

UNIVERSIDADE NOVE DE JULHO
PROGRAMA DE PÓS-GRADUAÇÃO DOUTORADO EM MEDICINA

ROBSON JOSÉ DE ALMEIDA

**ANÁLISE DE TRANSCRIPTOMA COMBINADA COM ESTUDO DE
CONJUNTOS DE DADOS MULTIÔMICOS NA IDENTIFICAÇÃO DE
TRANSCRITOS REGULADOS PELO HORMÔNIO TIROIDIANO**

SÃO PAULO - SP

2020

ROBSON JOSÉ DE ALMEIDA

ANÁLISE DE TRANSCRIPTOMA COMBINADA COM ESTUDO DE
CONJUNTOS DE DADOS MULTIÔMICOS NA IDENTIFICAÇÃO DE
TRANSCRITOS REGULADOS PELO HORMÔNIO TIROIDIANO

Tese de doutorado apresentado ao
Programa de Pós-Graduação em Medicina
da Universidade Nove de Julho, como
requisito para obtenção de título de Doutor
em Ciências da Saúde.

Orientador: Prof. Dr. Cléber Pinto Camacho

**São Paulo - SP
2020**

Almeida, Robson José de.

Análise de transcriptoma combinada com estudo de conjuntos de dados multiômicos na identificação de transcritos regulados pelo hormônio tireoideano. / Robson José de Almeida. 2020.

126 f.

Tese (Doutorado) - Universidade Nove de Julho - UNINOVE, São Paulo, 2020.

Orientador (a): Prof. Dr. Cléber Pinto Camacho.

São Paulo, 14 de dezembro de 2020.



São Paulo 14 de dezembro de 2020

TERMO DE APROVAÇÃO

Aluno: **ROBSON JOSÉ DE ALMEIDA**

Titulo da tese: **META-ANÁLISE MULTIÔMICA COMBINADA COM SEQUENCIAMENTO DE LARGA ESCALA NA IDENTIFICAÇÃO DE TRANSCRITOS REGULADOS PELO HORMÔNIO TIROIDIANO.**

Presidente: PROF. DR. CLEBER PINTO CAMACHO

Membro: PROF. DR. JOSÉ ANTÔNIO SILVA JÚNIOR

Membro: PROFA. DRA. SORAIA MICAELA

Membro: PROF. DR. ALEXANDRE ROSSI PASCHOAL

DEDICATÓRIA

Dedico esta tese à minha esposa e toda minha família, que incondicionalmente me apoiaram a seguir meus sonhos.

Ao irmão Paulo José de Almeida (*in memoriam*), ser humano iluminado, que teve como missão nos ensinar o amor, simplicidade e a alegria de viver. A evolução do seu espírito sempre me inspirará.

AGRADECIMENTOS

A Universidade Nove de Julho, pela oportunidade de realização do meu Doutorado, assim como no passado da minha graduação e Mestrado, em especial a nossa coordenadora Profa. Dra. Fernanda Marciano Consolim Colombo.

Ao Prof. Dr. Cléber Pinto Camacho, pela sabedoria e paciência em lapidar meu intelecto. Pelo apoio motivador e ensinamentos que transformaram minha vida. Que eu possa honrá-lo sempre, como digno ser humano e mestre que é. Gratidão eterna!

Ao Prof. Dr. Humberto Dellê, pelos aconselhamentos, redirecionamento profissional, aprendizados, preocupação e ajuda.

Ao Prof. Dr. Daniel Giannella Neto (*in memoriam*), pela forma simples, descontraída, porém amplamente experiente de transmitir conhecimento. Seus alvíres jamais serão esquecidos!

Ao Centro de Pesquisa e Diagnóstico Molecular de Doenças Genéticas da Escola Paulista de Medicina, departamento de biofísica, em especial ao Prof. Dr. João Bosco Pesquero, pela cordialidade em “abrir” seu laboratório em apoio e parceria a nossa pesquisa, disponibilizando de forma irrestrita seus colaboradores e equipamentos, bem como a Dra. Patrícia Varella Lima Teixeira (Escola Paulista de Medicina), que com muita gentileza e prestatividade nos auxiliou com sua larga experiência laboratorial, e o Dr. Leonardo Martins Silva (Escola Paulista de Medicina) pela atenção, solicitude e execução dos protocolos necessários ao transcriptoma. A todos, o meu absoluto respeito e gratidão.

Ao Laboratório de Endocrinologia Molecular e Translacional (LEMT) da disciplina de Endocrinologia da Escola Paulista de Medicina, departamento de endocrinologia, em particular a Msc. Teresa Kasamatsu, pela bondade de me ensinar e executar provas bioquímicas necessárias em nossa pesquisa.

Aos colegas Luiz Antônio de Jesus Rocha, Valdelena Alessandra da Silva, Andréa Harumy de Lima Hirata, Albertina Gomes Rodrigues, Mariana de Almeida Camargo Lautenschlager, Yone de Oliveira Di Sarli e Miriam Duarte de Arruda Motta pela a honra e a oportunidade de trabalharmos juntos, pela parceria e os momentos de mútuas motivações.

À Fundação de amparo à pesquisa do estado de São Paulo (FAPESP) sob o número 2014/04193-0 e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) sob o número 422445/2016-9, pelo apoio financeiro, certo que sem a

aquisição do fomento esse trabalho não seria concretizado, e também a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio ao programa de pós graduação *Strictu Sensu* em Medicina da Universidade Nove de Julho . Em nome de toda equipe envolvida nesse trabalho, o nosso muito obrigado.

A todos que de forma direta e indireta participaram para a concretização dessa pesquisa, o meu muito obrigado.

RESUMO

A tireoide é uma glândula endócrina que sintetiza dois hormônios denominados de hormônios tireoidianos (HT), T3 e T4, suas funções sistêmicas, são essenciais para o desenvolvimento e manutenção do organismo, atuando de forma efetiva no metabolismo celular. A investigação clínica por NGS fornece uma alta sensibilidade em uma plataforma de alto rendimento, produzindo dados torrenciais para uma ampla análise genômica. A metabolômica é uma técnica que avalia a concentração de moléculas e substratos em uma determinada situação de interesse. As mitocôndrias regulam seus mecanismos de controle de qualidade afim de manter a homeostase energética do tecido em qualquer condição. Em nosso trabalho, o sangue periférico foi o material escolhido dada à capacidade de amplitude de informações que o transcriptoma de nova geração pode oferecer. Este trabalho foi dividido em dois objetivos: Objetivo primário: Caracterização dos genes mitocondriais diferencialmente expressos no hipotireoidismo e sua correlação com o quadro laboratorial por análises dos transcriptomas, metabolômas, metilômas e identificação de TREs e objetivo secundário: Identificação de genes diferencialmente expressos buscando qualitativamente e quantitativamente identificar o melhor método analítico em diferentes ferramentas e bibliotecas. Cada sequenciamento é um universo particular, fator que define estratégias em bioinformática mais adequadas para diferentes estudos, e para nossa biblioteca a melhor estratégia foi a dispersão comum, dispersão de tendência e teste de razão de verossimilhança. Observamos que os miRNAs e snoRNAs exercem influência sobre os mRNAs diferencialmente expressos. Além disso, os dados da metabolômica e metiloma já publicados, sugerem uma regulação extra genômica que apontam uma tendência biológica do organismo buscar o equilíbrio à uma possível ineficiência do hormônio exógeno. A quantificação de TREs podem ser importantes legitimadores da expressão gênica na eficiência ou deficiência do T3. Genes mitocondriais sofrem alterações de expressões no hipotireoidismo diferentemente do conceito de genes estáveis.

Palavras chaves: Tireoide. Hipotireoidismo. Transcriptoma. RNAseq. Sequenciamento. Biomarcador. Metilômica. Metabolômica.

ABSTRACT

The thyroid is an endocrine gland that synthesizes two hormones called thyroid hormones (HT), T3 and T4, their systemic functions, are essential for the development and maintenance of the organism, acting effectively in cell metabolism. Clinical research by NGS provides high sensitivity on a high-throughput platform, producing torrential data for extensive genomic analysis. Metabolomics is a technique that assesses the concentration of molecules and substrates in a given situation of interest. Mitochondria regulate their quality control mechanisms in order to maintain tissue energy homeostasis in any condition. In our work, peripheral blood was the material of choice given the capacity of breadth of information that the new generation transcriptome can offer. This work was divided into two objectives: Primary objective: Characterization of mitochondrial genes differentially expressed in hypothyroidism and their correlation with the laboratory condition by analysis of transcriptomes, metabolomes, methylomes and identification of TREs and secondary objective: Identification of differentially expressed genes seeking qualitatively and quantitatively identify the best analytical method in different tools and libraries. Each sequencing is a particular universe, a factor that defines strategies in bioinformatics that are most suitable for different studies, and for our library the best strategy was the common dispersion, trend dispersion and likelihood ratio test. We observed that miRNAs and snoRNAs exert an influence on differentially expressed mRNAs. In addition, the published metabolomics and methylome data suggest an extra genomic regulation that points to a biological tendency of the organism to seek equilibrium with a possible inefficiency of the exogenous hormone. The quantification of TREs can be important to legitimize gene expression in T3 efficiency or deficiency. Mitochondrial genes undergo changes in expression in hypothyroidism unlike the concept of stable genes.

Key words: Thyroid. Hypothyroidism. Transcriptome. RNAseq. Sequencing. Biomarker. Methylomic. Metabolomics.

LISTA DE TABELAS

Tabela 1. Distribuições antropométricas, laboratoriais e clínicas dos grupos CTL e HT. Sexo, Idade: anos, Peso: Kg, TSH valores de referência >20 anos 0,270 - 4,50 UI/mL, T4 Livre valores de referência 0,93 - 1,7 ng/dL, TgAb valores de referência <115 UI/, TPOAb valores de referência < 34 UI/mL e Tireoidectomia.

Tabela 2. Valores de N, medianas, valor mínimo, máximo e percentis dos grupos Controle e Hipotireoidismo.

Tabela 3. Dados moleculares e qualitativos pré, peri e pós confecções das bibliotecas no NGS. [RNA] concentração de ácido ribonucleico ideal: >10 ng/ μ L, % IQ [RNA] integridade e qualidade da [RNA] ideal : 1 - 10%, qPCR quantificação da reação em cadeia da polimerase ideal: >100pM, % *Beads* esferas magnéticas conjugadas ao DNA ideal: 10 - 30%, *Id Barcode*: Identificação da amostra, *Id Chip*: Identificação do *CHIP* de sequenciamento, Comprimento médio: pb (pares de bases), Leituras de fitas de DNA: %, *Output* BAM tamanho: Saída do arquivo de múltiplo alinhamento binário em GB (*Gigabytes*).

Tabela 4. Lista DGE filtrada para transcritos não codificantes, por ordem crescente da expressão do logFC, identificação Refseq, símbolo do gene, valores de logCPM, LR e FDR.

Tabela 5. Genes não codificantes, descrições e biotipos.

Tabela 6. Lista com os 8 genes mitocondriais conforme os resultados de logFC, logCPM, LR, FDR, Refseq e símbolo do gene da lista DGE (HT comparado ao CTL).

Tabela 7. Descrição dos genes mitocondriais da lista DGE (HT comparado ao CTL).

Tabela 8. Lista de DGE mitocondriais da biblioteca HT comparado a LT-4 no alvo terapêutico, conforme as expressões das duas análises.

Tabela 9. Comparação das expressões dos mitGenes do grupo HT e banco GSE131411 no T2 e T3.

Tabela 10: Avaliação da possível influência das sequências responsivas ao T3. A sequência de maior destaque é a "TCCAGT" representada pelo gene *CMC1*.

Tabela 11: Contagem da densidade das ilhas CPGs e o seu tamanho em nucleotídeos.

Tabela 12: Comparação entre RNAs e metabólitos relacionados que estão na mesma via biológica. O * indica valor retirado de trabalhos externos.

LISTA DE GRÁFICOS

Gráfico 1. Conjunto dos 8 genes mitocondriais e a expressão em tecidos alvos do hormônio tireoideano, apresentados em porcentagem e $-\log_{10}$ (*p-value*).

Gráfico 2. Conjunto dos 8 genes mitocondriais e a relação com funções moleculares, apresentados em porcentagem e $-\log_{10}$ (*p-value*).

Gráfico 3. Conjunto dos 8 genes mitocondriais e a relação com processos biológicos, apresentados em porcentagem e $-\log_{10}$ (*p-value*).

Gráfico 4. Conjunto dos 8 mitGenes e a relação com os fenótipos clínicos catalogados na literatura que são classicamente detectáveis no hipotireoidismo.

Gráfico 5: *Network* entre RNA não codificante snoRNA e miRNAs.

Gráfico 6: *Network* entre RNAs mitocondriais e miRNAs.

Gráfico 7. Correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH e T4L do grupo CTL, todas correlações em $p < 0,05$.

Gráfico 8. Correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH, T4L, Anti TgAb e Anti TPOAb do grupo HT, todas correlações em $p < 0,05$.

LISTA DE FIGURAS

Figura 1. Ilustração do mecanismo de regulação da síntese do hormônio tireoideano e seus efeitos intracelulares para controle de expressão gênica e manutenção de organelas.

Figura 2. Fluxograma das etapas da pesquisa resumindo as estratégias adotadas para que fosse possível a realização dos objetivos propostos no presente trabalho.

Figura 3. Comparação entre os pacotes EdgeR, DESeq2 e NOISeq, na detecção de DGE nas bibliotecas do grupo HT, Hipotireoidismo leve, GSE103305 e a concordância dos DGEs.

Figura 4. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*.

Figura 5. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão

Figura 6. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão.

Figura 7. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão.

Figura 8. Capacidade de detecção de DGE em cada estratégia.

Figura 9. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste exato nas bibliotecas do estudo estado da tiroide.

Figura 10. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo estado da tiroide.

Figura 11. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste exato nas bibliotecas do estudo estado da tiroide

Figura 12. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste de razão por verossimilhança nas bibliotecas do estudo estado da tiroide.

Figura 13. Capacidade de detecção de DGE em cada estratégia na biblioteca do estudo estado da tiroide

Figura 14. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste exato nas bibliotecas do estudo transtorno bipolar.

Figura 15. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo transtorno bipolar.

Figura 16. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*; D: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste exato nas bibliotecas do estudo transtorno bipolar.

Figura 17. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*; D: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste de razão por verossimilhança nas bibliotecas do estudo transtorno bipolar.

Figura 18. Capacidade de detecção de DGE em cada estratégia nas bibliotecas do estudo transtorno bipolar.

Figura 19. *Network* baseada na lista de transcritos não codificantes.

Figura 20. Genes não codificantes comuns entre HT comparado ao CTL e a biblioteca do estudo estado da tiroide.

Figura 21. Identificação de genes mitocondriais entre base *MitoMiner*, *MITOCARTA* e *Mitochondrial Biology Unit*, com a lista DGE (HT comparado ao CTL).

Figura 22. *Network* com 5 genes mitocôndrias (vermelhos) que interagem com outras vias metabólicas.

Figura 23. Identificação de genes mitocôndrias entre HT comparado ao CTL e lista DGE do estudo estado da tiroide.

LISTA DE ABREVIATURAS E SIGLAS

LEMT - Laboratório de Endocrinologia Molecular e Translacional
CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
HT – Hormônios tireoidianos
3,5,3',5'-tetraiodo-L-tironina - Tiroxina ou T4
3,5,3'-triiodo-L-tironina - Triiodotironina ou T3
HHT - Eixo das glândulas hipotálamo, hipófise e tireoide
TRH - Hormônio liberador de tirotrófina
TSH - Hormônio estimulador da tireoide
MIT - Iodotirosinas 3-moniiodo-L-tirosina
DIT - 3,5-diiiodo-L-tirosina
TSHR – Receptores de TSH
DIO - Iodotironinas desiodases
TR - Receptores de triiodotironina
TER - Elementos de resposta ao hormônio tireoidiano
MCT8 – Transportador de monocarboxilatos tipo 8
MCT10 – Transportador de monocarboxilatos tipo 10
DIO1 - Desiodases tipo 1
DIO2 - Desiodases tipo 2
DIO3 - Desiodases tipo 3
rT3 – T3 reverso
T2 - 3,3'-diiiodo-L-tironina
L-T4 - Levotiroxina
TPOAb - Anticorpo Anti-Tiroperoxidase
TgAb - Anticorpo Anti-Tiroglobulina
RNA – Ácido ribonucleico
mRNA – RNA mensageiro
miRNA – Micro RNA
NGS – Sequenciamento de nova geração
DNA – Ácido desoxirribonucleico
WGBS - *Whole Genome Bisulfite Sequencing*
lncRNA – RNA longo não codificante

lincRNA – RNA longo não codificante intergênicos

PCR-RT - Reação de cadeia da polimerase em tempo real

mtDNA – DNA mitocondrial

ATP – Adenosina trifosfato

FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

TCLE - Termo de consentimento livre e esclarecido

qPCR – PCR quantitativa

DGE – Genes diferencialmente expressos

NCBI - Centro Nacional de Informações sobre Biotecnologia

GO - *Gene ontology*

ncRNA – RNA não codificantes

IMPI - Índice de Proteína Mitocondrial Integrado

scaRNA – RNA específico para pequenos corpos de Cajal

snoRNA – Pequeno RNA nucleolar

mitGene – Gene mitocondrial

WGBS - *Whole Genome Bisulfite Sequencing*

LISTA DE SÍMBOLOS

kDa - Kilodalton

mU/L – Mili unidades por litro

mcg – Micrograma

Kg/dia – Quilograma por dia

μ UI/mL – Micro unidades internacionais por mililitros

UI/mL – Unidades internacionais por mililitros

mcg/kg/dia – Micrograma por quilograma por dia

SUMÁRIO

DEDICATÓRIA.....	4
RESUMO.....	Erro! Indicador não definido.
ABSTRACT	8
LISTA DE TABELAS	9
LISTA DE GRÁFICOS.....	11
LISTA DE FIGURAS	12
LISTA DE ABREVIATURAS E SIGLAS	15
LISTA DE SÍMBOLOS.....	17
SUMÁRIO.....	18
1 INTRODUÇÃO.....	20
1.1 Aspectos moleculares da fisiologia tiroidiana	21
1.2 Hipotireoidismo.....	23
1.3 Manifestações clínicas da deficiência dos HT	23
1.4 Diagnóstico laboratorial do hipotireoidismo	25
1.5 Tratamento com reposição hormonal no hipotireoidismo	26
1.6 Diagnóstico por biologia molecular	26
1.7 Sequenciamento de nova geração	26
1.8 Transcriptoma.....	27
1.9 Biotipos de RNA	28
1.9.1 Genes mitocondriais	28
1.9.2 Metilômica	29
1.9.3 Metabolômica	30
2 JUSTIFICATIVA.....	31
3 OBJETIVOS.....	31
4 MATERIAIS E MÉTODOS	32
4.1 Coleta de RNA circulante e exames complementares.....	33

4.2	Extração do RNA total	34
4.3	Análises bioquímicas	34
4.4	Execução do transcriptoma	35
4.5	Bioinformática	35
4.6	Análise bioestatística	36
4.7	Bibliotecas externas.....	37
5	RESULTADOS	40
5.1	Comparação do melhor método analítico qualitativo e quantitativo para identificação de genes diferencialmente expressos	42
5.2	Identificação de genes não codificantes no hipotireoidismo	68
5.3	Caracterização de genes mitocondriais diferencialmente expressos no hipotireoidismo e relação com o quadro clínico e laboratorial	72
5.4	Multiômica	88
6	DISCUSSÃO.....	90
7	CONCLUSÃO	96
	REFERÊNCIAS BIBLIOGRÁFICAS	97
	ANEXO 1.....	103
	ANEXO 2.....	106
	ANEXO 3.....	107
	ANEXO 4.....	114

1 INTRODUÇÃO

A glândula tiroide localizada a região cervical anterior, possui dois lóbulos, esquerdo e direito, separados por um istmo. A sua anatomia faz alusão ao formato de escudo, e sua origem etimológica vem do latim *thyreós*. Atualmente os termos tiroide e tireoide são cientificamente corretos e aceitos (1). A tiroide é uma glândula endócrina que sintetiza dois hormônios denominados de hormônios tireoidianos (HT), o 3,5,3',5'-tetraiodo-L-tironina (tiroxina ou T4) e o 3,5,3'-triiodo-L-tironina (triiodotironina ou T3), de funções sistêmicas, modulam de forma direta e indireta os processos fisiológicos primordiais para o desenvolvimento e manutenção do organismo, atuando de forma efetiva no metabolismo celular (2). A produção e secreção dos HT é realizada por um mecanismo de retroalimentação negativa, que envolve o hipotálamo, hipófise e tiroide (HHT), essa atuação neuroendócrina conhecida como eixo HHT fornece uma integração das complexas funções dos tecidos gerando respostas coordenadas às específicas necessidades do organismo (3). Frente às demandas metabólicas, o organismo sinaliza (via eixo HHT) os receptores das glândulas hipotálamo e hipófise a diminuição das concentrações plasmáticas circulantes de T3 e T4. Essa redução promove a secreção do hormônio liberador de tirotrofina (TRH) pelo hipotálamo diretamente ao sistema portal, sensibilizando a glândula hipófise à secreção do hormônio estimulador da tiroide (TSH) na circulação sistêmica, excitando às células foliculares da tiroide para sintetizar e secretar T3 e T4, para que em todos os tecidos esses HT possam promover sua ação intracelular que resumidamente resulta na regulação da velocidade metabólica do maquinário celular via ativação dos processos transcricionais representados na Figura 1 (4).

Para a síntese de T3 e T4 faz-se necessário a captação do iodo, íon adquirido pela alimentação, e nas células foliculares da tiroide o iodeto é adicionado ao núcleo tirosil da tiroglobulina, formando as iodotirosinas 3-moniodo-L-tirosina (MIT) e as 3,5-diiodo-L-tirosina (DIT), já que para a formação do T4 é necessário a ligação de duas moléculas de DIT e a do T3 a ligação de uma molécula MIT e uma DIT e posteriormente são armazenados no coloide, sendo secretadas posteriormente na corrente sanguínea proporções aproximadas de 80% de T4 e 20% de T3 (5).

1.1 Aspectos moleculares da fisiologia tiroidiana

O TSH é uma glicoproteína heterodimérica, com massa molecular de 28 kDa composta por duas subunidades designadas alfa e beta e sua síntese na hipófise é ativada pela ação do TRH (6). A produção dos HT é regulada principalmente pelo hormônio TSH via ligação em receptores de TSH (TSHR) nos tirócitos foliculares, essa ativação dos TSHR proporciona uma ampla variedade de processos metabólicos nas sínteses dos HT, resultando na produção e liberação do T3 e T4 e seu transporte plasmático acontece pela ligação às proteínas plasmáticas tiroglobulina e albumina (7). O T3 é considerado a forma bioativa pois ele exerce o papel efetivo de ativar o metabolismo celular, enquanto o T4 necessita da ação das enzimas iodotironinas desidases (DIO) transformando-o em T3. Nas células o T3 deve se ligar aos receptores de triiodotironina (TR), que são membros da família de receptores nucleares e sua estrutura proteica consiste de diferentes domínios, ou seja, da ativação da função N-terminal função 1, domínio de ligação a molécula de DNA, região de dobradiça e domínio C-terminal função 2 (2). Os TR são codificados por dois genes; *TRH α* e *TRH β* e devido ao *splincing* alternativo e uso alternativo do promotor, o gene *TRH α* pode dar origem a seis isoformas; TR α 1, TR α 2, TRD α 1, TRD α 2, P46 e P28, o gene *TRH β* também pode produzir isoformas via uso de promotor alternativo, sendo elas: TR β 1, TR β 2 e TRH β , porém apenas os TR α 1, TR β 1, TR β 2 são receptores efetivamente úteis na ligação de um domínio ao DNA. Sabe-se que o T3 tem maior afinidade aos receptores TR α 1 do que no TR β 1 e a conexão do domínio do receptor com a ligação ao DNA que modula a transcrição de sequências específicas são conhecidas como elementos de responsivos ao hormônio tiroidiano (TRE) (8). A via clássica da atuação do T3 é genômica, embora estudos atuais já mostrem que ações não genômicas envolvendo diferentes vias de transduções de sinalizações dos HT ainda necessitam de melhor caracterização e aprofundamento das pesquisas (9, 10).

A entrada celular dos HT nos tecidos é dependente de proteínas específicas de membrana como o transportador de monocarboxilatos tipo 8 (MCT8) com alta expressão no fígado, rim, coração, cérebro, osso, placenta, pulmão e músculo

esquelético e o transportador de monocarboxilatos tipo 10 (MCT10) com alta expressão no intestino, fígado, músculo esquelético, coração e placenta (11, 12).

Outra ação primordial para a biotransformação dos HT é realizada pelas enzimas da família desidase (DIO) cuja função é converter o T4 para T3, e se classificam em desidases tipo 1 (DIO1) do tecido hepático e renal, desidases tipo 2 (DIO2) que atuam na hipófise, hipotálamo, cérebro, tecido adiposo marrom, cóclea, osso e músculos esqueléticos e as desidases tipo 3 (DIO3) normalmente presentes no sistema nervoso central, placenta, pele e tecido fetal, que converte o T4 em T3 reverso (rT3) e T3 em 3,3'-diiodo-L-tironina (T2) limitando assim a ação dos HT (13). Atualmente, há evidências convincentes de que o controle dos HT realizados pelas desidases refletem significativamente na fisiologia dos diferentes órgãos e tecidos, garantindo ou não os efeitos do T3 de forma adequada em todas as fases da vida (14, 15).

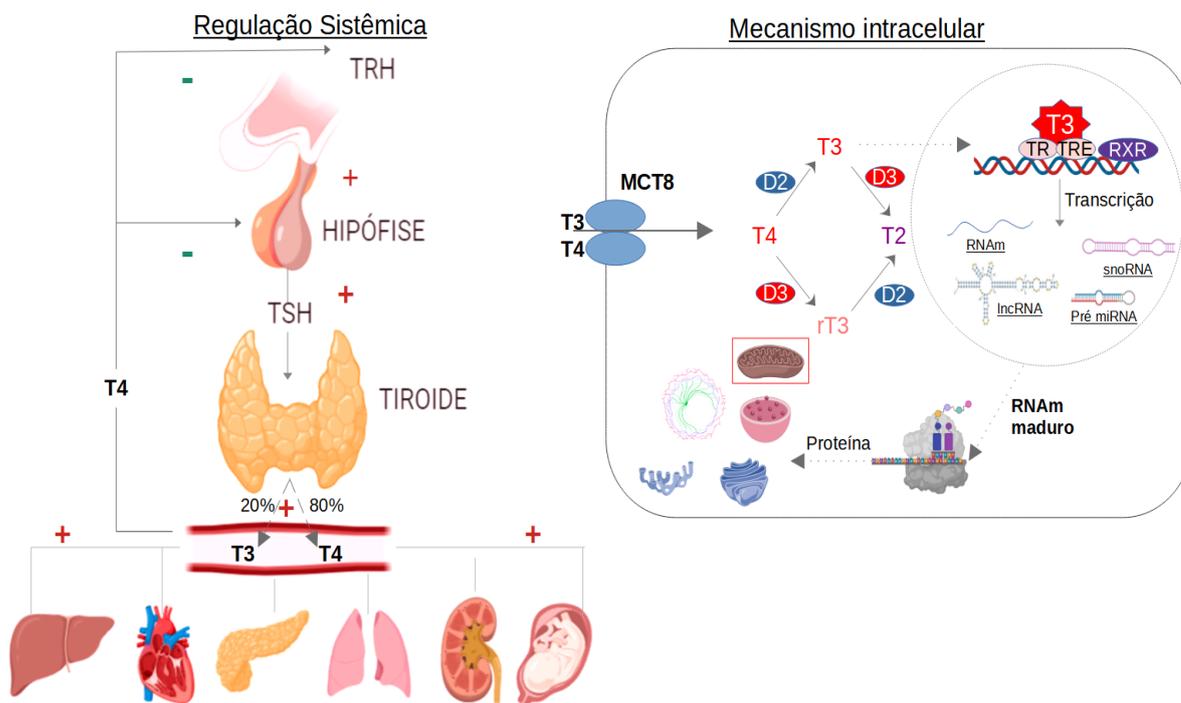


Figura 1. Ilustração do mecanismo de regulação da síntese do hormônio tireoideano e seus efeitos intracelulares para controle de expressão gênica e manutenção de organelas.

1.2 Hipotiroidismo

As doenças da tireoide, especialmente as autoimunes, atingem cerca de 5% da população mundial. As disfunções tiroidianas são detectadas 5 a 10 vezes mais em mulheres do que em homens, e a sua frequência aumenta com a idade (6). O hipotiroidismo é definido pela incapacidade da tireoide produzir HT suficientes para suprir as necessidades metabólicas do organismo, e pode acontecer por falha primária da glândula ou por deficiente estimulação do TSH, decorrentes de anormalidades congênitas, destruição autoimune (Tiroidite Crônica de *Hashimoto*), deficiência de iodo ou doenças infiltrativas (5, 6).

1.3 Manifestações clínicas da deficiência dos HT

O hipotiroidismo pode resultar numa ampla apresentação de sinais clínicos e sintomas, muitas vezes inespecíficos, sendo os mais comuns: bradicardia, pele grossa e seca, letargia, fala lenta, fraqueza, adinamia, diminuição da sudorese, pele fria, edema palpebral, intolerância ao frio, macroglossia, edema facial, mixedema, cabelo seco e sem brilho, aumento da área cardíaca (notável apenas na imagem radiológica), palidez da pele, perturbações da memória, constipação, ganho de peso, perda de cabelo, dispneia, rouquidão, anorexia, nervosismo, menorragia, surdez, palpitações, abafamento de bulhas cardíacas, dor precordial, baixa acuidade visual, entre muitos outros. A característica inespecífica dificulta a determinação da origem exata desses sintomas, entretanto, dada a influência global da glândula sobre o organismo, é possível atribuir a preponderância dos distúrbios da produção dos HT à uma piora desses efeitos (16).

O coração é particularmente vulnerável à diminuição sérica do T3, sendo esse hormônio essencial para a morfologia e bom desempenho cardíaco na vida adulta (17). A relação do HT e o sistema cardiovascular é um fenômeno bem estabelecido, sendo que uma disfunção tiroidiana pode levar o indivíduo a um risco elevado de desenvolver insuficiência cardíaca (18-20). A recomendação é de que os médicos precisam examinar os resultados dos exames da função tiroidiana como uma variável contínua de modo semelhante aos fatores de riscos coronarianos atualmente aceitos, tais como elevação da glicemia, pressão arterial e colesterol (21).

Nas repercussões para o sistema nervoso, o hipotireoidismo é uma causa frequente de alteração do humor, da cognição, da memória e depressão sendo uma causa comum, porém reversível quando tratadas em adultos e idosos (22).

Nos estudos relacionados ao metabolismo ósseo, as ações do T3 são complexas e ainda parcialmente compreendidas. Sabe-se que o T3 estimulam vias diretas e indiretas do ciclo da remodelação óssea, reabsorção, formação, proliferação de osteoclastos, diferenciação celular, apoptose, aumenta a expressão de osteocalcina, colágeno tipo 1, fosfatase alcalina e metaloproteínas, além disso, o T3 atua de forma sinérgica com os hormônios osteoclastogênicos como o hormônio da paratiroide (PTH) e vitamina D (23).

Em idosos as disfunções tireoidianas são amplamente prevalentes, especialmente associadas a outras morbidades onde os sinais e sintomas de disfunções tireoidianas são mais sutis ou ausentes em idosos, tornando o diagnóstico mais difícil, onde normalmente apresentam alterações cardiovasculares (fibrilação atrial, insuficiência cardíaca e outros), osteoporose, oftalmopatias, deficiência cognitiva, nódulos benignos ou malignos na tiroide (24).

Estudos demonstraram que no hipotireoidismo existe uma diminuição na capacidade renal, caracterizados pela redução da reabsorção de sódio dos segmentos distais e proximais dos túbulos renais em muitos relatos de casos documentam o aumento dos níveis séricos de creatinina no hipotireoidismo (25).

Nas mulheres o HT está envolvido no controle do ciclo menstrual, bem como na fertilidade, e o T3 possui ação específica sobre os ovócitos e sua deficiência pode afetar todos os aspectos da reprodução, sendo o segundo distúrbio endócrino mais comum entre as mulheres em idade reprodutiva e normalmente quando não tratada durante a gestação, aumentam significativamente os riscos de aborto espontâneo, descolamento prematuro da placenta, pré-eclâmpsia e restrição do crescimento fetal (26, 27).

A síndrome do doente eutiroidiano é muito observada em várias doenças agudas ou crônicas como um impactante fator de risco para o agravamento ou mortalidade em pacientes com doenças graves, com menor ação e concentração triiodotironina (T3). Essa condição é observada ao nível celular e tecidual e não é um distúrbio associado a função tireoidiana ou ao eixo HHT (28).

