



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE BIOCIENCIAS  
DEPARTAMENTO DE FISIOLOGIA E FARMACOLOGIA  
CIÊNCIAS BIOLÓGICAS – BACHARELADO

PEDRO HENRIQUE SANTOS DE JESUS

**Determinação dos efeitos do clorotalonil em células do córtex pré-frontal de camundongo (*Mus musculus*) a partir de análise transcriptômica**

Recife  
2025

PEDRO HENRIQUE SANTOS DE JESUS

**Determinação dos efeitos do clorotalonil em células do córtex pré-frontal de camundongo (*Mus musculus*) a partir de análise transcriptômica**

Trabalho de Conclusão de Curso apresentado ao Curso de Ciências Biológicas – Bacharelado da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Ciências Biológicas.

Orientadora: Mariana Leivas Müller Hoff

Recife

2025

**Ficha de identificação da obra elaborada pelo autor,  
através do programa de geração automática do SIB/UFPE**

Jesus, Pedro Henrique Santos de.

Determinação dos efeitos do clorotalonil em células do córtex pré-frontal de camundongo (*Mus musculus*) a partir de análise transcriptômica / Pedro Henrique Santos de Jesus. - Recife, 2025.

52 p. : il., tab.

Orientador(a): Mariana Leivas Müller Hoff

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Biociências, Ciências Biológicas - Bacharelado, 2025.  
Inclui referências.

1. pesticidas. 2. RNA-Seq. 3. agrotóxicos. 4. bioinformática. 5. genes. I. Hoff, Mariana Leivas Müller. (Orientação). II. Título.

570 CDD (22.ed.)

PEDRO HENRIQUE SANTOS DE JESUS

**Determinação dos efeitos do clorotalonil em células do córtex pré-frontal de camundongo (*Mus musculus*) a partir de análise transcriptômica**

TCC apresentado ao Curso de Ciências Biológicas – Bacharelado da Universidade Federal de Pernambuco, campus Recife, como requisito para a obtenção do título de Bacharel em Ciências Biológicas.

Aprovado em: 04/12/2025.

**BANCA EXAMINADORA**

---

Profa. Dra. Mariana Leivas Müller Hoff  
Universidade Federal de Pernambuco

---

Profa. Dra. Belmira Lara da Silveira Andrade da Costa  
Universidade Federal de Pernambuco

---

Prof. Dr. Paulo Sergio Martins de Carvalho  
Universidade Federal de Pernambuco

Dedico este trabalho a Maria Edlene de Jesus,  
à vovó Gerusa Helena dos Santos e a Eva  
Luiza Souza de Jesus (Evinha).

“Assim como nós estamos hoje vivendo o desastre do nosso tempo, ao qual algumas seletas pessoas chamam Antropoceno. A grande maioria está chamando de caos social, desgoverno geral, perda de qualidade no cotidiano, nas relações e estamos todos jogados nesse abismo.” (Krenank, 2020, p. 72).

## RESUMO

O Brasil é o maior consumidor de agrotóxicos do mundo desde 2008 e em 2021 consumiu cerca de 719 mil toneladas. O clorotalonil (CTL) é um pesticida de amplo espectro e não sistêmico, e foi o quinto pesticida mais vendido do Brasil em 2023. O CTL está enquadrado na classificação 2 de toxicidade: altamente tóxico. Neste trabalho buscamos investigar os efeitos do CTL em células do córtex pré-frontal do camundongo através de dados públicos de sequenciamento de RNA. Duas amostras de CTL, foram selecionadas do bioprojeto com o código de acesso PRJNA288135, no banco de dados SRA (*Sequence Read Archive*) do National Center for Biotechnology information (NCBI). O biorjeto é proveniente de um artigo no qual os autores investigam alterações transpcionais semelhantes ao autismo e à neurodegeneração de 294 pesticidas em culturas celulares do córtex pré-frontal de camundongo. As culturas foram tratadas com a dose não-citotóxica de 1 µM de CTL (grupo exposto ao CTL), ou com dimetilsulfóxido na concentração final  $\leq 0,5\%$  (grupo veículo, condição controle). As amostras selecionadas para os grupos controle e CTL foram baixadas diretamente no ambiente Galaxy e seguimos dois fluxos de trabalho: (a) o fluxo 1, no qual as amostras tiveram as 10 bases iniciais trimadas no pré-processamento; e o fluxo 2 (b), no qual essas bases foram mantidas. Realizamos análises de qualidade, pré-processamento, mapeamento, análise de similaridade, análise de expressão diferencial e análise das vias metabólicas. Três genes foram estatisticamente inibidos (*False Discovery Rate*  $< 0,05$ ) – subunidade beta-1 da proteína de ligação a nucleotídeos de guanina (Gnb1), placofilina 4 (Pkp4) e proteína de domínio triplo funcional (PTPRF Trio). O fluxo 2 foi robusto o suficiente para agrupar os transcriptomas das amostras de CTL e identificar o efeito do CTL na repressão desses genes. Nosso trabalho aponta que o CTL afeta genes envolvidos no citoesqueleto e na sinalização celular e os seus efeitos na expressão destes genes pode estar associada à carcinogenicidade deste pesticida.

