

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

ALYNNE OYA E CHIROMATZO

**Abordagem computacional para identificar
vias metabólicas afetadas por miRNAs**

Ribeirão Preto – SP

2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Catálogo na Publicação
Serviço de Documentação
Faculdade de Medicina de Ribeirão Preto

Chiromatzo, Alynne Oya

Abordagem computacional para identificar vias metabólicas afetadas por miRNAs / Alynne Oya Chiromatzo; orientador: Wilson Araújo da Silva Jr. – Ribeirão Preto -SP, 2010.

90 f. : il.

Tese (Doutorado – Programa de Pós-Graduação em Genética. Área de Concentração: Genética) – Faculdade de Medicina de Ribeirão Preto.

1. MiRNAs, genes alvos, vias metabólicas. 2. Bibliotecas de SAGE.

ALYNNE OYA E CHIROMATZO

Abordagem computacional para identificar vias metabólicas afetadas por miRNA

Tese apresentada ao Programa de Pós-Graduação em Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para a obtenção do título de Doutora em Ciências.

Área de Concentração: Genética
Orientador: Prof. Dr. Wilson Araújo da Silva Júnior

Ribeirão Preto – SP

2010

Dedicatória

Dedico esta Tese primeiramente a Deus.

*Depois aos meus pais, Nanami e José
e meus irmãos Ha e Xan que sempre estão ao meu lado.*

*E ao meu namorado Kita que está todos esses dias comigo me
ajudando a ver o mundo com olhos diferentes.*

Agradecimentos

Ao Prof. Dr. Wilson Araújo da Silva Junior por ter me concedido esta oportunidade, pela paciência e compreensão em todos os momentos.

À Prof. Dra. Silvana Giuliatti que sempre contribuiu para o meu crescimento.

Ao doutorando "irmãozinho" Thiago Kikuchi Oliveira que sempre me dá um força nas horas difíceis.

A todos Laboratório de Genética Molecular e Bioinformática pela colaboração, mas principalmente pela amizade. Cristiane Ayres, Willys Tristão, Nathalia Moreno Curi, Viviane Aline, Carla Kaneto, Julio, Greice Molfetta, Daniel residente, Daniel, Maria Fernanda, Dalila e Adriana. E aos meninos do Bit Israel, Daniel, Marcelo, Mudinho e "Brown".

À Meire Tarla que está presente desde o começo até o fim, ajudando em tudo que é possível.

À Dalvinha e a Gislaíne da Silva Pimentel Pereira que me ajudam a ser uma pessoa melhor.

Ao pessoal voluntário do Caribe que tem o coração maior que eles mesmos.

À minha família que me ama mesmo estando tão longe. E ao meu namorado Kita, Dona Rioko e Seu Atsushi que tem me ajudado muito com o seu carinho.

Aos funcionários do hemocentro, sempre alegres e educados.

A todos que não citei por motivo de esquecimento mas que têm o meu profundo agradecimento por terem contribuído direta ou indiretamente com este projeto.

*“Tudo é possível se crês,
nada é possível se não tens fé!*

*Ele abre caminhos
onde na vida já não há.
Tudo é possível se crês!.”*

Neil Velez

Resumo

MiRNAs são pequenas moléculas de RNAs endógenos não codificantes com aproximadamente 23nt que atuam na regulação da expressão gênica. A sua função é inibir a tradução de genes transcritos através de um mecanismo que viabiliza a ligação do miRNA com o mRNA alvo levando à inibição da tradução ou a degradação do RNA mensageiro. Estudos evidenciam a relação dos miRNAs com diversos processos biológicos como proliferação celular, diferenciação, desenvolvimento e doenças. Uma vez que estão envolvidos na regulação gênica, também alteram as vias metabólicas. Atualmente, as ferramentas computacionais disponíveis para o estudo dos miRNAs são o miRBase, microCosm, o miRGen e o miRNAmip. Elas possuem informações sobre as sequências dos miRNAs, genes alvos e sobre elementos que estão próximos à região dos miRNAs. Embora o avanço até o momento, não existia que relacionasse os miRNAs com as vias metabólicas, para isso foi construída a plataforma miRNAmip que auxilia no estudo da função dos miRNAs por meio da análise dos seus alvos dentro das vias metabólicas. De modo semelhante, também não existia uma abordagem que relacione dados de expressão miRNAs e seus alvos dentro de um mesmo experimento. Para tanto, neste trabalho foi feita uma abordagem utilizando bibliotecas de SAGE (*Serial Analysis of Gene Expression*) que será incorporada no miRNAmip. O miRNAmip encontra-se disponível em <http://lgmb.fmrp.usp.br/mirnapath>.

Abstract

Computational approach for identification of metabolic pathways affected by miRNAs.

MiRNAs are small molecules of endogenous non-coding RNAs with approximately 23nt in length that acts over gene expression regulation. Its function is inhibit the translation of gene transcripts through a mechanism that links the miRNA with its mRNA target leading to a translational repression or degradation. Studies show the relation of RNAs in many biological processes like cell proliferation, differentiation and development of diseases. Since they are involved in gene regulation, they also change the metabolic pathways. Currently, the available computational tools for the study of miRNAs are miRBase, microCosm, miRGen and miRNAmap. They have information about miRNAs sequences, targets and features. Despite the the advances, until now, there is no tool that correlates the miRNAs with metabolic pathways, therefore we developed the miRNAPATH platform that helps in the analysis of miRNAs function through the study of its targets that are into the metabolic pathway. In the same way, there is no approach that put together information of expression of miRNAs and its targets in the same experiment. In this work we develop an approach with SAGE (*Serial Analysis of Gene Expression*) libraries that will be integrated to miRNAPATH. The platform is available at <http://lgmb.fmrp.usp.br/mirnapath>.

Sumário

Lista de Figuras

Lista de Tabelas

| | | |
|----------|--|-------|
| 1 | Introdução | p. 12 |
| 1.1 | miRNAs: da origem à função | p. 12 |
| 1.2 | Perfis de expressão dos miRNAs | p. 17 |
| 1.3 | Bancos de dados biológicos | p. 17 |
| 2 | Objetivos | p. 22 |
| 3 | Material e métodos | p. 23 |
| 3.1 | Modelagem do banco de dados | p. 23 |
| 3.2 | Diagrama de navegação | p. 24 |
| 3.3 | Processamento do miRNApath | p. 25 |
| 3.4 | Bibliotecas de SAGE | p. 28 |
| 3.5 | Processamento das análises | p. 29 |
| 3.6 | Validação dos resultados | p. 29 |
| 4 | Resultados e discussão | p. 30 |
| 4.1 | miRNApath | p. 30 |
| 4.1.1 | <i>Quick Search</i> | p. 30 |
| 4.1.2 | <i>Advanced Analysis</i> | p. 34 |
| 4.2 | Avaliação e identificação de <i>tags</i> de miRNAs | p. 37 |

| | |
|--|-------|
| 5 Conclusões | p. 42 |
| Referências Bibliográficas | p. 44 |
| Apêndice A – Artigo que descreve a plataforma miRNAPath | p. 48 |

Lista de Figuras

- 1 Representação do processo envolvendo a biogênese dos miRNAs. Adaptado de <http://uranus.mtveurope.org/wklim/research.html>. p. 14
- 2 Representação esquemática dos transcritos de pri-miRNAs de humanos. A) Os pre-miRNAs estão em vermelho, regiões não codificantes em verde, introns em azul e regiões codificantes em laranja e PA é ponto alternativo de poli(A); B) estrutura de grampo. Locais de clivagem da Drosha são apontados pelas setas e o ponto de clivagem da Dicer é apontado pelos triângulos. Adaptado de (CULLEN, 2004). p. 15
- 3 Representação simplificada da técnica de SAGE. Figura adaptada de <http://www.scq.ubc.ca/painless-gene-expression-profiling-sage-serial-analysis-of-gene-expression/> p. 18
- 4 Site do miRBase e sua ferramenta *Genomics* p. 19
- 5 miRGen e suas ferramentas: *Genomics*, *Targets* e *Clusters*. p. 20
- 6 miRNAmapp e sua ferramenta para busca de alvos de miRNAs e padrões de expressão. p. 21
- 7 Descrição da estrutura do acesso às informações do KEGG sobre vias metabólicas e seus genes. p. 21
- 8 Representação das relações entre as tabelas do miRNAPath. p. 24
- 9 Figura representando o diagrama do miRNAPath. p. 25
- 10 Esquema do processo de desenvolvimento da plataforma miRNAPath. p. 26
- 11 Estratégia desenvolvida para identificar *tags* de miRNAs. p. 29
- 12 Página principal das buscas com as sessões *Quick Search* e *Advanced Analysis*. p. 31
- 13 Página da busca por genes *Search by gene*. p. 31

| | | |
|----|--|-------|
| 14 | Páginas de resultado da busca. A seleção em verde é o link que foi clicado. A) Página com os genes da lista de entrada. B) Lista de todos os miRNAs que regulam o gene AXIN2. C) Lista de todas as vias metabólicas que são reguladas pelo miRNA hsa-mir-107. A via metabólica destacada em vermelho, é uma regulada pelo miRNA e possui o gene AXIN2. | p. 32 |
| 15 | Página da busca por miRNAs <i>Search by miRNA</i> | p. 33 |
| 16 | Página da busca por vias metabólicas <i>Search by Pathway</i> | p. 33 |
| 17 | Primeiro nível do resultado das buscas. A) Por miRNAs e B) por vias metabólicas. | p. 34 |
| 18 | Primeiro nível do resultado das buscas da sessão <i>Advanced Analysis</i> | p. 35 |
| 19 | Página do resultado da busca, ordenada por genes. A) Lista de genes e a contagem de vias metabólicas e miRNAs. A caixa ao lado da quantidade de miRNAs mostra a listagem com os nomes dos miRNAs. B) Seguindo o link de vias metabólicas (<i>Pathways</i>). Tem-se a informação sobre o gene, seus alvos e as vias onde está envolvido. | p. 35 |
| 20 | Página do resultado da busca, ordenado por miRNAs. | p. 36 |
| 21 | Página do resultado da busca, ordenado por vias metabólicas. | p. 37 |
| 22 | Distribuição dos miRNAs dentro de genes e intergênicos. | p. 37 |
| 23 | MiRNAs expressos no estudo de SAGE e seus genes alvos na via metabólica do glioma. | p. 39 |
| 24 | Expressão de SAGE para o miR-155 e do gene SOS1. | p. 40 |
| 25 | Expressão do miR-21 A) Expressão dos das <i>tags</i> de miRNA em SAGE e B) expressão validada por RT-PCT. | p. 40 |
| 26 | Expressão da tag de miR-21 e dos genes PTEN e PDCD4 nas bibliotecas de SAGE para o estudo de glioblastoma. | p. 41 |

Lista de Tabelas

| | | |
|---|--|-------|
| 1 | miRNAs e suas respectivas funções dentro de patologias. | p. 16 |
| 2 | Lista de organismos e seus respectivos códigos do NCBI. | p. 23 |
| 3 | Lista dos programas utilizados pela plataforma miRNApath. | p. 27 |
| 4 | Bibliotecas de sage e suas respectivas quantidade de <i>tags</i> | p. 28 |
| 5 | miRNAs identificados pela técnica de SAGE. | p. 38 |

1 *Introdução*

Os microRNAs (miRNAs) são uma classe dentro dos RNAs não codificantes com um tamanho de aproximadamente 23pbs que regulam a expressão gênica no nível pós-transcricional. Essas moléculas são altamente conservadas em diversos organismos e estão envolvidas em diferentes processos celulares como a diferenciação celular, proliferação e apoptose (Di Leva; CALIN; CROCE, 2006).

A identificação de novos miRNAs é difícil de ser feita experimentalmente, devido à especificidade espaço-temporal da sua expressão e pela baixa expressão de alguns. Porém as suas características específicas possibilitam que sejam desenvolvidas ferramentas computacionais de predição de novos microRNAs. Atualmente dos 721 miRNAs em humanos existentes, a maioria é predita computacionalmente. Muitos estudos utilizam inicialmente as predições para depois fazer a sua validação experimental (HOU; YING; LI, 2008; BATTUWITA; PALADE, 2009).

Embora a função de alguns miRNAs esteja bem definida, entender o seu papel nas vias metabólicas é um processo bastante complexo, pois envolvem redes com muitos genes. Com o intuito de esclarecer as funções dos miRNAs, pesquisas recentes utilizam ferramentas de predição para selecionar os melhores candidatos a genes alvos para serem avaliados e validados experimentalmente (GREENE *et al.*, 2010). Porém essas ferramentas não são capazes de oferecer uma correlação entre os genes e as vias metabólicas de interesse, fazendo com que este trabalho seja feito de forma manual.

1.1 **miRNAs: da origem à função**

O primeiro miRNA identificado foi o lin-4 através de um estudo realizado por Lee, Wightman e colaboradores (1993) que ao caracterizarem os genes envolvidos no desenvolvimento larval do verme *Caenorhabditis elegans* descobriram que o gene lin-4 não codificava proteína, mas produzia dois transcritos de tamanhos 22pb e 61pb que continham sequen-

cias complementares à região 3' UTR do transcrito gene *lin-14* regulando negativamente a produção de sua proteína.

Ao passo que se estudava o mecanismo envolvendo os miRNAs, descobria-se que eles estavam presentes em outros organismos e em grande número. Pasquinelli *et al.* (2000) detectaram os primeiros homólogos do *lin-4* em vertebrados, anelídeos, artrópodes e moluscos, os quais mostraram uma expressão temporal semelhante a de *C. elegans*, sugerindo que eles pudessem desempenhar um papel ortólogo.

Lagos-Quintana *et al.* (2001) utilizando PCR (*polymerase chain reaction*) reversa descobriu novos miRNAs cujas expressões foram verificadas por *Northern blot*. Eles mostraram que existiam outros miRNAs com tamanhos parecidos com o do *lin-4* (23 pbs) que também estavam presentes em vertebrados e invertebrados com um alto grau de conservação. E sugeriram que a regulação pós-transcricional destes poderia ser mais abrangente do que se conhecia previamente.

Eumeros estudos estão sendo realizados a fim de esclarecer o seu mecanismo de regulação (LEE *et al.*, 2004; HAN *et al.*, 2004; BORCHERT; LANIER; DAVIDSON, 2006). Embora este processo ainda não esteja completamente desvendado, o conceito simplificado da biogênese do miRNA se aplica à maior parte dos organismos (WINTER *et al.*, 2009).

O processamento do miRNA (Figura 1) é realizado inicialmente no núcleo onde a transcrição dos miRNAs é feita pelas enzimas polII ou polIII, dando origem ao transcrito primário chamado (pri-miRNA) o qual pode conter um ou mais miRNAs. Este é então processado pelo complexo formado pela Droscha que reconhece o dobramento do RNA em formato de grampo (*hairpin*) e o processa, gerando o pre-miRNA (miRNA precursor) cujo o tamanho é de 70-100pbs. Em seguida, este é então exportado para fora do núcleo pela Exportin-5 (KIM, 2004).

No citoplasma, a Dicer se liga ao pre-miRNA retirando o duplex miRNA:miRNA* que se encontra na haste da estrutura de grampo (CHENDRIMADA *et al.*, 2005). A fita funcional ou miRNA maduro é carregado no complexo RISC (*RNA-induced silencing complex*) que o direciona até o mRNA alvo para reprimir a tradução ou em alguns casos para levar a degradação do mRNA (TANG, 2005).