1.4 Diagnóstico laboratorial do hipotireoidismo

A melhor avaliação laboratorial da função tireoidiana para o diagnóstico de hipotireoidismo primário é a dosagem do TSH, já que a interação TSH, T3 e T4 pode facilitar uma abordagem mais individualizada, frente as amplas variações interindividuais de cada marcador. Quando os níveis plasmáticos do TSH estão elevados, o exame seguinte a ser realizado é a dosagem da fração livre do T4 no plasma, pois o hipotireoidismo primário é caracterizado pela elevação da concentração sérica do hormônio TSH e pela diminuição da concentração do hormônio T4 livre, podendo também apresentar outra situação distinta, quando as concentrações de TSH estão aumentados e as concentrações de T4 livre estão normais, caracterizando o hipotireoidismo subclínico (29). Os valores de referência para o TSH sérico considerados como referências de normalidade, são 0,45 mU/L à 4,5 mU/L para a maioria das populações adultas, exceto na gestação, infância e adolescência. A repetição do teste para dosagem do TSH é importante para descartar a possibilidade de aumentos transitórios e resultados entre 4,6mU/L à 10mU/L normalmente tendem a se normalizar nas repetições do exame (16).

Sobre o hipotireoidismo subclínico a prevalência é especialmente mais alta em indivíduos caucasianos, mulheres, idosos e em populações com alta ingestão de iodo. O debate ainda é amplo sobre a melhor intervenção, de um lado, uma frente considera a reposição exógena de levotiroxina (L-T4), e do outro lado, não. Estudos clínicos randomizados sobre o efeito da reposição de L-T4 no hipotireoidismo subclínico não conseguiram demonstrar evidências consistentes na redução da morbidade cardiovascular, melhor qualidade de vida ou melhora nos sintomas nos grupos tratados (30).

A dosagem dos anticorpos antitireoidianos: Anticorpo Anti-Tiroperoxidase - TPOAb e Anticorpo Anti-Tiroglobulina – TgAb auxilia na definição etiológica da tireoidite autoimune, sendo o TPOAb o anticorpo mais prevalente e sensível (31).

1.5 Tratamento com reposição hormonal no hipotireoidismo

A reposição de L-T4 é o tratamento de escolha para o hipotireoidismo, em média a L-T4 tem uma meia vida de 7 dias. Convencionalmente os pacientes fazem uso da levotiroxina pela manhã meia hora antes do café com o estômago vazio a fim de evitar o comprometimento da absorção do fármaco com alimentos, sendo assim o padrão de cuidados para o tratamento do hipotireoidismo (32-34).

1.6 Diagnóstico por biologia molecular

Os avanços na biologia molecular revelam que pelo menos 87,3% do genoma humano são ativamente transcritos e que menos de 3% desses transcritos efetivamente codificam proteínas, essa conclusão foi uma surpresa, e sugere que a maior parte do genoma são transcritos na forma de ácido ribonucleico (RNA) não codificantes (35).

Estudos no âmbito da biologia molecular foram revelados que a comunicação intercelular não é realizada apenas pelas vias clássicas já conhecidas como citocinas, fatores de crescimento, neurotransmissores ou hormônios em geral, mas também através da troca de pequenas moléculas de RNA, e essas moléculas podem estar inseridas em pequenas estruturas chamadas exossomos, que são constituídos de proteínas, ácido ribonucleico mensageiro (mRNA) e micro ácido ribonucleico (miRNA) ou até mesmo na forma livre no plasma, os quais podem ser utilizados como estratégia diagnóstica não invasiva com potencial de avaliar o funcionamento dos tecidos (36, 37).

1.7 Sequenciamento de nova geração

O primeiro sequenciamento completo do genoma humano demorou 12 anos para ser concluído a um custo de quase 3 milhões de Euros (38). O esforço envolveu centenas de pesquisadores do mundo inteiro, com a única plataforma disponível na época para sequenciamento, o *Sanger* de primeira geração (39). O término desse projeto ambicioso aconteceu em 2002 e de lá pra cá o desenvolvimento de novas plataformas de sequenciamento se tornaram totalmente necessárias, e sua evolução foi vertiginosa (40).

Desde a década de 1990, a matriz de micro arranjos (*microArray*) de ácido desoxirribonucleico (DNA) tem sido a tecnologia de escolha para estudos em larga escala em níveis de expressão de genes (41). A capacidade dessas matrizes em analisar simultaneamente milhares de transcrições levou a significativos avanços das amplas variedades de questões e problemas biológicos incluindo expressão gênica entre indivíduos saudáveis e doentes ou respostas farmacogenômicas (42, 43). Durante os últimos anos a tecnologia utilizada para análise de expressão gênica mudou drasticamente, os “velhos” *microArrays* vêm sendo substituídos por plataformas de sequenciamento de nova geração (NGS). O único aspecto negativo sobre a tendência NGS é a análise e interpretação dos resultados, pois pode ser morosa e dependente de uma alta tecnologia computacional e bioinformatas capacitados (44).

A investigação clínica por NGS fornece uma alta sensibilidade em uma plataforma de alto rendimento, produzindo dados torrenciais para uma ampla análise genômica (45, 46). Os custos de sequenciamento NGS diminuíram nos últimos anos, e o sequenciamento do genoma agora estão relativamente dentro das possibilidades dos custos dos exames médicos (47-49).

1.8 Transcriptoma

O transcriptoma é a avaliação simultânea da expressão de transcritos em larga escala, e sua expressão reflete especificamente uma condição fisiológica ou fisiopatológica, facilitando a compreensão dos constituintes moleculares de uma célula ou tecido, assim, conseguimos por ele entender o desenvolvimento de doenças, possibilitando catalogar as sequencias de transcritos, avaliar agentes de modificações pós transcricionais e quantificar as alterações dos níveis de cada expressão sob diferentes condições (50).

O transcriptoma engloba o sequenciamento de todos os tipos de moléculas de RNA, a saber: miRNA, mRNA, scaRNA, snoRNA e lncRNA. Quando analisamos de maneira conjunta, os RNAs refletem genes que estão ativamente expressos em um tecido em particular (51).

1.9 Biotipos de RNA

A molécula de RNA tem biotipos e funções que vão muito além da codificação de proteínas por mRNA, e fazem parte de um leque de vários biotipos com características bem específicas em seus mecanismos pois transcritos não codificantes interferem de forma regulatória toda constituição celular (52).

Os micros RNAs (miRNA) possuem em média de 23 a 30 nucleotídeos e silenciam processos transcricionais e pós transcricionais e são encontrados em biofluidos extracelulares na forma de exossomos, que são vesículas que transportam essas moléculas e são fagocitadas por outros tecidos tendo como alvo os mRNA para o processo de silenciamento, e atualmente são explorados pela pesquisa científica como biomarcadores de doenças (53).

Um outro biotipo de RNA são os longos não codificantes (lncRNA), constituído por mais de 200 nucleotídeos, não codificam proteínas mas participa de forma regulatória na fisiologia celular, como a apoptose, carcinogênese, respostas imunológicas, diferenciação muscular e desenvolvimento neuronal, onde podem se integrar a proteínas heteronucleares, muitas vezes modificando a cromatina, desestabilizam o mRNA ou promovem a fosforilação de proteínas (54). Também descritos, temos os RNAs longos não codificantes intergênicos (lincRNA) que também tem características semelhantes aos lncRNAs, pois se acumulam no citoplasma servindo como “isca” de miRNAs, ou se armazenam no núcleo participando da regulação gênica remodelando a cromatina e modificações epigenéticas atuando como ativadores ou repressores transcricionais. Os genes snoRNA são responsáveis por controlar a ordem de *splicing* e modular a saída do mRNA do núcleo para o citoplasma, sendo importantes reguladores pós transcrição (55).

1.9.1 Genes mitocondriais

As mitocôndrias surgiram de uma alfa-pretobactéria fagocitada por uma célula eucarionte, sendo constituídas de duas membranas: externas e internas separadas, o seu genoma é caracterizado como circular (mtDNA), sintetizam mais de 1500 proteínas conforme o tecido específico, e tem essencial função de produzir energia para a célula na forma de adenosina trifosfato (ATP), intermediários biossintéticos, contribuem para

respostas contra o estresse celular, autofagia, apoptose e comunicação celular. Suas disfunções se caracterizam em doenças genéticas, como neurodegenerativas ou metabólicas (56).

A essencial importância das mitocôndrias faz com que elas possuam um rigoroso controle de qualidade regulados pela dinâmica funcional mitocondrial, mitofagia e fusão, que permitem que mitocôndrias danificadas sejam segregadas, facilitando a equalização de seus componentes como o DNA, proteínas e metabólitos, estes processos são as principais vias deste controle para reciclar os danos nesta organela (57).

As mitocôndrias regulam seus mecanismos de controle de qualidade afim de manter a homeostase energética do tecido em qualquer condição, muito embora em como já caracterizado, algumas doenças metabólicas levam à incapacidade das mitocôndrias a se adaptarem às alterações que são denominadas de inflexibilidade metabólica. Fenômeno presente na obesidade e diabetes mellitus tipo 2 (58). Embora o genoma mitocondrial seja transcrito e sintetizado em proteínas dentro das mitocôndrias, quase todas as proteínas necessárias para estes processos são sintetizadas pelo genoma nuclear, demonstrando uma forte ligação entre estes dois genomas (59).

1.9.2 Metilômica

A metilação é uma forma de interação química de um grupo metil ao DNA, sem que aconteça uma mudança estrutural nas sequências de nucleotídeos, conhecida também como mecanismo epigenético. O *Whole Genome Bisulfite Sequencing* (WGBS) é uma técnica capaz de rastrear as regiões promotoras dos genes com maior nível de metilação através das ilhas CpGs (*CpGs Islands*) e dessa forma podemos inferir a tendência de transcrição e atividade do RNA codificado, que de forma geral podem influenciar na estrutura da cromatina e regular a atividade de um determinado gene (60).

1.9.3 Metabolômica

A metabolômica é uma técnica que avalia a concentração de moléculas e substratos em uma determinada situação de interesse. As moléculas podem ser os produtos de reações ou componentes essenciais para que as mesmas ocorram. São capazes de proverem informações finais de um processo metabólico (*end-point*), ou seja, informações que indicam a funcionalidade de um processo biológico (61). Essas informações podem indicar se uma via biológica se desenvolveu na forma esperada ou indicar a possível presença de algum mecanismo de controle que antes não era observado. Nós utilizamos dados publicados em artigos revisados por pares para comparar com os nossos dados transcriptômicos. A comparação foi feita com o uso da plataforma *Integrated Molecular Pathway Level Analysis* (IMPaLA) (62).

2 JUSTIFICATIVA

A identificação de marcadores moleculares tem se tornado alvos importantes para o diagnóstico e prognóstico de doenças, em especial pela relativa facilidade de aquisição de material biológico. Em nosso trabalho o sangue periférico foi o material escolhido, dada à capacidade de amplitude de informações que o transcriptoma de nova geração pode nos proporcionar diante das disfunções tiroidianas que são altamente prevalentes no Brasil e no mundo. Produzir um painel de genes mitocondriais que refletem a disfunção tecidual em pacientes com hipotireoidismo é uma ferramenta útil que pode favorecer um manejo pontual das disfunções específicas objetivando um melhor prognóstico.

3 OBJETIVOS

Este trabalho foi dividido em dois principais objetivos:

Objetivo primário: Caracterização dos genes mitocondriais diferencialmente expressos no hipotireoidismo e sua correlação com o quadro laboratorial por análises dos transcriptomas, metabolômas, metilômas e identificação de TREs

Objetivo secundário: Identificação de genes diferencialmente expressos buscando qualitativamente e quantitativamente identificar o melhor método analítico em diferentes ferramentas e bibliotecas.

4 MATERIAIS E MÉTODOS

No período de 1 ano e 6 meses (2014 – 2016) foram atendidos no ambulatório de endocrinologia da Universidade Nove de Julho 320 pacientes adultos, dos quais, 120/320 se enquadraram no perfil da pesquisa e se voluntariaram ao estudo, obedecendo os trâmites legais como a assinatura do termo de consentimento livre e esclarecido (TCLE) com aprovação do comitê de ética em pesquisa da Universidade Nove de Julho de número 665.331 – 28/05/2014. Após uma análise criteriosa do perfil clínico e laboratorial de cada voluntário (120 indivíduos) chegamos à definição dos melhores indivíduos que atendiam os objetivos da pesquisa para a realização do transcriptoma, sendo 4 para o grupo eutiroidianos e 4 para o grupo hipotiroídianos que nunca foram tratados com L-T4 (CTL x HT), concretizando o número (n) de 8 indivíduos, sendo 6 mulheres e 2 homens.

Realizamos também outro transcriptoma com indivíduos dos seguintes perfis: 4 hipotiroídianas que fazem uso de LT-4 e estavam no alvo terapêutico com TSH em níveis normais formando o grupo L-T4 no alvo terapêutico (LEVAT) e 4 (3 mulheres e 1 homem) hipotiroídianas(o) que fazem uso de L-T4 e estão fora do alvo terapêutico com TSH em níveis alterados formando o grupo L-T4 fora do alvo terapêutico (LEVOUT).

Numa segunda etapa entre 2017 e 2019, também por atendimento médico no ambulatório de endocrinologia da Universidade Nove de Julho e no Complexo Hospitalar Mandaqui, respectivamente sob os números de aprovação do comitê de ética em pesquisa 2.136.486 e 2229867, aderimos 58 indivíduos crianças e adolescentes que obedeceram à análise de exames laboratoriais e clínicos e necessariamente tinham que obter a dosagem de TSH dentro dos limites de referência do método utilizado em duas dosagens em período inferior a 60 dias (dentro das suas faixas etárias), assinaram o termo de assentimento livre e esclarecido e seus responsáveis assinaram o TCLE, e assim foram constituídos os grupos hipotiroidismo e eutiroidianos para análise de validação de genes diferencialmente expressos nesta população, baseada nos resultados das bibliotecas de hipotiroídianos construídas neste trabalho.

A adesão à pesquisa para todos os participantes que formaram o grupo controle (grupo de pessoas eutiroidianas) obedeceu à análise de exames laboratoriais e

clínicos e necessariamente tinham que obter a dosagem de TSH dentro dos limites de referência do método utilizado em duas dosagens em período inferior a 60 dias.

A adesão à pesquisa para todos os participantes que formaram o grupo hipotiroídianos primários (grupo de pessoas hipotiroídianas que nunca foram tratadas com L-T4) obedeceu à análise de exames laboratoriais e clínicos e necessariamente tinham que obter a dosagem de TSH superiores a 10 μ UI/mL do método utilizado em duas dosagens em período inferior a 60 dias.

A adesão à pesquisa para todos os participantes que formaram o grupo hipotiroídianos no alvo terapêutico (grupo L-T4 no alvo terapêutico) obedeceu à análise de exames laboratoriais e clínicos e necessariamente tinham que obter a dosagem de TSH dentro dos limites de referência do método utilizado em duas dosagens em período inferior a 60 dias.

A adesão à pesquisa para todos os participantes que formaram o grupo hipotiroídianos fora do alvo terapêutico (grupo L-T4 fora do alvo terapêutico) obedeceu à análise de exames laboratoriais e clínicos e necessariamente tinham que obter a dosagem de TSH superiores a 10 μ UI/mL do método utilizado em duas dosagens em período inferior a 60 dias.

4.1 Coleta de RNA circulante e exames complementares

O procedimento de coleta seguiu as recomendações da Sociedade Brasileira de Patologia Clínica e Medicina Laboratorial para coleta de sangue venoso, 2009, 2ª Edição. A venopunção foi realizada com escalpe de segurança *push button* em agulha 25G (Calibre 5) e canhão (tubo de vinil) de 7 polegadas de comprimento. Na punção venosa, a colheita dos tubos seguia a seguinte ordem: 1º tubo para preservação do sangue total, contendo EDTA, colhidos 4 mL de sangue; 2º tubo para ativação de coágulo e obtenção do soro sanguíneo, contendo ativador de coágulo e gel separador, colhidos 5mL de sangue; 3º tubo para preservação das moléculas de RNA (*PAXgene blood RNA*), contendo 6,9 mL de aditivo de preservação de RNA, colhidos 2,5 mL de sangue em aspiração à vácuo. A utilização de escalpe obedeceu a instrução do fabricante (*Qiagen, NL, DE*), de forma que o tubo *PAXgene blood RNA* ficasse em posição vertical durante o procedimento para evitar refluir o aditivo na veia puncionada.

O tubo com sangue total após suave homogeneização foi armazenado imediatamente em temperatura refrigerada (2° à 8°C) a fim de mantermos material para uma eventual análise molecular do RNA/DNA. O tubo para obtenção do soro sanguíneo, foi submetido à centrifugação em 1500 RPM (rotações por minuto) por 15 minutos, separando pelo gel a fase sólida (células e coágulo) da líquida (soro) e armazenado em congelamento de -20° C para posterior realização dos exames TSH, T4 Livre, Anti TGAb, Anti TPOAb. No tubo *PAXgene* blood RNA após a aspiração por vácuo de 2,5mL de sangue, realizava-se a homogeneização com o aditivo por inversão em 180° por 10 vezes e deixamos o tubo exposto por 24 horas em temperatura ambiente (15°C à 25°C), para que, de acordo com as recomendações do fabricante, houvesse a completa lise celular, otimizando assim o rendimento das moléculas de RNA ao final da extração, e após esse período, foi armazenado em congelamento a -20°C para posterior realização do transcriptoma.

Os pacientes que atendiam aos critérios de inclusão da pesquisa realizaram inicialmente a dosagem plasmática do hormônio estimulador da tireoide (TSH), dosagem plasmática do hormônio tiroxina (T4 Livre), dosagem plasmática dos anticorpos anti-tiroglobulina (Anti TGAg), dosagem plasmática dos anticorpos anti-peroxidase da tireoide (Anti TPOAg), e selecionados dentro do perfil de nossa pesquisa, realizaram o transcriptoma do RNA total, a partir do seu sangue periférico.

4.2 Extração do RNA total

O RNA total obtido a partir do sangue periférico foi extraído pelo *kit* do fabricante Qiagen (NL, DE), *PAXgene blood RNA kit*, (Ref. 762174).

A extração do RNA preservado seguiu o protocolo do fabricante, conforme o anexo 1.

4.3 Análises bioquímicas

As análises bioquímicas de TSH, T4 Livre, Anti TPOAb, Anti TGAg foram executadas no equipamento *Elecsys 2010* do fabricante *Roche Diagnostics* (IN, USA),

com os *Kits* do mesmo fabricante, seguindo protocolos automatizados específicos para cada teste, de acordo com o anexo 2.

Os intervalos de referência e a linearidade dos *kits* seguem os seguintes valores: TSH valores de referência >20 anos 0,270 - 4,50 $\mu\text{UI/mL}$ e intervalo de detecção do *kit* 0,005 - 100 $\mu\text{UI/mL}$, T4 Livre valores de referência 0,93 - 1,7 ng/dL e intervalo de detecção do *kit* 0,23 - 77,68 ng/dL, Anti TGAg valores de referência <115 UI/mL e intervalo de detecção do *kit* 10,0 - 4000 UI/mL, Anti TPOAg valores de referência < 34 UI/mL e intervalo de detecção do *kit* 5,00 - 600 UI/mL.

4.4 Execução do transcriptoma

A realização do transcriptoma foi realizada pelo equipamento sequenciador do fabricante *Thermo Fisher Scientific* (MA, USA), *Ion Proton System*, através dos protocolos *Ion AmpliSeq Transcriptome Human Gene Expression Kit*, pelo catálogo número: A26325, A26326 e A26327. Todas as etapas: quantificação do RNA total, amplificação das sequências alvos, preparo das sequências iniciadores, ligação dos adaptadores (*barcodes*) aos *amplicons*, purificação da biblioteca, quantificação da biblioteca (qPCR), realização da PCR emulsão, enriquecimento das bibliotecas com as esferas magnéticas (*beads*), calibração do sequenciador e preparo dos chips estão detalhadamente descritos no anexo 3.

Toda biblioteca constituída nesse protocolo, como também as amostras extraídas de RNA total, estão armazenadas em nossa soroteca em *freezer* -80°C.

4.5 Bioinformática

Os dados gerados pelo *Ion Proton System* são em arquivos de extensão BAM (*Binary Alignment Map*), e são armazenados por *upload* no servidor *online* do fabricante da plataforma NGS. Executamos o *download* dos dados e os arquivamos em plataformas locais de fácil acesso e em *hard disks* portáteis. As análises de dados foram realizadas pelo R *Statistics* versão 4.0.2 (63) com a interface do RStudio versão 0.99.49 (64). Para a expressão diferencial foram utilizados os pacotes *edgeR* versão 3.32.0, *DESeq2*

versão 1.20.0 e NOISEq versão 2.34.0 disponibilizado pela *Bioconductor open source software for bioinformatics* (65-67).

Dos arquivos BAM obtidos da máquina, extraímos as contagens pela função *idxstats* disponível no *software Samtools* versão 1.2 (68) com o terminal padrão disponível no sistema operacional Ubuntu versão 14.04. As contagens foram armazenadas em um arquivo .txt, de forma que cada coluna representava um paciente e cada linha representava um gene, totalizando 22.786 transcritos (linhas) formando uma matriz com 8 pacientes (colunas). A anotação utilizada foi a RefSeq (69).

A identificação e caracterização dos mitRNAs foi produzida pela base de dados *MitoMiner* 4.0 (70), ferramenta *online* que fornece banco de dados de evidências de localização mitocondrial em mamíferos, fenótipos e doenças, em conjunto com a ferramenta *online* MITOCARTA (71) que fornece um inventário de genes mitocondriais de mamíferos. Utilizamos a lista oficial do ano de 2020 contém 1329 genes genuinamente mitocondriais. No anexo 4 estão presentes todos os *pipelines* para as produzidos para as análises em R.

A identificação das sequências de TREs foram realizadas pela consulta de sequências de nucleotídeos de referência FASTA no NCBI.

O enriquecimento de análises genéticas foram realizadas pelos *softwares* *FunRich* versão 3.1.3, *Cytoscape* versão 3.8.0 e *GSEA* versão 4.0.3 (72-74).

4.6 Análise bioestatística

A análise bioestatística foi realizada com o software R versão 4.0.2 (63) para o cálculo das medianas, percentis e valores máximos e mínimos. A representação da correlação *spearman* foi feita pelo pacote *corrplot* Versão 0.84 (75).

Para produção dos diagramas de intersecção *Venn*, utilizamos a ferramenta online *InteractiVenn* (76).

4.7 Bibliotecas externas

Para métodos comparativos foram utilizadas bibliotecas de estudos publicados na base do Centro Nacional de Informações sobre Biotecnologia (NCBI - <https://pubmed.ncbi.nlm.nih.gov/>), disponíveis sob os números GSE103305, GSE124326, GSE131411 e GSE127306 (32, 77, 78).

GSE103305 *"Thyroid State Regulates Gene Expression in Human Whole Blood Cells"*

O presente estudo caracteriza-se por um transcriptoma realizado na plataforma Illumina HiSeq 2500. Devido à similaridade da condição avaliada, variações do hormônio tireoidiano e o material utilizado sangue total, este estudo foi considerado como viável para identificar genes comumente diferencialmente expressos e ampliar o nosso número de grupos avaliados.

GSE124326 *"Whole blood transcriptome analysis in bipolar disorder reveals strong lithium effect"*

Este estudo caracteriza-se por um transcriptoma, também realizado na plataforma Illumina HiSeq 2500, com o diferencial de um número maior, possuindo 240 pacientes controles e 240 pacientes acometidos pelo transtorno bipolar, totalizando 480. O objetivo principal foi traçar um perfil de expressão em desordem bipolar através do sangue total. Considerando os efeitos do hormônio tireoidiano no sistema nervoso central (SNC), este estudo foi considerado uma forma viável de avaliar os nossos achados em outras condições.

GSE131411 *"Whole blood transcriptome analysis of Septic shock and Cardiogenic shock patients over one week of ICU stay"*

Este estudo caracteriza-se por um transcriptoma, realizado na plataforma Illumina HiSeq 2500, com 21 pacientes hospitalizados na unidade de terapia intensiva (UTI), com choque séptico. O transcriptoma foi realizado pela coleta do sangue total em tempos diferentes: T1: 16 horas após admissão na UTI; T2: 2º dia e T3: 7º dia. Estas

bibliotecas foram incluídas em nosso estudo para avaliar a atividade dos mitGenes diante de doença em estado grave.

GSE127306 "*WGBS from thyroid gland (ENCSR601MHU)*"

Diferente dos bancos anteriores, o GSE127306 trata-se de um metiloma realizado em tecido tireoideano com o uso do Whole Genome Bisulfite Sequencing (WGBS) em uma mulher adulta com 51 anos de idade. No WGBS é possível rastrear as regiões promotoras dos genes com maior nível de metilação através das CpGs Islands (Ilhas CpGs), dessa forma, utilizamos este banco como uma estratégia para detectar o possível aumento ou diminuição da expressão futura do RNA. Quanto maior a concentração das ilhas CpGs, menor tende a ser a transcrição deste gene. O UCSC Genome Browser foi utilizado para confirmar o fenômeno regulatório observado.

Em metabolômica, utilizamos como referência os achados de *Piras, C. Et al* (79) para rastrear o perfil comportamental dos metabólitos que sofrem influência dos genes selecionados para estudo.

Uma síntese representativa de nosso trabalho é representada na figura abaixo:

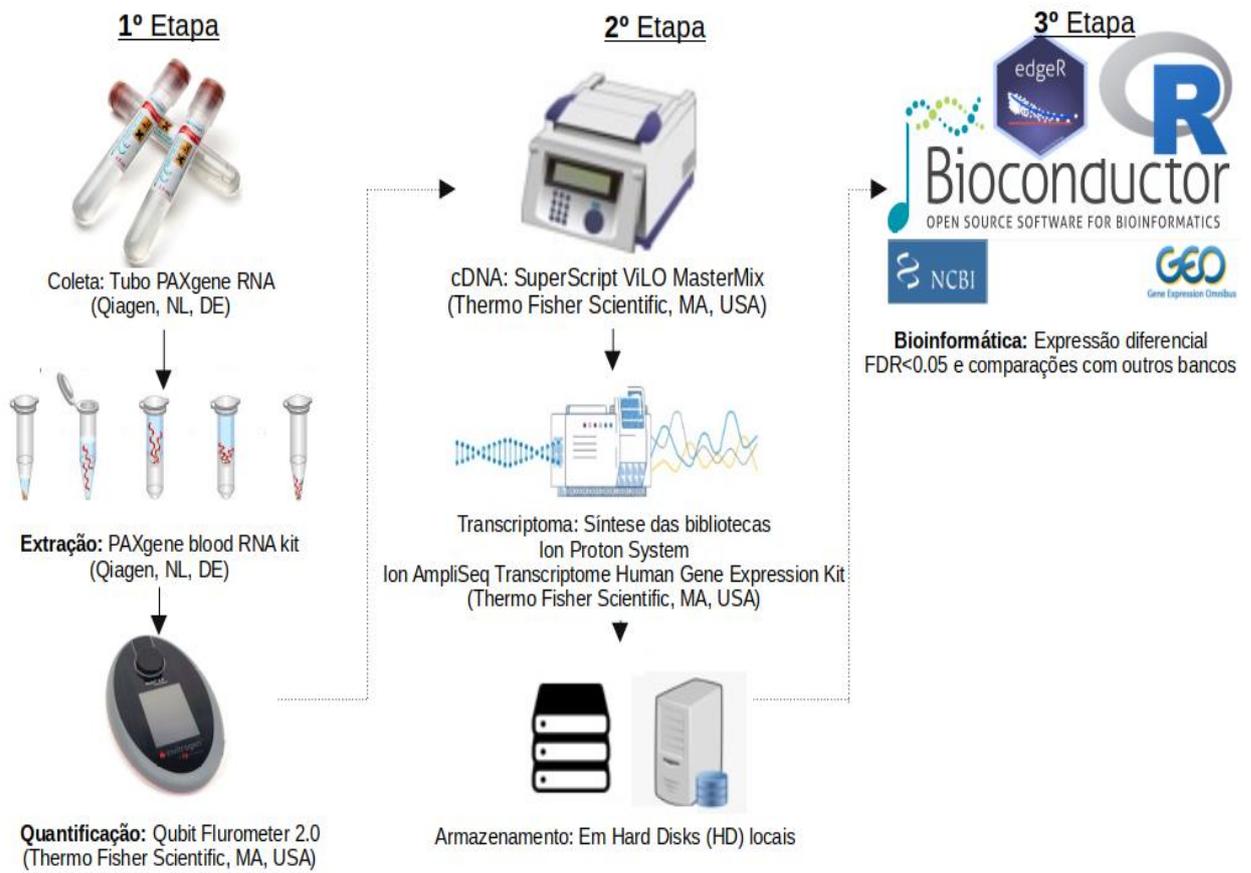


Figura 2. Fluxograma das etapas da pesquisa resumindo as estratégias adotadas para que fosse possível a realização dos objetivos propostos no presente trabalho.

5 RESULTADOS

Apresentamos as distribuições antropométricas, laboratoriais e clínicas dos grupos dos indivíduos selecionados para confecção das bibliotecas do transcriptoma nos indivíduos adultos com hipotireoidismo (HT - nunca tratados com reposição de L-T4) e indivíduos saudáveis (CTL - controle).

Tabela 1. Distribuições antropométricas, laboratoriais e clínicas dos grupos CTL e HT. Sexo, Idade: anos, Peso: Kg, TSH valores de referência >20 anos 0,270 - 4,50 UI/mL, T4 Livre valores de referência 0,93 - 1,7 ng/dL, TgAb valores de referência <115 UI/, TPOAb valores de referência < 34 UI/mL e Tireoidectomia.

Grupo	Identificação	Sexo	Idade	Peso	TSH	T4L	TgAb	TPOAb	Tireoidectomia
CTL	CTL1	Feminino	40	73,4	3	1,2	10	32	Não
	CTL2	Masculino	36	65,25	1,7	1,2	10	5	Não
	CTL3	Feminino	36	83,5	1,3	1,3	10	5	Não
	CTL4	Feminino	40	69,3	4,2	1,1	10	14	Não
HT	HT1	Feminino	38	62,9	65	0,4	668	27	Não
	HT2	Masculino	34	110,5	100	0,1	1195	515	Não
	HT3	Feminino	52	47,7	59	0,1	10	5	Sim
	HT4	Feminino	51	83,2	100	0,1	864	368	Não

Análises descritivas das distribuições dos grupos controle e hipotireoidismo.

Tabela 2. Valores de N, medianas, valor mínimo, máximo e percentis dos grupos Controle de Hipotireoidismo.

	GRUPO CTL						GRUPO HT						
	Idade	Peso	TSH	FT4	TgAb	TPOAb	Idade	Peso	TSH	FT4	TgAb	TPOAb	
Mediana	38	71,4	2,4	1,2	10	9,5	44,5	73,05	82,3	0,1	766	197,5	
Mínimo	36	65,3	1,3	1,1	10	5	34	47,7	58,7	0,1	10	5	
Máximo	40	83,5	4,2	1,3	10	32	52	110,5	100	0,4	1195	515	
25	36	66,3	1,4	1,1	10	5	35	51,5	60,2	0,1	174,5	10,5	
%	50	38	71,4	2,4	1,2	10	9,5	44,5	73,05	82,3	0,1	766	197,5
75	40	81	3,9	1,3	10	27,5	51,8	103,6	100	0,3	1112,3	478,25	

Os dados técnicos laboratoriais, moleculares e computacionais são necessários para comprovar a qualidade das fases pré e pós analíticas conforme.

Tabela 3. Dados moleculares e qualitativos pré, peri e pós confecções das bibliotecas no NGS. [RNA] concentração de ácido ribonucleico ideal: >10 ng/μL, % IQ [RNA] integridade e qualidade da [RNA] ideal : 1 - 10%, qPCR quantificação da reação em cadeia da polimerase ideal: >100pM, % *Beads* esferas magnéticas conjugadas ao DNA ideal: 10 - 30%, *Id Barcode*: Identificação da amostra, *Id Chip*: Identificação do *CHIP* de sequenciamento, Comprimento médio: pb (pares de bases), Leituras de fitas de DNA: %, *Output BAM* tamanho: Saída do arquivo de múltiplo alinhamento binário em GB (*Gigabytes*).

Grupo	Grupo	[RNA]	% IQ [RNA]	qPCR pré sequencimento	% Beads	Id Bar-code	Id Chip	Comprimento	Leituras	Output BAM
CTL	CTL1	42	6,2	9110	17	26	1	113	99,1	8,12
	CTL2	604	6,9	2230	18	33	2	108	98,8	5,08
	CTL3	682	7,5	750	18	34	2	108	98,8	3,06
	CTL4	882	6,7	1350	18	32	2	108	98,8	5,45
HT	HT1	23	6,7	1140	17	25	1	112	99,1	3,81
	HT2	14	7,2	5830	17	27	1	112	99,1	4,16
	HT3	98	6,9	790	18	36	2	109	98,8	3,27
	HT4	204	6,4	2010	18	35	2	108	98,8	5,09

5.1 Comparação do melhor método analítico qualitativo e quantitativo para identificação de genes diferencialmente expressos

Os resultados a seguir referem-se à comparação entre os pacotes EdgeR, DESeq2 e NOISeq na detecção de DGE nas bibliotecas do grupo HT, Hipotiroidismo leve e nas bibliotecas do estudo estado da tiroide.