Palavras-chave: pesticidas; RNA-Seq; agrotóxicos; bioinformática; genes.

## ABSTRACT

Brazil has been the world's largest consumer of pesticides since 2008, and in 2021 it used approximately 719 thousand tons. Chlorothalonil (CTL) is a broad-spectrum, non-systemic pesticide and was the fifth best-selling pesticide in Brazil in 2023. CTL is classified as toxicity class 2: highly toxic. In this study, we sought to investigate the effects of CTL on prefrontal cortex cells of the mouse using public RNA sequencing data. Two CTL samples from the bioproject with accession code PRJNA288135 were selected from the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI). The bioproject originates from a study in which the authors investigated transcriptional alterations resembling autism and neurodegeneration induced by 294 pesticides in cultured mouse prefrontal cortex cells. The cultures were treated with a non-cytotoxic dose of 1  $\mu$ M CTL (CTL-exposed group) or with dimethyl sulfoxide at a final concentration  $\leq$  0.5% (vehicle group, control condition). The selected samples for the control and CTL groups were downloaded directly into the Galaxy environment, and we followed two workflow strategies: (a) workflow 1, in which the first 10 bases of the reads were trimmed during preprocessing; and (b) workflow 2, in which these bases were retained. We performed quality control, preprocessing, read mapping, similarity analysis, differential expression analysis, and metabolic pathway analysis. Three genes were statistically downregulated (False Discovery Rate  $<$  0.05): guanine nucleotide-binding protein subunit beta-1 (Gnb1), plakophilin-4 (Pkp4), and triple functional domain protein (PTPRF Trio). Workflow 2 was sufficiently robust to cluster the CTL transcriptomes and detect the CTL-induced repression of these genes. Our findings indicate that CTL affects genes involved in cytoskeletal organization and cell signaling, and its effects on the expression of these genes may be associated with the carcinogenic potential of this pesticide.

Keywords: pesticides; RNA-Seq; agrochemicals; bioinformatics; genes.

## SUMÁRIO

<b>RESUMO.....</b>	<b>7</b>
ABSTRACT.....	8
<b>SUMÁRIO.....</b>	<b>9</b>
<b>1. INTRODUÇÃO.....</b>	<b>17</b>
1.1. Pesticidas no mundo e no Brasil.....	17
1.2. Problemas do uso dos pesticidas.....	17
1.2.1. Saúde humana.....	17
1.2.2. Ambiente.....	17
1.3. Clorotalonil.....	18
1.3.1. Descrição do composto.....	18
1.3.2. Mecanismo de ação.....	18
1.3.3. Uso no Brasil.....	19
1.3.4. Efeitos do CTL.....	19
1.4. Modelo animal: camundongo e células do córtex pré-frontal.....	20
1.5. Sequenciamento de nova geração (NGS) e bioinformática.....	20
<b>2. OBJETIVOS.....</b>	<b>21</b>
2.1. Objetivo Geral.....	21
2.2. Objetivos Específicos.....	21
<b>3. METODOLOGIA.....</b>	<b>22</b>
3.1. Obtenção e caracterização das amostras de RNA-Seq.....	22
3.2. Análise de qualidade.....	23
3.3. Pré-processamento.....	24
3.4. Mapeamento e quantificação de transcritos.....	25
3.5. Similaridade de expressão gênica entre os transcriptomas.....	26
3.6. Anotação dos transcritos em genes.....	30
3.7. Preparação dos dados para as análises de expressão gênica.....	32
3.8. Análise de expressão gênica diferencial (DEG).....	33
3.9. Análise da expressão de genes envolvidos no mecanismo de ação do CTL.....	34
<b>4. RESULTADOS.....</b>	<b>37</b>
4.1. Análise de qualidade dos dados brutos.....	37
4.2. Análise de qualidade dos dados trimados.....	37
4.3. Similaridade da expressão gênica entre os transcriptomas.....	40
4.3.1. Análise de genes diferencialmente expressos.....	41
4.4. Análise dos genes mais diferencialmente expressos no CTL em relação ao controle..	42
4.5. Análise de genes envolvidos no mecanismo de ação do CTL.....	43
<b>5. DISCUSSÃO.....</b>	<b>50</b>
5.1. Efeito do pré-processamento.....	50
5.1.1. Comparação entre o fluxo 1 e 2.....	50
5.1.2. Análise de genes diferencialmente expressos (DEG).....	51

5.1.3. Análise da expressão de genes mais e menos expressos no grupo clorotalonil em relação ao controle.....	52
5.1.4. Análise da expressão de genes de interesse.....	53
<b>6. CONCLUSÃO.....</b>	<b>55</b>
<b>REFERÊNCIAS.....</b>	<b>56</b>