Atualmente foram identificados 721 miRNAs em humanos, 579 em camundongos, 325 em ratos, 321 em cães e conforme a atualização 14 do banco de dados miRBase pode-se verificar a quantidade de miRNAs para mais de 100 outros organismos. Ainda

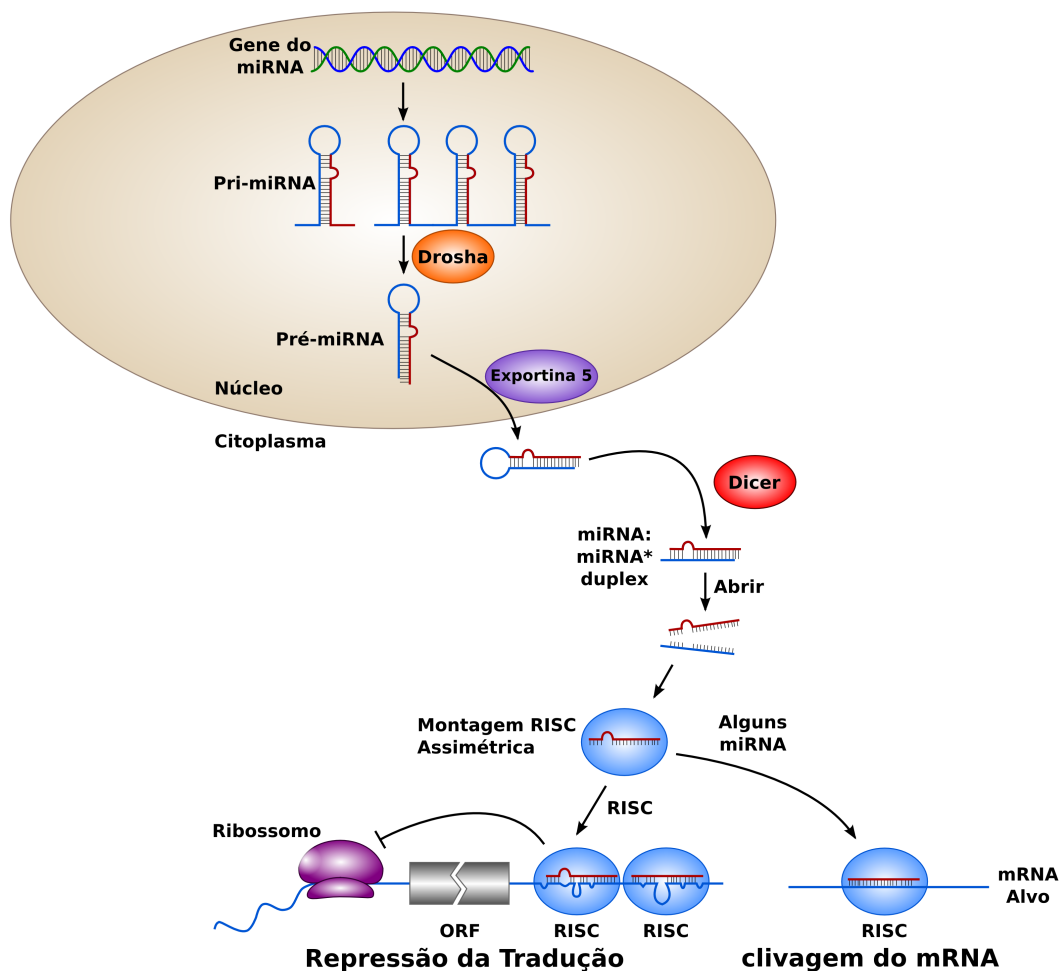


Figura 1: Representação do processo envolvendo a biogênese dos miRNAs. Adaptado de <http://uranus.mtveurope.org/wklim/research.html>.

não se sabe ao certo quantos miRNAs existem, mas estima-se que existam por volta de 800 miRNAs (BENTWICH *et al.*, 2005). Embora sejam muitos, a sua maioria é obtida computacionalmente e pouco se sabe sobre a sua função, os seus genes hospedeiros e a regulação das suas unidades de transcrição.

Uma forma de elucidar a questão das unidades de transcrição do miRNA é o estudo de alguns pri-miRNAs já identificados. Os pri-miRNAs gerados por seus genes estão organizados em (Figura 2):

- i) Transcritos não-codificadores e sem introns (pri-miR-21 e pri-miR-23a-27a-24.2);
- ii) Transcrito não codificador (pri-miR-155);
- iii) Intron (pri-miR-26b);
- iv) 3'UTR (pri-miR-198);

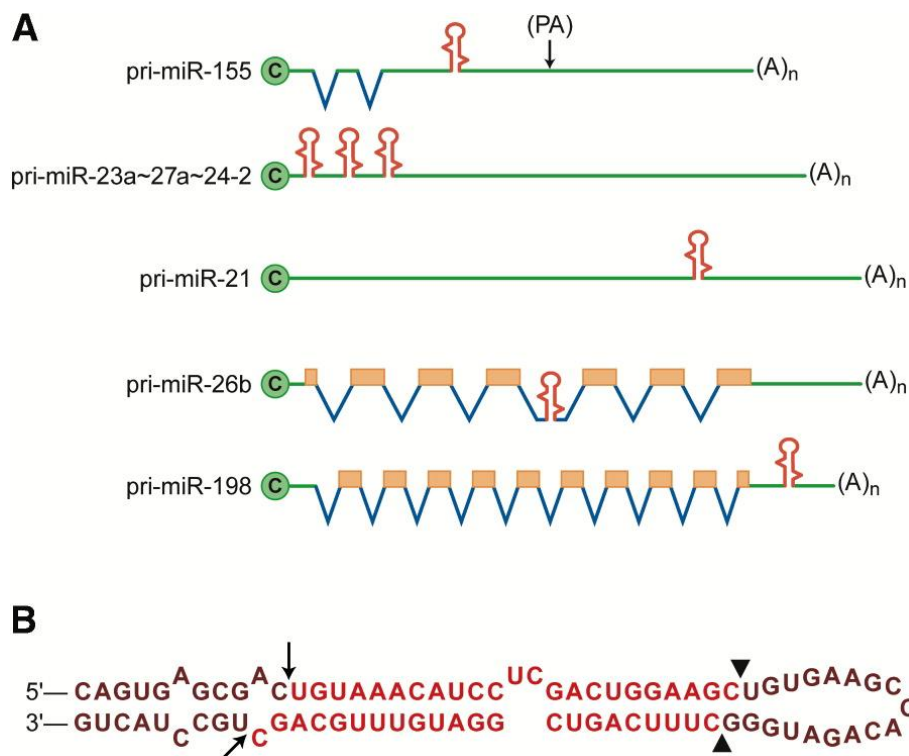


Figura 2: Representação esquemática dos transcritos de pri-miRNAs de humanos. A) Os pre-miRNAs estão em vermelho, regiões não codificantes em verde, introns em azul e regiões codificantes em laranja e PA é ponto alternativo de poli(A); B) estrutura de grampo. Locais de clivagem da Drosha são apontados pelas setas e o ponto de clivagem da Dicer é apontado pelos triângulos. Adaptado de (CULLEN, 2004).

Em i) e ii) são casos de genes próprios de miRNAs, eles possuem característica de genes comuns como a cauda poli(A), mas não produzem proteína. Os seus miRNAs maduros estão envolvidos em desenvolvimento de tumores (KULSHRESHTHA *et al.*, 2007). Já os pri-miRNAs dos itens iii) e iv) são miRNAs que estão dentro de genes. As abordagens computacionais mostram que diversos miRNAs encontram-se dentro de genes.

O pri-miR-26b está localizado no intron do gene CTDSP1 (*carboxy-terminal domain, RNA polymerase II, polypeptide A small phosphatase 1*) cuja proteína tem função desconhecida, porém ela é descrita como possível reguladora de expressão (YEO *et al.*, 2003). Enquanto que o pri-miR-198 está localizado na região 3'UTR do gene FLST1 (folistatin-like 1), uma proteína envolvida na produção do folículo do ovário e a baixa expressão do miRNA está ligado à anencefalia humana (HERRERA *et al.*, 2005; ZHANG *et al.*, 2010). Essas características mostram que os pri-miRNAs não seguem um padrão, exigindo um modelo complexo para aplicação na abordagem computacional.

A identificação de genes de miRNAs é difícil de ser realizada tanto em laboratório quanto utilizando programas. Já os precursores de miRNAs não são fáceis de detectar em experimentos, devido à especificidade espaço-temporal da expressão dos miRNAs e pela

baixa expressão de alguns deles. Porém os miRNAs possuem algumas características permitem uma investigação computacional. As principais são: i) os precursores possuem tamanho de 70-100nts; ii) estes deve estar estrutura secundária em forma de grampo; iii) o miRNA maduro localizado na haste do grampo e iv) os miRNAs possuem um alto grau de conservação entre os organismos.

Dentre os programas para predição utilizados no início da descoberta de novos miRNAs estão o miRseeker em *Drosophila* (LAI *et al.*, 2003), mirScan em *Caenorhabditis elegans* (LIM *et al.*, 2003) e *Phylogenetic shadowing* em humanos (BEREZIKOV *et al.*, 2005). Porém outras abordagens surgem melhorando e atualizando os algoritmos de predição.

Dentre as pesquisas recentes que envolvem funções atribuídas aos miRNAs, os tópicos mais abordados são as neoplasias e patologias humanas. Alguns exemplos de funções já identificadas são infarte, obesidade, Alzheimer e glioblastoma (Tabela 1).

Tabela 1: miRNAs e suas respectivas funções dentro de patologias.

| miRNA | Função | Referência |
|-------------------------------------|--------------------------|--------------------------------------|
| miR-1, miR-133a, miR-133b e miR-208 | Infarte do miocárdio | (BOVSTJANVCIVC <i>et al.</i> , 2009) |
| miR-519d | Obesidade | (MARTINELLI <i>et al.</i> , 2010) |
| miR-107 | Alzheimer | (WANG <i>et al.</i> , 2008) |
| miR-124, miR137 | Glioblastoma multiforme | (SILBER <i>et al.</i> , 2008) |
| miR-126 | Câncer de pulmão | (SUN <i>et al.</i> , 2009) |
| miR-30 | Carcinoma Hepatocelular | (YAO <i>et al.</i> , 2009) |
| miR-146a | Câncer de pâncreas | (LI <i>et al.</i> , 2010) |
| miR-141 | Naso-faringial carcinoma | (ZHANG <i>et al.</i> , 2010) |

Uma forma de investigar a função dos miRNAs é por meio do estudo de seus alvos verificando os seus papéis dentro das células, mas da mesma forma que a abordagem computacional é necessária para a predição de novos miRNAs, a elaboração de programas de predição de alvos também se tornou importante uma vez que poucos foram identificados experimentalmente e ainda não existe uma alternativa em larga escala (MAZIÈRE; ENRIGHT, 2007).

O primeiro programa foi desenvolvido por Stark *et al.* (2003) que aplicavam cadeias de Markov para identificar sequencias reversamente complementares aos miRNAs da *Drosophila melanogaster*. Logo surgiram outras abordagens para predição de alvos em outros organismos. Os mais conhecidos são o miRanda, TargetScan e TargetScanS, Pictar (KREK *et al.*, 2005), Diana-microT (KIRIAKIDOU *et al.*, 2004) e RNAHybrid (REHMSMEIER *et al.*, 2004). Estes são mantidos por seus laboratórios e conforme avançam as pesquisas sobre os miRNAs que são atualizados afim de obter predições mais precisas.

1.2 Perfis de expressão dos miRNAs

Os perfis de expressão são utilizados para a investigação de miRNAs que apresentam padrões anormais. Esses estudos investigam o envolvimento dos miRNAs nos processos biológicos, principalmente em doenças. O intuito é que estes possam ser potencialmente utilizados como ferramentas para diagnóstico, prognóstico ou terapia (CHO, 2009).

Esses perfis de expressão são obtidos a partir de técnicas como o *microarray* com miRNAs conhecidos ou com as novas tecnologias de sequenciamento unido com ferramentas de bioinformática (MASRI *et al.*, 2010). Porém estudos realizados por Ge, Wu e Wang (2006) e Rupert (2008) mostraram que perfis gerados por *Serial Analysis of Gene Expression* (SAGE) também podem ser utilizados.

Explicando a técnica de SAGE uma forma resumida, ela começa com a preparação do RNA (Figura 3). *Primers* com uma sequência de timinas marcados com biotina se ligam à partículas magnéticas com streptavidina. E a partir da hibridação com a cauda poli(A) dos transcritos se inicia a síntese dos cDNA da amostra. Depois são clivados pela enzima de ancoragem (*AE-Anchoring Enzyme*), ou seja, são cortados em um sítio específico.

Normalmente, é utilizada a enzima *NlaIII* que reconhece o sítio (CATG). Então elas são separadas em duas alíquotas. Um adaptador com sítio de reconhecimento da enzima de etiquetagem (*TE-Tagging Enzyme*), unido à um *primer* utilizado na amplificação por PCR (*Polymerase Chain Reaction*), é adicionado a cada uma com *primers* diferentes. Elas são digeridas por uma endonuclease de restrição do tipo IIS, usualmente a *BsmF1*, que corta abruptamente em 20bps a partir de seu sítio, a sequência que não é sítio de restrição é denominada *tag*.

As duas alíquotas são reunidas para que suas sequências se liguem formando uma *ditag*, as quais são amplificadas por PCR, digeridas novamente pela AE e reunidas em uma só, formando um concatêmero para ser sequenciado. Ao final do procedimento tem-se um conjunto de *tags* únicas provenientes da transcrição de vários genes.

1.3 Bancos de dados biológicos

Devido ao rápido crescimento na identificação de novo miRNAs preditos computacionalmente, foi criada uma base de dados chamada *The mirBase Registry* (GRIFFITHS-JONES, 2004). Ela é baseada nas diretrizes de anotação de Ambros *et al.* (2003) que oferece a garantia da unicidade do nome a partir de um pré-cadastramento das sequen-

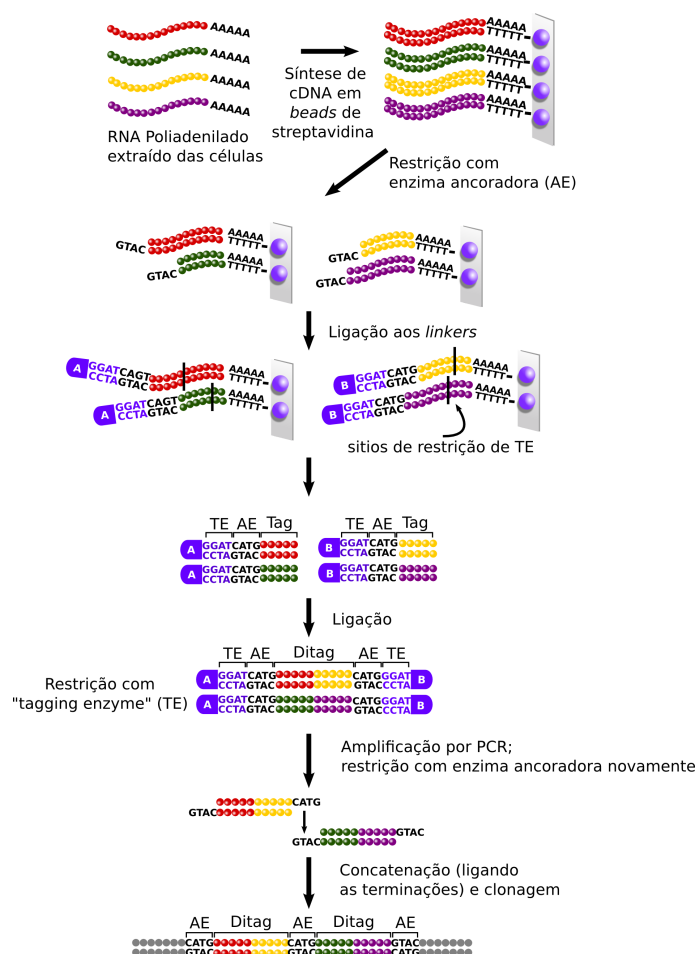


Figura 3: Representação simplificada da técnica de SAGE. Figura adaptada de <http://www.scq.ubc.ca/painless-gene-expression-profiling-sage-serial-analysis-of-gene-expression/>

cias. Após a publicação, o seu nome se torna definitivo. Isto é realizado a partir de uma interface *web* que também possibilita a obtenção das sequências dos miRNAs maduros e seus precursores.

