



Instituto de
Biociências



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de Botucatu



PG-BGA

UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

Abordagem in-sílico para seleção de miRNAs
relacionados a carne de Tilápia para fins de melhoramento genético.

Ravi Augustus Nogueira

Orientador: Prof. Dr. Ney Lemke

Coorientador: Luiz Augusto Bovolenta

Dissertação apresentada ao Instituto de Biociências,
Campus de Botucatu, UNESP, para obtenção do título de
Mestre no Programa de Pós-Graduação em Biologia
Geral e Aplicada, Área de concentração Biomoléculas.

Prof. Dr. Ney Lemke

**BOTUCATU – SP
2020**

Programa de Pós-graduação em Biologia Geral e Aplicada
Instituto de Biociências de Botucatu/ UNESP
Rua Professor Doutor Antonio Celso Wagner Zanin, 250 - CEP 18618-689 - Botucatu - SP - Brasil
Tel (14) 3880-0781 posgraduacao@ibb.unesp.br

N778a Nogueira, Ravi Augustus
Abordagem in-silico para seleção de miRNAs
relacionados a carne de Tilápia para fins de melhoramento
genético. / Ravi Augustus Nogueira. -- Botucatu, 2020
31 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista
(Unesp), Instituto de Biociências, Botucatu
Orientador: Ney Lemke

1. Micro RNA. 2. Oreochromis niloticus. 3.
Transcriptomes. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do
Instituto de Biociências, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Dedico às pessoas que me deram oportunidade para chegar até aqui.

Agradecimentos

Agradeço ao meu orientador, Prof. Dr. Ney Lemke por me dar a oportunidade de ingressar no curso de mestrado, e por ser uma pessoa compreensiva.

Ao meu Coorientador, Luiz Augusto Bovolenta por ceder dados de seu trabalho, me possibilitando realizar meu projeto.

Ao Dr. Guilherme Targino Valente por ceder o espaço de seu laboratório para realização de minha pesquisa e também ao seu tempo dedicado dando suporte a meu trabalho de mestrado.

Ao Prof. Dr. José Maurício Sforcin e a minha amiga Eliza de Oliveira Cardoso por me permitirem ingressar no curso de mestrado.

À Agência de fomento CAPES, cuja bolsa nos permite seguir em frente na pesquisa.

À equipe Pós-graduação da Unidade da UNESP, campus de Botucatu, pelos serviços prestados à comunidade científica.

A todos os meus colegas.

Ao meu Pai, Edgar José Nogueira.

A meu tio, Rivaldo Nogueira.

E a minha falecida mãe, Sônia da Paixão Gonçalves.

Índice

Resumo	6
Abstract	8
Introdução	9
Objetivos	15
Materiais e Métodos	16
Resultados e Discussão	21
Material Suplementar	29
Referências Bibliográficas	30

Índice de Tabelas e Figuras

Tabela 1	10
Figura 1	11
Figura 2	12
Figura 3	13
Figura 4	14
Figura 5	14
Figura 6	17
Figura 7	17
Figura 8	18
Figura 9	18
Figura 10	19
Figura 11	19
Tabela 2	22
Tabela 3	22
Tabela 4	23
Tabela 5	24
Tabela 6	25
Figura 12	26

Resumo

O melhoramento genético proporciona aumento na produtividade em diversas áreas, devido a seleção convencional e a organismos geneticamente modificados (OGMs). A maioria dos animais melhorados e linhagens de peixes “melhoradas” tem origem por “seleção convencional” o que é mais dispendioso, tanto em tempo como financeiro. Usando ferramentas de bioinformática (gprofiler, Weka, R entre outros) e banco de dados (Ensembl, KEGG, NCBI), foram realizadas análises de dados de transcriptoma provenientes da tilápia do nilo (*Oreochromis niloticus*), usando o sistema operacional Linux. Dentre mais de 15.000 miRNAs provenientes dos dados analisados, foi selecionado um miRNA, sendo ele o miR-214-3p atuante em 3 vias metabólicas de interesse no presente trabalho, ao mesmo tempo: mTOR/Akt, via Hippo e Ubiquitin Ligase todas relacionadas ao crescimento e desenvolvimento muscular.

Abstract

The genetic improvement due to conventional selection and genetically modified organisms (OGMs) generates an increase in productivity on several areas. Most of the “improved” animals and fish lineages originate from “conventional selection”, which is costlier: both in time and financially. Using tools from Bioinformatics (g:Profiler, Weka, R, amongst others) and various data banks (Ensembl, KEGG, NCBI), data analyses of transcriptome from Nile tilapia (*Oreochromis niloticus*) using the Linux operational system. Amongst more than 15,000 miRNAs a miRNA was selected, being the miR-214-3p active being active on 3 pathways of interest in the present work: mTOR/Akt, Hippo and Ubiquitin Ligase, all related to muscle development and growth.

Introdução

A tilápia do Nilo (*Oreochromis niloticus*) é a espécie de tilapina mais comumente produzidas no mundo (CORREA et al., 2019). Em 2008, a cultura da tilápia do Nilo ficou em quinto lugar entre as espécies mais cultivadas do mundo, com uma produção total 2,3 milhões de toneladas (EKNATH; HULATA, 2009). A China é o maior produtor e consumidor (cerca de 53% da produção global) de tilápia, demonstrando a grande importância econômica da espécie, com uma produção estimada em 3,15 milhão de toneladas em 2016 (LAGO et al., 2017).

O melhoramento genético é uma área de pesquisa com seus próprios métodos e procedimentos utilizados a fins de modificar as frequências de alelos desejáveis em uma população animal ou vegetal (MITRA, 2001). O melhoramento genético é praticado pelo homem, e iniciou-se simultaneamente à fase de domesticação das plantas, há cerca de 9.000 anos atrás (MITRA, 2001). A domesticação e o conseqüente cultivo de um grande número de indivíduos de uma espécie permitiu a aproximação de muitas espécies selvagens, que passaram a trocar genes com as plantas cultivadas, resultando daí uma variação muito grande nos descendentes, devido aos cruzamentos espontâneos que então ocorriam (MITRA, 2001). Tem sua importância em diversas áreas, uma de suas vertentes são os estudos relacionados ao aumento de produtividade (FORABOSCO et al., 2013), que tem grande impacto econômico.

Com o decorrer do tempo passou-se de técnicas de cruzamento (inter e intraespecífica), para também técnicas laboratoriais, onde se possibilitou maior compreensão do que estava ocorrendo de fato na maquinaria biológica do organismo que influenciava e/ou gerava as características desejadas (FORABOSCO et al., 2013). Usando a técnica de microinjeção, (introdução de material genético exógeno em uma célula somática, ovócito ou embrião, via micropipeta de vidro), Hammer et al. (1985), iniciaram-se trabalhos relacionados a inserção de genes do Hormônio de crescimento (GH) em porcos, aumentando sua taxa de crescimento, o volume muscular, a porcentagem de carne magra e reduzindo a quantidade de ração utilizada por kg de carne produzida. A atuação de melhoristas em relação a moléculas, no princípio visava a produção de certas proteínas como produto direto do gene inserido (FORABOSCO et al., 2013), sendo a principal delas as proteínas mostradas na Tabela 1.