## 1. INTRODUÇÃO

### 1.1. Pesticidas no mundo e no Brasil

Pesticidas são formulações químicas com compostos ativos que servem, primariamente, para eliminar plantas, insetos, fungos e outros agentes indesejados na agricultura. A quantidade de pesticidas utilizados na agricultura mundialmente, em 2022, foi de 3,7 milhões de toneladas (Mt), 4% maior que em 2021, 13% a mais que na última década e o dobro desde 1990 (FAO, 2024, p. 3). Segundo a Organização das Nações Unidas para Alimentação e Agricultura (FAO) (2024), as Américas são as maiores consumidoras de pesticidas dentre todos os continentes desde meados da década de 90. No Brasil, o uso de pesticidas mais que dobrou nos últimos 15 anos, atingindo a marca de 719 mil toneladas em 2021 e é considerado, desde 2008, o maior consumidor de agrotóxicos do mundo (Bombardi, 2023). Enquanto cultivos de alimentos-chave da dieta brasileira declinaram nas últimas décadas, a soja e a área destinada à criação de gado cresceu colossalmente.

### 1.2. Problemas do uso dos pesticidas

#### 1.2.1. Saúde humana

A exposição aos pesticidas desencadeia uma série de problemas para a saúde humana, de acordo com a classificação de toxicidade, o tempo de exposição e a quantidade absorvida pelo organismo (INCA, 2022). Segundo o boletim epidemiológico da Secretaria de Vigilância em Saúde e Ambiente (2023), entre janeiro de 2013 e junho de 2022, 124.294 casos de intoxicação exógena por agrotóxicos foram notificados em todo o Brasil. A exposição a pesticidas aumentou a incidência de cânceres, depressão, abortos e óbitos fetais, problemas respiratórios e problemas renais (Corrêa *et al.*, 2020).

#### 1.2.2. Ambiente

O uso de pesticidas costuma ser específico para os seus alvos de acordo com seu mecanismo de ação. Entretanto, organismos não-alvo do ecossistema também sofrem com a exposição a esses pesticidas, que podem estar ligados a perda global de biodiversidade (Wan *et al.*, 2025). Rachel Carlson, em sua obra pioneira “Primavera silenciosa”, alertou sobre o uso indiscriminado de pesticidas, sobretudo o

DDT (diclorodifeniltricloroetano), e suas terríveis consequências para o meio natural, como a contaminação do ar, do solo e da água (Carlson, 2010).

### 1.3. Clorotalonil

#### 1.3.1. Descrição do composto

O clorotalonil (CTL) é um pesticida de amplo espectro e não sistêmico, utilizado principalmente como fungicida para controlar doenças foliares de hortaliças, campos e plantas ornamentais. É, também, utilizado como protetor de madeira, agente antimofe e em tintas anti-incrustantes. Apesar de seu vasto uso, seu mecanismo de ação não é completamente conhecido (USEPA, 1999). Seu nome químico é 2,4,5,6-Tetracloroisofalonitrila, sua fórmula molecular é  $C_8Cl_4N_2$ , sua massa molar é 265,9 g/mol, seu ponto de fusão é 253 °C e sua fórmula estrutural é a apresentada na figura 1.

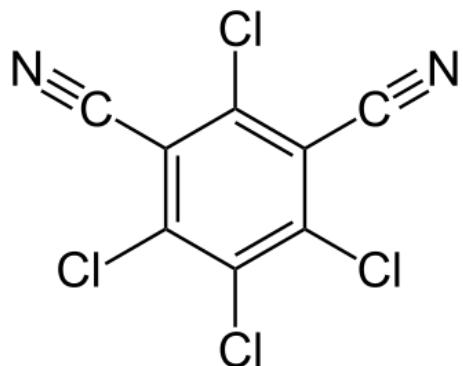


Figura 1 – Fórmula estrutural do clorotalonil

#### 1.3.2. Mecanismo de ação

O clorotalonil (CTL) é um fungicida clorado cuja ação tóxica principal decorre da reatividade eletrofílica do anel aromático com nucleófilos tiol ou grupos sulfidrila ( $-SH$ ). Essa reação forma conjugados com pequenas moléculas, como a glutationa, e com grupos tiol de proteínas, levando à inibição de enzimas dependentes de tiol, depleção de glutationa, estresse oxidativo e disfunção mitocondrial — efeitos que explicam sua toxicidade em fungos e também em animais (Tillman; Siegel; Long, 1973; Arvanites; Boerth, 2001). O CTL, durante a sua biotransformação, reage com a glutationa, formando um aduto de glutationa com eliminação de ácido clorídrico.

Seu mecanismo de ação é semelhante ao dos fungicidas sulfenil triclorometil (Tillman; Siegel; Long, 1973). O efeito antifúngico do CTL é atribuído à inibição da glicólise e da respiração celular pela reação com grupos tiol funcionais de enzimas dessas vias (Tillman *et al.*, 1973), como a gliceraldeído-3-fosfato desidrogenase (Long; Siegel, 1975). O CTL é capaz de reagir com tiois de baixo peso molecular (Vincent; Sisler, 1968).

Como um composto aromático eletrofílico, o clorotalonil [...] entra em reações de substituição com componentes nucleofílicos de proteínas, por exemplo, com resíduos tiol livres (Enoch *et al.*, 2009; Natsch *et al.*, 2011; Roberts *et al.*, 2015) como resultados dos quais estruturas exógenas imunologicamente processáveis são formadas. (Hartwig; Mak Comission, 2025, tradução nossa).

