sequenciadas foram unidas em um *pool* e submetidas à eletroforese em gel de agarose, para separação dos miRNAs dos demais RNAs e posterior clusterização. Após a purificação da biblioteca, o cDNA foi precipitado em etanol para obter maior concentração. Em seguida foi realizada a quantificação da biblioteca usando *Qubit 2.0 Fluorometer* (Invitrogen by Life Technologies, Carlsbad, CA, USA).

*Etapa 2. Clusterização.* O processo de clusterização consiste na produção de múltiplas cópias de fragmentos específicos através da amplificação pela PCR em fase sólida. Nesse processo, os fragmentos ligados aos adaptadores são desnaturados para separação da fita dupla; as moléculas de fita simples então reconhecem os *primers* fixados na superfície da *flow cell*, e a enzima DNA polimerase incorpora nucleotídeos, formando pontes de fita dupla. Ao final da reação, cada fragmento foi representado inúmeras vezes, agrupados em *clusters*.

*Etapa 3: Sequenciamento.* A corrida de sequenciamento foi realizada utilizando o MiSeq Reagent Kit v2 (50 ciclos). O protocolo de sequenciamento seguiu as recomendações do fabricante. Essa etapa consiste na adição de terminadores de cadeia (ddNTPs - dideoxinucleotídeos) marcados com fluoróforos, *primers* e enzima DNA polimerase na *flow cell*, resultando na incorporação de um nucleotídeo por ciclo de sequenciamento. Os sinais fluorescentes emitidos em cada cluster e correspondentes a cada base foram registrados e interpretados utilizando um software específico, possibilitando a construção e a leitura das sequências geradas.

### 3.3.5. Análises de sequenciamento

Após o sequenciamento, milhões de fragmentos de RNA foram gerados e a qualidade das sequências (aqui definidas como *reads*) foi visualizada pela ferramenta FastQC [63]. A limpeza dos *reads* foi realizada pelo *CutAdapt* [64], o qual permite retirar as sequências de adaptadores e selecionar os *reads* de boa qualidade. Os *reads* que não passaram o controle de qualidade, com média de qualidade inferior ao *score* do programa de análise, foram removidos do alinhamento e não foram utilizados nas análises subsequentes. O mapeamento/alinhamento das amostras de adenocarcinoma e das amostras histologicamente normais foi realizado com base no genoma de referência mais recente, hg19 (<https://genome.ucsc.edu/cgi-bin/hgTracks?hgsid=12832096&chromInfoPage=>), utilizando o

programa *Bowtie1* [65]. O arquivo resultante do alinhamento foi submetido ao *script HT-Seq* [66], o qual realiza anotações e quantificações dos *reads*, a fim de identificar os miRNAs mapeados. Esta etapa de anotação e identificação foi realizada utilizando os arquivos contendo as sequências de referência: miRNA maduro, snoRNA e hairpins. A normalização dos dados [67] e a análise de expressão diferencial foram realizadas através do *edgeR* (Bioconductor/R) [68-70]. Para a identificação de polimorfismos de base única (SNPs), as sequências de miRNAs anotadas/mapeadas foram submetidas ao programa *SAMtools* [71, 72] e a ferramenta *VCFtools* [73] (disponível neste programa), para filtrar as variantes genômicas. Em seguida, as anotações dos SNPs foram realizadas através da ferramenta *ANNOVAR* [74].

As sequências resultantes foram alinhadas com o genoma humano de referência. Os dados obtidos permitiram a identificação de alterações na sequência e na expressão de miRNAs em amostras de tumores (AD) comparado com tecido histologicamente normal (amostras controle).

As etapas da análise dos dados de sequenciamento estão apresentadas no **Anexo 3**.

### **3.3.6. Validação da expressão de miRNAs**

O objetivo foi validar o perfil de expressão de miRNAs identificados como diferencialmente expressos em AD comparado com tecido histologicamente normal (N). A análise de validação utilizou dados obtidos na plataforma *TaqMan® Array Human MicroRNA A Cards v2.0* (Applied Biosystem, Foster City, CA, USA), a qual consiste em um cartão com sondas liofilizadas, que permite quantificar a expressão de 377 miRNAs humanos e é otimizada para análise de amostras FFPE. A plataforma inclui três controles internos (miRNAs endógenos) para normalização dos dados e um controle negativo, sendo que o protocolo permite amplificar uma amostra por placa. Em resumo, o protocolo inclui a transcrição reversa para converter moléculas de miRNAs em dupla fita, seguida da amplificação e quantificação da expressão pela RT-QPCR.

Os perfis de expressão de miRNAs foram determinados utilizando o programa *RQ Manager v1.2* (Life Technologies), disponível no equipamento *QuantStudio 12K* (Life Technologies). Foi realizada normalização global dos dados no programa *Expression Suite* (Life Technologies). Os resultados foram analisados pelo método Delta Delta Ct [75]. Esse método permitiu a quantificação da expressão de miRNAs em AD relativa às amostras controle.



### **3.3.7. Análise estatística**

Foram utilizados diferentes algoritmos de análise estatística em testes de correlação entre a expressão dos miRNAs alterados com os dados clínicos e prognóstico dos pacientes com AD. As análises estatísticas dos dados utilizaram os programas PRISMA e SAS 9.1 (*Statistical Analyses System*). Os dados de expressão e presença ou ausência de mutações em miRNAs foram associados com os dados clínicos dos pacientes, utilizando o teste de Mann-Whitney e o teste exato de Fisher. Essa etapa do trabalho foi realizada com auxílio do Prof. Dr. Rogério Antônio de Oliveira, Depto. de Bioestatística, Instituto de Biociências de Botucatu (IBB), UNESP.

### 3.4. Parte II

#### 3.4.1. Meta-análise em NSCLC

A meta-análise foi realizada com o objetivo de determinar as alterações moleculares envolvendo miRNAs e genes-alvo regulados por miRNAs em NSCLC, independente de subtipo histológico, bem como comparar os dados dos estudos em AD com os nossos dados de miRNA-Seq. A **Figura 2**, abaixo, mostra o delineamento da meta-análise.

Foram incluídos 19 estudos publicados [76-94]. Os estudos selecionados, contendo dados brutos disponíveis de expressão de miRNAs em NSCLC estão apresentados na **Tabela 1**, abaixo. Dados de estudos publicados (2000-2014) foram pesquisados e obtidos dos bancos de dados públicos *PubMed* ([www.ncbi.nlm.nih.gov/pubmed](http://www.ncbi.nlm.nih.gov/pubmed)) e *Gene Expression Omnibus* (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>).

Os critérios de inclusão de estudos foram os seguintes: trabalhos em NSCLC, de todos os subtipos histológicos, amostras de tumor primário, humano, inclusão de amostras de tecido pulmonar normal dos mesmos pacientes, mesmo que parcialmente, amostras provenientes de tecido coletado a fresco e ou congelado ou fixado em parafina, contendo dados de expressão global de miRNAs e que foram validados, acesso aos dados brutos de expressão de miRNAs. As palavras-chave utilizadas na pesquisa foram: *global gene expression, global microRNA expression, microRNA expression, microRNA signature, non-small cell lung cancer, tissue, human*.

Os miRNAs identificados na meta-análise foram utilizados em análises de predição de genes-alvo potencialmente regulados por miRNAs. Para tal, utilizamos a ferramenta mirDIP - microRNA Data Integration Portal (<http://ophid.utoronto.ca/mirDIP/>) [95], a qual integra dados de miRNAs disponíveis nos principais bancos de dados públicos; DIANA-microT [96, 97], microRNA.org/miRanda [98, 99], microCosm Targets (<http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/>), picTar (<http://pictar.mdc-berlin.de/>), PITA ([http://genie.weizmann.ac.il/pubs/mir07/mir07\\_dyn\\_data.html](http://genie.weizmann.ac.il/pubs/mir07/mir07_dyn_data.html)), RNA22 microRNA target detection (<http://cbcsrv.watson.ibm.com/rna22.html>) e TargetScan [100-102].

A seguir, análises de integração dos dados de expressão de miRNAs e dos genes-alvo preditos foram realizadas utilizando a ferramenta I2D - *Interologous Interaction Database* (v.2.3, <http://ophid.utoronto.ca/ophidv2.204/>) [103, 104] para identificar redes de interação proteica e vias moleculares relevantes em AD pulmonar. A interpretação e visualização desses dados foram realizadas no sistema NAViGaTOR (v.2.3.0, <http://ophid.utoronto.ca/navigator/>) [105]. Com base nessas análises, genes-alvo preditos foram selecionados para análise da expressão gênica em amostras primárias de NSCLC e tecido pulmonar normal.

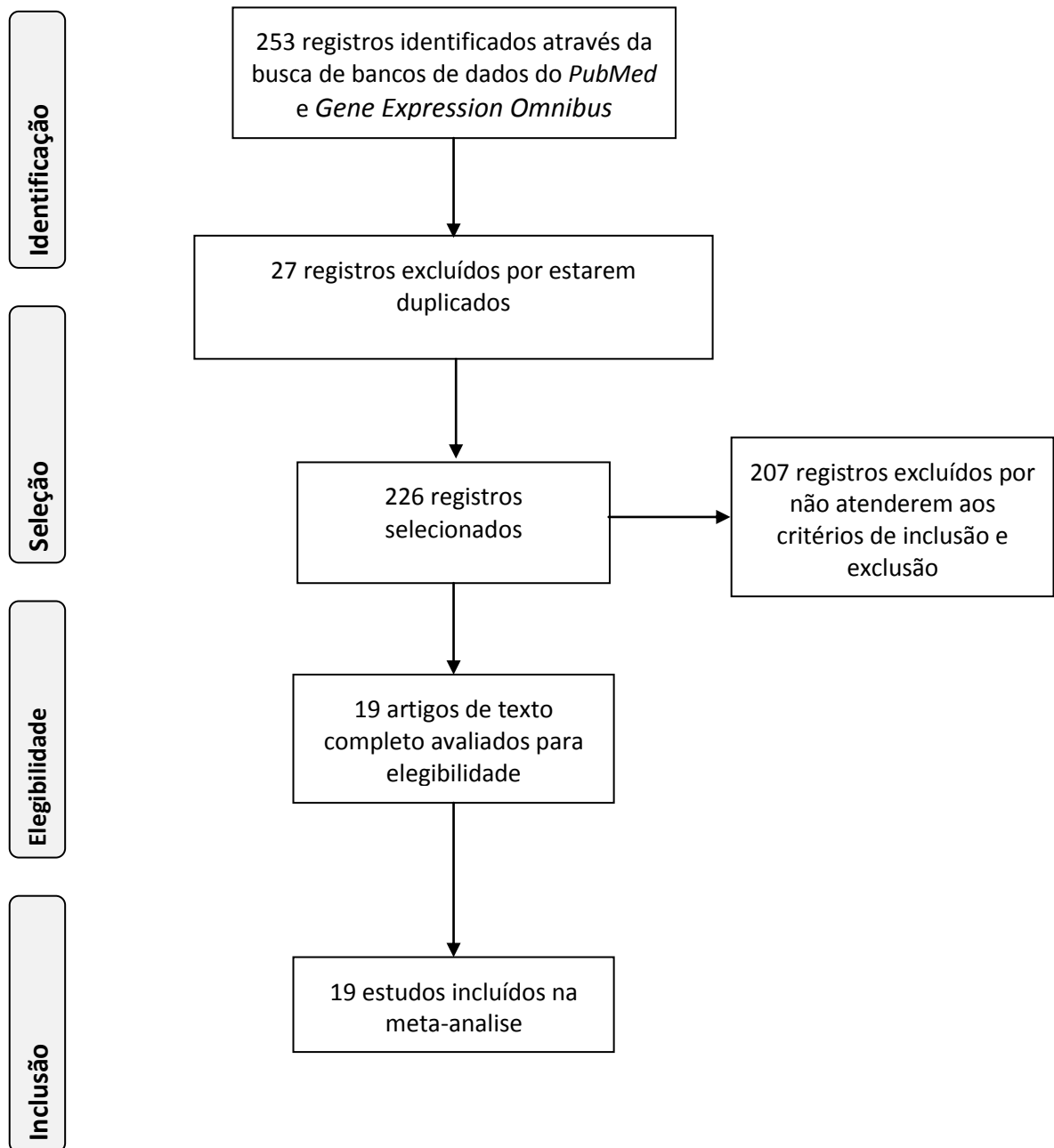


Figura 2. Delineamento da Meta-análise

**Tabela 1.** Descrição dos 19 artigos utilizados na meta-análise.

<b>Public Data Set</b>	<b>Sample Size</b>	<b>Platform</b>
<b>Bandi et al. 2009</b>	23 NSCLC, 3 N	7500 Real-time PCR System (ABI)
<b>Boeri et al. 2011</b>	28 NSCLC, 24 N	miRNA array, Ohio State University
<b>Cho et al. 2009</b>	10 NSCLC, 10 N	miRNA array (Agilent)
<b>Cho et al. 2011</b>	10 NSCLC, 10 N	NimbleGen 4-plex array (Roche)
<b>Gao et al. 2010</b>	8 NSCLC, 8 N	LNA miRNA array (Exiqon)
<b>Gao et al. 2011</b>	4 NSCLC, 4 N	LNA miRNA array (Exiqon)
<b>Incoronato et al. 2010</b>	18 NSCLC, 18 N	7900 HT RT-QPCR (ABI)
<b>Liu et al. 2012</b>	6 NSCLC, 6 N	miRNA array (Agilent)
<b>Ma et al. 2011 GSE29250</b>	6 NSCLC, 6 N	miRNA array (Illumina)
<b>Saito et al. 2011</b>	317 NSCLC, 317 N	7900 HT RT-QPCR (ABI)
<b>Seike et al. 2009 GSE14936</b>	28 NSCLC, 28 N	miRNA array, v. 3.0, Ohio University
<b>Solomides et al. 2012</b>	16 NSCLC, 4 N	miRNA array (Agilent)
<b>Shen et al 2011</b>	28 NSCLC, 28 N	Bio-Red IQ5 Multi-color RT-QPCR Detection System
<b>Tan X et al 2011</b>	60 NSCLC, 60 N	miRNA array (CapitalBio)
<b>Takamizawa et al. 2004</b>	143 NSCLC, 120 N	7900 RT-QPCR (ABI)
<b>Volinia et al. 2006</b>	123 NSCLC and 123 N	miRNA array (custom platform, Kimmel Cancer Center)
<b>Xing et al. 2010</b>	15 NSCLC, 15 N	Gene-Chip miRNA array (Affymetrix)
<b>Yanaihara et al. 2006</b>	104 NSCLC, 104 N	miRNA array (custom platform, Kimmel Cancer Center)
<b>Yu et al. 2010</b>	20 NSCLC, 20 N	miRNA array A (System Biosciences)

AD: adenocarcinoma, SCC: carcinoma de células escamosas, N: tecido de pulmão histologicamente normal, LNA: *locked-nucleic acid*

### 3.4.2. Amostras utilizadas na análise da expressão gênica

As amostras utilizadas para análise de expressão gênica foram coletadas a partir de cirurgias realizadas no Hospital das Clínicas da Faculdade de Medicina de Botucatu e no Hospital AC Camargo, São Paulo. Foram coletadas 55 amostras de tecido provenientes de 40 pacientes com NSCLC (32 tumores e 23 tecidos pulmonares histologicamente normais dos mesmos pacientes).

Todas as amostras de tumores e tecidos normais coletadas foram analisadas microscopicamente pelo patologista colaborador do estudo, Dr. Júlio Defaveri, anteriormente à extração de RNA, para assegurar a presença de um mínimo de 80% de células tumorais nas amostras de tumores e 80% de células normais nas amostras de tecido normal. Para tal, todas as amostras coletadas foram submetidas à análise de congelação. Os tecidos foram congelados em meio de congelação Tissue Tek® e colocados em cassetes histológicos. Esse procedimento utilizou o Criostato Leica CM 3050. Para confirmação dos tecidos tumorais e dos normais, foram realizados cortes dos tecidos congelados, seguido de coloração histológica por Hematoxilina-Eosina (H&E). A análise de congelação confirmou a presença de células tumorais ou células normais, nas respectivas amostras de tumor e normal coletadas.

Os critérios de inclusão dos pacientes foram baseados em doença primária, diagnóstico de carcinoma pulmonar do tipo não-pequenas células (NSCLC), incluindo os subtipos histológicos adenocarcinoma (AD) e carcinoma de células escamosas (CCE), não tratados previamente com quimioterapia e radioterapia, com tumores de qualquer localização anatômica nos pulmões e de qualquer estadiamento e grau histológico. As amostras de tecido coletadas foram congeladas em freezer -80 °C.

Os dados clínicos-histopatológicos dos pacientes incluídos na Parte II do trabalho estão apresentados no **Anexo 4**.

### 3.4.3. Extração de RNA

As amostras de tecido tumoral e normal foram cortadas (50-100 mg), imersas em 500 µL de reagente *Trizol* [106] (Life Technologies, Carlsbad, CA, USA), transferidas para tubos *Precellys Lysing kit* (Berting Technologies, Rockville, MD, USA) e lisadas no equipamento *Precellys 24 lysis and homogenization* (Berting Technologies, Rockville, MD, USA) por 10 segundos à 6.500 rpm. Após o processamento da amostra, o conteúdo foi transferido para tubos de 1,5 mL e adicionado mais 500 µL de *Trizol* e incubados por 3 minutos. A seguir, foram adicionados 200 µL de clorofórmio, seguido da centrifugação a 12.000g por 15 minutos, resultando na separação de três fases: vermelho-claro (fenol-clorofórmio), interfase

(resíduo do congelamento) e o sobrenadante (fase aquosa contendo RNA). A seguir, foi adicionado um volume de 500 µL de isopropanol para precipitação do RNA. A seguir, foi realizada a purificação do RNA utilizando *RNeasy Mini Kit* (Qiagen, Hilden, Germany), de acordo com as instruções do fabricante. As amostras de RNA foram quantificadas utilizando *NanoDrop 8000* (Thermo Fisher Scientific, Waltham, MA, USA) e a qualidade foi determinada utilizando *Bionalyzer 2100* (Agilent Technologies, Santa Clara, CA, USA). Após esses procedimentos, as amostras de RNA foram armazenadas em freezer -80°C.

#### **3.4.4. Transcrição Reversa (RT)**

A metodologia de RT utilizou o *cDNA SuperScript III First-Strand Synthesis SuperMix* (Invitrogen, Life Technologies, Van Allen Carlsbad, CA, USA). Resumidamente, 500ng a 1µg de RNA total foi adicionado à reação contendo 10µL de 2X First Strand Reaction mix, 2µL de *SuperScript III/RNase out Enzyme mix*, ddH<sub>2</sub>O (Sigma-Aldrich, Munich, Germany). A seguir, foram adicionados 1µL de oligodT e 1µL de *Annealing Buffer*, fornecidos com o kit. As reações foram incubadas no termociclador a 65°C, por 20 min. A seguir, foram adicionados 12 µL de Master Mix, em cada tubo e as reações incubadas a 50°C por 50 min., seguido de 85°C por 5 min. Após o término da reação, as amostras de cDNA foram imediatamente congeladas a -20°C. Controles negativos foram incluídos em cada reação de síntese de cDNA, para garantir a ausência de contaminação.

#### **3.4.5. Testes de eficiência dos iniciadores (*primers*)**

A eficiência da amplificação de sequências foi calculada a partir da construção de uma curva padrão usando diluições em série de quantidades conhecidas de cDNA. O modelo de regressão linear de Log(fluorescência) pelo número de ciclos de amplificação pela RT-QPCR foi proposto como um método para calcular a concentração inicial de cDNA e a eficiência da amplificação em cada amostra. Foram utilizadas diluições seriadas (1000, 100, 10, 1 e 0,1 µg/µL) de cDNA de dois pools, cada um composto por pelo menos 10 amostras de tecido pulmonar normal (N) e tumoral (T), respectivamente. Os pools de cDNA foram testados em triplicata, na placa de RT-QPCR. A análise dos dados utilizou os programas descritos abaixo. O protocolo utilizou mix de reação contendo 11,75µL de ddH<sub>2</sub>O livre de DNase/RNase (Sigma-Aldrich, Munich, Germany), 0,5µL do *primer forward* (10µM), 0,5µL *primer reverse* (10µM), 6,25 µL de Power SYBR Green mix (Applied Biosystems, Warrington, UK). 1µL de cDNA foi adicionado a cada reação. O equipamento utilizado foi o *7300 Real-Time PCR*

*System* (Applied Biosystems). Curvas de dissociação foram incluídas em todas as amplificações para assegurar a produção de produtos específicos. Os valores de Ct (*threshold cycle*), que constitui o número de ciclos na RT-QPCR quando uma amplificação estatisticamente significativa é detectada pela primeira vez na reação, foram exportados para posterior análise.

A eficiência da amplificação (E) durante a RT-QPCR foi calculada utilizando a fórmula  $E = 10^{(-1/\text{slope})}$ , no qual o valor do *slope* corresponde à inclinação da reta da curva de calibração, resultante da análise de variação do valor de Ct do gene alvo e do(s) controle(s) endógeno(s), em função do *log* das diferentes concentrações de cDNA na diluição seriada. O valor esperado do *slope* é de 3,32, o qual corresponde à eficiência de 100%, indicando que a quantidade de produto amplificado dobra a cada ciclo de amplificação. Utilizamos os programas *LinRegPCR* (versão 11.1, Amsterdam, The Netherlands, UE), o qual determina a quantidade do mRNA alvo amplificado e a eficiência da reação (E).

#### **3.4.6. Seleção dos genes endógenos**

Para seleção dos controles endógenos mais estáveis em tecido pulmonar normal e tumoral, tais amostras foram amplificadas para genes candidatos, selecionados a partir da literatura em NSCLC. Os candidatos a genes endógenos testados foram os seguintes: *TBP*, *GAPDH*, *B2M*, *RPS13*, *NDUFA1*, *HPRT*, *GUSB*, *ACTINB*, *HBMS*. Esses genes foram selecionados após extensa pesquisa nos dados da literatura em NSCLC, a qual identificou genes endógenos com resultados mais estáveis na maioria dos estudos publicados (**Tabela 2**), abaixo. Nessa análise, o fator de normalização foi obtido através da média geométrica dos valores de Ct dos genes controles endógenos.

**Tabela 2.** Lista de controles endógenos comumente utilizados na literatura para análise de expressão gênica em NSCLC.

<b>Símbolo do gene</b>	<b>Referências</b>
<i>TBP</i>	[107]
<i>RPII</i>	[108]
<i>TBP</i>	[109]
<i>ACTINB</i>	[110, 111]
<i>TBP, ACTINB, B2M, BAT1</i>	[112]
<i>ACTINB, PBGD, B2M</i>	[113]
<i>ACTINB</i>	[114]
<i>GAPDH</i>	[84, 115]
<i>RPS11</i>	[116]
<i>GAPDH, r18s</i>	[117]
<i>TBP, BAT1, B2M, ACTB</i>	[118-120]
<i>TBP, BAT1, ACTB</i>	[119]
<i>TBP</i>	[121]
<i>RPS13</i>	[122]
<i>NDUFA1, RPS18, RPS19</i>	[123]



### 3.4.7. Análise dos dados de genes endógenos pela RT-QPCR

A análise dos dados para seleção dos genes endógenos mais estáveis utilizou o programa LinReg PCR, descrito abaixo.

O *LinReg PCR* é um programa de análise de dados de RT-QPCR, resultantes do monitoramento da reação da PCR com corantes fluorescentes tais como o SYBR Green I. O programa usa uma linha de base de fluorescência para cada amostra, determina um valor de linearidade e em seguida utiliza uma análise de regressão linear para ajustar uma linha reta (originada a partir do valor do *slope*) a partir do conjunto de dados da RT-QPCR. A eficiência da PCR foi então calculada para cada amostra, individualmente.

Os resultados mostraram que os endógenos de maior estabilidade foram os genes de referência *RPS13* e *B2M*, os quais foram incluídos como controles nos experimentos de validação da expressão gênica.

### 3.4.8. Análise de expressão gênica

Com base nas análises de bioinformática, 10 genes desregulados em NSCLC (*ERBB3*, *KLF4*, *MUC5B*, *NOTCH1*, *PI3KR1*, *SRC*, *SATB2*, *ZEB1*, *EGFR*, *KRAS*) foram selecionados para análise pela metodologia RT-QPCR com o corante fluorescente SYBR Green I, utilizando o sistema de detecção *7300 Real-Time Detection System*. Os genes e as sequências dos *primers* utilizados estão representados no **Anexo 5**. O nível de expressão gênica foi avaliado em 32 amostras tumorais e 23 normais. Os dados de expressão foram calculados pelo método Delta Delta CT [75].

Análises estatísticas foram realizadas para correlacionar os dados de expressão gênica com os dados clínicos e histopatológicos dos pacientes.

### 3.4.9. Análise Estatística

Os dados de expressão gênica foram avaliados utilizando o teste não paramétrico de Mann Whitney, no programa *GraphPad Prism*. A significância estatística foi determinada como  $p < 0,05$ .

## 4. RESULTADOS

### 4.1. Parte I

#### 4.1.1. Sequenciamento de miRNAs (miRNA-Seq)

A análise de sequenciamento identificou mutações e alterações na expressão de miRNAs em amostras de tumores comparado com tecido pulmonar histologicamente normal.

O total de leituras (aqui definidas como *reads*) brutas obtidas no sequenciamento foi de 13.135.522 (13 milhões, cento e trinta e cinco mil e quinhentos e vinte e dois), com uma média de 547.313 *reads* por amostra. Os *reads* foram submetidos à análise de qualidade na ferramenta FastQC; o teste de qualidade mostrou que 96,1% (12.623.236,642) dos *reads* passaram no teste de qualidade, com valor de *Q-score*  $\geq 30$ .

#### 4.1.2. miRNAs com expressão alterada em adenocarcinoma

As análises foram realizadas de forma a identificar as alterações em: (1) todos os tumores comparado com os tecidos normais (amostras não pareadas, N=17 AD e 7N) e (2) nos tumores comparado com os tecidos normais dos mesmos pacientes (amostras pareadas, N=5 AD e 5N).

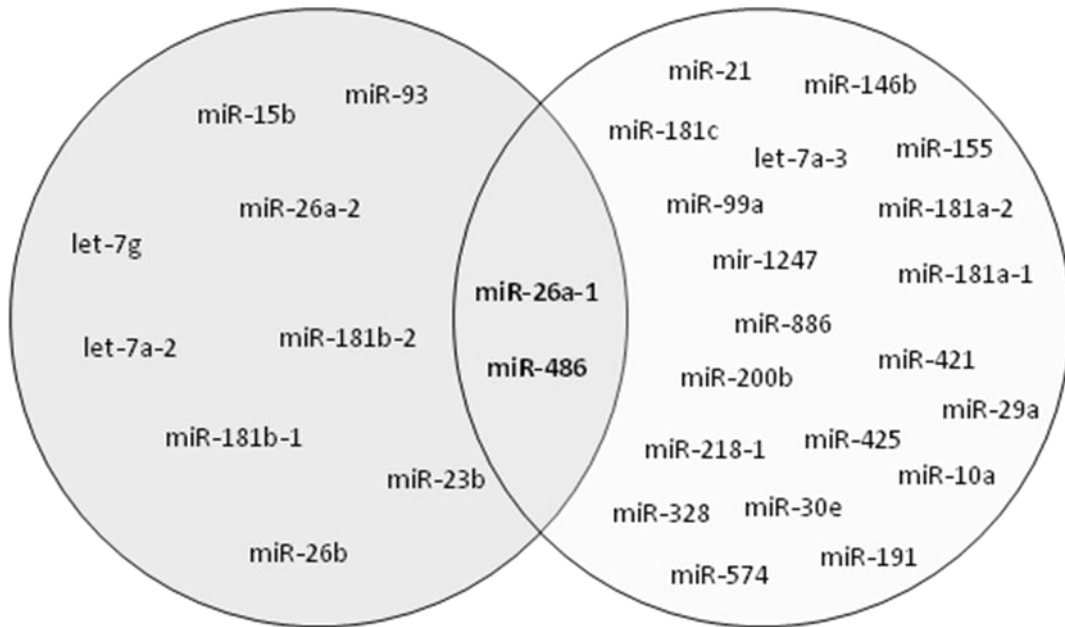
#### 4.1.3. Resultados da análise nas amostras não pareadas:

Os resultados mostraram que o miRNA miR-486 apresentou expressão diminuída e os miRNAs (let-7a-2, let-7g, miR-15b, miR-181b-1, miR-181b-2, miR-23b, miR-26a-1, miR-26a-2, miR-26b, miR-93) apresentaram expressão aumentada, estatisticamente significativa ( $p \leq 0,01$  e  $FDR \leq 0,1$ ) em AD comparado com tecidos pulmonares histologicamente normais.

#### 4.1.4. Resultados da análise nas amostras pareadas:

Essa análise identificou 22 miRNAs com expressão diferencial estatisticamente significativa ( $p \leq 0,01$  e  $FDR \leq 0,1$ ), sendo que 8 miRNAs (miR-486, miR-1247, miR-218-1, miR-181a-1, miR-181a-2, miR-328, miR-574 e miR-886) estavam com expressão diminuída e 14 miRNAs (let-7a-3, miR-146b, miR-26a-1, miR-200b, miR-191, miR-181c, miR-10a, miR-155, miR-99a, miR-30e, miR-21, miR-425, miR-29a e miR-421) estavam com expressão aumentada em AD comparado com o tecido normal do mesmo paciente. Interessantemente, expressão desregulada dos miRNAs (miR-486 e miR-26a-1) foi detectada em ambas as análises (amostras pareadas e não-pareadas) considerando os miRNAs com  $p \leq 0,01$  e  $FDR \leq 0,1$ .

Os miRNAs diferencialmente expressos em amostras pareadas e não pareadas estão ilustrados na **Figura 3**, abaixo e descritos na **Tabela 3**.



**Figura 3.** Diagrama de Venn ilustrando os miRNAs diferencialmente expressos nos grupos de amostras não-pareadas (cinza escuro) e pareadas (cinza claro). miRNAs incluídos nessa figura apresentaram  $p \leq 0,01$  e  $FDR \leq 0,1$ .

**Tabela 3.** miRNAs diferencialmente expressos em adenocarcinoma comparado com tecido normal, conforme dados obtidos por miRNA-Seq.

<b>miRNA</b>	<b>LogFC</b>	<b>P-value</b>	<b>FDR</b>
<b>Amostras não-pareadas</b>			
miR-486	-1,635	0,0004	0,0689
let-7g	0,8879	0,0015	0,0689
miR-15b	0,9831	0,0004	0,0689
miR-181b-2	0,9887	0,0009	0,0689
miR-23b	0,9479	0,0014	0,0689
miR-26a-1	0,9756	0,0014	0,0689
miR-26a-2	0,9595	0,0013	0,0689
miR-26b	1,0565	0,0009	0,0689
miR-93	1,203	0,0005	0,0689
let-7a-2	0,8844	0,0025	0,1052
miR-181b-1	0,9221	0,0027	0,1052
<b>Amostras Pareadas</b>			
miR-486	-2,5430	8,8276.E-12	1,805.E-09
miR-1247	-2,6582	3,2911.E-08	3,3651.E-06
miR-218-1	-1,6175	0,0001	0,0043
miR-181a-2	-1,1145	0,0002	0,0079
miR-328	-1,4708	0,0004	0,0108
miR-181a-1	-1,1444	0,0006	0,0147
miR-574	-1,1719	0,0036	0,0503
miR-886	-1,3187	0,0046	0,0575
let-7a-3	0,4856	3,2638.E-264	5,1692.E-262
miR-146b	0,7012	0,0005	0,0147
miR-26a-1	0,7382	0,0043	0,0551
miR-200b	1,0898	0,0072	0,0864
miR-191	1,0935	0,0018	0,0307
miR-181c	1,1230	0,0041	0,0551
miR-10a	1,2858	0,0009	0,0208
miR-155	1,3089	0,0029	0,0418
miR-99a	1,4447	0,0028	0,0418
miR-30e	1,4926	0,00002	0,0011
miR-21	1,5219	0,0028	0,0418
miR-425	1,6206	7,6888.E-06	0,0005
miR-29a	1,6488	0,0001	0,0043
miR-421	1,7648	0,0085	0,0993

Amostras não-pareadas (N=17 AD e 7 N); amostras pareadas (N=5 AD e 5 N). LogFC = log de *fold change* (valores negativos indicam expressão diminuída e valores positivos indicam expressão aumentada).