Anos mais tarde, surgiu o *mirBase Database* (GRIFFITHS-JONES *et al.*, 2006), um repositório que além do *mirBase Registry*, também armazena e disponibiliza informações sobre as sequências e suas anotações (Figura 4).

Em 2008, o *mirBase* foi atualizado com uma nova ferramenta chamada *Genomics*. Ela utiliza programas de predição para mostrar elementos que estão na região próxima aos precursores dos miRNAs, tal como os locais de iniciação de transcrição, ilhas CPG, cauda polyA, locais de ligação de fatores de transcrição entre outras. Ela é composta por quatro tipos de busca: i) *By miRNA identifier or keyword*: direciona o usuário para a página com todas as informações do banco sobre o miRNA; ii) *By genomic location*: mostra todos os miRNAs que estão na localização desejada; iii) *For clusters*: mostra todos os agrupamentos

Figura 4: Site do miRBase e sua ferramenta *Genomics*

de miRNA de determinado organismo e iv) *By Sequence*: onde se entra com as sequencias e ele retorna todas os resultados que se alinham com a entrada (GRIFFITHS-JONES *et al.*, 2008).

As abordagens computacionais tem desempenhado um papel importante no estudo dos miRNAs principalmente na identificação de possíveis biomarcadores para prognósticos (CATALUCCI; GALLO; CONDORELLI, 2009; EACKER; DAWSON; DAWSON, 2009) e no seu envolvimento na regulação de processos patogênicos (MAZIÈRE; ENRIGHT, 2007; CHO, 2009). Algumas ferramentas que auxiliam no estudo da função dos miRNAs encontram-se disponíveis no miRBase¹, no microCosm², no miRGen³ (MEGRAW *et al.*, 2007; ALEXIOU *et al.*, 2010) e no miRNAMap⁴ (HSU *et al.*, 2006, 2008).

Cada um dos repositórios oferecem ferramentas diferentes. O microCosm oferece a busca por miRNAs e seus alvos, ele possui quatro tipos de entrada: i) nome do miRNA; ii) nome do gene; iii) código do gene no EnSEMBL⁵; iv) termos do *Gene Ontology*⁶ (GO) e v) palavra-chave.

O miRGen disponibiliza três ferramentas que são: i) *Genomics*; ii) *Targets* e iii) *Clusters*. A primeira possibilita a visualização da coleção de miRNAs em genomas completos dos organismos cadastrados. Como entrada é necessário selecionar o organismo, a localização do miRNA, a relação entre as fitas e base de genes que será usado como referência e o tipo de ordenação. Segunda possibilita o estudo de alvos de miRNA e cuja entrada é a

¹<http://www.mirbase.org/>

²<http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/>

³<http://www.diana.pcbi.upenn.edu/miRGen.html>

⁴<http://mirnamap.mbc.nctu.edu.tw/>

⁵<http://www.ensembl.org/index.html>

⁶<http://www.geneontology.org/>



Figura 5: miRGen e suas ferramentas: *Genomics*, *Targets* e *Clusters*.

seleção de organismo, a origem dos dados de genes alvos, o nome do miRNA ou do gene. A última permite o estudo de miRNAs que se encontram em grupo, sua entrada é a seleção do organismo, a distância entre os miRNAs do grupo e o conjunto de genes (Figura 5).

Já o miRNAMap fornece informações sobre miRNAs e alvos validados experimentalmente, alvos preditos e padrões de expressão que ajudam a fazer análise tanto do miRNA quanto do gene alvo (Figura 6).

Outro repositório que pode ser utilizado no estudo das funções dos miRNAs é o KEGG⁷ (*Kyoto Encyclopedia of Gene and Genomes*) que fornece informações sobre as vias metabólicas e seus respectivos genes.

Ele mantém em seu servidor uma API (*Application Programming Interface*) para disponibilizar suas informações que pode ser acessado dentro de um programa por meio da utilização do SOAP (*Simple Object Access Protocol*) sobre o protocolo HTTP juntamente

⁷<http://www.genome.jp/kegg/soap/>



Figura 6: miRNAmap e sua ferramenta para busca de alvos de miRNAs e padrões de expressão.

com o WSDL (*Web Service Definition Language*)⁸ (Figura 7).

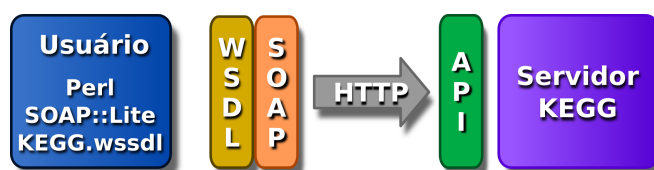


Figura 7: Descrição da estrutura do acesso às informações do KEGG sobre vias metabólicas e seus genes.

Até o presente momento não existe uma ferramenta que utilize dados de expressão para identificar miRNAs e seus alvos dentro de um mesmo experimento. A primeira ferramenta que permite correlacionar miRNAs, genes e vias metabólicas, chamada miRNApath, foi publicada por nosso grupo (CHIROMATZO *et al.*, 2007). Ela possibilita a identificação de vias metabólicas que podem ser afetadas por um determinado miRNA, baseado-se somente na informação de seus alvos preditos. Portanto, no presente estudo, é descrito a implementação de uma abordagem computacional adicional ao miRNApath que utiliza dados de expressão de SAGE para identificar de forma mais precisa as vias metabólicas afetadas por miRNAs dentro de um mesmo experimento.

⁸<http://soap.genome.jp/KEGG.wsdl>

2 *Objetivos*

Esta tese tem como objetivo principal desenvolver uma plataforma que possibilite identificar vias metabólicas afetadas por miRNAs através da análise de seus componentes que são os genes, as vias metabólicas e os miRNAs. Os objetivos específicos são:

- Desenvolver uma plataforma *web* que possibilite o estudo sobre a relação dos miRNAs com os seus genes alvos e as vias metabólicas;
- Desenvolver uma abordagem computacional para analisar a expressão de miRNAs e seus alvos em dados de SAGE de um mesmo experimento e avaliar a função dos miRNAs dentro da via metabólica;

3 *Material e métodos*

3.1 Modelagem do banco de dados

O banco de dados é baseado em repositórios públicos. Como a sua composição é heterogênea, não é possível carregá-lo com informações de todos os organismos. Afim de manter a consistência dos dados foram selecionados somente quatro deles. A Tabela 2 apresenta o nome dos organismos e os seus respectivos códigos de identificação no NCBI (*National Center for Biotechnology Information*). Com isto, as demais tabelas são carregadas da seguinte forma:

Tabela 2: Lista de organismos e seus respectivos códigos do NCBI.

| Organismo | Taxon id |
|--------------------------|----------|
| <i>Homo sapiens</i> | 9606 |
| <i>Mus musculus</i> | 10090 |
| <i>Rattus norvegicus</i> | 10116 |
| <i>Canis familiaris</i> | 9615 |

- **genes, tipogene, unigene, transcrito:** as informações foram extraídas do NCBI, elas são utilizadas principalmente como chaves de entrada. Somente a gene possui um detalhamento maior, por sua função dentro da plataforma;
- **mirna, gene_mirna:** os dados podem ser obtidos a partir do site miRBase¹. Já a relação dos miRNAs e seus alvos foram obtidos do *microCosm Targets*² e do miRGen³;
- **pathway, gene_path:** foi extraída a relação das vias metabólicas e seus genes a partir do repositório KEGG;

¹<ftp://mirbase.org/pub/mirbase/CURRENT/>

²<http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/>

³<http://www.diana.pcbi.upenn.edu/miRGen/v3/DownloadData.html>

- **tag,ref**: a análise das bibliotecas de SAGE também foi armazenada. A primeira tabela guarda as *tags* e seus respectivos genes e a segunda possui o gene ao qual a *tag* está associada.

A implementação do banco de dados relacional foi realizada utilizando o sistema de gerenciamento de bancos MySQL (v5.1.30) em servidor DELL PowerEdge 1500 SC com dois processadores Pentium III CPU 1.4GHz com 1Gb de memória. A Figura 8 apresenta o modelo do banco de dados relacional.

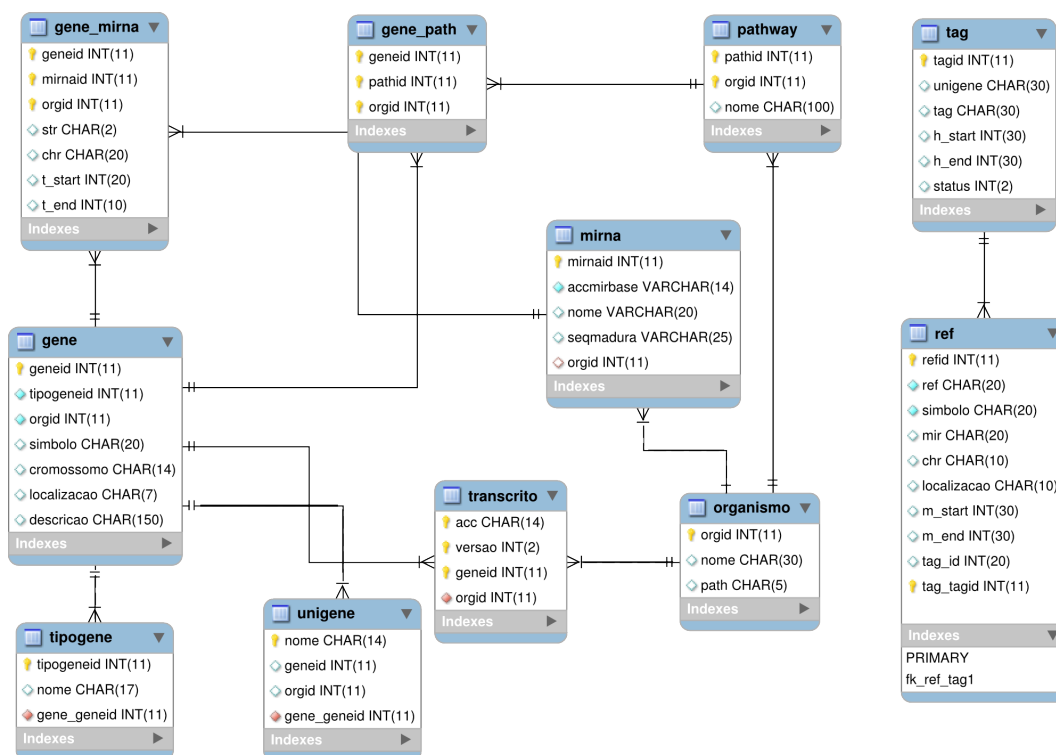


Figura 8: Representação das relações entre as tabelas do miRNAPath.

3.2 Diagrama de navegação

A Figura 9 descreve o diagrama de navegação proposto para a construção da plataforma de consulta miRNAPath. Inicialmente é apresentada a página principal (*home*) onde se encontram três outros links que direcionam para a página com os serviços de busca (*services*), para a lista das pessoas que colaboraram com o projeto (*credits*) e para o e-mail de contato (*contact*).

Dentro de *services* a sessão *Quick Search* e a *Advanced Analysis* direcionam para três tipos de buscas que viabilizam o estudo a partir ou do gene, ou do miRNA ou ainda da

via metabólica. Para a entrada de genes nas ferramentas de busca foi decidido utilizar o símbolo, ou o unigene, ou o código de acesso do NCBI ou o geneID, pois o NCBI é uma referência internacional para dados de biotecnologia.

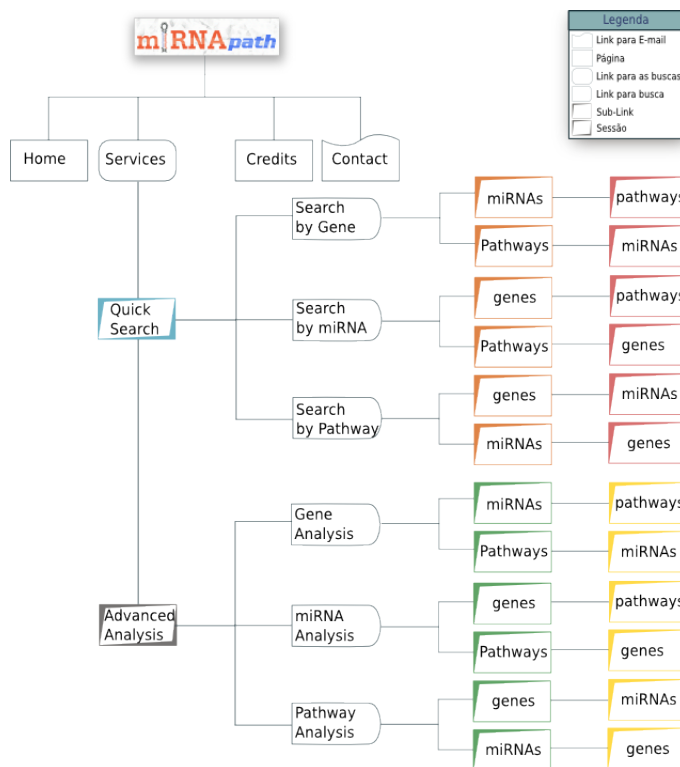


Figura 9: Figura representando o diagrama do miRNAPath.

3.3 Processamento do miRNAPath

Os dados dos repositórios públicos foram extraídos e formatados para a sua utilização no banco de dados do miRNAPath (Figura 10). Os programas desenvolvidos nesse processo utilizam a linguagem de programação Perl em conjunto com módulos específicos para conexão com banco de dados como o DBI e para a comunicação com APIs como o SOAP::Lite. A seguir está a listagem de cada um dos programas e sua função, eles são executados para cada organismo sendo que a Tabela 2 é fundamental para extrair informações específicas dos arquivos do NCBI.