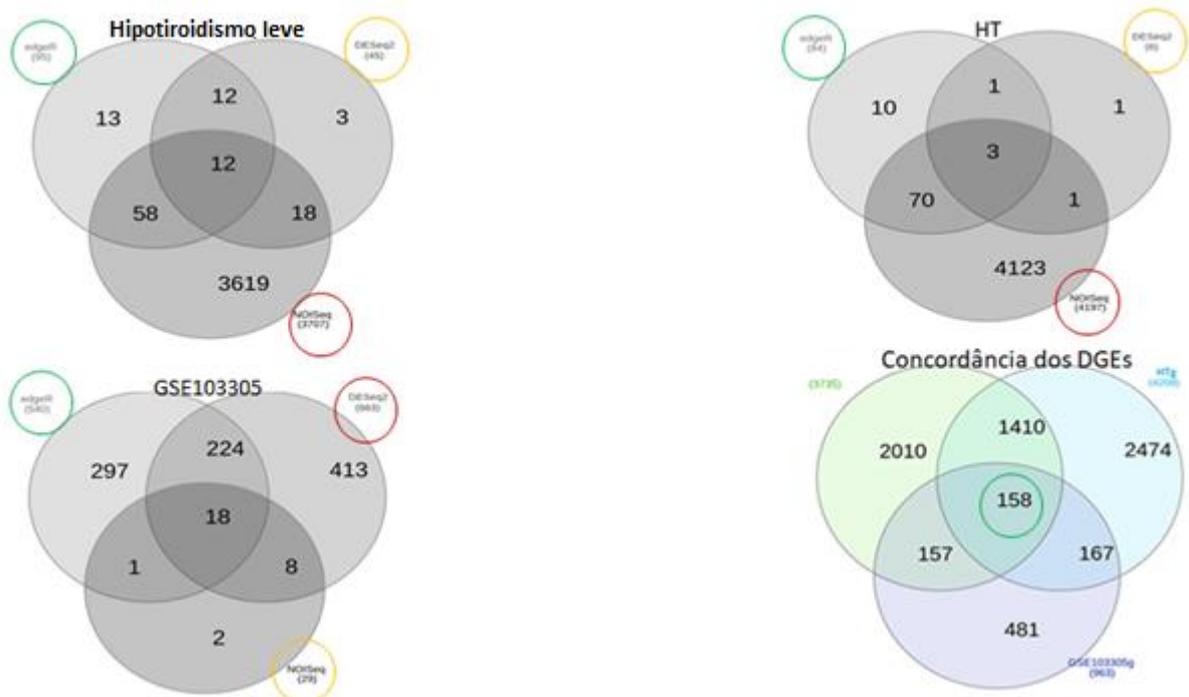


Figura 3. Comparação entre os pacotes EdgeR, DESeq2 e NOISeq, na detecção de DGE nas bibliotecas do grupo HT, Hipotiroidismo leve, GSE103305 e a concordância dos DGEs.

Os resultados a seguir referem-se à identificação de genes diferencialmente expressos, objetivando o melhor método analítico para maior detecção de DGE nas bibliotecas dos grupos CTL e HT e nas bibliotecas do estudo estado da tiroide e transtorno bipolar. Em todas as análises utilizamos filtro para *reads* >0.

Iniciamos a análise por dispersão comum e teste exato.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão comum e teste exato nas bibliotecas do grupo CTL e grupo HT, que produziu a lista de 100 DGE, sendo 14920 não significantes, 70 genes com aumento de expressão e 30 genes com diminuição de expressão.

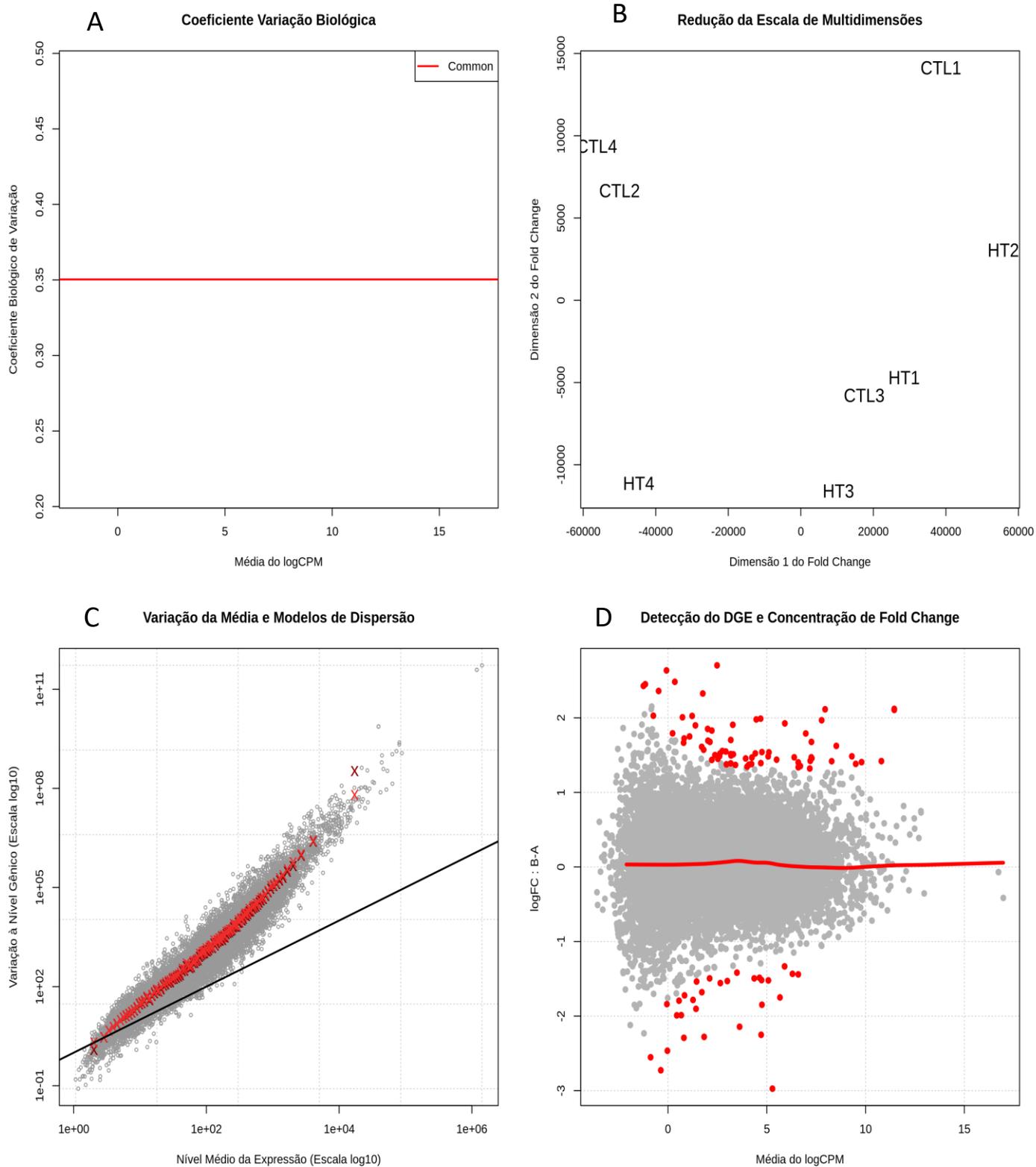


Figura 4. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*.

Após os resultados da estratégia em dispersão comum x teste exato, iniciamos a segunda análise de dispersão comum, dispersão por tendência e teste de hipótese por razão de verossimilhança.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia Dispersão comum, Dispersão de tendência e teste de razão de verossimilhança nas bibliotecas do grupo CTL e grupo HT, que produziu a lista de 105 DGE, sendo 14916 não significantes, 74 genes com aumento de expressão e 31 genes com diminuição de expressão.

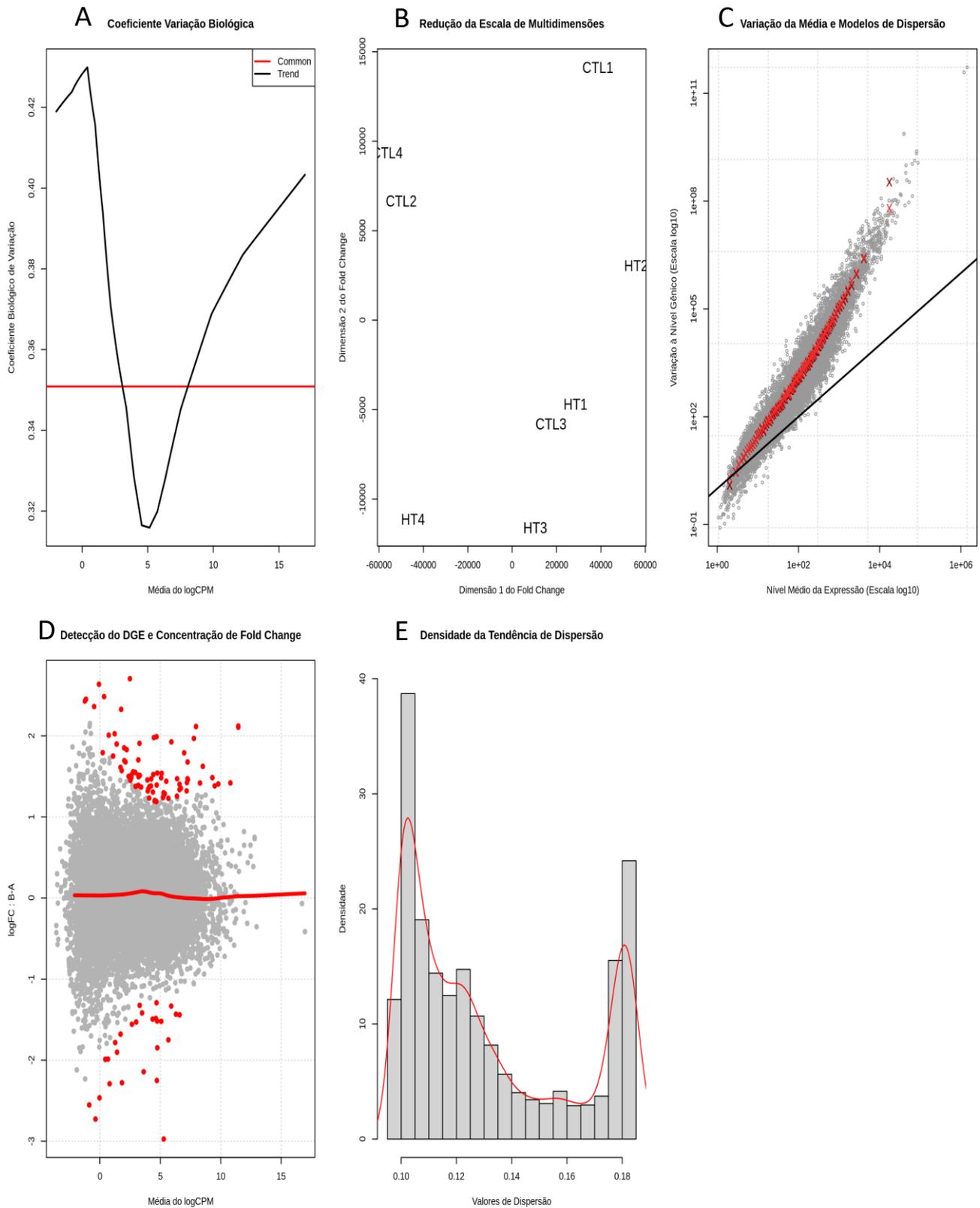


Figura 5. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão.

Com os resultados anteriores iniciamos a próxima etapa de análise que é dispersão de tendência e teste exato.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia Dispersão de tendência e teste exato nas bibliotecas do grupo CTL e grupo HT, que produziu a lista de 94 DGE, sendo 14926 não significantes, 71 genes com aumento de expressão e 23 genes com diminuição de expressão.

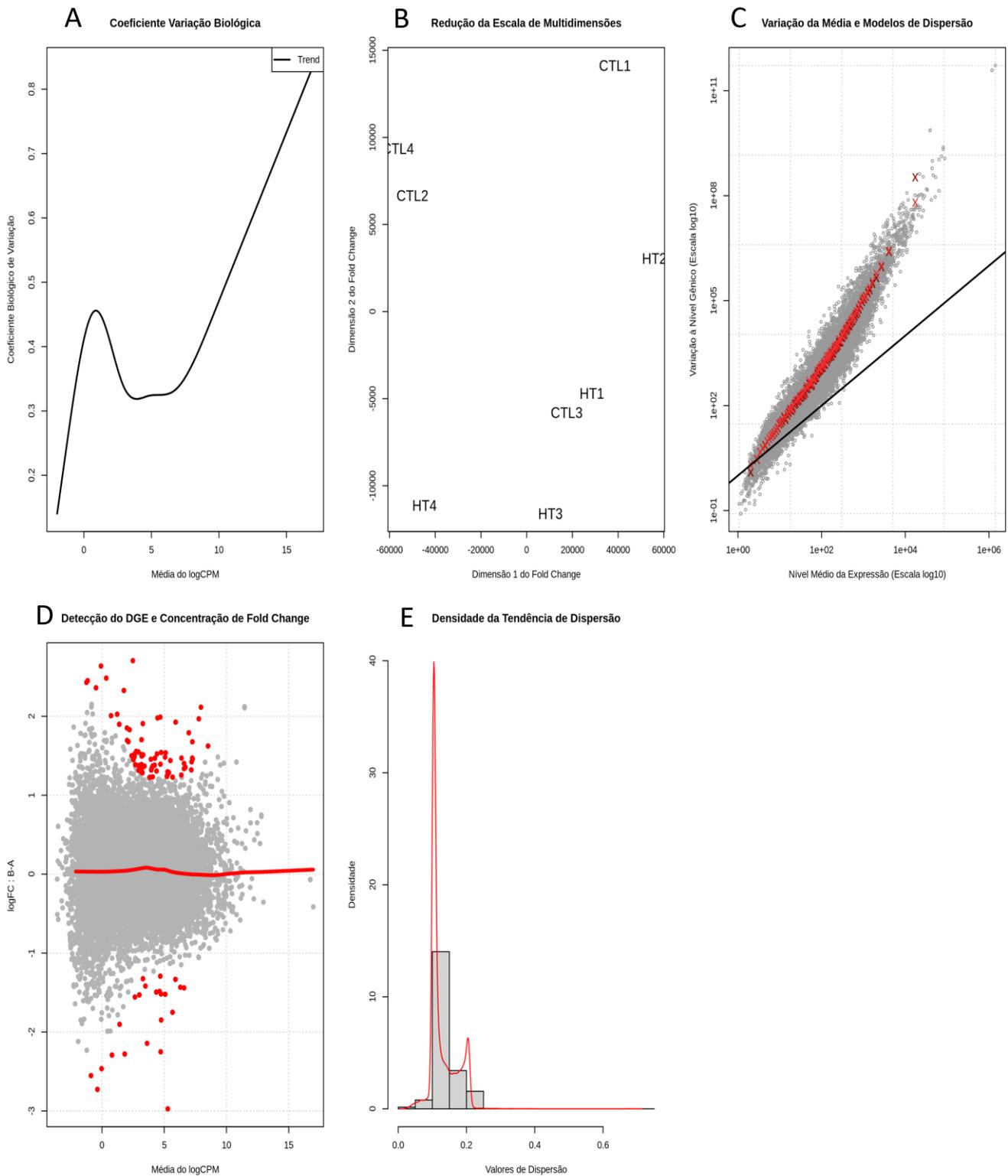


Figura 6. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão.

Com os 3 resultados realizamos a última estratégia que é dispersão de tendência e teste de razão por verossimilhança.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão de tendência e teste de razão de verossimilhança nas bibliotecas do grupo CTL e grupo HT, que produziu a lista de 98 DGE, sendo 14922 não significantes, 74 genes com aumento de expressão e 24 genes com diminuição de expressão.

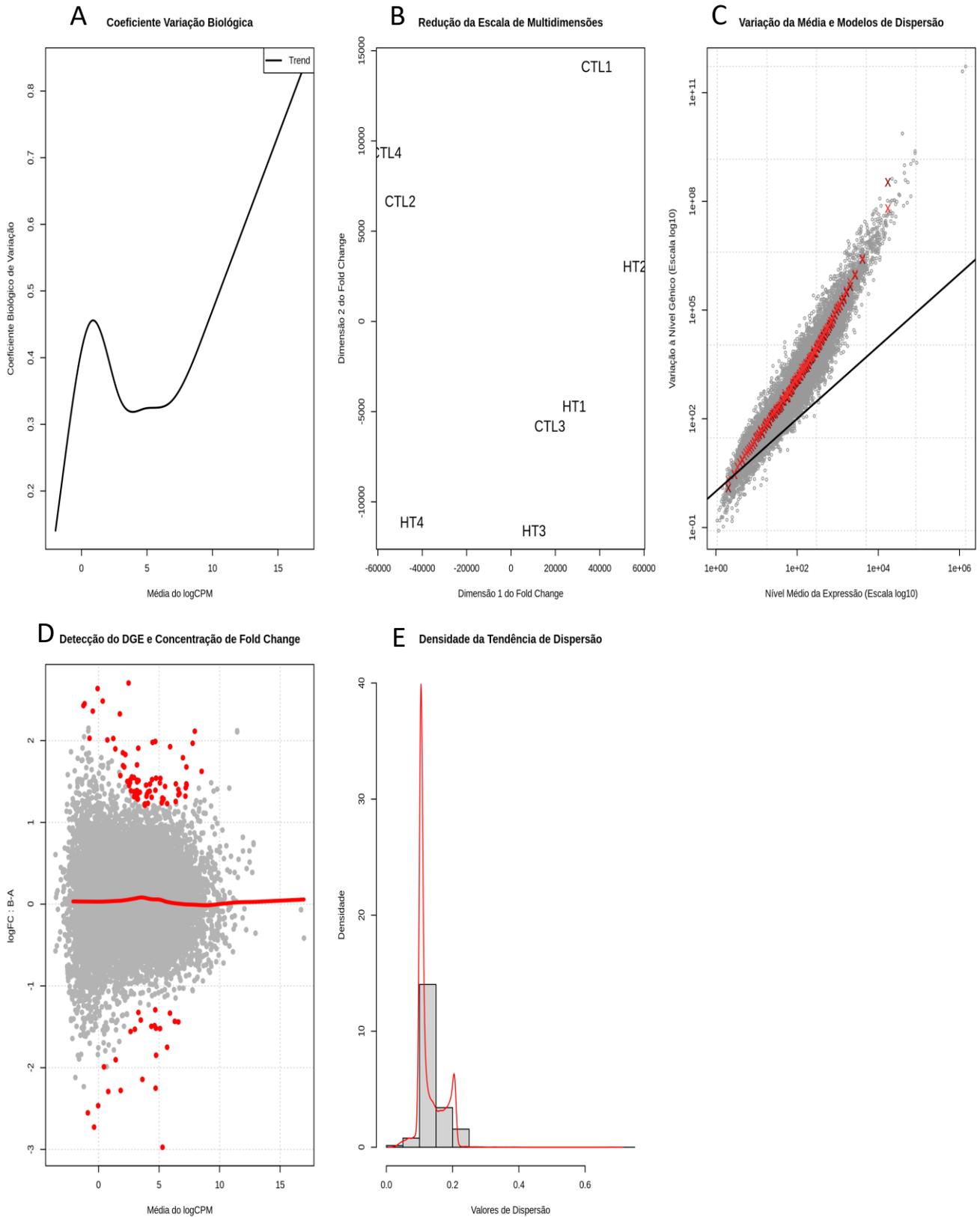


Figura 7. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão.

Para visualizar a capacidade e detecção de genes diferencialmente expressos produzimos um diagrama de *Venn* com as 4 estratégias.

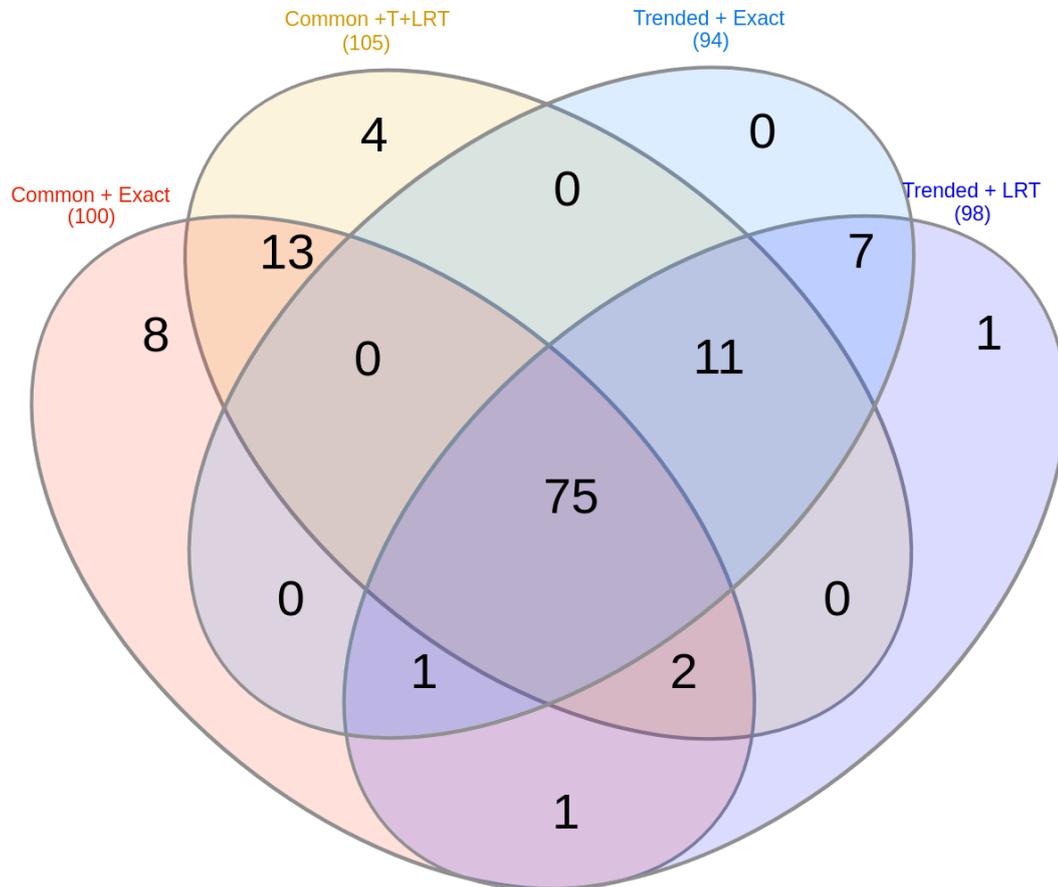


Figura 8. Capacidade de detecção de DGE em cada estratégia.

Um segundo cenário de análises, utilizamos a biblioteca do estudo GSE103305, iniciando pela estratégia dispersão comum e teste exato.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão comum e teste exato nas bibliotecas do estudo estado da tiroide, que produziu a lista de 621 DGE, sendo 14935 não significantes, 442 genes com aumento de expressão e 179 genes com diminuição de expressão.

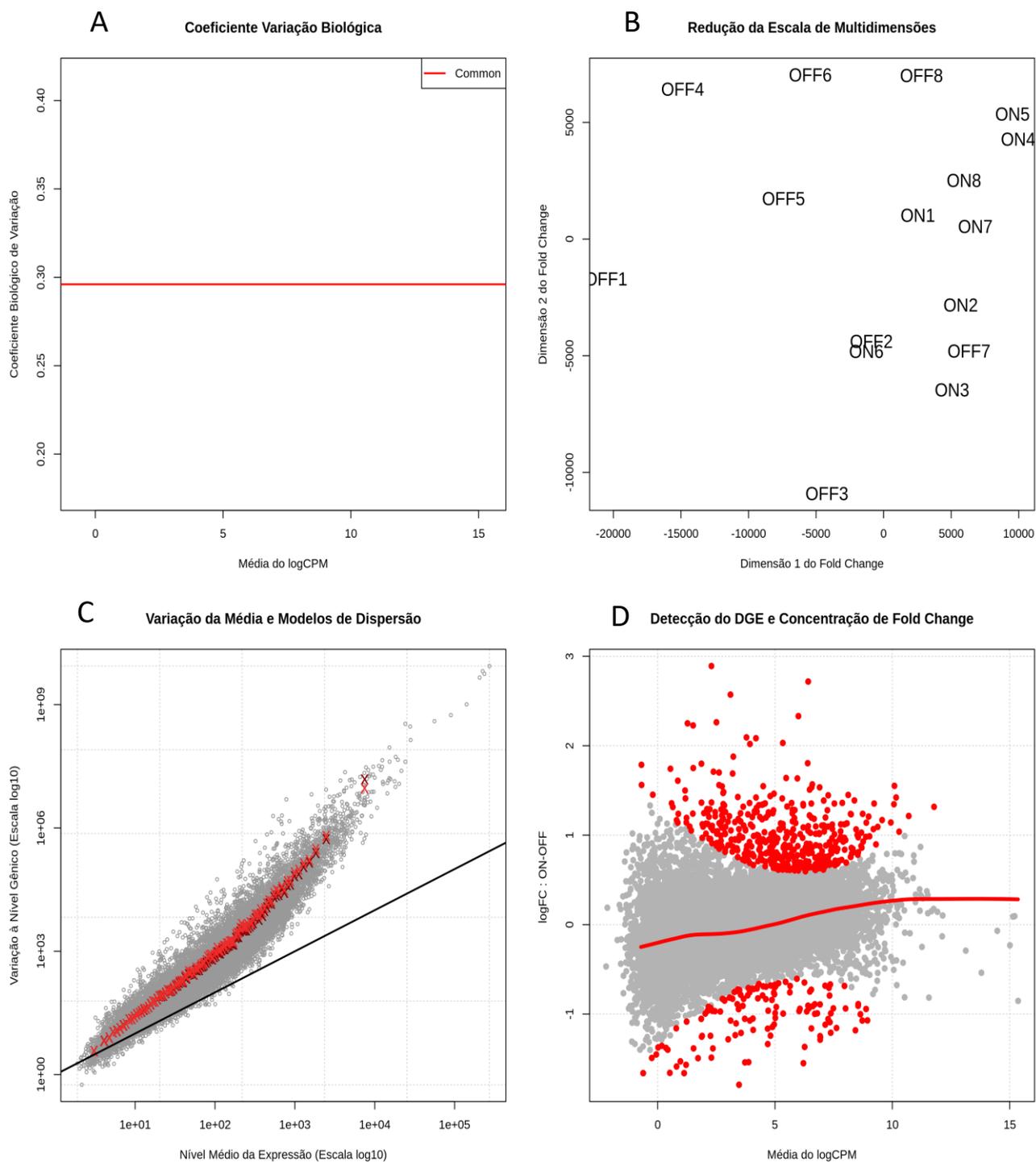


Figura 9. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste exato nas bibliotecas do estudo estado da tireoide.

Após os resultados da estratégia em dispersão comum e teste exato, iniciamos a segunda análise de dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo estado da tiroide.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão comum e teste de razão de verossimilhança nas bibliotecas do estudo estado da tiroide, que produziu a lista de 634 DGE, sendo 14922 não significantes, 450 genes com aumento de expressão e 184 genes com diminuição de expressão.

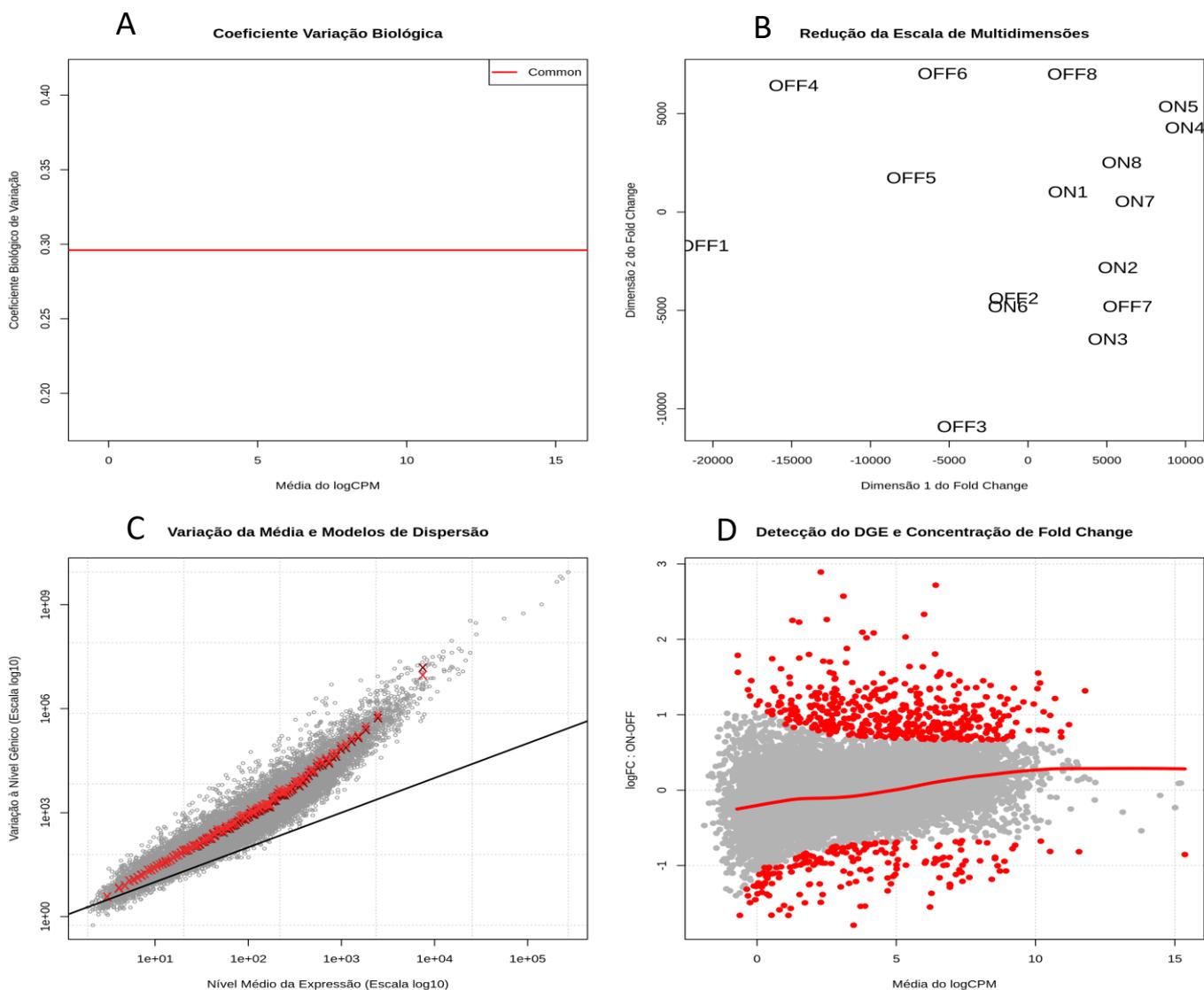


Figura 10. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo estado da tiroide.

Com os resultados anteriores iniciamos a próxima etapa de análise que é dispersão de tendência e teste exato das bibliotecas do estudo estado da tiroide.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão de tendência e teste exato nas bibliotecas do estudo estado da tiroide, que produziu a lista de 560 DGE, sendo 14996 não significantes, 430 genes com aumento de expressão e 130 genes com diminuição de expressão.