FDR = *false discovery rate*. FDR≤0,1; P≤0,01.

#### 4.1.5. Validação de alterações na expressão de miRNAs

Todos os miRNAs com expressão significativamente aumentada ou diminuída ( $p \leq 0,01$  e  $FDR \leq 0,1$ ) (33 miRNAs, descritos na **Tabela 3**) foram comparados com dados de expressão dos mesmos miRNAs obtidos utilizando ensaios *Taqman Low Density Arrays* (TLDA, Life Technologies). Os ensaios TLDA foram realizados em um conjunto independente de 22 amostras de AD e 12 amostras de tecido pulmonar normal (3 *pools* de 4 amostras cada).

Destes 33 miRNAs, 14 miRNAs estavam com nível de expressão concordante em ambas as análises (miRNA-Seq e TLDA). Destes 14 miRNAs, 5 miRNAs estavam com expressão diminuída (miR-486, miR-181a-1, miR-181a-2, miR-218-1 e miR-886) e 9 estavam com expressão aumentada (let-7a-2, let-7a-3, let-7g, miR-15b, miR-26b, miR-200b, miR-155, miR-21, miR-425). Os miRNAs let-7a-2, let-7a-3, miR-15b, miR-200b, miR-155 e miR-21 foram considerados validados, pois apresentaram diferença estatisticamente significativa ( $p \leq 0,01$  e  $FDR \leq 0,1$ ) no nível de expressão entre AD e N em ambas as análises (**Tabela 4**).

**Tabela 4.** miRNAs com expressão concordante nas análises de miRNA-Seq e TLDA.

miRNA-Seq			TLDA		
miRNA	logFC	<i>P-value</i>	miRNA	logFC	<i>P-value</i>
let-7a-2	0,8844	0,0025	let-7a	2,3601	<b>0,000*</b>
let-7a-3	0,4856	3,2638.E-264	let-7a	2,3601	<b>0,000*</b>
miR-15b	0,9831	0,0004	miR-15b	2,4022	<b>0,000*</b>
miR-200b	1,0898	0,0072	miR-200b	1,9404	<b>0,000*</b>
miR-21	1,5219	0,0028	miR-21	2,6311	<b>0,001*</b>
miR-155	1,3089	0,0029	miR-155	1,3294	<b>0,010*</b>
miR-486	-1,635	0,0004	miR-486	-3,9434	0,084
miR-181a-1	-1,1444	0,0006	miR-181a	-0,8625	0,140
miR-181a-2	-1,1145	0,0002	miR-181a	-0,8625	0,140
let-7g	0,8879	0,0015	let-7g	0,3918	0,143
miR-26b	1,0565	0,0009	miR-26b	0,5597	0,224
miR-218-1	-1,6175	0,0001	miR-218	-1,4699	0,386
miR-886	-1,3187	0,0046	miR-886-3p	-1,1746	0,482
miR-425	1,6206	7,6888.E-06	miR-425-5p	0,1190	0,906

Grupo teste: valor de  $P \leq 0,01$  e  $FDR \leq 0,1$ , conforme determinado pelo programa edgeR;

Grupo de validação: \* $P \leq 0,01$ , conforme determinado pelo programa *Expression Suite*.

#### 4.1.6. miRNAs com variações em nucleotídeos únicos

Os resultados das análises de miRNA-Seq mostraram a ocorrência de variações em nucleotídeos únicos (SNVs). SNVs foram identificadas em miRNAs, comparado ao genoma humano de referência (*National Center for Biotechnology Information build 37*, hg19).

Foram identificadas 30 SNVs em AD e 19 SNVs em tecido pulmonar normal adjacente ao tumor (N). As SNVs identificadas em miRNAs em amostras de AD e N estão descritas nas **Tabelas 5 e 6**, respectivamente.

**Tabela 5.** SNVs identificadas em miRNAs em adenocarcinoma.

Chr	Início	Fim	Ref	Alt	Localização	Gene
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	Intergenic	mir-4490
Chr21	9827315	9827315	T	C	Intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	let-7c
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr1	16861827	16861827	G	A	Intergenic	mir-3675
Chr11	122017274	122017274	T	C	ncRNA_exonic	let-7a-2
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-24-1
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-3074
Chr12	54731084	54731084	C	A	ncRNA_exonic	mir-148b
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200a
Chr17	46657243	46657243	A	G	ncRNA_exonic	mir-10a
Chr1	56633037	56633037	C	T	Intergenic	mir-4422
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr1	186105018	186105018	G	A	ncRNA_exonic	mir-548f-1
Chr1	186105020	186105020	G	A	ncRNA_exonic	mir-548f-1
Chr2	227230043	227230043	A	G	Intergenic	mir-5702
Chr8	143215944	143215944	G	C	Intergenic	mir-4472-1
Chr9	97847569	97847569	C	T	ncRNA_exonic	mir-23b
Chr14	53927750	53927750	T	G	Intergenic	mir-5580
Chr11	122022981	122022981	A	T	ncRNA_exonic	mir-100
Chr21	17911436	17911436	T	A	ncRNA_exonic	mir-99a
Chr3	48261476	48261476	C	A	Intergenic	mir-4443
Chr19	46522264	46522264	T	A	ncRNA_exonic	mir-769

Análises de bioinformática: *SAM tools*, *VCF tools* e *ANNOVAR*.

**Tabela 6.** SNVs identificadas em miRNAs em tecido normal.

<b>Chr</b>	<b>Início</b>	<b>Fim</b>	<b>Ref</b>	<b>Alt</b>	<b>Localização</b>	<b>Gene</b>
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200a
Chr4	10917451	10917451	T	G	intergenic	mir-572
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	intergenic	mir-4490
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr21	9827315	9827315	T	C	intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	let-7c
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr11	122017274	122017274	T	C	ncRNA_exonic	let-7a-2
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195896	52195896	C	A	ncRNA_exonic	mir-99b
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr8	143215944	143215944	G	C	intergenic	mir-4472-1
Chr9	96938260	96938260	A	G	ncRNA_exonic	let-7a-1
Chr9	96938262	96938262	A	G	ncRNA_exonic	let-7a-1

Análises de bioinformática: *SAM tools*, *VCF tools* e *ANNOVAR*.



Quanto à distribuição global de SNVs em miRNAs, observamos uma maior frequência de SNVs nas amostras de tumor comparado com os tecidos de pulmão normal. A ocorrência de todas as SNVs em AD e N está detalhada no **Anexo 6**. Do total de miRNAs com SNVs, 5 miRNAs (miR-24-1, miR-3074, miR-4422, miR-125b-2 e miR-150) apresentaram uma maior frequência de SNVs em AD. Os miRNAs miR-24-1, miR-3074 e miR-4422 apresentaram SNVs (A/T, A/T e C/T) exclusivamente no tumor, sendo observadas em 7/17, 7/17 e 3/17 tumores, respectivamente.

Adicionalmente, foram identificadas SNVs (T/A e G/T) nos miRNAs miR-125b-2 e miR-150, em 11/17 (65%) e 12/17 (70%) tumores, sendo que SNVs foram estatisticamente significativas no miR-150 em AD comparado com tecido N ( $P=0,0233$ , teste exato de Fisher).

SNVs nos miRNAs miR-125b-2 e miR-150 estavam presentes em ambas as amostras de tumor e normal em um caso (52T e 52N), sugerindo que essas SNVs podem representar um polimorfismo herdado.

As SNVs mais frequentes nos miRNAs (miR-24-1, miR-3074, miR-4422, miR-125b-2 e miR-150), foram comparadas com os dados clínicos e histopatológicos dos pacientes. Os resultados indicaram que a SNV T/A no miR-125b-2, foi correlacionada com o sexo, ocorrendo com maior frequência no sexo feminino 9/17 (52,94%) em relação ao sexo masculino 2/17 (11,76%) ( $p=0,0295$ ). As SNVs nos miRNAs miR-24-1 e miR-3074 estavam associadas com estadiamentos iniciais I e II do tumor (**Tabela 7**). Os outros parâmetros clínicos e histopatológicos analisados não mostraram associação estatisticamente significativa com a ocorrência de SNVs em miRNAs.

**Tabela 7.** Correlação de SNVs em miRNAs e dados clínicos e histopatológicos dos pacientes com AD.

SNVs	miR-24-1		miR-3074	
	0	1	0	1
<b>Estadiamento tumoral</b>				
<b>IA</b>	2 11,76	3 17,65	2 11,76	3 17,65
<b>IB</b>	6 35,29	0 0	6 35,29	0 0
<b>IIA</b>	0 0	2 11,76	0 0	2 11,76
<b>IIB</b>	0 0	1 5,88	0 0	1 5,88
<b>IIIA</b>	1 5,88	1 5,88	1 5,88	1 5,88
<b>IIIB</b>	1 5,88	0 0	1 5,88	0 0
<b><i>p-value</i></b>	0,0295		0,0295	

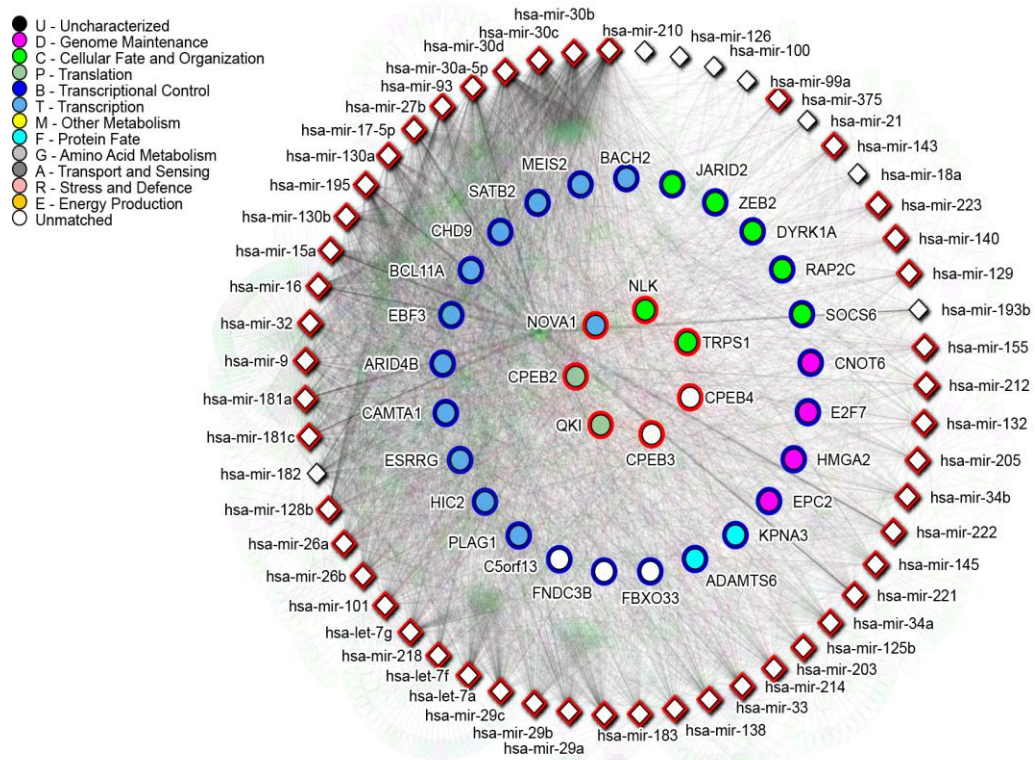
1 = presença ou 0 = ausência de SNVs nas sequências de miRNAs. Valor de *p* determinado de acordo com o teste exato de Fisher.

## 4.2. Parte II

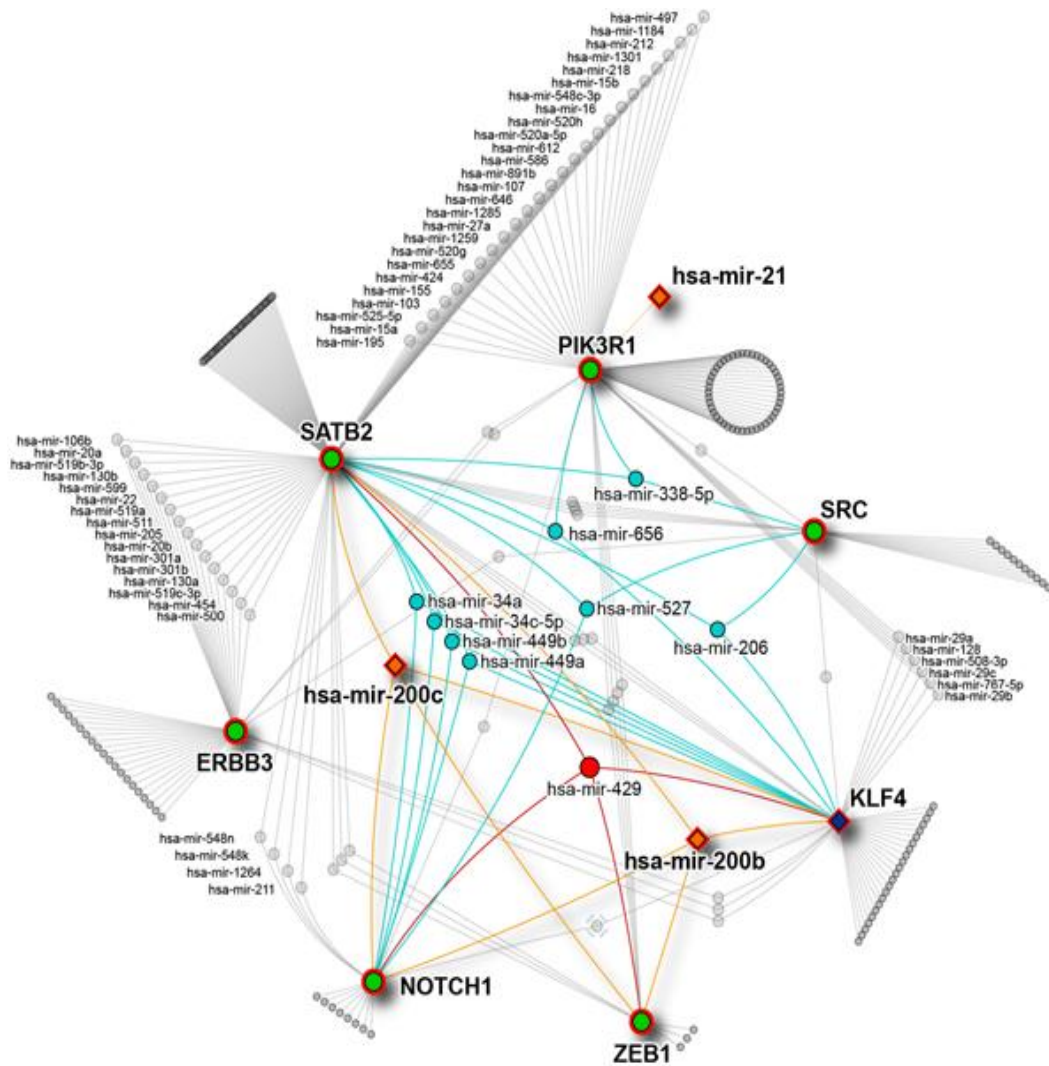
### 4.2.1. miRNAs desregulados em NSCLC: meta-análise de dados da literatura

Os resultados da meta-análise de 19 estudos previamente publicados identificaram 103 miRNAs desregulados em 901 tumores comparados com 776 tecidos pulmonares histologicamente normais. 20/103 miRNAs estavam com expressão aumentada ou diminuída em pelo menos 3 estudos diferentes: miR-21, família do miR-30 (miR-30a, b, c, d), miR-31, miR-126, miR-138, miR-139, miR-141, miR-143, miR-145, miR-155, miR-182, miR-200b, miR-200c, miR-205, miR-210, miR-429 e miR-451.

Os dados da análise de predição mostraram vários genes-alvo potencialmente regulados pelos miRNAs identificados, os quais codificam proteínas com funções biológicas importantes na tumorigênese, tais como controle da transcrição, tradução, morte celular, entre outras (**Figura 4**). Interessantemente, os genes-alvo preditos, na sua maioria, também estão frequentemente desregulados em câncer, incluindo NSCLC. Avaliamos o nível de expressão de 10 genes (*ERBB3*, *PIK3R1*, *KLF4*, *NOTCH1*, *MUC5B*, *SATB2*, *SRC*, *ZEB1*, *EGFR* e *KRAS*) em 34 amostras de NSCLC e 26 amostras pulmonares normais, pela RT-QPCR. Os genes avaliados foram selecionados a partir da rede de interação proteica (**Figura 5**).



**Figura 4.** Análise de interação entre microRNAs e genes-alvo em NSCLC. Rede de interação proteína-proteína e os 103 miRNAs desregulados em NSCLC comparado com pulmão normal. A rede de interação foi construída utilizando o programa I2D - *Interologous Interaction Database* e visualizada no programa NAViGaTOR. As anotações funcionais estão destacadas na figura, pela cor, para cada interação proteína-proteína.



**Figura 5.** Rede de interação entre miRNAs e proteínas codificadas por genes frequentemente alterados em NSCLC. Os miRNAs que potencialmente modulam a expressão de genes específicos são mostrados na rede de interação proteica. Os genes-alvo avaliados por RT-QPCR apresentam papel central na rede, conforme evidenciado pelo nó central interligado a outras proteínas, bem como função biológica importante no câncer.

#### 4.2.2. Expressão gênica desregulada em NSCLC

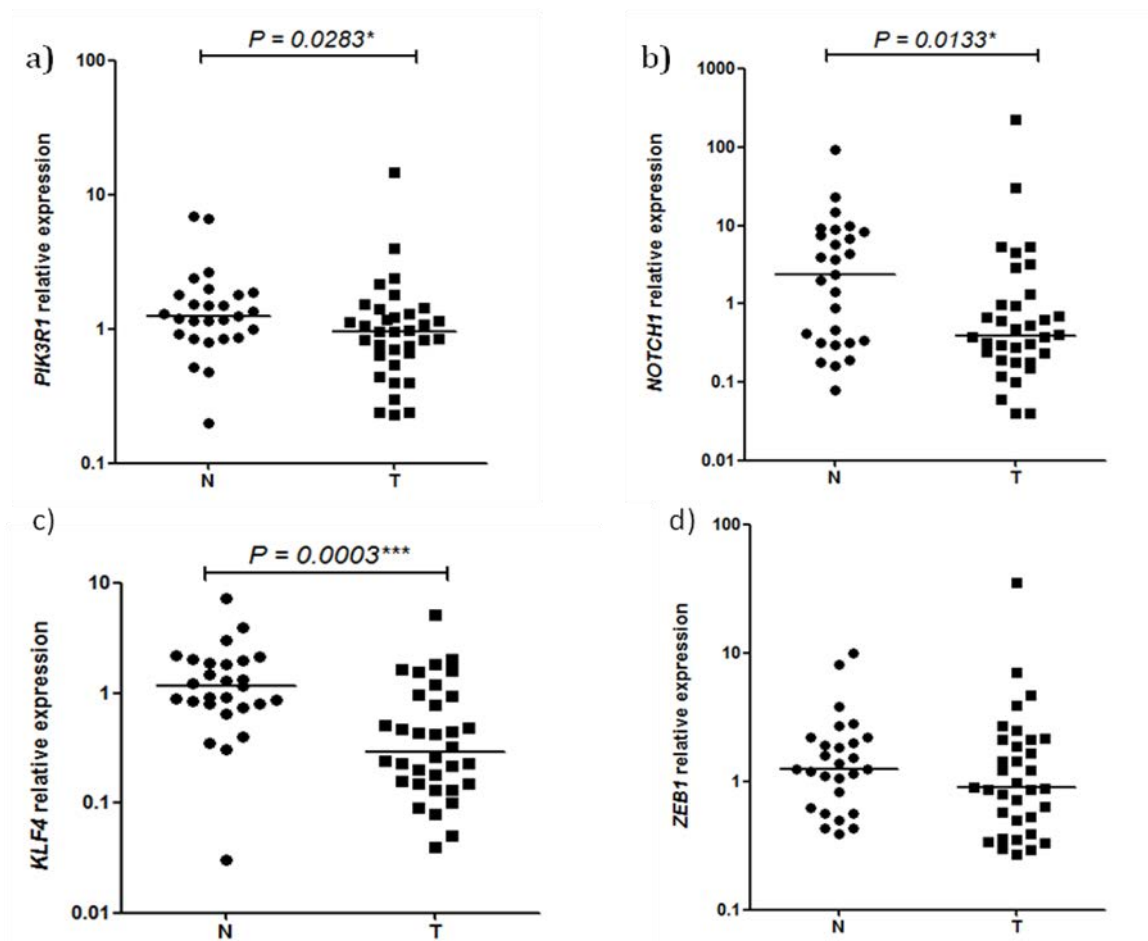
Os resultados confirmaram que todos os genes estavam frequentemente desregulados (com expressão aumentada ou diminuída) em tecido primário de NSCLC, comparado com tecido pulmonar normal. Os resultados de expressão gênica estão detalhados na **Tabela 8**, abaixo.

**Tabela 8.** Expressão gênica desregulada em NSCLC vs. normal conforme determinado por RT-QPCR.

	<i>ERBB3</i>	<i>PIK3R1</i>	<i>KLF4</i>	<i>NOTCH1</i>	<i>MUC5B</i>	<i>SATB2</i>	<i>SRC</i>	<i>ZEB1</i>	<i>EGFR</i>	<i>KRAS</i>
<b>Mediana</b>										
Normal	1,87	1,25	1,15	2,33	2,98	2,37	1,80	1,26	0,96	1,05
Tumor	2,04	0,96	0,29	0,39	3,83	1,13	0,84	0,90	1,16	0,87
<b>P-value</b>	0,9710	<b>0,0283*</b>	<b>0,0003*</b>	<b>0,0133*</b>	0,3876	0,9595	0,0610	0,2039	0,4589	0,4161

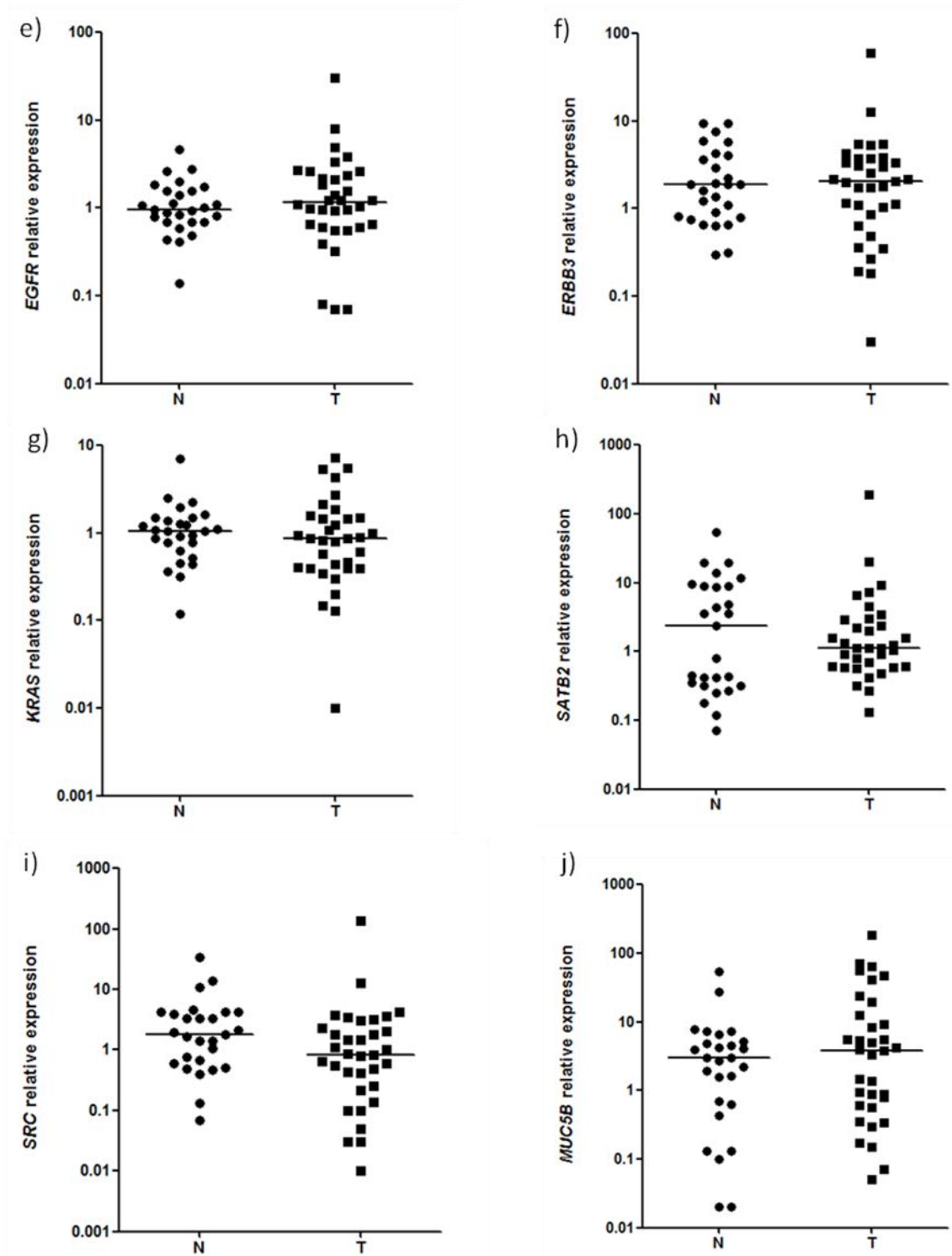
Mann Whitney test e intervalo de confiança (CI), *GraphPad Prism software* \*(p<0,05)

As **Figuras 6 e 7** mostram os níveis de expressão gênica ( $\log_{10}$ ) em tumor (T) comparado com normal (N). Dos 10 genes analisados, 3 genes (*PIK3R1*, *NOTCH1*, *KLF4*) mostraram diferença estatisticamente significativa ( $p < 0,05$ ) no nível da expressão entre tumor e normal (**Figure 6a-c**). Os genes *ZEB1*, *EGFR*, *ERBB3*, *KRAS*, *SATB2*, *SRC* e *MUC5B*, também mostraram expressão gênica desregulada no tumor, entretanto não houve diferença estatisticamente significativa entre tumor e normal (**Figure 6d e Figure 7e-j**).



**Figura 6.** Os gráficos mostram a comparação entre o nível de expressão gênica ( $\text{Log}_{10}$ ) no tumor (T) vs. normal (N). Variações em níveis de expressão gênica detectados são mostradas (Gráficos 3a-c) ( $p < 0,05$ , Mann Whitney test, GraphPad Prism software). O Gráfico 3d não mostrou valor estatisticamente significativo.

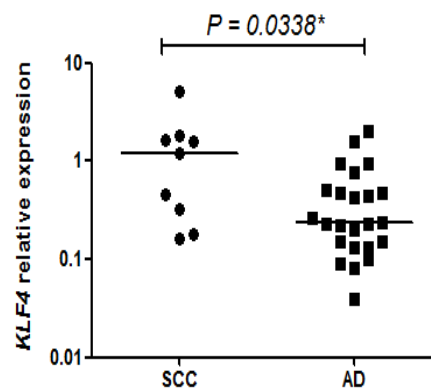




**Figura 7.** Os gráficos mostram a comparação entre o nível de expressão gênica (Log<sub>10</sub>) no tumor (T) vs. normal (N). Os Gráficos 3e-j não mostraram valor estatisticamente significativo ( $p < 0,05$ , Mann Whitney test, GraphPad Prism software).

A baixa expressão do *KLF4* foi estatisticamente correlacionada com o grau histológico avançado do tumor (Mediana: BD-MD=0,235 vs. PD=0,855,  $p=0,0371$ ).

Adicionalmente, os dados de expressão gênica foram analisados de acordo com o tipo histológico (AD vs. SCC). O único gene que mostrou diferença estatística significativa no nível de expressão entre AD e SCC foi o *KLF4* (Mediana: SSC=1,2 vs. AD=0,235,  $p=0,0338$ ) (**Figura 8**).



**Figura 8.** O gráfico mostra o nível de expressão gênica de acordo com o subtipo histológico, adenocarcinoma (AD) vs. carcinoma de células escamosas (SCC). As margens das diferenças estatisticamente significantes são mostradas ( $p < 0,05$ , Mann Whitney test, GraphPad Prism software).

Os dados de expressão gênica em amostras tumorais vs normais também foram analisados separadamente nos subtipos histológicos AD e SCC. Os resultados mostraram que ambos os genes *KLF4* e *NOTCH1* apresentaram expressão diminuída nas amostras de AD vs. normal (*KLF4*, Mediana: AD=0,235 vs. N=0,975,  $p=0,0022$ ; *NOTCH1*, Mediana: AD=0,435 vs. N=4,02,  $p=0,0041$ ). Em SCC, o gene *SATB2* estava com alta expressão comparado com normal (Mediana: SCC=3,40 vs. N=0,42,  $p=0,0216$ ).

O gene *KRAS* estava com expressão diminuída em tumores de pacientes com metástase em linfonodos (Mediana: N0=1,03 vs. N1/N2=0,535,  $p=0,0406$ ). O gene *MUC5B* estava com expressão aumentada no tumor de pacientes sem metástase em linfonodos (Mediana: N0=5,53 vs. N1-N2=1,17,  $p=0,0314$ ). Sugerimos que o gene *K-RAS* pode desempenhar um papel associado à progressão e o gene *MUC5B* pode estar preferencialmente envolvido no desenvolvimento, mas não na progressão da doença. Os outros parâmetros clínicos e histopatológicos analisados não mostraram associações estatisticamente significativas com a expressão gênica.

## 5. DISCUSSÃO

### 5.1. Parte I

Alterações na expressão de miRNAs levam à desregulação da expressão de genes-alvo, sendo que as vias moleculares alteradas podem ter papéis importantes no desenvolvimento e progressão tumoral. Os miRNAs identificados com expressão diminuída: miR-486, miR-181a-1, miR-181a-2, miR-218-1, miR-886 e com expressão aumentada: let-7a-2, let-7a-3, let-7g, miR-15b, miR-26b, miR-200b, miR-155, miR-21, miR-425 regulam, simultaneamente, um grande número de genes com funções celulares importantes no câncer.

Similar aos nossos achados, a família do miRNA let-7 (let-7a, b, c, d, e, f, g, i) foi relatada com expressão diferencial estatisticamente significativa entre os grupos histológicos de NSCLC, sendo altamente expressa em AD pulmonar [124]. Outros estudos indicam a diminuição da expressão de alguns membros da família let-7 em linhagens celulares de NSCLC associada a baixa sobrevida [90, 93]. A expressão diminuída, principalmente do let-7a-2 e let-7f-1 foi relatada correlacionada com a menor sobrevida de pacientes com adenocarcinoma. Entretanto, um subgrupo de adenocarcinomas apresentava expressão aumentada de alguns membros da família let-7, como o let-7a-1, let-7a-2 e let-7a-3 [90].