### 1.3.3. Uso no Brasil

O Brasil é um dos principais destinos de agrotóxicos proibidos na União Europeia (Bombardi, 2023). Dentre os 10 pesticidas mais vendidos aqui, 5 estão banidos na Europa. O clorotalonil foi o 5º pesticida mais vendido no Brasil em 2023, com 45.533,10 toneladas de ingrediente ativo vendido, segundo o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (Ibama, 2023). O seu uso é direcionado para diversas culturas, tais como, amendoim, banana, batata, batata-doce, berinjela, beterraba, caqui, figo, caju, carombola, figo, goiaba, kiwi, mangaba, cará, cenoura, feijão, ervilha, grão-de-bico, lentinha, gengibre, inhame, jiló, mandioca, melão, melancia, milho, nabo, rabanete, pepino, abóbora, abobrinha, chuchu, maxixe, pimenta, pimentão, soja, sorgo, tomate e trigo (Nortox S.A., 2024). Ainda, no Brasil o limite máximo de resíduos (LMR) dos pesticidas costuma ser dezenas, centenas e quiçá milhares de vezes maior do que na União Europeia (UE). É o caso do resíduo do CTL no alface, que tem LMR de 0,01 mg/kg na UE e 6 mg/kg no Brasil, 600 vezes maior (Bombardi, 2023).

### 1.3.4. Efeitos do CTL

A exposição oral de pequenos mamíferos ao clorotalonil é considerada “praticamente não tóxica”  $DL_{50} > 10.000$  mg/kg, porém a toxicidade é expressivamente maior através da inalação ( $CL_{50}$  0,094 mg/L) em ratos machos

(USEPA, 1999). Segundo Hartwig e Mak Comission (2025), a exposição ao CTL acarretou em danos pulmonares, como congestão pulmonar, bronquite e rinite.

#### 1.4. Modelo animal: camundongo e células do córtex pré-frontal

A utilização de *Mus musculus* como modelo animal para investigar os efeitos tóxicos do clorotalonil justifica-se pelo seu vasto uso dentro do campo das neurociências e da toxicologia (Araújo, 2023). Acredita-se que o uso de culturas celulares híbridas modelem melhor a variação genética associada às populações humanas (French *et al.*, 2015). Culturas de células do córtex pré-frontal, região altamente desenvolvida do cérebro e responsável pela tomada de decisão e planejamento (Chen; Colonna, 2021), são particularmente vantajosas pela sua reproduzibilidade e confiabilidade devido ao controle rigoroso das condições ambientais.

#### 1.5. Sequenciamento de nova geração (NGS) e bioinformática

O sequenciamento de nova geração (NGS) representa um avanço tecnológico significativo em relação à primeira geração e caracteriza-se por ser capaz de avaliar milhões de reações de sequenciamentos simultâneos, além de reduzir o custo e o tempo para o sequenciamento (Korpelainen *et al.*, 2015). O sequenciamento de RNA (RNA-Seq) é uma tecnologia robusta, que permite a quantificação de abundância de moléculas de RNA mensageiro (mRNA) em uma amostra. A quantidade de vezes que um gene é transcrito pelo mRNA indica o seu grau de expressão. Se muitas sequências de leituras codificantes de um mesmo gene são transcritas, a proteína/enzima codificada por ele será significativamente mais expressa (Samih; Ferreira; Nikoloski, 2026). Possibilita, assim, a identificação dos transcritos que foram muito expressos e também daqueles pouco expressos (Araújo, 2023). É uma ferramenta muito útil para analisar o transcriptoma celular e suas alterações e, dentre seus principais usos, destaca-se a identificação dos genes diferencialmente expressos entre duas ou mais condições biológicas (Batut *et al.*, 2021).

## **2. OBJETIVOS**

### **2.1. Objetivo Geral**

O objetivo geral do trabalho é determinar o efeito do clorotalonil em culturas de células do córtex pré-frontal do camundongo (*Mus musculus*) através de dados públicos de sequenciamento de RNA.

### **2.2. Objetivos Específicos**

Os objetivos específicos deste trabalho são:

1. Identificar os genes cuja expressão é afetada pela exposição ao clorotalonil;
2. Comparar os efeitos de dois protocolos diferentes de pré-processamento das amostras nos resultados obtidos; e
3. Investigar o efeito do clorotalonil na expressão de genes que codificam para enzimas das vias metabólicas relacionadas ao mecanismo de ação deste pesticida.

### 3. METODOLOGIA

#### 3.1. Obtenção e caracterização das amostras de RNA-Seq

Duas amostras de CTL, com os números de acesso: SRR2078355 e SRR2078356, foram selecionadas do bioprojeto com o código de acesso PRJNA288135, no banco de dados SRA (*Sequence Read Archive*) do National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>; NCBI), usando como palavra-chave “chlorothalonil”. Este bioprojeto foi aberto para o depósito de 405 amostras de RNA-Seq, provenientes de um estudo sobre o efeito de 294 pesticidas no sistema nervoso de mamíferos (Pearson, 2016). Neste mesmo estudo, foram selecionadas as amostras veículo SRR2078648, SRR2078649, SRR2078651, SRR2078654, SRR2078659 e SRR2078661, para compor o grupo controle.