- **parse_genBank.pl:** extrai as informações sobre os genes e o tipo de gene a partir do arquivo `gene_info`⁴ e as carrega na respectiva tabela;
- **parse_gene2refseq.pl:** retira do arquivo `gene2refseq`⁵ os dados sobre os transcritos

⁴ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz

⁵<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2refseq.gz>

e antes de inserir em sua tabela verifica se o gene já está cadastrado no banco de dados;

- **get_path_genes.pl**: conecta com a API do KEGG e retorna os dados sobre as vias, verifica se os genes estão inseridos e carrega a tabela de pathway e gene_path;
- **parse_unigene.pl**: extrai as informações sobre o UniGene⁶, verifica a existência do gene e insere em sua tabela;
- **parse_mirbase.pl**: carrega os dados sobre os miRNAs a partir do arquivo de maduros e faz a sua inserção;
- **sql_gene_mirna.pl**: carrega os arquivos de alvos de miRNAs do miRGen e do microCosm, verifica a existência do miRNA e a existência do gene e então faz a inserção no banco de dados.



Figura 10: Esquema do processo de desenvolvimento da plataforma miRNApath.

Os programas acima são responsáveis pela manutenção do banco de dados. Os demais apresentados na Tabela 3, foram desenvolvidos para a construção da plataforma miRNApath. Eles foram implementados utilizando linguagem PHP para geração de páginas HTML, juntamente com arquivo CSS *Cascading Style Sheets* para ajustes de *layout*. Eles encontram-se em um servidor Apache (v2.0.55).

⁶<ftp://ftp.ncbi.nih.gov/repository/UniGene>

Tabela 3: Lista dos programas utilizados pela plataforma *miRNApath*.

| Função | Programas |
|-----------------------------------|----------------------------------|
| Paginação | paginacao_analise.php |
| | paginacao_analysis_gene.php |
| | paginacao_gene.php |
| | paginacao_mirna2.php |
| | paginacao_mirna.php |
| | paginacao_path.php |
| Página de busca | paginacao_ret_from_mirna.php |
| | paginacao_ret_from_path.php |
| | query_analysis_mirna.php |
| | query_analysis.php |
| | query_gene.php |
| | query_mirna.php |
| Página de resultado | query_path.php |
| | query_path_title.php |
| | return_from_analysis_1_gene.php |
| | return_from_analysis_1_mirna.php |
| | return_from_analysis_1_path.php |
| | return_from_analysis_1.php |
| | return_from_analysis_gene.php |
| | return_from_analysis_mirna.php |
| | return_from_analysis_path.php |
| | return_from_analysis.php |
| | return_from_gene.php |
| | return_from_mirna.php |
| return_from_path.php | |
| Páginas de sub-busca no resultado | view_analysis_1_gene_mirna.php |
| | view_analysis_1_gene_path.php |
| | view_analysis_1_path_gene.php |
| | view_analysis_1_path_mirna.php |
| | view_analysis_gene.php |
| | view_analysis_mirna_path.php |
| | view_analysis_mirna.php |
| | view_analysis_path.php |
| | view_gene_mirna.php |
| | view_gene_path.php |
| | view_mirna_gene.php |
| | view_mirna_path.php |
| | view_path_gene.php |
| | view_path_mirna.php |
| | view_path.php |
| | path_gene.php |
| | path_mirna.php |
| | gene_mirna.php |
| gene_path.php | |
| path_gene.php | |
| path_mirna.php | |

3.4 Bibliotecas de SAGE

Esse método também tem sido utilizado para a detecção de precursores de miRNAs (GE; WU; WANG, 2006; RUPERT, 2008). Para exemplificar a utilização da plataforma miRNApath foram selecionadas as bibliotecas de tecido norma do sistema nervoso central e de astrocitomas de graus I, II, III e glioblastoma (Tabela 4). Elas foram obtidas do CGAP SAGE Genie⁷.

Tabela 4: Bibliotecas de sage e suas respectivas quantidade de *tags*.

| Cérebro | Biblioteca | Quantidade de Tags |
|----------------------|--|--------------------|
| Normal (5) | SAGE_Brain_normal_peds_cortex_B_H1571 | 83066 |
| | SAGE_Brain_normal_leptomeninges_B_AL2 | 77631 |
| | SAGE_Brain_normal_substantia_nigra_B_1 | 44592 |
| | SAGE_Brain_normal_cortex_B_pool6 | 66425 |
| | SAGE_Brain_normal_cortex_B_BB542 | 100476 |
| | Subtotal | 372190 |
| Astrocitoma I (1) | SAGE_Brain_astrocitoma_grade_I_B_H1043 | 80288 |
| Astrocitoma II (8) | SAGE_Brain_astrocitoma_grade_II_B_H563 | 94489 |
| | SAGE_Brain_astrocitoma_grade_II_B_H127 | 124052 |
| | SAGE_Brain_astrocitoma_grade_II_B_H518 | 124404 |
| | SAGE_Brain_astrocitoma_grade_II_B_H501 | 137766 |
| | SAGE_Brain_astrocitoma_grade_II_B_H359 | 113143 |
| | SAGE_Brain_astrocitoma_grade_II_B_H388 | 113304 |
| | SAGE_Brain_astrocitoma_grade_II_B_H516 | 114989 |
| | SAGE_Brain_astrocitoma_grade_II_B_H530 | 108972 |
| | Subtotal | 931119 |
| Astrocitoma III (11) | SAGE_Brain_astrocitoma_grade_III_B_H1020 | 53665 |
| | SAGE_Brain_astrocitoma_grade_III_B_H1055 | 117427 |
| | SAGE_Brain_astrocitoma_grade_III_B_R140 | 126988 |
| | SAGE_Brain_astrocitoma_grade_III_B_H272 | 102213 |
| | SAGE_Brain_astrocitoma_grade_III_B_407 | 115667 |
| | SAGE_Brain_astrocitoma_grade_III_B_343 | 107177 |
| | SAGE_Brain_astrocitoma_grade_III_B_R927 | 114831 |
| | SAGE_Brain_astrocitoma_grade_III_B_828 | 106923 |
| | SAGE_Brain_astrocitoma_grade_III_B_H970 | 114037 |
| | SAGE_Brain_astrocitoma_grade_III_B_584 | 110318 |
| | SAGE_Brain_astrocitoma_grade_III_B_439 | 114828 |
| | Subtotal | 1184074 |
| Glioblastoma (10) | SAGE_Brain_glioblastoma_B_R70 | 105325 |
| | SAGE_Brain_glioblastoma_B_H1353 | 133387 |
| | SAGE_Brain_glioblastoma_B_H1371 | 52131 |
| | SAGE_Brain_glioblastoma_B_R336 | 108853 |
| | SAGE_Brain_glioblastoma_B_pooled | 59907 |
| | SAGE_Brain_glioblastoma_B_R20 | 107239 |
| | SAGE_Brain_glioblastoma_B_H833 | 107448 |
| | SAGE_Brain_glioblastoma_B_GBM1062 | 63532 |
| | SAGE_Brain_glioblastoma_B_H1425C | 94821 |
| | SAGE_Brain_glioblastoma_B_H1110 | 72998 |
| | Subtotal | 905639 |
| Total | 35 | 3473310 |

⁷<http://cgap.nci.nih.gov/SAGE>

3.5 Processamento das análises

As análises realizadas para avaliação da plataforma miRNApath foram desenvolvidas separadamente. A primeira envolve dados de SAGE e a segunda dados de polimorfismo do gene HLA-G na população brasileira realizado por Castelli *et al.* (2008, 2009).

Antes de fazer a análise das bibliotecas de SAGE, é necessário fazer a extração e a preparação dos dados. Para este estudo foi realizada a seguinte abordagem (Figura 11):

- Obter a relação das *tags* com o seu respectivo código do UniGene através SAGEmap;
- Associar um miRNA a uma *tag*. Para tanto, foi necessário relacionar cada miRNA a um UniGene;
- Carregar as bibliotecas e fazer a normalização dos dados;
- Selecionar os miRNAs que não estão em genes conhecidos e possuem *tags* associadas. Isto porque a técnica de SAGE consegue identificar miRNAs precursores, mas não os maduros, devido a sua característica que necessita da presença da cauda polyA;
- Selecionar os miRNAs intergênicos que possuem *tags* associadas;

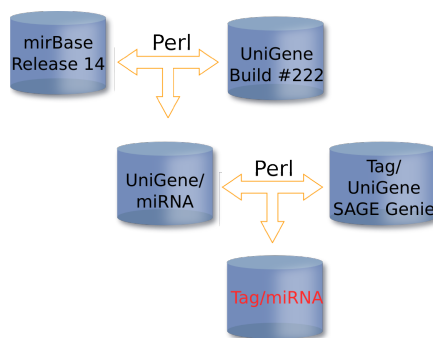


Figura 11: Estratégia desenvolvida para identificar *tags* de miRNAs.

3.6 Validação dos resultados

Os dados gerados pelas análises propostas foram validadas experimentalmente e *in Silico*. O resultado obtido da biblioteca de SAGE foi validado utilizando pelo Laboratório de Biologia Molecular e Bioinformática da Faculdade de Medicina de Ribeirão Preto-USP. Eles analisaram amostras de pacientes utilizando *Real-Time* PCR.

4 *Resultados e discussão*

A proposta deste trabalho é fornecer uma abordagem computacional que possibilite o estudo de vias metabólicas afetadas por miRNAs. A solução foi desenvolver um banco de dados que permitisse relacionar informações sobre genes, miRNAs e vias metabólicas. Sendo que para a realização de consultas foi implementada a plataforma de busca miRNAPATH (CHIROMATZO *et al.*, 2007) cuja publicação se encontra no apêndice A.

Em seguida, está a abordagem computacional para a identificação de vias metabólicas que podem ser afetadas por miRNAs utilizando dados de expressão gênica gerados por bibliotecas de SAGE.

4.1 miRNAPATH

Com a montagem do banco de dados, foi desenvolvida uma interface *web* para realizar consultas que facilitam o estudo sobre miRNAs. O miRNAPATH oferece duas sessões de buscas, a *Quick Search* e a *Advanced Analysis*, como pode ser visto na Figura 12. O primeiro permite um estudo mais abrangente sobre os dados de entrada, possibilitando a visualização de dados relacionados e o segundo mostra somente informações extritamente ligas à entrada. Ele encontra-se disponível em <http://lgmb.fmrp.usp.br/mirnapath> e o artigo da sua publicação está no apêndice .

4.1.1 *Quick Search*

Nesta sessão são possíveis três formas de buscas que podem ser feitas a partir de genes (*Search by Gene*), de miRNAs (*Search by miRNA*) ou vias metabólicas (*Search by Pathway*).

A página inicial da busca por gene (*Search by Gene*) é mostrada na Figura 13. Nela, pode-se inserir a lista de genes na caixa de texto ou carregar um arquivo. Depois se deve selecionar o organismo e o formato dos dados de entrada. Na Figura 14 está um exemplo

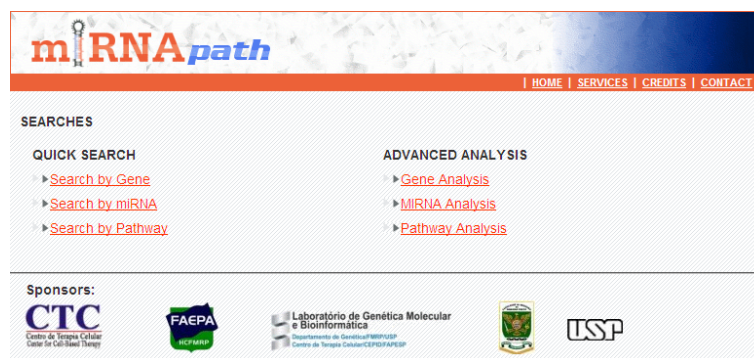


Figura 12: Página principal das buscas com as sessões *Quick Search* e *Advanced Analysis*.

do resultado desta busca, tendo como entrada uma lista de genes.

Em A) é apresentado o primeiro nível onde estão os genes da lista de entrada. Para cada gene é mostrado o organismo que foi selecionado, a localização cromossômica, a descrição, a quantidade de miRNAs que têm o gene como alvo e a quantidade de vias metabólicas nas quais o gene está envolvido. Os valores da quantidade de miRNA e de vias metabólicas são links para o nível seguinte mostrado em B).

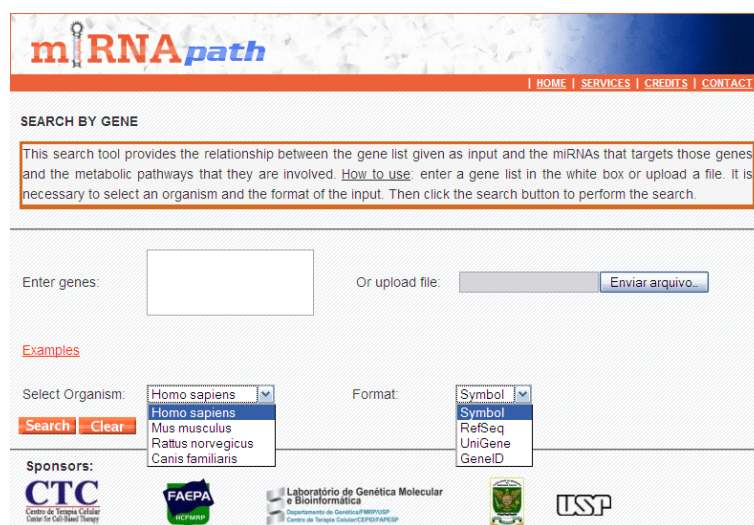



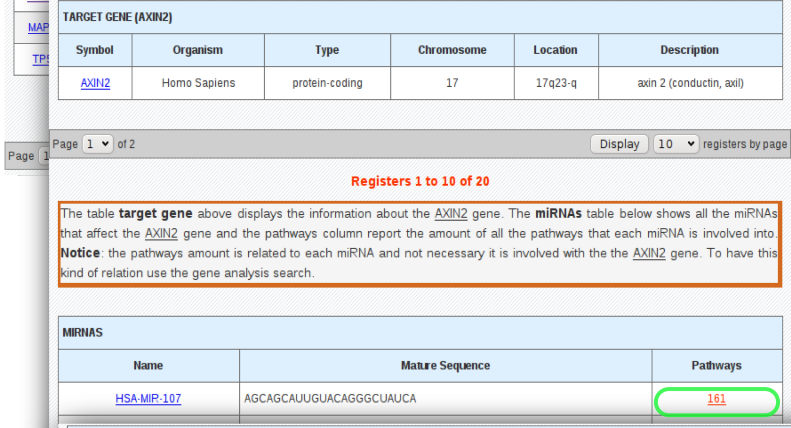
Figura 13: Página da busca por genes *Search by gene*.

Neste exemplo, o link que selecionado foi o da quantidade de miRNAs (marcado em verde), ele mostra na tabela *target gene*, o gene e suas respectivas informações e na tabela miRNAs, a lista completa dos miRNAs que tem o gene como alvo, a sequência do maduro e a quantidade de vias metabólicas que o respectivo miRNA também possui genes alvos. Este último também é um link para o próximo nível que pode ser visto em C).

Esta página apresenta três tabelas, a *target gene*, a *miRNA* e *pathways*. A primeira,

mostra as informações sobre o gene de entrada o qual se iniciou a navegação na busca. A seguinte, mostra o miRNA que foi selecionado para estudo e a última mostra todas as vias metabólicas que possuem genes que são afetados por ele. A via destacada em vermelho é aquela que possui o gene AXIN2, as demais são afetadas pelo miRNAs através de outros genes que também são alvos do miRNA.