Os miRNAs também têm sido explorados como marcadores circulantes devido à estabilidade e fácil detecção dessas moléculas na circulação sanguínea [125]. Nosso estudo identificou a expressão desregulada de alguns miRNAs, tais como miR-15b e miR-146b, os quais já foram detectados com expressão anormal no sangue de pacientes com NSCLC. Em particular, a expressão aumentada do miR-146b foi detectada em amostras de soro de pacientes com AD pulmonar [126]. Por outro lado, a expressão diminuída do miR-146b foi observada, independente da histologia (SCC ou AD) e do estadiamento tumoral [127].

Um estudo de fase I/II realizou um *screening* em amostras de soro de pacientes com NSCLC comparado com indivíduos saudáveis e identificou a expressão diferencial do miR-15b/27b em pacientes com NSCLC comparado com indivíduos saudáveis [128]. A expressão aumentada de outra assinatura de miRNAs; miR-15b/16-2 induziu a expressão de genes envolvidos nos mecanismos de sinalização da proteína ATM e RAD3, ciclo celular, reparo ao DNA e apoptose; esse estudo demonstrou que o miR-15b tem como alvo direto o *PPM1D*, também denominado *WIP1* (*wild-type p53-induced phosphatase 1*), reduzindo sua expressão gênica e protéica em células pulmonares. Os autores sugeriram que a indução de resposta a danos ao DNA pelo miR-15b deve ser parcialmente dependente da supressão de *PPM1D* e pode representar um alvo terapêutico potencial em doenças pulmonares induzidas pela

exposição à radiação [129]. *PPMID* codifica uma serina/treonina fosfatase que desempenha um papel na desfosforilação de várias proteínas de resposta a danos no DNA, tais como a ATM, ATR, p38MAPK, CHK1 e CHK2 [130].

A expressão diminuída do miR-486 e a expressão aumentada do miR-21, miR-200b e miR-375 foi observada em amostras de tecido e de escarro de pacientes com estadios I de AD pulmonar; nesse estudo, a combinação de expressão desregulada dos 4 miRNAs permitiu distinguir pacientes com AD e SCC dos indivíduos sem doença, sendo que houve maior sensibilidade e especificidade da expressão dessa assinatura de 4 miRNAs em AD, sugerindo que esse grupo de miRNAs pode ser útil para a detecção precoce de AD pulmonar [94]. Concordante com os dados da literatura, a expressão aumentada do miR-21 e do miR-200b foi estatisticamente significativa e considerada válida em nossas amostras de adenocarcinoma quando comparado com tecido pulmonar histologicamente normal.

O efeito de sensibilidade ou resistência a tratamentos como a quimioterapia também tem sido associado a perfis de expressão desregulados de miRNAs. A expressão aumentada do miR-21 e a expressão diminuída da proteína PTEN, foi correlacionada com baixa sobrevida e baixa sensibilidade ao inibidor TKI em pacientes com NSCLC tratados com Gefitinibe ou Erlotinibe. Da mesma forma, a expressão aumentada do miR-21 e a inibição do PTEN foi observada em linhagens resistentes ao Gefitinibe. Ensaios *in vitro* usando miR-21 mimics mostraram uma redução da sensibilidade ao Gefitinibe em linhagens celulares sensíveis ao TKI pela inibição do PTEN e ativação de AKT/ERK, enquanto que a inibição do miR-21 resultou em aumento da expressão do PTEN e inativação da via AKT/ERK, restaurando a sensibilidade ao tratamento. Os autores sugeriram que a alteração na expressão miR-21/PTEN pode explicar a resistência ao tratamento com TKI em NSCLC [131].

Os miRNAs também regulam vias importantes no processo metastático, tal como a via de sinalização TGF- $\beta$ , a qual ativa múltiplos fatores de transcrição indutores do programa EMT (transição epitélio mesênquima). A família do miR-200 apresenta importante papel no processo de EMT, por meio da regulação de repressores de transcrição, tais como *CDHI* (E-caderina), *ZEB1* e *ZEB2* (*Zinc finger E-box binding homeobox*) [132]. Ensaios *in vivo* e *in vitro* mostraram que o miR-200b foi capaz de inibir a expressão de *ZEB2* em nível de mRNA e proteína, e de *CDHI* em câncer de mama. *ZEB2* é um repressor transcricional e participa da regulação da via de sinalização do TGF- $\beta$  e do processo de EMT pela regulação de *CDHI* [133]. Entretanto, outros estudos mostraram que, apesar da modulação de *ZEB1/CDHI* pelo miR-200b, a restauração de *ZEB1* em linhagem de câncer de mama positiva para a expressão do miR-200b não foi capaz de alterar o potencial metastático, sugerindo uma possível

regulação de outros genes-alvo que contribuem para o processo de metástase, como por exemplo genes associados à actina [134].

Outra via molecular relevante na progressão tumoral, a via de sinalização Wnt/ $\beta$ -catenina, foi associada à regulação do miR-155. A sinalização desregulada da via Wnt/ $\beta$ -catenina promove o crescimento celular e alterações na expressão de Wnt/ $\beta$ -catenina foram associadas ao pior prognóstico de pacientes com NSCLC [135, 136]. O miR-155 apresenta importante papel na biologia do câncer, sendo que validamos o aumento da expressão desse miRNA em AD. Outro estudo também detectou expressão aumentada do miR-155 em amostras de tumor primário de pacientes com AD [93] e em soro de pacientes com estadiamento avançado (IV) de NSCLC [137], em ambos os estudos essa alteração foi correlacionada com baixa sobrevida dos pacientes [93, 137]. A expressão aumentada do miR-155 também foi identificada em lipossarcoma; ensaios *in vitro* e *in vivo* mostraram que o CK1 $\alpha$  (caseína quinase 1 $\alpha$ ), um regulador chave da via Wnt/ $\beta$ -catenina, é um alvo direto do miR-155, induzindo a sinalização da  $\beta$ -catenina e a expressão da ciclina D1, resultando na proliferação celular e progressão tumoral [138]. Interessantemente, o miR-155 tem sido estudado como alvo terapêutico; ensaios *in vitro* e *in vivo* mostraram a entrega eficiente do anti-miR-155 utilizando nanopartículas (*lactosylated gramicidin-containing lipid nanoparticles*) em linhagem celular de carcinoma hepatocelular e em camundongos [139].

Entre os miRNAs com expressão diminuída em AD, identificamos os miRNAs miR-181a-1 e miR-181a-2. Outro estudo relatou a expressão diminuída do miR-181a (embora não especificou as variantes 1 e 2) em NSCLC associada a baixa sobrevida dos pacientes, independentemente do estadiamento tumoral e da presença de metastases em linfonodos [80]. Em células de leucemia linfocítica crônica, o miR-181a/b mostrou níveis baixos de expressão em pacientes com pior prognóstico e associado a menor sobrevida global dos pacientes [140].

Identificamos um nível de expressão significativamente aumentado do miR-425 em AD, concordante com estudos prévios em vários tipos de câncer, como por exemplo, câncer gástrico [141]. Entretanto, o significado clínico dessa alteração em AD ainda não está bem estabelecido. Recentemente, foi observado o envolvimento do miR-425 na via de sinalização NF-kappaB (*nuclear factor kappa B*) [141], ativada por citocinas pró-inflamatórias como IL-1 $\beta$  [142, 143], a qual, por sua vez, pode inibir a expressão do supressor tumoral PTEN, frequentemente desregulado em câncer [144, 145]. Ensaios *in vitro*, utilizando linhagem de câncer gástrico humano, mostraram que células que expressam o miR-425 apresentaram diminuição da expressão do PTEN em nível de mRNA e de proteína, promovendo a sobrevivência de células IL-1 $\beta$ -induzidas. Adicionalmente, foi observado que a inibição do

PTEN diminuiu a ativação da caspase-3, inibindo o processo de apoptose dessas células. A transfecção de anti-miR-425 aumentou significativamente a ativação da caspase-3 e a apoptose em resposta ao tratamento com IL-1 $\beta$ . Interessantemente, ensaios *in vivo* mostraram que tumores tratados com anti-miR-425 apresentaram um aumento no nível da proteína PTEN, além de ocorrer redução no tamanho do tumor [141].

Estudo prévio mostrou que o miR-218 estava com expressão diminuída em uma linhagem celular de NSCLC, corroborando nossos dados em adenocarcinoma primário. O miR-218 regula negativamente o HMGB1 (*high mobility group box-1*) em nível de mRNA e proteína e inibe a migração e invasão celular, sendo sugerido pelos autores como potencial biomarcador em carcinoma pulmonar metastático [146]. A expressão aumentada do gene HMGB1 promove invasão celular pela regulação da *MMP-9 (matrix metallo-peptidase 9)* em câncer de pulmão. A indução da expressão de MMP-9 pelo HMGB1, ocorre pela ativação das vias de sinalização PI3K/Akt e NF-kB [147]. Por outro lado, o nível de expressão do miR-218 estava aumentado em linhagem de câncer de mama e metástases ósseas de carcinoma mamário [148]. O miR-218 atua na via de sinalização do Wnt, promovendo a diferenciação de osteoblastos; além disso o miR-218 influencia a progressão tumoral, contribuindo para a formação de metástases ósseas [148].

O miR-886 foi caracterizado como supressor tumoral em carcinoma de células escamosas do esôfago [149] e de carcinoma pulmonar de células pequenas (SCLC) [150]. A baixa expressão da forma precursora desse miRNA (pré-miR-886) em carcinoma de células escamosas do esôfago foi associada com baixa sobrevida livre de recorrência dos pacientes [149]. A perda da expressão do miR-886-3p foi correlacionada com hipermetilação da região promotora que codifica esse miRNA e estava significativamente associada com menor sobrevida apresentando um valor prognóstico em SCLC. Ensaios *in vitro* e *in vivo* mostraram que a indução da expressão do miR-886 (utilizando a estratégia de miR *mimics*) resultou na supressão da proliferação, migração e invasão celular [150].

Recentemente, foi demonstrado o envolvimento de miRNAs, particularmente o miR-328 em células tumorais progenitoras. Interessantemente, o miR-328 estava com expressão significativamente diminuída em tecido tumoral de pacientes com carcinoma colorretal e especificamente em células tumorais progenitoras (população lateral ou *side population*, SP) comparado com células não progenitoras. Células da população lateral ou SP foram descritas como um subconjunto de células que expressa transportadores ABC e possuem um fenótipo similar a células-tronco ou células tumorais progenitoras. Adicionalmente, a expressão do miR-328 inibiu a resistência a drogas e a capacidade de invasão das células tumorais

progenitoras (população SP). Esse estudo demonstrou a modulação da expressão gênica e protéica dos genes *ABCG2* e *MMP16* pelo miR-328 nas células tumorais progenitoras. A expressão do gene *ABCG2*, o qual codifica uma proteína de transporte por meio de ligação ao ATP, tem sido utilizada para definir a população SP de neoplasias humanas e está com expressão aumentada em glioblastoma e associado à resistência a múltiplas drogas, sendo que *ABCG2* é regulado pelo miR-328 [151]. Outro estudo também demonstrou que a expressão do miR-328 estava significativamente diminuída em gliomas de alto grau e associada a baixa sobrevida de pacientes com glioblastoma [152]. Concordante com esses achados, nossos dados de miRNA-Seq mostraram uma redução significativa do miR-328 em células tumorais de AD comparado com tecido pulmonar histologicamente normal, entretanto essa alteração não foi validada em amostras de tecido emblocado em parafina. Acreditamos que a análise detalhada da população SP nesses casos é válida para confirmar o envolvimento do miR-328 na tumorigênese do AD pulmonar.

Ao contrário dos nossos dados, outro estudo relatou níveis aumentados de expressão do miR-328 em tumor primário e metástase cerebral de pacientes com NSCLC, além de perda de adesão e aumento da migração celular detectadas em ensaios *in vitro* [153]. Entretanto, esse estudo utilizou um número limitado de amostras de pacientes com diversos tipos histológicos de NSCLC (AD, SCC e carcinoma de grandes células ou LCC) e os estudos funcionais foram realizados em duas linhagens celulares, sendo uma de AD e outra de SCC. Os autores relatam que estudos de validação com um maior número de amostras são necessários para elucidar o envolvimento do miR-328 em NSCLC.

Apesar dos estudos supracitados, ensaios funcionais são necessários para elucidar o papel dos miRNAs aqui validados (*let-7a-2*, *let-7a-3*, *miR-15b*, *miR-21*, *miR-155*, *miR-200b*) na modulação de vias moleculares associadas ao desenvolvimento e progressão do adenocarcinoma pulmonar.

## 5.2. Parte-II

Os genes-alvos regulados por miRNAs desregulados em NSCLC (*ERBB3*, *PIK3R1*, *KLF4*, *NOTCH1*, *MUC5B*, *SATB2*, *SRC*, *ZEB1*, *EGFR* e *KRAS*) apresentam papel na organização e destino celulares, como evidenciado nas funções biológicas identificadas na rede de interação protéica.

Vias de sinalização envolvendo os genes *NOTCH1*, *ZEB1*, *EGFR* e *ERBB3*, estão implicadas na modulação do mecanismo de transição epitélio-mesênquima (EMT) em células tumorais, resultando na perda de adesão celular e conseqüentemente invasão e metástase tumoral [157, 158]. Vias moleculares importantes no processo metastático, tal como a via de sinalização TGF- $\beta$ , ativam múltiplos fatores de transcrição indutores do programa EMT. A via *NOTCH*, depende da ativação do TGF- $\beta$  para induzir a EMT. Fatores de transcrição, como o *ZEB1* (*Zinc finger E-box binding homeobox 1*), reprime a transcrição de genes epiteliais, como a *E-caderina* [159] e induzem o fenótipo de EMT. Paralelamente, a ativação de vias moleculares por receptores tirosina quinase, tais como o EGFR, pode induzir a expressão de fatores de transcrição indutores de EMT (*Snail e Slug*), pela via Ras/MAPK [157].

Outra via que apresenta importante papel na transformação celular, invasão celular e apoptose é a do NF- $\kappa$ B; foi demonstrado que essa via pode interagir com a sinalização do *NOTCH1* em células de carcinoma de pâncreas [160]. Os membros da família *NOTCH* estão envolvidos em uma variedade de processos de desenvolvimento e controle do destino da célula. Estudos prévios evidenciaram o envolvimento da sinalização do *NOTCH1* em vários tipos de câncer, entretanto, seu papel na tumorigênese é complexo, uma vez que via de sinalização do *NOTCH1* pode apresentar papel duplo, promovendo crescimento e diferenciação ou inibição do crescimento celular tumoral. Recentemente, foi mostrado que a expressão positiva do *NOTCH1* controla os processos de proliferação e apoptose tanto em linhagens de carcinomas de células pequenas (SCLC) como em linhagens de carcinomas de células não pequenas (NSCLC), do subtipo histológico AD; nesses tumores o gene *NOTCH1* atuaria como supressor tumoral. Adicionalmente, a ativação do *NOTCH1* em SCLC induziu a apoptose e redução do crescimento celular, enquanto o bloqueio da sinalização em células de AD inibiu a apoptose. Quando a sinalização do *NOTCH1* foi bloqueada em linhagens de AD e SCLC, ocorreu um aumento da proliferação celular [161]. O bloqueio da via *NOTCH1* também resultou na diminuição do crescimento celular em NSCLC [162].

Os miRNAs também atuam no controle de sinalização da via *NOTCH1* apresentando um papel na transição epitélio-mesênquima [163]. Foi identificada a expressão aumentada do *NOTCH1* associada à diminuição da expressão dos miRNAs miR-200b e miR-200c em



células de carcinoma pancreático. Células com expressão positiva do *NOTCH1* apresentaram um aumento da expressão da proteína ZEB1 (zinc finger E-box binding homeobox 1), sugerindo que o gene *NOTCH1* atua controlando a expressão da proteína ZEB1. Adicionalmente, um aumento da expressão do miR-200b diminuiu a expressão gênica e proteica de ZEB1 e aumentou a expressão da *E-caderina* (proteína importante na adesão celular) em células com *NOTCH1* ativado [163]. As vias de interação proteica aqui identificadas mostraram as conexões entre o miR-200b e as proteínas codificadas pelos genes-alvo *NOTCH1* e *ZEB1*.

O *ERBB3* (*v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3*) consiste em um membro da família do receptor do fator de crescimento epidermal (EGFR). A heterodimerização do *ERBB3* com outros receptores da família ERBB leva à ativação de vias moleculares que regulam a proliferação e diferenciação celular [158]. Avaliando a expressão de genes envolvidos na transição epitélio mesênquima, um estudo mostrou que pacientes com neuroblastoma e com baixa expressão do *ERBB3* no tecido tumoral tiveram pior sobrevida em relação aos pacientes que apresentavam alta expressão do *ERBB3*. Adicionalmente, ensaios *in vitro* mostraram que a diminuição da expressão do *ERBB3* foi altamente correlacionada com a capacidade de invasão em linhagens celulares de neuroblastoma [158].

A desregulação do *PIK3R1* (*phosphoinositide-3-kinase, regulatory subunit 1 (alpha)*) foi identificada em nosso estudo. Tian *et al.* identificaram os miRNAs miR-19b, miR-23b, miR-26a e miR-92a com expressão aumentada em câncer de próstata, promovendo a proliferação celular pela regulação do *PTEN* e da via PIK3/Akt [164]. *PIK3R1* é um componente chave da via PI3K/Akt, regulado por vários miRNAs; estudos adicionais são necessários para melhor elucidar o seu papel na tumorigênese dos carcinomas pulmonares.

A expressão gênica e proteica de *KLF4* (*Kruppel-like factor 4*) foi identificada como diminuída em tecido primário de NSCLC, incluindo SCC e AD e em linhagens de carcinoma pulmonar. Adicionalmente, foi observado que o silenciamento do *KLF4* suprime a expressão do p21 (regulador negativo da fase G1-S do ciclo celular) e promove a expressão do gene *CCND1* (Ciclina D1). Quando a expressão do *KLF4* foi induzida em células de câncer de pulmão, a expressão do p21 foi aumentada e a expressão de *CCND1* foi inibida, levando à saída do ciclo celular na fase G1; entretanto foi observado um efeito mínimo na indução da apoptose; tais achados indicam que a restauração da expressão do *KLF4* pode modificar a expressão de moléculas essenciais no controle da progressão do ciclo celular [165].

Outras proteínas importantes na função de células pulmonares são as mucinas. Mucinas são glicoproteínas de alto peso molecular sintetizadas por células epiteliais e

constituem importantes promotores da invasão de células tumorais [166]. O gene *MUC5B* é um dos principais genes expressos nas células produtoras de mucosa, tais como as células do epitélio pulmonar, apresentando importante papel na regulação das funções normais do pulmão. Devido ao seu papel, alterações nesse gene são comumente observadas em doenças pulmonares [167, 168] incluindo o câncer [169, 170]. No câncer, a mucina apresenta importante papel no crescimento da célula tumoral, invasão e metástase [166]. No presente estudo, a expressão do gene *MUC5B* estava aumentada em NSCLC e associada à metástases em linfonodos. Outro estudo mostrou que a expressão aumentada do gene *MUC5B* em NSCLC estava associada com recorrência pós-operatória em um período de um ano após a cirurgia e com um pior prognóstico dos pacientes. Além disso, a expressão aumentada de *MUC5B* ocorria preferencialmente em tumores de pacientes com histórico de consumo excessivo de tabaco [169]. Um nível aumentado de expressão da proteína MUC5B em AD pulmonar foi sugerido como um potencial biomarcador para detecção de AD [170].

Dentre os fatores de transcrição com papel importante na tumorigênese, o gene *SATB2* (special AT-rich sequence binding protein 2) é um fator de transcrição associado à matriz nuclear e membro da família de proteínas de ligação rica em AT. A expressão diminuída do *SATB2* detectada em nosso estudo, já foi descrita em apenas dois outros tipos de câncer; no carcinoma de laringe e no carcinoma coloretal [171, 172]. O nível de expressão gênica e proteica do *SATB2* estava diminuído no tecido tumoral primário e em linhagens de carcinoma de células escamosas de laringe (LSCC). A diminuição da expressão do *SATB2* foi correlacionada com estadiamento clínico e grau histológico avançados e recorrência do tumor. Adicionalmente, ensaios *in vitro* mostraram que a expressão positiva do *SATB2* inibiu a proliferação celular e a capacidade de transformação maligna das células de LSCC; enquanto que o bloqueio da expressão do *SATB2* nas células resultou no aumento da capacidade de transformação maligna e proliferação celular; esses achados sugerem um papel importante desse fator de transcrição no desenvolvimento de LSCC [171]. Em outros carcinomas de cabeça e pescoço, tais como câncer oral, o *SATB2* estava com expressão aumentada e associada à invasão celular [173]. De acordo com os achados na literatura, sugere-se que o *SATB2* pode desempenhar funções diferentes de acordo com o contexto tecidual e da doença.

Considerando que os miRNAs regulam vários genes-alvo e vias moleculares importantes, conforme descrito acima, os achados desse estudo devem contribuir para a identificação e validação de alvos moleculares com potencial terapêutico em NSCLC.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

1. **Globocan 2012 Stats - Global Cancer Rates Continue to Soar**
2. <http://www2.inca.gov.br/>
3. Jordan RC, Macabeo-Ong M, Shiboski CH, Dekker N, Ginzinger DG, Wong DT, Schmidt BL: **Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer.** *Clin Cancer Res* 2004, **10**(19):6460-6465.
4. Hecht SS: **Tobacco smoke carcinogens and lung cancer.** *J Natl Cancer Inst* 1999, **91**(14):1194-1210.
5. Nattey C, Urban M, Kielkowski D: **0342 Lung cancer risk attributable to occupation: in a case control study in black South Africans, 2001-2008.** *Occup Environ Med* 2014, **71 Suppl 1**:A110-111.
6. Schmid K, Kuwert T, Drexler H: **Radon in indoor spaces: an underestimated risk factor for lung cancer in environmental medicine.** *Dtsch Arztebl Int* 2010, **107**(11):181-186.
7. Herbst RS, Heymach JV, Lippman SM: **Lung cancer.** *N Engl J Med* 2008, **359**(13):1367-1380.
8. Dienstmann R PH, Zukin M: **Tratamento atual do cancer de pulmao nao-pequenas celulas avancado.** *Pratica Hospitalar* 2007; 51: 93-8.; 2007.
9. Wahbah M, Boroumand N, Castro C, El-Zeky F, Eltorkey M: **Changing trends in the distribution of the histologic types of lung cancer: a review of 4,439 cases.** *Ann Diagn Pathol* 2007, **11**(2):89-96.
10. DeVita H, and Rosemberg's: **Cancer Principles & Praticce of Oncology.** In: *Screening for lung cancer.* Edited by Jett DEMaJR; 2011: 625.
11. **NCCN Clinical Practices Guidelines in Oncology (NCCN Guidelines). Non-Small Cell Lung Cancer, version 4.2014**
12. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG *et al*: **Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.** *N Engl J Med* 2004, **350**(21):2129-2139.
13. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ *et al*: **EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.** *Science* 2004, **304**(5676):1497-1500.
14. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L *et al*: **EGF receptor gene mutations are common in lung**

- cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib.** *Proc Natl Acad Sci U S A* 2004, **101**(36):13306-13311.
15. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR: **The biology of epidermal growth factor receptor in lung cancer.** *Clin Cancer Res* 2004, **10**(12 Pt 2):4227s-4232s.
  16. Sharma SV, Bell DW, Settleman J, Haber DA: **Epidermal growth factor receptor mutations in lung cancer.** *Nat Rev Cancer* 2007, **7**(3):169-181.
  17. Cooper WA, Lam DC, O'Toole SA, Minna JD: **Molecular biology of lung cancer.** *J Thorac Dis* 2013, **5 Suppl 5**:S479-490.
  18. Larsen JE, Minna JD: **Molecular biology of lung cancer: clinical implications.** *Clin Chest Med* 2011, **32**(4):703-740.
  19. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, Ladanyi M, Zakowski MF, Heelan RT, Kris MG, Varmus HE: **KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib.** *PLoS Med* 2005, **2**(1):e17.
  20. Zhu CQ, da Cunha Santos G, Ding K, Sakurada A, Cutz JC, Liu N, Zhang T, Marrano P, Whitehead M, Squire JA *et al*: **Role of KRAS and EGFR as biomarkers of response to erlotinib in National Cancer Institute of Canada Clinical Trials Group Study BR.21.** *J Clin Oncol* 2008, **26**(26):4268-4275.
  21. Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK: **Non-small-cell lung cancers: a heterogeneous set of diseases.** *Nat Rev Cancer* 2014, **14**(8):535-546.
  22. Shtivelman E, Hensing T, Simon GR, Dennis PA, Otterson GA, Bueno R, Salgia R: **Molecular pathways and therapeutic targets in lung cancer.** *Oncotarget* 2014, **5**(6):1392-1433.
  23. Liu P, Cheng H, Roberts TM, Zhao JJ: **Targeting the phosphoinositide 3-kinase pathway in cancer.** *Nat Rev Drug Discov* 2009, **8**(8):627-644.
  24. Beck JT, Ismail A, Tolomeo C: **Targeting the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway: An emerging treatment strategy for squamous cell lung carcinoma.** *Cancer Treat Rev* 2014.
  25. Network CGAR: **Comprehensive genomic characterization of squamous cell lung cancers.** *Nature* 2012, **489**(7417):519-525.
  26. Han SS, Kim WJ, Hong Y, Hong SH, Lee SJ, Ryu DR, Lee W, Cho YH, Lee S, Ryu YJ *et al*: **RNA sequencing identifies novel markers of non-small cell lung cancer.** *Lung Cancer* 2014, **84**(3):229-235.
  27. Pérez-Soler R: **Individualized therapy in non-small-cell lung cancer: future versus current clinical practice.** *Oncogene* 2009, **28 Suppl 1**:S38-45.
  28. Shenouda SK, Alahari SK: **MicroRNA function in cancer: oncogene or a tumor suppressor?** *Cancer Metastasis Rev* 2009, **28**(3-4):369-378.

29. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN: **MicroRNA genes are transcribed by RNA polymerase II.** *EMBO J* 2004, **23**(20):4051-4060.
30. Iorio MV, Croce CM: **MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review.** *EMBO Mol Med* 2012, **4**(3):143-159.
31. Bhayani MK, Calin GA, Lai SY: **Functional relevance of miRNA sequences in human disease.** *Mutat Res* 2012, **731**(1-2):14-19.
32. Di Leva G, Calin GA, Croce CM: **MicroRNAs: fundamental facts and involvement in human diseases.** *Birth Defects Res C Embryo Today* 2006, **78**(2):180-189.
33. Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC: **Mammalian mirtron genes.** *Mol Cell* 2007, **28**(2):328-336.
34. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC: **The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila.** *Cell* 2007, **130**(1):89-100.
35. Ruby JG, Jan CH, Bartel DP: **Intronic microRNA precursors that bypass Drosha processing.** *Nature* 2007, **448**(7149):83-86.
36. Yang JS, Lai EC: **Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants.** *Mol Cell* 2011, **43**(6):892-903.
37. Westholm JO, Lai EC: **Mirtrons: microRNA biogenesis via splicing.** *Biochimie* 2011, **93**(11):1897-1904.
38. Havens MA, Reich AA, Duelli DM, Hastings ML: **Biogenesis of mammalian microRNAs by a non-canonical processing pathway.** *Nucleic Acids Res* 2012, **40**(10):4626-4640.
39. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**(2):281-297.
40. Harfe BD: **MicroRNAs in vertebrate development.** *Curr Opin Genet Dev* 2005, **15**(4):410-415.
41. Cummins JM, Velculescu VE: **Implications of micro-RNA profiling for cancer diagnosis.** *Oncogene* 2006, **25**(46):6220-6227.
42. Rinaldi A, Poretti G, Kwee I, Zucca E, Catapano CV, Tibiletti MG, Bertoni F: **Concomitant MYC and microRNA cluster miR-17-92 (C13orf25) amplification in human mantle cell lymphoma.** *Leuk Lymphoma* 2007, **48**(2):410-412.
43. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH: **Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells.** *J Biol Chem* 2008, **283**(2):1026-1033.

44. Reis PP, Tomenson M, Cervigne NK, Machado J, Jurisica I, Pintilie M, Sukhai MA, Perez-Ordenez B, Grénman R, Gilbert RW *et al*: **Programmed cell death 4 loss increases tumor cell invasion and is regulated by miR-21 in oral squamous cell carcinoma.** *Mol Cancer* 2010, **9**:238.
45. Kent OA, Mendell JT: **A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes.** *Oncogene* 2006, **25**(46):6188-6196.
46. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**(2):126-139.
47. Ha M, Kim VN: **Regulation of microRNA biogenesis.** *Nat Rev Mol Cell Biol* 2014, **15**(8):509-524.
48. Iorio MV, Croce CM: **MicroRNAs in cancer: small molecules with a huge impact.** *J Clin Oncol* 2009, **27**(34):5848-5856.
49. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K *et al*: **Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.** *Proc Natl Acad Sci U S A* 2002, **99**(24):15524-15529.
50. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, Liang S, Naylor TL, Barchetti A, Ward MR *et al*: **microRNAs exhibit high frequency genomic alterations in human cancer.** *Proc Natl Acad Sci U S A* 2006, **103**(24):9136-9141.
51. Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, Zeng Y, Miao R, Jin G, Ma H *et al*: **Genetic variants of miRNA sequences and non-small cell lung cancer survival.** *J Clin Invest* 2008, **118**(7):2600-2608.
52. Liu J, Lu KH, Liu ZL, Sun M, De W, Wang ZX: **MicroRNA-100 is a potential molecular marker of non-small cell lung cancer and functions as a tumor suppressor by targeting polo-like kinase 1.** *BMC Cancer* 2012, **12**:519.
53. Khodayari N, Mohammed KA, Goldberg EP, Nasreen N: **EphrinA1 inhibits malignant mesothelioma tumor growth via let-7 microRNA-mediated repression of the RAS oncogene.** *Cancer Gene Ther* 2011, **18**(11):806-816.
54. Miao H, Wei BR, Peehl DM, Li Q, Alexandrou T, Schelling JR, Rhim JS, Sedor JR, Burnett E, Wang B: **Activation of EphA receptor tyrosine kinase inhibits the Ras/MAPK pathway.** *Nat Cell Biol* 2001, **3**(5):527-530.
55. Kinch MS, Moore MB, Harpole DH: **Predictive value of the EphA2 receptor tyrosine kinase in lung cancer recurrence and survival.** *Clin Cancer Res* 2003, **9**(2):613-618.
56. Nasreen N, Mohammed KA, Antony VB: **Silencing the receptor EphA2 suppresses the growth and haptotaxis of malignant mesothelioma cells.** *Cancer* 2006, **107**(10):2425-2435.