No estudo de Pearson (2016), o modelo de sistema nervoso utilizado foi a cultura primária de neurônios do córtex de camundongo (*Mus musculus*). Fêmeas C57BL/6J foram cruzadas com machos CAST/EiJ e os embriões foram coletados no décimo quarto dia de gravidez. O córtex dos embriões foi dissecado e tripsinizado com TrypLE express a 37 °C por 10 minutos. Os neurônios dissociados foram semeados em placas revestidas com poli-D-lisina (0,1 mg/mL) com meio Neurobasal (Life Technologies) contendo 5% de soro fetal bovino (Gibco), B27 (17504-044, Invitrogen) e o agente antimicrobiano Antibiotic-Antimycotic (15240-062, Invitrogen) e GlutaMAX (35050-061, Invitrogen). No terceiro dia *in vitro* (DIV 3), foi realizada troca de metade do meio, sendo o meio adicionado idêntico ao meio de plaqueamento, exceto pela omissão de soro fetal bovino e inclusão de 4,84 mg/mL de uridina 50-trifosfato (U6625, Sigma-Aldrich) e 2,46 mg/mL de 5-fluoro-20-deoxiuridina (F0503, Sigma-Aldrich) para evitar mitose nas células em divisão. No DIV 7, as culturas foram tratadas com a dose não-citotóxica de 1 µM de CTL (grupo exposto ao clorotalonil), ou com DMSO (dimetilsulfóxido) na concentração final ≤ 0,5% (grupo veículo - condição controle), por 24 horas em placas de 12 poços, na densidade de  $5 \times 10^5$  células por poço. Ao final da exposição, as células foram coletadas, e o RNA foi isolado usando o kit *RNeasy plus mini* (Cat. #74134, Qiagen) e quantificado no espectrofotômetro Nanodrop 1000 (Thermo Scientific). A qualidade do RNA foi determinada por Nanodrop e a partir do RNA

*Integrity Number* (RIN), usando Agilent Bioanalyzer 2100 ou TapeStation 2200. Amostras com valores de RIN superiores a 7 foram usadas para sequenciamento. As bibliotecas de cDNA foram construídas a partir de amostras de RNA usando o kit *TruSeq RNA Library Preparation* e sequenciadas na plataforma Illumina HiSeq 2500 em leituras (*reads*) com 50 pares de base (bp) de comprimento e extremidades pareadas na *University of North Carolina High Throughput Sequencing Facility* (Chapel Hill, NC/ Estados Unidos da América).

As amostras selecionadas para os grupos controle (veículo) e CTL (tabela 1) foram baixadas diretamente no ambiente Galaxy (The Galaxy Community, 2014; <https://usegalaxy.eu/>). A plataforma Galaxy é um ambiente digital gratuito que permite a realização de uma vasta quantidade de análises computacionais via servidores na nuvem, fornece uma interface amigável ao usuário e possui mais de 9000 ferramentas científicas, permitindo uma grande variedade de análises aplicadas às ciências da vida.

Tabela 1 – Tabela de amostras de RNA-Seq

GRUPO	NÚMERO DE ACESSO	CÓDIGO DA AMOSTRA NESTE TRABALHO
Controle (veículo - VH)	SRR2078648	VH48
	SRR2078649	VH49
	SRR2078651	VH51
	SRR2078654	VH54
	SRR2078659	VH59
	SRR2078661	VH61
Clorotalonil (CTL)	SRR2078355	CTL55
	SRR2078356	CTL56

### 3.2. Análise de qualidade

Primeiramente, a qualidade das amostras foi analisada utilizando a ferramenta Falco versão 1.2.4 (Brandine; Smith, 2021), no ambiente Galaxy. Os

parâmetros verificados nesta análise são estatísticas básicas, como: total de sequências, comprimento da sequência e taxa do conteúdo de guanina-citosina (GC), qualidade por base da sequência, qualidade média por leitura, conteúdos de sequência por base, proporção de bases GC em cada sequência, proporção de bases N em cada posição da sequência, distribuição do tamanho das leituras, nível de duplicação das leituras, sequências super-representadas e sequências de adaptadores nas leituras.

### 3.3. Pré-processamento

Ainda no ambiente Galaxy, as leituras brutas (*raw reads*) são filtradas com a ferramenta *Trimmomatic* Versão 0.39 (Bolger; Lohse; Usadel, 2014). O *Trimmomatic* é uma ferramenta baseada em Java que pode eliminar adaptadores e cortar as leituras de diferentes maneiras de acordo com a qualidade delas (Britton *et al.*, 2015). Duas diferentes trimagens foram realizadas, dando origem a dois fluxos de trabalho, e os parâmetros utilizados para cada uma estão dispostos na tabela 2. Após a trimagem, uma nova análise de qualidade é feita para avaliar o efeito do pré-processamento nas amostras.