A) 

B) 

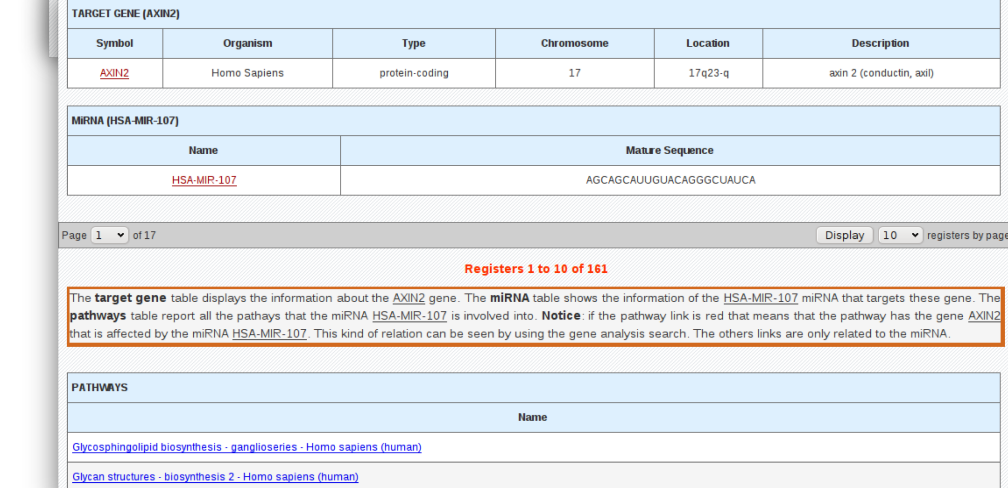
C) 

Figura 14: Páginas de resultado da busca. A seleção em verde é o link que foi clicado. A) Página com os genes da lista de entrada. B) Lista de todos os miRNAs que regulam o gene AXIN2. C) Lista de todas as vias metabólicas que são reguladas pelo miRNA hsa-mir-107. A via metabólica destacada em vermelho, é uma regulada pelo miRNA e possui o gene AXIN2.

A busca por miRNAs (*Search by miRNA*) possui um entrada parecida com a anterior, como mostra a Figura 15. Nesta pode-se entrar com a lista de miRNAs na caixa de texto ou carregar um arquivo e depois o selecionar o organismo.

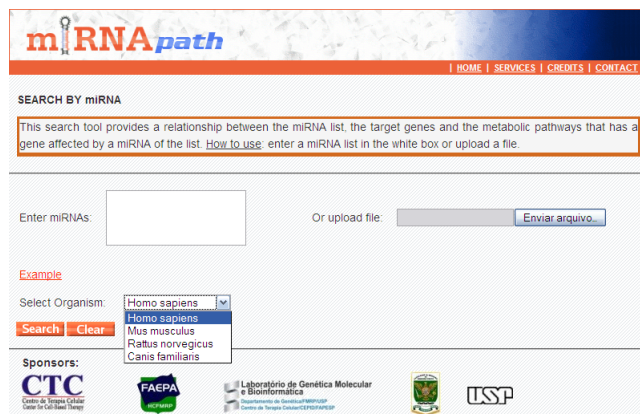


Figura 15: Página da busca por miRNAs *Search by miRNA*.

Já a busca por vias metabólicas (*Search by Pathway*), pode ser feita em duas etapas. Primeiro seleciona-se o organismo (Figura 16 A), depois é apresentada uma lista com todas as vias metabólicas do banco de dados (Figura 16 B). Diante do nome de cada via, está um quadrado que, ao ser clicado, seleciona o item que aparece na caixa de texto no final da página. Cada letra do alfabeto direciona para a região da página que contém as vias cujos nomes começam com a respectiva letra. Os links *top* e *down* apontam para o início e para o final da página, facilitando a navegação.

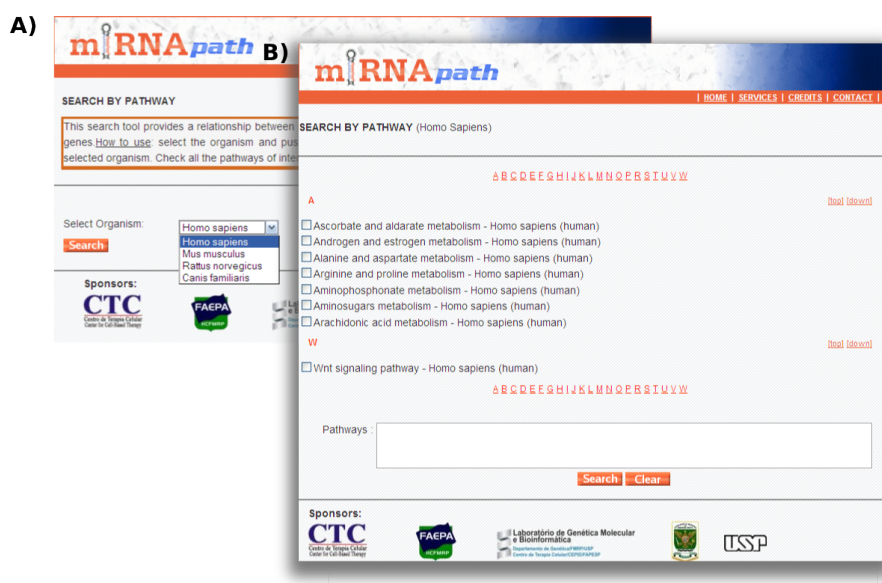
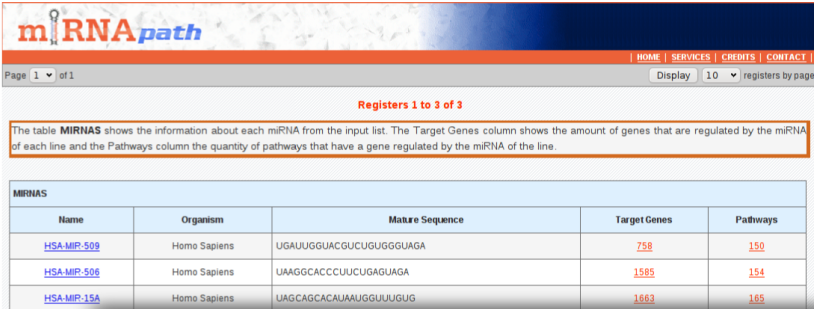


Figura 16: Página da busca por vias metabólicas *Search by Pathway*.

O resultado das buscas *Search by miRNA* e *Search by Pathway* utilizam um raciocínio análogo à busca anterior. Na Figura 17 estão os primeiros níveis de cada uma das buscas. Em A) Mostra para cada miRNA, os dados e os links para o segundo nível que são a quantidade de genes alvos para aquele miRNA e a quantidade vias metabólicas que aquele miRNA possui um gene alvo. Em B) Estão as informações de cada via metabólica e os links para o próximo nível que são a quantidade de miRNAs que afetam a via e a quantidade de genes na via.

A)



miRNApath

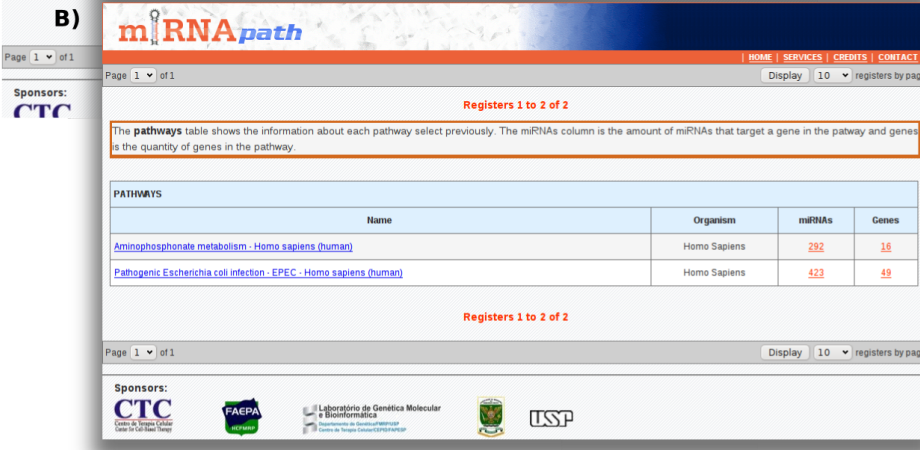
Page 1 of 1 | HOME | SERVICES | CREDITS | CONTACT | Display 10 registers by page

Registers 1 to 3 of 3

The table **MIRNAS** shows the information about each miRNA from the input list. The Target Genes column shows the amount of genes that are regulated by the miRNA of each line and the Pathways column the quantity of pathways that have a gene regulated by the miRNA of the line.

| MIRNAS | Name | Organism | Mature Sequence | Target Genes | Pathways |
|--------|-----------------------------|--------------|------------------------|--------------|----------|
| | HSA-MIP-509 | Homo Sapiens | UGAUUGGUACGUCUGUGGUAGA | 758 | 150 |
| | HSA-MIP-506 | Homo Sapiens | UAAGGCACCCUUCUGAGUAGA | 1585 | 154 |
| | HSA-MIP-154 | Homo Sapiens | UAOCAGCACAAUGGUUUUGUG | 1663 | 165 |

B)



miRNApath

Page 1 of 1 | HOME | SERVICES | CREDITS | CONTACT | Display 10 registers by page

Sponsors: CTC

Registers 1 to 2 of 2

The **pathways** table shows the information about each pathway select previously. The miRNAs column is the amount of miRNAs that target a gene in the pathway and genes is the quantity of genes in the pathway.

| PATHWAYS | Name | Organism | miRNAs | Genes |
|----------|---|--------------|--------|-------|
| | Aminophosphonate metabolism - Homo sapiens (human) | Homo Sapiens | 292 | 16 |
| | Pathogenic Escherichia coli infection - EPEC - Homo sapiens (human) | Homo Sapiens | 423 | 49 |

Registers 1 to 2 of 2

Page 1 of 1 | Display 10 registers by page

Sponsors: CTC, FAPEA, Laboratório de Genética Molecular e Bioinformática, USP

Figura 17: Primeiro nível do resultado das buscas. A) Por miRNAs e B) por vias metabólicas.

4.1.2 Advanced Analysis

A sessão *Advances Analysis* possui três tipos de buscas com entradas semelhantes à sessão anterior, que podem ser feitas por genes (*Gene Analysis*), por miRNAs (*miRNA Analysis*) e por vias metabólicas (*Pathway Analysis*). As diferenças estão na forma de apresentação e no conteúdo do resultado da busca, uma vez que só apresenta informações extritamente relacionadas com a entrada.

Figura 18: Primeiro nível do resultado das buscas da sessão *Advanced Analysis*.

Para cada uma das buscas do primeiro nível é o mesmo o qual está na Figura 18. Este é composto por três formas de visualização do resultado. A *Order by Gene*, que mostra o resultado da busca ordenado pelos genes, a *Order by miRNA* que ordenado por miRNAs e a *Order by Pathway* que faz o mesmo para as vias metabólicas.

O exemplo utilizado nas figuras seguintes é baseado em uma busca cuja a entrada é uma lista de genes. A Figura 19 mostra o resultado da busca ordenado por genes (*Order by Gene*). Em A) é possível verificar que o gene TP53 está presente em 16 vias metabólicas e é afetado por 38 miRNAs e em B) pode-se observar todos os nomes das vias e dos miRNAs para o gene.

A)

| GENES | Symbol | Number of Pathways | List of miRNAs that affect the gene |
|-------|--------|--------------------|-------------------------------------|
| | TP53 | 16 | 18 HSA-MIP-143 HSA-LET-7I |
| | AXIN2 | 4 | 20 HSA-MIP-600 |
| | MAPK7 | | |

B)

| GENE TP53 | | | | miRNAs | |
|-------------|--------------|----------|--|--|--|
| Symbol | Organism | Location | Description | | |
| TP53 | Homo Sapiens | 17p13.1 | tumor protein p53 (L1-Fraumeni syndrome) | HSA-MIP-143 HSA-LET-7I HSA-MIP-600 HSA-MIP-620 HSA-MIP-92 HSA-MIP-22 HSA-MIP-604 HSA-MIP-381 HSA-MIP-150 HSA-MIP-550 HSA-MIP-189 HSA-MIP-433 HSA-MIP-387 HSA-MIP-452 HSA-MIP-375 HSA-MIP-928 HSA-MIP-152 HSA-MIP-634 | |


PATHWAYS

- MAPK signaling pathway - Homo sapiens (human)
- Cell cycle - Homo sapiens (human)
- Apoptosis - Homo sapiens (human)
- Wnt signaling pathway - Homo sapiens (human)
- Amnophic lateral sclerosis (ALS) - Homo sapiens (human)
- Huntington's disease - Homo sapiens (human)
- Colorectal cancer - Homo sapiens (human)

Figura 19: Página do resultado da busca, ordenada por genes. A) Lista de genes e a contagem de vias metabólicas e miRNAs. A caixa ao lado da quantidade de miRNAs mostra a listagem com os nomes dos miRNAs. B) Seguindo o link de vias metabólicas (*Pathways*). Tem-se a informação sobre o gene, seus alvos e as vias onde está envolvido.


A ordenação por miRNA (Figura 20) permite fazer a análise de quantos genes da entrada podem ser regulados por cada miRNA e quantas vias envolvem aqueles genes. Para uma visão mais detalhada das informações é só seguir os links. E por último, a ordenação por vias metabólicas (Figura 21) permite fazer a análise sobre uma via específica, vendo quais os genes da entrada que são regulados por miRNA e podendo verificar se existe um miRNA que regule mais que um gene ou que seja específico.

A)



| NAME | Number of Genes of the List | Number of pathways |
|-----------------------------|-----------------------------|--------------------|
| HSA-MIR-556 | 2 | 7 |
| HSA-MIR-92B | 2 | 16 |

B)



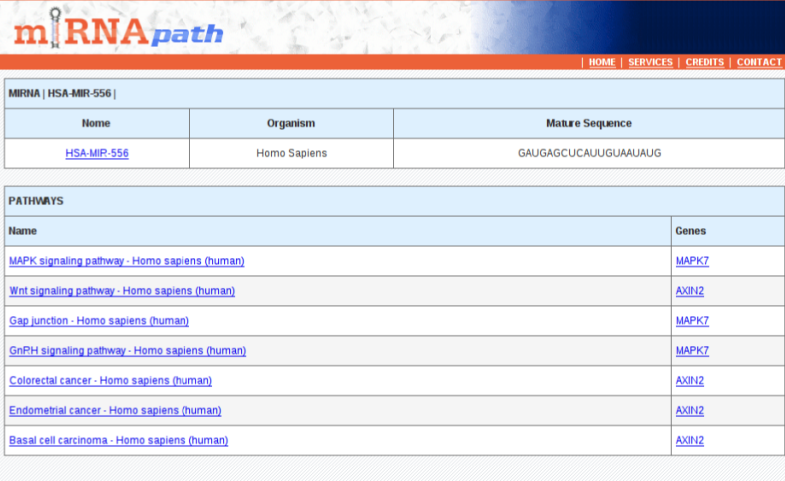
MIRNA | HSA-MIR-556 |

| Name | Organism | Mature Sequence |
|-----------------------------|--------------|---------------------|
| HSA-MIR-556 | Homo Sapiens | GAUGAGCUCAUUGUAAUUG |

GENES

| Symbol | Location | Description | Pathways |
|-----------------------|----------|------------------------------------|---|
| MAPK7 | 17p11.2 | mitogen-activated protein kinase 7 | MAPK signaling pathway - Homo sapiens (human) Gap junction - Homo sapiens (human) GnRH signaling pathway - Homo sapiens (human) Wnt signaling pathway - Homo sapiens (human) |
| AXIN2 | | | |

C)



MIRNA | HSA-MIR-556 |

| Name | Organism | Mature Sequence |
|-----------------------------|--------------|---------------------|
| HSA-MIR-556 | Homo Sapiens | GAUGAGCUCAUUGUAAUUG |

PATHWAYS

| Name | Genes |
|---|-----------------------|
| MAPK signaling pathway - Homo sapiens (human) | MAPK7 |
| Wnt signaling pathway - Homo sapiens (human) | AXIN2 |
| Gap junction - Homo sapiens (human) | MAPK7 |
| GnRH signaling pathway - Homo sapiens (human) | MAPK7 |
| Colorectal cancer - Homo sapiens (human) | AXIN2 |
| Endometrial cancer - Homo sapiens (human) | AXIN2 |
| Basal cell carcinoma - Homo sapiens (human) | AXIN2 |

Figura 20: Página do resultado da busca, ordenado por miRNAs.