57. Lee HY, Mohammed KA, Kaye F, Sharma P, Moudgil BM, Clapp WL, Nasreen N: **Targeted delivery of let-7a microRNA encapsulated ephrin-A1 conjugated liposomal nanoparticles inhibit tumor growth in lung cancer.** *Int J Nanomedicine* 2013, **8**:4481-4494.
58. Li B, Ren S, Li X, Wang Y, Garfield D, Zhou S, Chen X, Su C, Chen M, Kuang P *et al*: **MiR-21 overexpression is associated with acquired resistance of EGFR-TKI in non-small cell lung cancer.** *Lung Cancer* 2014, **83**(2):146-153.
59. Okudela K, Suzuki T, Umeda S, Tateishi Y, Mitsui H, Miyagi Y, Ohashi K: **A comprehensive search for microRNAs with expression profiles modulated by oncogenic KRAS: Potential involvement of miR-31 in lung carcinogenesis.** *Oncol Rep* 2014.
60. Wang LK, Hsiao TH, Hong TM, Chen HY, Kao SH, Wang WL, Yu SL, Lin CW, Yang PC: **MicroRNA-133a suppresses multiple oncogenic membrane receptors and cell invasion in non-small cell lung carcinoma.** *PLoS One* 2014, **9**(5):e96765.
61. Ling H, Fabbri M, Calin GA: **MicroRNAs and other non-coding RNAs as targets for anticancer drug development.** *Nat Rev Drug Discov* 2013, **12**(11):847-865.
62. Bouchie A: **First microRNA mimic enters clinic.** *Nat Biotechnol* 2013, **31**(7):577.
63. Andrews S: **FastQC A Quality Control tool for High Throughput Sequence Data.** In. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; 2010.
64. MARTIN M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** . In., vol. 17. **EMBnet.journal:** <http://journal.embnet.org/index.php/embnetjournal/article/view/200>; 2011: 10-12.
65. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**(3):R25.
66. Simon Anders PTP, Wolfgang Huber: **HTSeq - A Python framework to work with high-throughput sequencing data.** In. bioRxiv preprint; 2014.
67. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data.** *Genome Biol* 2010, **11**(3):R25.
68. Robinson MD, Smyth GK: **Small-sample estimation of negative binomial dispersion, with applications to SAGE data.** *Biostatistics* 2008, **9**(2):321-332.
69. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139-140.
70. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J *et al*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**(10):R80.

71. Li H: **A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.** *Bioinformatics* 2011, **27**(21):2987-2993.
72. Li H: **Improving SNP discovery by base alignment quality.** *Bioinformatics* 2011, **27**(8):1157-1158.
73. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al*: **The variant call format and VCFtools.** *Bioinformatics* 2011, **27**(15):2156-2158.
74. Wang K, Li M, Hakonarson H: **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.** *Nucleic Acids Res* 2010, **38**(16):e164.
75. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.** *Methods* 2001, **25**(4):402-408.
76. Bandi N, Zbinden S, Gugger M, Arnold M, Kocher V, Hasan L, Kappeler A, Brunner T, Vassella E: **miR-15a and miR-16 are implicated in cell cycle regulation in a Rb-dependent manner and are frequently deleted or down-regulated in non-small cell lung cancer.** *Cancer Res* 2009, **69**(13):5553-5559.
77. Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, Calabrò E, Croce CM, Pastorino U, Sozzi G: **MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer.** *Proc Natl Acad Sci U S A* 2011, **108**(9):3713-3718.
78. Cho WC, Chow AS, Au JS: **Restoration of tumour suppressor hsa-miR-145 inhibits cancer cell growth in lung adenocarcinoma patients with epidermal growth factor receptor mutation.** *Eur J Cancer* 2009, **45**(12):2197-2206.
79. Cho WC, Chow AS, Au JS: **MiR-145 inhibits cell proliferation of human lung adenocarcinoma by targeting EGFR and NUDT1.** *RNA Biol* 2011, **8**(1):125-131.
80. Gao W, Yu Y, Cao H, Shen H, Li X, Pan S, Shu Y: **Deregulated expression of miR-21, miR-143 and miR-181a in non small cell lung cancer is related to clinicopathologic characteristics or patient prognosis.** *Biomed Pharmacother* 2010, **64**(6):399-408.
81. Gao W, Shen H, Liu L, Xu J, Shu Y: **MiR-21 overexpression in human primary squamous cell lung carcinoma is associated with poor patient prognosis.** *J Cancer Res Clin Oncol* 2011, **137**(4):557-566.
82. Incoronato M, Garofalo M, Urso L, Romano G, Quintavalle C, Zanca C, Iaboni M, Nuovo G, Croce CM, Condorelli G: **miR-212 increases tumor necrosis factor-related apoptosis-inducing ligand sensitivity in non-small cell lung cancer by targeting the antiapoptotic protein PED.** *Cancer Res* 2010, **70**(9):3638-3646.



83. Liu XG, Zhu WY, Huang YY, Ma LN, Zhou SQ, Wang YK, Zeng F, Zhou JH, Zhang YK: **High expression of serum miR-21 and tumor miR-200c associated with poor prognosis in patients with lung cancer.** *Med Oncol* 2012, **29**(2):618-626.
84. Ma L, Huang Y, Zhu W, Zhou S, Zhou J, Zeng F, Liu X, Zhang Y, Yu J: **An integrated analysis of miRNA and mRNA expressions in non-small cell lung cancers.** *PLoS One* 2011, **6**(10):e26502.
85. Saito M, Schetter AJ, Mollerup S, Kohno T, Skaug V, Bowman ED, Mathé EA, Takenoshita S, Yokota J, Haugen A *et al*: **The association of microRNA expression with prognosis and progression in early-stage, non-small cell lung adenocarcinoma: a retrospective analysis of three cohorts.** *Clin Cancer Res* 2011, **17**(7):1875-1882.
86. Seike M, Goto A, Okano T, Bowman ED, Schetter AJ, Horikawa I, Mathe EA, Jen J, Yang P, Sugimura H *et al*: **MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers.** *Proc Natl Acad Sci U S A* 2009, **106**(29):12085-12090.
87. Solomides CC, Evans BJ, Navenot JM, Vadigepalli R, Peiper SC, Wang ZX: **MicroRNA profiling in lung cancer reveals new molecular markers for diagnosis.** *Acta Cytol* 2012, **56**(6):645-654.
88. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, Guarnera M, Liao J, Chou A, Lu CL *et al*: **Plasma microRNAs as potential biomarkers for non-small-cell lung cancer.** *Lab Invest* 2011, **91**(4):579-587.
89. Tan X, Qin W, Zhang L, Hang J, Li B, Zhang C, Wan J, Zhou F, Shao K, Sun Y *et al*: **A 5-microRNA signature for lung squamous cell carcinoma diagnosis and hsa-miR-31 for prognosis.** *Clin Cancer Res* 2011, **17**(21):6802-6811.
90. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y *et al*: **Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival.** *Cancer Res* 2004, **64**(11):3753-3756.
91. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M *et al*: **A microRNA expression signature of human solid tumors defines cancer gene targets.** *Proc Natl Acad Sci U S A* 2006, **103**(7):2257-2261.
92. Xing L, Todd NW, Yu L, Fang H, Jiang F: **Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers.** *Mod Pathol* 2010, **23**(8):1157-1164.
93. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.** *Cancer Cell* 2006, **9**(3):189-198.

94. Yu L, Todd NW, Xing L, Xie Y, Zhang H, Liu Z, Fang H, Zhang J, Katz RL, Jiang F: **Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers.** *Int J Cancer* 2010, **127**(12):2870-2878.
95. Shirdel EA, Xie W, Mak TW, Jurisica I: **NAViGaTing the microne--using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs.** *PLoS One* 2011, **6**(2):e17429.
96. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA *et al*: **Accurate microRNA target prediction correlates with protein repression levels.** *BMC Bioinformatics* 2009, **10**:295.
97. Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K *et al*: **DIANA-microT web server: elucidating microRNA functions through target prediction.** *Nucleic Acids Res* 2009, **37**(Web Server issue):W273-276.
98. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: **Human MicroRNA targets.** *PLoS Biol* 2004, **2**(11):e363.
99. Betel D, Wilson M, Gabow A, Marks DS, Sander C: **The microRNA.org resource: targets and expression.** *Nucleic Acids Res* 2008, **36**(Database issue):D149-153.
100. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**(1):15-20.
101. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA targeting specificity in mammals: determinants beyond seed pairing.** *Mol Cell* 2007, **27**(1):91-105.
102. Friedman RC, Farh KK, Burge CB, Bartel DP: **Most mammalian mRNAs are conserved targets of microRNAs.** *Genome Res* 2009, **19**(1):92-105.
103. Brown KR, Jurisica I: **Online predicted human interaction database.** *Bioinformatics* 2005, **21**(9):2076-2082.
104. Brown KR, Jurisica I: **Unequal evolutionary conservation of human protein interactions in interologous networks.** *Genome Biol* 2007, **8**(5):R95.
105. Brown KR, Otasek D, Ali M, McGuffin MJ, Xie W, Devani B, Toch IL, Jurisica I: **NAViGaTOR: Network Analysis, Visualization and Graphing Toronto.** *Bioinformatics* 2009, **25**(24):3327-3329.
106. Chomczynski P, and Sacchi, N.: **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** . In., vol. 162: *Anal. Biochem*; 1987: 156-159.

107. Angulo B, Suarez-Gauthier A, Lopez-Rios F, Medina PP, Conde E, Tang M, Soler G, Lopez-Encuentra A, Cigudosa JC, Sanchez-Cespedes M: **Expression signatures in lung cancer reveal a profile for EGFR-mutant tumours and identify selective PIK3CA overexpression by gene amplification.** *J Pathol* 2008, **214**(3):347-356.
108. Boelens MC, van den Berg A, Fehrmann RS, Geerlings M, de Jong WK, te Meerman GJ, Sietsma H, Timens W, Postma DS, Groen HJ: **Current smoking-specific gene expression signature in normal bronchial epithelium is enhanced in squamous cell lung cancer.** *J Pathol* 2009, **218**(2):182-191.
109. Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, Cheng CL, Wang CH, Terng HJ, Kao SF *et al*: **A five-gene signature and clinical outcome in non-small-cell lung cancer.** *N Engl J Med* 2007, **356**(1):11-20.
110. Dehan E, Ben-Dor A, Liao W, Lipson D, Frimer H, Rienstein S, Simansky D, Krupsky M, Yaron P, Friedman E *et al*: **Chromosomal aberrations and gene expression profiles in non-small cell lung cancer.** *Lung Cancer* 2007, **56**(2):175-184.
111. Lonergan KM, Chari R, Coe BP, Wilson IM, Tsao MS, Ng RT, Macaulay C, Lam S, Lam WL: **Transcriptome profiles of carcinoma-in-situ and invasive non-small cell lung cancer as revealed by SAGE.** *PLoS One* 2010, **5**(2):e9162.
112. Barsyte-Lovejoy D, Lau SK, Boutros PC, Khosravi F, Jurisica I, Andrulis IL, Tsao MS, Penn LZ: **The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis.** *Cancer Res* 2006, **66**(10):5330-5337.
113. Lee ES, Son DS, Kim SH, Lee J, Jo J, Han J, Kim H, Lee HJ, Choi HY, Jung Y *et al*: **Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression.** *Clin Cancer Res* 2008, **14**(22):7397-7404.
114. Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, Sun Z, Szoke J, Gerald WL, Watson M *et al*: **A gene expression signature predicts survival of patients with stage I non-small cell lung cancer.** *PLoS Med* 2006, **3**(12):e467.
115. Toh SH, Prathipati P, Motakis E, Kwoh CK, Yenamandra SP, Kuznetsov VA: **A robust tool for discriminative analysis and feature selection in paired samples impacts the identification of the genes essential for reprogramming lung tissue to adenocarcinoma.** *BMC Genomics* 2011, **12 Suppl 3**:S24.
116. Rohrbeck A, Neukirchen J, Rosskopf M, Pardillos GG, Geddert H, Schwalen A, Gabbert HE, von Haeseler A, Pitschke G, Schott M *et al*: **Gene expression profiling for molecular distinction and characterization of laser captured primary lung cancers.** *J Transl Med* 2008, **6**:69.
117. Zhu CQ, Ding K, Strumpf D, Weir BA, Meyerson M, Pennell N, Thomas RK, Naoki K, Ladd-Acosta C, Liu N *et al*: **Prognostic and predictive gene signature for**

- adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol* 2010, **28**(29):4417-4424.**
118. Boutros PC, Lau SK, Pintilie M, Liu N, Shepherd FA, Der SD, Tsao MS, Penn LZ, Jurisica I: **Prognostic gene signatures for non-small-cell lung cancer.** *Proc Natl Acad Sci U S A* 2009, **106**(8):2824-2828.
  119. Zhu CQ, Strumpf D, Li CY, Li Q, Liu N, Der S, Shepherd FA, Tsao MS, Jurisica I: **Prognostic gene expression signature for squamous cell carcinoma of lung.** *Clin Cancer Res* 2010, **16**(20):5038-5047.
  120. Hai J, Zhu CQ, Bandarchi B, Wang YH, Navab R, Shepherd FA, Jurisica I, Tsao MS: **L1 cell adhesion molecule promotes tumorigenicity and metastatic potential in non-small cell lung cancer.** *Clin Cancer Res* 2012, **18**(7):1914-1924.
  121. Corson TW, Zhu CQ, Lau SK, Shepherd FA, Tsao MS, Gallie BL: **KIF14 messenger RNA expression is independently prognostic for outcome in lung cancer.** *Clin Cancer Res* 2007, **13**(11):3229-3234.
  122. Lockwood WW, Chari R, Coe BP, Thu KL, Garnis C, Malloff CA, Campbell J, Williams AC, Hwang D, Zhu CQ *et al*: **Integrative genomic analyses identify BRF2 as a novel lineage-specific oncogene in lung squamous cell carcinoma.** *PLoS Med* 2010, **7**(7):e1000315.
  123. Chari R, Lonergan KM, Pikor LA, Coe BP, Zhu CQ, Chan TH, MacAulay CE, Tsao MS, Lam S, Ng RT *et al*: **A sequence-based approach to identify reference genes for gene expression analysis.** *BMC Med Genomics* 2010, **3**:32.
  124. Landi MT, Zhao Y, Rotunno M, Koshiol J, Liu H, Bergen AW, Rubagotti M, Goldstein AM, Linnoila I, Marincola FM *et al*: **MicroRNA expression differentiates histology and predicts survival of lung cancer.** *Clin Cancer Res* 2010, **16**(2):430-441.
  125. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A *et al*: **Circulating microRNAs as stable blood-based markers for cancer detection.** *Proc Natl Acad Sci U S A* 2008, **105**(30):10513-10518.
  126. Rani S, Gately K, Crown J, O'Byrne K, O'Driscoll L: **Global analysis of serum microRNAs as potential biomarkers for lung adenocarcinoma.** *Cancer Biol Ther* 2013, **14**(12):1104-1112.
  127. Heegaard NH, Schetter AJ, Welsh JA, Yoneda M, Bowman ED, Harris CC: **Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer.** *Int J Cancer* 2012, **130**(6):1378-1386.
  128. Hennessey PT, Sanford T, Choudhary A, Mydlarz WW, Brown D, Adai AT, Ochs MF, Ahrendt SA, Mambo E, Califano JA: **Serum microRNA biomarkers for detection of non-small cell lung cancer.** *PLoS One* 2012, **7**(2):e32307.

129. Rahman M, Lovat F, Romano G, Calore F, Acunzo M, Bell EH, Nana-Sinkam P: **miR-15b/16-2 regulates factors that promote p53 phosphorylation and augments the DNA damage response following radiation in the lung.** *J Biol Chem* 2014, **289**(38):26406-26416.
130. Lowe J, Cha H, Lee MO, Mazur SJ, Appella E, Fornace AJ: **Regulation of the Wip1 phosphatase and its effects on the stress response.** *Front Biosci (Landmark Ed)* 2012, **17**:1480-1498.
131. Shen H, Zhu F, Liu J, Xu T, Pei D, Wang R, Qian Y, Li Q, Wang L, Shi Z *et al*: **Alteration in Mir-21/PTEN Expression Modulates Gefitinib Resistance in Non-Small Cell Lung Cancer.** *PLoS One* 2014, **9**(7):e103305.
132. Korpala M, Kang Y: **The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis.** *RNA Biol* 2008, **5**(3):115-119.
133. Christoffersen NR, Silahtaroglu A, Orom UA, Kauppinen S, Lund AH: **miR-200b mediates post-transcriptional repression of ZFH1B.** *RNA* 2007, **13**(8):1172-1178.
134. Li X, Roslan S, Johnstone CN, Wright JA, Bracken CP, Anderson M, Bert AG, Selth LA, Anderson RL, Goodall GJ *et al*: **MiR-200 can repress breast cancer metastasis through ZEB1-independent but moesin-dependent pathways.** *Oncogene* 2014, **33**(31):4077-4088.
135. Uematsu K, He B, You L, Xu Z, McCormick F, Jablons DM: **Activation of the Wnt pathway in non small cell lung cancer: evidence of dishevelled overexpression.** *Oncogene* 2003, **22**(46):7218-7221.
136. MacDonald BT, Tamai K, He X: **Wnt/beta-catenin signaling: components, mechanisms, and diseases.** *Dev Cell* 2009, **17**(1):9-26.
137. Cui EH, Li HJ, Hua F, Wang B, Mao W, Feng XR, Li JY, Wang X: **Serum microRNA 125b as a diagnostic or prognostic biomarker for advanced NSCLC patients receiving cisplatin-based chemotherapy.** *Acta Pharmacol Sin* 2013, **34**(2):309-313.
138. Zhang P, Bill K, Liu J, Young E, Peng T, Bolshakov S, Hoffman A, Song Y, Demicco EG, Terrada DL *et al*: **MiR-155 is a liposarcoma oncogene that targets casein kinase-1 $\alpha$  and enhances  $\beta$ -catenin signaling.** *Cancer Res* 2012, **72**(7):1751-1762.
139. Zhang M, Zhou X, Wang B, Yung BC, Lee LJ, Ghoshal K, Lee RJ: **Lactosylated gramicidin-based lipid nanoparticles (Lac-GLN) for targeted delivery of anti-miR-155 to hepatocellular carcinoma.** *J Control Release* 2013, **168**(3):251-261.
140. Zhu DX, Zhu W, Fang C, Fan L, Zou ZJ, Wang YH, Liu P, Hong M, Miao KR, Xu W *et al*: **miR-181a/b significantly enhances drug sensitivity in chronic lymphocytic leukemia cells via targeting multiple anti-apoptosis genes.** *Carcinogenesis* 2012, **33**(7):1294-1301.

141. Ma J, Liu J, Wang Z, Gu X, Fan Y, Zhang W, Xu L, Zhang J, Cai D: **NF-kappaB-dependent microRNA-425 upregulation promotes gastric cancer cell growth by targeting PTEN upon IL-1 $\beta$  induction.** *Mol Cancer* 2014, **13**:40.
142. Pikarsky E, Porat RM, Stein I, Abramovitch R, Amit S, Kasem S, Gutkovich-Pyest E, Urieli-Shoval S, Galun E, Ben-Neriah Y: **NF-kappaB functions as a tumour promoter in inflammation-associated cancer.** *Nature* 2004, **431**(7007):461-466.
143. Ben-Neriah Y, Karin M: **Inflammation meets cancer, with NF- $\kappa$ B as the matchmaker.** *Nat Immunol* 2011, **12**(8):715-723.
144. Vasudevan KM, Gurumurthy S, Rangnekar VM: **Suppression of PTEN expression by NF-kappa B prevents apoptosis.** *Mol Cell Biol* 2004, **24**(3):1007-1021.
145. Yu M, Trobridge P, Wang Y, Kannurn S, Morris SM, Knoblauch S, Grady WM: **Inactivation of TGF- $\beta$  signaling and loss of PTEN cooperate to induce colon cancer in vivo.** *Oncogene* 2014, **33**(12):1538-1547.
146. Zhang C, Ge S, Hu C, Yang N, Zhang J: **MiRNA-218, a new regulator of HMGB1, suppresses cell migration and invasion in non-small cell lung cancer.** *Acta Biochim Biophys Sin (Shanghai)* 2013, **45**(12):1055-1061.
147. Liu PL, Tsai JR, Hwang JJ, Chou SH, Cheng YJ, Lin FY, Chen YL, Hung CY, Chen WC, Chen YH *et al*: **High-mobility group box 1-mediated matrix metalloproteinase-9 expression in non-small cell lung cancer contributes to tumor cell invasiveness.** *Am J Respir Cell Mol Biol* 2010, **43**(5):530-538.
148. Hassan MQ, Maeda Y, Taipaleenmaki H, Zhang W, Jafferji M, Gordon JA, Li Z, Croce CM, van Wijnen AJ, Stein JL *et al*: **miR-218 directs a Wnt signaling circuit to promote differentiation of osteoblasts and osteomimicry of metastatic cancer cells.** *J Biol Chem* 2012, **287**(50):42084-42092.
149. Lee HS, Lee K, Jang HJ, Lee GK, Park JL, Kim SY, Kim SB, Johnson BH, Zo JI, Lee JS *et al*: **Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis.** *Oncotarget* 2014, **5**(11):3472-3481.
150. Cao J, Song Y, Bi N, Shen J, Liu W, Fan J, Sun G, Tong T, He J, Shi Y *et al*: **DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer.** *Cancer Res* 2013, **73**(11):3326-3335.
151. Li WQ, Li YM, Tao BB, Lu YC, Hu GH, Liu HM, He J, Xu Y, Yu HY: **Downregulation of ABCG2 expression in glioblastoma cancer stem cells with miRNA-328 may decrease their chemoresistance.** *Med Sci Monit* 2010, **16**(10):HY27-30.
152. Wu Z, Sun L, Wang H, Yao J, Jiang C, Xu W, Yang Z: **MiR-328 expression is decreased in high-grade gliomas and is associated with worse survival in primary glioblastoma.** *PLoS One* 2012, **7**(10):e47270.

153. Arora S, Ranade AR, Tran NL, Nasser S, Sridhar S, Korn RL, Ross JT, Dhruv H, Foss KM, Sibenaller Z *et al*: **MicroRNA-328 is associated with (non-small) cell lung cancer (NSCLC) brain metastasis and mediates NSCLC migration.** *Int J Cancer* 2011, **129**(11):2621-2631.
154. Gu W, Fang S, Gao L, Tan Y, Yang Z: **Clinic significance of microRNA-99a expression in human lung adenocarcinoma.** *J Surg Oncol* 2013, **108**(4):248-255.
155. Yu S, Zhang C, Deng F, Zhang Y: **miR-99a suppresses the metastasis of human non-small cell lung cancer cells by targeting AKT1 signaling pathway.** *J Cell Biochem* 2014.
156. Fennell DA: **Caspase regulation in non-small cell lung cancer and its potential for therapeutic exploitation.** *Clin Cancer Res* 2005, **11**(6):2097-2105.
157. Yang J, Weinberg RA: **Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis.** *Dev Cell* 2008, **14**(6):818-829.
158. Nozato M, Kaneko S, Nakagawara A, Komuro H: **Epithelial-mesenchymal transition-related gene expression as a new prognostic marker for neuroblastoma.** *Int J Oncol* 2013, **42**(1):134-140.
159. Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, Berx G, Cano A, Beug H, Foisner R: **DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells.** *Oncogene* 2005, **24**(14):2375-2385.
160. Wang Z, Zhang Y, Banerjee S, Li Y, Sarkar FH: **Notch-1 down-regulation by curcumin is associated with the inhibition of cell growth and the induction of apoptosis in pancreatic cancer cells.** *Cancer* 2006, **106**(11):2503-2513.
161. Wael H, Yoshida R, Kudoh S, Hasegawa K, Niimori-Kita K, Ito T: **Notch1 signaling controls cell proliferation, apoptosis and differentiation in lung carcinoma.** *Lung Cancer* 2014, **85**(2):131-140.
162. Ji X, Wang Z, Geamanu A, Sarkar FH, Gupta SV: **Inhibition of cell growth and induction of apoptosis in non-small cell lung cancer cells by delta-tocotrienol is associated with notch-1 down-regulation.** *J Cell Biochem* 2011, **112**(10):2773-2783.
163. Bao B, Wang Z, Ali S, Kong D, Li Y, Ahmad A, Banerjee S, Azmi AS, Miele L, Sarkar FH: **Notch-1 induces epithelial-mesenchymal transition consistent with cancer stem cell phenotype in pancreatic cancer cells.** *Cancer Lett* 2011, **307**(1):26-36.
164. Tian L, Fang YX, Xue JL, Chen JZ: **Four microRNAs promote prostate cell proliferation with regulation of PTEN and its downstream signals in vitro.** *PLoS One* 2013, **8**(9):e75885.

165. Hu W, Hofstetter WL, Li H, Zhou Y, He Y, Pataer A, Wang L, Xie K, Swisher SG, Fang B: **Putative tumor-suppressive function of Kruppel-like factor 4 in primary lung carcinoma.** *Clin Cancer Res* 2009, **15**(18):5688-5695.
166. Hollingsworth MA, Swanson BJ: **Mucins in cancer: protection and control of the cell surface.** *Nat Rev Cancer* 2004, **4**(1):45-60.
167. Chen Y, Zhao YH, Di YP, Wu R: **Characterization of human mucin 5B gene expression in airway epithelium and the genomic clone of the amino-terminal and 5'-flanking region.** *Am J Respir Cell Mol Biol* 2001, **25**(5):542-553.
168. Rose MC, Voynow JA: **Respiratory tract mucin genes and mucin glycoproteins in health and disease.** *Physiol Rev* 2006, **86**(1):245-278.
169. Yu CJ, Yang PC, Shun CT, Lee YC, Kuo SH, Luh KT: **Overexpression of MUC5 genes is associated with early post-operative metastasis in non-small-cell lung cancer.** *Int J Cancer* 1996, **69**(6):457-465.
170. Li Y, Wang X, Ao M, Gabrielson E, Askin F, Zhang H, Li QK: **Aberrant Mucin5B expression in lung adenocarcinomas detected by iTRAQ labeling quantitative proteomics and immunohistochemistry.** *Clin Proteomics* 2013, **10**(1):15.
171. Liu TR, Xu LH, Yang AK, Zhong Q, Song M, Li MZ, Hu LJ, Chen FJ, Hu ZD, Han P *et al*: **Decreased expression of SATB2: a novel independent prognostic marker of worse outcome in laryngeal carcinoma patients.** *PLoS One* 2012, **7**(7):e40704.
172. Wang S, Zhou J, Wang XY, Hao JM, Chen JZ, Zhang XM, Jin H, Liu L, Zhang YF, Liu J *et al*: **Down-regulated expression of SATB2 is associated with metastasis and poor prognosis in colorectal cancer.** *J Pathol* 2009, **219**(1):114-122.
173. Chung J, Lau J, Cheng LS, Grant RI, Robinson F, Ketela T, Reis PP, Roche O, Kamel-Reid S, Moffat J *et al*: **SATB2 augments  $\Delta$ Np63 $\alpha$  in head and neck squamous cell carcinoma.** *EMBO Rep* 2010, **11**(10):777-783.



## ANEXOS

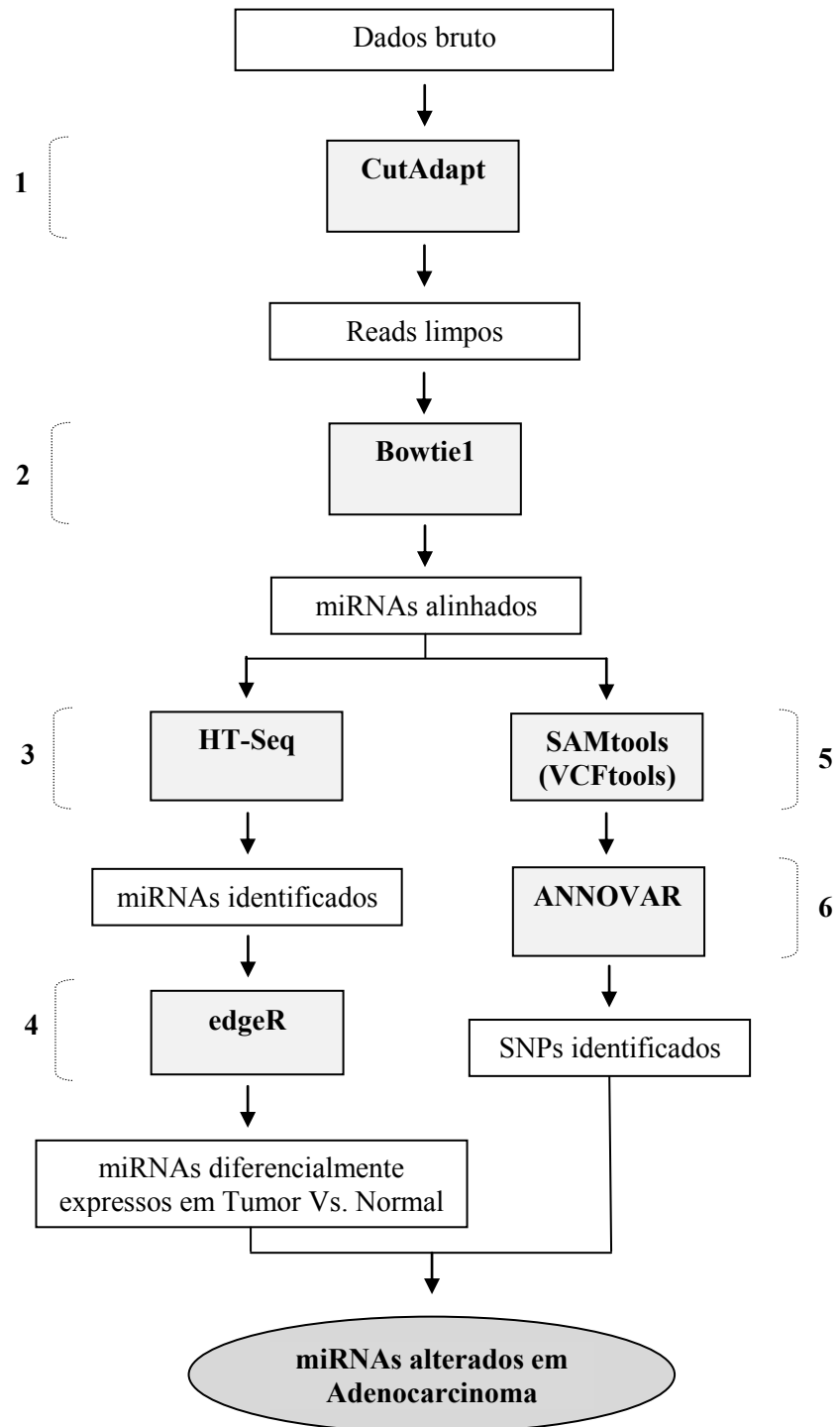
**Anexo 1.** Dados demográficos, clínicos e histopatológicos dos pacientes com adenocarcinoma pulmonar do grupo teste.

	<b>Número Total</b>	<b>%</b>
<b>Idade (anos)</b>		
Mediana (intervalo)	66,8(43-83 anos)	
Média	65	
<b>Sexo</b>		
Masculino	8	47
Feminino	9	53
<b>Tabagismo</b>		
Sim	10	59
Não	7	41
<b>Etilismo</b>		
Sim	1	6
Não	16	94
<b>Tipo Histológico</b>		
Adenocarcinoma	17	100
<b>Grau Histológico do tumor</b>		
Bem diferenciado	5	29
Moderadamente diferenciado	10	59
Pobrememente diferenciado	2	12
<b>Tamanho do Tumor</b>		
T1-T2	15	88
T3-T4	2	12
<b>Metástase Linfonodal</b>		
Negativo (N0)	12	71
Positivo (N1, N2)	5	29
<b>Metástase à distancia</b>		
Sim	5	29
Não	12	71
<b>Estadiamento do tumor</b>		
Ia; Ib; IIa; IIb	14	82
IIIa; IIIb	3	18
<b>Sobrevida</b>		
Vivo com doença	12	71
Óbito	2	12

**Anexo 2.** Dados demográficos, clínicos e histopatológicos dos pacientes com adenocarcinoma pulmonar do grupo de validação.