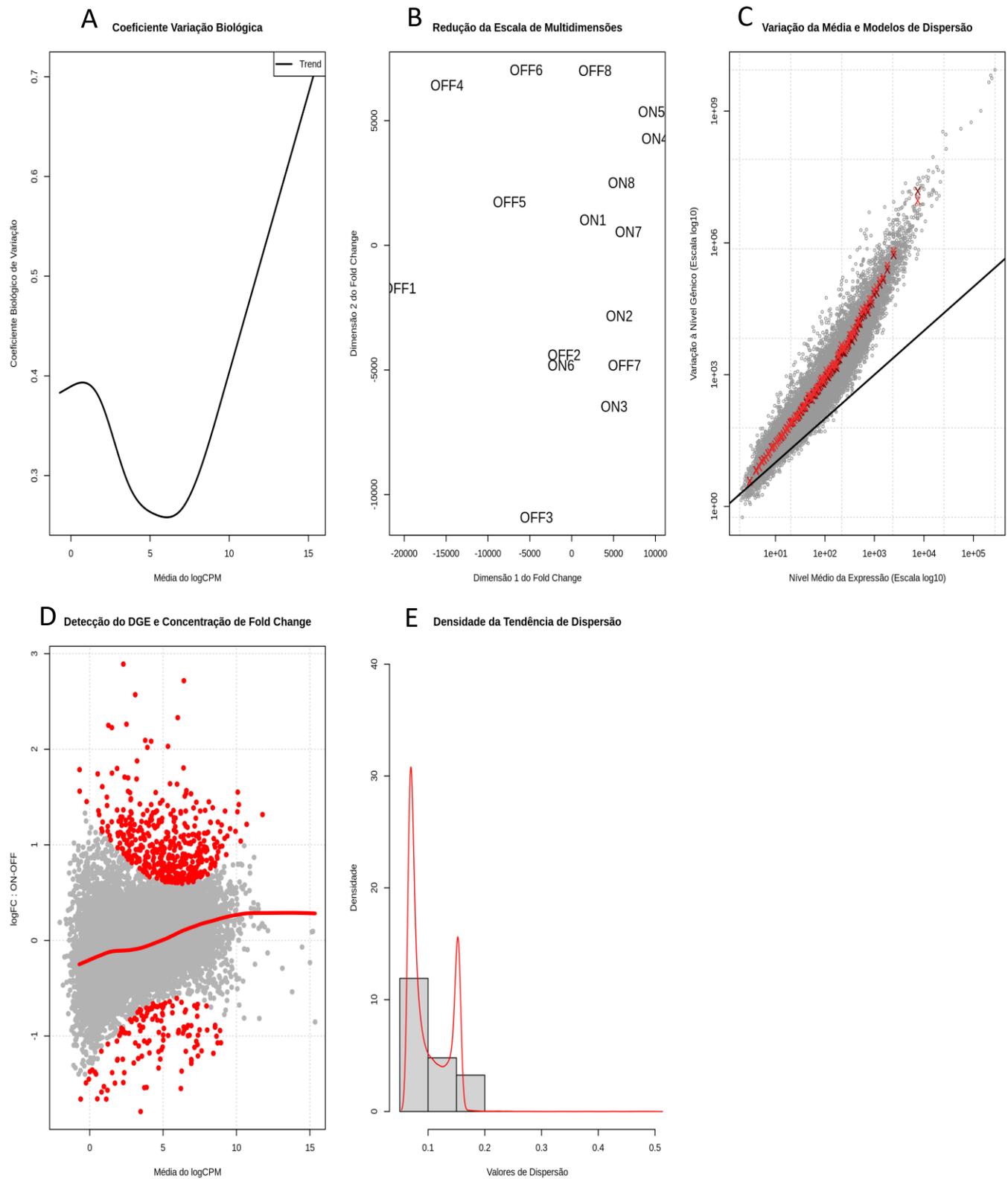


Figura 11. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste exato nas bibliotecas do estudo estado da tiroide.

Com os 3 resultados realizamos a última estratégia que é dispersão de tendência e teste de razão por verossimilhança das bibliotecas do estudo estado da tiroide.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão de tendência e teste de razão de verossimilhança nas bibliotecas do estudo estado da tiroide, que produziu a lista de 576 DGE, sendo 14980 não significantes, 441 genes com aumento de expressão e 135 genes com diminuição de expressão.

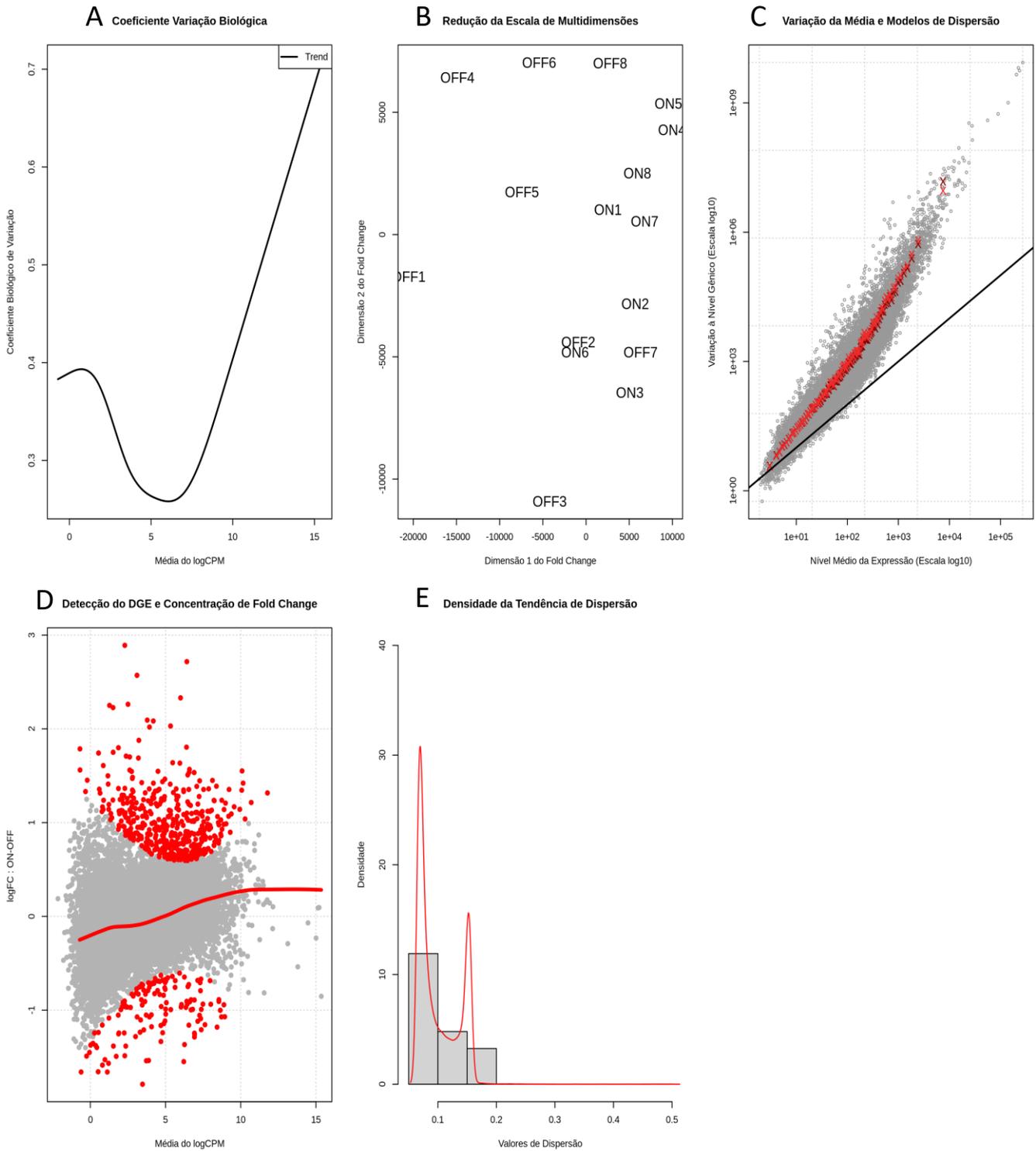


Figura 12. A: Coeficiente de variação biológica; B: Redução da escala de multidimensões; C: Variação média e modelos de dispersão; D: Detecção do DGE e concentração de *Fold Change*; E: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste de razão por verossimilhança nas bibliotecas do estudo estado da tiroide.

Para visualizar a capacidade e detecção de genes diferencialmente expressos produzimos um *Venn* com as 4 estratégias das bibliotecas do estudo estado da tiroide.

Na Figura 12 apresentamos o diagrama de *Venn* com as quatro estratégias edgeR para visualizar a capacidade de detecção de DGE em cada uma no estudo estado da tiroide.

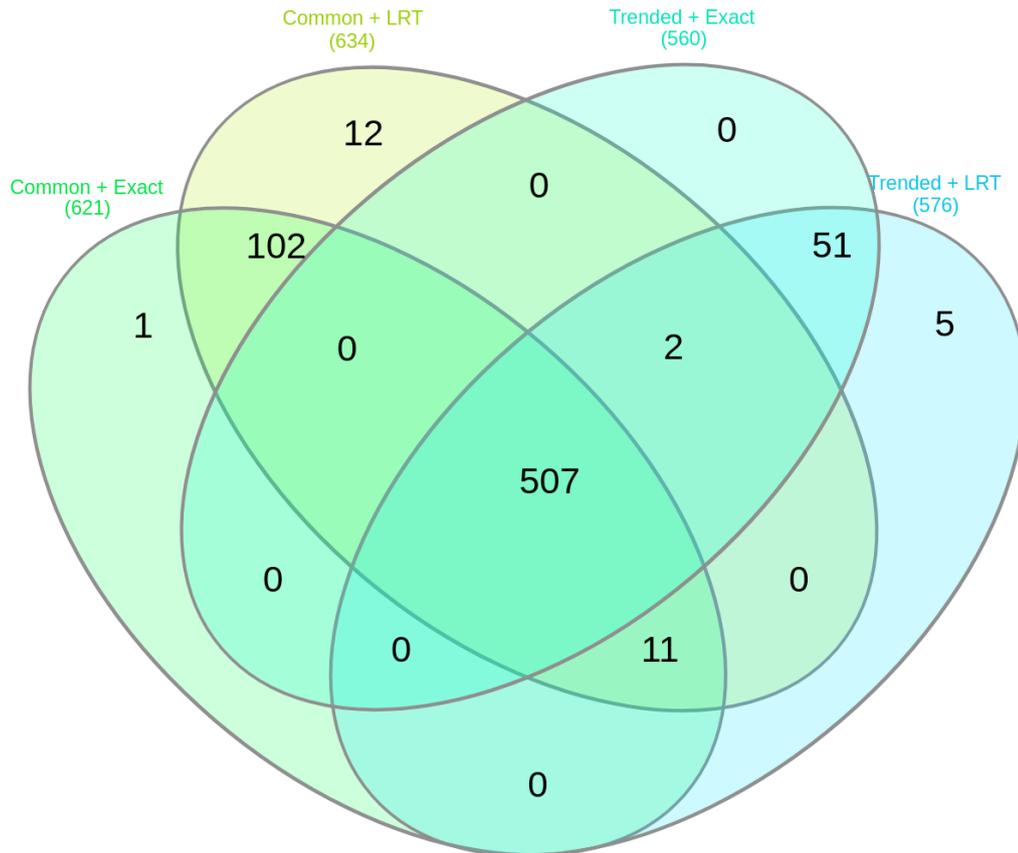


Figura 13. Capacidade de detecção de DGE em cada estratégia na biblioteca do estudo estado da tiroide.

No terceiro cenário de análises, utilizamos a biblioteca do estudo GSE124326 "*Whole blood transcriptome analysis in bipolar disorder reveals strong lithium effect*", iniciando pela estratégia dispersão comum e teste exato.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão comum e teste exato nas bibliotecas do estudo transtorno bipolar, que produziu a lista de 906 DGE, sendo 2625 não significantes, 571 genes com aumento de expressão e 335 genes com diminuição de expressão.

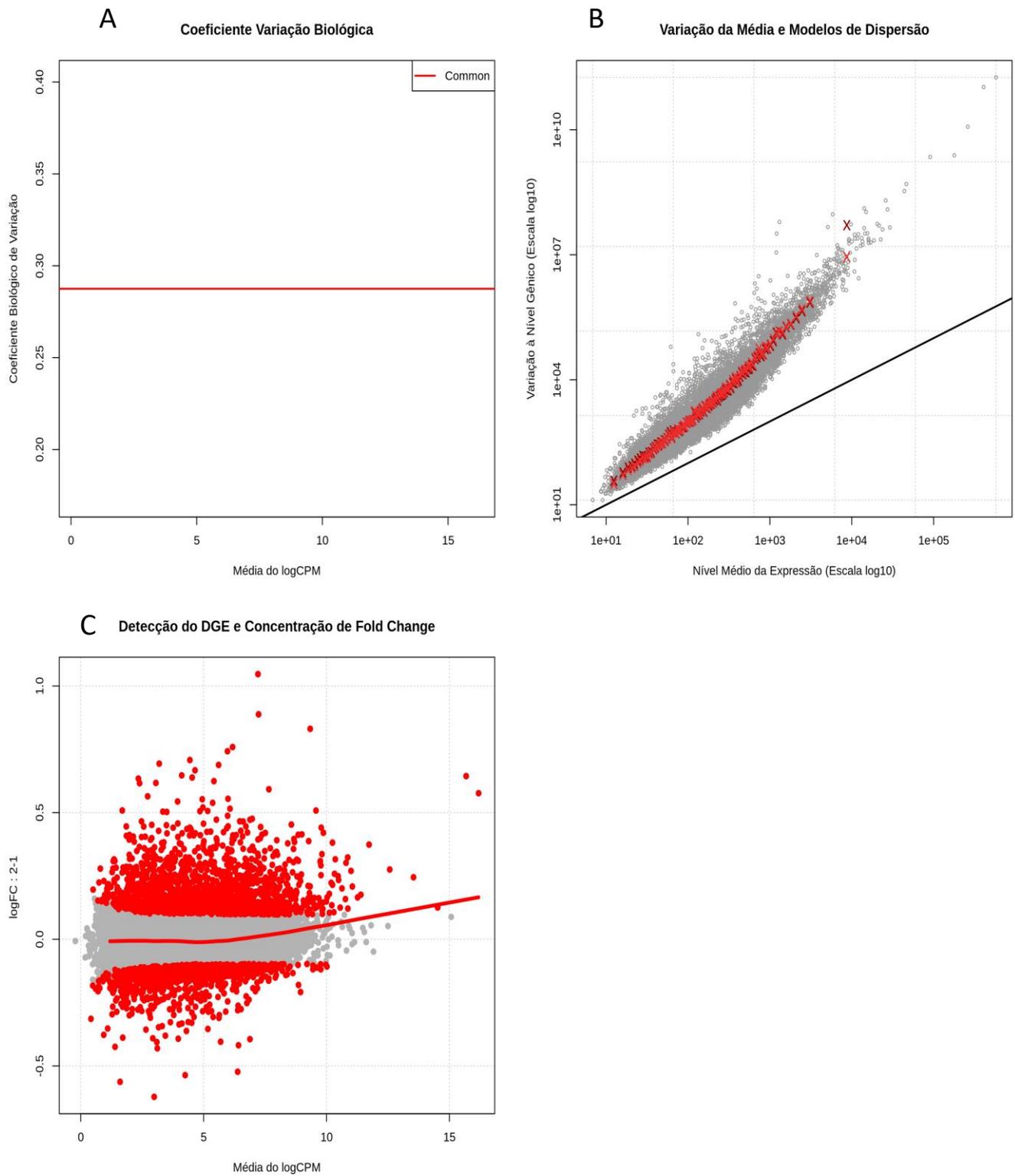


Figura 14. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste exato nas bibliotecas do estudo transtorno bipolar.

Após os resultados da estratégia em dispersão comum x teste exato, iniciamos a segunda análise de dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo transtorno bipolar.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão comum e teste de razão de verossimilhança nas bibliotecas do estudo transtorno bipolar, que produziu a lista de 893 DGE, sendo 2638 não significantes, 568 genes com aumento de expressão e 325 genes com diminuição de expressão.

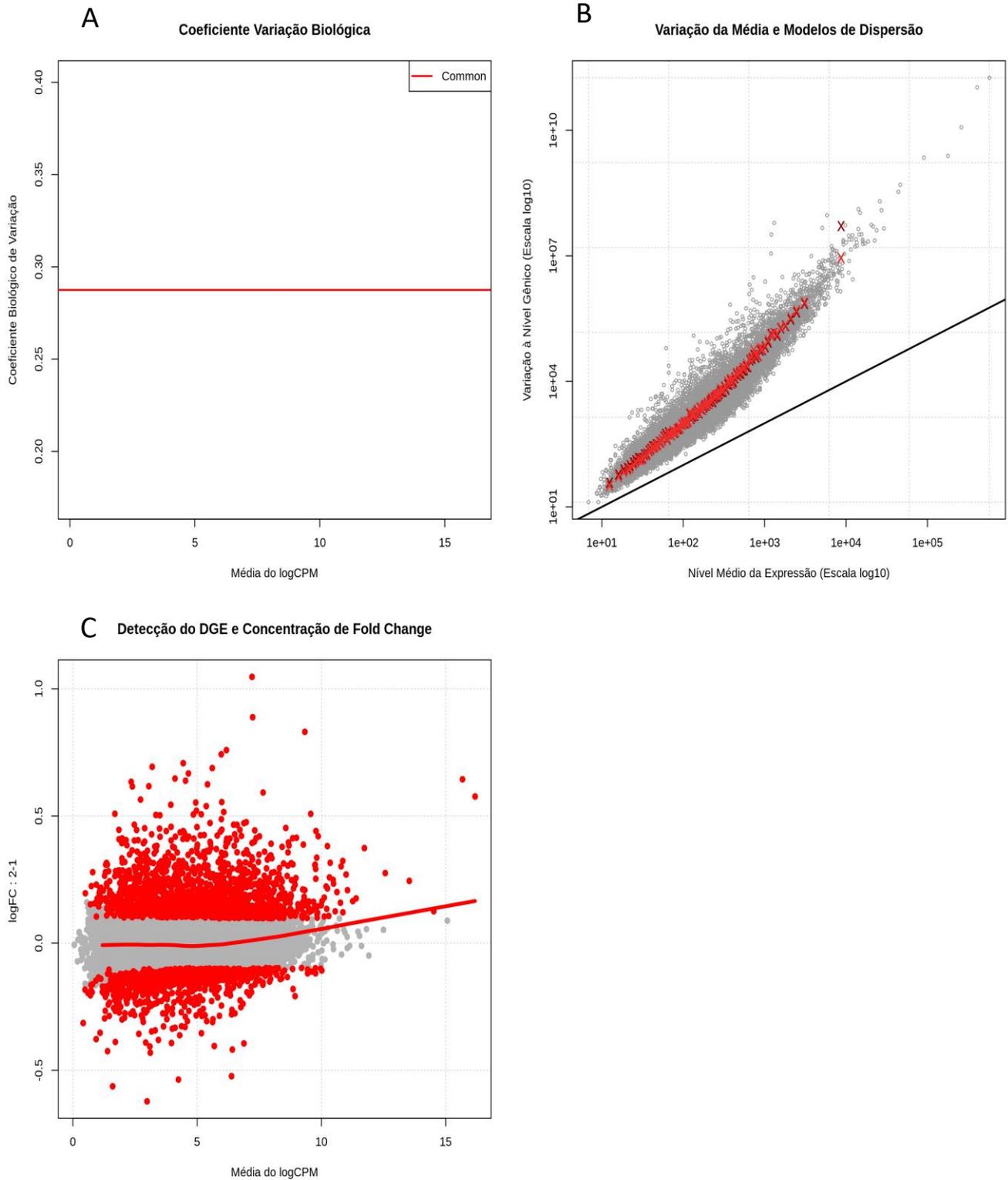


Figura 15. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*. Estratégia de análise edgeR em dispersão comum e teste de razão por verossimilhança nas bibliotecas do estudo transtorno bipolar.

Com os resultados anteriores iniciamos a próxima etapa de análise que é dispersão de tendência e teste exato das bibliotecas do estudo transtorno bipolar.

Apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão de tendência e teste exato nas bibliotecas do estudo transtorno bipolar, que produziu a lista de 911 DGE, sendo 2620 não significantes, 567 genes com aumento de expressão e 344 genes com diminuição de expressão.

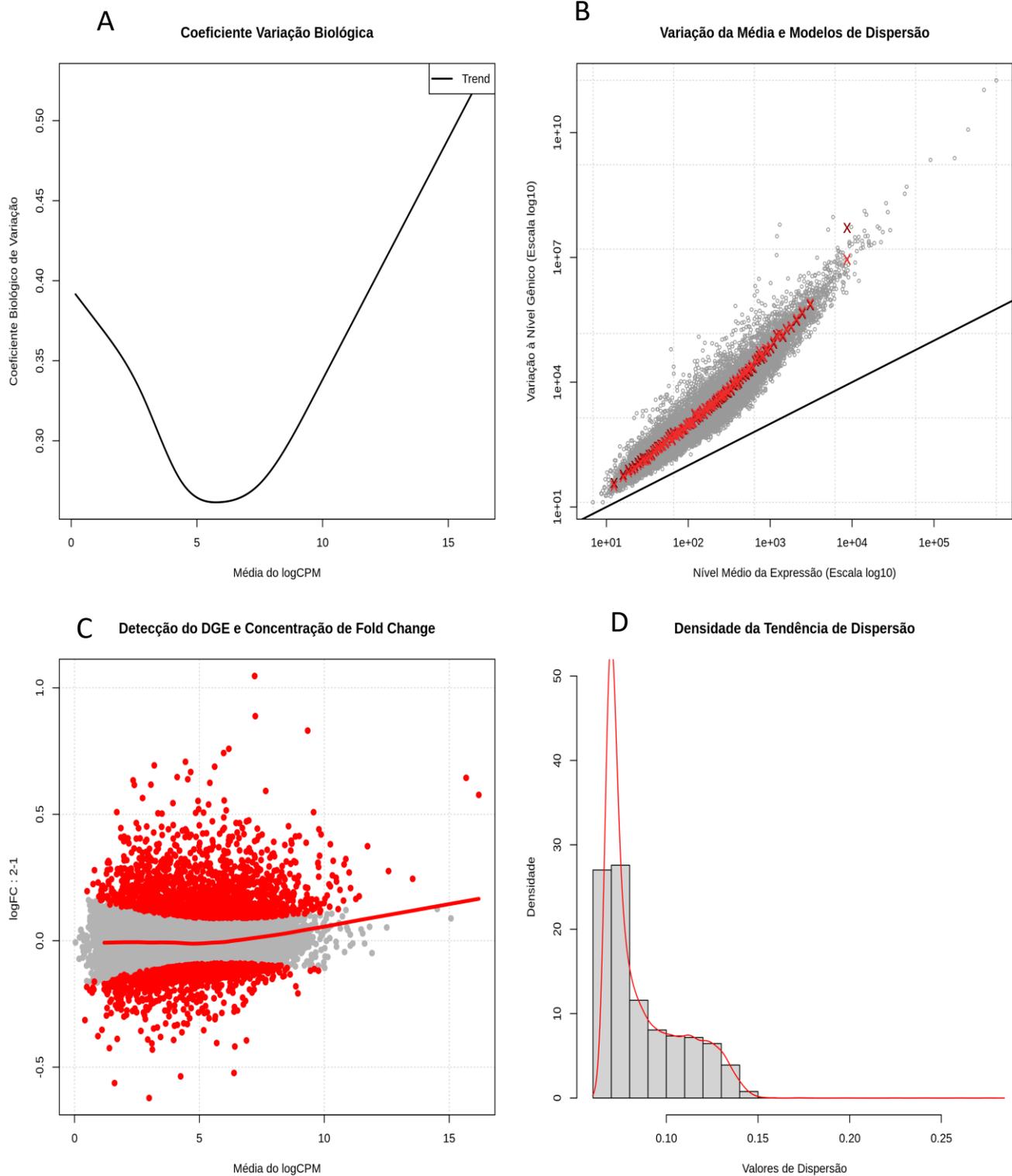


Figura 16. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*; D: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste exato nas bibliotecas do estudo transtorno bipolar.

Com os 3 resultados realizamos a última estratégia que é dispersão de tendência e teste de razão por verossimilhança das bibliotecas do estudo transtorno bipolar.

Na figura 16 apresentamos o quadro com os resultados das análises realizadas pelo pacote edgeR utilizando a estratégia dispersão de tendência e teste de razão de verossimilhança nas bibliotecas do estudo transtorno bipolar, que produziu a lista de 892 DGE, sendo 2639 não significantes, 563 genes com aumento de expressão e 329 genes com diminuição de expressão.

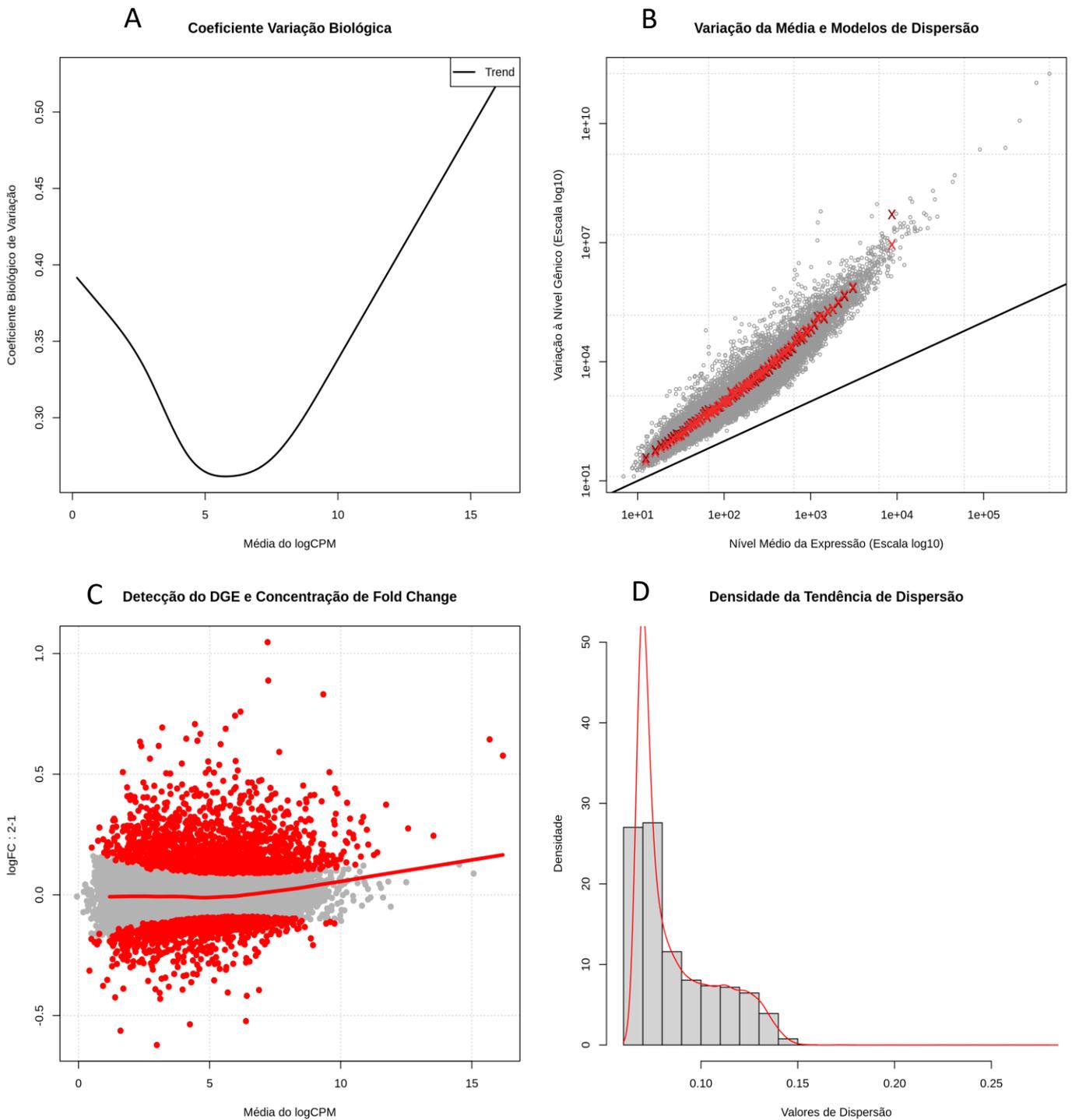


Figura 17. A: Coeficiente de variação biológica; B: Variação média e modelos de dispersão; C: Detecção do DGE e concentração de *Fold Change*; D: Densidade da tendência de dispersão. Estratégia de análise edgeR em dispersão de tendência e teste de razão por verossimilhança nas bibliotecas do estudo transtorno bipolar.

Para visualizar a capacidade e detecção de genes diferencialmente expressos produzimos um *Venn* com as 4 estratégias das bibliotecas do estudo transtorno bipolar.

Na Figura 17 apresentamos o diagrama de *Venn* com as quatro estratégias edgeR para visualizar a capacidade de detecção de DGE em cada uma no estudo transtorno bipolar.

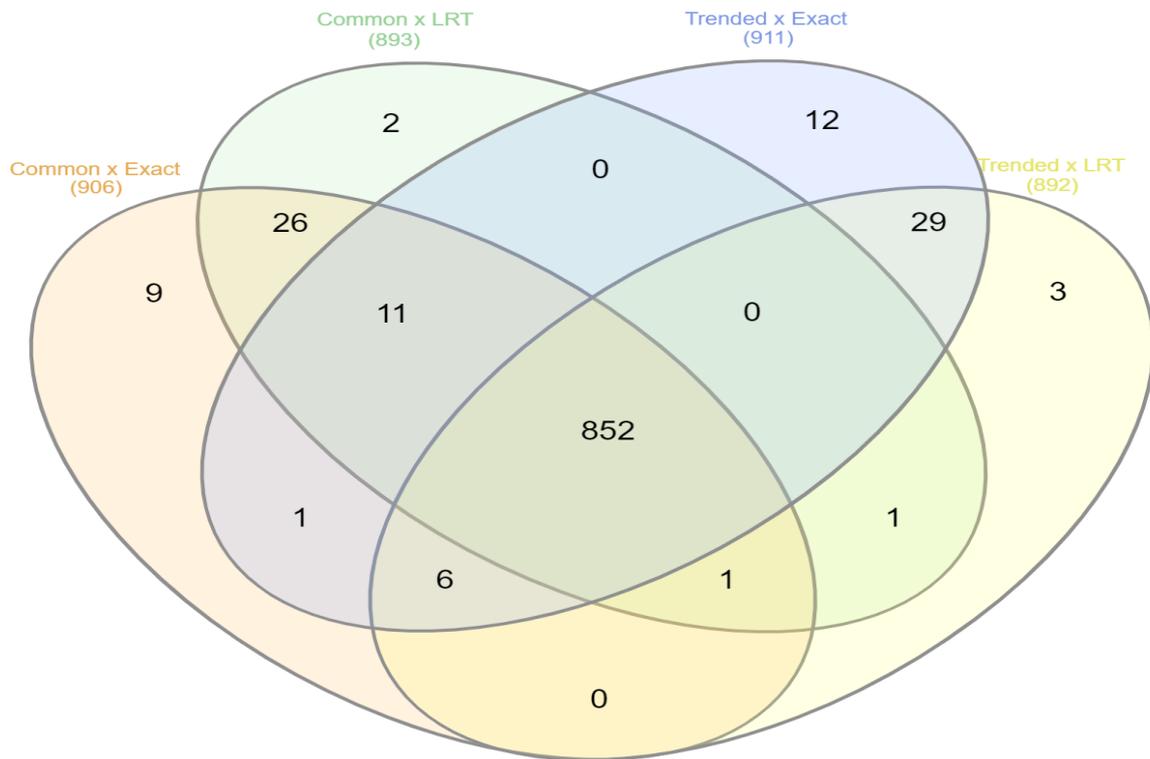


Figura 18. Capacidade de detecção de DGE em cada estratégia nas bibliotecas do estudo transtorno bipolar.

5.2 Identificação de genes não codificantes no hipotiroidismo

Com os dados produzidos no NGS iniciamos a análise para produção de uma lista DGE por dispersão comum e teste de razão por verossimilhança.

No anexo 5 apresentamos a tabela com os resultados das análises realizadas pelo pacote edgeR utilizando estratégia dispersão comum, dispersão de tendência e teste de razão de verossimilhança nas bibliotecas do grupo CTL e grupo HT, que produziu a lista de 105 DGE, sendo 14916 não significantes, 74 genes com aumento de expressão e 31 genes com diminuição de expressão. Após a produção da lista total de DGE, identificamos somente os genes não codificantes.

Está representado na Tabela 4 a lista DGE filtrada para transcritos não codificantes, por ordem crescente da expressão do logFC, onde temos 18 DGE não codificantes sendo 14 com aumento de expressão e 4 com diminuição de expressão.

Tabela 4. Lista DGE filtrada para transcritos não codificantes, por ordem crescente da expressão do logFC e identificação Refseq e símbolo do gene, valores de logCPM, LR e FDR.