A)

B)

Figura 21: Página do resultado da busca, ordenado por vias metabólicas.

4.2 Avaliação e identificação de tags de miRNAs

A primeira etapa do processamento dos dados do SAGE é mapear os miRNAs com os seus respectivos códigos do UniGene e tags. Para tanto foram utilizadas 35 bibliotecas de SAGE depositadas no SAGE Genie¹. O resultado revelou 351 tags únicas associadas aos miRNAs, a maior parte dos miRNAs (241) encontra-se dentro de genes conhecidos sendo que 131 estão dentro de introns, 104 em exons, 4 na junção intron/exon e 2 na junção exon/intron (Figura 22). Os 110 restantes são intergênicos.

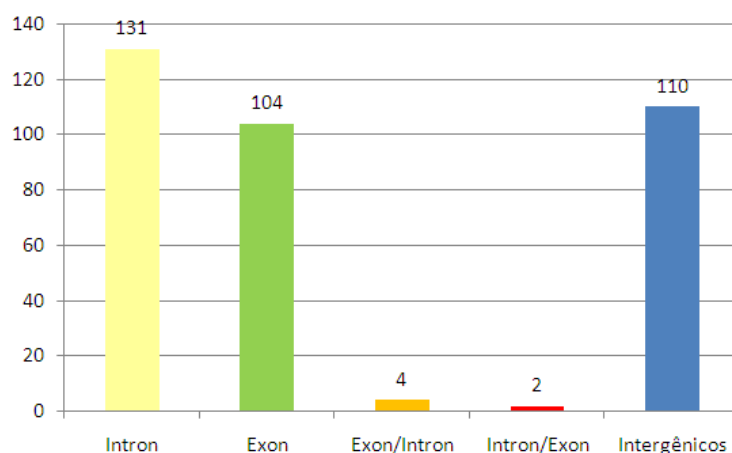


Figura 22: Distribuição dos miRNAs dentro de genes e intergênicos.

¹ftp://ftp1.nci.nih.gov/pub/SAGE/HUMAN/Hs_short.best_gene.gz

O estudo concentrou a análise em *tags* intergênicas, pela maior chance de se identificar precursores próprios de miRNAs. A avaliação das bibliotecas de astrocitoma identificou a expressão de três deles, o miR-17, o miR-21 e o miR-155. A Tabela 5 mostra as *tags* encontradas, os UniGenes associados, os miRNAs e seus níveis de expressão nos diferentes graus de glioblastoma.

Tabela 5: miRNAs identificados pela técnica de SAGE.

| Tag | UniGene | miRNA | Normal | Grau I | Grau II | Grau III | Glioblastoma |
|------------|-----------|-------------|--------|--------|---------|----------|--------------|
| TTTACTTTTA | Hs.652877 | hsa-miR-17 | 0.00 | 0.00 | 0.27 | 0.26 | 0.56 |
| ATGGCGGGCA | Hs.697120 | hsa-miR-155 | 6.72 | 26.16 | 3.16 | 5.44 | 3.18 |
| ATCTCAAAGA | Hs.444569 | hsa-miR-21 | 6.75 | 26.16 | 18.95 | 20.15 | 35.44 |

Utilizando a plataforma miRNAPath foi observado que os três miRNAs possuem genes alvos na via metabólica do glioma² (Figura 23). Os genes da via que são alvos dos miRNAs somam 26%. Sendo que os genes E2F1, E2F3, PTEN e PIK3R1 podem estar sendo regulados ao mesmo tempo pelos miRNAs, miR-21 e miR-17.

O miRNA miR-17 é predito como regulador dos genes E2F1, EGF, HRAS, PIK3CD, AKT3, CDKN1A, E2F3, PDGFRA, PIK3R1, PIK3R2, PTEN, RB1 e CCND1 da via. Embora apresente pouca expressão, ele está descrito na literatura como altamente expresso (MALZKORN *et al.*, 2009) regulando negativamente a expressão do gene E2F1 (O'DONNELL *et al.*, 2005). Isto pode ser devido à limitação da técnica.

Já o miR-155 não foi validado experimentalmente como regulador de genes em glioblastoma, ele está correlacionado com certos tumores (TAM; DAHLBERG, 2006) e tem como alvos os genes KRAS, PIK3CA e SOS1 da via do glioma. A busca *Advanced Analysis* para miRNA ordenada por via metabólica mostra o envolvimento do miR-155 em diversas em diversas vias de tumores tais como o de pulmão, o de rim, a leucemia e o de pâncreas. Além disto participa da via apoptótica e das células exterminadoras celulares mediante à citotoxina.

²http://www.genome.jp/dbget-bin/www_bget?pathway:hsa05214

| PATHWAY Glioma - Homo sapiens (human) | | Name | | Organism | |
|---|----------------|-------------------------------|----------|---|---|
| | | Glioma - Homo sapiens (human) | | Homo Sapiens | |
| GENES | | | | | |
| Symbol | Type | Chromosome | Location | Description | MiRNAs |
| CCND1 | protein-coding | 11 | 11q13 | cyclin D1 | HSA-MIR-17-5P |
| CDKN1A | protein-coding | 6 | 6p21.2 | cyclin-dependent kinase inhibitor 1A (p21, Cip1) | HSA-MIR-17-5P |
| E2F1 | protein-coding | 20 | 20q11.2 | E2F transcription factor 1 | HSA-MIR-17-3P HSA-MIR-17-5P |
| E2F3 | protein-coding | 6 | 6p22 | E2F transcription factor 3 | HSA-MIR-21 HSA-MIR-17-5P |
| EGF | protein-coding | 4 | 4q25 | epidermal growth factor (beta-urogastrone) | HSA-MIR-17-3P |
| HRAS | protein-coding | 11 | 11p15.5 | v-Har-ras Harvey rat sarcoma viral oncogene homolog | HSA-MIR-17-3P HSA-MIR-21 |
| KIPAS | protein-coding | 12 | 12p12.1 | v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog | HSA-MIR-155 |
| PDGFRA | protein-coding | 4 | 4q11-q1 | platelet-derived growth factor receptor, alpha polypeptide | HSA-MIR-17-5P |
| PIK3CA | protein-coding | 3 | 3q26.3 | phosphoinositide-3-kinase, catalytic, alpha polypeptide | HSA-MIR-155 |
| PIK3CD | protein-coding | 1 | 1p36.2 | phosphoinositide-3-kinase, catalytic, delta polypeptide | HSA-MIR-17-3P |
| PIK3P1 | protein-coding | 5 | 5q13.1 | phosphoinositide-3-kinase, regulatory subunit 1 (p85 alpha) | HSA-MIR-21 HSA-MIR-17-5P |
| PIK3P2 | protein-coding | 19 | 19q13.2 | phosphoinositide-3-kinase, regulatory subunit 2 (p85 beta) | HSA-MIR-17-5P |
| PTEN | protein-coding | 10 | 10q23.3 | phosphatase and tensin homolog (mutated in multiple advanced cancers 1) | HSA-MIR-21 HSA-MIR-17-5P |
| PBL1 | protein-coding | 13 | 13q14.2 | retinoblastoma 1 (including osteosarcoma) | HSA-MIR-17-5P |
| SOS1 | protein-coding | 2 | 2p22-p2 | son of sevenless homolog 1 (Drosophila) | HSA-MIR-155 |
| TFDP1 | protein-coding | 13 | 13q34 | transcription factor Dp-1 | HSA-MIR-21 |
| AKT3 | protein-coding | 1 | 1q43-q4 | v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma) | HSA-MIR-17-5P |

Figura 23: MiRNAs expressos no estudo de SAGE e seus genes alvos na via metabólica do glioma.

Analisando a expressão dos genes e do miRNA, pode-se sugerir que o gene SOS1 seja regulado pelo miR-155 (Figura 24), pois eles apresentam uma expressão inversamente proporcional.

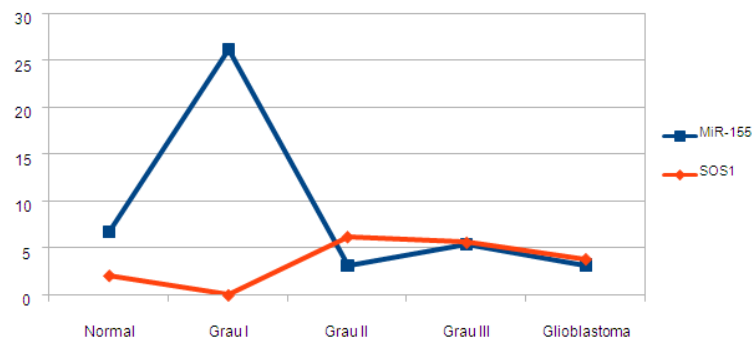


Figura 24: Expressão de SAGE para o miR-155 e do gene SOS1.

Para validação experimental da *tag* de miRNA foi utilizada a *tag* do miR-21 que já é descrito na literatura como altamente expresso em glioblastomas (CHEN *et al.*, 2008; CONTI *et al.*, 2009). A análise da *tag* do miR-21 também revela uma expressão elevada do miR-21 em glioblastoma (grau IV) e baixa expressão em tecido normal. Ele foi encontrado até 5 vezes mais expresso em glioblastoma do que em tecidos normais. A validação experimental da expressão do miR-21 para a *tag* ATCTCAAAGA foi feita por *real-time* e demonstrou que o miR-21 se encontra expresso nos diferentes graus do glioblastoma (Figura 25).

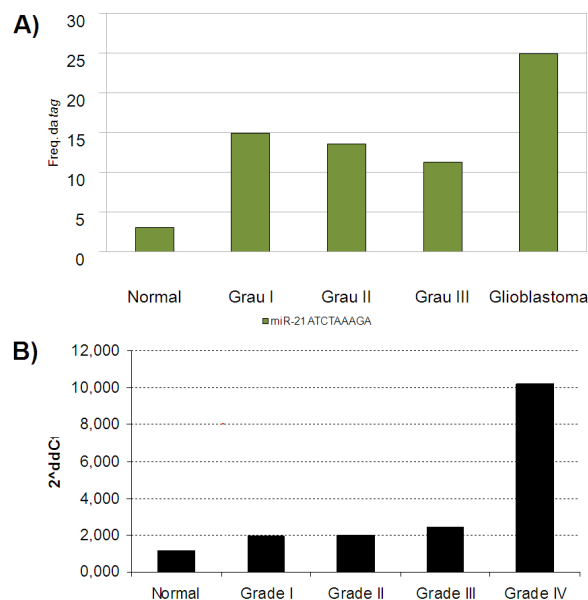


Figura 25: Expressão do miR-21 A) Expressão dos tags de miRNA em SAGE e B) expressão validada por RT-PCT.

Utilizando a plataforma miRNAPath para investigar o miR-21, foram encontrados

1012 genes alvos. Os seus genes alvos da via glioma são E2F3, HRAS, PIK3R1, PTEN e TFDP1. Dentre eles somente o PTEN está descrito na literatura como alvo validado (MENG *et al.*, 2007). Dos genes que não participam da via do glioma, foi selecionado o PDCD4 que é descrito como envolvido na apoptose celular. Embora não esteja associado a uma via metabólica, ele é um alvo comprovado pela literatura (CHEN *et al.*, 2008).

Os gráficos na Figura 26 mostram a expressão da *tag* do miR-21 e dos genes PTEN e PDCD4 nos diferentes graus de glioblastoma. O gráfico mostra que os genes possuem uma expressão inversamente proporcional miR-21 em cada estágio. O resultado confirma o que é descrito na literatura e ainda mostra que o gene PDCD4 parece ser mais modulado pelo miR-21 do que o PTEN.

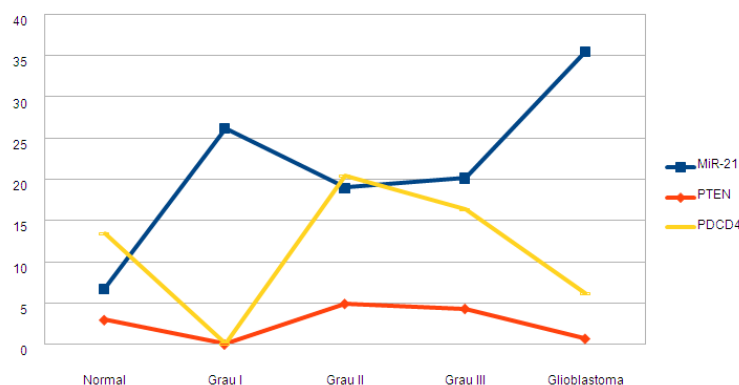


Figura 26: Expressão da tag de miR-21 e dos genes PTEN e PDCD4 nas bibliotecas de SAGE para o estudo de glioblastoma.

5 Conclusões

A plataforma miRNApath publicada em 2007 (CHIROMATZO *et al.*, 2007) possui um banco de dados baseada em repositórios públicos que associa os dados de miRNA, genes e vias metabólicas. Para viabilizar o estudo da relação entre os dados, foram implementados dois tipos de consultas: a *Quick Search* e a *Advanced Analysis*. A diferença principal entre elas é que segunda mostra informações restritas às que foram inseridas como entrada. Ambas possuem três formas de entrada: com miRNAs (*Search by mirna*), com genes (*Search by gene*) ou com vias metabólicas (*Search by metabolic pathways*).

Como segunda parte deste trabalho, foi desenvolvida uma abordagem computacional para identificar miRNAs e seus alvos em bibliotecas de SAGE, possibilitando assim, avaliar em um mesmo experimento a expressão dos miRNAs e seus alvos. Utilizando como exemplo as bibliotecas do tecido do sistema nervoso central, foi possível identificada a expressão dos miRNAs: miR-17, miR-21 e miR-155.

A busca das vias metabólicas pelo *Advanced Analysis* apresentou diversas vias metabólicas, porém a escolhida foi o glioma. Investigando os genes desta via pode-se verificar que 15 deles (26%) são alvos preditos para os miRNAs selecionados.

Dentre esses genes três (SOS1, PTEN e PDCD4) apresentaram uma relação de expressão inversamente proporcional a dos seus miRNAs reguladores. Observando os gráficos de expressão das *tags* de miRNAs e de seus genes alvos verificou-se que a expressão da *tag* do miR-155 é inversamente proporcional à do gene SOS1 (Figura 24). Atualmente, a literatura mostra que o miR-155 está expresso em tumores, mas não há genes alvos validados experimentalmente. Isto mostra que o gene SOS1 pode ser considerado como um provável gene regulado pelo miR-155 em tecidos do sistema nervoso.