Tabela 2 – Parâmetros de operação do *Trimmomatic*

ETAPA	Fluxo 1	Fluxo 2
Executar etapa inicial do ILLUMINACLIP?	Sim	Sim
Selecionar sequências de adaptadores padrão ou fornecer personalizadas?	Padrão	Padrão
Sequências adaptadoras a serem usadas	TruSeq3 (additional seqs) (paired-end, for MiSeq and HiSeq)	TruSeq3 (additional seqs) (paired-end, for MiSeq and HiSeq)
Contagem máxima de incompatibilidades que ainda permitirá que uma correspondência completa seja realizada	2	2
Quão precisa deve ser a correspondência entre as duas leituras 'adaptadas' para o alinhamento da leitura do palíndromo PE	30	30
Quão precisa deve ser a	10	10

correspondência entre qualquer sequência de adaptador etc. em relação a uma leitura		
Comprimento mínimo do adaptador que precisa ser detectado (modo específico de PE/palíndromo)	8	8
Sempre manter ambas as leituras (modo específico de PE/palíndromo)?	Sim	Sim
Operação 1: Cortar um número especificado de bases do início de uma leitura (HEADCROP)	10	Não
Operação 2: Corte leituras adaptativamente, equilibrando o comprimento da leitura e a taxa de erros para maximizar o valor de cada leitura (MAXINFO)		
Comprimento de leitura alvo	36	36
Rigor	0,8	0,8
Operação 3: Remover leituras abaixo de um comprimento especificado (MINLEN)		
Comprimento mínimo de leituras a serem mantidas	20	20

### 3.4. Mapeamento e quantificação de transcritos

Para quantificar os transcritos de cada amostra, utilizamos a ferramenta *Salmon quant* Versão. 1.10.1 (Patro *et al.*, 2017; Srivastava *et al.*, 2019), no ambiente Galaxy. Esta ferramenta mapeia as leituras trimadas em uma referência e estima a expressão gênica baseada na abundância das leituras mapeadas em uma dada sequência da referência. Nesta análise, utilizamos como referência o transcriptoma do camundongo, disponível no NCBI, bem como os parâmetros listados na tabela 3.

Tabela 3 – Parâmetros da operação *Salmon quant*

ETAPA	Entrada
Selecione o modo de quantificação do <i>Salmon</i>	Leituras
Selecione um transcriptoma de referência do seu histórico ou use um índice integrado?	Histórico
arquivo FASTA dos transcritos	GCF_000001635.27_GRCm39_rna.fna

Contagem de k-mer	31
Hash perfeito	Falso
Tipo de índice	quasi
Descartar quasi órfãos	Falso
Validar mapeamentos	
Fração de pontuação mínima	0,65
Pontuação de correspondência	2
Penalidade de divergência	4
Penalidade de abertura de lacuna	5
Penalidade de extensão de lacuna	3
Imitar Bowtie 2	Falso
Imitar Bowtie 2 estrito	Falso
Filtro rígido	Falso
Folga de consenso	0
Permitir Dovetail	Falso
Recuperar órfãos	Falso
Qualidade de escrita	Falso
Sucessos consistentes	Falso
Executar correção de viés específica de sequência	Sim
Executar correção de viés de GC de fragmento	Sim
incompatPrior	0
Meta	Falso

### 3.5. Similaridade de expressão gênica entre os transcriptomas

Os resultados do mapeamento e da quantificação foram extraídos do Galaxy e usados para construir uma única matriz, contendo todas as amostras. Para isso, os comandos abaixo listados, em linguagem de programação Python (v. 3.13.3), através da biblioteca de código aberto Pandas, foram utilizados através da interface do programa Visual Studio Code versão 1.105.1, a partir do código:

```

import pandas as pd
import os

os.chdir(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\ANOTAÇÃO
SEM
HEADCROP\\')
list_arch = []
for c in os.listdir():
    list_arch.append(c)

#Entrada do arquivo do clorotalonil CTL55
dfs = {}

print("ATENÇÃO: Para as amostras de clorotalonil, utilize o código 'CTL
+ o código da amostra'\nPara os veículos/controles, utilize 'VH + o código do
veículo.'")

for c in list_arch:
    print("-"*20)
    caminho = c
    print(c)
    print("-"*20)
    name = input("Qual o nome da sua amostra ou código?\n")
    if caminho.endswith(".tabular"):
        df = pd.read_table(caminho, sep="\t") #Lê o arquivo tabular e
atribui a variável
        elif caminho.endswith(".xlsx"):
            df = pd.read_excel(caminho)
            #contador = f"df{i}"
            df1 = df.drop(["Length", "EffectiveLength", "TPM"], axis=1)
            df1.columns = ["Name", f"NumReads {name}"]
            dfs[name] = df1
            print("\n", "\n")
            #[f'df{i}']

#print(dfs.keys())
keys_df = list(dfs.keys())
print(keys_df)
keys_CTL = []
keys_VH = []

for c in keys_df:
    if "CTL" in c:
        keys_CTL.append(c)