Gene	logFC	logCPM	LR	FDR	RefSeq
<i>CASP17P</i>	2,68	-0,411	19,943	0,004	NR_034078
<i>SCARNA7</i>	2,116	8,31	32,201	0	NR_003001
<i>OR2L1P</i>	2,006	1,627	22,631	0,001	NR_002145
<i>MT1L</i>	1,969	3,574	25,688	0	NR_001447
<i>SNORA12</i>	1,679	7,491	20,825	0,003	NR_002954
<i>SNORA53</i>	1,541	5,309	17,371	0,012	NR_003015
<i>SNORA70F</i>	1,531	3,148	16,072	0,016	NR_033309
<i>SCARNA5</i>	1,473	7,433	16,207	0,016	NR_003008
<i>USP17L6P</i>	1,465	3,391	14,97	0,022	NR_027279
<i>SCARNA10</i>	1,462	7,458	15,977	0,017	NR_004387
<i>LOC145474</i>	1,442	4,119	14,963	0,022	NR_027046
<i>SNORD15B</i>	1,426	7,37	15,23	0,02	NR_000025
<i>SNORA63</i>	1,407	9,963	14,901	0,023	NR_002586
<i>SNORA7B</i>	1,338	6,745	13,433	0,039	NR_002992
<i>SERPINB9P1</i>	-1,723	0,966	14,609	0,025	NR_033851
<i>RNF5P1</i>	-1,914	1,079	17,853	0,01	NR_003129
<i>HOTAIRM1</i>	-1,984	0,431	15,555	0,019	NR_038366
<i>ZNF890P</i>	-2,36	-0,154	16,561	0,016	NR_034163

Com a lista DGE de genes não codificantes, iniciamos a busca de processos biológicos e metabólicos de todos, e nas bases de dados online e apenas 1 gene foi sinalizado nessas vias, obviamente por muitos ainda não serem validados e caracterizados em pesquisas e literatura.

Apresentamos o *Network* (70) que identificou um único processo biológico envolvendo a lista de transcritos não codificantes para apenas um gene *MT1L*, via número GO: 0010273, que se caracteriza por qualquer processo que reduza ou remova a toxicidade do íon cobre. Isso inclui o transporte de cobre para longe de áreas sensíveis e para compartimentos ou complexos cuja finalidade é o sequestro de íons de cobre (71).

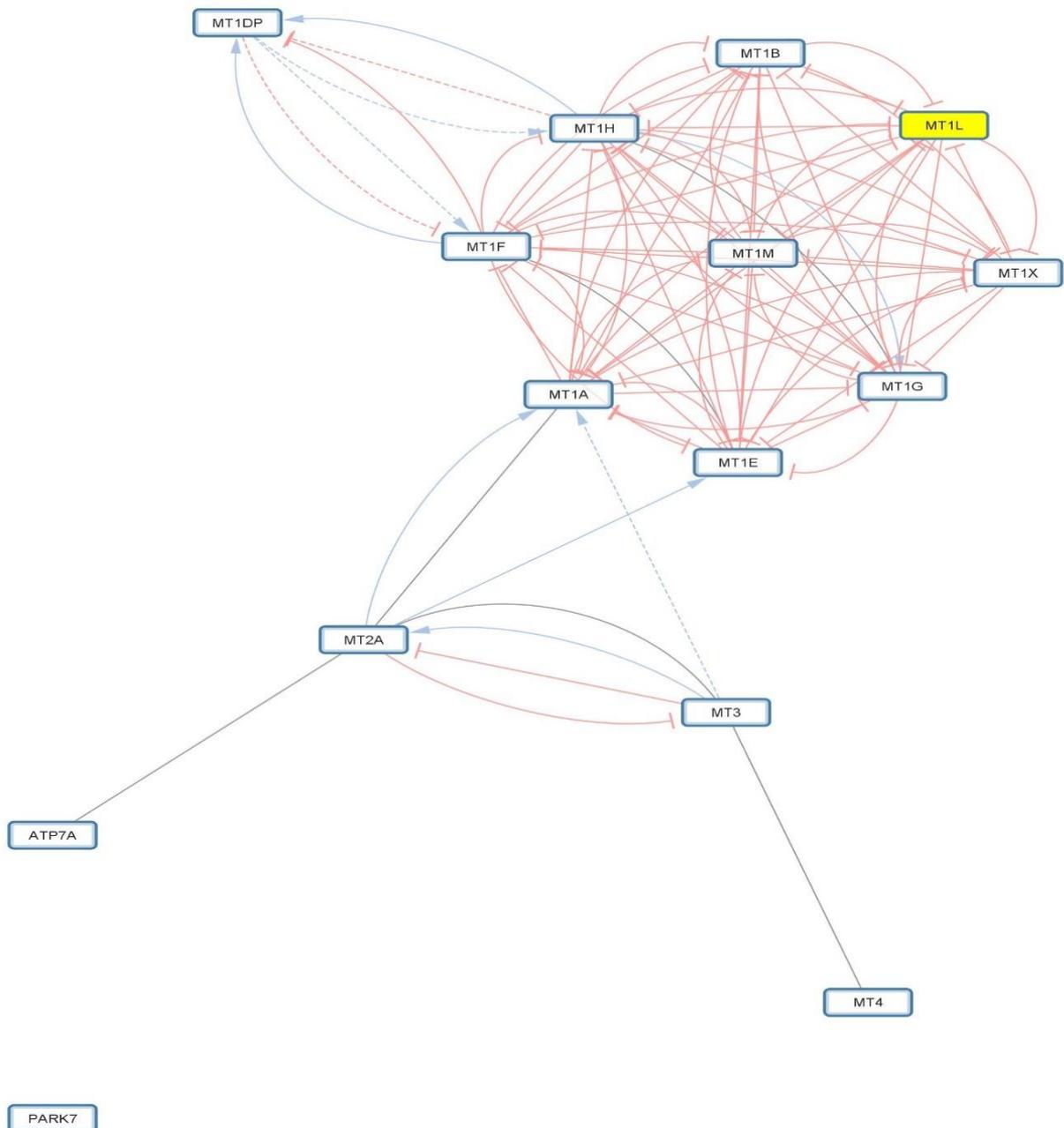


Figura 19. *Network* baseada na lista de transcritos não codificantes.

Realizamos a identificação de genes não codificantes comuns entre nossa biblioteca CTL e HT e a biblioteca do estudo estado da tiroide.

Na figura 19, apresentamos o diagrama de *Venn* que identificou a presença de 7 genes não codificantes entre a lista DGE da nossa biblioteca (HT comparada ao CTL) com a biblioteca do estudo estado da tiroide. São os 7 genes: *SCARNA7*, *SCARNA5*, *SNORD15B*, *SNORA7B*, *SCARNA10*, *SNORA53* e *SNORA12*.

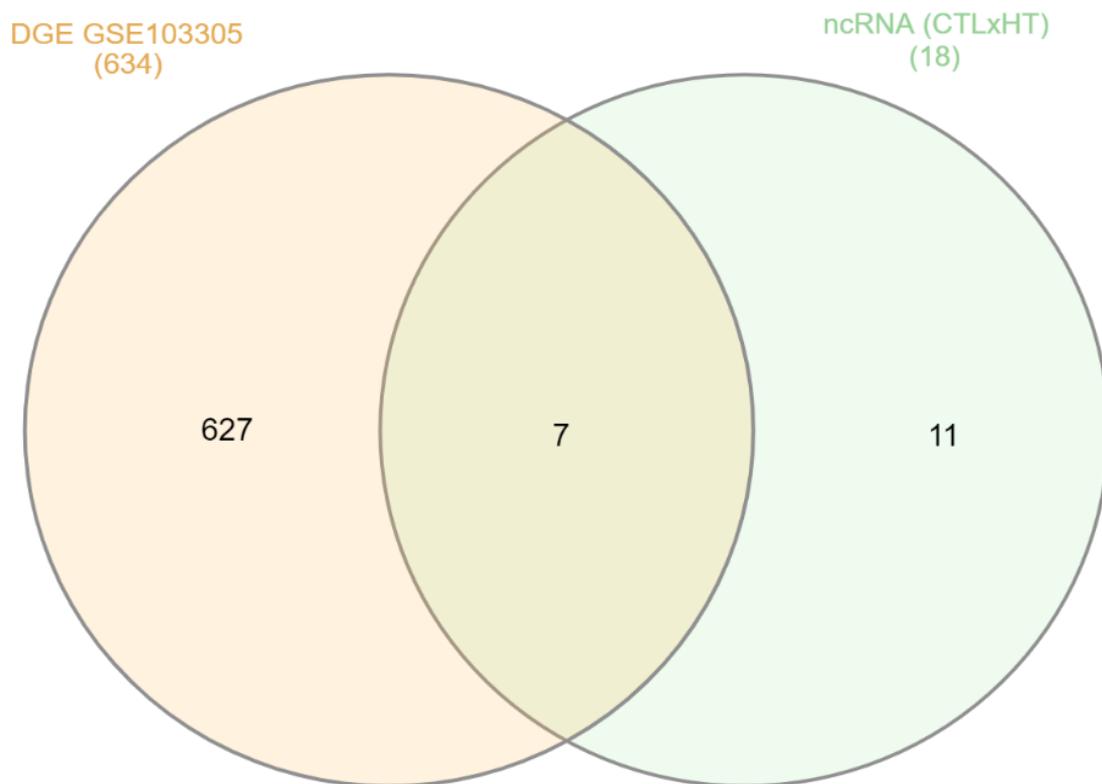


Figura 20. Genes não codificantes comuns entre HT comparado ao CTL e a biblioteca do estudo estado da tiroide.

O próximo passo foi realizar a anotação dos RNAs e suas descrições básicas. Na tabela 5 apresento a lista dos genes não codificantes, suas descrições e biotipos.

Tabela 5. Genes não codificantes e suas descrições e biotipos.

Símbolo do Gene	Descrição e Biotipo
<i>SCARNA10</i>	small Cajal body-specific RNA 10
<i>SCARNA5</i>	small Cajal body-specific RNA 5
<i>SCARNA7</i>	small Cajal body-specific RNA 7
<i>SNORA12</i>	small nucleolar RNA, H/ACA box 12
<i>SNORA53</i>	small nucleolar RNA, H/ACA box 53
<i>SNORA7B</i>	small nucleolar RNA, H/ACA box 7B
<i>SNORD15B</i>	small nucleolar RNA, C/D box 15B

5.3 Caracterização de genes mitocondriais diferencialmente expressos no hipotireoidismo e relação com o quadro clínico e laboratorial

Os resultados a seguir refere-se à caracterização dos genes mitocondriais (mitGenes) diferencialmente expressos nas bibliotecas produzidas em nosso estudo HT em comparação ao CTL e sua relação com o quadro laboratorial.

Produzimos conforme figura 20 cruzando a lista de genes mitocondriais com a nossa lista DGE (HT comparado ao CTL) para identificá-los.

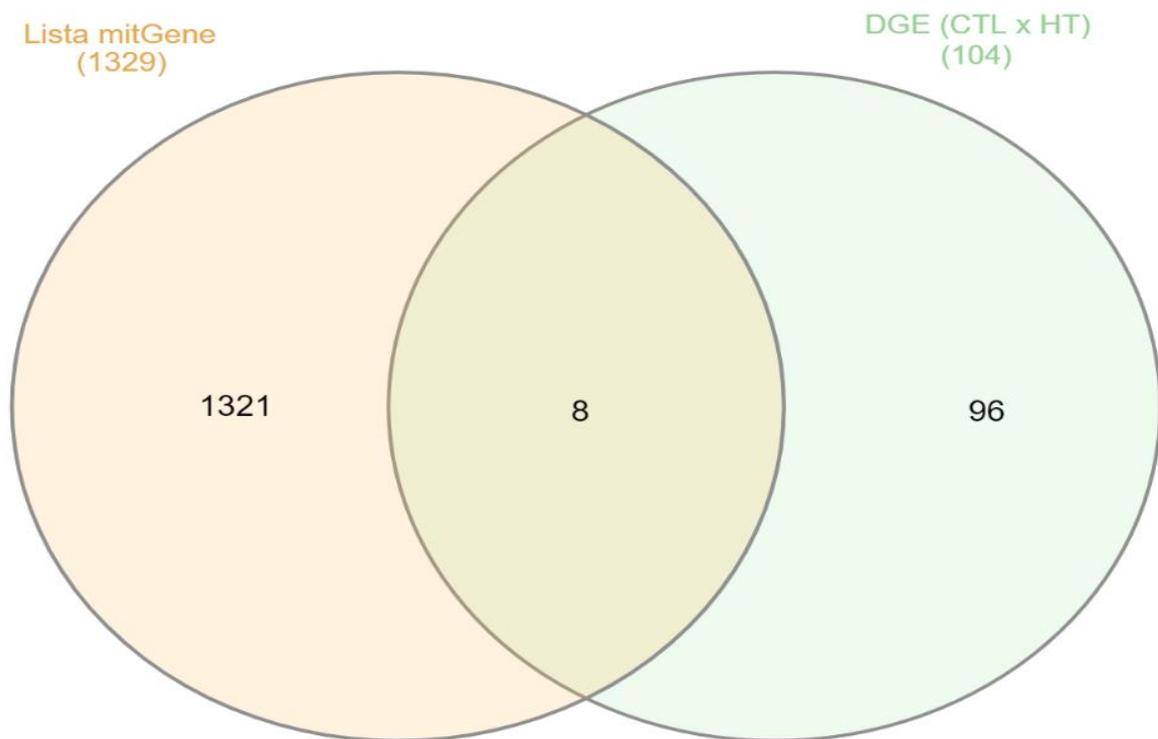


Figura 21. Identificação de genes mitocondriais entre base *MitoMiner*, *MITOCARTA* e *Mitochondrial Biology Unit*, com a lista DGE (HT comparado ao CTL).

Realizamos a filtragem das expressões obtidas na análise de DGE total em nossa biblioteca.

Apresentamos a lista com os 8 genes mitocondriais conforme os resultados de logFC, logCPM, LR, FDR, Refseq e símbolo do gene da lista DGE (HT comparado ao CTL).

Tabela 6. Lista com os 8 genes mitocondriais conforme os resultados de logFC, logCPM, LR, FDR, Refseq e símbolo do gene da lista DGE (HT comparado ao CTL).

Gene	logFC	logCPM	LR	FDR	RefSeq
<i>CMC1</i>	1,811	2,536	21,095	0,003	NM_182523
<i>SLIRP</i>	1,734	3,438	20,46	0,003	NM_031210
<i>DBI</i>	1,55	4,958	17,447	0,011	NM_001178017
<i>TMEM126B</i>	1,506	2,864	15,354	0,02	NM_018480
<i>RPL11</i>	1,42	10,969	15,179	0,021	NM_000975
<i>COX7B</i>	1,399	4,399	14,146	0,03	NM_001866
<i>CMPK2</i>	-1,492	4,563	16,088	0,016	NM_207315
<i>IFI27</i>	-1,87	1,797	19,954	0,004	NM_005532

Produzimos uma lista com descrições básicas desses mitGenes.

Tabela 7. Descrição dos genes mitocondriais da lista DGE (HT comparado ao CTL).

Gene	Descrição	Tecidos
<i>IFI27</i>	proteína induzível por interferon alfa 27	Estômago; Baço; Pulmões
<i>COX7B</i>	subunidade 7B da citocromo c oxidase	Coração ; Rins; Cólon
<i>TMEM126B</i>	proteína transmembrana 126B	Rins; Tecido Adiposo
<i>SLIRP</i>	proteína de ligação de RNA de interação haste-alça SRA	Cólon; Rins; Tecido Adiposo
<i>CMPK2</i>	citidina / uridina monofosfato quinase 2	Medula Óssea; Baço
<i>CMC1</i>	Motivo C-X9-C contendo 1	Tecido Adiposo; Rins; Cólon
<i>RPL11</i>	proteína ribossomal L11	Ovário; Medula Óssea
<i>DBI</i>	inibidor de ligação de diazepam, proteína de ligação de acil-CoA	Próstata; Fígado; Tecido Adiposo

A próxima etapa foi o enriquecimento das análises.

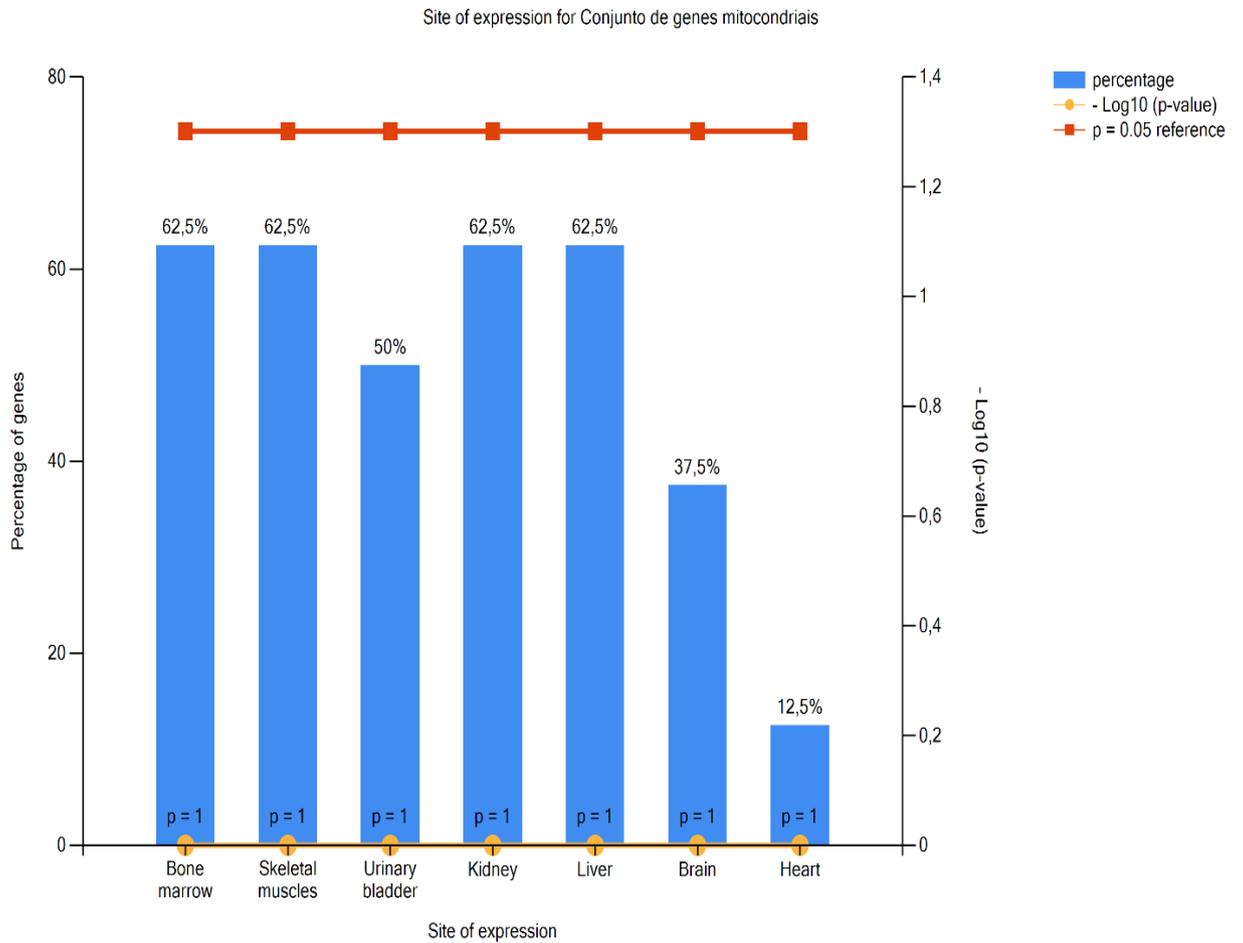


Gráfico 1. Conjunto dos 8 genes mitocondriais e a expressão em tecidos alvos do hormônio tiroídiano, apresentados em porcentagem e $-\log_{10}(p\text{-value})$.

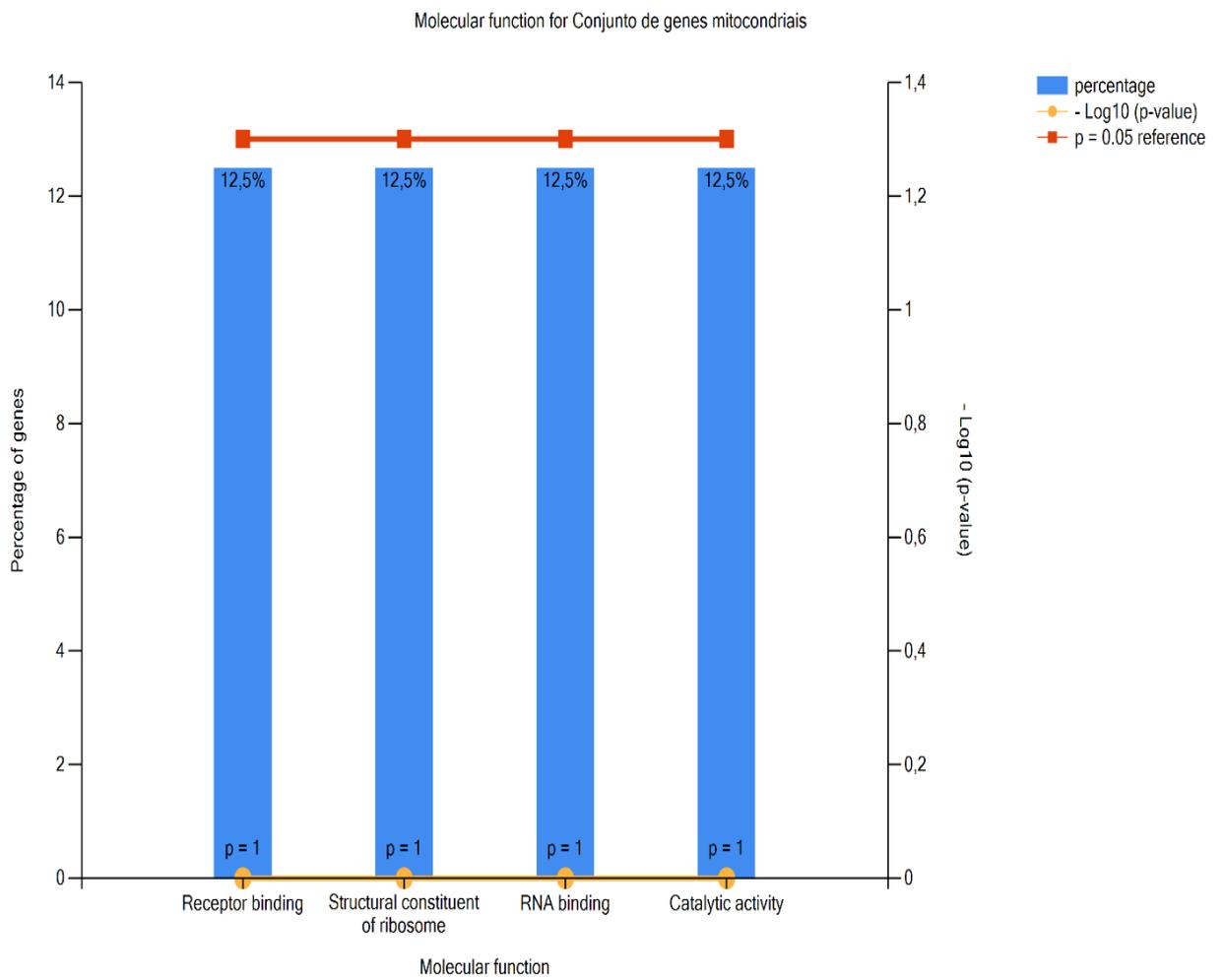


Gráfico 2. Conjunto dos 8 genes mitocondriais e a relação com funções moleculares, apresentados em porcentagem e $-\log_{10}(p\text{-value})$.

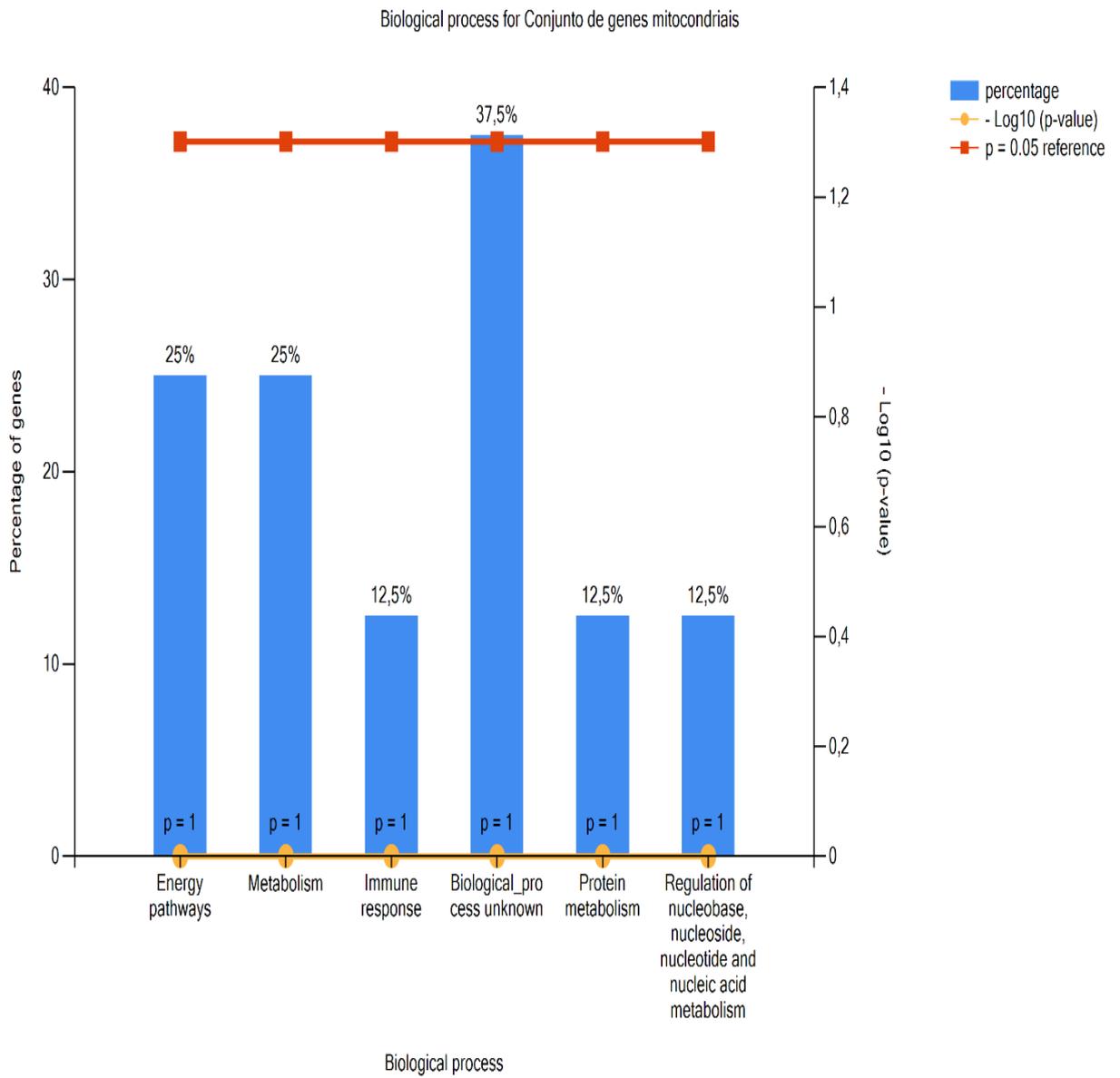


Gráfico 3. Conjunto dos 8 genes mitocondriais e a relação com processos biológicos, apresentados em porcentagem e $-\log_{10}(p\text{-value})$.

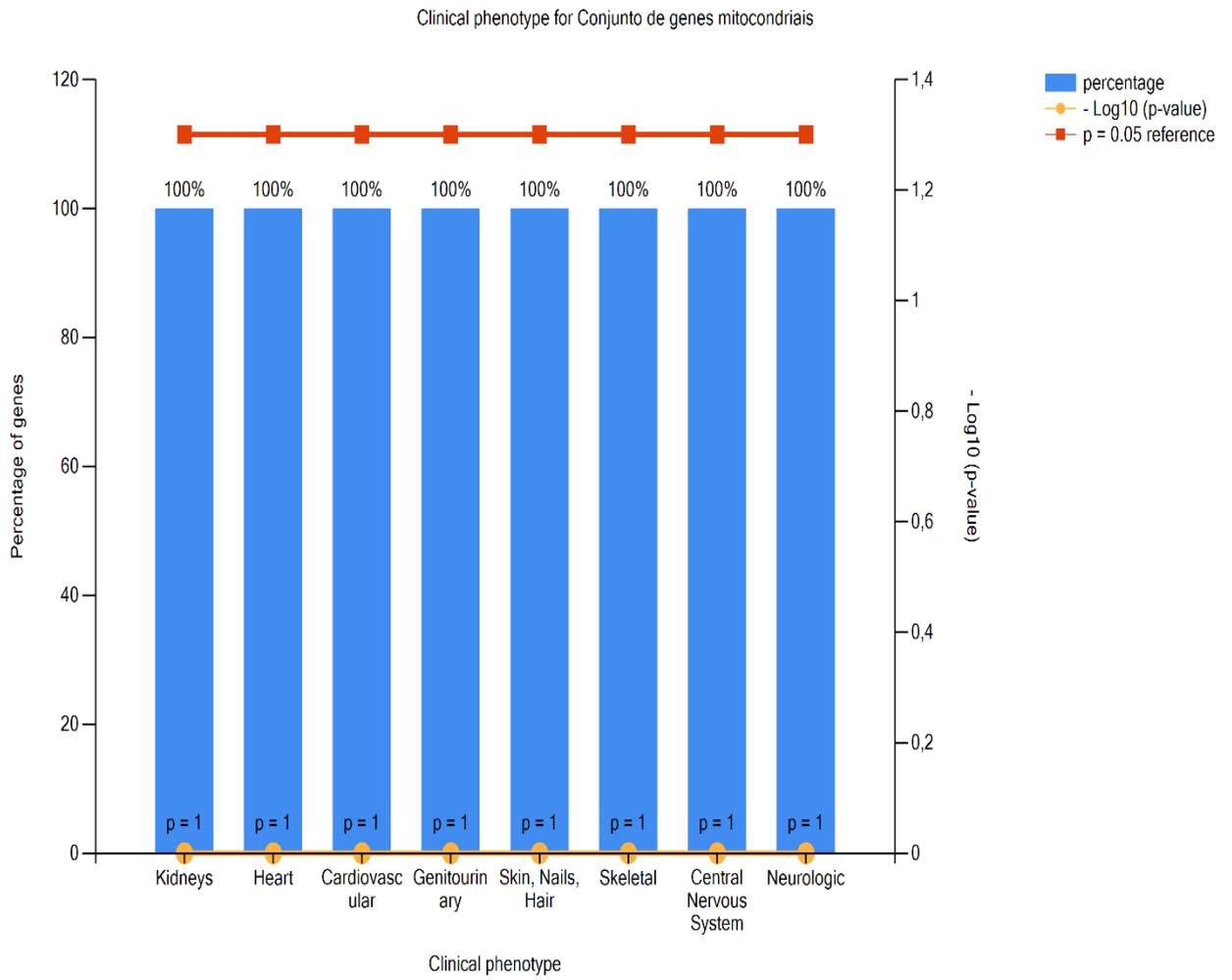


Gráfico 4. Conjunto dos 8 mitGenes e a relação com os fenótipos clínicos catalogados na literatura que são classicamente detectáveis no hipotireoidismo.