Species	Target trait	Transgenic gene	Origin of the gene	Reference
Swine	Growth	GH cSKI	Human Chicken	Nottle et al. (1999) Pursel et al. (1992)
	Milk	α -Lactalbumin Bacteria resistance	Bovine Mouse	Bleck et al. (1998), Wheeler et al. (2001) Lo et al. (1991)
	Health	Lysozyme Unsat. fatty acid Omega-3	Human Spinach Nematode	Tong et al. (2011) Saeki et al. (2004) Lai et al. (2006)
	Feed eff.	Phytase	Bacterium	Forsberg et al. (2006), Phillips et al. (2006)
Sheep	Growth	IGF-1	Ovine	Damak et al. (1996)
	Health	Visna virus resistance BSE Bacteria resistance	Virus Deletion Mouse	Clements et al. (1994) Denning et al. (2001) Lo et al. (1991)
Goat	Health	Lysozyme	Bovine	Scharfen et al. (2007)
	Udder	Monosat. fatty acid	Bovine	Reh et al. (2004)
Chicken	Health	Leukosis resistance H5N1 resistance lacZ	Virus Virus Bacteria	Salter and Crittenden (1989) Lyll et al. (2011) Mozdziak et al. (2003)
	Growth	GH	Bovine	Chen et al. (1990)
Cattle	Health	Lysozyme BSE	Human Knockout	Yang et al. (2011) Richt et al. (2007)
	Udder	Omega-3 α and κ -Casein Antimicrobial	Nematode Bovine Bacteria	Wu et al. (2012) Brophy et al. (2003) Wall et al. (2005)
	Growth	GH	Bovine	Chen et al. (1990)
Fish	Health	Cecropin Lactoferrin Lysozyme	Insect Human Fish	Dunham et al. (2002) Mao et al. (2004) Fletcher et al. (2011)
	Envir. tolerance	Antifreeze GH	Fish Fish	Hew et al. (1999) Saunders et al. (1998)

Tabela 1. Visão geral dos genes em animais e peixes geneticamente modificados na produção de alimentos. Pode-se reparar que, em organismos geneticamente modificados a intenção da maior parte dos pesquisadores é a inserção de genes, com o intuito da produção de certas proteínas específicas. Adaptado de (FORABOSCO et al., 2013).

No caso das galinhas, os principais objetivos do melhorista visava a resistência a doenças. Foram produzidas galinhas transgênicas com cópias de genes do vírus da leucose aviária, onde as mesmas mostraram resistência à leucose e uma maior resistência a infecções e oncógenes (SALTER; CRITTENDEN, 1989). Trabalhos mais recentes atuam em outras modificações, não sendo somente focadas em produção direta de proteínas. A inserção de plasmídeos, visando a produção de shRNA (short-hairpin RNA), moléculas de RNA de interferência, que atuam silenciando genes específicos, sendo um exemplo as galinhas geneticamente modificadas, garantindo maior resistência a H5N1 (gripe do frango) (LYALL et al., 2011).

Os miRNAs, são um subconjunto de RNAs não codificadores de proteínas, de aproximadamente 22-nt de comprimento com produção endógena, tendo como característica principal a regulação pós-transcricional de moléculas de mRNA por clivagem ou ligação na mesma inibindo-a física e quimicamente, a tradução em proteínas (BARTEL, 2004). Os miRNAs possuem uma região conservada entre diversos organismos, conhecida por "região seed", possui entre 6-8 nt localizada na região 5', tendo papel importante no reconhecimento dos mRNAs alvos se ligando, principalmente na região 3'UTR do mRNA (DENG et al., 2008).

A região “seed” possui pouca especificidade em relação aos seus mRNAs alvos, devida a sua pequena região de reconhecimento (6-8 nt), com isso um único miRNA pode regular diversos genes (GREGORY; SHIEKHATTAR, 2005).

O processamento e formação do miRNA maduro, começa no núcleo com a transcrição em uma molécula de pri-miRNA (~80 nt) pela RNA-polimerase II, formando uma estrutura em dsRNA em “steen loop”, que posteriormente é reconhecida pela proteína nuclear DGCR8/“Pasha” em invertebrados e associados a proteína Drosha (RNase III), que cliva o pri-miRNA formando o complexo Microprocessador, e o pré-miRNA (CAI; HAGEDORN; CULLEN, 2004) (Figura 1).

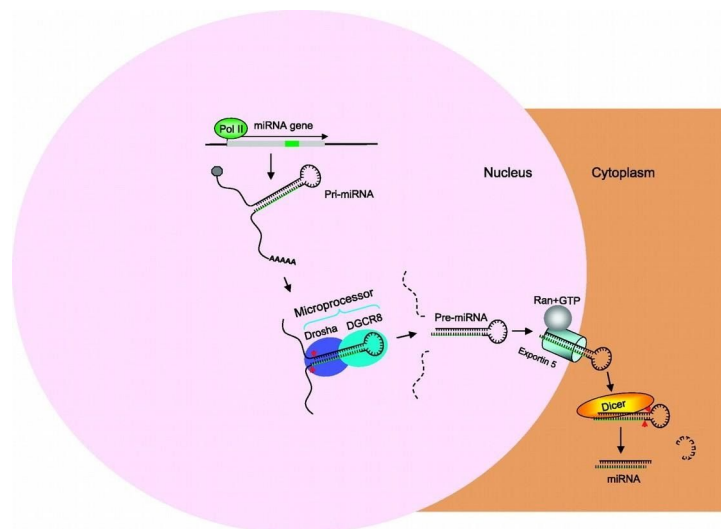


Figura 1. Processamento e Biogênese de miRNA maduro, adaptado de (GREGORY; SHIEKHATTAR, 2005).

Após ser exportado para o citoplasma pela exportina-5, o pré-miRNA é clivado pela RNase III (Dicer) tirando o loop e formando duas fitas conhecidas como miRNA/miRNA* ou braço 3'/5' incorporado pelo complexo “RISC” (RNA-induced silencing complex) onde há uma seleção de um dos “braços” (3’ ou 5’) (GREGORY; CHENDRIMADA; SHIEKHATTAR, 2006).

A identificação de genes diferencialmente expressos em condições específicas é de grande interesse para o melhoramento, sendo a RNA-seq (tecnologia de sequenciamento de nova geração), uma poderosa ferramenta para esta análise (COSTA-SILVA; DOMINGUES; LOPES, 2017).

Com a grande quantidade de dados gerados pela técnica, é imprescindível o uso da Bioinformática para a interpretação dos mesmos, conferindo sentido biológico e uma maior compreensão para possíveis aplicações dos resultados das análises in-silico, em experimentos in-vivo.

Vias metabólicas relacionadas a anabolismo e catabolismo muscular.

As vias metabólicas possuem alta sinergia, tendo diversos compostos intermediários, enzimas e miRNAs se conectando. Com a alteração na proporção de moléculas que participam das vias, há um deslocamento em sua função tendo como exemplo, vias de catabolismo e anabolismo muscular (ZERHOUNI, 2004).

Foram selecionadas 3 vias de importância ao crescimento (hipertrofia e hiperplasia) e diferenciação muscular: (Akt/mTOR, via Hippo e via ubiquitina ligase específica muscular) onde foram selecionadas as seguintes proteínas (TSC1/2, Murf1, TRAF6, Set7, Mst1/2, Sav1, Mob, Lats1/2), essas tendo importância na regulação do crescimento muscular.

A via ubiquitina-proteassoma está envolvida na regulação da massa muscular (SISHI et al., 2011). A via contribui para a degradação de proteínas, pois as proteínas direcionadas para a degradação são substratos que podem ser identificados e ligados à ubiquitina, que posteriormente podem ser substratos para serem poli-ubiquitinados e então direcionados à degradação por proteassomas (Figura 2) (ATTAIX et al., 2008). A ativação da via ubiquitina-proteassoma pode ter um papel importante na mediação da atrofia do músculo esquelético (CHAI; WU; SHENG, 2003).