Os outros dois, PTEN e o PDCD4, possuem expressão inversamente proporcional à suas *tags*. Sendo que a literatura confirma que os genes PTEN e PDCD4 são genes alvos deste miRNA. Porém, o estudo mostra que o PDCD4 é mais sensível à modulação do miR-21 do que o PTEN.

O miR-21 é reconhecido na literatura como envolvido em glioblastoma, para validar a sua expressão com a *tag*(ATCTCAAAGA) do miR-21, foi realizado o *real-time PCR* que comprovou a expressão encontrada nas bibliotecas de SAGE.

Em resumo, o presente trabalho descreve uma abordagem computacional consistente para identificar vias metabólicas que podem ser afetadas por miRNAs, avaliando a sua expressão e de seus alvos em um mesmo experimento de expressão gênica. O estudo também auxilia na validação de alvos de miRNAs.

Referências Bibliográficas

- ALEXIOU, P. *et al.* mirgen 2.0: a database of microRNA genomic information and regulation. *Nucleic Acids Res*, v. 38, n. Database issue, p. D137–41, jan. 2010. ISSN 1362-4962.
- AMBROS, V. *et al.* A uniform system for microRNA annotation. *RNA*, v. 9, n. 3, p. 277–9, mar. 2003. ISSN 1355-8382.
- BATUWITA, R.; PALADE, V. micropred: effective classification of pre-mirnas for human mirna gene prediction. *Bioinformatics*, v. 25, n. 8, p. 989–95, abr. 2009. ISSN 1367-4811.
- BENTWICH, I. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet*, v. 37, n. 7, p. 766–70, jul. 2005. ISSN 1061-4036.
- BEREZIKOV, E. *et al.* Phylogenetic shadowing and computational identification of human microRNA genes. *Cell*, v. 120, n. 1, p. 21–4, jan. 2005. ISSN 0092-8674.
- BORCHERT, G. M.; LANIER, W.; DAVIDSON, B. L. Rna polymerase iii transcribes human microRNAs. *Nat Struct Mol Biol*, v. 13, n. 12, p. 1097–101, dez. 2006. ISSN 1545-9993.
- BOVSTJANVCIVC, E. *et al.* MicroRNAs mir-1, mir-133a, mir-133b and mir-208 are dysregulated in human myocardial infarction. *Cardiology*, v. 115, n. 3, p. 163–169, dez. 2009. ISSN 1421-9751.
- CASTELLI, E. C. *et al.* Hla-g polymorphism and transitional cell carcinoma of the bladder in a brazilian population. *Tissue Antigens*, v. 72, n. 2, p. 149–57, ago. 2008. ISSN 0001-2815.
- CASTELLI, E. C. *et al.* In silico analysis of microRNAs targeting the hla-g 3' untranslated region alleles and haplotypes. *Hum Immunol*, v. 70, n. 12, p. 1020–5, dez. 2009. ISSN 1879-1166.
- CATALUCCI, D.; GALLO, P.; CONDORELLI, G. MicroRNAs in cardiovascular biology and heart disease. *Circ Cardiovasc Genet*, v. 2, n. 4, p. 402–8, ago. 2009. ISSN 1942-3268.
- CHEN, Y. *et al.* MicroRNA-21 down-regulates the expression of tumor suppressor pcdcd4 in human glioblastoma cell t98g. *Cancer Lett*, v. 272, n. 2, p. 197–205, dez. 2008. ISSN 1872-7980.
- CHENDRIMADA, T. P. *et al.* Trbp recruits the dicer complex to ago2 for microRNA processing and gene silencing. *Nature*, v. 436, n. 7051, p. 740–4, ago. 2005. ISSN 1476-4687.

- CHIROMATZO, A. O. *et al.* mirnapath: a database of mirnas, target genes and metabolic pathways. *Genet Mol Res*, v. 6, n. 4, p. 859–65, 2007. ISSN 1676-5680.
- CHO, W. C. S. Micrnas: Potential biomarkers for cancer diagnosis, prognosis and targets for therapy. *Int J Biochem Cell Biol*, dez. 2009. ISSN 1878-5875.
- CONTI, A. *et al.* mir-21 and 221 upregulation and mir-181b downregulation in human grade ii-iv astrocytic tumors. *J Neurooncol*, v. 93, n. 3, p. 325–32, jul. 2009. ISSN 1573-7373.
- CULLEN, B. R. Transcription and processing of human microRNA precursors. *Mol Cell*, v. 16, n. 6, p. 861–5, dez. 2004. ISSN 1097-2765.
- Di Leva, G.; CALIN, G. A.; CROCE, C. M. Micrnas: fundamental facts and involvement in human diseases. *Birth Defects Res C Embryo Today*, v. 78, n. 2, p. 180–9, jun. 2006. ISSN 1542-975X.
- EACKER, S. M.; DAWSON, T. M.; DAWSON, V. L. Understanding micrnas in neurodegeneration. *Nat Rev Neurosci*, v. 10, n. 12, p. 837–41, dez. 2009. ISSN 1471-0048.
- GE, X.; WU, Q.; WANG, S. M. Sage detects microRNA precursors. *BMC Genomics*, v. 7, p. 285, 2006. ISSN 1471-2164.
- GREENE, S. B. *et al.* A putative role for microRNA-205 in mammary epithelial cell progenitors. *J Cell Sci*, jan. 2010. ISSN 1477-9137.
- GRIFFITHS-JONES, S. The microRNA registry. *Nucleic Acids Res*, v. 32, n. Database issue, p. D109–11, jan. 2004. ISSN 1362-4962.
- GRIFFITHS-JONES, S. *et al.* mirbase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, v. 34, n. Database issue, p. D140–4, jan. 2006. ISSN 1362-4962.
- GRIFFITHS-JONES, S. *et al.* mirbase: tools for microRNA genomics. *Nucleic Acids Res*, v. 36, n. Database issue, p. D154–8, jan. 2008. ISSN 1362-4962.
- HAN, J. *et al.* The drosha-dgcr8 complex in primary microRNA processing. *Genes Dev*, v. 18, n. 24, p. 3016–27, dez. 2004. ISSN 0890-9369.
- HERRERA, L. *et al.* Mouse ovary developmental rna and protein markers from gene expression profiling. *Dev Biol*, v. 279, n. 2, p. 271–90, mar. 2005. ISSN 0012-1606.
- HOU, Y.-Y.; YING, X.-M.; LI, W.-J. [computational approaches to microRNA discovery]. *Yi Chuan*, v. 30, n. 6, p. 687–96, jun. 2008. ISSN 0253-9772.
- HSU, P. W. C. *et al.* mirnamap: genomic maps of microRNA genes and their target genes in mammalian genomes. *Nucleic Acids Res*, v. 34, n. Database issue, p. D135–9, jan. 2006. ISSN 1362-4962.
- HSU, S.-D. *et al.* mirnamap 2.0: genomic maps of micrnas in metazoan genomes. *Nucleic Acids Res*, v. 36, n. Database issue, p. D165–9, jan. 2008. ISSN 1362-4962.
- KIM, V. N. MicroRNA precursors in motion: exportin-5 mediates their nuclear export. *Trends Cell Biol*, v. 14, n. 4, p. 156–9, abr. 2004. ISSN 0962-8924.

- KIRIAKIDOU, M. *et al.* A combined computational-experimental approach predicts human microRNA targets. *Genes Dev*, v. 18, n. 10, p. 1165–78, maio 2004. ISSN 0890-9369.
- KREK, A. *et al.* Combinatorial microRNA target predictions. *Nat Genet*, v. 37, n. 5, p. 495–500, maio 2005. ISSN 1061-4036.
- KULSHRESHTHA, R. *et al.* A microRNA signature of hypoxia. *Mol Cell Biol*, v. 27, n. 5, p. 1859–67, mar. 2007. ISSN 0270-7306.
- LAGOS-QUINTANA, M. *et al.* Identification of novel genes coding for small expressed RNAs. *Science*, v. 294, n. 5543, p. 853–8, out. 2001. ISSN 0036-8075.
- LAI, E. C. *et al.* Computational identification of drosophila microRNA genes. *Genome Biol*, v. 4, n. 7, p. R42, 2003. ISSN 1465-6914.
- LEE, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*, v. 23, n. 20, p. 4051–60, out. 2004. ISSN 0261-4189.
- LI, Y. *et al.* mir-146a suppresses invasion of pancreatic cancer cells. *Cancer Res*, fev. 2010. ISSN 1538-7445.
- LIM, L. P. *et al.* The microRNAs of *Caenorhabditis elegans*. *Genes Dev*, v. 17, n. 8, p. 991–1008, abr. 2003. ISSN 0890-9369.
- MALZKORN, B. *et al.* Identification and functional characterization of microRNAs involved in the malignant progression of gliomas. *Brain Pathol*, ago. 2009. ISSN 1750-3639.
- MARTINELLI, R. *et al.* mir-519d overexpression is associated with human obesity. *Obesity (Silver Spring)*, jan. 2010. ISSN 1930-7381.
- MASRI, S. *et al.* The role of microRNA-128a in regulating TGFβ signaling in letrozole-resistant breast cancer cells. *Breast Cancer Res Treat*, jan. 2010. ISSN 1573-7217.
- MAZIERE, P.; ENRIGHT, A. J. Prediction of microRNA targets. *Drug Discov Today*, v. 12, n. 11-12, p. 452–8, jun. 2007. ISSN 1359-6446.
- MEGRAW, M. *et al.* miRGEN: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res*, v. 35, n. Database issue, p. D149–55, jan. 2007. ISSN 1362-4962.
- MENG, F. *et al.* MicroRNA-21 regulates expression of the p53 tumor suppressor gene in human hepatocellular cancer. *Gastroenterology*, v. 133, n. 2, p. 647–58, ago. 2007. ISSN 0016-5085.
- O'DONNELL, K. A. *et al.* c-myc-regulated microRNAs modulate E2F1 expression. *Nature*, v. 435, n. 7043, p. 839–43, jun. 2005. ISSN 1476-4687.
- PASQUINELLI, A. E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, v. 408, n. 6808, p. 86–9, nov. 2000. ISSN 0028-0836.
- REHMSMEIER, M. *et al.* Fast and effective prediction of microRNA/target duplexes. *RNA*, v. 10, n. 10, p. 1507–17, out. 2004. ISSN 1355-8382.

- RUPERT, J. L. Genomics and environmental hypoxia: what (and how) we can learn from the transcriptome. *High Alt Med Biol*, v. 9, n. 2, p. 115–22, 2008. ISSN 1527-0297.
- SILBER, J. *et al.* mir-124 and mir-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. *BMC Med*, v. 6, p. 14, 2008. ISSN 1741-7015.
- STARK, A. *et al.* Identification of drosophila microRNA targets. *PLoS Biol*, v. 1, n. 3, p. E60, dez. 2003. ISSN 1545-7885.
- SUN, Y. *et al.* mir-126 inhibits non-small cell lung cancer cells proliferation by targeting egfl7. *Biochem Biophys Res Commun*, dez. 2009. ISSN 1090-2104.
- TAM, W.; DAHLBERG, J. E. mir-155/bic as an oncogenic microRNA. *Genes Chromosomes Cancer*, v. 45, n. 2, p. 211–2, fev. 2006. ISSN 1045-2257.
- TANG, G. sirna and mirna: an insight into riscs. *Trends Biochem Sci*, v. 30, n. 2, p. 106–14, fev. 2005. ISSN 0968-0004.
- WANG, W.-X. *et al.* The expression of microRNA mir-107 decreases early in alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci*, v. 28, n. 5, p. 1213–23, jan. 2008. ISSN 1529-2401.
- WINTER, J. *et al.* Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol*, v. 11, n. 3, p. 228–34, mar. 2009. ISSN 1476-4679.
- YAO, J. *et al.* MicroRNA-30d promotes tumor invasion and metastasis by targeting galphai2 in hepatocellular carcinoma. *Hepatology*, nov. 2009. ISSN 1527-3350.
- YEO, M. *et al.* A novel rna polymerase ii c-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem*, v. 278, n. 28, p. 26078–85, jul. 2003. ISSN 0021-9258.
- ZHANG, Z. *et al.* MicroRNAs: Potential regulators involved in human anencephaly. *Int J Biochem Cell Biol*, v. 42, n. 2, p. 367–374, fev. 2010. ISSN 1878-5875.

*APÊNDICE A – Artigo que descreve a
plataforma miRNPath*

miRNPath: a database of miRNAs, target genes and metabolic pathways

A.O. Chiromatzo¹, T.Y.K. Oliveira¹, G. Pereira¹, A.Y. Costa², C.A.E. Montesco³, D.E. Gras¹,
F. Yonetake⁴, J.B. Vilar⁵, M. Cervato⁴, P.R.R. Prado¹, R.G.C.C.L. Cardenas⁴, R. Cerri⁶,
R.L. Borges⁴, R.N. Lemos⁴, S.M. Alvarenga⁷, V.R.C. Perallis⁶, D.G. Pinheiro¹, I.T. Silva¹,
R.M. Brandão¹, M.A.V. Cunha¹, S. Giuliatti¹ and W.A. Silva Jr¹

¹Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

²Departamento de Bioinformática, Universidade de Ribeirão Preto, Ribeirão Preto, SP, Brasil

³Departamento de Física e Matemática, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

⁴Departamento de Informática Biomédica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

⁵Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil

⁶Departamento de Estatística, Matemática Aplicada e Computação, Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, SP, Brasil

⁷Instituto de Biotecnologia Aplicada à Agricultura, Universidade Federal de Viçosa, Viçosa, MG, Brasil

Corresponding author: A.O. Chiromatzo

E-mail: alynne@lgmb.fmrp.usp.br

Genet. Mol. Res. 6 (4): 859-865 (2007)

Received August 03, 2007

Accepted September 25, 2007

Published October 05, 2007

ABSTRACT. MicroRNAs (miRNAs) are small non-coding RNAs that regulate target gene expression and hence play important roles in metabolic pathways. Recent studies have evidenced the interrelation of miRNAs with cell proliferation, differentiation, development, and diseases. Since they are involved in gene regulation, they are intrinsically related to metabolic pathways. This leads to questions that are particularly interesting for investigating medical and laboratorial applications. We developed an miRNApath online database that uses miRNA target genes to link miRNAs to metabolic pathways. Currently, databases about miRNA target genes (DIANA miRGen), genomic maps (miRNAMap) and sequences (miRBase) do not provide such correlations. Additionally, miRNApath offers five search services and a download area. For each search, there is a specific type of input, which can be a list of target genes, miRNAs, or metabolic pathways, which results in different views, depending upon the input data, concerning relationships between the target genes, miRNAs and metabolic pathways. There are also internal links that lead to a deeper analysis and cross-links to other databases with more detailed information. miRNApath is being continually updated and is available at <http://lgmb.fmrp.usp.br/mirnapath>.

Key words: MicroRNA, Metabolic pathways, Target genes, Database, Interaction, Investigation

INTRODUCTION

MicroRNAs (miRNAs) are a class of small, single-stranded non-coding RNAs about ~22 nt in length that regulate gene expression at the post-transcriptional level. Biogenesis starts in the nucleus, where the Drosha enzyme extracts from a long transcript, called primary miRNA (pri-miRNA), with one or more miRNA precursors (pre-miRNA) of ~70 nt. Each of these pre-miRNA is then exported to the cytoplasm, where they are processed by the Dicer enzyme, usually producing a mature miRNA (Bartel, 2004; Kim, 2005; Kim and Nam, 2006).