Network entre snRNA e miRNA

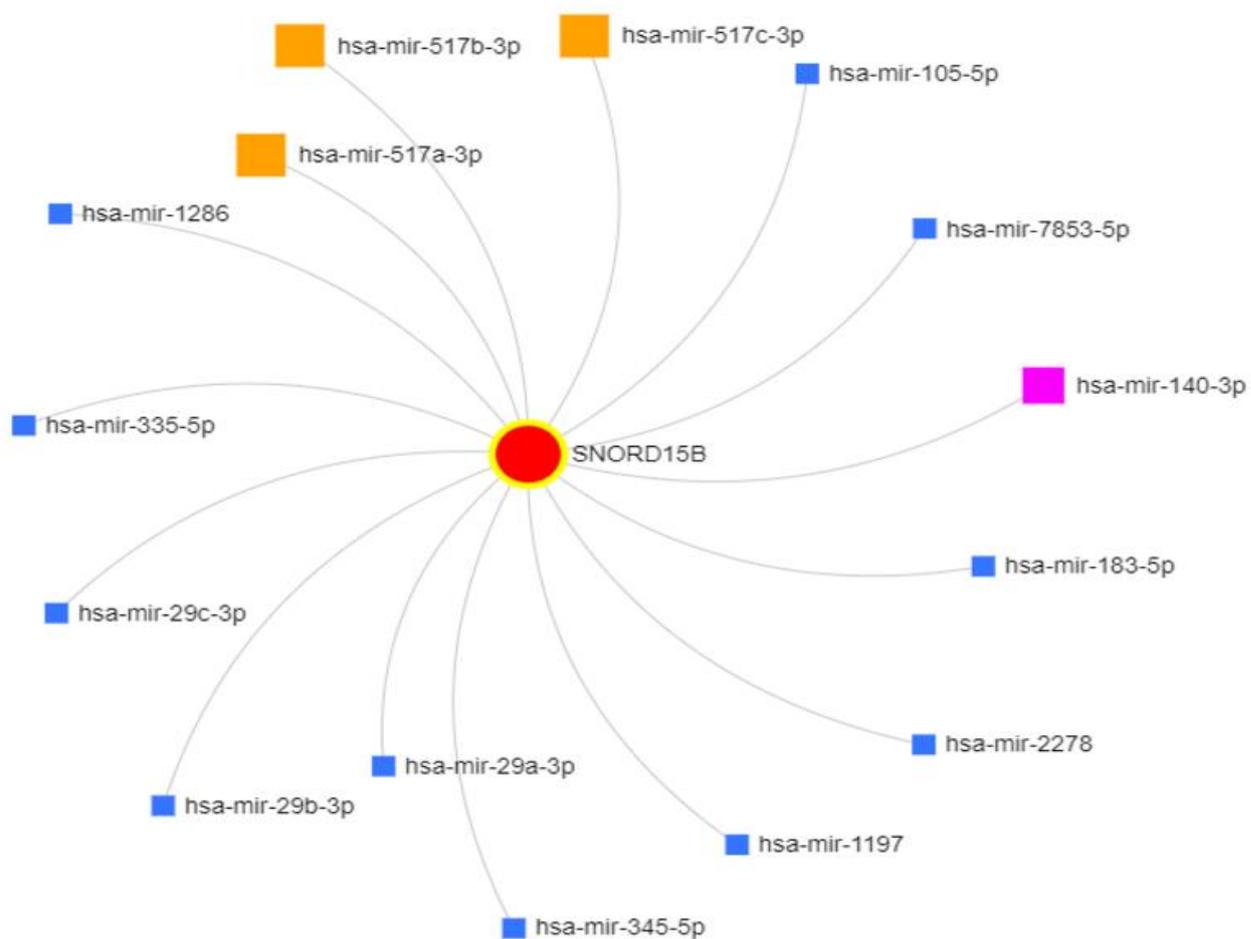


Gráfico 5: Network entre RNA não codificante (snoRNA) e miRNAs;



SNORD15B eventos associados ao hipotireoidismo



MIR140 é comumente visto quando há a manifestação de doença coronariana cardíaca



MIR517 está relacionada à pré-eclâmpsia



miRNAs

Regulação de RNAs mitocondriais por miRNAs

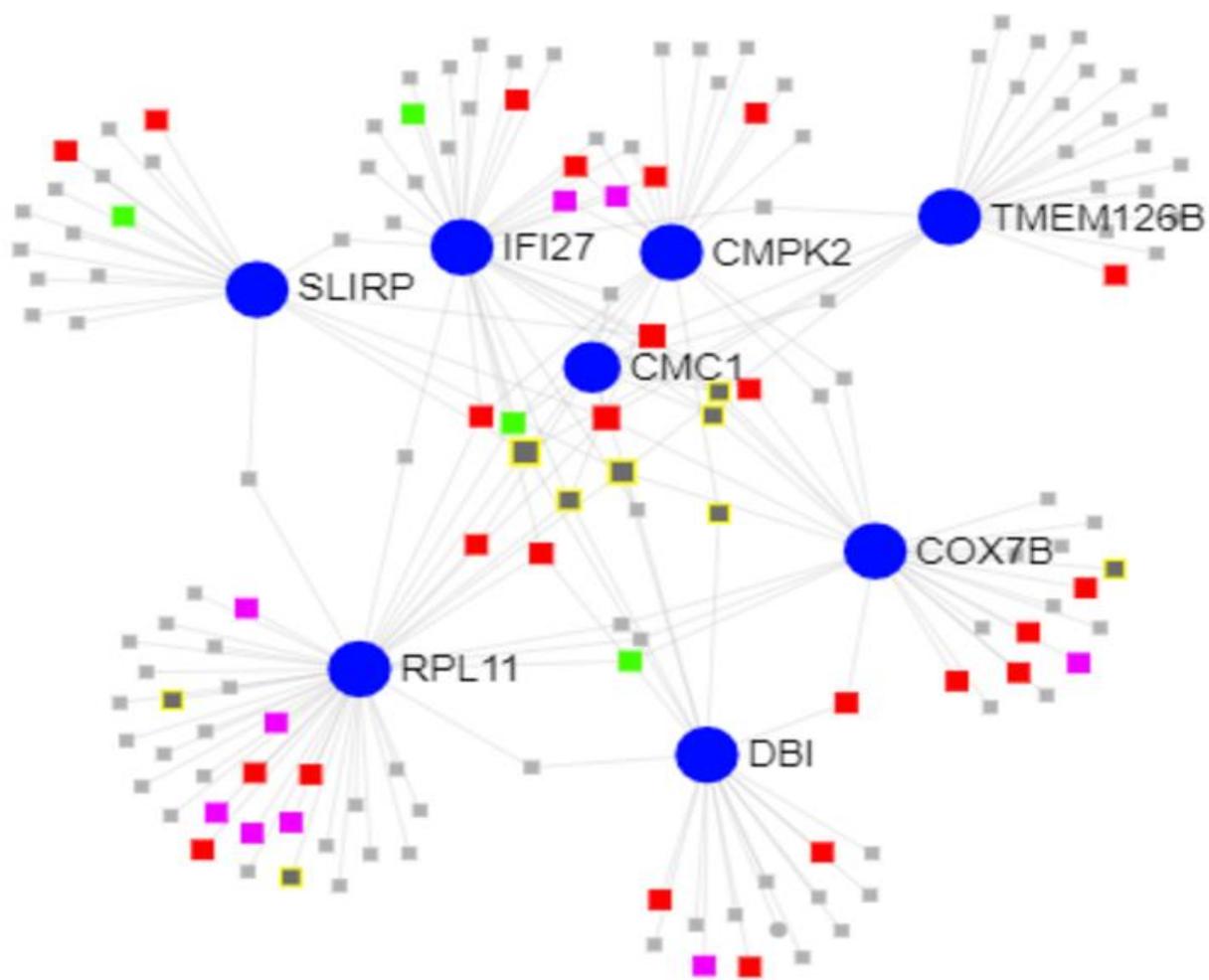


Gráfico 6: Network entre RNAs mitocondriais e miRNAs;

-  mRNAs
-  miRNAs reguladores de sinalização hormonal
-  miRNAs reguladores de remodelação cardíaca
-  miRNAs reguladores do desenvolvimento embrionário
-  miRNAs reguladores do metabolismo, incluindo o ciclo de krebs

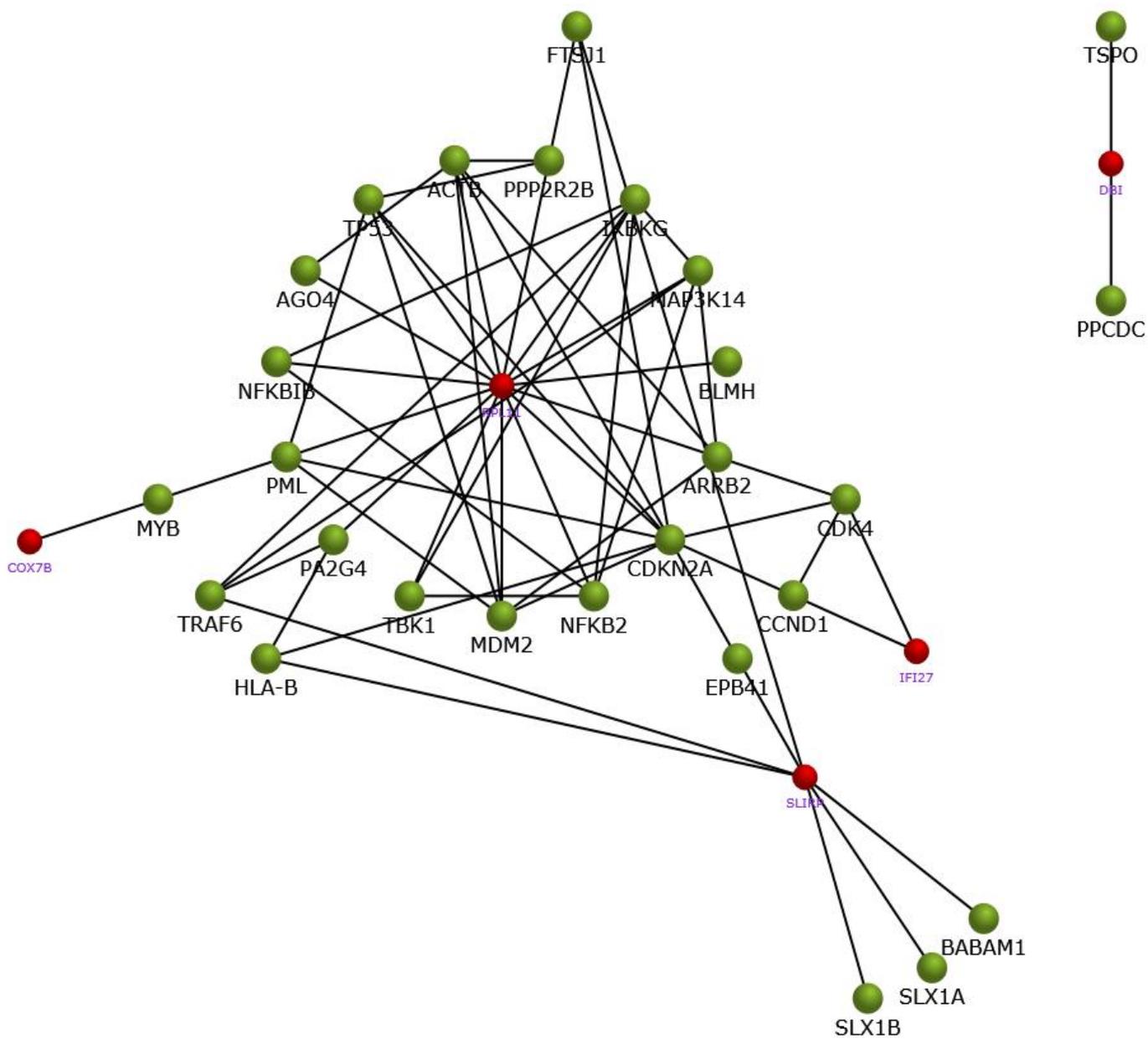


Figura 22. Network com 5 genes mitocôndrias (vermelhos) que interagem com outras vias metabólicas.

A próxima etapa foi correlacionar os genes mitocondriais com os dados laboratoriais dos grupos CTL e HT.

Resultados da correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH e T4L do grupo CTL. A significância de todos os testes de correlações obteve um $p < 0,05$.

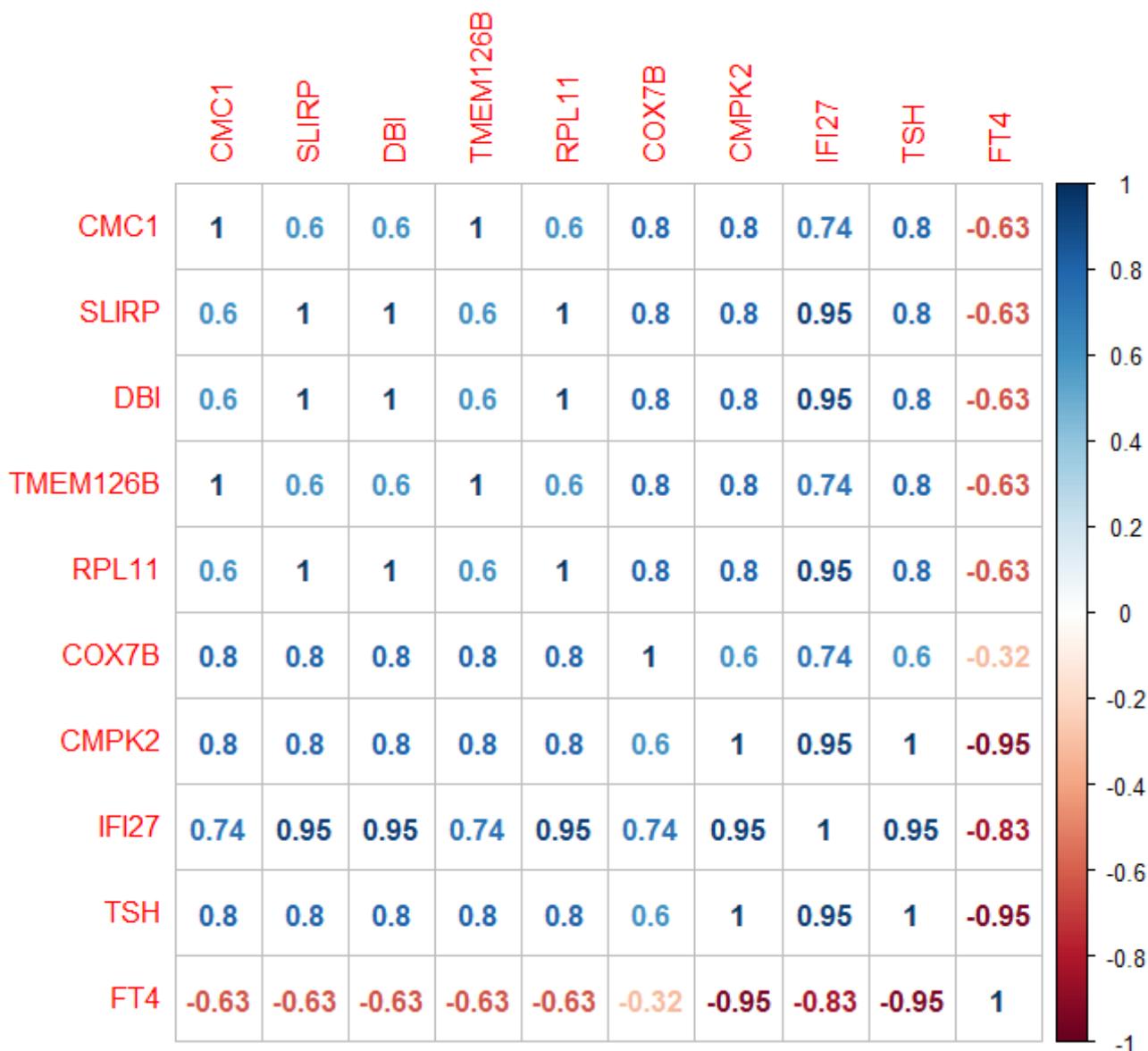


Gráfico 7. Correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH e T4L do grupo CTL, todas correlações em $p < 0,05$.

Resultados da correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH, T4L, Anti TgAb e Anti TPOAb do grupo HT. A significância de todos os testes de correlações obteve um $p < 0,05$.

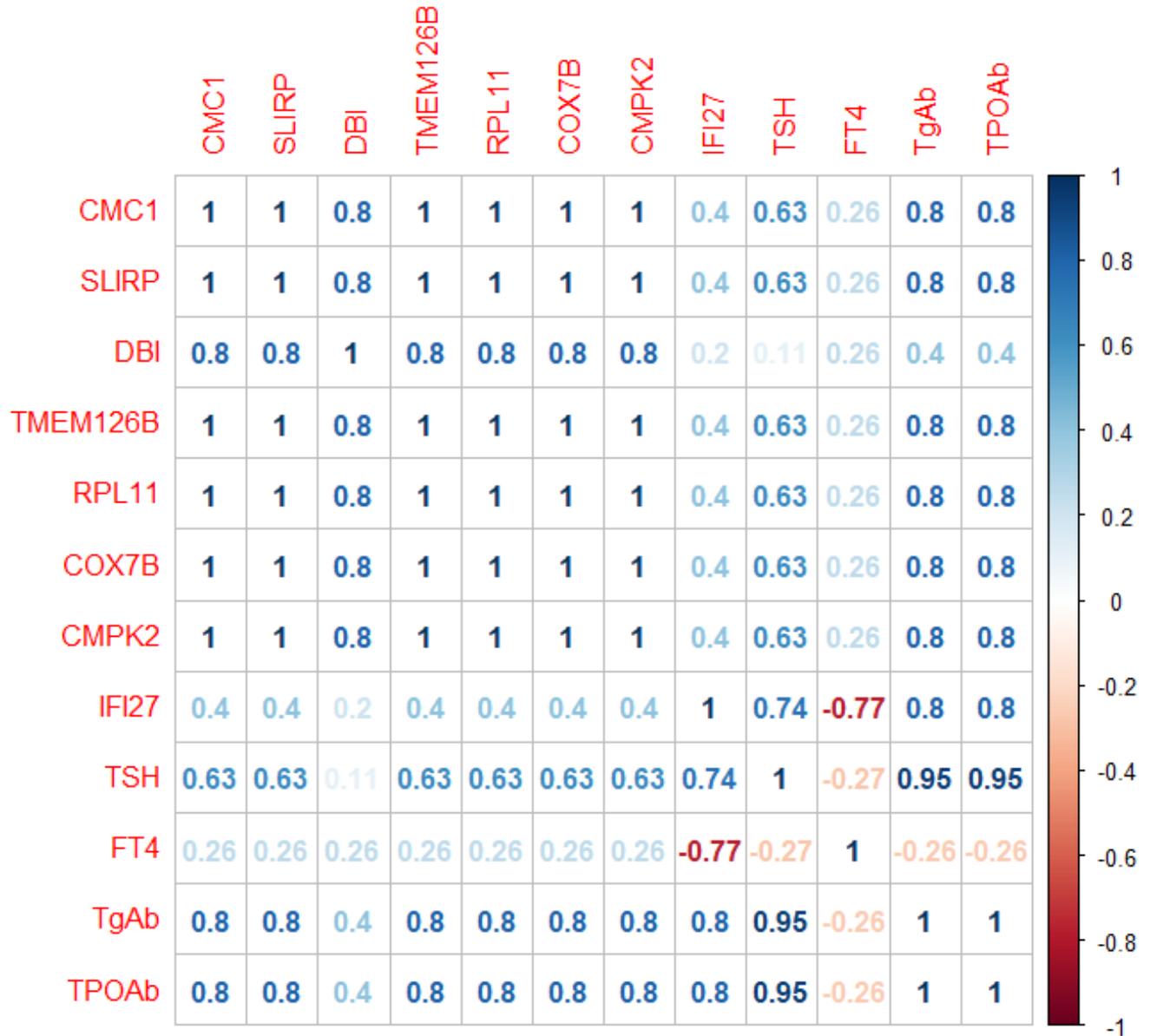


Gráfico 8. Correlação entre os *reads* dos genes mitocondriais com os resultados laboratoriais de TSH, T4L, Anti TgAb e Anti TPOAb do grupo HT, todas correlações em $p < 0,05$.

A análise seguinte foi a identificação de genes mitocondriais comuns entre as bibliotecas CTL e HT e as do estudo do estado da tiroide.

Na Figura 22, temos o diagrama *Venn* comparando a presença de DGE mitocondrias de nossas bibliotecas (HT comparada ao CTL) com a lista DGE do estudo estado da tiroide. Identificados 4 genes mitocondriais: *RPL11*, *COX7B*, *CMPK2* e *IFI27*.

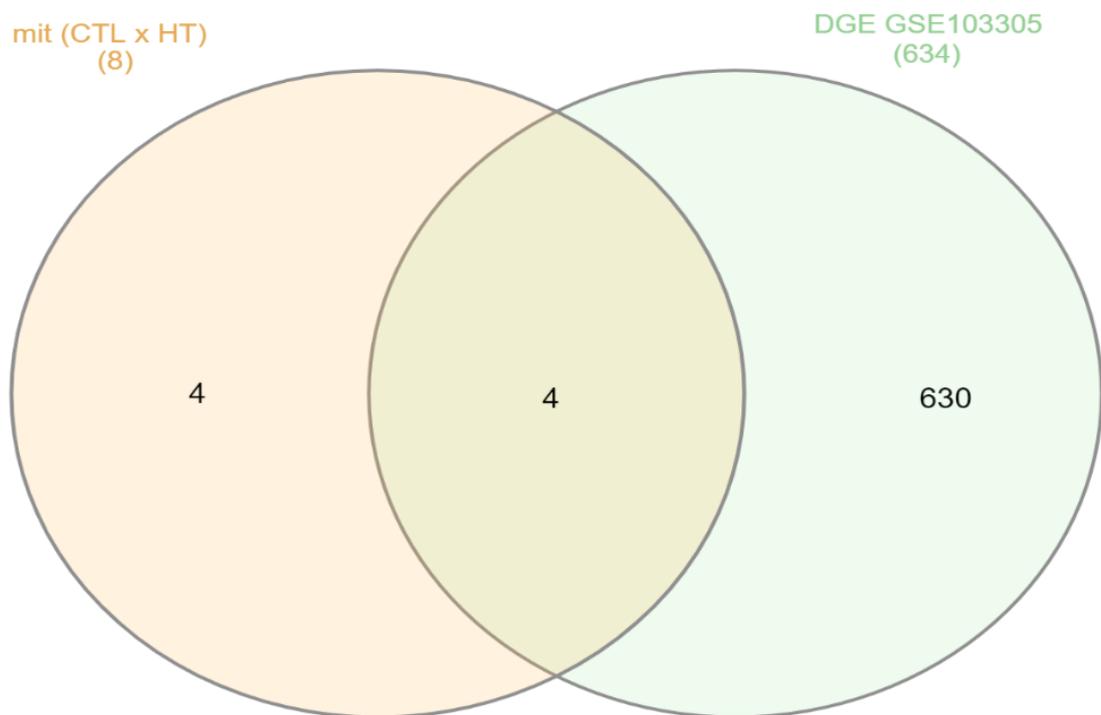


Figura 23. Identificação de genes mitocôndrias entre HT comparado ao CTL e lista DGE do estudo estado da tiroide.

A seguir apresentamos a lista de DGE mitocondriais das bibliotecas dos grupos HT e LT-4 no alvo terapêutico, conforme as expressões das duas análises.

Podemos observar uma inversão de valores nos genes *TMEM126B*, *COX7B*, *CMPK2* e *IFI27*.

Tabela 8. Lista de DGE mitocondriais da biblioteca HT comparado a LT-4 no alvo terapêutico, conforme as expressões das duas análises.

Gene	LT-4 no alvo terapêutico	HT
	logFC	logFC
<i>CMC1</i>	0,941	1,811
<i>SLIRP</i>	Não presente no resultado	1,734
<i>DBI</i>	1,417	1,55
<i>TMEM126B</i>	-3,722	1,506
<i>RPL11</i>	Não presente no resultado	1,42
<i>COX7B</i>	-5,73	1,399
<i>CMPK2</i>	-6,421	-1,492
<i>IFI27</i>	6,58	-1,87

A seguir apresentamos comparação das expressões dos mitGenes entre o grupo HT e banco GSE131411 no T2 e T3.

Tabela 9. Comparação das expressões dos mitGenes do grupo HT e banco GSE131411 no T2 e T3.

mitGene	HT	T2 (2º dia UTI)	T3 (7º dia UTI)
	logFC	logFC	logFC
<i>CMC1</i>	1,811	Não presente	Não presente
<i>SLIRP</i>	1,734	Não presente	Não presente
<i>DBI</i>	1,55	Não presente	Não presente
<i>TMEM126B</i>	1,506	Não presente	Não presente
<i>RPL11</i>	1,42	Não presente	Não presente
<i>COX7B</i>	1,399	Não presente	Não presente
<i>CMPK2</i>	-1,492	Não presente	Não presente
<i>IFI27</i>	-1,87	-1,262	-3,342

A maior presença de sequências responsivas ao T3 em um gene, sugere uma maior capacidade de responder à presença do hormônio. Na tabela 9 buscamos explorar o número de TREs em RNAs mitocondriais.

Tabela 10: Avaliação da possível influência das sequências responsivas ao T3. A sequência de maior destaque é a “TCCAGT” representada pelo gene *CMC1*.

Gene	Cromossomo	Mitocôndria	Núcleo	TRE	AGGTCA	TCCAGT	TGACCT	ACTGGA
<i>CMC1</i>	3p24.1	Presente	Ausente	95	21	32	26	16
<i>SLIRP</i>	14q24.3	Presente	Ausente	74	22	16	27	9
<i>DBI</i>	2q14.2	Presente	Ausente	11	3	2	3	3
<i>TMEM126B</i>	11q14.1	Presente	Ausente	8	2	3	1	2
<i>RPL11</i>	1p36.11	Presente	Presente	6	1	0	3	2
<i>COX7B</i>	Xq21.1	Presente	Ausente	11	2	2	6	1
<i>CMPK2</i>	2p25.2	Presente	Ausente	41	8	16	7	10
<i>IFI27</i>	14q32.12	Presente	Presente	18	7	4	3	4

5.4 Multiômica

Metilômica

A análise metilômica revelou que dos 8 genes de interesses, 3 não possuíam regulação próxima da sua região promotora, sugerindo alta atividade transcricional (produção de mRNA)

Tabela 11: Contagem da densidade das ilhas CPGs e o seu tamanho em nucleotídeos.

Gene	CpG UCSC	CpG Tiroide NCBI GSE127306	Tamanho (NT - Tiroide) NCBI GSE127306	Tamanho (NT - UCSC)
<i>TMEM126B</i>	0	0	0	0
<i>COX7B</i>	0	0	0	0
<i>DBI</i>	85	61	995	1046
<i>CMC1</i>	60	60	889	631
<i>SLIRP</i>	47	56.4	784	496
<i>RPL11</i>	17	56.4	892	212
<i>CMPK2</i>	0	72.1	1400	0
<i>IFI27</i>	0	0	0	0

Metabolômica

Dos 8 genes observados anteriormente como comuns entre as análises, 3 deles possuíam associações metabólicas na mesma via biológica, são eles: *DBI*, *TMEM126B* e *COX7B*. Por sua vez, estes genes participam do mesmo processo biológico que depende da presença de citrato, glicerol e myo-inositol. Os detalhes da expressão e quais vias podem ser encontrados na tabela 10.

Tabela 12: Comparação entre RNAs e metabólitos relacionados que estão na mesma via biológica. O * indica valor retirado de trabalhos externos.

Gene	RNA	Via Metabólica	Nome	Metabólito	
	logFC			Expressão Hipo Antes do T4 (logFC)	Expressão Hipo Após T4 (logFC)
<i>DBI</i>	1.54 (-1.5*)	Beta Oxidação de Ácidos Graxos	Glicerol;	0.368;	0.606;
			Citrato;	0.391;	0.383;
			Myo-Inositol	0.364	0.393
<i>TMEM126B</i>	1.5	Cíclo do Ácido Cítrico e Cadeia Respiratória; Biogênese do Complexo I	Citrato	0.391	0.383
<i>COX7B</i>	1.39 (-0.88*)	Síntese de ATP e Produção de Calor Pelo Desacoplamento de Proteínas	Glicerol	0.368;	0.606;

6 DISCUSSÃO

No presente trabalho, realizamos um transcriptoma a partir do sangue total de pacientes hipotiroídianos com a finalidade de construir um painel contendo o quadro fisiológico de diversos tecidos de uma maneira simultânea. O transcriptoma foi realizado na plataforma *Ion Proton Torrent*, uma plataforma capaz de detectar os nucleotídeos presentes na reação através da alteração de Potencial Hidrogeniônico (pH), dado que cada componente da sequência ATCG possui um pH distinto. A estratégia capaz de selecionar o maior número de genes relacionados ao hormônio tireoídiano e que fosse capaz de excluir genes de menor interesse diminuindo o Erro Tipo I (Falso Positivo), foi considerada a melhor estratégia. O edgeR utiliza uma adaptação de Poisson para controlar a variação, posteriormente é possível utilizar modelos da família *Generalized Linear Model* (GLM) para ajustar e coletar informações a respeito da variação do estudo. Os modelos podem ser classificados como dispersão comum, onde a variação global (de todos os genes) é ajustada em uma única reta; Dispersão de tendência (tendência geral da variação) e a variação gene-a-gene ou *tagwise*. A variação global, aqui classificada como dispersão comum, serve como referência para a dispersão de tendência, e portanto, uma combinação das duas do ponto de vista estatístico é o recomendado. A variação gene-a-gene pode ser mais conservadora, pois leva em consideração a variação gênica do nível mais baixo ao mais alto, além da variação entre os próprios genes. Em nossas análises, a variação gene-a-gene produziu 0 genes diferencialmente expressos, indicando uma barreira para corrigir a variação encontrada. Por esse motivo, optamos por comparar a dispersão comum e a dispersão de tendência.

Além das características singulares de cada situação, a variação encontrada entre as contagens mudou drasticamente, como podemos observar nas dispersões de tendência presente nas figuras A das composições 4, 5, 6, 10, 11, 15, 16. De acordo com esses gráficos, quando removemos a dispersão comum, a dispersão de tendência torna-se excessiva, fugindo dos limites que a dispersão comum é capaz de conceder para controle interno. Além disso, observamos que a intervenção com a levotiroxina favorece a diminuição da dispersão e da variação da expressão gênica, sugerindo uma maior estabilidade ou o desligamento da expressão de determinados RNAs. Outro ponto a ser considerado a respeito das comparações é o número de amostras, o que evidencia um

grande impacto da capacidade de cobertura. A cobertura em sequenciamento de nova geração é o evento que caracteriza a capacidade de uma tecnologia de coletar informações do material estudado, independentemente de ser RNA ou DNA. Acidentes no processo podem fazer com que a cobertura não seja suficiente, como uma amplificação abaixo do esperado, vencimento do kit utilizado ou o painel selecionado, de forma que a cobertura e profundidade são tópicos de constantes investigações (80-82). Dependendo do número de amostras utilizadas avaliações transcriptômicas, essa falta de cobertura pode ser compensada na estatística. Os nossos maiores números de genes diferencialmente expressos foram obtidos na população bipolar (GSE124326), que possui ao todo 480 amostras, demonstrando uma contribuição que o tamanho amostral pode exercer sobre o teste de hipótese. Dentro dos testes de hipóteses explorados, estão o exactTest e o Teste de Verossimilhança. O exactTest técnica assume que as contagens possuem uma distribuição binomial negativa, ou seja, dado um número total de eventos, estima-se o número de sucessos antes que ocorra um erro. No transcriptoma, o fenômeno é traduzido como: dado um número total de genes, consideramos a variação de um determinado gene caso ela esteja próxima de certa faixa, estando presente nessa faixa, o pesquisador estipula um ponto de corte, como $P < 0.05$ para atribuir uma significância e classificá-lo como diferencialmente expresso. O Teste de Verossimilhança possui outra maneira de considerar as informações, nesse teste, não apenas as variações individuais são levadas em consideração, mas a área em que ela se encontra e os genes que estão próximos dos outros. Como resultado, diferente do exactTest, obtemos uma coluna adicional com a Taxa de Similaridade (LR – *Likelihood Ratio*), e essa taxa de similaridade deve acompanhar uma curva similar à curva gaussiana. Por fim, comparando os resultados gerados por cada modelo, observamos que no nosso estudo de hipotireoidismo, a combinação das dispersões comum e tendência e o teste de hipótese de Verossimilhança produziram o maior número de genes diferencialmente expressos (Figura 7). Entretanto, o mesmo evento não foi observado na população de levotiroxina (Figura 12), onde a dispersão comum teve vantagem sobre a dispersão de tendência. A dispersão de tendência encontrou genes que a dispersão de comum não encontrou em todos os cenários, possibilitando um número de genes diferencialmente expressos ligeiramente conservador, assim como a possibilidade de identificar novas descobertas.

As regiões codificadoras de genes que não produzem proteína, classificados como RNAs não codificantes, tem sido alvo de grande atenção na literatura (83-85). Esses RNAs atuam como grandes reguladores da função de RNAm, e podem empregar papel importante na doença e na homeostasia. Para detectar RNAs não codificantes que atuam sobre os mensageiros, optamos por realizar um transcriptoma que capture diversos tipos de RNAs. Como demonstrado nas tabelas 4 e 5, a maior parte dos nossos RNAs não codificantes presentes no sangue e considerados diferencialmente expressos são da classe sno/scaRNA. Esses RNAs são responsáveis por controlar o *splicing* e a maturação do RNAm antes que esta saia do núcleo para traduzir uma proteína. O snoRNA de destaque foi o *SNORD15B*, que é controlado por diversos miRNAs, sendo os mais importantes o *MIR140* e os *MIR517*. O *MIR140* atua na regulação cardíaca e desenvolvimento cardíaco, enquanto os *MIR517* possuem implicância na pré eclâmpsia e controle placentário (86-92). A respeito do *SNORD15B*, *Saffari et al 2019* (93) demonstrou através de RNA-Seq em sangue total coletado de pacientes gêmeos idênticos autistas uma diferença de expressão do *SNORD15B* com uma perda de -0.85, os nossos achados apontam um ganho de 1.4 através do sangue total de hipotiroídianos, o grupo de *Saffari* utilizou o edgeR as mesmas instruções que utilizamos. Outro aspecto dos RNAs não codificantes explorados foram os miRNAs e a sua regulação sobre RNAs mitocondriais. Observamos uma variedade ampla de regulações, e classificamos em cores sendo vermelho para os reguladores da sinalização hormonal, verde para os miRNAs reguladores de remodelação cardíaca, roxo para os miRNAs reguladores do desenvolvimento embrionário e cinza para os miRNAs reguladores do metabolismo (Gráfico 7). Os genes *COX7B* e *CMC1* concentram a maior parte dos vermelhos.