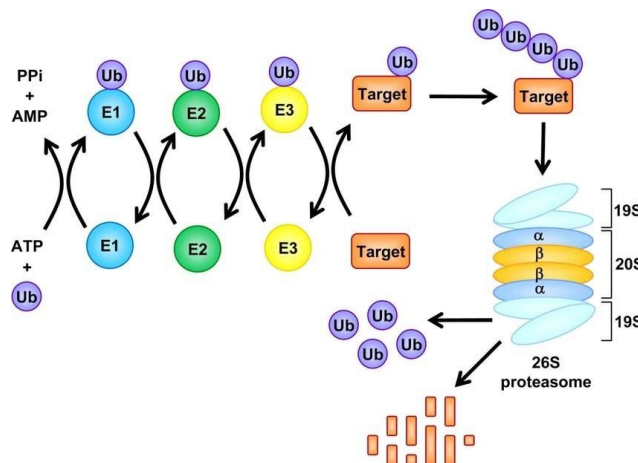


Figura 2. A ubiquitina (Ub) é ativada pelas enzimas ativadoras da Ub (E1) através da adenilação e formação de ligação éster tiol de alta energia e depois transferida para Ub (E2) enzimas, (E3) então transfere a Ub para um reativo resíduo de lisina de uma proteína alvo. As proteínas ubiquitinadas são reconhecidas pela 19S do proteassoma 26S e alimentado em seu núcleo catalítico 20S para degradação em oligopeptídeos. A Ub é então liberada e reciclada.

Durante a atrofia muscular, MuRF1, uma ubiquitina ligase específica de músculo, é superexpressada no músculo, e inibindo a função de MuRF1 demonstrou-se que suprime a perda muscular e atenua a atrofia muscular (EDDINS et al., 2011).

A TRAF6 é altamente expressado no desenvolvimento do músculo esquelético de animais jovens, mas seus níveis são reduzidos em animais adultos (PAUL et al., 2010).

TRAF6 promove a ubiquitinação de Beclin-1, que é fundamental para a indução da autofagia, em resposta à sinalização do receptor Toll-like 4 (SHI; KEHRL, 2010). Como muitas dessas vias de sinalização e sistemas proteolíticos, ativados por mecanismos dependentes do TRAF6, também são ativados na atrofia do músculo esquelético. Foi hipotetizado que o TRAF6 desempenhe um papel crítico na regulação da massa muscular esquelética em diferentes condições catabólicas (SHI; KEHRL, 2010).

A via Akt/mTOR (Figura 3), está intimamente ligada ao crescimento celular, a proliferação, motilidade e sobrevivência, bem como a transcrição de genes e síntese de proteínas que são ativadas em resposta a hormônios, fatores de crescimento e nutrientes (MEMMOTT; DENNIS, 2009).

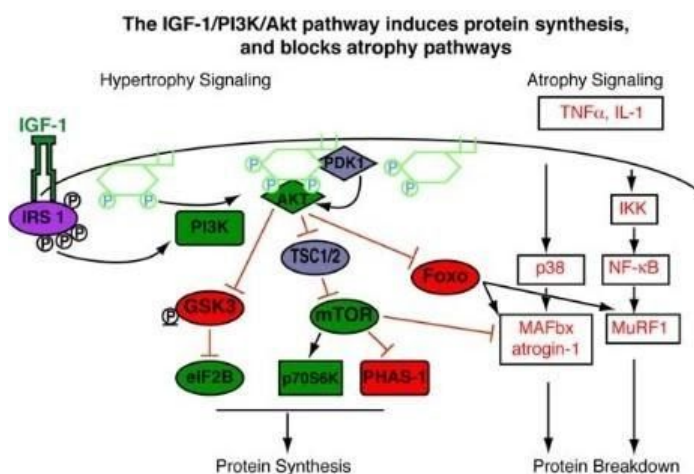


Figura 3. As moléculas com efeito negativo na hipertrofia (vermelho), as proteínas que induzem a hipertrofia (verde). GSK3, glicogênio sintase quinase 3 beta; mTOR, alvo de rapamicina em mamíferos; PI3K, fosfatidilinositol-3 quinase. TNFalpha, induz NF-B, a ativação de NF-B é o gatilho, que induz a regulação positiva da transcrição da MuRF1 (E3 ubiquitina-ligase). A ativação de Akt inibe a regulação positiva de MAFbx e MuRF1, através da inibição da família FOXO (DIENSTMANN et al., 2014).

A proteína TSC1 forma um complexo proteico com TSC2 e outras proteínas onde regula negativamente a sinalização da mTORC1 ativando uma pequena GTPase (Rheb) e interferindo com a via catabólica (ZERHOUNI, 2004).

A via Hippo medeia o controle da proliferação celular por diversos mecanismos, que detectam a organização espacial e física das células, como resultado as células nas bordas de uma colônia, uma ferida em um tecido ou um tumor são mais sensíveis aos fatores de crescimento e mais propensos a proliferar, migrar ou diferenciar (GUMBINER; KIM, 2014).

A via possui uma proteína chave conhecida como Yap (yes-associated protein). A proteína Set7 metila a Yap na posição k494, sequestrando a proteína Yap para o citoplasma (OUDHOFF et al., 2013)(figura 3).

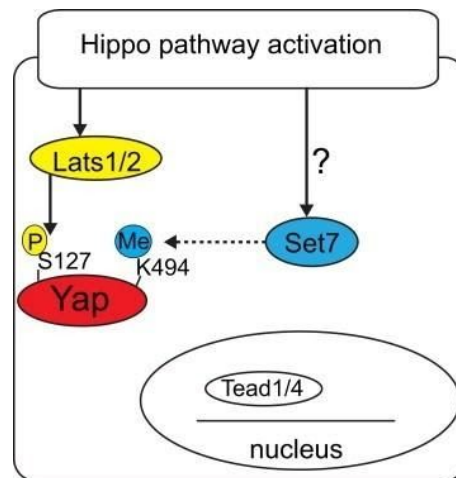


Figura 4. Metilação de Yap via Set7. Adaptado de (OUDHOFF et al., 2013).

O complexo proteico formado por Mst1/2, Sav1, Mob, Lats1/2 promove a fosforilação da proteína Yap, fazendo com que a mesma fique retida no citoplasma e posteriormente degradada, inibindo assim sua atividade (GUMBINER; KIM, 2014)(Figura 4).

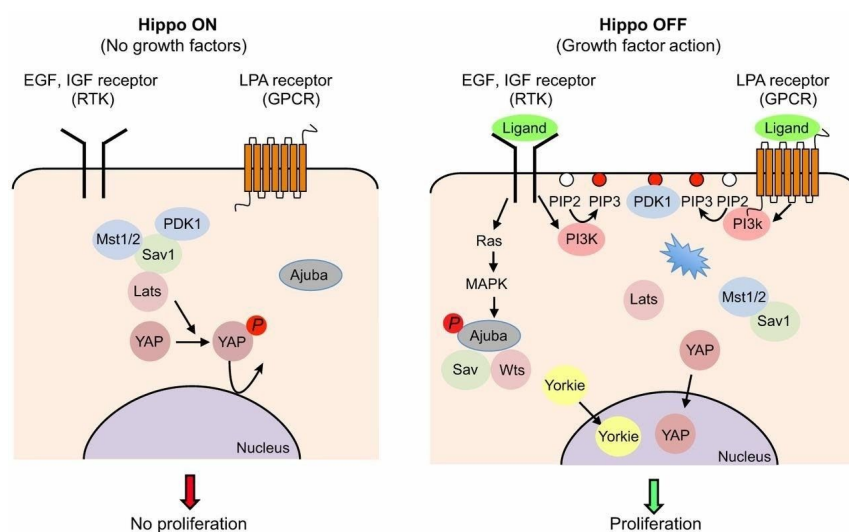


Figura 5. A fosforilação da proteína Yap, pelo complexo proteico Mst1/2, Sav1, Mob (não mostrado na figura), Lats1/2, inibe a proliferação celular Adaptado de (GUMBINER; KIM, 2014).

Objetivos

Analisar os dados proveniente do trabalho de Bovolenta et al., 2016, usando ferramentas de bioinformática para selecionar os miRNAs que interagem com mRNAs produtores das proteínas TSC1/2, Murf1, TRAF6, Set7, Mst1/2, Sav1, Mob e Lats1/2, diretamente ligadas a vias metabólicas relacionadas ao desenvolvimento, crescimento (hipertrofia e hiperplasia) e diferenciação muscular como, mTOR/Akt, via Hippo e Ubiquitina Ligase com objetivo de melhoramento genético da carne de tilápia do Nilo.