```

```

print(keys_CTL)
for c in keys_df:
    if "VH" in c:
        keys_VH.append(c)

print(keys_VH)
print("Fazendo a junção de todos...")
df_final = pd.merge(dfs[keys_df[0]], dfs[keys_df[1]], on="Name")
for key in list(dfs.keys())[2:]:
    df_final = pd.merge(df_final, dfs[key], on="Name")
print("Finalizado\n")

print("Começando a formação do Clorotalonil")
df_finalCTL = pd.merge(dfs[keys_CTL[0]], dfs[keys_CTL[1]], on="Name")
print("Finalizado :D\n")

df_finalVH = pd.merge(dfs[keys_VH[0]], dfs[keys_VH[1]], on="Name")
for key in keys_VH[2:]:
    df_finalVH = pd.merge(df_finalVH, dfs[key], on="Name")

print("Iniciando a exportação...\n")
df_final.to_excel(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\NumReads SEM
HEADCROP\\NumReadsTodos_Sem_Headcrop.xlsx', index=False, index_label=False)
df_finalCTL.to_excel(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\NumReads SEM
HEADCROP\\NumReadsCTL_Sem_Headcrop.xlsx', index=False, index_label=False)
df_finalVH.to_excel(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\NumReads SEM
HEADCROP\\NumReadsVH_Sem_Headcrop.xlsx', index=False, index_label=False)
print('CTL + VH')
print('*'*80)
print(df_final.head(5))
print('*'*80)
print('CTL')
print(df_finalCTL.head(5))
print('VH')
print('*'*80)
print(df_finalVH.head(5))
print('*'*40)
print('Exportação concluída com êxito.')
print('*'*40)

```

Nesta matriz, as linhas determinavam os códigos de identificação dos transcritos e as colunas o número de leituras de cada amostra. A partir desta matriz, foi calculado o coeficiente de correlação entre cada duas amostras, através do seguinte código:

```

import seaborn as sb
import pandas as pd
import os
import matplotlib.pyplot as plt
import numpy as np

def MEGA_file(matriz):
    #MEGA
    with open('matriz.meg', 'a') as arc:
        nome = input('Qual o nome da matriz? ')
        arc.write('#MEGA\n')
        arc.write(f'!Title: Matriz Correlação{nome};\n')
        arc.write('!Format      DataType=Distance
DataFormat=UpperRight;\n\n')

        spcs = []
        for c in matriz.columns:
            spcs.append(c)
            especie = '#' + c + '\n'
            arc.write(especie)
        arc.write('\n')

        for c in spcs:
            param = False
            valores = reversed(list(matriz.loc[c]))
            print(valores)
            for c in valores:
                if c == 1:
                    break
                local = valores
                values = str(c) + " "
                arc.write(values)

            arc.write("\n")

    os.chdir(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\NumReads COM HEADCROP')

```

```

df = pd.read_excel('NumReadsTodosComHeadcrop.xlsx')
df = df.set_index(['Name'])
print(df)
df = df.corr()
print(df)
matriz = df.where(np.triu(np.ones(df.shape)).astype(bool))
print(matriz)
#matriz.to_excel("MatrizNumRead_Sem_Headcrop.xlsx", index=False,
index_label=False)
#MEGA_file(matriz)

plot = sb.heatmap(df, annot=True)
plt.show()
print('Cabosse')
df = 1-df
df.to_excel(r'D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\CORRELAÇÃO COM
HEADCROP\\matriz_com_headcrop_corrigida.xlsx')

```

Os coeficientes de correlação foram transformados em coeficientes de distância de acordo com a equação:

$$\text{Coeficiente de distância} = 1 - \text{coeficiente de correlação}$$

Em seguida, os coeficientes de distância foram utilizados no programa MEGA (*Molecular Evolutionary Genetics Analysis*, v. 12.0.11) para representar a similaridade entre as amostras através de uma árvore filogenética *Neighbor-joining*.

### 3.6. Anotação dos transcritos em genes

A anotação dos transcritos foi extraída do próprio arquivo fasta do transcriptoma de *Mus musculus* usado como referência, através da linguagem de programação Python. Para isso, o nome e símbolo do gene referente a cada transcrito foi identificado no arquivo fasta de RNA de *Mus musculus* e inserida na tabela. Os comandos para esta operação foram os abaixo descritos:

```

from Bio import SeqIO #Análise de dados de sequências
import pandas as pd #Criação de data frames

#Abrir o arquivo com sua referência
arch = r"Galaxy35-GCF_000001635.27_GRCm39_rna.fna" #arquivo entrada

```

```

sequencias = SeqIO.parse(f"{arch}.fasta", "fasta") #acessar cada sequência

#Extrair dados das sequências (ids, descrição e código)
def description(descricao):
    num0 = descricao.find("musculus")
    desc_filtrada = descricao[(num0+9):]
    num1 = desc_filtrada.rfind("(")
    num2 = desc_filtrada.rfind(")")
    ofcname = (desc_filtrada[(num1+1):num2])
    return desc_filtrada, ofcname

#Abertura de listas para adicionar os dados de cada id:
ids = []
desc = []
ofc_name = []

#Extrair dados e inserir nas listas
for c in sequencias:
    ids.append(c.id)
    desc.append(description(c.description)[0])
    ofc_name.append(description(c.description)[1])

#Criação de um dataframe com as colunas e dados
df = pd.DataFrame({
    "Name" : ids,
    "Descrição" : desc,
    "Nome Oficial" : ofc_name
})

#Exportar o dataframe para um arquivo excel
df.to_excel("IdsAnotadas.xlsx", index=False)