Os gráficos 1, 2, 3 e 4 trazem uma caracterização geral dos 8 genes mitocondriais identificados. Observamos que o seu tecido de expressão é bastante amplo, nos chamando a atenção para cérebro, fígado e coração. A influência do hormônio tireoídiano no cérebro, conforme relatada anteriormente, é especialmente íntima da sinapse, enquanto no fígado é conhecida a sua relação com citocromo P450 e no coração, a sua relação com o nodo sino atrial para estimular o batimento cardíaco. Além disso, são genes com capacidade de atuar como receptores e enzimas, portanto, possuem relação com a geração de energia e metabolismo.

Como o hormônio tireoideano possui um efeito global sobre os tecidos, ele também exerce uma influência importante sobre as organelas, em especial, as mitocôndrias (94-97). Os genes mitocondriais identificados em nossas estratégias estão descritos na tabela 7. Através da síntese de ATP no ciclo de Krebs, a mitocôndria é capaz de manter a atividade do transporte ativo das bombas localizadas na membrana plasmática celular, alimentar reações enzimáticas e de dar suporte para a divisão celular, que por sua vez, utiliza Guanosina Trifosfato (GTP). Devido ao seu papel essencial para o desenvolvimento da célula eucariótica, os genes com grande impacto na função mitocondrial se tornaram o principal objetivo do estudo. Apesar de todos os 8 genes possuírem relação com miRNAs, apenas 5 possuem conexão com outros mensageiros presentes na análise, sendo eles: *SLIRP*, *DBI*, *RPL11*, *COX7B* e *IFI27* (Figura 21), a hipótese que consideramos é que os restantes: *CMC1*, *TMEM126B* e *CMPK2* estejam mais confinados na matriz mitocondrial e retículo endoplasmático.

Observamos que todos os genes mitocondriais presentes no grupo Eutiroidiano (Controle - CTL) apresentaram um R negativo em relação ao T4, com destaque para o *CMPK2*; Enquanto o Hipotireoidismo (HT) apresentou um R de mesmo valor para todos os genes, sendo $R=0.26$. Valores positivos indicam que as variáveis clínicas (TSH, FT4, TPO e Tg) seguem em uma mesma direção na reta, enquanto valores negativos indicam que as direções são opostas, de forma que, a medida que um aumenta, o outro diminui. Assim sendo, no Eutiroidiano presume-se que a suficiência do hormônio tem potencial para diminuir a capacidade de expressão de alguns genes, quando comparados com o *Fold Change* (FC), observamos uma correspondência entre o valor de R e o valor de FC do gene *IFI27*, sendo *down-regulated*. O *IFI27* é um gene com funções inflamatórias e a sua diminuição possui implicações em pré eclampsia (98, 99). O *IFI27* apresenta perda de expressão no hipotireoidismo quando comparado ao indivíduo eutiroidiano, no entanto, quando comparado ao indivíduo que faz uso da levotiroxina em dosagens recomendadas “no alvo”, o gene apresenta ganho de expressão, indicando uma piora da sua função quando o tratamento é realizado.

Além do transcriptoma, 3 dos 8 genes identificados na mitocôndria possuem relações com metabólitos que tem funções importantes no metabolismo mitocondrial e sofrem influência do hormônio tireoideano (Tabela 10). Para comparar o resultado do gene *DBI*, utilizamos o estudo de levotiroxina. Quando submetidos ao tratamento com T3, o

RNA produzido pelo gene *DBI* diminuiu expressão para -1.5 *fold change* no hipotálamo de ratos (100), enquanto em uma situação de falta de hormônio tireoidiano, o gene tem um ganho para 1.54 em sangue total. O RNA produzido pelo gene *COX7B* faz parte da síntese de ATP e desacoplamento de proteínas na matriz mitocondrial, gerando calor. Podemos observar inversão semelhante quando os indivíduos são submetidos à levotiroxina. Os resultados sugerem que o hormônio tireoidiano exógeno tem potencial para diminuir a expressão de RNAs mitocondriais, enquanto as vias metabólicas permanecem ativas. Outra possibilidade apresentada ocorre pela mudança da expressão de Glicerol, que sob efeito do hormônio tireoidiano, pode dobrar. O Glicerol é um importante componente para o metabolismo de glicose, e o mesmo possui influência da expressão de *DBI* e *COX7B*. Quando observado o ciclo metabólico da glicose, o glicerol é importante no metabolismo de ácidos graxos, e o seu aumento sugere uma maior atividade desse metabolismo. A diminuição do *DBI* e *COX7B* implicam que uma compensação ocorre no ciclo do ácido cítrico ou ciclo de krebs, de forma que, apesar da entrada de glicose na célula, o seu metabolismo e a geração de ATP pode ser ficar mais lenta. A expressão do *TMEM126B* sugere um aumento da capacidade de transporte do meio extracelular para o meio intracelular.

Complementando os achados anteriores, o tamanho das ilhas CpGs demonstradas na Tabela 11, sugerem uma maior ou menor atividade transcricional de um gene. Observamos que o *TMEM126B*, *COX7B* e *IFI27* não possuem ilhas CpGs tanto no UCSC Browser quanto no metiloma realizado em tecido tireoidiano, contrastando com o gene *DBI*, que possui uma função importante no metabolismo da glicose a partir dos ácidos graxos, e de acordo com o metiloma, uma regulação maior da sua expressão.

Diante do caso grave de choque séptico (pacientes em UTI) o transcriptoma pareado (T2: 2º dia e T3: 7º dia) observamos que os mitGenes não apresentaram expressões (exceto *IFI27*), caracterizando uma diminuição importante das funções mitocondriais. Doenças em estágio grave, podem alterar o controle de qualidade mitocondrial. O mitGene *IFI27* apresentou cronologicamente uma diminuição de expressão progressiva no choque séptico (GSM131411). Este gene é intimamente ligado aos processos inflamatórios, e na sua função mitocondrial, a sua perda de expressão relaciona sua pouca participação na resposta imunológica diante o agravamento da sepse (78).

A ação genômica do T3 ocorre por ligação com receptores nucleares específicos que estimulam o processo de transcrição gênica (101). Observando as subclasses específicas desses receptores, encontramos: receptores do retinoide X (RXR), receptor do hormônio tireoidiano (TR), receptores do ácido retinóico (RAR), receptores da Vitamina D (VDR) e receptores proliferadores peroxissomais (PPAR) (102-104). Esses receptores se ligam em sequências específicas do gene, denominadas de elementos responsivos do hormônio tireoidiano (TREs), que podem estar na forma de palíndromos ou palíndromos invertidos, são hexâmeros constituídos pela sequência AGGTCA separados por 4 ou 6 nucleotídeos diferentes, formando as seguintes configurações: AGGTCANNNNAGGTCA, TCCAGTNNNNTCCAGT, TGACCTNNNNNNAGGTCA, ACTGGANNNNNNTCCAGT, AGGTCATGACCT ou TCCAGTACTGGA (o N indica o número de nucleotídeos diversos) e normalmente encontram-se na região promotora do gene (105). Embora essas configurações sejam quase um consenso, podemos observar em certos casos que essas sequências podem estar presentes em 3' flanqueando o curso da região codificadora (106-108). Conforme Willians GR e Brent GA, podem existir uma variação considerável entre essas sequências de TREs bem como o número de espaçamento entre elas (109). Rastreando as sequências de nucleotídeos de cada mitGene encontramos regiões atípicas de TREs, localização cromossômicas e compartimento subcelular dos mitGenes, mas é notório a associação com o número de TREs atípicos (Tabela 9) como aumento da expressão como nos genes *CMC1*, *SLIRP* e *DBI*, como também a mesma associação aos genes que obtiveram uma queda em suas expressões, como o *CMPK2* e *IFI27*. O pensamento lógico de quanto maior as possibilidades de TREs, maior serão suas expressões, porém um fato que chama atenção é em genes que tem um relativo número de TREs atípicos estarem com baixas expressões, caracterizando-os como um perda real, ou até mesmo por mecanismos regulatórios diferenciando-os do que tecnicamente a fisiologia mostra.

7 CONCLUSÃO

Entre as comparações dos pacotes para análises de genes diferencialmente expressos, observamos que a utilização da ferramenta EdgeR configurada para dispersão comum, dispersão de tendência e teste de razão de verossimilhança, identifica os DGEs em quantidade mediana, evitando uma possível hipossensibilidade ou hipersensibilidade demonstradas nos outros pacotes.

Genes mitocôndrias sofrem alterações de expressões no hipotireoidismo diferente do conceito de genes estáveis.

A análise do Banco GSE131411 (UTI – Choque Séptico) revelou que a função mitocondrial fica comprometida, sugerindo um estado de eutiroidiano doente (nível tecidual), e o mitGene *IFI27* apresentou uma diminuição progressiva, refletindo a sua ineficiência mitocondrial nas respostas imunológicas, neste cenário clínico.

Os dados da metabolômica e metilômica sugerem uma regulação extra genômica que apontam uma tendência biológica do organismo buscar o equilíbrio à uma possível ineficiência/deficiência do hormônio exógeno ou endógeno.

A presença dos TREs pode ser importante legitimador da expressão gênica, tanto na presença e ação do T3 como na sua co-repressão.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Rezende JM. ["Tireóide" or "tiróide"]. *Arquivos brasileiros de endocrinologia e metabologia*. 2004;48(3):432-4.
2. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism. *Physiological reviews*. 2014;94(2):355-82.
3. Dumitrescu AM, Refetoff S. The syndromes of reduced sensitivity to thyroid hormone. *Biochimica et biophysica acta*. 2013;1830(7):3987-4003.
4. Duarte GC, Cendoroglo MS, Araújo LM, Almada Filho Cde M. Association between increased serum thyrotropin concentration and the oldest old: what do we know? *Einstein (Sao Paulo, Brazil)*. 2015;13(1):117-21.
5. Campbell NR, Dary O, Cappuccio FP, Neufeld LM, Harding KB, Zimmermann MB. Need for coordinated programs to improve global health by optimizing salt and iodine intake. *Revista panamericana de salud publica = Pan American journal of public health*. 2012;32(4):281-6.
6. Matyjaszek-Matuszek B, Pyzik A, Nowakowski A, Jarosz MJ. Diagnostic methods of TSH in thyroid screening tests. *Annals of agricultural and environmental medicine : AAEM*. 2013;20(4):731-5.
7. Estrada JM, Soldin D, Buckey TM, Burman KD, Soldin OP. Thyrotropin isoforms: implications for thyrotropin analysis and clinical practice. *Thyroid : official journal of the American Thyroid Association*. 2014;24(3):411-23.
8. Fliers E, Kalsbeek A, Boelen A. Beyond the fixed setpoint of the hypothalamus-pituitary-thyroid axis. *European journal of endocrinology*. 2014;171(5):R197-208.
9. Maia AL, Goemann IM, Meyer EL, Wajner SM. Deiodinases: the balance of thyroid hormone: type 1 iodothyronine deiodinase in human physiology and disease. *The Journal of endocrinology*. 2011;209(3):283-97.
10. Senese R, Cioffi F, de Lange P, Goglia F, Lanni A. Thyroid: biological actions of 'nonclassical' thyroid hormones. *The Journal of endocrinology*. 2014;221(2):R1-12.
11. Karapanou O, Papadimitriou A. Thyroid hormone transporters in the human. *Hormones (Athens, Greece)*. 2011;10(4):270-9.
12. Williams GR, Bassett JH. Deiodinases: the balance of thyroid hormone: local control of thyroid hormone action: role of type 2 deiodinase. *The Journal of endocrinology*. 2011;209(3):261-72.
13. Dong YH, Fu DG. Autoimmune thyroid disease: mechanism, genetics and current knowledge. *European review for medical and pharmacological sciences*. 2014;18(23):3611-8.
14. Arrojo EDR, Bianco AC. Type 2 deiodinase at the crossroads of thyroid hormone action. *The international journal of biochemistry & cell biology*. 2011;43(10):1432-41.
15. Dentice M, Marsili A, Zavacki A, Larsen PR, Salvatore D. The deiodinases and the control of intracellular thyroid hormone signaling during cellular differentiation. *Biochimica et biophysica acta*. 2013;1830(7):3937-45.
16. Brenta G, Vaisman M, Sgarbi JA, Bergoglio LM, Andrada NC, Bravo PP, et al. Clinical practice guidelines for the management of hypothyroidism. *Arquivos brasileiros de endocrinologia e metabologia*. 2013;57(4):265-91.
17. Baumgartner C, Blum MR, Rodondi N. Subclinical hypothyroidism: summary of evidence in 2014. *Swiss medical weekly*. 2014;144:w14058.

18. Biondi B. Mechanisms in endocrinology: Heart failure and thyroid dysfunction. *European journal of endocrinology*. 2012;167(5):609-18.
19. Lioudaki E, Mavroei NG, Mikhailidis DP, Ganotakis ES. Subclinical hypothyroidism and vascular risk: an update. *Hormones (Athens, Greece)*. 2013;12(4):495-506.
20. Ning N, Gao D, Triggiani V, Iacoviello M, Mitchell JE, Ma R, et al. Prognostic Role of Hypothyroidism in Heart Failure: A Meta-Analysis. *Medicine*. 2015;94(30):e1159.
21. Coceani M. Heart disease in patients with thyroid dysfunction: hyperthyroidism, hypothyroidism and beyond. *Anadolu kardiyoloji dergisi : AKD = the Anatolian journal of cardiology*. 2013;13(1):62-6.
22. Samuels MH. Psychiatric and cognitive manifestations of hypothyroidism. *Current opinion in endocrinology, diabetes, and obesity*. 2014;21(5):377-83.
23. Cardoso LF, Maciel LM, Paula FJ. The multiple effects of thyroid disorders on bone and mineral metabolism. *Arquivos brasileiros de endocrinologia e metabologia*. 2014;58(5):452-63.
24. Papaleontiou M, Haymart MR. Approach to and treatment of thyroid disorders in the elderly. *The Medical clinics of North America*. 2012;96(2):297-310.
25. Mariani LH, Berns JS. The renal manifestations of thyroid disease. *Journal of the American Society of Nephrology : JASN*. 2012;23(1):22-6.
26. Carney LA, Quinlan JD, West JM. Thyroid disease in pregnancy. *American family physician*. 2014;89(4):273-8.
27. Medenica S, Nedeljkovic O, Radojevic N, Stojkovic M, Trbojevic B, Pajovic B. Thyroid dysfunction and thyroid autoimmunity in euthyroid women in achieving fertility. *European review for medical and pharmacological sciences*. 2015;19(6):977-87.
28. Wang Y-F, Heng J-F, Yan J, Dong L. Relationship between disease severity and thyroid function in Chinese patients with euthyroid sick syndrome. *Medicine*. 2018;97(31):e11756-e.
29. Hoermann R, Midgley JE, Larisch R, Dietrich JW. Is pituitary TSH an adequate measure of thyroid hormone-controlled homeostasis during thyroxine treatment? *European journal of endocrinology*. 2013;168(2):271-80.
30. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function in subclinical hypothyroidism: importance of clinical follow-up and therapy. *European journal of endocrinology*. 2011;164(3):317-23.
31. Carvalho GA, Perez CL, Ward LS. The clinical use of thyroid function tests. *Arquivos brasileiros de endocrinologia e metabologia*. 2013;57(3):193-204.
32. Duntas LH, Jonklaas J. Levothyroxine Dose Adjustment to Optimise Therapy Throughout a Patient's Lifetime. *Advances in therapy*. 2019;36(Suppl 2):30-46.
33. Jonklaas J, Bianco AC, Bauer AJ, Burman KD, Cappola AR, Celi FS, et al. Guidelines for the treatment of hypothyroidism: prepared by the american thyroid association task force on thyroid hormone replacement. *Thyroid : official journal of the American Thyroid Association*. 2014;24(12):1670-751.
34. Maseeh uz Z, Fatima N, Sajjad Z. Amiodarone therapy: don't forget thyroid. *JPMA The Journal of the Pakistan Medical Association*. 2012;62(3):268-72.
35. Ricciuti B, Mencaroni C, Pagliarunga L, Paciullo F, Crinò L, Chiari R, et al. Long noncoding RNAs: new insights into non-small cell lung cancer biology, diagnosis and therapy. *Medical oncology (Northwood, London, England)*. 2016;33(2):18.
36. Alvarez ML, Khosroheidari M, Kanchi Ravi R, DiStefano JK. Comparison of protein, microRNA, and mRNA yields using different methods of urinary exosome

isolation for the discovery of kidney disease biomarkers. *Kidney international*. 2012;82(9):1024-32.

37. Mittelbrunn M, Sánchez-Madrid F. Intercellular communication: diverse structures for exchange of genetic information. *Nature reviews Molecular cell biology*. 2012;13(5):328-35.

38. Ruddle FH. Mapping and sequencing of the human genome. *Japanese journal of cancer research : Gann*. 1998;89(12):inside front cover.

39. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.

40. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931-45.

41. Muzzey D, Evans EA, Lieber C. Understanding the Basics of NGS: From Mechanism to Variant Calling. *Current genetic medicine reports*. 2015;3(4):158-65.

42. Passador-Gurgel G, Hsieh WP, Hunt P, Deighton N, Gibson G. Quantitative trait transcripts for nicotine resistance in *Drosophila melanogaster*. *Nature genetics*. 2007;39(2):264-8.

43. Rifkin SA, Kim J, White KP. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nature genetics*. 2003;33(2):138-44.

44. Pilarsky C, Nanduri LK, Roy J. Gene Expression Analysis in the Age of Mass Sequencing: An Introduction. *Methods in molecular biology (Clifton, NJ)*. 2016;1381:67-73.

45. Gao J, Wu H, Wang L, Zhang H, Duan H, Lu J, et al. Validation of targeted next-generation sequencing for RAS mutation detection in FFPE colorectal cancer tissues: comparison with Sanger sequencing and ARMS-Scorpion real-time PCR. *BMJ open*. 2016;6(1):e009532.

46. Lim EC, Brett M, Lai AH, Lee SP, Tan ES, Jamuar SS, et al. Next-generation sequencing using a pre-designed gene panel for the molecular diagnosis of congenital disorders in pediatric patients. *Human genomics*. 2015;9:33.

47. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2010;32(6):524-36.

48. Schatz MC, Langmead B. The DNA Data Deluge: Fast, efficient genome sequencing machines are spewing out more data than geneticists can analyze. *IEEE spectrum*. 2013;50(7):26-33.

49. Souilmi Y, Lancaster AK, Jung JY, Rizzo E, Hawkins JB, Powles R, et al. Scalable and cost-effective NGS genotyping in the cloud. *BMC medical genomics*. 2015;8:64.

50. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews Genetics*. 2009;10(1):57-63.

51. Moreton J, Izquierdo A, Emes RD. Assembly, Assessment, and Availability of De novo Generated Eukaryotic Transcriptomes. *Frontiers in genetics*. 2015;6:361.

52. Menck CFM. A nova grande promessa da inovação em fármacos: RNA interferência saindo do laboratório para a clínica. *Estudos Avançados*. 2010;24:99-108.

53. Godoy PM, Bhakta NR, Barczak AJ, Cakmak H, Fisher S, MacKenzie TC, et al. Large Differences in Small RNA Composition Between Human Biofluids. *Cell reports*. 2018;25(5):1346-58.

54. Wang ZD, Shen LP, Chang C, Zhang XQ, Chen ZM, Li L, et al. Long noncoding RNA Inc-RI is a new regulator of mitosis via targeting miRNA-210-3p to release PLK1 mRNA activity. *Scientific reports*. 2016;6:25385.

55. Sakakibara I, Santolini M, Ferry A, Hakim V, Maire P. Six homeoproteins and a linc-RNA at the fast MYH locus lock fast myofiber terminal phenotype. *PLoS genetics*. 2014;10(5):e1004386.
56. Nunnari J, Suomalainen A. Mitochondria: in sickness and in health. *Cell*. 2012;148(6):1145-59.
57. Anzell AR, Maizy R, Przyklenk K, Sanderson TH. Mitochondrial Quality Control and Disease: Insights into Ischemia-Reperfusion Injury. *Molecular neurobiology*. 2018;55(3):2547-64.
58. Yu SB, Pekkurnaz G. Mechanisms Orchestrating Mitochondrial Dynamics for Energy Homeostasis. *Journal of molecular biology*. 2018;430(21):3922-41.
59. Ali AT, Boehme L, Carbajosa G, Seitan VC, Small KS, Hodgkinson A. Nuclear genetic regulation of the human mitochondrial transcriptome. *eLife*. 2019;8.
60. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes & development*. 2011;25(10):1010-22.
61. Zhang X-w, Li Q-h, Xu Z-d, Dou J-j. Mass spectrometry-based metabolomics in health and medical science: a systematic review. *RSC Advances*. 2020;10(6):3092-104.
62. Kamburov A, Cavill R, Ebbels TM, Herwig R, Keun HC. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. *Bioinformatics (Oxford, England)*. 2011;27(20):2917-8.
63. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
64. RStudio Team. Integrated Development for R. RStudio, PBC. Boston, MA2020.
65. Robinson MD, Smyth GK. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics (Oxford, England)*. 2008;9(2):321-32.
66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014;15(12):550.
67. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: a matter of depth. *Genome research*. 2011;21(12):2213-23.
68. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*. 2009;25(16):2078-9.
69. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic acids research*. 2007;35(Database issue):D61-5.
70. Smith AC, Robinson AJ. MitoMiner v3.1, an update on the mitochondrial proteomics database. *Nucleic acids research*. 2016;44(D1):D1258-61.
71. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic acids research*. 2016;44(D1):D1251-7.
72. Pathan M, Keerthikumar S, Chisanga D, Alessandro R, Ang CS, Askenase P, et al. A novel community driven software for functional enrichment analysis of extracellular vesicles data. *Journal of extracellular vesicles*. 2017;6(1):1321455.
73. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003;13(11):2498-504.
74. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-

wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(43):15545-50.

75. Taiyun Wei and Viliam Simko. R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). 2017.

76. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC bioinformatics*. 2015;16(1):169.

77. Krebs CE, Ori APS, Vreeker A, Wu T, Cantor RM, Boks MPM, et al. Whole blood transcriptome analysis in bipolar disorder reveals strong lithium effect. *Psychological medicine*. 2019:1-12.

78. Braga D, Barcella M, Herpain A, Aletti F, Kistler EB, Bollen Pinto B, et al. A longitudinal study highlights shared aspects of the transcriptomic response to cardiogenic and septic shock. *Critical care (London, England)*. 2019;23(1):414.

79. Piras C, Pibiri M, Leoni VP, Balsamo A, Tronci L, Arisci N, et al. Analysis of metabolomics profile in hypothyroid patients before and after thyroid hormone replacement. *Journal of endocrinological investigation*. 2020.

80. LaDuca H, Farwell KD, Vuong H, Lu HM, Mu W, Shahmirzadi L, et al. Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. *PloS one*. 2017;12(2):e0170843.

81. Lim JS, Choi BS, Lee JS, Shin C, Yang TJ, Rhee JS, et al. Survey of the Applications of NGS to Whole-Genome Sequencing and Expression Profiling. *Genomics & informatics*. 2012;10(1):1-8.

82. Sun Y, Ruivenkamp CA, Hoffer MJ, Vrijenhoek T, Kriek M, van Asperen CJ, et al. Next-generation diagnostics: gene panel, exome, or whole genome? *Human mutation*. 2015;36(6):648-55.

83. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*. 2014;157(1):77-94.

84. Hombach S, Kretz M. Non-coding RNAs: Classification, Biology and Functioning. *Advances in experimental medicine and biology*. 2016;937:3-17.

85. Wei JW, Huang K, Yang C, Kang CS. Non-coding RNAs as regulators in epigenetics (Review). *Oncology reports*. 2017;37(1):3-9.

86. Boštjančič E, Zidar N, Glavač D. MicroRNAs and cardiac sarcoplasmic reticulum calcium ATPase-2 in human myocardial infarction: expression and bioinformatic analysis. *BMC genomics*. 2012;13:552.

87. Hromadnikova I, Kotlabova K, Doucha J, Dlouha K, Krofta L. Absolute and relative quantification of placenta-specific microRNAs in maternal circulation with placental insufficiency-related complications. *The Journal of molecular diagnostics : JMD*. 2012;14(2):160-7.

88. Hromadnikova I, Kotlabova K, Ondrackova M, Kestlerova A, Novotna V, Hymanova L, et al. Circulating C19MC microRNAs in preeclampsia, gestational hypertension, and fetal growth restriction. *Mediators of inflammation*. 2013;2013:186041.

89. Li J, Li Y, Jiao J, Wang J, Li Y, Qin D, et al. Mitofusin 1 is negatively regulated by microRNA 140 in cardiomyocyte apoptosis. *Molecular and cellular biology*. 2014;34(10):1788-99.

90. Rustagi Y, Jaiswal HK, Rawal K, Kundu GC, Rani V. Comparative Characterization of Cardiac Development Specific microRNAs: Fetal Regulators for Future. *PloS one*. 2015;10(10):e0139359.

91. Yang S, Li H, Chen L. MicroRNA-140 attenuates myocardial ischemia-reperfusion injury through suppressing mitochondria-mediated apoptosis by targeting YES1. *Journal of cellular biochemistry*. 2019;120(3):3813-21.
92. Yang S, Li H, Ge Q, Guo L, Chen F. Deregulated microRNA species in the plasma and placenta of patients with preeclampsia. *Molecular medicine reports*. 2015;12(1):527-34.
93. Saffari A, Arno M, Nasser E, Ronald A, Wong CCY, Schalkwyk LC, et al. RNA sequencing of identical twins discordant for autism reveals blood-based signatures implicating immune and transcriptional dysregulation. *Molecular autism*. 2019;10:38.
94. Lanni A, Moreno M, Goglia F. Mitochondrial Actions of Thyroid Hormone. *Comprehensive Physiology*. 2016;6(4):1591-607.
95. Noli L, Khorsandi SE, Pyle A, Giritharan G, Fogarty N, Capalbo A, et al. Effects of thyroid hormone on mitochondria and metabolism of human preimplantation embryos. *Stem cells (Dayton, Ohio)*. 2020;38(3):369-81.
96. Wrutniak-Cabello C, Casas F, Cabello G. Thyroid Hormone Action: The p43 Mitochondrial Pathway. *Methods in molecular biology (Clifton, NJ)*. 2018;1801:163-81.
97. Yau WW, Singh BK, Lesmana R, Zhou J, Sinha RA, Wong KA, et al. Thyroid hormone (T₃) stimulates brown adipose tissue activation via mitochondrial biogenesis and MTOR-mediated mitophagy. *Autophagy*. 2019;15(1):131-50.
98. Kronborg CS, Knudsen UB, Martensen PM. Differential regulation of the interferon induced gene ISG12A by serum from healthy and preeclamptic pregnancies. *Cytokine*. 2008;42(1):105-12.
99. Ugolini-Lopes MR, Torrezan GT, Gândara APR, Olivieri EHR, Nascimento IS, Okazaki E, et al. Enhanced type I interferon gene signature in primary antiphospholipid syndrome: Association with earlier disease onset and preeclampsia. *Autoimmunity reviews*. 2019;18(4):393-8.
100. Herwig A, Campbell G, Mayer CD, Boelen A, Anderson RA, Ross AW, et al. A thyroid hormone challenge in hypothyroid rats identifies T₃ regulated genes in the hypothalamus and in models with altered energy balance and glucose homeostasis. *Thyroid : official journal of the American Thyroid Association*. 2014;24(11):1575-93.
101. Hennemann G, Docter R, Friesema EC, de Jong M, Krenning EP, Visser TJ. Plasma membrane transport of thyroid hormones and its role in thyroid hormone metabolism and bioavailability. *Endocrine reviews*. 2001;22(4):451-76.
102. Arai K, Chrousos GP. Syndromes of glucocorticoid and mineralocorticoid resistance. *Steroids*. 1995;60(1):173-9.
103. Mangelsdorf DJ, Evans RM. The RXR heterodimers and orphan receptors. *Cell*. 1995;83(6):841-50.
104. Olefsky JM. Nuclear receptor minireview series. *The Journal of biological chemistry*. 2001;276(40):36863-4.
105. Ribeiro RC, Apriletti JW, Wagner RL, West BL, Feng W, Huber R, et al. Mechanisms of thyroid hormone action: insights from X-ray crystallographic and functional studies. *Recent progress in hormone research*. 1998;53:351-92; discussion 92-4.
106. Bigler J, Eisenman RN. Novel location and function of a thyroid hormone response element. *The EMBO journal*. 1995;14(22):5710-23.
107. Yen PM. Physiological and molecular basis of thyroid hormone action. *Physiological reviews*. 2001;81(3):1097-142.

108. Zhang W, Brooks RL, Silversides DW, West BL, Leidig F, Baxter JD, et al. Negative thyroid hormone control of human growth hormone gene expression is mediated by 3'-untranslated/3'-flanking DNA. *The Journal of biological chemistry*. 1992;267(21):15056-63.
109. Brent GA, Williams GR, Harney JW, Forman BM, Samuels HH, Moore DD, et al. Effects of varying the position of thyroid hormone response elements within the rat growth hormone promoter: implications for positive and negative regulation by 3,5,3'-triiodothyronine. *Molecular endocrinology (Baltimore, Md)*. 1991;5(4):542-8.

ANEXO 1

Extração do RNA total

- 1 Centrifugação dos tubos *PAXgene blood RNA*, por 10 minutos à 3500 Rotações por minutos (RPM, sempre utilizando um rotor de balanço exterior com adaptadores de tubos de mesmo volume (balanceamento da centrífuga).
- 2 Retirado o sobrenadante por meio de decantação e visualizado o *pellet* ao fundo do tubo.
- 3 Adicionado 4mL de água livre de RNase (RNFW), e fechado o tubo com uma nova tampa *Hemogard BD* secundária, fornecidos pelo *Kit*.
- 4 Misturado ao vórtex até a visível dissolução do *pellet*.
- 5 Centrifugado os tubos *PAXgene blood RNA*, por 10 minutos à 3500RPM, sempre utilizando um rotor de balanço exterior com adaptadores de tubos de mesmo volume (balanceamento da centrífuga).
- 6 Retirado e eliminado o sobrenadante, sem interferir com o *pellet*.
- 7 Adicionado 350µL de tampão de ressuspensão (BR1), fornecido pelo *Kit*.
- 8 Misturado ao vórtex até a visível dissolução do *pellet*.
- 9 Transferido a amostra (350µL) para um microtubo (MCT) de 1,5mL.
- 10 Adicionado 300µL de tampão de ligação (BR2), e 40µL proteinase K (PK), fornecidos pelo *Kit*.
- 11 Misturado ao vórtex por 5 segundos.
- 12 Incubado por 10 minutos à 55°C em agitador incubador, sob velocidade de aproximadamente 400RPM.
- 13 Transferido o lisado diretamente para uma coluna de rotação do homogeneizador *PAXgene* (coluna de cor lilás PSC) e colocada em um tubo de processamento (PT) de 2mL, fornecidos pelo *Kit*.
- 14 Centrifugado por 3 minutos à 15000RPM em microcentrífuga, adequadamente balanceada.
- 15 Transferido cuidadosamente a totalidade do sobrenadante da fração residual para um MCT de 1,5mL, sem interferir com o *pellet* do PT.
- 16 Adicionado 350µL de etanol 99% grau de pureza.
- 17 Misturado ao vórtex por 5 segundos.
- 18 Realizado uma breve centrifugação para remover as gotas presentes no interior da tampa do MCT. Centrifugado em microcentrífuga por 3 segundos (SPIN).

- 19 Transferido 700µL da amostra para outra coluna de rotação de RNA *PAXgene* (coluna de cor vermelha PRC), colocada em um tubo de processamento (PT) de 2mL.
- 20 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 21 Colocado a coluna (PRC) em um novo tubo de processamento (PT) de 2mL e eliminado o PT antigo contendo o produto residual.
- 22 Transferido a amostra restante à coluna PRC.
- 23 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 24 Colocado a coluna PRC em um novo PT de 2mL, e eliminado o PT antigo contendo o produto residual.
- 25 Transferido 350µL de tampão de lavagem 1 (BR3) para a coluna PRC.
- 26 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 27 Colocado a coluna PRC em um novo PT de 2mL, e eliminado o PT antigo contendo o produto residual.
- 28 Preparado o *mix* de DNase I e tampão de digestão de DNA (RDD), fornecidos pelo *Kit*. Em um MCT de 1,5mL transferimos 10µL de solução de DNase com 70µL de tampão de digestão de DNA (RDD) por amostra, totalizando um volume de *mix* de 80µL. Homogeneizando suavemente com a mão, pois a DNase I é especialmente sensível a desnaturação física.
- 29 Transferido os 80µL do *mix* diretamente para a membrana da coluna PRC e aguardarmos por 15 minutos, com a amostra em temperatura ambiente.
- 30 Transferido 350µL de tampão de lavagem 1 (BR3) para a coluna PRC, fornecidos pelo *Kit*.
- 31 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 32 Colocado a coluna PRC em um novo PT de 2mL, e eliminado o PT antigo contendo o produto residual.
- 33 Transferido 500µL de tampão de lavagem 2 (BR4) para a coluna de rotação PRC, fornecidos pelo *Kit*.