Materiais e Métodos

Os dados utilizados para identificar as interações miRNA/mRNA foram obtidos previamente pelo trabalho de (BOVOLENTA, 2016), e estão nomeados com o padrão do Ensembl (BIRNEY et al., 2004), onde foram encontrados os nomes neste formato. Manualmente, uma a uma, das proteínas de interesse para esse estudo sendo elas nomeadas em *Homo sapiens* como: TSC1/2, Murf1, TRAF6, Set7, Mst1/2, Sav1, Mob e Lats1/2 por homologia e conservação de genes ortólogos entre as espécies.

Com o uso do sistema operacional Ubuntu v18.04(Linux), foram utilizados comandos pelo “terminal” (grep, awk, cat, print entre outros), o comando “grep” procura dentro do arquivo pela palavra entre aspas, e o comando “awk” é utilizado para “focalizar” a busca, principalmente de informações em colunas e o comando “print” mostra na tela os resultados em vermelho, com isso utilizado como filtros e buscando informações dentro dos dados provenientes do trabalho de Bovolenta et. al., localizando os mRNAs precursores das seguintes proteínas: TSC1, set7, Mst1/2, Mob, Lats1, os miRNAs que tem como alvo os mRNAs das proteínas acima, os miRNAs que regulam mRNAs de genes constitutivos (esses excluídos pois, tem grande importância para manutenção do organismo, explicado melhor em “Resultados e Discussões”).

Ainda possuindo grande quantidade de dados a serem filtrados, calculou-se medidas de “importância”, usando um scripts da Figura 6, em R, do pacote “Igraph”, sendo essas medidas: Segundo Kolaczyk (2014, p. 44-48) grau (quantidade de ligações/interações entre um vértice e seus vizinhos, sendo os vértices mais “importantes” aqueles que possuem mais número de interações), centralidade de proximidade (medida que dá a noção de que um vértice é “central” se estiver “próximo” de muitos outros vértices), centralidade de intermediação (até que ponto o vértice está localizado “entre” outros pares de vértices, sendo baseadas na perspectiva de que, “importância” refere-se a, onde um vértice está localizado com respeito aos caminhos no gráfico da rede gerada), centralidade de autovetor (Medida da importância de um vértice em função da importância de seus vizinhos (BONACICH, 1987).

```
> A<- get.adjacency (x, sparse=FALSE)
> library (network)
> g <- network: : as.network.matrix(A)
>library (sna)
>sna: :gplot.target (g, degree (g), main="Degree",
+   circ.lab = FALSE, circ.col="skyblue",
+   usearrows = FALSE,
```

```

+ vertex.col=c ("blue", rep ("red", 32), "yellow"),
+ edge.col="darkgray")

```

Figura 6. Onde a variável “x” é o nome do arquivo, e o termo em amarelo a “grandeza” a ser calculada, o script foi usado para calculo das 4 “grandezas”, grau, centralidade de proximidade, centralidade de intermediação e autovetores, imagem adaptada de (NORDHAUSEN, 2015).

Os valores encontrados foram normalizados (LibreOffice Calc), e usados como arquivo de entrada para o programa Weka versão 3.6.14 (HALL et al., 2009), o Weka procede à análise computacional e estatística, recorrendo a técnicas de mineração, onde foi usado o modo de clusterização “Simple Expectation Maximisation”, que é uma generalização natural da estimativa de máxima verossimilhança para o caso de dados incompletos (DO; BATZOGLOU, 2008), gerando 3 clusters e diversos gráficos, sendo utilizados os gráficos de autovetores (eixo y) estático vs os demais valores (centralidade de proximidade, centralidade de intermediação, grau e autovetores) no eixo x (Figura 7, 8 e 9).

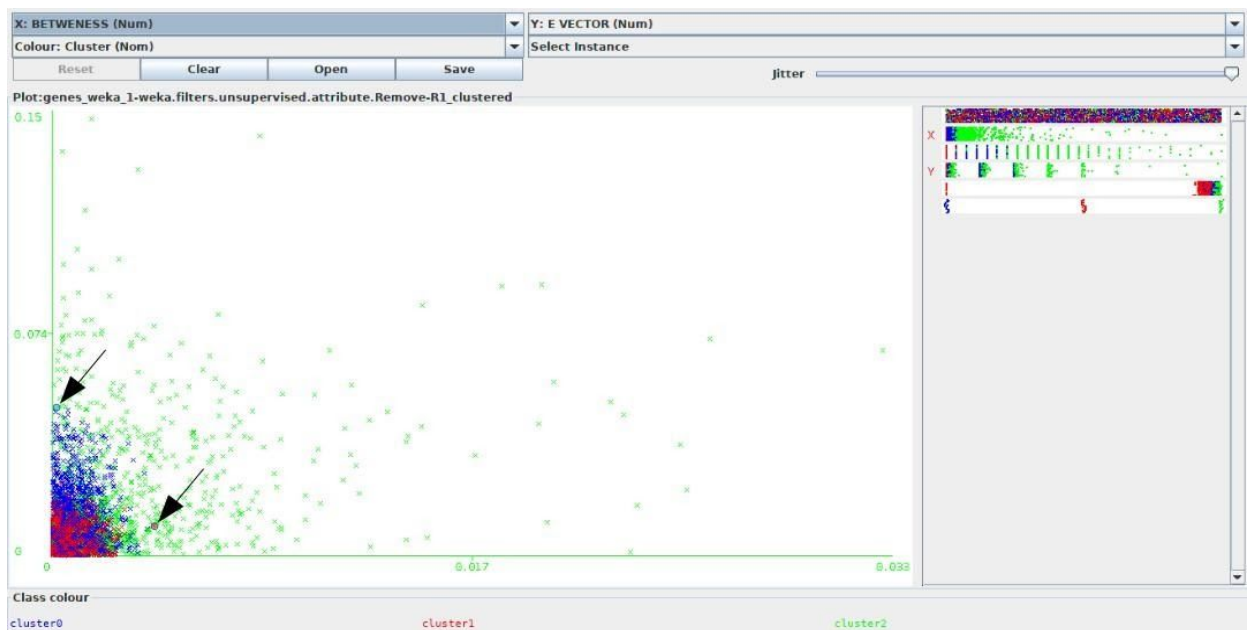


Figura 7. Gráfico de centralidade de intermediação (eixo x) vs autovetores (eixo y), os 3 clusters formados pelo programa Weka, versão 3.6.14 se encontram em cores diferentes sendo o cluster 2 (verde) o grupo de “maior importância”, o cluster 1 (azul) de “importância mediana” e o cluster 0 (vermelho) de “menor importância”. Os locais apontados pelas flechas são os valores máximos de x (ciano) no valor de 0.0000894900565008639 e y (laranja) no valor de 0.0366076. Fonte: Os autores.