	<b>Número Total</b>	<b>%</b>
<b>Idade (anos)</b>		
Mediana (intervalo)	58,0(40-84 anos)	
Média	60,5	
<b>Sexo</b>		
Masculino	12	55
Feminino	10	45
<b>Tabagismo</b>		
Sim	15	68
Não	7	32
<b>Etilismo</b>		
Sim	7	32
Não	15	68
<b>Tipo Histológico</b>		
Adenocarcinoma	22	100
<b>Grau Histológico do tumor</b>		
Bem diferenciado	3	14
Moderadamente diferenciado	12	55
Pobrememente diferenciado	7	32
<b>Tamanho do Tumor</b>		
T1-T2	13	59
T3-T4	9	41
<b>Metástase Linfonodal</b>		
Negativo (N0)	12	55
Positivo (N1, N2)	10	45
<b>Metástase à distancia</b>		
Sim	2	9
Não	20	91
<b>Estadiamento do tumor</b>		
Ia; Ib; IIa; IIb	13	59
IIIa; IIIb	9	41
<b>Sobrevida</b>		
Vivo com doença	5	23
Vivo sem evidencia de doença	9	41
Óbito pela doença	3	14
Óbito por outra causa	4	18

## Anexo 3



**Anexo 3. Passos utilizados nas análises de sequenciamento:** 1.Filtragem; 2.Alinhamento/Mapeamento; 3. Anotações; 4.Análises de expressão diferencial; 5.Filtragem das variantes genômicas; 6.Anotações dos SNPs.

**Anexo 4** : Dados demográficos, clínicos e histopatológicos dos pacientes com NSCLC do grupo de validação da meta-análise.

	<b>Número Total</b>	<b>%</b>
<b>Idade em anos (média)</b>	62,99	
<b>Mediana (intervalo)</b>	65,25 (43-82 anos)	
<b>Sexo</b>		
Masculino	16	53,3
Feminino	14	46,7
<b>Tabagismo</b>		
Sim	22	73,3
Não	8	26,7
<b>Etilismo</b>		
Sim	5	16,7
Não	25	83,3
<b>Tipo Histológico</b>		
Adenocarcinoma	21	70,0
CEC	9	30,0
<b>Grau Histológico</b>		
Bem diferenciado	10	33,3
Moderadamente diferenciado	12	40,0
Pobremente diferenciado	8	26,7
<b>Tamanho do tumor</b>		
T1-T2	25	83,3
T3-T4	5	16,7
<b>Metástase em linfonodos</b>		
Negativo (N0)	19	63,3
Positivo (N1, N2)	11	36,7
<b>Estadiamento do tumor</b>		
Ia; Ib; IIa; IIb	22	73,3
IIIa; IIIb, IV	8	26,7
<b>Sobrevida</b>		
Vivo com a doença	21	70,0
Óbito	9	30,0

**Anexo 5.** Genes e sequência dos primers utilizados na validação.

<b>Símbolo do Gene</b>	<b>Sequência dos primers</b>
<b><i>ERBB3</i></b>	F-5'-gccgaaacgtgctactcaa-3' R-5'- aggtcagccacacccaaaatctg-3'
<b><i>KLF4</i></b>	F-5'-tggggtctgtgactggatct-3' R-5'-tcacccccttggcattttgt-3'
<b><i>MUC5B</i></b>	F-5'-cctttgacggcacctttacac-3' R-5'-gagattcccaaagcgtgcat-3'
<b><i>NOTCH1</i></b>	F-5'-tgcgagaccaacatcaacga-3' R-5'-aagcagaggtaggcgttgtc-3'
<b><i>PIK3R1</i></b>	F-5'-gcttgccgagccctataact-3' R-5'-tgcacaaggaggtgtgttg-3'
<b><i>SRC</i></b>	F-5'-agcctcagtttatcacgcaa-3' R-5'-tcctgggaaggttgaactg-3'
<b><i>SATB2</i></b>	F-5'-tctcccagcttagtcacaa-3' R-5'-gctaacggcaatctgttggtt-3'
<b><i>ZEB 1</i></b>	F-5'-tgtgaatgggcgaccaaga-3' R-5'-gtgggactgcctggtgatg-3'
<b><i>EGFR</i></b>	F-5'-tccgtgagttgatcatcgaatt-3' R-5'-ccacgtcgtccatgttctt-3'
<b><i>KRAS</i></b>	F-5'-cattggtgagggagatccga-3' R-5'-agaaggcatcatcaacacca-3'

(F: forward; R: reverse)

### Anexo 6. Distribuição global das SNVs em miRNAs

AMOSTRA	miR-24-1	miR-3074	miR-4422	miR-125b2	miR-150	miR-3675	miR-103a1	miR-103a1	miR-4490	let-7a2	miR-140	miR-99b	miR-3687	miR-5702	let-7c	let-7c	miR-23b	miR-23b	miR-148b	miR-200a	miR-199a2	miR-548f1	miR-548f-1	miR-4472-1	miR-5580	miR-100	miR-99a	miR-4443	miR-769	miR-572	let-7a1	let-7a1					
TUMOR																																					
31T																																					
33T					+	+	+	+	+	+	+	+	+		+	+																					
52T					+		+	+	+		+		+		+	+	+																				
59T	+	+			+		+	+	+				+		+	+	+			+																	
65T	+	+		+	+		+	+	+	+	+		+		+	+	+																				
66T	+	+		+			+	+	+	+			+		+	+	+				+																
69T				+	+	+	+		+						+	+	+																				
72T				+	+		+	+	+		+		+		+	+	+				+																
73T	+	+	+	+	+		+	+	+		+			+	+	+	+	+				+	+	+	+												
74T					+		+	+	+		+				+	+	+				+					+											
78T	+	+	+	+	+		+	+	+	+	+		+		+	+	+					+	+	+			+	+									
81T	+	+	+				+	+	+		+		+		+	+		+			+																
82T	+	+		+	+		+	+	+	+	+		+		+	+	+														+	+					
95T				+	+		+	+	+	+			+		+	+	+																				
171T				+	+		+	+	+				+		+	+	+					+															
172T				+			+	+	+	+	+	+	+		+	+	+																				
174T							+	+	+	+	+	+	+		+	+	+																				
NORMAL																																					
33N							+	+	+		+	+	+		+	+	+				+												+	+			
52N					+		+	+	+	+		+	+		+	+	+				+	+												+			
55N							+		+			+	+		+	+					+																
59N							+	+	+	+	+	+	+		+	+					+																
72N							+	+	+				+		+	+	+																				
73N				+			+		+	+	+	+			+	+	+																				
173N									+	+			+		+	+										+								+	+	+	

+: presença de SNVs

# **MANUSCRITO**

**miRNA transcriptome sequencing identifies potential biomarkers in lung adenocarcinoma from Brazilian patients.**

**Sequenciamento de miRNAs identifica potenciais biomarcadores de adenocarcinoma pulmonar de pacientes Brasileiros**

## **miRNA transcriptome sequencing identifies potential biomarkers in lung adenocarcinoma from Brazilian patients.**

Cinegaglia NC<sup>1</sup>, Andrade SCS<sup>2</sup>, Pinheiro M<sup>3</sup>, Severino FE<sup>1</sup>, Oliveira RA<sup>4</sup>, Hasimoto E<sup>1</sup>, Cataneo DC<sup>1</sup>, Cataneo AJM<sup>1</sup>, Defaveri J<sup>5</sup>, Souza CP<sup>1,6</sup>, Marques MMC<sup>6,7</sup>, Carvalho RF<sup>8</sup>, Coutinho LL<sup>2</sup>, Gross JL<sup>9</sup>, Rogatto SR<sup>9,10</sup>, Lam WL<sup>11</sup>, Jurisica I<sup>12,13</sup>, Reis PP<sup>1</sup>.

<sup>1</sup>Department of Surgery and Orthopedics, Faculty of Medicine, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>2</sup>Department of Animal Biotechnology, University of São Paulo (USP), Piracicaba, SP, Brazil

<sup>3</sup>Department of Genetics, Biosciences Institute, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>4</sup>Department of Biostatistics, Biosciences Institute, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>5</sup>Department of Pathology, Faculty of Medicine, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>6</sup>Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, SP, Brazil

<sup>7</sup>Barretos School of Health Sciences, Barretos, SP, Brazil

<sup>8</sup>Department of Morphology, Biosciences Institute, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>9</sup>International Center of Research and Training (CIPE), A. C. Camargo Cancer Center, São Paulo, SP

<sup>10</sup>Department of Urology, Faculty of Medicine, São Paulo State University, UNESP, Botucatu, SP, Brazil

<sup>11</sup>Department of Integrative Oncology, British Columbia Cancer Research Centre, Vancouver, British Columbia V5Z 1L3, Canada

<sup>12</sup>University Health Network Division of Cancer Informatics, Ontario Cancer Institute and Princess Margaret Hospital, Toronto, Ontario, Canada

<sup>13</sup>Department of Computer Science, University of Toronto, Toronto, Ontario, Canada

### **Corresponding author**

Patricia Reis, PhD

Assistant Professor

Dept. of Surgery and Orthopedics, Faculty of Medicine,  
São Paulo State University (UNESP)

Av. Prof. Montenegro, S/N. CEP 18618-970

Botucatu, SP.

BRAZIL

Phone: +55 (14) 3880-1451

Email: [preis@fmb.unesp.br](mailto:preis@fmb.unesp.br)



## ABSTRACT

**Background:** Lung cancer is a leading cause of cancer death, with an incidence in Brazil of >27,000 new cases/year. Although advances in molecularly-targeted treatment, high mortality rates are still observed. There is thus a need for identification of new, clinically useful biomarkers/deregulated molecular pathways to further improve patient survival. microRNAs (miRNAs) are important gene expression regulators involved in oncogenesis.

**Objectives:** To identify global transcriptome miRNA changes, target genes and molecular pathways modulated by miRNAs in adenocarcinoma from Brazilian patients.

**Patients and Methods:** Patient samples were selected as test (N=24) and independent validation (N=34) sets, including tumors and paired histologically normal lung tissues. Test samples were subjected to miRNA transcriptome sequencing followed by TaqMan PCR validation.

**Results:** Unpaired sample analysis showed that 11 miRNAs were statistically significantly ( $p \leq 0.01$ ,  $FDR \leq 0.1$ ) deregulated, including miR-486 under-expression and let-7a-2, let-7g, miR-15b, miR-181b-1, miR-181b-2, miR-23b, miR-26a-1, miR-26a-2, miR-26b and miR-93 over-expression in the tumor compared to normal tissues. In additional, a paired-sample analysis showed that 22 miRNAs were significantly deregulated ( $p \leq 0.01$ ,  $FDR \leq 0.1$ ); 8 miRNAs (miR-486, miR-1247, miR-218-1, miR-181a-1, miR-181a-2, miR-328, miR-574 and miR-886) were down-regulated and 14 miRNAs (let-7a-3, miR-146b, miR-26a-1, miR-200b, miR-191, miR-181c, miR-10a, miR-155, miR-99a, miR-30e, miR-21, miR-425, miR-29a and miR-421) were up-regulated in the tumor compared to the normal tissue from the same patient. TaqMan PCR validation showed that let-7a-2, let-7a-3, miR-15b, miR-21, miR-155 and miR-200b were significantly deregulated ( $p \leq 0.01$ ,  $FDR \leq 0.1$ ) in adenocarcinoma in both miRNA-Seq and TaqMan PCR analyses.

**Conclusions:** let-7a-2, let-7a-3, miR-15b, miR-21, miR-155 and miR-200b regulate the expression of several target genes interconnected into pathways relevant to tumor development, progression and metastasis. Our findings shed light into future functional characterization of such miRNAs, which may have a potential role for the development of clinically applicable therapeutic strategies in lung adenocarcinoma.

## INTRODUCTION

Lung carcinoma is a leading cause of cancer death with >1.5 million deaths every year, worldwide [1]. In developing countries such as Brazil, incidence data from the local National Cancer Institute (INCA) show ~27,000 new cases/year [2]. In Brazil, São Paulo State has one of the greatest number of cancers affecting the trachea, bronchus and lungs, with an incidence of ~22.9/100,000 man and ~11.8/100,000 women [3]. Lung cancer is frequently diagnosed in advanced disease stages, which is associated with lower survival rates of 13-21% in developed countries and 7-10% in developing countries [2]. Risk factors include tobacco smoking in the majority of cases, besides passive exposure to smoking, radiation, occupational and environmental factors, among others [4-6].

Currently, lung cancer is classified in different histological subtypes, which are used to help guide treatment decisions [7]. Non Small Cell Lung Cancer (NSCLC) comprise the majority (~85%) of cases, classified into several histological subtypes; the most common are adenocarcinoma, squamous cell carcinoma and large cell carcinoma [7, 8]. Among these, adenocarcinoma is the most frequent subtype, occurring in ~47% of women and 37% of men diagnosed with NSCLC [7, 9].

Recent advances in treatment of patients with NSCLC include molecularly targeted approaches, such as the use of tyrosine-kinase inhibitors Gefitinib and Erlotinib for patients with *EGFR* positive tumors (mainly exon 19 and 21 deletions) [10-14] and Crizotinib for tumors containing ALK rearrangements [15, 16]; current treatment protocols are described in the National Comprehensive Cancer Network Guidelines ([http://www.nccn.org/professionals/physician\\_gls/f\\_guidelines.asp](http://www.nccn.org/professionals/physician_gls/f_guidelines.asp)) [17].

Although several studies have identified genomic and transcriptome changes including mutations in NSCLC, still ~50% of driver changes are unknown in adenocarcinoma. Considering that molecularly targeted approaches have benefited a small fraction of patients with adenocarcinoma, there is still a need for the identification of new driver changes and molecular pathways that may be useful as therapeutic targets [18].

In the past years, the role of microRNAs (miRNAs) has emerged in several diseases, including cancer. miRNAs are small, non-coding RNAs (~18-22 nucleotides) transcribed from DNA and with an important role in post-transcriptional gene expression regulation leading to translational repression [19]. Interestingly, miRNAs may also bind ribonucleoproteins in the seed sequence, independent of the RNA Induced Silencing Complex (RISC), influencing miRNA:mRNA binding and leading to translational activation. miRNAs may also modulate gene expression at the transcriptional level by

direct binding to DNA [20]. miRNA biogenesis and processing through canonical and non-canonical pathways demonstrates how multiple functions can be performed by RNA transcripts in the cell.

It is well known that miRNAs play important roles in multiple biological processes, such as embryonic development, cell proliferation and differentiation, apoptosis [19, 21, 22] and oncogenesis mechanisms [23-25]. According to miRBase, there are 2,588 mature miRNA sequences identified in the human genome (<http://www.mirbase.org/cgi-bin/browse.pl?org=hsa>); miRNAs are mapped to distinct genomic regions including introns or exons of both non-coding and coding transcripts or may cover full introns; the latter named miRtrons [26, 27].

In cancer, miRNAs may act as oncogenes and tumor suppressor genes [28-30]. Different mechanisms can lead to cellular miRNA deregulation including genomic deletions, amplifications, mutations and polymorphisms, transcriptional deregulation and alterations in the biogenesis machinery or epigenetic modifications [31].

Deregulated miRNA expression has been associated with lung tumorigenesis [32-35]. Currently, molecular pathways targeted by miRNAs are being investigated as therapeutic approaches [36]. For example, a phase I clinical trial is testing the potential of miR-34 mimics in different solid tumors; miR-34 has tumor-suppressive functions inhibiting the expression of >20 oncogenes with roles in cell cycle and proliferation, apoptosis, metastasis and chemoresistance of cancer cells [37].

Considering the importance of miRNAs in tumorigenesis, characterization of transcriptome global expression changes in miRNAs contributes to the development of improved therapeutic strategies in NSCLC. To our knowledge, ours is the first study on the global transcriptome miRNA sequencing of tumors from Brazilian patients. In addition, our study includes profiling of histologically normal lung tissues, which allowed us to identify miRNA changes in paired tumor and normal, as well as the overall miRNA transcriptome changes in lung adenocarcinoma in this patient group.

## **MATERIAL AND METHODS**

### **Patients and Samples**

This study was approved by the Research Ethics Boards of Faculty of Medicine, UNESP, Botucatu, SP (4319/2012), AC Camargo Hospital, São Paulo, SP (1573/11) and Barretos Cancer Hospital, Barretos, SP (75907). Informed patient consent was obtained from all eligible patients before sample collection. Inclusion criteria were patients >18 years old,

histopathological diagnosis of lung adenocarcinoma, primarily treated by surgery. Exclusion criteria were patients <18 years old and with diagnosis of other diseases.

Patient samples were selected as test and validation sets. Test set samples (N=24) were prospectively collected from surgeries performed at AC Camargo Hospital, SP and comprised tumors (lung adenocarcinoma, N=17) and histologically normal lung tissue from same patients (N=7). All prospectively collected tissue samples were immediately frozen in liquid nitrogen and kept at -80°C until RNA extraction. Validation set samples (N=34) were retrospectively obtained in the period of 2000-2012 from the Department of Pathology of Faculty of Medicine, UNESP, Botucatu, SP and Barretos Cancer Hospital, Barretos, SP. Validation set samples thus comprised formalin-fixed, paraffin embedded (FFPE) tissue blocks from tumors (lung adenocarcinoma, N=22) and histologically normal lung tissue from same patients (N=12). Patient samples were thus obtained from three different hospitals all located within São Paulo State and the patient population had similar demographic characteristics.

We aimed at identifying a panel of miRNA expression changes in lung adenocarcinoma through deep sequencing followed by validation by quantitative real-time RT-PCR (RT-QPCR) using a TaqMan miRNA array platform. **Figure 1** shows the experimental design and samples. Detailed clinical and histopathological data of patients comprising the test and validation set samples are shown in **Tables 1 and 2**, respectively.

### **RNA Extraction**

Fresh-frozen tissue samples were subjected to frozen section analysis to ensure the presence of >90% tumor or normal cells in cancer or normal lung samples collected by surgery, respectively. Samples were fragmented and lysed using the Precellys 24 lysing and homogenization system (Berting Technologies, Rockville, MD, USA) for 10s at 6,500 rpm. RNA extraction was then performed using the Trizol reagent [38] (Life Technologies, Carlsbad, CA, USA) followed by purification using the RNeasy Mini Kit (Qiagen, Hilden, Germany), as previously described [39].

RNA from FFPE samples was isolated using the *RecoverAll Total Nucleic Acid Isolation* (Ambion/Life Technologies, Carlsbad, CA, USA), following a previously reported protocol with modifications to improve RNA yield [40]. RNA samples were quantified using *NanoDrop 8000* (Thermo Fisher Scientific, Waltham, MA, USA) and RNA quality was assessed using *Bionalyzer 2100* (Agilent Technologies, Santa Clara, CA, USA), following the manufacturer's protocol.

RNA samples were stored at -80°C until use for library preparation.

### **miRNA Transcriptome Sequencing**

RNA (1µg) from test set samples (N=24) was used for library preparation, cluster generation and sequencing using the MiSeq system (Illumina, San Diego, CA, USA). Deep sequencing was performed at the Laboratory of Biotechnology, University of São Paulo (USP), Piracicaba, SP. Sequencing consisted in *in vitro* cloning of RNA fragments in a solid platform. MiSeq platform generated 50 bp single read fragments. All steps including library preparation, clusterization and sequencing followed the manufacturer's instructions. Library preparation used the *TruSeq Small RNA Sample Preparation* kit (Illumina, San Diego, CA, USA). Briefly, 1µg RNA was used for adaptor ligation, which contain a ligation site for the sequencing primer, used to identify samples comprising an RNA pool and another ligation site for the primers in the flow cell, which are used for fragment amplification by PCR. cDNA library was obtained by PCR amplification following 11 cycles of 98°C for 30s, 98°C for 10s; 60°C for 30s; 72°C for 15s and 72°C for 10 min. Library samples were subjected to agarose gel electrophoresis for miRNA isolation; cDNA samples were then ethanol precipitated and quantified using Qubit 2.0 Fluorometer (Invitrogen by Life Technologies, Carlsbad, CA, USA). In the clusterization step, fragments ligated to adaptors were denatured for double strand separation, allowing single strand molecules to bind primers in the flow cell and to produce multiple copies of specific fragments by solid phase PCR amplification. Transcriptome sequencing was performed using the MiSeq Reagent Kit v2 (50 cycles) following the manufacturer's protocol.

### **miRNA Transcriptome Sequencing Data Analysis**

Reads quality was assessed using FastQC [41] and reads cleaning performed by *CutAdapt* [42] in order to select good quality reads for further analysis. Reads that did not pass quality control and cleaning steps were thus removed. Alignment of reads was performed based on hg19 reference genome (<https://genome.ucsc.edu/cgi-bin/hgTracks?hgsid=12832096&chromInfoPage=>) using *Bowtie1* [43] followed by *HT-Seq* [44] for annotation and quantification of aligned miRNA sequences. Data normalization [45] and miRNA differential expression analysis were performed using *edgeR* (Bioconductor/R) [46-48]. Single nucleotide variations in miRNA sequences were identified by sequence mapping using *SAMtools* [49, 50] and *VCFtools* [51] followed by genomic annotation using *ANNOVAR* [52]. Data analyses steps are shown in **Figure 2**.

### **Validation of miRNA expression**

miRNA expression was validated using a TaqMan® Array Human MicroRNA platform (Life Technologies, Foster City, CA, USA), as previously described [53]. We used the QuantStudio 12K system (Life Technologies, Foster City, CA, USA). Global data normalization was performed in Expression Suite software (Life Technologies, Foster City, CA, USA) and miRNA expression profiles were determined using RQ Manager v1.2 software (Life Technologies, Foster City, CA, USA).

### **Statistical Analyses**

Statistical analyses were performed to correlate deregulated miRNA expression with clinical and histopathological data of patients using Mann-Whitney test and Fisher's exact test. Statistical tests were applied using SAS 9.1 (Statistical Analyses System). Statistical significance was defined as  $p < 0.05$ .

## **RESULTS**

### **Transcriptome miRNA-Sequencing (miRNA-Seq)**

miRNA-Seq generated 13,135,522 reads with an average of 547,313 reads/sample. FastQC quality test showed that 96,1% (12,623,236,642) of reads had a Q-score  $\geq 30$  and were thus considered for further analyses. Overall unpaired sample analysis showed that 11 miRNAs were statistically significantly ( $p \leq 0,01$  and  $FDR \leq 0,1$ ) deregulated, including miR-486 under-expression and let-7a-2, let-7g, miR-15b, miR-181b-1, miR-181b-2, miR-23b, miR-26a-1, miR-26a-2, miR-26b and miR-93 over-expression in the tumor compared to normal tissues. In addition, a paired-sample analysis showed that 22 miRNAs were statistically significantly deregulated ( $p \leq 0,01$ ;  $FDR \leq 0,1$ ); 8 miRNAs (miR-486, miR-1247, miR-218-1, miR-181a-1, miR-181a-2, miR-328, miR-574 and miR-886) were down-regulated and 14 miRNAs (let-7a-3, miR-146b, miR-26a-1, miR-200b, miR-191, miR-181c, miR-10a, miR-155, miR-99a, miR-30e, miR-21, miR-425, miR-29a and miR-421) were up-regulated in the tumor compared to the normal tissue from the same patient. Notably, deregulated expression of miR-486 and miR-26a-1 were detected in both analyses (unpaired and paired samples), considering filtering criteria of  $p \leq 0,01$  and  $FDR \leq 0,1$ . Statistically significantly deregulated miRNAs identified in unpaired and paired samples are shown in **Table 3**.

### **Validation of miRNA expression**

TaqMan quantitative real-time PCR validation was performed for all significantly deregulated miRNAs (**Table 3**) in an independent set of 22 adenocarcinoma samples and

12 normal lung tissues; normal samples were used as controls and combined into 3 pools of 4 samples each. Fourteen miRNAs showed concordant levels of expression (although not all miRNAs were statistically significant) between miRNA-Seq and TaqMan PCR data (**Table 4**). Of these 14 miRNAs, 5 were under-expressed (miR-486, miR-181a-1, miR-181a-2, miR-218-1 e miR-886) and 9 were over-expressed (let-7a-2, let-7a-3, let-7g, miR-15b, miR-26b, miR-200b, miR-155, miR-21, miR-425). 6/14 miRNAs; let-7a-2, let-7a-3, miR-15b, miR-200b, miR-155 and miR-21 were statistically significantly deregulated ( $p \leq 0,01$ ;  $FDR \leq 0,1$ ) in tumor compared to normal lung tissues in both miRNA-Seq and TaqMan PCR analysis (**Table 4**).

### **Single nucleotide variants**

miRNA transcriptome sequencing data analysis showed single nucleotide variants (SNVs) in miRNA mature sequences. SNVs were identified compared to GRCh37/hg19 Genome Reference Consortium. We identified 30 SNVs in adenocarcinoma samples and 19 SNVs in histologically normal lung tissues. SNVs identified in tumor and normal tissues are described in **Tables 5 and 6**, respectively. Three miRNAs (miR-24-1, miR-3074 and miR-4422) had SNVs (A/T, A/T and C/T) exclusively in tumor samples, being detected in 7/17 (41%), 7/17 (41%) and 3/17 (17%) tumors, respectively. Highly frequent SNVs (T/A and G/T) were detected in the miRNAs; miR-125b-2 in 11/17 (65%) tumors and miR-150 in 12/17 (70%) tumors; the latter being significant when compared tumor to normal tissue ( $P=0,0233$ , Fisher's exact test). SNVs observed in miR-125b-2 and miR-150 were detected in both tumor and normal in one case (52T, 52N); such SNVs may represent inherited polymorphisms rather than mutations. Although, SNVs detected in tumor and normal samples require further validation by DNA sequencing analysis, to verify whether the genomic DNA sequencing encoding the miRNA harbors nucleotide changes. SNVs for both miR-24-1 and miR-3074 were associated with early tumor stage (I/II) ( $p=0.0295$ ).

## **DISCUSSION**

miRNA expression changes lead to target gene deregulation, which in turn affects molecular pathways important to tumor development and progression. Alterations in miRNA expression may be the result of genomic changes such as amplifications, deletions or alterations in miRNA transcription and processing machinery; single nucleotide variations or mutations may result in a change in mRNA targets [31]. Therefore, the identification of pathways modulated by miRNAs may lead to better

characterization of clinically applicable molecular targets that may ultimately result in improved cancer therapeutics [54].

By applying stringent criteria to our miRNA-Seq analysis, we identified statistically significant under-expressed miRNAs, namely, miR-486, miR-181a-1, miR-181a-2, miR-218-1 and miR-886 and over-expressed miRNAs; let-7a-2, let-7a-3, let-7g, miR-15b, miR-21, miR-26b, miR-155, miR-200b and miR-425 in lung adenocarcinoma compared to histologically normal lung tissues. Such miRNAs regulate, simultaneously, a large number of genes with important cellular functions. Indeed, miRNAs are known to play roles in the hallmarks of cancer modulating the expression of genes and proteins during disease development to progression [20].

In agreement with our data, miRNA let-7 family (let-7a, b, c, d, e, f, g and i) have been reported as significantly over-expressed in lung adenocarcinoma [124]. Other reports have shown reduced expression of let-7 family members in NSCLC cell lines and associated with lower patient survival [90, 93]. Specifically, low expression of let-7a-2 and let-7f-1 was associated with low survival of patients with lung adenocarcinoma, however this same study showed that a subgroup of adenocarcinomas had over-expression of let-7a-1, let-7a-2 and let-7a-3 [90].

miRNAs have also been explored as circulating biomarkers due to their stability and relatively easy detection in body fluids such as blood [125]. Herein, we identified deregulated expression of miRNAs such as miR-15b, which have been previously detected in the blood of patients with NSCLC [126]. Distinct expression levels of miR-15b/27b were identified in patients with NSCLC compared to healthy individuals in a phase I/II study [128]. Increased expression of a miRNA signature, miR-15b/16-2 induced expression of genes involved in ATM and RAD3 signaling, cell cycle, DNA repair and apoptosis; this study demonstrated that *PPM1D* (*WIP1*; *wild-type p53-induced phosphatase 1*) is a direct target of miR-15b, reducing its mRNA and protein expression in lung cells, suggesting that DNA damage response by miR-15b may be partially modulated by *PPM1D* inhibition and could be a potential therapeutic target in radiation-induced lung cancer [129]. *PPM1D* encodes a serine/threonine phosphatase that plays a role in dephosphorylation of several DNA damage response proteins such as ATM, ATR, p38MAPK, CHK1 and CHK2 [130].

Deregulated expression of miR-486 (under) and miR-21, mir-200b and miR-375 (over-expressed) was detected in tumor tissue and sputum of patients with early stage (I) lung adenocarcinoma; in this study, combined expression of 4 miRNAs in the sputum



distinguished patients with adenocarcinoma and squamous cell carcinoma histology from healthy individuals, having this 4-miRNA signature higher specificity and sensitivity in the adenocarcinoma subtype [94]. In our study, we have shown consistent data with statistically significant over-expression of miR-21 and miR-200b in most adenocarcinoma samples compared to histologically normal lung tissues.

Deregulated miRNA expression profiles have also been associated with resistance or sensibility to chemotherapy. Indeed, miR-21 over-expression and PTEN protein under-expression were associated with low sensibility to the tyrosine kinase inhibitors Gefitinib or Erlotinib and low survival of patients with NSCLC. Increased miR-21 and decrease PTEN expression has been shown in Gefitinib-resistant cell lines; *in vitro* miR-21 mimics data showed a reduced sensibility to Gefitinib due to PTEN inhibition and AKT/ERK activation while miR-21 inhibition was able to restore sensibility to treatment [131]. Over-expression of both miR-21 and miR-200b in our adenocarcinoma samples may explain, at least in part, the low rate of adjuvant chemotherapy response of our patients.

miRNAs also regulate important pathways associated with metastasis, such as TGF- $\beta$  signaling, which activates multiple transcription factors that induce epithelial to mesenchymal transition (EMT). In particular, miR-200 family plays an important role in EMT through regulation of genes such as *CDH1*, *ZEB1* and *ZEB2* (*Zinc finger E-box binding homeobox*) [132]. *In vivo* and *in vitro* data demonstrated that miR-200b inhibited *ZEB2* and *CDH1* expression in breast cancer cells. *ZEB2* is a transcriptional repressor that cooperates with regulation of TGF- $\beta$  signaling and EMT through *CDH1* [133]. However, other studies have shown that although *ZEB1/CDH1* are regulated by miR-200b, restoration of *ZEB1* expression in breast cancer cell lines that express miR-200b was unable to alter metastatic potential, suggesting a mechanism of regulation including other target genes which contribute to metastasis, such as actin-associated genes [134].

Wnt/ $\beta$ -catenin signaling, a molecular pathway important to tumor progression, has been associated with miR-155 regulation. Deregulated Wnt/ $\beta$ -catenin signaling was able to promote cell growth and was associated with poor prognosis of patients with NSCLC [135, 136]. In our study, miR-155 over-expression was validated in the independent series of adenocarcinoma samples. An additional study has also detected miR-155 over-expression in primary adenocarcinoma tissues [93] and in serum from patients with advanced-stage (IV) NSCLC [137]; in both studies this alteration was associated with low patient survival [93, 137]. miR-155 over-expression has also been reported in other tumor types, such as liposarcoma; *in vitro* and *in vivo* assays showed that CK1 $\alpha$  (casein

kinase 1 $\alpha$ ), a key regulator of Wnt/ $\beta$ -catenin pathway, is a direct target of miR-155, leading to  $\beta$ -catenin signaling and CCND1 expression, which contributes to cell proliferation and tumor progression [138]. Interestingly, miR-155 has been studied as a therapeutic target; *in vitro* and *in vivo* data showed efficient delivery of nti-miR-155 using nanoparticles in a hepatocellular carcinoma cell line [139].