- 34 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 35 Colocado a coluna PRC em um novo PT de 2mL, e eliminado o PT antigo contendo o produto residual.
- 36 Transferido mais 500µL de tampão de lavagem 2 (BR4) para a coluna de rotação PRC, fornecidos pelo *Kit*.
- 37 Centrifugado por 3 minutos à 15000RPM em microcentrífuga, adequadamente balanceada.
- 38 Colocado a coluna PRC em um novo PT de 2mL, e eliminado o PT antigo contendo o produto residual.
- 39 Centrifugado por 1 minuto em microcentrífuga adequadamente balanceada, com 15000RPM.
- 40 Eliminado o PT antigo contendo o produto residual.
- 41 Colocado a coluna PRC em um novo MCT de 1,5mL.
- 42 Transferido 40µL de tampão de eluição (BR5) diretamente para membrana da coluna PRC, fornecidos pelo *Kit*.
- 43 Centrifugado por 1 minuto à 15000RPM em microcentrífuga, adequadamente balanceada.
- 44 Repetido novamente o passo da eluição, transferindo mais 40µL de tampão de eluição (BR5) diretamente para membrana da coluna PRC, fornecidos pelo *Kit*, utilizando o mesmo MCT de 1,5mL.
- 45 Incubado o MCT contendo o eluído por 5 minutos à 65°C.
- 46 Após a incubação, arrefecemos imediatamente o MCT em gelo por 3 minutos.
Imediatamente após o arrefecimento armazenamos as amostras (todas devidamente identificadas) com o extraído de RNA em *freezer* -80°C.

ANEXO 2

Testes Bioquímicos

1 - *Kit Roche Diagnostics* para dosagem de anticorpo dirigido contra a tiroglobulina (anti-TG) Referência: 6368697 190, resultados em UI/mL, detecção do *Kit*: 10.0-4000 UI/mL, valores teóricos normais: menor que 115 UI/mL

2 - *Kit Roche Diagnostics* para dosagem de anticorpos anti-peroxidase (anti-TPO) da tiroide Referência: 06368590 190, resultados em UI/mL, detecção do *Kit*: 5.00 - 600 UI/mL, valores teóricos normais: menor que 34 UI/mL.

3 - *Kit Roche Diagnostics* para dosagem do hormônio estimulante da tiroide (TSH) Referência: 11731459 122, resultados em >20 anos 0,270 - 4,50 μ UI/mL e intervalo de detecção do *kit* 0,005 - 100 μ UI/mL.

4 - *Kit Roche Diagnostics* para dosagem do hormônio tiroxina livre T4, Referência: 06437281 190, valores em ng/dL, detecção do *Kit*: 0.3-100 pmol/L, factores de conversão das unidades: pmol/L x 0.077688 = ng/dL, ng/dL x 12.872 = pmol/L, valores teóricos normais: 0.93 - 1.7 ng/dL.

ANEXO 3

Sequenciamento plataforma NGS

Etapa 1: Quantificação do RNA Total pelo *Kit Qubit RNA Assay Kit*, catálogo números Q32852 e Q32855, mensurado no equipamento *Qubit 2.0 Fluorometer*, ambos do fabricante *Thermo Fisher Scientific*. Protocolo utilizado:

1 – Separado dois microtubos para as amostras padrões e um microtubo para cada amostra dos RNA a serem quantificados.

2 – Preparado a solução de trabalho diluindo o *Qubit RNA* reagente 1:200 em tampão *Qubit RNA*. Preparado 200mL de solução de trabalho para cada padrão e para cada amostra de RNA a ser quantificada.

3 – Preparado os microtubos transferindo os volumes de acordo com as seguintes recomendações:

Volume	Padrão	Amostras
Solução de trabalho	190µL	180 - 199µL
Amostra padrão	10µL	---
Amostra	----	1 - 20µL
Total em cada tubo	200µL	200µL

4 – Agitado no vórtex e deixado incubado por 2 minutos em temperatura ambiente, todas as amostras e os padrões.

5 – Efetuado a leitura das amostras padrões e as amostras de RNA no equipamento *Qubit Fluorometer 2.0*, de acordo com as normas de calibração.

Etapa 2: Iniciado o protocolo do transcriptoma com o isolamento, quantificação do RNA e transcriptase reversa RNA ;

1 – Efetuado a diluição das amostras de RNA para uma concentração final de 10ng em volume final de 5µL.

2 – Preparado o *maxter mix* para a quantidade de amostras adicionando os seguintes volumes conforme as recomendações para cada amostra:

Componente	Volume
5X VILO™ RT Reaction Mix	1,0µL
10X SuperScript® III Enzyme Mix	0,5
DNase-treated total RNA (10 ng)	≤3,5
Nuclease-Free water	Para 5µL

Total	5µL
--------------	-----

2 – Agitado no vórtex e efetuado um spin de 2 segundos.

3 – Colocado às amostras no termociclador obedecendo as seguintes orientações de programação:

Temperatura	Tempo
42°C	30 minutos
85°C	5 minutos
4°C	Período de espera

Etapa 3 – Amplificação dos alvos:

1 – Manipulado as amostras em recipiente com gelo para manter a temperatura refrigerada de 4°C

2 – Preparado o *mastermix* de acordo com a quantidade de amostras, seguindo as seguintes recomendações:

Componente	Volume por reação
<i>5X Ion AmpliSeq HiFi Mix (red cap)</i>	4µL
<i>Ion AmpliSeq Transcriptome Human Gene Expression Core Panel</i>	8µL
<i>Nuclease-Free Water</i>	3µL
Total	15µL

3 – Agitado o PCR *mastermix* no vórtex e efetuado um spin para retirar as gotículas do MCT.

4 – Transferido 15µL de *mastermix* em cada amostra da reação de transcrição.

5 - Agitado no vórtex e efetuado um *spin*.

6 – Colocado às amostras novamente no termociclador obedecendo as seguintes programações:

Estágio	Temperatura	Tempo
Espera	99°C	2 minutos

Ciclo: Programar de acordo com a tabela	99°C	15 segundos
	60°	16 minutos
Espera	10°C	Espera

Concentração de RNA da amostra de entrada	Valor	Número de ciclos
	0,1 – 1ng	16
	10ng	12
	100ng	10

Etapa 4 – Digerir parcialmente as sequencias iniciadoras, obedecendo as seguintes recomendações:

1 – Transferido 2mL do reagente *FuPa Reagent* para cada amostra amplificada.

2 – Agitado no vórtex e efetuado spin.

3 – Colocado às amostras novamente no termociclador e efetuado a seguinte programação:

Temperatura	Tempo
50°C	10 minutos
55°C	10 minutos
60°C	20 minutos
10°C	Espera

Etapa 5 – Ligando os adaptadores aos *amplicons* e purificando-os, de acordo com as seguintes recomendações:

1 - Preparado para cada código de barras escolhido uma mistura de *Ion P1 Adapter* e *Ion Xpress Barcode X* em uma diluição final de 1:4 para cada adaptador, conforme os volumes indicados:

Exemplo de <i>Mix</i> adaptador para até 4 reações
--

Componente	Volume
<i>Ion P1 Adapter</i>	2 μ L
<i>Ion Xpress Barcode X</i>	2 μ L
<i>Nuclease-Free water</i>	4 μ L
Total	8 μ L

Realizado a reação de ligação seguindo as seguintes recomendações:

2 – Transferido para cada amostra digerida os componentes adaptadores preparados conforme os volumes indicados:

Componente	Volume
<i>Switch Solution</i>	4 μ L
<i>Mix adaptador Barcode diluído</i>	2 μ L
Total: incluindo os 22 μ L do <i>amplicon</i> (amostra)	28 μ L

3 – Agitado em vórtex e realizado spin

4 – Adicionado 2mL de DNA ligase a cada MCT (30 μ L de volume total)

5 – Agitado em vórtex e realizado *spin*

6 – Colocado os MCT no termociclador e efetuado a seguinte programação:

Temperatura	Tempo
22°C	30 minutos
72°C	5 minutos
10°C	Espera

Etapa 6 – Purificando a biblioteca não amplificada, conforme as seguintes recomendações:

1 – Transferido 5 μ L do reagente *Agencourt AMPure XP* nos MCT e homogeneizados 5 vezes com a própria pipeta para misturar a suspensão de esferas com o DNA.

2 – Incubado por 5 minutos em temperatura ambiente

3 – Colocado os MCT em uma rack magnética com o *DynaMag* imã e incubado por 2 minutos . Após os 2 minutos retirado e desprezado o sobrenadante sem interferir com o sedimento.

4 – Adicionado 150µL de etanol 70%, movendo os MCT em um giro de 180° de sua posição inicial para lavar o *pellet*, após desprezado o sobrenadante sem interferir com o sedimento.

5 – Repetido novamente o passo anterior para uma segunda lavagem.

6 – Removido todo etanol com pipeta de 20µL, mantendo os MCT na *rack* com imã abertos por 2 minutos para secagem em temperatura ambiente.

Etapas 7 – Quantificar as bibliotecas por qPCR, conforme as etapas recomendadas:

1 – Retirado os MCT da rack imã e adicionado 50µL de *Low TE* ao *pellet* para dispensar as esferas. Agitado em vórtex e efetuado spin.

2 – Colocar novamente os MCT a *rack* magnética e aguardar por 2 minutos. Transferido o 45µL do sobrenadante para o poço de uma placa nova.

3 – Preparado cinco diluições em séries de 10 vezes do *E. coli* DH10B *Ion Control Library* a 6,8pM, 0,68pM, 0,068pM e 0,0068pM e 0,00068pM e marcados como padrão e usado a concentrações no *software* do equipamento qPCR.

4 – Diluído cada biblioteca *Ion AmpliSeq* usando as seguintes recomendações;

RNA de entrada	Concentração	Diluição recomendada
	10ng	1:10.000

5 – Preparado as misturas de reação em 3 poços para cada biblioteca padrão e cada amostra, seguindo as recomendações para calcular o volume necessário do *mastermix*;

Componente	Volume por reação	
	Placa de 96 poços	Placa de 384 poços
<i>2X TaqMan Master Mix</i>	10µL	5µL
<i>20X Ion TaqMan Assay</i>	1µL	0,5µL

Total	11 μ L	5,5 μ L
-------	------------	-------------

6 – Realizado na placa de 96 poços, transferido 11 μ L do *mastermix* em cada poço e adicionado 9 μ L da das bibliotecas padrão e as amostras.

7 – Programado o equipamento da *Applied Biosystems 7500 Fast Real-Time PCR System* nas seguintes configurações:

Estágio	Temperatura	Tempo
Espera	50°C	2 minutos
Espera	95°C	20 segundos
40 ciclos	95°C	1 segundo
	60°C	20 segundos

8 – Seguindo o qPCR, calculado a concentração média da biblioteca AmpliSeq transcriptoma não diluída pela multiplicação determinada por qPCR da biblioteca diluída utilizada no ensaio.

9 – As amostras que tiveram concentração maior de 100pM, efetuado diluição com *Nuclease-Free water* para o valor de 100pM em volume de 8 μ L.

Etapa 8 – Realizado PCR emulsão no equipamento *Ion PI HI-Q OT2 200 Kit*, sob o catálogo: A26434, seguindo as instruções:

1 – Acoplado os reagentes necessários no equipamento; *Ion OneTouch Reaction Oil* (25mL) e *Nuclease-free Water*.

2 – Preparado os reagentes de trabalho; *Ion PI Master Mix*, *Ion PI Enzyme Mix*, *Ion PI Ion Sphere Particles*, agitados em vórtex e efetuado spin.

3 – Diluído as amostras conforme orientação de 100pM iniciais em volume de 8 μ L adicionando 92 μ L de *Nuclease free Water*, para volume total de 100 μ L.

4 – Agitado por vórtex e efetuado spin.

5 – Num MCT de 2mL, foi adicionado os seguintes reagentes na seguinte ordem:

Ordem	Reagente	Volume
1	<i>Nuclease-free Water</i>	80 μ L

2	<i>Ion PI Enzyme Mix</i>	120µL
3	<i>Ion PI ISPs</i>	100µL
4	Biblioteca diluída	100µL
	Volume total incluindo <i>mastermix</i>	2400µL

6 – Agitado em vórtex e centrifugado por 2 segundos e acondicionado o tubo na sua posição no equipamento.

7 – Alimentado o equipamento com o *Ion OneTouch Reaction Filter*.

8 – Inicializado a configuração do *Ion OneTouch 2 Instrument*, obedecendo as seguintes configurações na tela *touch*; seleciona *run*, selecionado a opção *Proton: Ion PI Hi-Q OT2 200 Kit*, selecionado *next*, selecionado *assisted*, selecionado *yes*, e esperado o termino do processo, em média 5 horas.

Etapa 9 – Realizado o enriquecimento das bibliotecas com partículas de esferas (*beads*) no equipamento *Ion OneTouch ES*.

Etapa 10 – Lavado as bibliotecas e as partículas de esferas (*beads*) e quantificado a amostra no *Qubit Fluorometer*, seguindo as recomendações de valores: 10 a 30% de DNA ligado as esferas é o valor ideal.

Etapa 11 – Calibrado automaticamente do Equipamento *Ion Proton System*.

Etapa 12 - Preparado os chips que receberam as bibliotecas com as *beads*. A solução produto de todos os protocolos anteriores foi injetada no *Ion PI Chip Kit V3* de forma que o liquido se espalhasse por toda matriz do componente e por consequência nos poços microscópicos do chip. Utilizamos 1 *chip* para sequenciar 4 amostras do grupo controle e outro chip para sequenciar 4 amostras do grupo hipotiroidianos clínicos sem de tratamento com L-T4.

1 – Centrifugado o *Ion Chip Kit V3* por 5 minutos em centrífuga especial para *Ion PI Chip*, eliminando toda bolha de ar que existia na matriz.

2 – Acondicionado o *Ion PI Chip* no soquete do equipamento *Ion Proton System*, efetuado o cadastro da corrida e o protocolo utilizado RNA-seq.

ANEXO 4

Pipeline das análises CTL x HT.

```
#####  
# Libraries  
require(edgeR)  
require(org.Hs.eg.db)  
require(e1071)  
#Import and filtering  
data=read.table("CTL x HT.txt",header=T)  
data=dplyr::filter(data,CTL1>0,CTL2>0,CTL3>0,CTL4>0,HT1>0,HT2>0,HT3>0,HT4>0)  
data=data.frame(data,row.names=1)  
print(head(data))  
#Set contrast  
group <- as.factor(rep(c("A","B"),each=4))  
groups <- data.frame(samples=rep(c("A","B"), each=4),  
row.names=c("A1","A2","A3","A4","B1","B2","B3","B4"))  
design <- model.matrix(~0+group)  
colnames(design) <- levels(group)  
row.names(design) <- row.names(groups)  
contrast <- makeContrasts(B_A=B-A, levels=design)  
#Normalization  
norm <- calcNormFactors(data)  
DGL1=DGEList(data, norm.factors=norm,group=group)  
DGL2=DGEList(data, norm.factors=norm,group=group)  
DGL3=DGEList(data, norm.factors=norm,group=group)  
DGL4=DGEList(data, norm.factors=norm,group=group)  
DGL5=DGEList(data, norm.factors=norm,group=group)  
#MODEL ONE (Common + exact)  
#Dispersions  
DGL1=estimateCommonDisp(DGL1)  
#Hypothesis Test and Plot output  
top1 <- exactTest(DGL1)
```

```

top1 <- topTags(top1, n=NULL)
top1 <- top1@.Data[[1]]
top1 <- subset(top1, FDR<0.05)
png("MODEL1.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL1, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variabilidade Biológica", ylab="Coeficiente Biológico de
Variabilidade", xlab="Módulo do logCPM")
plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão 1
do Fold Change",ylab="Dimensão 2 do Fold Change",cex=1.5)
plotMeanVar(DGL, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variabilidade da Média e Modelos de Dispersão", xlab="Nível Médio da
Expressão (Escala log10)",ylab="Variabilidade e Nível Genético (Escala log10)")
plotSmear(DGL1, de.tags=row.names(top1), col="grey70",cex=1, lowess=TRUE,
main="Detecção do DGE e Concentração de Fold Change", xlab="Módulo do
logCPM")
dev.off()

#MODEL TWO (Trend + Exact)
#Dispersions
DGL2 <- estimateTrendedDisp(DGL2)
#Hypothesis Test and Plot output
top2 <- exactTest(DGL2)
top2 <- topTags(top2, n=NULL)
top2 <- top2@.Data[[1]]
top2 <- subset(top2, FDR<0.05)
png("MODEL2.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,3))
plotBCV(DGL2, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variabilidade Biológica", ylab="Coeficiente Biológico de
Variabilidade", xlab="Módulo do logCPM")

```

```

plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão 1
do Fold Change", ylab="Dimensão 2 do Fold Change", cex=1.5)
plotMeanVar(DGL2, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variação da Média e Modelos de Dispersão", xlab="Nível Médio da
Expressão (Escala log10)", ylab="Variação e Nível Genético (Escala log10)")
plotSmear(DGL2, de.tags=row.names(top2), col="grey70", cex=1, lowess=TRUE,
main="Detecção do DGE e Concentração de Fold Change", xlab="Média do
logCPM")
hist(DGL2$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,40))
lines(density(DGL2$trended.dispersion), col="red", lwd=1.2)
dev.off()

```

```

#MODEL THREE (Trend + LRT)
#Dispersions
DGL3 <- estimateTrendedDisp(DGL3)
#Hypothesis Test and Plot output
fit <- glmFit(DGL3, design=design)
fit <- glmLRT(fit, contrast=contrast)
top <- topTags(fit, n=NULL)
top3 <- top@.Data[[1]]
top3 <- subset(top3, FDR<0.05)
png("MODEL3.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,3))
plotBCV(DGL3, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variação Biológica", ylab="Coeficiente Biológico de
Variação", xlab="Média do logCPM")
plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão 1
do Fold Change", ylab="Dimensão 2 do Fold Change", cex=1.5)

```

```

plotMeanVar(DGL3, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Varia  o da M dia e Modelos de Dispers o", xlab="N vel M dio da
Express o (Escala log10)",ylab="Varia  o   N vel G nico (Escala log10)")
plotSmear(DGL3, de.tags=row.names(top3), col="grey70",cex=1, lowess=TRUE,
main="Detec o do DGE e Concentra o de Fold Change", xlab="M dia do
logCPM")
hist(DGL3$trended.dispersion, freq=FALSE, main="Densidade da Tend ncia de
Dispers o", xlab="Valores de Dispers o", ylab="Densidade", ylim=c(0,40))
lines(density(DGL3$trended.dispersion), col="red", lwd=1.2)
dev.off()

#MODEL FOUR (Trend + Common + LRT)
#Dispersions
DGL4 <- estimateDisp(DGL4, design=design, tagwise=FALSE)
fit <- glmFit(DGL4, design=design)
fit <- glmLRT(fit, contrast=contrast)
top <- topTags(fit, n=NULL)
DGE <- top@.Data[[1]]
DGE <- subset(DGE, FDR<0.05)
png("MODEL4.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,3))
plotBCV(DGL4, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Varia  o Biol gica", ylab="Coeficiente Biol gico de
Varia  o", xlab="M dia do logCPM")
plotMDS(data, main="Redu o da Escala de Multidimens es", xlab="Dimens o 1
do Fold Change",ylab="Dimens o 2 do Fold Change", cex=1.5)
plotMeanVar(DGL4, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Varia  o da M dia e Modelos de Dispers o", xlab="N vel M dio da
Express o (Escala log10)",ylab="Varia  o   N vel G nico (Escala log10)")
plotSmear(DGL4, de.tags=row.names(DGE), col="grey70",cex=1, lowess=TRUE,
main="Detec o do DGE e Concentra o de Fold Change", xlab="M dia do
logCPM")

```

```

hist(DGL4$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,40))
lines(density(DGL4$trended.dispersion), col="red", lwd=1.2)
dev.off()

```

Pipeline das análises estudo estado da tiroide.

```

#####
# Libraries
require(edgeR)
require(org.Hs.eg.db)
require(e1071)
#Import and filtering
data=read.csv("onoff.csv", header=T)
data=dplyr::filter(data,OFF1>0,OFF2>0,OFF3>0,OFF4>0,OFF5>0,OFF6>0,OFF7>0,OFF
F8>0,ON1>0,ON2>0,ON3>0,ON4>0,ON5>0,ON6>0,ON7>0,ON8>0)
data=data.frame(data,row.names=1)
print(head(data))
#Set contrast
group <- as.factor(rep(c("OFF","ON"),each=8))
groups <- data.frame(samples=rep(c("OFF","ON"), each=8),
row.names=c("OFF1","OFF2","OFF3","OFF4","OFF5","OFF6","OFF7","OFF8","ON1","O
N2","ON3","ON4","ON5","ON6","ON7","ON8"))
design <- model.matrix(~0+group)
colnames(design) <- levels(group)
row.names(design) <- row.names(groups)
contrast <- makeContrasts(ON_OFF=ON-OFF, levels=design)
print(contrast)
#Normalization
norm <- calcNormFactors(data)
DGL1=DGEList(data, norm.factors=norm,group=group)

```

```

DGL2=DGEList(data, norm.factors=norm,group=group)
DGL3=DGEList(data, norm.factors=norm,group=group)
DGL4=DGEList(data, norm.factors=norm,group=group)
DGL5=DGEList(data, norm.factors=norm,group=group)
#MODEL ONE (Common + exact)
#Dispersions
DGL1=estimateCommonDisp(DGL1)
#Hypothesis Test and Plot output
top1 <- exactTest(DGL1)
summary(decideTestsDGE(top1))
top1 <- topTags(top1, n=NULL)
top1 <- top1@.Data[[1]]
top1 <- subset(top1, FDR<0.05)
png("MODEL1.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL1, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variaco Biolgica", ylab="Coeficiente Biolgico de
Variaco", xlab="Mdia do logCPM")
plotMDS(data, main="Reduo da Escala de Multidimenses", xlab="Dimenso 1
do Fold Change", ylab="Dimenso 2 do Fold Change", cex=1.5)
plotMeanVar(DGL1, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variaco da Mdia e Modelos de Disperso", xlab="N-vel Mdio da
Expresso (Escala log10)", ylab="Variaco  N-vel Gnico (Escala log10)")
plotSmear(DGL1, de.tags=row.names(top2), col="grey70",cex=1, lowess=TRUE,
main="Deteco do DGE e Concentraco de Fold Change", xlab="Mdia do
logCPM")
dev.off()

#MODEL TWO (Trend + Exact)
#Dispersions
DGL2 <- estimateTrendedDisp(DGL2)
#Hypothesis Test and Plot output

```

```

top2 <- exactTest(DGL2)
summary(decideTestsDGE(top2))
top2 <- topTags(top2, n=NULL)
top2 <- top2@.Data[[1]]
top2 <- subset(top2, FDR<0.05)
png("MODEL2.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,3))
plotBCV(DGL2, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente VariaÃ§Ã£o BiolÃ³gica", ylab="Coeficiente BiolÃ³gico de
VariaÃ§Ã£o", xlab="MÃ©dia do logCPM")
plotMDS(data, main="ReduÃ§Ã£o da Escala de MultidimensÃµes", xlab="DimensÃ£o 1
do Fold Change", ylab="DimensÃ£o 2 do Fold Change", cex=1.5)
plotMeanVar(DGL2, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="VariaÃ§Ã£o da MÃ©dia e Modelos de DispersÃ£o", xlab="NÃvel MÃ©dio da
ExpressÃ£o (Escala log10)", ylab="VariaÃ§Ã£o Ã NÃvel GÃ©nico (Escala log10)")
plotSmear(DGL2, de.tags=row.names(top2), col="grey70", cex=1, lowess=TRUE,
main="DetecÃ§Ã£o do DGE e ConcentraÃ§Ã£o de Fold Change", xlab="MÃ©dia do
logCPM")
hist(DGL2$trended.dispersion, freq=FALSE, main="Densidade da TendÃªncia de
DispersÃ£o", xlab="Valores de DispersÃ£o", ylab="Densidade", ylim=c(0,40))
lines(density(DGL2$trended.dispersion), col="red", lwd=1.2)
dev.off()

#MODEL THREE (Trend + LRT)
#Dispersions
DGL3 <- estimateTrendedDisp(DGL3)
#Hypothesis Test and Plot output
fit <- glmFit(DGL3, design=design)
fit <- glmLRT(fit, contrast=contrast)
summary(decideTestsDGE(fit))
top <- topTags(fit, n=NULL)

```

```

top3 <- top@.Data[[1]]
top3 <- subset(top3, FDR<0.05)
png("MODEL3.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,3))
plotBCV(DGL3, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variabilidade Biológica", ylab="Coeficiente Biológico de
Variabilidade", xlab="Média do logCPM")
plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão 1
do Fold Change", ylab="Dimensão 2 do Fold Change", cex=1.5)
plotMeanVar(DGL3, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variabilidade da Média e Modelos de Dispersão", xlab="N-vel Média da
Expressão (Escala log10)", ylab="Variabilidade N-vel Genético (Escala log10)")
plotSmear(DGL3, de.tags=row.names(top3), col="grey70", cex=1, lowess=TRUE,
main="Detecção do DGE e Concentração de Fold Change", xlab="Média do
logCPM")
hist(DGL3$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,40))
lines(density(DGL3$trended.dispersion), col="red", lwd=1.2)
dev.off()

#MODEL FOUR (Common + LRT)
#Dispersions
DGL4 <- estimateCommonDisp(DGL4)
fit <- glmFit(DGL4, design=design)
fit <- glmLRT(fit, contrast=contrast)
summary(decideTestsDGE(fit))
top <- topTags(fit, n=NULL)
DGE <- top@.Data[[1]]
DGE <- subset(DGE, FDR<0.05)
png("MODEL4.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))

```

```

plotBCV(DGL4, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente VariaÃ§Ã£o BiolÃ³gica", ylab="Coeficiente BiolÃ³gico de
VariaÃ§Ã£o", xlab="MÃ©dia do logCPM")
plotMDS(data, main="ReduÃ§Ã£o da Escala de MultidimensÃµes", xlab="DimensÃ£o 1
do Fold Change", ylab="DimensÃ£o 2 do Fold Change", cex=1.5)
plotMeanVar(DGL4, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="VariaÃ§Ã£o da MÃ©dia e Modelos de DispersÃ£o", xlab="NÃ-vel MÃ©dio da
ExpressÃ£o (Escala log10)", ylab="VariaÃ§Ã£o Ã NÃ-vel GÃnico (Escala log10)")
plotSmear(DGL4, de.tags=row.names(DGE), col="grey70", cex=1, lowess=TRUE,
main="DetecÃ§Ã£o do DGE e ConcentraÃ§Ã£o de Fold Change", xlab="MÃ©dia do
logCPM")
hist(DGL4$trended.dispersion, freq=FALSE, main="Densidade da TendÃncia de
DispersÃ£o", xlab="Valores de DispersÃ£o", ylab="Densidade", ylim=c(0,40))
lines(density(DGL4$trended.dispersion), col="red", lwd=1.2)
#hist(DGL4$trended.dispersion, freq=FALSE, main="Density of Individual Dispersion",
xlab="Dispersion Value", ylim=c(0,40))
#lines(density(DGL4$trended.dispersion), col="red", lwd=1.2)
dev.off()

```

Pipeline para o estudo transtorno bipolar.

```

#libraries
require(edgeR)
require(dplyr)
require(Hmisc)

#import
data=read.table("MATRIX 240x240.txt",header=T,row.names=1)
row_sub=apply(data,1,function(row)all(row!=0))
data=data[row_sub,]

```

```

#norm and groups
norm=calcNormFactors(data)
group=c(rep("1",240),rep("2",240))
group=as.factor(group)
#design <- model.matrix(~group)
#contrast <- makeContrasts(1_2=2-1, levels=design)

#set DGLs
DGL1=DGEList(data, norm.factors=norm,group=group)
DGL2=DGEList(data, norm.factors=norm,group=group)
DGL3=DGEList(data, norm.factors=norm,group=group)
DGL4=DGEList(data, norm.factors=norm,group=group)

#MODEL ONE (Common + exact)
#Dispersions
DGL1=estimateCommonDisp(DGL1)
#Hypothesis Test and Plot output
top1 <- exactTest(DGL1)
top1 <- topTags(top1, n=NULL)
top1 <- top1@.Data[[1]]
top1 <- subset(top1, FDR<0.05)
png("MODEL1.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL1, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variabilidade Biológica", ylab="Coeficiente Biológico de
Variabilidade", xlab="Módulo do logCPM")
#plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão
1 do Fold Change",ylab="Dimensão 2 do Fold Change",cex=1.5)

```

```

plotMeanVar(DGL, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Varia  o da M dia e Modelos de Dispers o", xlab="N vel M dio da
Express o (Escala log10)",ylab="Varia o   N vel G nico (Escala log10)")
plotSmear(DGL1, de.tags=row.names(top1), col="grey70",cex=1, lowess=TRUE,
main="Detec o do DGE e Concentra o de Fold Change", xlab="M dia do
logCPM")
dev.off()

```

```

#MODEL TWO (Trend + Exact)
#Dispersions
DGL2 <- estimateTrendedDisp(DGL2)
#Hypothesis Test and Plot output
top2 <- exactTest(DGL2)
top2 <- topTags(top2, n=NULL)
top2 <- top2@.Data[[1]]
top2 <- subset(top2, FDR<0.05)
png("MODEL2.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL2, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Varia o Biol gica", ylab="Coeficiente Biol gico de
Varia o", xlab="M dia do logCPM")
#plotMDS(data, main="Redu o da Escala de Multidimens es", xlab="Dimens o
1 do Fold Change",ylab="Dimens o 2 do Fold Change", cex=1.5)
plotMeanVar(DGL2, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Varia o da M dia e Modelos de Dispers o", xlab="N vel M dio da
Express o (Escala log10)",ylab="Varia o   N vel G nico (Escala log10)")
plotSmear(DGL2, de.tags=row.names(top2), col="grey70",cex=1, lowess=TRUE,
main="Detec o do DGE e Concentra o de Fold Change", xlab="M dia do
logCPM")

```

```

hist(DGL2$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,50))
lines(density(DGL2$trended.dispersion), col="red", lwd=1.2)
dev.off()

```

```

#MODEL THREE (Trend + LRT)
#Dispersions
DGL3 <- estimateTrendedDisp(DGL3)
#Hypothesis Test and Plot output
fit <- glmFit(DGL3)
fit <- glmLRT(fit)
top <- topTags(fit, n=NULL)
top3 <- top@.Data[[1]]
top3 <- subset(top3, FDR<0.05)
png("MODEL3.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL3, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variável Biológica", ylab="Coeficiente Biológico de
Variável", xlab="Média do logCPM")
#plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão
1 do Fold Change", ylab="Dimensão 2 do Fold Change", cex=1.5)
plotMeanVar(DGL3, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variável da Média e Modelos de Dispersão", xlab="Nível Médio da
Expressão (Escala log10)", ylab="Variável Média Nível Genético (Escala log10)")
plotSmear(DGL3, de.tags=row.names(top3), col="grey70", cex=1, lowess=TRUE,
main="Detecção do DGE e Concentração de Fold Change", xlab="Média do
logCPM")
hist(DGL3$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,50))
lines(density(DGL3$trended.dispersion), col="red", lwd=1.2)
dev.off()

```

```

#MODEL FOUR (Common + LRT)
#Dispersions
DGL4 <- estimateCommonDisp(DGL4)
fit <- glmFit(DGL4)
fit <- glmLRT(fit)
top <- topTags(fit, n=NULL)
DGE <- top@.Data[[1]]
DGE <- subset(DGE, FDR<0.05)
png("MODEL4.png", width=13, height=13, units="in", res=300)
par(mfrow=c(2,2))
plotBCV(DGL4, col.tagwise="cadet4blue4", cex=0.8, pch=19, col.trend="black",
lwd=1.8, main="Coeficiente Variacional Biológica", ylab="Coeficiente Biológico de
Variacional", xlab="Média do logCPM")
#plotMDS(data, main="Redução da Escala de Multidimensões", xlab="Dimensão
1 do Fold Change", ylab="Dimensão 2 do Fold Change", cex=1.5)
plotMeanVar(DGL4, show.raw.vars=TRUE, show.binned.common.disp.vars=TRUE,
main="Variacional da Média e Modelos de Dispersão", xlab="Nível Médio da
Expressão (Escala log10)", ylab="Variacional vs Nível Genético (Escala log10)")
plotSmear(DGL4, de.tags=row.names(DGE), col="grey70", cex=1, lowess=TRUE,
main="Detecção do DGE e Concentração de Fold Change", xlab="Média do
logCPM")
hist(DGL4$trended.dispersion, freq=FALSE, main="Densidade da Tendência de
Dispersão", xlab="Valores de Dispersão", ylab="Densidade", ylim=c(0,50))
lines(density(DGL4$trended.dispersion), col="red", lwd=1.2)
dev.off()

```