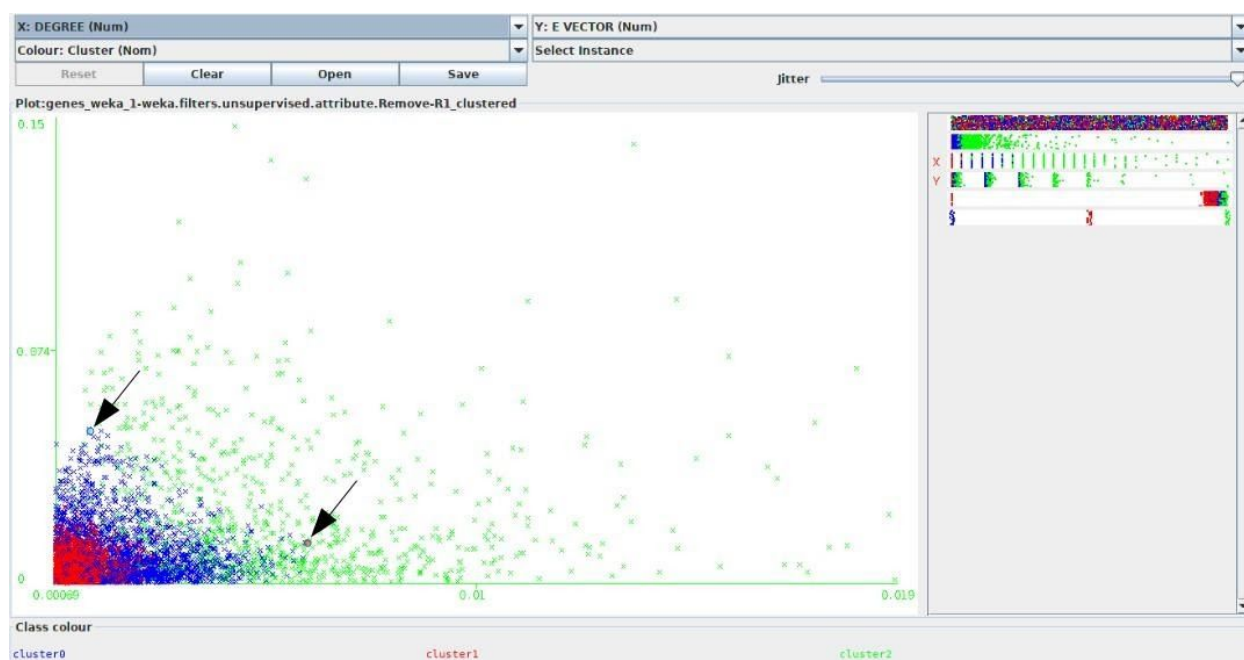


Figura 8. Gráfico de grau (eixo x) vs autovetores (eixo y), os 3 clusteres formados pelo programa Weka, versão 3.6.14 se encontram em cores diferentes sendo o cluster 2 (verde) o grupo de “maior importância”, o cluster 1 (azul) de “importância mediana” e o cluster 0 (vermelho) de “menor importância”. Os locais apontados pelas flechas são os valores máximos de x (ciano) no valor de 0.0048644892 e y (laranja) no valor de 0.0366076. Fonte: Os autores.

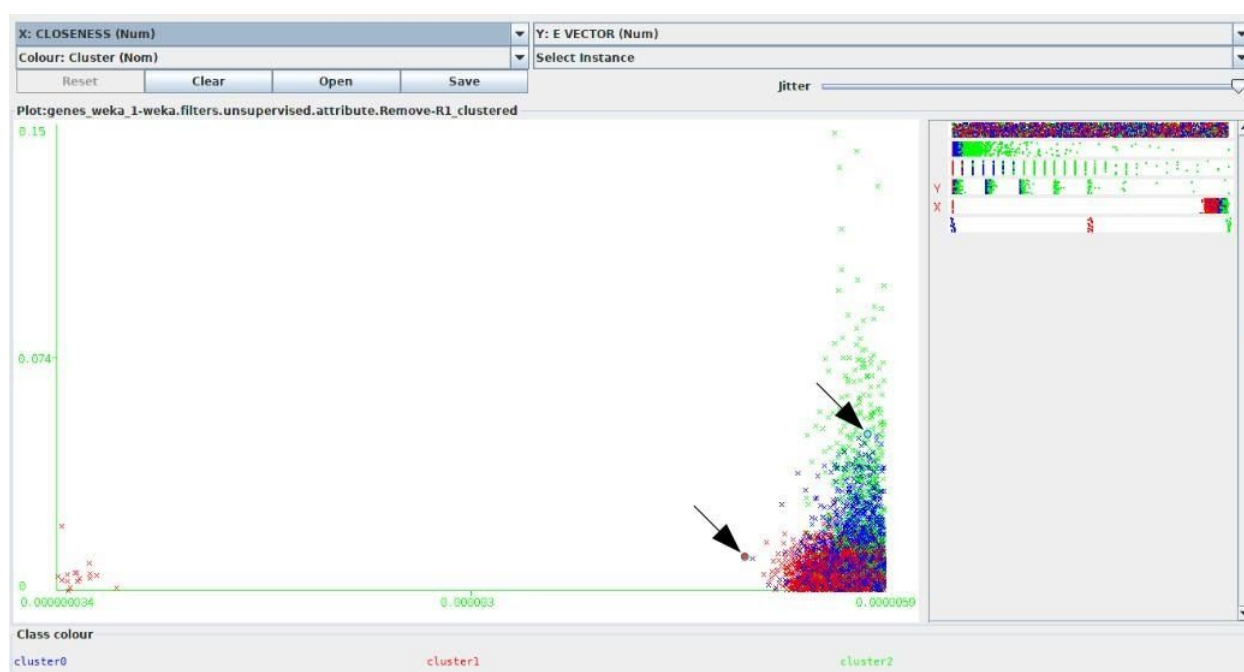


Figura 9. Gráfico de centralidade de proximidade (eixo x) vs autovetores (eixo y), os 3 clusteres formados pelo programa Weka, versão 3.6.14 se encontram em cores diferentes sendo o cluster 2 (verde) o grupo de “maior importância”, o cluster 1 (azul) de “importância mediana” e o cluster 0 (vermelho) de “menor importância”. Os locais apontados pelas flechas são os valores máximos de x (ciano) no valor de 0.000005642 e y (laranja) no valor de 0.0366076. Fonte: Os autores.

Pegando o valor máximo obtido de cada gráfico para cada medida foi criado um script (Figura 10), excluindo valores acima de um determinado limiar (esse obtido pelo programa Weka para cada medida).

```
cat "cluster_EM_test.arff" | awk '{if (($4) > "0.0366076"
&& ($2) > 0.0000894900565008639") print}' > ev_btw.csv
cat "cluster_EM_test.arff" | awk '{if (($4) > "0.0366076"
&& ($3) > 0.0048644892") print}' > ev_dg.csv
cat "cluster_EM_test.arff" | awk '{if (($4) > "0.0366076"
&& ($2) > 0.000005642") print}' > ev_cl.csv
```

Figura 10. Script usado para filtrar os pontos de “maior importância”, provenientes dos dados gerados pelo programa Weka, versão 3.6.14, uso das condições (if) aceitando apenas pontos com valores de autovetores > 0.0366076 e centralidade de proximidade > 0.000005642 para o arquivo “ev_cl.csv” e autovetores > 0.0366076 e centralidade de intermediação > 0.0000894900565008639 para o arquivo “ev_btw.csv” e autovetores > 0.0366076 e grau > 0.0048644892 . Fonte: Os autores.

Para a identificação dos pontos de intersecção entre os pontos obtidos nos 3 gráficos utilizou-se a ferramenta web Venny, usando como arquivo de entrada as tabelas geradas pelo script da Figura 6 contidas nos links: [1](#), [2](#), [3](#), gerando o diagrama de Venn mostrado na figura 11.

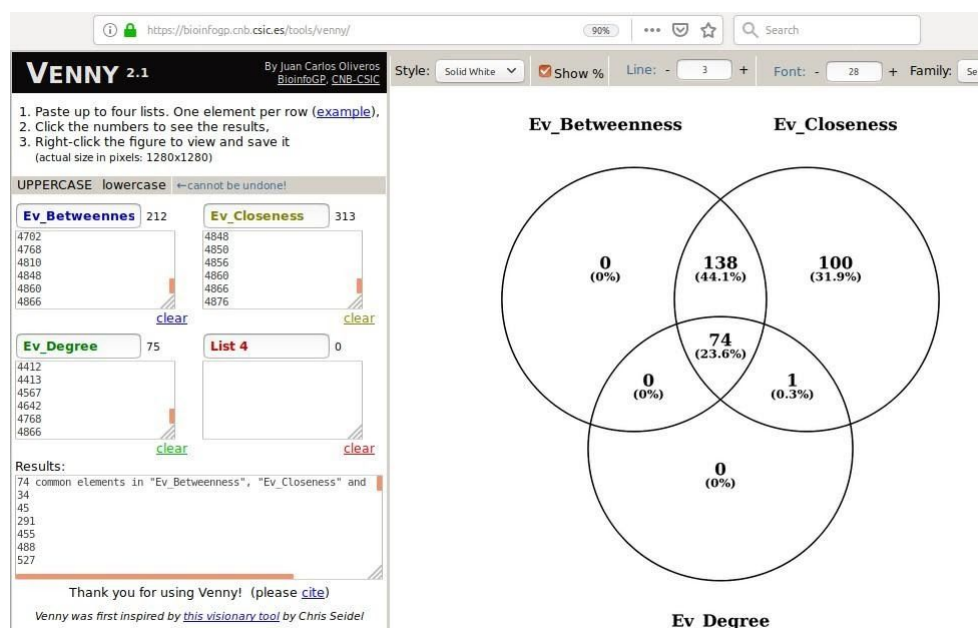


Figura 11. Diagrama de Venn contendo 74 pontos de “maior importância” na intersecção entre as 3 grandezas (centralidade de proximidade, centralidade de intermediação e autovetores) mostrados na aba “Results”. Fonte: Os autores.