The mature miRNA enters into an RNA-induced silencing complex and then controls gene expression by cleavage, if they have sufficient complementarity with the mRNA, or by translation repression, if suitable complementary sites are present. Recent research has shown that miRNAs play key roles in many regulatory pathways, such as control of developmental timing, cell proliferation, differentiation, and apoptosis (Wu and Belasco, 2005; Hurst, 2006). Consequently, their deregulation is involved in a series of human diseases, including leukemia, cardiac hypertrophy, and Tourette's syndrome (Di Leva et al., 2006; Cheng et al., 2007).

Due to the complexity of genomic and genetic information, bioinformatics has been widely used for the identification of possible drug targets (Jiang and Zhou, 2005; Ricke et al., 2006). A common approach used to determine the pathway signaling components is to interfere in gene expression. Since miRNAs are important regulators, they can be used for the investigation of potential drugs.

Although it is important to know the influence of miRNA on metabolic pathways, there is no information about their interrelation (Hsu et al., 2006; Shahi et al., 2006; Sethupathy et al., 2006). The Sanger Institute elaborated the miRBase database, which provides information about the genomic coordinates, predicted precursor sequence, mature sequence, and target gene prediction of the mature miRNAs (John et al., 2004; Griffiths-Jones et al., 2006). Other databases, such as miRNAMap, offer new predicted targets (Hsu et al., 2006) and miRGen centralizes genomic and target gene information (Megraw et al., 2007).

The main purpose of the miRNApath online database is to provide information about the influence of miRNA on metabolic pathways. This kind of study is made possible by using one of the six web services made available. The search result includes cross-links to other databases (miRBase, KEGG and Entrez) for detailed information about the gene, metabolic pathway and miRNA. The links within miRNApath include further information about the search.

MATERIAL AND METHODS

The miRNApath database is built based on well-known databases. Its main stored information is the miRNA data from the miRBase ftp site (<http://microrna.sanger.ac.uk/sequences/>), the target genes from the miRGen assembly (<http://www.diana.pcbi.upenn.edu/miRGen.html>) and the metabolic pathways network obtained from the KEGG database, using the SOAP/WSDL interface (<http://www.genome.jp/kegg/soap/>).

Nowadays, the most important institutions involved are the National Center for Biological Information (NCBI) and the University of California at Santa Cruz. Usually, each institution establishes its own code(s) to identify and retrieve information from their databases. Since miRNApath works with other databases, there is a problem with data denomination heterogeneity. Hence, miRNApath adopted the gene symbol, the miRNA name, the metabolic pathway name, and the NCBI access number (GenBank database), as well as geneid (Refseq database) and unigene (Unigene database) to represent the data. miRNApath handles only those organisms that are in all databases. The selected organisms are *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, and *Canis familiaris*.

Integration of all database information is made using the Application Programming Interfaces supplied by the institutions in Perl language to parse the downloaded data. This is a meticulous task that demands caution in order to maintain miRNApath consistency. In miRNApath, all the target genes from the miRGen and KEGG metabolic network genes must have equivalent code identifications from NCBI databases. Also, all the miRNAs of miRGen must be registered in the miRBase. The miRNApath is stored in a Mysql database and the web interface was developed using PHP language.

RESULTS

The miRNApath offers an online interface with six web services separated into the categories: QUICK SEARCH, which returns superficial information and ADVANCED ANALYSIS, which presents a detailed result. Figure 1A illustrates the main page. QUICK SEARCH has three options: Search by Gene, Search by miRNA and Search by Pathways. The difference between them is the object of study; for example, if the user wants to analyze a target gene list, he must click the link "Search by Gene", and so on. Figure 1B shows the Search by Gene site. To make a search, the user has to write a target gene list or upload a file, specify which organism and format what the input refers to and then press the search button to submit the data. The figures below show an example of how to make a search and how the result will appear. The other searches are analogous.

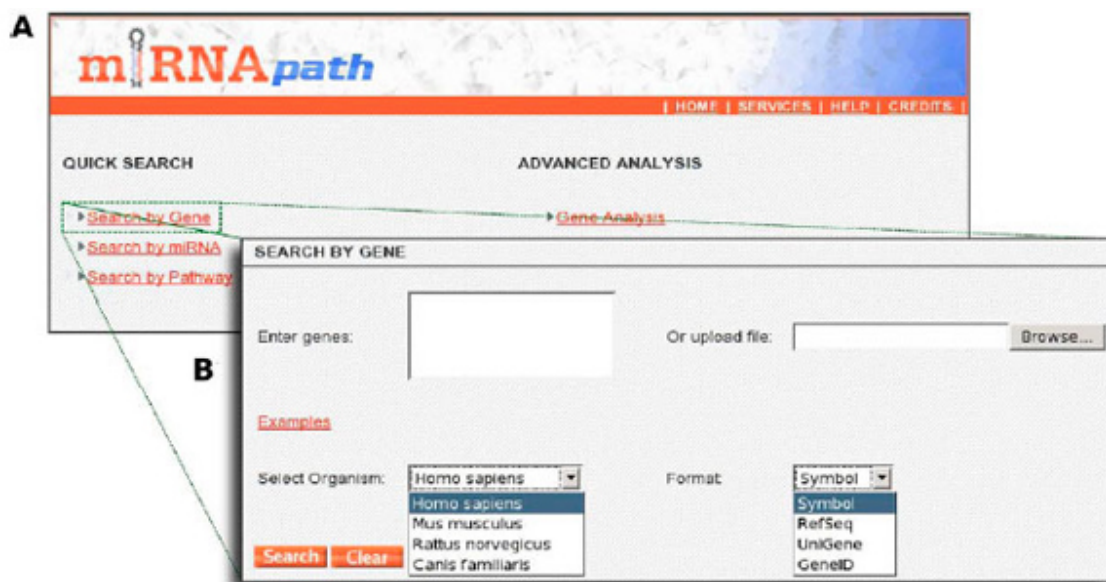


Figure 1. A. Main page. The green box links to the Search by Gene site. B. Search by Gene input page. The user can write a list of target genes or upload a file and then select the corresponding organism and input data type.

The Search by Gene result is given in Figure 2A. It shows, for each target

gene entered, the quantity of miRNAs that target the gene and the number of pathways in which the gene is involved. The orange links lead the user to internal sites with detailed information, and the blue links direct the user to outside databases. If the user clicks on miRNAs amount (blue box), it will open another site (Figure 2B) with target gene AXIN2 information, including all the 17 miRNAs that affect this gene and the number of pathways for which each miRNA has a target gene. If the user clicks on the pathways amount (green box), he will indicate a site (Figure 2C) with information about the AXIN2 target gene, the hsa-miR-449 miRNA and the name of all the 94 pathways in which the miRNA has a target gene. The links in dark red mean that the gene AXIN2 is regulated by the hsa-miR-449 miRNA and it participates in the Wnt signaling pathway.

Finally, we present an example of the ADVANCED ANALYSIS. The gene analysis search and the gene search have the same input interface, but with different results. Figure 3A shows the search result site, which has three ways of displaying data that are ordered by pathway, gene and miRNA. Figure 3B shows the order by pathway view. It displays the pathways in which the target genes of the input list are involved, the number of target genes from the input list that are in the pathway and the number of miRNAs that affect these target genes. In the example, the MAPK signaling pathway has two genes from the input target list and there are 44 miRNAs that regulate those genes. If the user clicks on amount of genes, a site like Figure 3C will appear. This site has information about the pathway, the genes from the input list that belong to the pathway and the respective miRNAs that repress the genes.

A

miRNApath | HOME | SERVICES | HELP | CREDITS

Page 1 of 1 | Display 10 registers by page

Registers 1 to 4 of 4

| Symbol | Organism | Location | Description | miRNAs | Pathways |
|-----------------------|--------------|----------|--------------------------|--------|----------|
| AXIN2 | Homo Sapiens | 17q23-q | axin 2 (conductin, axil) | 17 | 1 |

B

TARGET GENE (AXIN2)

| Symbol | Organism | Type | Chromosome | Location | Description |
|-----------------------|--------------|----------------|------------|----------|--------------------------|
| AXIN2 | Homo Sapiens | protein-coding | 17 | 17q23-q | axin 2 (conductin, axil) |

Page 1 of 2 | Display 10 registers by page

Registers 1 to 10 of 17

| Name | Mature Sequence | Pathways |
|-----------------------------|------------------------|----------|
| HSA-MIR-449 | UGGCAGUGUAUUGUUAGCUGGU | 94 |

C

TARGET GENE (AXIN2)

| Symbol | Organism | Type | Chromosome | Location | Description |
|-----------------------|--------------|----------------|------------|----------|--------------------------|
| AXIN2 | Homo Sapiens | protein-coding | 17 | 17q23-q | axin 2 (conductin, axil) |

MIRNA (HSA-MIR-449)

| Name | Mature Sequence |
|-----------------------------|------------------------|
| HSA-MIR-449 | UGGCAGUGUAUUGUUAGCUGGU |

Page 1 of 10 | Display 10 registers by page

Registers 1 to 10 of 94

| NAME |
|---|
| Wnt signaling pathway - Homo sapiens (human) |
| Cell cycle - Homo sapiens (human) |
| ABC transporters - General - Homo sapiens (human) |

Figure 2. A. The result of the Search by Gene. B. Detailed site with information about the AXIN2 gene and all the miRNAs that target this gene. C. Information about the target gene AXIN2, the hsa-miR-449 miRNA and all the metabolic pathways in which this miRNA targets a gene. The links in dark red mean that the gene AXIN2 is regulated by the hsa-miR-449 miRNA and it participates in the Wnt signaling pathway.

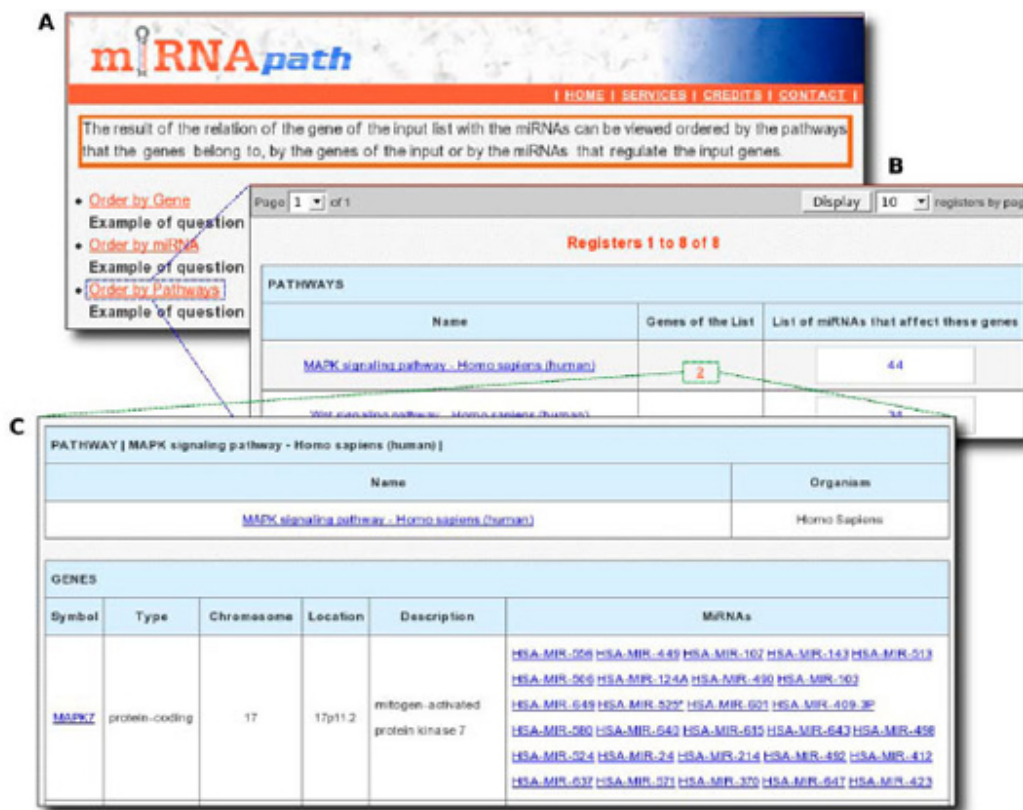


Figure 3. A. Query result with the three viewing options. B. Order by pathway view shows the quantity of target genes from the input list and the quantity of miRNAs that affect these genes. C. Information about the genes of the list that belong to the selected pathway.

DISCUSSION

The miRNApath handles many questions about the relationship between miRNAs, target genes and metabolic pathways. Investigating miRNAs, one can discover the pathways affected by these miRNAs and possibly determine a disease caused by them. Or, by analyzing a list of target genes, it is possible to find an miRNA (or a set of them), which when deregulated could be the cause of a certain symptom. In conclusion, miRNApath provides an easy and friendly means of investigating a large number of questions about this relationship. Soon, we intend to allow the user to download the results from the searches, to implement various searches, to add new input formats, and to provide other useful information.

ACKNOWLEDGMENTS

Research supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento Pessoal de Nível

Superior (CAPES), Financiadora de Estudos e Projetos (FINEP), and Fundação de Apoio ao Ensino, Pesquisa e Assistência do Hospital das Clínicas FMRP-USP (FAEPA).

REFERENCES

- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.
- Cheng Y, Ji R, Yue J, Yang J, et al. (2007). MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy? *Am. J. Pathol.* 170: 1831-1840.
- Di Leva G, Calin GA and Croce CM (2006). MicroRNAs: fundamental facts and involvement in human diseases. *Birth Defects Res. C. Embryo Today* 78: 180-189.
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, et al. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34: D140-D144.
- Hsu PW, Huang HD, Hsu SD, Lin LZ, et al. (2006). miRNAMap: genomic maps of microRNA genes and their target genes in mammalian genomes. *Nucleic Acids Res.* 34: D135-D139.
- Hurst LD (2006). Preliminary assessment of the impact of microRNA-mediated regulation on coding sequence evolution in mammals. *J. Mol. Evol.* 63: 174-182.
- Jiang Z and Zhou Y (2005). Using bioinformatics for drug target identification from the genome. *Am. J. Pharmacogenomics* 5: 387-396.
- John B, Enright AJ, Aravin A, Tuschl T, et al. (2004). Human MicroRNA targets. *PLoS. Biol.* 2: e363.
- Kim VN (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* 6: 376-385.
- Kim VN and Nam JW (2006). Genomics of microRNA. *Trends Genet.* 22: 165-173.
- Megraw M, Sethupathy P, Corda B and Hatzigeorgiou AG (2007). miRGen: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res.* 35: D149-D155.
- Ricke DO, Wang S, Cai R and Cohen D (2006). Genomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* 10: 303-308.
- Sethupathy P, Corda B and Hatzigeorgiou AG (2006). TarBase: a comprehensive database of experimentally supported animal microRNA targets. *RNA* 12: 192-197.
- Shahi P, Loukianiouk S, Bohne-Lang A, Kenzelmann M, et al. (2006). Argonaute - a database for gene regulation by mammalian microRNAs. *Nucleic Acids Res.* 34: D115-D118.
- Wu L and Belasco JG (2005). Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. *Mol. Cell Biol.* 25: 9198-9208.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)