Herein, we have also identified Single Nucleotide Variants (SNVs) in miRNA sequences, namely, miR-let-7 family (let-7a-1, let-7a-2 and let-7c) were identified in tumor samples. Variants in miRNA sequences may map to distinct genomic locations, however their sequences are highly homologous. Although variants have unexplored functions, they may have biological roles in tumorigenesis; such variations have been detected for let-7f miRNAs [72] and miR-196a2 in NSCLC [73] and may lead to weak miRNA:mRNA binding or alterations in mRNA targets [31].

Deregulated miRNAs identified, especially let-7a-2, let-7a-3, miR-15b, miR-21, miR-155 and miR-200b regulate the expression of several target genes interconnected into pathways relevant to tumor development, progression and metastasis. Our findings shed light into future functional characterization of such miRNAs, which may have a potential role for the development of clinically applicable therapeutic strategies in lung adenocarcinoma.

## REFERENCES

1. **Globocan 2012 Stats - Global Cancer Rates Continue to Soar**
2. <http://www2.inca.gov.br/>
3. Jordan RC, Macabeo-Ong M, Shiboski CH, Dekker N, Ginzinger DG, Wong DT, Schmidt BL: **Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer.** *Clin Cancer Res* 2004, **10**(19):6460-6465.
4. Hecht SS: **Tobacco smoke carcinogens and lung cancer.** *J Natl Cancer Inst* 1999, **91**(14):1194-1210.
5. Nattey C, Urban M, Kielkowski D: **0342 Lung cancer risk attributable to occupation: in a case control study in black South Africans, 2001-2008.** *Occup Environ Med* 2014, **71 Suppl 1**:A110-111.
6. Schmid K, Kuwert T, Drexler H: **Radon in indoor spaces: an underestimated risk factor for lung cancer in environmental medicine.** *Dtsch Arztebl Int* 2010, **107**(11):181-186.
7. Herbst RS, Heymach JV, Lippman SM: **Lung cancer.** *N Engl J Med* 2008, **359**(13):1367-1380.
8. Dienstmann R PH, Zukin M: **Tratamento atual do cancer de pulmao nao-pequenas celulas avancado.** *Pratica Hospitalar* 2007; 51: 93-8.; 2007.
9. Wahbah M, Boroumand N, Castro C, El-Zeky F, Eltorkey M: **Changing trends in the distribution of the histologic types of lung cancer: a review of 4,439 cases.** *Ann Diagn Pathol* 2007, **11**(2):89-96.
10. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG *et al*: **Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.** *N Engl J Med* 2004, **350**(21):2129-2139.
11. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ *et al*: **EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.** *Science* 2004, **304**(5676):1497-1500.
12. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L *et al*: **EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib.** *Proc Natl Acad Sci U S A* 2004, **101**(36):13306-13311.
13. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR: **The biology of epidermal growth factor receptor in lung cancer.** *Clin Cancer Res* 2004, **10**(12 Pt 2):4227s-4232s.
14. Sharma SV, Bell DW, Settleman J, Haber DA: **Epidermal growth factor receptor mutations in lung cancer.** *Nat Rev Cancer* 2007, **7**(3):169-181.
15. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Jänne PA, Costa DB *et al*: **Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer.** *N Engl J Med* 2010, **363**(18):1693-1703.
16. Camidge DR, Bang YJ, Kwak EL, Iafrate AJ, Varella-Garcia M, Fox SB, Riely GJ, Solomon B, Ou SH, Kim DW *et al*: **Activity and safety of crizotinib in patients with ALK-positive non-small-cell lung cancer: updated results from a phase 1 study.** *Lancet Oncol* 2012, **13**(10):1011-1019.
17. **NCCN Clinical Practices Guidelines in Oncology (NCCN Guidelines). Non-Small Cell Lung Cancer, version 4.2014**
18. Pérez-Soler R: **Individualized therapy in non-small-cell lung cancer: future versus current clinical practice.** *Oncogene* 2009, **28 Suppl 1**:S38-45.

19. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**(2):281-297.
20. Iorio MV, Croce CM: **MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review.** *EMBO Mol Med* 2012, **4**(3):143-159.
21. Harfe BD: **MicroRNAs in vertebrate development.** *Curr Opin Genet Dev* 2005, **15**(4):410-415.
22. Cummins JM, Velculescu VE: **Implications of micro-RNA profiling for cancer diagnosis.** *Oncogene* 2006, **25**(46):6220-6227.
23. Rinaldi A, Poretti G, Kwee I, Zucca E, Catapano CV, Tibiletti MG, Bertoni F: **Concomitant MYC and microRNA cluster miR-17-92 (C13orf25) amplification in human mantle cell lymphoma.** *Leuk Lymphoma* 2007, **48**(2):410-412.
24. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH: **Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells.** *J Biol Chem* 2008, **283**(2):1026-1033.
25. Reis PP, Tomenson M, Cervigne NK, Machado J, Jurisica I, Pintilie M, Sukhai MA, Perez-Ordóñez B, Grénman R, Gilbert RW *et al*: **Programmed cell death 4 loss increases tumor cell invasion and is regulated by miR-21 in oral squamous cell carcinoma.** *Mol Cancer* 2010, **9**:238.
26. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**(2):126-139.
27. Ha M, Kim VN: **Regulation of microRNA biogenesis.** *Nat Rev Mol Cell Biol* 2014, **15**(8):509-524.
28. Kent OA, Mendell JT: **A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes.** *Oncogene* 2006, **25**(46):6188-6196.
29. Shenouda SK, Alahari SK: **MicroRNA function in cancer: oncogene or a tumor suppressor?** *Cancer Metastasis Rev* 2009, **28**(3-4):369-378.
30. Iorio MV, Croce CM: **MicroRNAs in cancer: small molecules with a huge impact.** *J Clin Oncol* 2009, **27**(34):5848-5856.
31. Kasinski AL, Slack FJ: **Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy.** *Nat Rev Cancer* 2011, **11**(12):849-864.
32. Joshi P, Middleton J, Jeon YJ, Garofalo M: **MicroRNAs in lung cancer.** *World J Methodol* 2014, **4**(2):59-72.
33. Liu J, Lu KH, Liu ZL, Sun M, De W, Wang ZX: **MicroRNA-100 is a potential molecular marker of non-small cell lung cancer and functions as a tumor suppressor by targeting polo-like kinase 1.** *BMC Cancer* 2012, **12**:519.
34. Vucic EA, Thu KL, Pikor LA, Enfield KS, Yee J, English JC, MacAulay CE, Lam S, Jurisica I, Lam WL: **Smoking status impacts microRNA mediated prognosis and lung adenocarcinoma biology.** *BMC Cancer* 2014, **14**:778.
35. Garofalo M, Jeon YJ, Nuovo GJ, Middleton J, Secchiero P, Joshi P, Alder H, Nazaryan N, Di Leva G, Romano G *et al*: **MiR-34a/c-Dependent PDGFR- $\alpha/\beta$  Downregulation Inhibits Tumorigenesis and Enhances TRAIL-Induced Apoptosis in Lung Cancer.** *PLoS One* 2013, **8**(6):e67581.
36. Ling H, Fabbri M, Calin GA: **MicroRNAs and other non-coding RNAs as targets for anticancer drug development.** *Nat Rev Drug Discov* 2013, **12**(11):847-865.
37. Bouchie A: **First microRNA mimic enters clinic.** *Nat Biotechnol* 2013, **31**(7):577.
38. Chomczynski P, and Sacchi, N.: **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** . In., vol. 162: *Anal. Biochem*; 1987: 156-159.

39. Reis PP, Waldron L, Perez-Ordonez B, Pintilie M, Galloni NN, Xuan Y, Cervigne NK, Warner GC, Makitie AA, Simpson C *et al*: **A gene signature in histologically normal surgical margins is predictive of oral carcinoma recurrence.** *BMC Cancer* 2011, **11**:437.
40. Goswami RS, Atenafu EG, Xuan Y, Waldron L, Reis PP, Sun T, Datti A, Xu W, Kuruvilla J, Good DJ *et al*: **MicroRNA signature obtained from the comparison of aggressive with indolent non-Hodgkin lymphomas: potential prognostic value in mantle-cell lymphoma.** *J Clin Oncol* 2013, **31**(23):2903-2911.
41. Andrews S: **FastQC A Quality Control tool for High Throughput Sequence Data.** In. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; 2010.
42. MARTIN M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** . In., vol. 17. **EMBnet.journal:** <http://journal.embnet.org/index.php/embnetjournal/article/view/200>; 2011: 10-12.
43. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**(3):R25.
44. Simon Anders PTP, Wolfgang Huber: **HTSeq - A Python framework to work with high-throughput sequencing data.** In. bioRxiv preprint; 2014.
45. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data.** *Genome Biol* 2010, **11**(3):R25.
46. Robinson MD, Smyth GK: **Small-sample estimation of negative binomial dispersion, with applications to SAGE data.** *Biostatistics* 2008, **9**(2):321-332.
47. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139-140.
48. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J *et al*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**(10):R80.
49. Li H: **A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.** *Bioinformatics* 2011, **27**(21):2987-2993.
50. Li H: **Improving SNP discovery by base alignment quality.** *Bioinformatics* 2011, **27**(8):1157-1158.
51. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al*: **The variant call format and VCFtools.** *Bioinformatics* 2011, **27**(15):2156-2158.
52. Wang K, Li M, Hakonarson H: **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.** *Nucleic Acids Res* 2010, **38**(16):e164.
53. Goswami RS, Waldron L, Machado J, Cervigne NK, Xu W, Reis PP, Bailey DJ, Jurisica I, Crump MR, Kamel-Reid S: **Optimization and analysis of a quantitative real-time PCR-based technique to determine microRNA expression in formalin-fixed paraffin-embedded samples.** *BMC Biotechnol* 2010, **10**:47.
54. Di Leva G, Garofalo M, Croce CM: **MicroRNAs in cancer.** *Annu Rev Pathol* 2014, **9**:287-314.
55. Markou A, Sourvinou I, Vorkas PA, Yousef GM, Lianidou E: **Clinical evaluation of microRNA expression profiling in non small cell lung cancer.** *Lung Cancer* 2013, **81**(3):388-396.

56. Heegaard NH, Schetter AJ, Welsh JA, Yoneda M, Bowman ED, Harris CC: **Circulating micro-RNA expression profiles in early stage non-small cell lung cancer.** *Int J Cancer* 2012, **130**(6):1378-1386.
57. Perry MM, Moschos SA, Williams AE, Shepherd NJ, Larner-Svensson HM, Lindsay MA: **Rapid changes in microRNA-146a expression negatively regulate the IL-1beta-induced inflammatory response in human lung alveolar epithelial cells.** *J Immunol* 2008, **180**(8):5689-5698.
58. Engels EA: **Inflammation in the development of lung cancer: epidemiological evidence.** *Expert Rev Anticancer Ther* 2008, **8**(4):605-615.
59. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A *et al*: **Circulating microRNAs as stable blood-based markers for cancer detection.** *Proc Natl Acad Sci U S A* 2008, **105**(30):10513-10518.
60. Arora S, Ranade AR, Tran NL, Nasser S, Sridhar S, Korn RL, Ross JT, Dhruv H, Foss KM, Sibenaller Z *et al*: **MicroRNA-328 is associated with (non-small) cell lung cancer (NSCLC) brain metastasis and mediates NSCLC migration.** *Int J Cancer* 2011, **129**(11):2621-2631.
61. Ulivi P, Foschi G, Mengozzi M, Scarpi E, Silvestrini R, Amadori D, Zoli W: **Peripheral Blood miR-328 Expression as a Potential Biomarker for the Early Diagnosis of NSCLC.** *Int J Mol Sci* 2013, **14**(5):10332-10342.
62. Foss KM, Sima C, Ugolini D, Neri M, Allen KE, Weiss GJ: **miR-1254 and miR-574-5p: serum-based microRNA biomarkers for early-stage non-small cell lung cancer.** *J Thorac Oncol* 2011, **6**(3):482-488.
63. Gu W, Fang S, Gao L, Tan Y, Yang Z: **Clinic significance of microRNA-99a expression in human lung adenocarcinoma.** *J Surg Oncol* 2013, **108**(4):248-255.
64. Yu S, Zhang C, Deng F, Zhang Y: **miR-99a suppresses the metastasis of human non-small cell lung cancer cells by targeting AKT1 signaling pathway.** *J Cell Biochem* 2014.
65. Fennell DA: **Caspase regulation in non-small cell lung cancer and its potential for therapeutic exploitation.** *Clin Cancer Res* 2005, **11**(6):2097-2105.
66. Uematsu K, He B, You L, Xu Z, McCormick F, Jablons DM: **Activation of the Wnt pathway in non small cell lung cancer: evidence of dishevelled overexpression.** *Oncogene* 2003, **22**(46):7218-7221.
67. MacDonald BT, Tamai K, He X: **Wnt/beta-catenin signaling: components, mechanisms, and diseases.** *Dev Cell* 2009, **17**(1):9-26.
68. Tan M, Wu J, Cai Y: **Suppression of Wnt signaling by the miR-29 family is mediated by demethylation of WIF-1 in non-small-cell lung cancer.** *Biochem Biophys Res Commun* 2013, **438**(4):673-679.
69. Zhu W, Shan X, Wang T, Shu Y, Liu P: **miR-181b modulates multidrug resistance by targeting BCL2 in human cancer cell lines.** *Int J Cancer* 2010, **127**(11):2520-2529.
70. Colamaio M, Borbone E, Russo L, Bianco M, Federico A, Califano D, Chiappetta G, Pallante P, Troncone G, Battista S *et al*: **miR-191 down-regulation plays a role in thyroid follicular tumors through CDK6 targeting.** *J Clin Endocrinol Metab* 2011, **96**(12):E1915-1924.
71. Nagpal N, Kulshreshtha R: **miR-191: an emerging player in disease biology.** *Front Genet* 2014, **5**:99.
72. Silva J, García V, Zaballos Á, Provencio M, Lombardía L, Almonacid L, García JM, Domínguez G, Peña C, Diaz R *et al*: **Vesicle-related microRNAs in plasma of**

- nonsmall cell lung cancer patients and correlation with survival. *Eur Respir J* 2011, **37**(3):617-623.
73. Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, Zeng Y, Miao R, Jin G, Ma H *et al*: **Genetic variants of miRNA sequences and non-small cell lung cancer survival.** *J Clin Invest* 2008, **118**(7):2600-2608.
74. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y *et al*: **Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival.** *Cancer Res* 2004, **64**(11):3753-3756.
75. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.** *Cancer Cell* 2006, **9**(3):189-198.
76. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D, Slack FJ: **RAS is regulated by the let-7 microRNA family.** *Cell* 2005, **120**(5):635-647.
77. He XY, Chen JX, Zhang Z, Li CL, Peng QL, Peng HM: **The let-7a microRNA protects from growth of lung carcinoma by suppression of k-Ras and c-Myc in nude mice.** *J Cancer Res Clin Oncol* 2010, **136**(7):1023-1028.
78. Khodayari N, Mohammed KA, Goldberg EP, Nasreen N: **EphrinA1 inhibits malignant mesothelioma tumor growth via let-7 microRNA-mediated repression of the RAS oncogene.** *Cancer Gene Ther* 2011, **18**(11):806-816.
79. Miao H, Wei BR, Peehl DM, Li Q, Alexandrou T, Schelling JR, Rhim JS, Sedor JR, Burnett E, Wang B: **Activation of EphA receptor tyrosine kinase inhibits the Ras/MAPK pathway.** *Nat Cell Biol* 2001, **3**(5):527-530.
80. Lee HY, Mohammed KA, Kaye F, Sharma P, Moudgil BM, Clapp WL, Nasreen N: **Targeted delivery of let-7a microRNA encapsulated ephrin-A1 conjugated liposomal nanoparticles inhibit tumor growth in lung cancer.** *Int J Nanomedicine* 2013, **8**:4481-4494.

**Table 1.** Clinical and histopathological data of patients of the test sample set.

<b>Variables</b>	<b>Total Number</b>	<b>N (%)</b>
<b>Ages (years)</b>		
Median (range)	66,8(43-83)	
Mean	65	
<b>Gender</b>		
Male	8	47
Female	9	53
<b>Tobacco use</b>		
Yes	10	59
No	7	41
<b>Alcohol use</b>		
Yes	1	6
No	16	94
<b>Histology</b>		
Adenocarcinoma	17	100
<b>Tumor grade</b>		
Well differentiated	5	29
Moderately differentiated	10	59
Poorly differentiated	2	12
<b>T category</b>		
T1-T2	15	88
T3-T4	2	12
<b>Nodal status</b>		
Negative (N0)	12	71
Positive (N1, N2)	5	29
<b>Distance metastasis</b>		
Yes	5	29
No	12	71
<b>Tumor stage</b>		
Ia; Ib; IIa; IIb	14	82
IIIa; IIIb	3	18
<b>Outcome</b>		
Alive with disease	15	88
Dead of disease	2	12



**Table 2.** Clinical and histopathological data of patients of the validation sample set.

<b>Variables</b>	<b>Total Number</b>	<b>N (%)</b>
<b>Ages (years)</b>		
Median (range)	58,0(40-84 anos)	
Mean	60,5	
<b>Gender</b>		
Male	12	55
Female	10	45
<b>Tobacco use</b>		
Yes	15	68
No	7	32
<b>Alcohol use</b>		
Yes	7	32
No	15	68
<b>Histology</b>		
Adenocarcinoma	22	100
<b>Tumor grade</b>		
Well differentiated	3	14
Moderately differentiated	12	55
Poorly differentiated	7	32
<b>T category</b>		
T1-T2	13	59
T3-T4	9	41
<b>Nodal status</b>		
Negative (N0)	12	55
Positive (N1, N2)	10	45
<b>Distance metastasis</b>		
Yes	2	9
No	20	91
<b>Tumor stage</b>		
Ia; Ib; IIa; IIb	13	59
IIIa; IIIb	9	41
<b>Outcome</b>		
Alive with disease	15	68
Dead of disease	7	32

**Table 3.** Differentially expressed miRNAs in adenocarcinoma (AD) compared to normal (N) in unpaired samples and paired samples.

<b>miRNA</b>	<b>LogFC</b>	<b><i>P</i>-value</b>	<b>FDR</b>
<b>Unpaired samples</b>			
miR-486	-1,635	0,0004	0,0689
let-7g	0,8879	0,0015	0,0689
miR-15b	0,9831	0,0004	0,0689
miR-181b-2	0,9887	0,0009	0,0689
miR-23b	0,9479	0,0014	0,0689
miR-26a-1	0,9756	0,0014	0,0689
miR-26a-2	0,9595	0,0013	0,0689
miR-26b	1,0565	0,0009	0,0689
miR-93	1,203	0,0005	0,0689
let-7a-2	0,8844	0,0025	0,1052
miR-181b-1	0,9221	0,0027	0,1052
<b>Paired samples</b>			
miR-486	-2,5430	8,8276.E-12	1,805.E-09
miR-1247	-2,6582	3,2911.E-08	3,3651.E-06
miR-218-1	-1,6175	0,0001	0,0043
miR-181a-2	-1,1145	0,0002	0,0079
miR-328	-1,4708	0,0004	0,0108
miR-181a-1	-1,1444	0,0006	0,0147
miR-574	-1,1719	0,0036	0,0503
miR-886	-1,3187	0,0046	0,0575
let-7a-3	0,4856	3,2638.E-264	5,1692.E-262
miR-146b	0,7012	0,0005	0,0147
miR-26a-1	0,7382	0,0043	0,0551
miR-200b	1,0898	0,0072	0,0864
miR-191	1,0935	0,0018	0,0307
miR-181c	1,1230	0,0041	0,0551
miR-10a	1,2858	0,0009	0,0208
miR-155	1,3089	0,0029	0,0418
miR-99a	1,4447	0,0028	0,0418
miR-30e	1,4926	0,00002	0,0011
miR-21	1,5219	0,0028	0,0418
miR-425	1,6206	7,6888.E-06	0,0005
miR-29a	1,6488	0,0001	0,0043
miR-421	1,7648	0,0085	0,0993

LogFC = log fold change; FDR = false discovery rate. FDR $\leq$ 0,1,  $p \leq$ 0,01.

**Table 4.** List of miRNAs with concordant expression between transcriptome sequencing and TaqMan PCR analyses.

<b>miRNA</b>	<b>miRNA-Seq</b>		<b>miRNA</b>	<b>TLDA</b>	
	<b>logFC</b>	<b>P-value</b>		<b>logFC</b>	<b>P-value</b>
let-7a-2	0,8844	0,0025	let-7a	2,3601	<b>0,000*</b>
let-7a-3	0,4856	3,2638.E-264	let-7a	2,3601	<b>0,000*</b>
miR-15b	0,9831	0,0004	miR-15b	2,4022	<b>0,000*</b>
miR-200b	1,0898	0,0072	miR-200b	1,9404	<b>0,000*</b>
miR-21	1,5219	0,0028	miR-21	2,6311	<b>0,001*</b>
miR-155	1,3089	0,0029	miR-155	1,3294	<b>0,010*</b>
miR-486	-1,635	0,0004	miR-486	-3,9434	0,084
miR-181a-1	-1,1444	0,0006	miR-181a	-0,8625	0,140
miR-181a-2	-1,1145	0,0002	miR-181a	-0,8625	0,140
let-7g	0,8879	0,0015	let-7g	0,3918	0,143
miR-26b	1,0565	0,0009	miR-26b	0,5597	0,224
miR-218-1	-1,6175	0,0001	miR-218	-1,4699	0,386
miR-886	-1,3187	0,0046	miR-886-3p	-1,1746	0,482
miR-425	1,6206	7,6888.E-06	miR-425-5p	0,1190	0,906

Test sample set:  $p \leq 0,01$  and  $FDR \leq 0,1$ , as determined by EdgeR software.

Validation sample set:  $p \leq 0,01$  as determined by Expression Suite software.

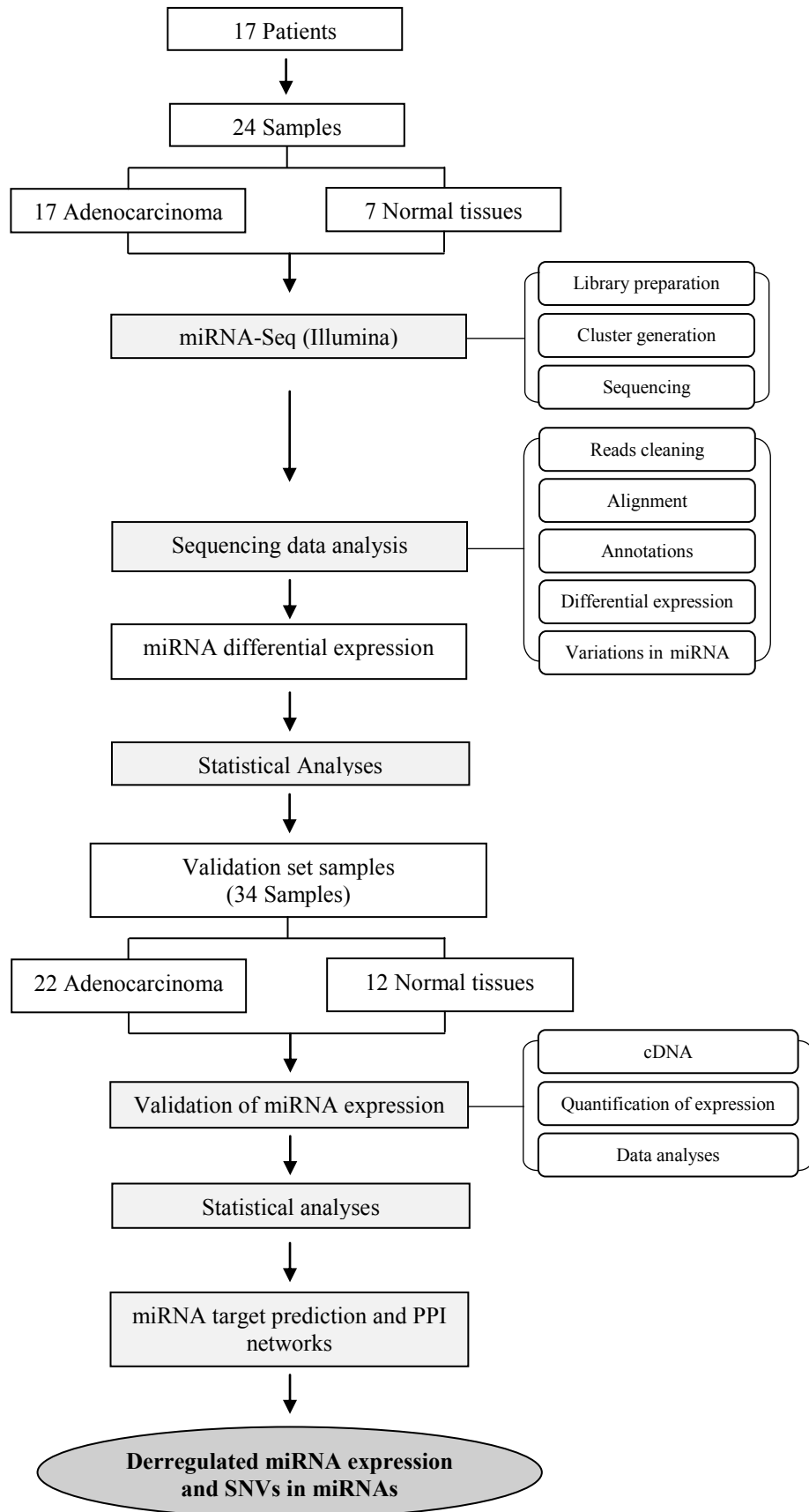
\*Statistically significantly deregulated miRNAs identified in both test and validation sets by transcriptome sequencing and TaqMan PCR analyses.

**Table 5.** SNVs in miRNAs identified in adenocarcinoma.

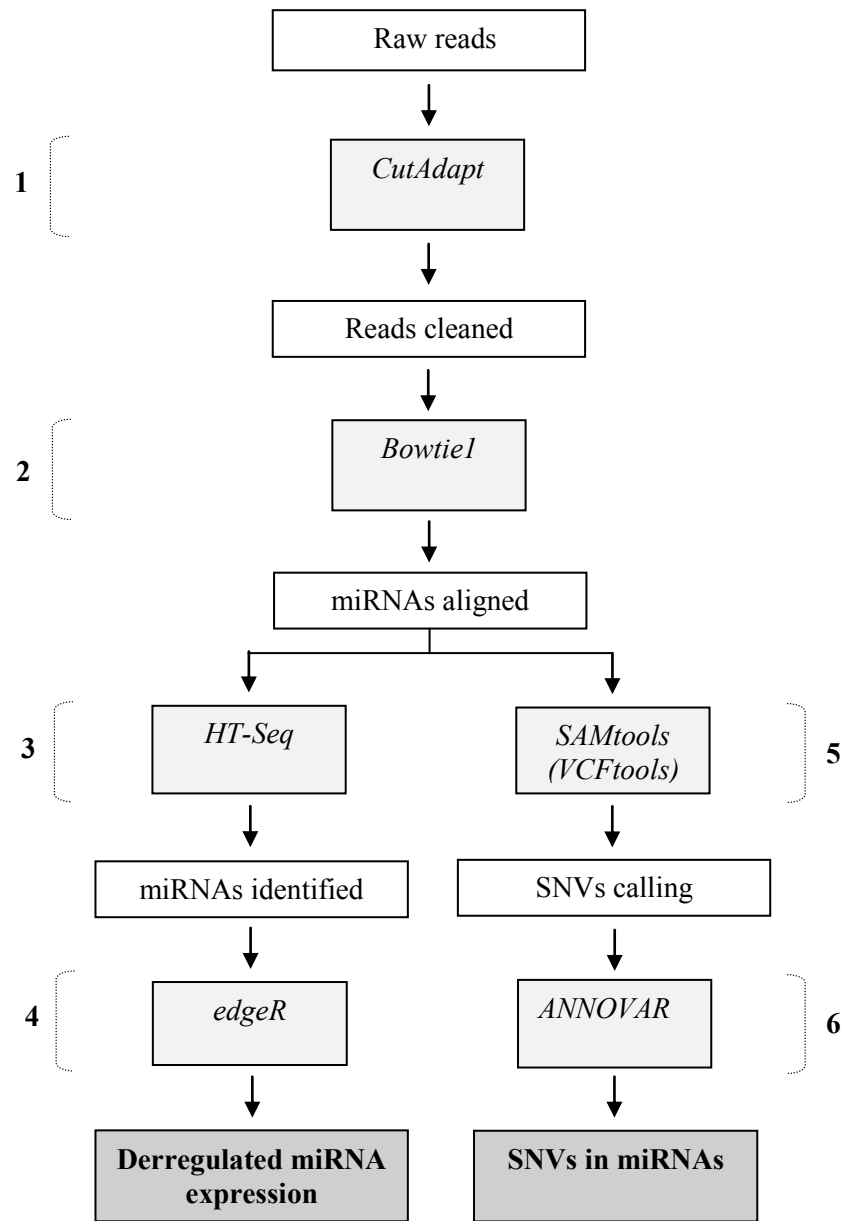
<b>Chromosome</b>	<b>Start</b>	<b>End</b>	<b>WT</b>	<b>SNV</b>	<b>miRNA genomic site</b>	<b>miRNA</b>
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	Intergenic	mir-4490
Chr21	9827315	9827315	T	C	Intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	mir-let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	mir-let-7c
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr1	16861827	16861827	G	A	Intergenic	mir-3675
Chr11	122017274	122017274	T	C	ncRNA_exonic	mir-let-7a-2
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-24-1
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-3074
Chr12	54731084	54731084	C	A	ncRNA_exonic	mir-148b
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200a
Chr17	46657243	46657243	A	G	ncRNA_exonic	mir-10a
Chr1	56633037	56633037	C	T	Intergenic	mir-4422
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr1	186105018	186105018	G	A	ncRNA_exonic	mir-548f-1
Chr1	186105020	186105020	G	A	ncRNA_exonic	mir-548f-1
Chr2	227230043	227230043	A	G	Intergenic	mir-5702
Chr8	143215944	143215944	G	C	Intergenic	mir-4472-1
Chr9	97847569	97847569	C	T	ncRNA_exonic	mir-23b
Chr14	53927750	53927750	T	G	Intergenic	mir-5580
Chr11	122022981	122022981	A	T	ncRNA_exonic	mir-100
Chr21	17911436	17911436	T	A	ncRNA_exonic	mir-99a
Chr3	48261476	48261476	C	A	Intergenic	mir-4443
Chr19	46522264	46522264	T	A	ncRNA_exonic	mir-769

**Table 6.** SNVs in miRNAs identified in histologically normal lung tissue.

<b>Chromosome</b>	<b>Start</b>	<b>End</b>	<b>WT</b>	<b>SNV</b>	<b>miRNA genomic site</b>	<b>miRNA</b>
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200A
Chr4	10917451	10917451	T	G	intergenic	mir-572
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	intergenic	mir-4490
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr21	9827315	9827315	T	C	intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	mir-let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	mir-let-7c
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr11	122017274	122017274	T	C	ncRNA_exonic	mir-let-7a-2
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195896	52195896	C	A	ncRNA_exonic	mir-99b
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr8	143215944	143215944	G	C	intergenic	mir-4472-1
Chr9	96938260	96938260	A	G	ncRNA_exonic	mir-let-7a-1
Chr9	96938262	96938262	A	G	ncRNA_exonic	mir-let-7a-1



**Figure 1.** Experimental design and samples



**Figure 2.** Sequencing data analyses steps: 1. Reads cleaning; 2. Alignment; 3. Annotations; 4. Differential expression analyses; 5. SNVs calling; 6. SNVs annotations.