Com os dados obtidos na aba “results”, pôde-se encontrar, manualmente, o nome dos genes e, via comando “grep” no terminal do Linux os miRNAs que regulam mRNAs produzidos por esses genes, gerando a tabela do link: [4](#)

Resultados e Discussão

Foram utilizados os dados de transcriptoma de *Oreochromis niloticus* Bovolenta et. al., onde foram selecionados, com uso de filtros no terminal usando comando “grep” do Linux os miRNAs que regulam os genes: TSC1/2 (Tuberous sclerosis protein formando um complexo (GTPase) atuante na via mTOR ativando a mTORC1), Murf1 (Muscle RING-finger protein-1, uma Ubiquitina Ligase (E3), atuando na marcação e degradação proteica), TRAF6 (TNF receptor associated factors 6, NF-κB um fator de transcrição que regula diversos genes sendoB um fator de transcrição que regula diversos genes sendo alguns deles fatores de crescimento), Set7 (histone H3 lysine 4 methyltransferase, atua na regulação do nível de expressão de Yap (Yes-associated protein) de modo negativo, sendo a proteína YAP um fator de transcrição ligado a proliferação celular e apoptose e importante na via Hippo), 2 tipos de serine/threonine kinases: Mst1/2 (Mammalian Ste20-like 1 and 2, responsáveis por dar início na cascata de sinalização formando um complexo com Sav1, na via Hippo), e Lats1/2 (Large Tumor Suppressor 1 and 2, formando junta a MOB1 o complexo LATS1/2–MOB1 que fosforila a proteína YAP ou o complexo YAP/TAZ, e fosforilada o complexo e/ou a proteína ficam retidas no citoplasma e degradadas) e 2 proteínas adaptadoras: Sav1 (Protein salvador homolog 1, formando junta a MST1/2 um complexo central na via Hippo, fosforilando e ativando o complexo LATS1/2–MOB1), Mob1 (Mps One Binder 1).

Usando o comando “grep”, no terminal do Linux, foram encontrados os miRNAs que regulam mRNAs provenientes dos genes de interesse entre outros genes (Tabela 2).

Nome miRNAs	Número de Alvos
miR-200a-3p	47
LG7_17451390-17451451_plus-5p	136
miR-130c-5p	119
miR-153a-3p/miR-153b-3p/miR-153c-3p	81
miR-25-3p	113
miR-26a-3p	4
miR-92a-3p/miR-92b-3p	113
miR-19b-5p	54
miR-19b-3p/miR-19d-3p	160
miR-25-3p	113
miR-9-3p	157
LG11_2906115-2906174_plus-3p	90
LG1_13952949-13953025_minus-5p	113
LG20_31197041-31197103_minus-3p	21
LG20_6198518-6198578_plus-3p	52
LG7_17451390-17451451_plus-5p	136

miR-142a-3p	30
miR-153b-5p	299
miR-214-3p	25
miR-2187-5p	126
miR-219-3p-2/miR-219-3p-3	20
LG6_6354367-6354432_minus-3p	25

Tabela 2. Foram encontrados 15 miRNAs e 7 linked Group (LG), o nome dos alvos estão localizados no link a seguir do material suplementar: [5](#). Fonte: Os autores.

Então excluídos os miRNAs (usando comando grep novamente) que regulam genes constitutivos (Tabela 3).

Nome miRNAs

miR-200a-3p
 LG7_17451390-17451451_plus-5p
 miR-130c-5p
 miR-25-3p
 miR-26a-3p
 miR-92a-3p
 miR-92b-3p
 miR-19b-5p
 miR-1-3p
 miR-19a-3p
 miR-9-3p
 LG20_31197041-31197103_minus-3p
 LG6_6354367-6354432_minus-3p
 miR-142a-3p
 miR-214-3p
 miR-219-3p-2
 miR-219-3p-3

Tabela 3. Foram encontrados 14 miRNAs e 3 linked group. Fonte: Os autores.

A tabela dos genes constitutivos possui 17 genes sendo eles:

Genes Constitutivos

18S ribossomal RNA

Beta actin

GAPDH

Phosphoglycerate kinase 1

Peptidylprolyl isomerase A

Ribosomal protein L13a

Ribosomal protein large P0

Acidic ribosomal phosphoprotein PO

Beta-2-microglobulin

Tyrosine 3-monooxygenase/tryptophan5-monooxygenase activation protein

zeta polypeptide

Succinate dehydrogenase complex subunit A flavoprotein (Fp)

Transferrin receptor

beta Glucuronidase

Hydroxymethylbilane synthase

TATA box binding protein

Hypoxanthine phosphoribosyltransferase 1

Tabela 4. Tabela obtida na literatura (THELLIN et al., 1999).

Ainda com relativamente grande quantidade de miRNAs e LG (linked group), no caso 17, com intuito de usar combinações de alterações na expressão de miRNAs que possuam uma certa “sinergia” para formação do fenótipo desejado (maior produtividade de carne), foram feitos testes estatísticos usando scripts no pacote “Igraph” no R version 3.4.4 usando como arquivo de entrada os dados brutos provenientes da tese de Bovolenta et. al., calculando-se medidas de “importância” de pontos: grau, centralidade de proximidade, centralidade de intermediação, autovetores, os valores normalizados obtidos se encontram no link do material suplementar a seguir: [6](#).

Os valores encontrados foram usados como arquivo de entrada para o programa Weka versão 3.6.14 (HALL et al., 2009), encontrando os valores máximos de autovetores (eixo y) estático pelos demais valores (centralidade de proximidade, centralidade de intermediação e grau) no eixo “x”.

Foram utilizados valores de importância para serem selecionados e excluídos miRNAs que regulam genes de aparente importância para o organismo, seguindo o mesmo intuito da exclusão de miRNAs que regulam genes constitutivos.

Após a obtenção dos valores máximos, foi criado um script (detalhado em Materiais e Métodos), obtendo-se no total de 600 pontos de “maior importância” onde no gráfico de centralidade de intermediação vs autovetores temos 212 pontos, no gráfico de centralidade de proximidade vs autovetores 313 pontos e no gráfico de grau vs autovetores 75 pontos, esses avaliados pela ferramenta web [Venny](#), com os valores resultantes (74 pontos de intersecção) foram encontrados os nomes dos genes (formato Ensembl) e os miRNAs que regulam mRNAs produzidos por esses genes se utilizando o comando “grep”, e manualmente pesquisando no banco de dados Ensembl pelo nome mais utilizado na literatura para cada gene.

Os miRNAs encontrados foram:

Nome miRNAs

miR-9-3p
miR-19b-3p
miR-19a-3p
miR-153a-3p
miR-153b-3p
miR-153b-5p
miR-153c-3p
miR-2187-5p
LG1_13952949-13953025_minus-5p
LG3_6766487-6766556_plus-5p
LG7_17451390-17451451_plus-5p
LG8-24_23675661-23675722_plus-3p
LG11_2906115-2906174_plus-3p
LG20_6198518-6198578_plus-3p

Tabela 5. Foram encontrados 8 miRNAs e 6 linked group. Fonte: Os autores.

Foram excluídos, manualmente, os miRNAs também presentes na tabela 3, os miRNAs excluídos foram: miR-9-3p, miR-19a-3p, LG7_17451390-17451451_plus-5p pois, não queremos miRNAs que atuam em alvos de grande importância na manutenção do organismo.

O banco de dados utilizado para se encontrar as possíveis vias de interações dos miRNAs filtrados (não sendo possível de se analisar os linked group pois, são miRNAs novos e não se encontram na literatura) foi o miRPathDB v 2.0 no site: [miRPathDB](http://miRPathDB.com), que se utiliza de diversos bancos de dados como: KEGG, GO (Gene Ontology), WikiPathways, Reactome entre outros (KEHL et al., 2019).

A principal tabela gerada foi a tabela de miRNAs e sua atuação em cada uma das vias de interesse no presente trabalho, sendo elas nomeadas da seguinte maneira: Hippo signaling pathway, mTOR, ubiquitin like protein ligase activity 5, ubiquitin ligase complex 4, positive regulation of protein ubiquitination involved in ubiquitin dependent protein catabolic process 8, tendo como destaque o miR-214-3p que possui influência em todas as vias (Tabela 6).

	Hippo pathway	signaling mTOR	ubiquitin like protein ligase activity 5	ubiquitin ligase complex 4
miR-200a-3p	X		X	X
miR-130c-5p				
miR-25-3p			X	X
miR-26a-3p				
miR-92a-3p			X	X
miR-92b-3p			X	X
miR-19b-5p				
miR-1-3p		X		
miR-142a-3p				
miR-214-3p	X	X	X	X
miR-219-3p-2/miR-219-3p-3				
LG6_6354367-6354432_minus-3p				
LG20_31197041-31197103_minus-3p				

Tabela 6. miRNAs e sua interação com componentes das vias metabólicas: Hippo, mTOR, e interações com complexos e regulação de vias relacionadas com proteínas ubiquitina ligase.

Utilizando-se do comando “grep”, no terminal do Linux, foi possível encontrar os alvos dos 12 miRNAs citados acima, a lista gerada se encontra no link: [8](#).

O miRNAs mais interessante analisado foi o miR-214-3p, sendo visível sua importância para com as vias responsáveis para desenvolvimento muscular.

Possíveis papéis do miR-214 revisados na literatura.

Quando a expressão de miR-214 foi aumentada em células HeLa, cancerígenas, verificou-se a redução significativa do crescimento do tumor (YANG et al., 2009), em contrapartida o aumento da expressão do miR-214 no câncer de pâncreas pode trazer maior resistência à quimioterapia. (ZHANG et al., 2010). Na linha celular de glioma humano, foi demonstrado que a expressão de miR-214 suprime a expressão da enzima conjugadora de ubiquitina 9 (UBC9) e reduz a proliferação de células tumorais (ZHAO et al., 2012). No peixe-zebra, a expressão do miR-214 é observada principalmente em torno do desenvolvimento de elementos esqueléticos (DESVIGNES; CONTRERAS; POSTLETHWAIT, 2014). O miR-214 é expresso durante os estágios iniciais da segmentação dos somitos no peixe-zebra (FLYNT et al., 2007). A inibição do miR-214 resulta em uma redução ou perda de tipos de células musculares lentas (FLYNT et al., 2007). Participa da miogênese da musculatura esquelética, atuando em diversos alvos importantes na via Akt/mTOR como por exemplo: INSR, IRS2 e AKT2 (HONARDOOST et al., 2015). A expressão do miR-214 em mioblastos é significativamente reduzida após a indução da diferenciação demonstrado em análises de qRT-PCR mostrou que os níveis de expressão de INSR, IRS2 e ATK2 aumentaram (demonstrando sua importância para o processo de diferenciação muscular) devido ao declínio da expressão do miR-214 que inibe a expressão desses genes (Figura 12) (HONARDOOST et al., 2015).

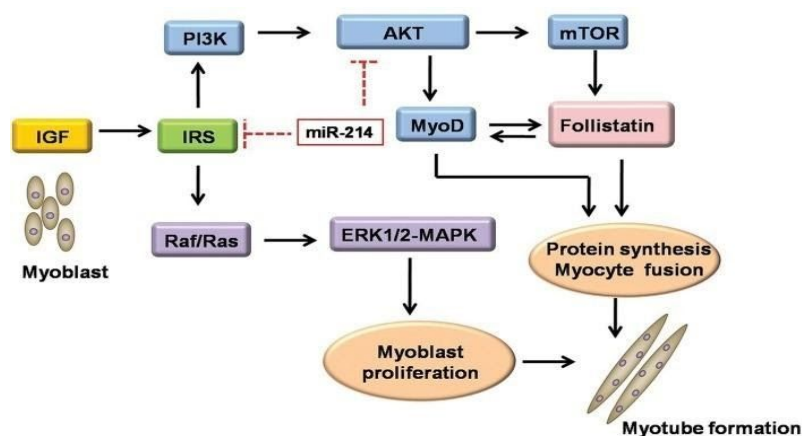


Figura 12. Miogênese pela via IGF / AKT / mTOR sendo o miR-214 um inibidor de AKT molécula importante para impulsionar a diferenciação aumentando a atividade transcricional do MyoD e ativando a via mTOR. IGF inicia uma sinalização através do receptor IRS ativando a cascata de MAPK com miR-214 inibindo IRS regulando negativamente a cascata. A atividade do mTOR aumenta durante a diferenciação, levando a um aumento na atividade da follistatina, que impede a miostatina (inibidor do crescimento muscular) de executar seu efeito inibitório. AKT;

serina/treonina quinase, ERK1/2; Quinases reguladas por sinal extracelular 1/2, IGF; Fator de crescimento semelhante à insulina, IRS; Substrato receptor de insulina, mTOR; Alvo da rapamicina em mamíferos, MyoD; Proteína de diferenciação miogênica, PI3K; Fosfatidilinositol 3-cinase, Raf; Fibrosarcoma e Ras; Vírus do sarcoma de rato. Adaptado de (HONARDOOST et al., 2015).

Conclusão

Com o uso de ferramentas de bioinformática foi encontrado o miR-214, importante regulador da miogenese através da via AKT/mTOR e também atuante em outras vias de importância na proliferação, autofagia, degradação e manutenção celular como a via Hippo e a via das Ubiquitina-Ligases tendo grande importância na hiperplasia e hipertrofia muscular sendo um candidato para futuros trabalhos in vivo relacionados a ganho de massa muscular em Tilapia do Nilo (*Oreochromis niloticus*).