## **Sequenciamento de miRNAs identifica potenciais biomarcadores de adenocarcinoma pulmonar de pacientes Brasileiros**

Cinegaglia NC<sup>1</sup>, Andrade SCS<sup>2</sup>, Pinheiro M<sup>3</sup>, Severino FE<sup>1</sup>, Oliveira RA<sup>4</sup>, Hasimoto E<sup>1</sup>, Cataneo DC<sup>1</sup>, Cataneo AJM<sup>1</sup>, Defaveri J<sup>5</sup>, Souza CP<sup>1,6</sup>, Marques MMC<sup>6,7</sup>, Carvalho RF<sup>8</sup>, Coutinho LL<sup>2</sup>, Gross JL<sup>9</sup>, Rogatto SR<sup>9,10</sup>, Lam WL<sup>11</sup>, Jurisica I<sup>12,13</sup>, Reis PP<sup>1</sup>.

<sup>1</sup>Departamento de Cirurgia e Ortopedia, Faculdade de Medicina, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>2</sup>Departamento de Biotecnologia Animal, Universidade de São Paulo (USP), Piracicaba, SP, Brasil

<sup>3</sup>Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>4</sup>Departamento de Bioestatística, Instituto de Biociências, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>5</sup>Departamento de Patologia, Faculdade de Medicina, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>6</sup>Centro de Oncologia Molecular, Hospital de Câncer de Barretos, Barretos, SP, Brasil

<sup>7</sup>Escola de Ciências da Saúde de Barretos, Barretos, SP, Brasil

<sup>8</sup>Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>9</sup>Centro Internacional de Pesquisa e Ensino (CIPE), Hospital A. C. Camargo, São Paulo, SP

<sup>10</sup>Departamento de Urologia, Faculdade de Medicina, Universidade Estadual Paulista, UNESP, Botucatu, SP, Brasil

<sup>11</sup>Departamento de Oncologia Integrada, Centro de Pesquisa em Câncer British Columbia, Vancouver, British Columbia V5Z 1L3, Canada

<sup>12</sup>University Health Network Divisão de Informática no Câncer, Ontario Cancer Institute e Princess Margaret Hospital, Toronto, Ontario, Canada

<sup>13</sup>Departamento de Ciências da Computação, University of Toronto, Toronto, Ontario, Canada

### **Autor correspondente**

Dra. Patricia Reis, PhD

Professor Assistente Doutor

Dept. de Cirurgia e Ortopedia, Faculdade de Medicina

Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)

Av. Prof. Montenegro, S/N. CEP 18618-970

Botucatu, SP.

BRASIL

Fone: +55 (14) 3880-1451

Email: [preis@fmb.unesp.br](mailto:preis@fmb.unesp.br)



## RESUMO

**Introdução:** O câncer de pulmão é uma causa líder de mortalidade por câncer, com uma incidência, no Brasil, de >27.000 novos casos/ano. Embora os avanços existentes nos tratamentos com alvos moleculares, altas taxas de mortalidade ainda são observadas. Portanto, há uma necessidade de identificação de biomarcadores e vias moleculares desreguladas clinicamente úteis para melhorar a sobrevida dos pacientes. microRNAs (miRNAs) são importantes reguladores da expressão gênica envolvidos no processo de oncogênese.

**Objetivos:** Identificar alterações no transcrito global de miRNAs, genes-alvo e vias moleculares moduladas por miRNAs em adenocarcinoma de pacientes brasileiros.

**Pacientes e Métodos:** Amostras de pacientes foram selecionadas como grupo teste (N=24) e grupo independente de validação (N=34), incluindo tumores e tecidos pulmonares histologicamente normais. As amostras teste foram submetidas ao sequenciamento do transcrito de miRNAs seguida da validação utilizando TaqMan PCR.

**Resultados:** A análise de amostras não-pareadas mostrou que 11 miRNAs estavam significativamente desregulados ( $p \leq 0,01$  e  $FDR \leq 0,1$ ) incluindo o miRNA miR-486 com expressão diminuída e os miRNAs (let-7a-2, let-7g, miR-15b, miR-181b-1, miR-181b-2, miR-23b, miR-26a-1, miR-26a-2, miR-26b, miR-93) com expressão aumentada no tumor comparado com tecido normal. Adicionalmente, a análise de amostras pareadas mostrou que 22 miRNAs estavam significativamente desregulados ( $p \leq 0,01$  e  $FDR \leq 0,1$ ); 8 miRNAs (miR-486, miR-1247, miR-218-1, miR-181a-1, miR-181a-2, miR-328, miR-574 e miR-886) estavam com expressão diminuída e 14 miRNAs (let-7a-3, miR-146b, miR-26a-1, miR-200b, miR-191, miR-181c, miR-10a, miR-155, miR-99a, miR-30e, miR-21, miR-425, miR-29a e miR-421) estavam com expressão aumentada no tumor comparado com tecido normal do mesmo paciente.

**Conclusões:** let-7a-2, let-7a-3, miR-15b, miR-21, miR-155 e miR-200b regulam a expressão de vários genes-alvo interconectados com vias moleculares relevantes ao desenvolvimento tumoral, progressão e metástase. Nossos achados sugerem miRNAs candidatos à caracterização funcional futura, os quais podem ter um papel potencial no desenvolvimento de novas estratégias terapêuticas em adenocarcinoma pulmonar.

## INTRODUÇÃO

O carcinoma pulmonar é uma das principais causas de morte por câncer com >1,5 milhões de mortes mundialmente [1]. Em países em desenvolvimento como o Brasil, dados de incidência do Instituto Nacional de Câncer (INCA) mostram ~27.000 novos casos/ano [2]. No Brasil, o Estado de São Paulo apresenta um dos maiores números de casos de cânceres que afetam a traquéia, brônquios e pulmões, com incidência de ~22,9/100.000 homens e ~11,8 /100.000 mulheres [3].

O carcinoma pulmonar é frequentemente diagnosticado em estádios avançados da doença e associado a baixas taxas de sobrevida de 13-21% nos países desenvolvidos e de 7-10% nos países em desenvolvimento [2]. Os fatores de risco incluem o uso do tabaco na maioria dos casos, além da exposição à fumaça do cigarro, radiação, fatores ocupacionais e ambientais, entre outros [4-6].

Atualmente, o câncer de pulmão é classificado em dois grupos histológicos principais, que são utilizados para auxiliar as decisões de tratamento [7]. O carcinoma de pulmão de células não pequenas (da sigla em inglês NSCLC, *Non-Small Cell Lung Carcinoma*) compreende a maioria dos casos, classificados em vários subtipos histológicos; os mais comuns são adenocarcinoma, carcinoma de células escamosas e carcinoma de grandes células [7, 8]. Entre estes, adenocarcinoma é o subtipo mais frequente, ocorrendo em ~47% das mulheres e 37% dos homens diagnosticados com NSCLC [7, 9].

Recentes avanços no tratamento dos pacientes com NSCLC incluem estratégias alvo moleculares, como o uso de inibidores tirosina-quinase Gefitinibe e Erlotinibe para pacientes com tumores *EGFR* positivo (principalmente deleções nos éxons 19 e 20) [10-14] e Crizotinibe para pacientes com tumores contendo rearranjo *ALK* [15, 16]; protocolos atuais de tratamento estão descritos no *National Comprehensive Cancer Network Guidelines* ([http://www.nccn.org/professionals/physician\\_gls/f\\_guidelines.asp](http://www.nccn.org/professionals/physician_gls/f_guidelines.asp)) [17].

Embora vários estudos tenham identificado alterações genômicas e transcrito incluindo mutações em NSCLC, ~50% das alterações que dirigem a tumorigênese (*drivers*) ainda são desconhecidas em adenocarcinoma. Considerando que alvos moleculares beneficiam uma pequena parcela de pacientes com adenocarcinoma, ainda é necessário identificar novas alterações do tipo *driver* e vias moleculares que podem ser utilizadas como alvos terapêuticos [18].

Nos últimos anos, têm-se evidenciado o papel dos microRNAs (miRNAs) em várias doenças, incluindo o câncer. miRNAs são RNAs pequenos (~18-22 nucleotídeos), não

codificadores de proteínas, transcritos a partir do DNA e com importante papel na regulação pós-transcricional da expressão gênica levando a repressão da tradução [19]. Interessantemente, miRNAs também podem se ligar a ribonucleoproteínas na sequência *seed* de modo independente de RISC (aqui definido pela sigla em inglês, *RNA Induced Silencing Complex*) e conseqüentemente, influenciando na ligação miRNA:mRNA e levando à ativação da tradução. miRNAs também podem modular a expressão gênica em nível transcricional, através da ligação direta com o DNA [20]. A biogênese e processamento do miRNA através das vias canônicas e não canônicas demonstra como funções múltiplas podem ser realizadas pelo RNA transcrito na célula.

Sabe-se que os miRNAs desempenham papéis importantes em múltiplos processos biológicos, assim como o desenvolvimento embrionário e diferenciação, apoptose, proliferação celular [19, 21, 22] e em mecanismos de oncogênese [23-25]. De acordo com o miRBase, foram identificados 2.588 miRNAs maduros no genoma humano (<http://www.mirbase.org/cgi-bin/browse.pl?org=hsa>); miRNAs encontram-se mapeados em regiões genômicas distintas incluindo íntrons ou éxons de transcritos não codificadores e codificadores ou podem constituir íntrons completos nomeados miRtrons [26, 27]. No câncer, os miRNAs atuam como oncogenes e genes supressores tumorais [28-30].

Diferentes mecanismos podem levar a desregulação de miRNA celular incluindo deleções genômicas, ampliações, mutações e polimorfismos, desregulação transcricional e alterações na maquinaria da biogênese ou modificações epigenéticas [31]. A expressão desregulada de miRNAs tem sido associada com a tumorigênese pulmonar [32-35]. Atualmente, vias moleculares alvos de miRNAs estão sendo investigadas como abordagens terapêuticas [36]. Por exemplo, ensaios clínicos de fase I estão testando o potencial do miR-34 *mimics* em diferentes tumores sólidos; o miR-34 funciona como supressor tumoral inibindo a expressão de >20 oncogenes com papel no ciclo celular e proliferação, apoptose, metástases e quimioresistência de células tumorais [37].

Considerando a importância dos miRNAs na tumorigênese, a caracterização das alterações na expressão do transcrito de miRNAs contribuem com o desenvolvimento de estratégias terapêuticas em NSCLC. Para nosso conhecimento, esse é o primeiro estudo de sequenciamento global de miRNAs em tumores de pacientes brasileiros. Adicionalmente, nosso estudo inclui análise de amostras de tecidos pulmonares histologicamente normais, o que nos permitiu identificar alterações em tecido tumoral e normal dos mesmos pacientes, bem como o perfil global de alterações em miRNAs em adenocarcinoma pulmonar nesse grupo de pacientes.

## **MATERIAL E MÉTODOS**

### **Pacientes e amostras**

Este estudo foi aprovado pelos Comitês de Ética em Pesquisa (CEP) da Faculdade de Medicina, UNESP, Botucatu, SP (4319/2012), Hospital AC Camargo, São Paulo, SP (1573/11) e Hospital de Câncer de Barretos, Barretos, SP (75907). O Termo de Consentimento foi obtido de todos os pacientes elegíveis ao estudo antes da coleta da amostra. Os critérios de inclusão foram pacientes maiores de 18 anos, diagnóstico de adenocarcinoma de pulmão, previamente tratados por cirurgia. Os critérios de exclusão foram pacientes menores de 18 anos e com diagnóstico de outras doenças.

As amostras dos pacientes foram selecionadas como grupo teste e grupo de validação. Amostras do grupo teste (N=24) foram respectivamente coletadas de cirurgias realizadas no Hospital AC Camargo, SP e constituindo de tumores (adenocarcinoma pulmonar, N=17) e tecido de pulmão histologicamente normal dos mesmos pacientes (N=7). Todas as amostras de tecidos respectivamente coletadas foram imediatamente congeladas em nitrogênio líquido e mantidos em freezer -80°C até a extração do RNA. As amostras do grupo de validação (N=34) foram retrospectivamente obtidas no período de 2000-2012 do Departamento de Patologia da Faculdade de Medicina, UNESP, Botucatu, SP e Hospital de Câncer de Barretos, Barretos, SP. Amostras do grupo de validação constituiu em blocos de tecido de tumores (adenocarcinoma pulmonar) fixadas em formalina e embebidas em parafina (FFPE), e tecido de pulmão histologicamente normal dos mesmos pacientes (N=12). As amostras dos pacientes foram então obtidas de três hospitais diferentes localizados no Estado de São Paulo e a população de pacientes apresentou características demográficas similares.

Tivemos como objetivo identificar um painel de modificações na expressão de miRNAs em adenocarcinoma pulmonar através de sequenciamento de alto desempenho seguido de validação por quantificação em tempo-real RT-PCR (RT-QPCR) usando a plataforma Taqman miRNA array. A **Figura 1** mostra desenho experimental e amostras. Os dados clínicos e histopatológicos detalhados dos pacientes compondo o grupo teste e o grupo de validação estão apresentados nas **Tabelas 1 e 2**, respectivamente.

### **Extração de RNA**

As amostras de tecido congelado foram submetidas a análises de corte de congelamento para garantir a presença de >90% de células normal ou tumoral em amostras de câncer ou de pulmão normal coletadas através da cirurgia, respectivamente. As amostras foram fragmentadas e lisadas utilizando o sistema de homogeneização e lise Precellys 24

(Berting Technologies, Rockville, MD, USA) por 10s e 6.500 rpm. A extração foi então realizada usando o reagente Trizol [38] (Life Technologies, Carlsbad, CA, USA) seguido da purificação usando o kit *RNeasy Mini Kit* (Qiagen, Hilden, Germany), como descrito previamente [39].

O RNA das amostras FFPE foi isolado utilizando o kit *RecoverAll Total Nucleic Acid Isolation* (Ambion/Life Technologies, Carlsbad, CA, USA), seguindo um protocolo previamente relatado com modificações para melhorar o rendimento de RNA [40]. As amostras de RNA foram quantificadas usando *NanoDrop 8000* (Thermo Fisher Scientific, Waltham, MA, USA) e a qualidade do RNA foi determinada utilizando *Bionalyzer 2100* (Agilent Technologies, Santa Clara, CA, USA), seguindo o protocolo do fabricante. As amostras de RNA foram estocadas a -80°C até o uso para a preparação da biblioteca.

### **Sequenciamento de miRNAs (miRNA-Seq)**

1µg de RNA de amostra do grupo teste (N=24) foi utilizada para a preparação da biblioteca, geração de cluster e sequenciamento usando o sistema MiSeq (Illumina, San Diego, CA, USA). O sequenciamento de alto desempenho foi realizado no Laboratório de Biotecnologia Animal, Universidade de São Paulo (USP), Piracicaba, SP. O sequenciamento consistiu na clonagem, *in vitro*, dos fragmentos de RNA em uma plataforma sólida. A plataforma MiSeq gerou fragmentos de leituras únicas de 50 pb. Todos os passos incluindo o preparo da biblioteca, clusterização e sequenciamento seguiram recomendações do fabricante. O sequenciamento ocorre em uma lâmina de fluxo ou superfície de clonagem (*flow cell*), a qual apresenta uma canaleta única (*lane*), possibilitando sequenciar até 12 amostras simultaneamente. A preparação da biblioteca utilizou o kit *TruSeq Small RNA Sample Preparation kit* (Illumina, San Diego, CA, USA). Brevemente, 1µg de RNA foi usado para ligação do adaptador, o qual contém um sítio de ligação para o *primer* de sequenciamento, usado para identificar amostras contendo um *pool* de amostras de RNA, e outro sítio de ligação aos iniciadores (*primers*) presentes na *flow cell* (superfície de clonagem) utilizados na amplificação dos fragmentos pela PCR. A biblioteca de cDNA foi obtida através da amplificação pela PCR, seguindo 11 ciclos de 98°C por 30 segundos, 98°C por 10 segundos; 60 °C, 30 segundos; 72°C, 15 segundos seguido de 72°C, 10 minutos. As amostras da biblioteca foram submetidas à eletroforese em gel de agarose para isolamento do miRNA; as amostras de cDNA foram precipitadas em etanol e quantificadas usando *Qubit 2.0 Fluorometer* (Invitrogen by Life Technologies, Carlsbad, CA, USA). No passo de clusterização, fragmentos ligados a adaptadores foram desnaturados para separação da fita dupla, permitindo que as

moléculas de fita simples então reconheçam os *primers* fixados na *flow cell*, produzindo múltiplas cópias de fragmentos específicos por amplificação em PCR de fase sólida. O sequenciamento de miRNA foi realizado utilizando o kit *MiSeq Reagent Kit v2* (50 ciclos), seguindo o protocolo do fabricante.

### **Análises dos Dados de miRNA-Seq**

A qualidade das leituras geradas (*reads*) foi visualizada pela ferramenta FastQC [41] e a limpeza dos *reads* foi realizada pelo *CutAdapt* [42], o qual permite selecionar os *reads* de boa qualidade para as análises posteriores. Os *reads* que não passaram o controle de qualidade foram então removidos. O alinhamento dos *reads* foi realizado com base no genoma de referência hg19 (<https://genome.ucsc.edu/cgi-bin/hgTracks?hgsid=12832096&chromInfoPage=>), utilizando o *Bowtie1* [43] seguido do *HT-Seq* [44] para anotações e quantificações das sequencias de miRNAs alinhadas. A normalização dos dados [45] e a análise de expressão diferencial foram realizadas através do *edgeR* (Bioconductor/R) [46-48].

Variações em nucleotídeos únicos nas sequencias de miRNAs foram identificadas pelo mapeamento das sequencias usando *SAMtools* [49, 50] e *VCFtools* [51], seguido da anotação genômica utilizando *ANNOVAR* [52]. As etapas da análise dos dados estão apresentadas na **Figura 2**.

### **Validação da expressão de miRNAs**

A expressão de miRNA foi validada utilizando a plataforma *TaqMan® Array Human MicroRNA* (Life Technologies, Foster City, CA, USA), como previamente descrito [53]. Nós usamos o sistema QuantStudio 12K system (Life Technologies, Foster City, CA, USA). A normalização global dos dados foi realizada no programa *Expression Suite* (Life Technologies, Foster City, CA, USA) e o perfil de expressão de miRNAs foi determinado utilizando o programa *RQ Manager v1.2* (Life Technologies, Foster City, CA, USA).

### **Análises estatísticas**

As análises estatísticas foram realizadas para correlacionar a expressão dos miRNAs alterados com os dados clínicos e histopatológico dos pacientes utilizando o teste de Mann-Whitney e o teste exato de Fisher. Testes estatísticos foram aplicados usando os programas PRISMA e SAS 9.1 (*Statistical Analyses System*). A significância estatística foi definida como  $p < 0.05$ .

## RESULTADOS

### Sequenciamento de microRNAs (miRNA-Seq)

A análise de miRNA-Seq gerou 13.135.522 leituras ou *reads* com uma média de 547.313 *reads*/amostra. o teste de qualidade FastQC mostrou que 96,1% (12.623.236,642) dos *reads* tinham um valor de *Q-score*  $\geq 30$ , e foram então consideradas para as análises posteriores. A análise global (não pareada) de todas as amostras mostrou que 11 miRNAs estavam com expressão desregulada estatisticamente significativa ( $p \leq 0,01$  and  $FDR \leq 0,1$ ), incluindo a expressão diminuída do miR-486 e a expressão aumentada dos miRNAs let-7a-2, let-7g, miR-15b, miR-181b-1, miR-181b-2, miR-23b, miR-26a-1, miR-26a-2, miR-26b e miR-93 em adenocarcinoma comparado com tecidos pulmonares normais. Adicionalmente, a análise nas amostras pareadas mostrou que 22 miRNAs apresentaram expressão diferencial estatisticamente significativa ( $p \leq 0,01$ ;  $FDR < 0,1$ ); 8 miRNAs (miR-486, miR-1247, miR-218-1, miR-181a-1, miR-181a-2, miR-328, miR-574 e miR-886) estavam com expressão diminuída e 14 miRNAs (let-7a-3, miR-146b, miR-26a-1, miR-200b, miR-191, miR-181c, miR-10a, miR-155, miR-99a, miR-30e, miR-21, miR-425, miR-29a e miR-421) estavam com expressão aumentada no tumor comparado com o tecido normal do mesmo paciente. Interessantemente, a expressão desregulada do miR-486 e do miR-26a-1 foi detectada em ambas as análises (amostras pareadas e não-pareadas) considerando o critério estatístico de  $p \leq 0,01$  e  $FDR \leq 0,1$ .

### Validação dos dados de expressão de miRNAs

A validação utilizando TaqMan RT-QPCR foi realizada para todos os miRNAs com expressão significativamente desregulada (**Tabela 3**) em um grupo independente de 22 amostras de adenocarcinoma e 12 tecidos de pulmonares normais; as amostras normais foram utilizadas como controle e combinadas em 3 *pools* de 4 amostras cada. Quatorze miRNAs mostraram níveis concordantes de expressão nos dados de miRNA-Seq e Taqman PCR (**Tabela 4**). Desses 14 miRNAs, 5 estavam com expressão diminuída (miR-486, miR-181a-1, miR-181a-2, miR-218-1 e miR-886) e 9 estavam com expressão aumentada (let-7a-2, let-7a-3, let-7g, miR-15b, miR-26b, miR-200b, miR-155, miR-21, miR-425). 6/14 miRNAs; let-7a-2, let-7a-3, miR-15b, miR-200b, miR-155 e miR-21 estavam com expressão diferencial estatisticamente significativa ( $p \leq 0,01$  e  $FDR \leq 0,1$ ) no tumor comparado com tecidos pulmonares normais em ambas as análises (miRNA-Seq e TaqMan PCR) (**Tabela 4**).

### **Variações em nucleotídeos únicos**

Análises dos dados de sequenciamento de miRNAs mostraram variações em nucleotídeos únicos (SNVs) nas sequências de miRNAs maduros. SNVs identificadas em comparação com o banco de dados do Consórcio de Referência do Genoma GRCh37/hg19. Foram identificadas 30 SNVs nas amostras de adenocarcinoma e 19 SNVs em tecido pulmonar histologicamente normal. As SNVs identificadas em tecidos tumorais e normais estão descritas nas Tabelas 5 e 6, respectivamente. Três miRNAs (miR-24-1, miR-3074 and miR-4422) apresentaram SNVs (A/T, A/T and C/T) exclusivamente nas amostras tumorais, sendo detectado em 7/17 (41%), 7/17 (41%) e 3/17 (17%) tumores, respectivamente. A alta frequência de SNVs (T/A e G/T) foram detectadas em miRNAs; o miR-125b-2 em 11/17 (65%) tumores e o miR-150 em 12/17 (70%) tumores; sendo este último significativo no tumor comparado com tecido normal ( $p=0,0233$ , teste exato de Fisher). As SNVs observadas no miR-125b-2 e no miR-150 estavam presentes em ambas as amostras de tumor e normal em um caso (52T, 52N); sugerindo que essas SNVs detectadas no tumor e nas amostras normais podem representar um polimorfismo herdado em vez de mutações. Entretanto, as SNVs detectadas nas amostras tumorais e normais precisam ser submetidas à validações adicionais por meio de análises de sequenciamento de DNA a fim de para verificar se as sequencias de DNA genômico que codificam miRNAs apresentam nucleotídeos alterados. SNVs nos miRNAs miR-24-1 and miR-3074 foram associadas com estadiamentos iniciais do tumor (I/II) ( $p=0,0295$ ).

### **DISCUSSÃO**

Alterações na expressão de miRNAs podem ser resultado de alterações genômicas, tais como amplificações, deleções ou alterações na transcrição de miRNAs e na maquinaria de processamento de miRNAs; variações em nucleotídeos únicos ou mutações podem resultar em alterações nos mRNAs-alvo dos miRNAs [31]. Portanto, a identificação de vias moduladas por miRNAs pode levar à melhor caracterização de alvos moleculares clinicamente aplicáveis, ultimamente resultando na melhora terapêutica do câncer [54].

Utilizando critérios rigorosos nas análises de miRNA-Seq, identificamos miRNAs com expressão diminuída estatisticamente significativa, incluindo miR-486, miR-181a-1, miR-181a-2, miR-218-1, miR-886 e com expressão aumentada: let-7a-2, let-7a-3, let-7g, miR-15b, miR-26b, miR-200b, miR-155, miR-21, miR-425 em adenocarcinoma pulmonar comparado com tecidos pulmonares histologicamente normais. Tais miRNAs regulam, simultaneamente, um grande número de genes com funções celulares importantes. De fato,



miRNAs desempenham papéis importantes modulando a expressão de genes e proteínas durante o desenvolvimento e progressão da doença [20].

Similar aos nossos achados, a família do miRNA let-7 (let-7a, b, c, d, e, f, g, i) foi relatada com expressão diferencial estatisticamente significativa entre os grupos histológicos de NSCLC, sendo altamente expressa em AD pulmonar [55]. Outros estudos indicam a diminuição da expressão de alguns membros da família let-7 em linhagens celulares de NSCLC associada a baixa sobrevida [56, 57]. A expressão diminuída, principalmente do let-7a-2 and let-7f-1 foi relatada correlacionada com a menor sobrevida de pacientes com adenocarcinoma. Entretanto, um subgrupo de adenocarcinomas apresentava expressão aumentada de alguns membros da família let-7, como o let-7a-1, let-7a-2 e let-7a-3 [56].

miRNAs também têm sido explorados como marcadores circulantes devido à estabilidade e fácil detecção dessas moléculas na circulação sanguínea [58]. Nosso estudo identificou a expressão desregulada de alguns miRNAs, tal como o miR-15b, o qual já foi detectados com expressão anormal no sangue de pacientes com NSCLC. Um estudo de fase I/II realizou um *screening* em amostras de soro de pacientes com NSCLC comparado com indivíduos saudáveis e identificou a expressão diferencial do miR-15b/27b em pacientes com NSCLC comparado com indivíduos saudáveis [59]. A expressão aumentada de outra assinatura de miRNAs; miR-15b/16-2 induziu a expressão de genes envolvidos nos mecanismos de sinalização da proteína ATM e RAD3, ciclo celular, reparo ao DNA e apoptose; esse estudo demonstrou que o miR-15b tem como alvo direto o *PPM1D*, também denominado *WIP1* (*wild-type p53-induced phosphatase 1*), reduzindo sua expressão gênica e protéica em células pulmonares. Os autores sugeriram que a indução de resposta a danos ao DNA pelo miR-15b deve ser parcialmente dependente da supressão de *PPM1D* e pode representar um alvo terapêutico em doenças pulmonares induzidas pela radiação [60]. *PPM1D* codifica uma serina/treonina fosfatase que desempenha um papel na desfosforilação de várias proteínas de resposta a danos no DNA, tais como a ATM, ATR, p38MAPK, CHK1 e CHK2 [61].

A expressão diminuída do miR-486 e a expressão aumentada do miR-21, miR-200b e miR-375 foi observada em amostras de tecido e de escarro de pacientes com estadio I de AD pulmonar; nesse estudo, a combinação de expressão desregulada dos 4 miRNAs permitiu distinguir pacientes com AD e SCC dos indivíduos sem doença, sendo que houve maior sensibilidade e especificidade da expressão dessa assinatura de 4 miRNAs em AD, sugerindo que esse grupo de miRNAs pode ser útil para a detecção precoce de AD pulmonar [62]. Concordante com os dados da literatura, a expressão aumentada do miR-21 e do miR-200b foi estatisticamente

significativa e considerada validada em nossas amostras de adenocarcinoma quando comparado com tecido pulmonar histologicamente normal.

O efeito de sensibilidade ou resistência a tratamentos como a quimioterapia também tem sido associado a perfis de expressão desregulados de miRNAs. A expressão aumentada do miR-21 e a expressão diminuída da proteína PTEN, foi correlacionada com baixa sobrevida e baixa sensibilidade ao inibidor TKI em pacientes com NSCLC tratados com Gefitinibe ou Erlotinibe. Da mesma forma, a expressão aumentada do miR-21 e a inibição do PTEN foi observada em linhagens resistentes ao Gefitinibe. Ensaios *in vitro* usando miR-21 mimics mostraram uma redução da sensibilidade ao Gefitinibe em linhagens celulares sensíveis ao TKI pela inibição do PTEN e ativação de AKT/ERK, enquanto que a inibição do miR-21 resultou em aumento da expressão do PTEN e inativação da via AKT/ERK, restaurando a sensibilidade ao tratamento [63]. A expressão aumentada de ambos miR-21 e miR-200b em nossas amostras de adenocarcinoma pode explicar, pelo menos em parte, a baixa taxa de resposta a quimioterapia adjuvante de nossos pacientes.

Os miRNAs também regulam vias importantes no processo metastático, tal como a via de sinalização TGF- $\beta$ , a qual ativa múltiplos fatores de transcrição indutores do programa EMT (transição epitélio mesênquima). A família do miR-200 apresenta importante papel no processo de EMT, por meio da regulação de repressores de transcrição, tais como *CDH1* (E-caderina), *ZEB1* e *ZEB2* (*Zinc finger E-box binding homeobox*) [64]. Ensaios *in vivo* e *in vitro* mostraram que o miR-200b foi capaz de inibir a expressão de *ZEB2* em nível de mRNA e proteína, e de *CDH1* em câncer de mama. *ZEB2* é um repressor transcricional e participa da regulação da via de sinalização do TGF- $\beta$  e do processo de EMT pela regulação de *CDH1* [65]. Entretanto, outros estudos mostraram que, apesar da modulação de *ZEB1/CDH1* pelo miR-200b, a restauração de *ZEB1* em linhagem de câncer de mama positiva para a expressão do miR-200b não foi capaz de alterar o potencial metastático, sugerindo uma possível regulação de outros genes-alvo que contribuem para o processo de metástase, como por exemplo genes associados à actina [66].

Outra via molecular relevante na progressão tumoral, a via de sinalização Wnt/ $\beta$ -catenina, foi associada à regulação do miR-155. A sinalização desregulada da via Wnt/ $\beta$ -catenina promove o crescimento celular e alterações na expressão de Wnt/ $\beta$ -catenina foram associadas ao pior prognóstico de pacientes com NSCLC [67, 68]. O miR-155 apresenta importante papel na biologia do câncer, sendo que validamos o aumento da expressão desse miRNA em AD. Outro estudo também detectou expressão aumentada do miR-155 em amostras de tumor primário de pacientes com AD [57] e em soro de pacientes com estadio avançado (IV) de

NSCLC [69], em ambos os estudos essa alteração foi correlacionada com baixa sobrevida dos pacientes [57, 69]. A expressão aumentada do miR-155 também foi identificada em lipossarcoma; ensaios *in vitro* e *in vivo* mostraram que o CK1 $\alpha$  (caseína quinase 1 $\alpha$ ), um regulador chave da via Wnt/ $\beta$ -catenina, é um alvo direto do miR-155, induzindo a sinalização da  $\beta$ -catenina e a expressão da ciclina D1, resultando na proliferação celular e progressão tumoral [70]. Interessantemente, o miR-155 tem sido estudado como alvo terapêutico; ensaios *in vitro* e *in vivo* mostraram a entrega eficiente do anti-miR-155 utilizando nanopartículas em linhagem celular de carcinoma hepatocelular [71].

Adicionalmente, identificamos Variações em Nucleotídeos Únicos (SNVs) na família do miR-let-7 (let-7a-1, let-7a-2 e let-7c) em adenocarcinoma. Variações nas sequências de miRNAs podem mapear em localizações genômicas distintas, entretanto suas sequências são altamente homólogas. Apesar das funções das variantes não terem sido exploradas, elas podem ter papéis na tumorigênese; tais variações foram detectadas nos miRNAs let-7f [72] e miR-196a-2 [73] em NSCLC; tais variações podem levar a alterações nas propriedades de ligação em miRNA:mRNA ou alterações em mRNAs-alvo [31].

Os miRNAs aqui identificados, especialmente let-7a-2, let-7a-3, miR-15b, miR-21, miR-155 e miR-200b regulam a expressão de vários genes-alvo interconectados em vias moleculares importantes no desenvolvimento, progressão tumoral e metástase. Estudos funcionais são necessários para a caracterização desse grupo seletivo de miRNAs, os quais podem ter um papel no desenvolvimento de novas estratégias terapêuticas em adenocarcinoma pulmonar.

## REFERÊNCIAS

1. **Globocan 2012 Stats - Global Cancer Rates Continue to Soar**
2. <http://www2.inca.gov.br/>
3. Jordan RC, Macabeo-Ong M, Shiboski CH, Dekker N, Ginzinger DG, Wong DT, Schmidt BL: **Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer.** *Clin Cancer Res* 2004, **10**(19):6460-6465.
4. Hecht SS: **Tobacco smoke carcinogens and lung cancer.** *J Natl Cancer Inst* 1999, **91**(14):1194-1210.
5. Nattey C, Urban M, Kielkowski D: **0342 Lung cancer risk attributable to occupation: in a case control study in black South Africans, 2001-2008.** *Occup Environ Med* 2014, **71 Suppl 1**:A110-111.
6. Schmid K, Kuwert T, Drexler H: **Radon in indoor spaces: an underestimated risk factor for lung cancer in environmental medicine.** *Dtsch Arztebl Int* 2010, **107**(11):181-186.
7. Herbst RS, Heymach JV, Lippman SM: **Lung cancer.** *N Engl J Med* 2008, **359**(13):1367-1380.
8. Dienstmann R PH, Zukin M: **Tratamento atual do cancer de pulmao nao-pequenas celulas avancado.** *Pratica Hospitalar* 2007; 51: 93-8.; 2007.
9. Wahbah M, Boroumand N, Castro C, El-Zeky F, Eltorkey M: **Changing trends in the distribution of the histologic types of lung cancer: a review of 4,439 cases.** *Ann Diagn Pathol* 2007, **11**(2):89-96.
10. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG *et al*: **Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.** *N Engl J Med* 2004, **350**(21):2129-2139.
11. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ *et al*: **EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.** *Science* 2004, **304**(5676):1497-1500.
12. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L *et al*: **EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib.** *Proc Natl Acad Sci U S A* 2004, **101**(36):13306-13311.
13. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR: **The biology of epidermal growth factor receptor in lung cancer.** *Clin Cancer Res* 2004, **10**(12 Pt 2):4227s-4232s.
14. Sharma SV, Bell DW, Settleman J, Haber DA: **Epidermal growth factor receptor mutations in lung cancer.** *Nat Rev Cancer* 2007, **7**(3):169-181.
15. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Jänne PA, Costa DB *et al*: **Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer.** *N Engl J Med* 2010, **363**(18):1693-1703.
16. Camidge DR, Bang YJ, Kwak EL, Iafrate AJ, Varella-Garcia M, Fox SB, Riely GJ, Solomon B, Ou SH, Kim DW *et al*: **Activity and safety of crizotinib in patients with ALK-positive non-small-cell lung cancer: updated results from a phase 1 study.** *Lancet Oncol* 2012, **13**(10):1011-1019.
17. **NCCN Clinical Practices Guidelines in Oncology (NCCN Guidelines). Non-Small Cell Lung Cancer, version 4.2014**
18. Pérez-Soler R: **Individualized therapy in non-small-cell lung cancer: future versus current clinical practice.** *Oncogene* 2009, **28 Suppl 1**:S38-45.

19. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**(2):281-297.
20. Iorio MV, Croce CM: **MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review.** *EMBO Mol Med* 2012, **4**(3):143-159.
21. Harfe BD: **MicroRNAs in vertebrate development.** *Curr Opin Genet Dev* 2005, **15**(4):410-415.
22. Cummins JM, Velculescu VE: **Implications of micro-RNA profiling for cancer diagnosis.** *Oncogene* 2006, **25**(46):6220-6227.
23. Rinaldi A, Poretti G, Kwee I, Zucca E, Catapano CV, Tibiletti MG, Bertoni F: **Concomitant MYC and microRNA cluster miR-17-92 (C13orf25) amplification in human mantle cell lymphoma.** *Leuk Lymphoma* 2007, **48**(2):410-412.
24. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH: **Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells.** *J Biol Chem* 2008, **283**(2):1026-1033.
25. Reis PP, Tomenson M, Cervigne NK, Machado J, Jurisica I, Pintilie M, Sukhai MA, Perez-Ordenez B, Grénman R, Gilbert RW *et al*: **Programmed cell death 4 loss increases tumor cell invasion and is regulated by miR-21 in oral squamous cell carcinoma.** *Mol Cancer* 2010, **9**:238.
26. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**(2):126-139.
27. Ha M, Kim VN: **Regulation of microRNA biogenesis.** *Nat Rev Mol Cell Biol* 2014, **15**(8):509-524.
28. Kent OA, Mendell JT: **A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes.** *Oncogene* 2006, **25**(46):6188-6196.
29. Shenouda SK, Alahari SK: **MicroRNA function in cancer: oncogene or a tumor suppressor?** *Cancer Metastasis Rev* 2009, **28**(3-4):369-378.
30. Iorio MV, Croce CM: **MicroRNAs in cancer: small molecules with a huge impact.** *J Clin Oncol* 2009, **27**(34):5848-5856.
31. Kasinski AL, Slack FJ: **Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy.** *Nat Rev Cancer* 2011, **11**(12):849-864.
32. Joshi P, Middleton J, Jeon YJ, Garofalo M: **MicroRNAs in lung cancer.** *World J Methodol* 2014, **4**(2):59-72.
33. Liu J, Lu KH, Liu ZL, Sun M, De W, Wang ZX: **MicroRNA-100 is a potential molecular marker of non-small cell lung cancer and functions as a tumor suppressor by targeting polo-like kinase 1.** *BMC Cancer* 2012, **12**:519.
34. Vucic EA, Thu KL, Pikor LA, Enfield KS, Yee J, English JC, MacAulay CE, Lam S, Jurisica I, Lam WL: **Smoking status impacts microRNA mediated prognosis and lung adenocarcinoma biology.** *BMC Cancer* 2014, **14**:778.
35. Garofalo M, Jeon YJ, Nuovo GJ, Middleton J, Secchiero P, Joshi P, Alder H, Nazaryan N, Di Leva G, Romano G *et al*: **MiR-34a/c-Dependent PDGFR- $\alpha/\beta$  Downregulation Inhibits Tumorigenesis and Enhances TRAIL-Induced Apoptosis in Lung Cancer.** *PLoS One* 2013, **8**(6):e67581.
36. Ling H, Fabbri M, Calin GA: **MicroRNAs and other non-coding RNAs as targets for anticancer drug development.** *Nat Rev Drug Discov* 2013, **12**(11):847-865.
37. Bouchie A: **First microRNA mimic enters clinic.** *Nat Biotechnol* 2013, **31**(7):577.
38. Chomeczynski P, and Sacchi, N.: **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** . In., vol. 162: *Anal. Biochem*; 1987: 156-159.

39. Reis PP, Waldron L, Perez-Ordonez B, Pintilie M, Galloni NN, Xuan Y, Cervigne NK, Warner GC, Makitie AA, Simpson C *et al*: **A gene signature in histologically normal surgical margins is predictive of oral carcinoma recurrence.** *BMC Cancer* 2011, **11**:437.
40. Goswami RS, Atenafu EG, Xuan Y, Waldron L, Reis PP, Sun T, Datti A, Xu W, Kuruvilla J, Good DJ *et al*: **MicroRNA signature obtained from the comparison of aggressive with indolent non-Hodgkin lymphomas: potential prognostic value in mantle-cell lymphoma.** *J Clin Oncol* 2013, **31**(23):2903-2911.
41. Andrews S: **FastQC A Quality Control tool for High Throughput Sequence Data.** In. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; 2010.
42. MARTIN M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** . In., vol. 17. **EMBnet.journal:** <http://journal.embnet.org/index.php/embnetjournal/article/view/200>; 2011: 10-12.
43. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**(3):R25.
44. Simon Anders PTP, Wolfgang Huber: **HTSeq - A Python framework to work with high-throughput sequencing data.** In. bioRxiv preprint; 2014.
45. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data.** *Genome Biol* 2010, **11**(3):R25.
46. Robinson MD, Smyth GK: **Small-sample estimation of negative binomial dispersion, with applications to SAGE data.** *Biostatistics* 2008, **9**(2):321-332.
47. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139-140.
48. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J *et al*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**(10):R80.
49. Li H: **A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.** *Bioinformatics* 2011, **27**(21):2987-2993.
50. Li H: **Improving SNP discovery by base alignment quality.** *Bioinformatics* 2011, **27**(8):1157-1158.
51. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al*: **The variant call format and VCFtools.** *Bioinformatics* 2011, **27**(15):2156-2158.
52. Wang K, Li M, Hakonarson H: **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.** *Nucleic Acids Res* 2010, **38**(16):e164.
53. Goswami RS, Waldron L, Machado J, Cervigne NK, Xu W, Reis PP, Bailey DJ, Jurisica I, Crump MR, Kamel-Reid S: **Optimization and analysis of a quantitative real-time PCR-based technique to determine microRNA expression in formalin-fixed paraffin-embedded samples.** *BMC Biotechnol* 2010, **10**:47.
54. Di Leva G, Garofalo M, Croce CM: **MicroRNAs in cancer.** *Annu Rev Pathol* 2014, **9**:287-314.
55. Landi MT, Zhao Y, Rotunno M, Koshiol J, Liu H, Bergen AW, Rubagotti M, Goldstein AM, Linnoila I, Marincola FM *et al*: **MicroRNA expression differentiates histology and predicts survival of lung cancer.** *Clin Cancer Res* 2010, **16**(2):430-441.

56. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y *et al*: **Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival.** *Cancer Res* 2004, **64**(11):3753-3756.
57. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.** *Cancer Cell* 2006, **9**(3):189-198.
58. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A *et al*: **Circulating microRNAs as stable blood-based markers for cancer detection.** *Proc Natl Acad Sci U S A* 2008, **105**(30):10513-10518.
59. Hennessey PT, Sanford T, Choudhary A, Mydlarz WW, Brown D, Adai AT, Ochs MF, Ahrendt SA, Mambo E, Califano JA: **Serum microRNA biomarkers for detection of non-small cell lung cancer.** *PLoS One* 2012, **7**(2):e32307.
60. Rahman M, Lovat F, Romano G, Calore F, Acunzo M, Bell EH, Nana-Sinkam P: **miR-15b/16-2 regulates factors that promote p53 phosphorylation and augments the DNA damage response following radiation in the lung.** *J Biol Chem* 2014, **289**(38):26406-26416.
61. Lowe J, Cha H, Lee MO, Mazur SJ, Appella E, Fornace AJ: **Regulation of the Wip1 phosphatase and its effects on the stress response.** *Front Biosci (Landmark Ed)* 2012, **17**:1480-1498.
62. Yu L, Todd NW, Xing L, Xie Y, Zhang H, Liu Z, Fang H, Zhang J, Katz RL, Jiang F: **Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers.** *Int J Cancer* 2010, **127**(12):2870-2878.
63. Shen H, Zhu F, Liu J, Xu T, Pei D, Wang R, Qian Y, Li Q, Wang L, Shi Z *et al*: **Alteration in Mir-21/PTEN Expression Modulates Gefitinib Resistance in Non-Small Cell Lung Cancer.** *PLoS One* 2014, **9**(7):e103305.
64. Korpala M, Kang Y: **The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis.** *RNA Biol* 2008, **5**(3):115-119.
65. Christoffersen NR, Silahtaroglu A, Orom UA, Kauppinen S, Lund AH: **miR-200b mediates post-transcriptional repression of ZFX1B.** *RNA* 2007, **13**(8):1172-1178.
66. Li X, Roslan S, Johnstone CN, Wright JA, Bracken CP, Anderson M, Bert AG, Selth LA, Anderson RL, Goodall GJ *et al*: **MiR-200 can repress breast cancer metastasis through ZEB1-independent but moesin-dependent pathways.** *Oncogene* 2014, **33**(31):4077-4088.
67. Uematsu K, He B, You L, Xu Z, McCormick F, Jablons DM: **Activation of the Wnt pathway in non small cell lung cancer: evidence of dishevelled overexpression.** *Oncogene* 2003, **22**(46):7218-7221.
68. MacDonald BT, Tamai K, He X: **Wnt/beta-catenin signaling: components, mechanisms, and diseases.** *Dev Cell* 2009, **17**(1):9-26.
69. Cui EH, Li HJ, Hua F, Wang B, Mao W, Feng XR, Li JY, Wang X: **Serum microRNA 125b as a diagnostic or prognostic biomarker for advanced NSCLC patients receiving cisplatin-based chemotherapy.** *Acta Pharmacol Sin* 2013, **34**(2):309-313.
70. Zhang P, Bill K, Liu J, Young E, Peng T, Bolshakov S, Hoffman A, Song Y, Demicco EG, Terrada DL *et al*: **MiR-155 is a liposarcoma oncogene that targets casein kinase-1 $\alpha$  and enhances  $\beta$ -catenin signaling.** *Cancer Res* 2012, **72**(7):1751-1762.

71. Zhang M, Zhou X, Wang B, Yung BC, Lee LJ, Ghoshal K, Lee RJ: **Lactosylated gramicidin-based lipid nanoparticles (Lac-GLN) for targeted delivery of anti-miR-155 to hepatocellular carcinoma.** *J Control Release* 2013, **168**(3):251-261.
72. Silva J, García V, Zaballos Á, Provencio M, Lombardía L, Almonacid L, García JM, Domínguez G, Peña C, Diaz R *et al*: **Vesicle-related microRNAs in plasma of nonsmall cell lung cancer patients and correlation with survival.** *Eur Respir J* 2011, **37**(3):617-623.
73. Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, Zeng Y, Miao R, Jin G, Ma H *et al*: **Genetic variants of miRNA sequences and non-small cell lung cancer survival.** *J Clin Invest* 2008, **118**(7):2600-2608.



**Tabela 1.** Dados clínicos e histopatológicos dos pacientes com adenocarcinoma pulmonar do grupo teste

	<b>Número Total</b>	<b>%</b>
<b>Idade (anos)</b>		
Mediana (intervalo)	66,8(43-83 anos)	
Média	65	
<b>Sexo</b>		
Masculino	8	47
Feminino	9	53
<b>Tabagismo</b>		
Sim	10	59
Não	7	41
<b>Etilismo</b>		
Sim	1	6
Não	16	94
<b>Tipo Histológico</b>		
Adenocarcinoma	17	100
<b>Grau Histológico do tumor</b>		
Bem diferenciado	5	29
Moderadamente diferenciado	10	59
Pobrememente diferenciado	2	12
<b>Tamanho do Tumor</b>		
T1-T2	15	88
T3-T4	2	12
<b>Metástase Linfonodal</b>		
Negativo (N0)	12	71
Positivo (N1, N2)	5	29
<b>Metástase à distancia</b>		
Sim	5	29
Não	12	71
<b>Estadiamento do tumor</b>		
Ia; Ib; IIa; IIb	14	82
IIIa; IIIb	3	18
<b>Sobrevida</b>		
Vivo com doença	12	71
Óbito	2	12

**Tabela 2.** Dados clínicos e histopatológicos dos pacientes com adenocarcinoma pulmonar do grupo de validação.

	<b>Número Total</b>	<b>%</b>
<b>Idade (anos)</b>		
Mediana (intervalo)	58,0(40-84 anos)	
Média	60,5	
<b>Sexo</b>		
Masculino	12	55
Feminino	10	45
<b>Tabagismo</b>		
Sim	15	68
Não	7	32
<b>Etilismo</b>		
Sim	7	32
Não	15	68
<b>Tipo Histológico</b>		
Adenocarcinoma	22	100
<b>Grau Histológico do tumor</b>		
Bem diferenciado	3	14
Moderadamente diferenciado	12	55
Pobrememente diferenciado	7	32
<b>Tamanho do Tumor</b>		
T1-T2	13	59
T3-T4	9	41
<b>Metástase Linfonodal</b>		
Negativo (N0)	12	55
Positivo (N1, N2)	10	45
<b>Metástase à distancia</b>		
Sim	2	9
Não	20	91
<b>Estadiamento do tumor</b>		
Ia; Ib; IIa; IIb	13	59
IIIa; IIIb	9	41
<b>Sobrevida</b>		
Vivo com doença	5	23
Vivo sem evidencia de doença	9	41
Óbito pela doença	3	14
Óbito por outra causa	4	18

**Tabela 3.** miRNAs diferencialmente expressos em adenocarcinoma comparado com tecido normal.

<b>miRNA</b>	<b>LogFC</b>	<b>P-value</b>	<b>FDR</b>
<b>Amostras não-pareadas</b>			
miR-486	-1,635	0,0004	0,0689
let-7g	0,8879	0,0015	0,0689
miR-15b	0,9831	0,0004	0,0689
miR-181b-2	0,9887	0,0009	0,0689
miR-23b	0,9479	0,0014	0,0689
miR-26a-1	0,9756	0,0014	0,0689
miR-26a-2	0,9595	0,0013	0,0689
miR-26b	1,0565	0,0009	0,0689
miR-93	1,203	0,0005	0,0689
let-7a-2	0,8844	0,0025	0,1052
miR-181b-1	0,9221	0,0027	0,1052
<b>Amostras Pareadas</b>			
miR-486	-2,5430	8,8276.E-12	1,805.E-09
miR-1247	-2,6582	3,2911.E-08	3,3651.E-06
miR-218-1	-1,6175	0,0001	0,0043
miR-181a-2	-1,1145	0,0002	0,0079
miR-328	-1,4708	0,0004	0,0108
miR-181a-1	-1,1444	0,0006	0,0147
miR-574	-1,1719	0,0036	0,0503
miR-886	-1,3187	0,0046	0,0575
let-7a-3	0,4856	3,2638.E-264	5,1692.E-262
miR-146b	0,7012	0,0005	0,0147
miR-26a-1	0,7382	0,0043	0,0551
miR-200b	1,0898	0,0072	0,0864
miR-191	1,0935	0,0018	0,0307
miR-181c	1,1230	0,0041	0,0551
miR-10a	1,2858	0,0009	0,0208
miR-155	1,3089	0,0029	0,0418
miR-99a	1,4447	0,0028	0,0418
miR-30e	1,4926	0,00002	0,0011
miR-21	1,5219	0,0028	0,0418
miR-425	1,6206	7,6888.E-06	0,0005
miR-29a	1,6488	0,0001	0,0043
miR-421	1,7648	0,0085	0,0993

LogFC = log de fold change; FDR = false discovery rate. FDR $\leq$ 0,1, P $\leq$ 0,01.

**Tabela 4.** Lista de miRNAs com concordância entre a expressão identificada nas análises de sequenciamento de miRNA e TaqMan PCR.

miRNA	miRNA-Seq		miRNA	TLDA	
	logFC	P-value		logFC	P-value
let-7a-2	0,8844	0,0025	let-7a	2,3601	<b>0,000*</b>
let-7a-3	0,4856	3,2638.E-264	let-7a	2,3601	<b>0,000*</b>
miR-15b	0,9831	0,0004	miR-15b	2,4022	<b>0,000*</b>
miR-200b	1,0898	0,0072	miR-200b	1,9404	<b>0,000*</b>
miR-21	1,5219	0,0028	miR-21	2,6311	<b>0,001*</b>
miR-155	1,3089	0,0029	miR-155	1,3294	<b>0,010*</b>
miR-486	-1,635	0,0004	miR-486	-3,9434	0,084
miR-181a-1	-1,1444	0,0006	miR-181a	-0,8625	0,140
miR-181a-2	-1,1145	0,0002	miR-181a	-0,8625	0,140
let-7g	0,8879	0,0015	let-7g	0,3918	0,143
miR-26b	1,0565	0,0009	miR-26b	0,5597	0,224
miR-218-1	-1,6175	0,0001	miR-218	-1,4699	0,386
miR-886	-1,3187	0,0046	miR-886-3p	-1,1746	0,482
miR-425	1,6206	7,6888.E-06	miR-425-5p	0,1190	0,906

Grupo teste: valor de  $P \leq 0,01$  e  $FDR \leq 0,1$ , determinado pelo programa edgeR;

Grupo de validação:  $P \leq 0,01$  determinado pelo programa *Expression Suite*.

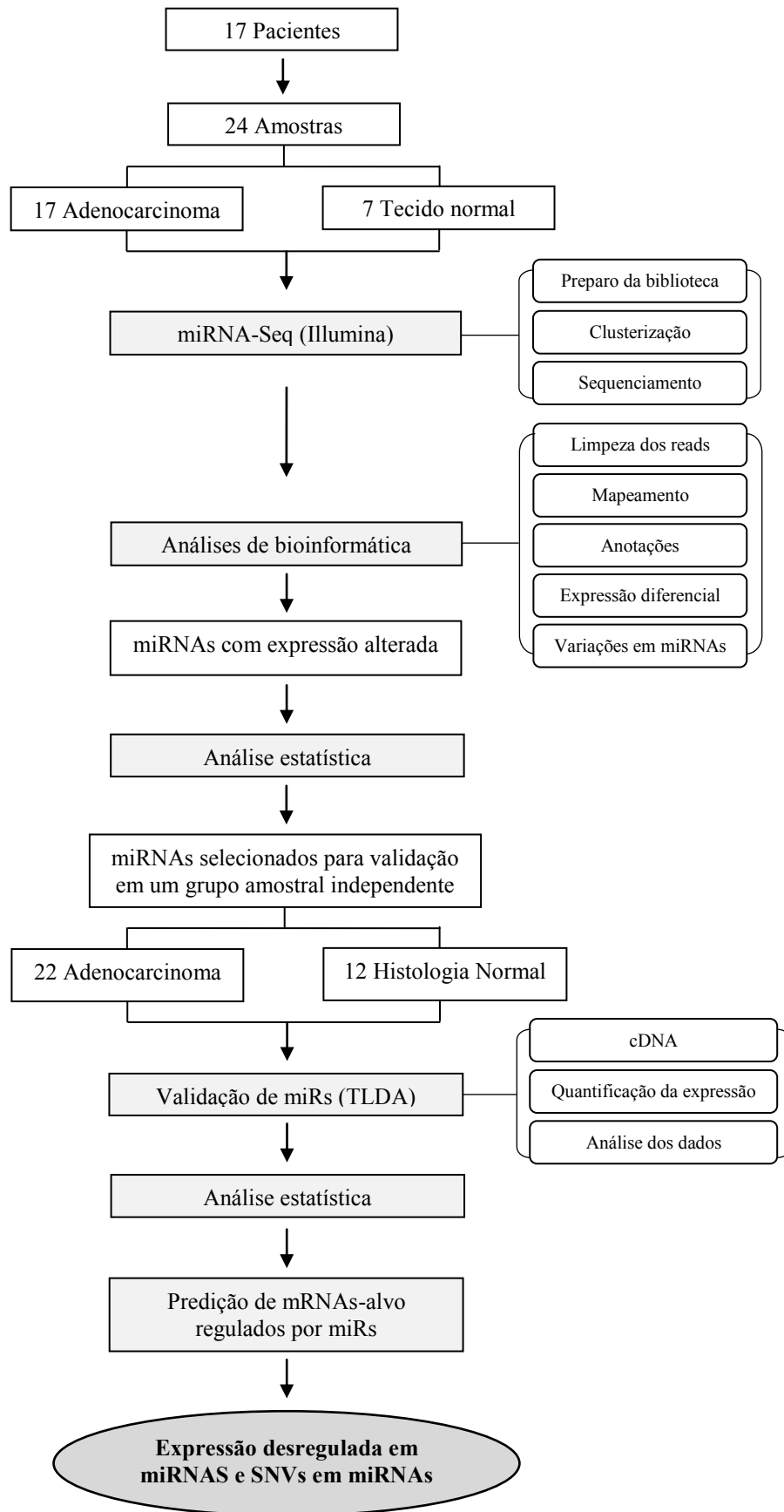
\*miRNAs desregulados estatisticamente significativos identificados em ambos os grupos teste e de validação através de sequenciamento e TaqMan PCR

**Tabela 5.** SNVs identificadas em miRNAs em adenocarcinoma.

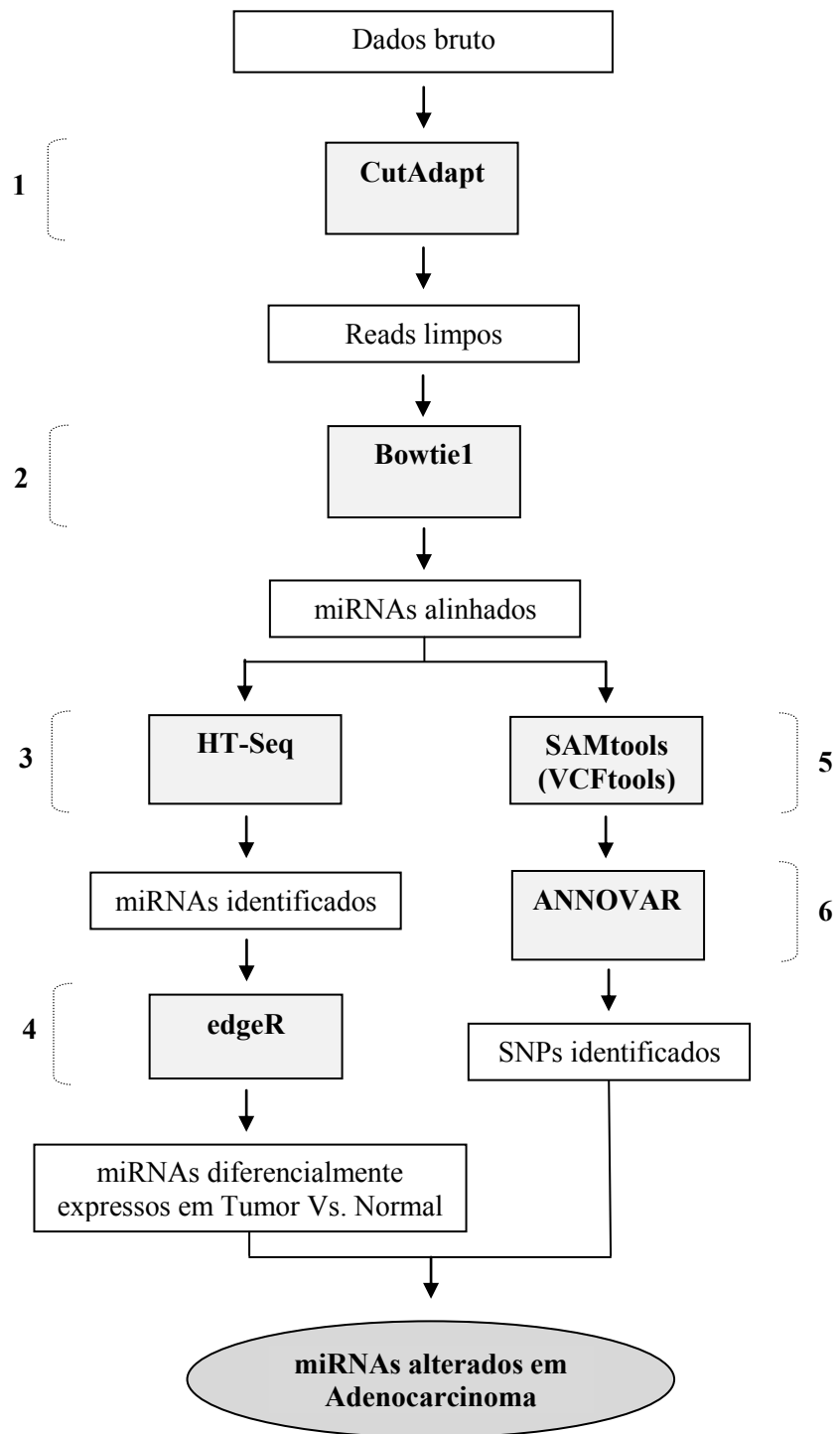
<b>Chr</b>	<b>Início</b>	<b>Fim</b>	<b>Ref</b>	<b>Alt</b>	<b>Localização</b>	<b>Gene</b>
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	Intergenic	mir-4490
Chr21	9827315	9827315	T	C	Intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	let-7c
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr1	16861827	16861827	G	A	Intergenic	mir-3675
Chr11	122017274	122017274	T	C	ncRNA_exonic	let-7a-2
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-24-1
Chr9	97848366	97848366	A	T	ncRNA_exonic	mir-3074
Chr12	54731084	54731084	C	A	ncRNA_exonic	mir-148b
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200a
Chr17	46657243	46657243	A	G	ncRNA_exonic	mir-10a
Chr1	56633037	56633037	C	T	Intergenic	mir-4422
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr1	186105018	186105018	G	A	ncRNA_exonic	mir-548f-1
Chr1	186105020	186105020	G	A	ncRNA_exonic	mir-548f-1
Chr2	227230043	227230043	A	G	Intergenic	mir-5702
Chr8	143215944	143215944	G	C	Intergenic	mir-4472-1
Chr9	97847569	97847569	C	T	ncRNA_exonic	mir-23b
Chr14	53927750	53927750	T	G	Intergenic	mir-5580
Chr11	122022981	122022981	A	T	ncRNA_exonic	mir-100
Chr21	17911436	17911436	T	A	ncRNA_exonic	mir-99a
Chr3	48261476	48261476	C	A	Intergenic	mir-4443
Chr19	46522264	46522264	T	A	ncRNA_exonic	mir-769

**Tabela 6.** SNVs identificadas em miRNAs em tecido normal.

<b>Chr</b>	<b>Início</b>	<b>Fim</b>	<b>Ref</b>	<b>Alt</b>	<b>Localização</b>	<b>Gene</b>
Chr1	1103312	1103312	C	A	ncRNA_exonic	mir-200 a
Chr4	10917451	10917451	T	G	intergenic	mir-572
Chr5	167987905	167987905	G	T	ncRNA_exonic	mir-103a-1
Chr5	167987908	167987908	T	A	ncRNA_exonic	mir-103a-1
Chr9	97847571	97847571	C	G	ncRNA_exonic	mir-23b
Chr11	91818572	91818572	T	C	intergenic	mir-4490
Chr16	69967068	69967068	A	T	ncRNA_exonic	mir-140
Chr19	52195893	52195893	G	T	ncRNA_exonic	mir-99b
Chr21	9827315	9827315	T	C	intergenic	mir-3687
Chr21	17912181	17912181	A	G	ncRNA_exonic	let-7c
Chr21	17912184	17912184	G	A	ncRNA_exonic	let-7c
Chr1	172113731	172113731	T	A	ncRNA_exonic	mir-199a-2
Chr11	122017274	122017274	T	C	ncRNA_exonic	let-7a-2
Chr19	50004088	50004088	G	T	ncRNA_exonic	mir-150
Chr19	52195896	52195896	C	A	ncRNA_exonic	mir-99b
Chr21	17962597	17962597	T	A	ncRNA_exonic	mir-125b-2
Chr8	143215944	143215944	G	C	intergenic	mir-4472-1
Chr9	96938260	96938260	A	G	ncRNA_exonic	let-7a-1
Chr9	96938262	96938262	A	G	ncRNA_exonic	let-7a-1



**Figura 1.** Delineamento experimental e amostras



**Figura 2. Passos utilizados nas análises de sequenciamento:** 1.Filtragem; 2.Alinhamento/Mapeamento; 3. Anotações; 4.Análises de expressão diferencial; 5.Filtragem das variantes genômicas; 6.Anotações dos SNPs.