Material Suplementar

- 1) <https://drive.google.com/open?id=1hUszIIGzLbXWGXhZwPW7DX175RWPows4>
- 2) https://drive.google.com/open?id=1LtsvhmHHbV_AxNhnZM4aeRKPrBQOaf6R
- 3) <https://drive.google.com/open?id=1JKq8Y2bv7j3VlKA-1a3v-g0vAUIWIKNz>
- 4) https://drive.google.com/open?id=1iCIm3WPotQrev1F9rRRI_9KCaRqbXVG8
- 5) https://drive.google.com/open?id=1tjjTXD98KGMf7pUyqEbZNB1_PvMOukHo
- 6) <https://drive.google.com/open?id=1XC5usQSZ8JCN6-tsN7OzPDB0V9b8n6FX>
- 7) https://drive.google.com/file/d/17Cf7a-aIzVA7TCEJyJ5uHFrfu7m_NvBS
- 8) <https://drive.google.com/open?id=1GdIch4BYV8JDZDaeQStTUsjCdQA341it>
- 9) https://drive.google.com/open?id=144FZi_vmJrdwxd4XoNSdZUkOa5800wc0
- 10) https://drive.google.com/open?id=1Uso7_MAhLcZkGpsaDbjDsoltIp-Hv6Q-
- 11) https://drive.google.com/open?id=16Y1norn--qQWBic_D2gbvUEpcf-c5eRv
- 12) https://drive.google.com/open?id=1YDlZJfhSNXATvodW3T3ouQxll45d_lbG
- 13) https://drive.google.com/open?id=1lu5kmWOeC1IE7PEiBcmkZKz8r_lTaaId
- 14) https://drive.google.com/open?id=1b5Y6UBwvAPO-ga_TVE5kU4I1ZDNGPQO-
- 15) <https://drive.google.com/open?id=1dT6eFwJkwip2Mz6m836w0z9DRV9wqWys>
- 16) <https://drive.google.com/open?id=1eYb6ecSRmfawp1On97eFLv2GXPICdk-J>
- 17) https://drive.google.com/open?id=1FDtz_v2jAbJH9C-4x_OCIEqMEf92Z4SS
- 18) <https://drive.google.com/open?id=1e-z2cHR78n58IFwoN7Yo1qipk73ERJZs>
- 19) <https://drive.google.com/open?id=1dfAcPQf-KiEf2aSPVnOcM7MKJogOAL1p>

Referências Bibliográficas

- ATTAIX, D. et al. Role of the ubiquitin-proteasome pathway in muscle atrophy in cachexia. **Current opinion in supportive and palliative care**, v. 2, n. 4, p. 262–266, dez. 2008.
- BARTEL, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. **Cell**, v. 116, n. 2, p. 281–297, 2004.
- BONACICH, P., 1987, “Power and Centrality: A Family of Measures”, **The American Journal of Sociology**, v. 92, n 5, pp. 1170-1182.
- BOVOLENTA, L. A. Análise exploratória em larga escala de microRNAs expressos em tilápia do Nilo utilizando ferramentas de bioinformática. 2016. Tese de Doutorado - Universidade Estadual Paulista “Júlio de Mesquita Filho” Instituto de Biociências (Campus de Botucatu).
- CHAI, J.; WU, Y.; SHENG, Z. Z. Role of ubiquitin-proteasome pathway in skeletal muscle wasting in rats with endotoxemia. **Critical care medicine**, v. 31, n. 6, p. 1802–1807, jun. 2003.
- CORREA, A. DE S. et al. Rearing of Nile tilapia (*Oreochromis niloticus*) juveniles in a biofloc system employing periods of feed deprivation. **Journal of Applied Aquaculture**, p. 1–18, 31 out. 2019.
- COSTA-SILVA, J.; DOMINGUES, D.; LOPES, F. M. RNA-Seq differential expression analysis: An extended review and a software tool. **PloS one**, v. 12, n. 12, p. e0190152–e0190152, 21 dez. 2017.
- DESVIGNES, T.; CONTRERAS, A.; POSTLETHWAIT, J. H. Evolution of the miR199-214 cluster and vertebrate skeletal development. **RNA biology**, v. 11, n. 4, p. 281–294, 2014.
- DIENSTMANN, R. et al. Picking the Point of Inhibition: A Comparative Review of PI3K/AKT/mTOR Pathway Inhibitors. **Molecular Cancer Therapeutics**, v. 13, n. 5, p. 1021 LP – 1031, 1 maio 2014.
- DO, C. B.; BATZOGLOU, S. What is the expectation maximization algorithm? **Nature Biotechnology**, v. 26, n. 8, p. 897–899, 2008.
- EDDINS, M. J. et al. Targeting the ubiquitin E3 ligase MuRF1 to inhibit muscle atrophy. **Cell biochemistry and biophysics**, v. 60, n. 1–2, p. 113–118, jun. 2011.
- EKNATH, A. E.; HULATA, G. Use and exchange of genetic resources of Nile tilapia (*Oreochromis niloticus*). **Reviews in Aquaculture**, v. 1, n. 3-4, p. 197–213, 1 set. 2009. 4, p. 197–213, 1 set. 2009.
- FLYNT, A. S. et al. Zebrafish miR-214 modulates Hedgehog signaling to specify muscle cell fate. **Nature Genetics**, v. 39, n. 2, p. 259–263, 2007.
- FORABOSCO, F. et al. Genetically modified farm animals and fish in agriculture: A review. **Livestock Science**, v. 153, n. 1, p. 1–9, 2013.
- GREGORY, R. I.; SHIEKHATTAR, R. MicroRNA Biogenesis and Cancer. **Cancer Research**, v. 65, n. 9, p. 3509 LP – 3512, 1 maio 2005.
- GUMBINER, B. M.; KIM, N.-G. The Hippo-YAP signaling pathway and contact inhibition of

growth. **Journal of cell science**, v. 127, n. Pt 4, p. 709–717, 15 fev. 2014.

HALL, M. et al. The WEKA Data Mining Software: An Update. **SIGKDD Explor. Newsl.**, v. 11, n. 1, p. 10–18, 2009.

HONARDOOST, M. et al. Expression Change of miR-214 and miR-135 during Muscle Differentiation. **Cell journal**, v. 17, n. 3, p. 461–470, 2015.

KEHL, T. et al. miRPathDB 2.0: a novel release of the miRNA Pathway Dictionary Database. **Nucleic Acids Research**, 6 nov. 2019.

LAGO, A. DE A. et al. The development of genetically improved red tilapia lines through the backcross breeding of two *Oreochromis niloticus* strains. **Aquaculture**, v. 472, p. 17–22, 2017.

LYALL, J. et al. Suppression of Avian Influenza Transmission in Genetically Modified Chickens. **Science**, v. 331, n. 6014, p. 223 LP – 226, 14 jan. 2011.

MEMMOTT, R. M.; DENNIS, P. A. Akt-dependent and -independent mechanisms of mTOR regulation in cancer. **Cellular Signalling**, v. 21, n. 5, p. 656–664, 2009.

MITRA, J. Genetics and genetic improvement of drought resistance in crop plants. **Current Science**, v. 80, n. 6, p. 758–763, 18 fev. 2001.

OUDHOFF, M. J. et al. Control of the Hippo Pathway by Set7-Dependent Methylation of Yap. **Developmental Cell**, v. 26, n. 2, p. 188–194, 2013.

PAUL, P. K. et al. Targeted ablation of TRAF6 inhibits skeletal muscle wasting in mice. **The Journal of cell biology**, v. 191, n. 7, p. 1395–1411, 27 dez. 2010.

SALTER, D. W.; CRITTENDEN, L. B. Artificial insertion of a dominant gene for resistance to avian leukosis virus into the germ line of the chicken. **Theoretical and Applied Genetics**, v. 77, n. 4, p. 457–461, 1989.

SHI, C.-S.; KEHRL, J. H. TRAF6 and A20 regulate lysine 63-linked ubiquitination of Beclin-1 to control TLR4-induced autophagy. **Science signaling**, v. 3, n. 123, p. ra42–ra42, 25 maio 2010.

SISHI, B. et al. Diet-induced obesity alters signalling pathways and induces atrophy and apoptosis in skeletal muscle in a prediabetic rat model. **Experimental physiology**, v. 96, n. 2, p. 179–193, fev. 2011.

YANG, Z. et al. MicroRNA-214 is aberrantly expressed in cervical cancers and inhibits the growth of HeLa cells. **IUBMB Life**, v. 61, n. 11, p. 1075–1082, 1 nov. 2009.

ZERHOUNI, E. A. Information access. NIH public access policy. **Science (New York, N.Y.)**, v. 306, n. 5703, p. 1895, 10 dez. 2004.

ZHANG, X. J. et al. Dysregulation of miR-15a and miR-214 in human pancreatic cancer. **Journal of hematology & oncology**, v. 3, p. 46, 24 nov. 2010.

ZHAO, Z. et al. microRNA-214-mediated UBC9 expression in glioma. **BMB reports**, v. 45, n. 11, p. 641–646, nov. 2012.