```

Com o Python e a biblioteca Pandas, unimos os dados de mapeamento e quantificação de cada amostra, obtidas conforme item anterior, com a planilha dos transcritos anotados, contendo a sigla do gene codificante e a descrição por extenso do que cada gene codifica. O resultado desta união foi uma tabela para cada amostra contendo as colunas “*Name*”, “*Nome Oficial*”, designação escolhida para se referir ao nome do gene, “*Descrição*” com o nome por extenso do gene, e dados de expressão “*TPM*”. O código usado foi o seguinte:

```

import pandas as pd

df = pd.read_excel(r"D:\\UFPE\\GRADUAÇÃO\\TCC\\REFERÊNCIAS\\IdsAnotadas.xlsx")
df1 = pd.read_table(r"D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\ANOTAÇÃO SEM
HEADCROP\\SRR2078355.tabular", sep='\t')

df_final = pd.merge(df[['Name', 'Nome Oficial', 'Descrição']],
df1[['Name', 'TPM']], on='Name', how='right')
print(df_final)

```

### 3.7. Preparação dos dados para as análises de expressão gênica

A linguagem de programação Python foi utilizada para preparar as tabelas de dados para a análise de expressão diferencial. A preparação dos dados consistiu da remoção das colunas “*Length*”, “*EffectiveLength*” e “*TPM*” das tabelas resultantes da operação de mapeamento e contagem, e do acréscimo do código identificador de cada amostra à coluna “*NumReads*” (ex.: “*NumReads CTL56*”, referindo-se à amostra de clorotalonil com terminação do código SRR em 56). Os comandos usados para tanto foram:

```

import pandas as pd
import os

os.chdir(r"D:\\UFPE\\GRADUAÇÃO\\TCC\\AMOSTRAS\\CTL SEM
HEADCROP\\SRR2078356 SEM HEADCROP")
df = pd.read_table(r'SRR2078356.tabular', sep='\t', decimal='.')
df = df[['Name', "NumReads"]]
df.rename(columns={'NumReads': 'NumReads CTL56'}, inplace=True)
df.to_csv(r'SRR2078356_tratado.tabular', sep='\t', index=False)
print('Concluído com êxito.')

```

### 3.8. Análise de expressão gênica diferencial (DEG)

Após a preparação das tabelas conforme item 3.7, as tabelas foram carregadas no ambiente Galaxy, onde foi realizada a análise de expressão gênica diferencial (*differentially expressed genes*, DEG) com a ferramenta edgeR versão 3.36.0 (Robinson; McCarthy; Smyth, 2010; Liu *et al.*, 2015). Nesta operação foram utilizados os parâmetros da tabela 4. A estatística do edgeR assumiu o limiar do valor de P em 0,05, o método de ajuste foi o Benjamini e Hochberg (1995). O valor de expressão foi considerado estatisticamente significativo se o FDR (*False Discovery Rate*) fosse menor que 0,05. O método de normalização foi o *Trimmed Mean of M-values* (TMM).

Tabela 4 – Parâmetros de entrada da análise de expressão gênica diferencial com o edgeR

ETAPA	Entrada
Contar arquivos ou matriz?	Arquivos
Nome	DGE_CTL_e_VH
Nome	CTL
Arquivos	SRR2078355_tratado.tabular
	SRR2078356_tratado.tabular
Nome	VH
Arquivos	SRR2078648_tratado.tabular
	SRR2078649_tratado.tabular
	SRR2078651_tratado.tabular
	SRR2078654_tratado.tabular
	SRR2078659_tratado.tabular
	SRR2078661_tratado.tabular
Usar anotações genéticas?	Falso
Contrastes de entrada manualmente ou por meio de um arquivo	Manualmente
Contraste de interesse 1	CTL-VH
Contraste de interesse 2	VH-CTL
Filtrar genes pouco expressos?	Não

Tabela de contagens normalizadas de saída?	Não
--	-----

### 3.9. Análise da expressão de genes envolvidos no mecanismo de ação do CTL

A expressão de genes que codificam para enzimas relacionadas ao mecanismo de ação do CTL foi investigada de duas formas. Primeiramente, foi construída uma lista de genes de interesse envolvidos em rotas metabólicas associadas, em estudos prévios, aos mecanismos de ação do CTL. À lista publicada no estudo de Hoff *et al.* (2017), abrangendo genes da via glicolítica e da fosforilação oxidativa, foram adicionados genes que codificam enzimas envolvidas no metabolismo de glutationa. Estes genes adicionados foram curados a partir da rota descrita como “00480 - *Glutathione metabolism*” na página Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.kegg.jp/>). Os genes de interesse listados foram selecionados e extraídos da tabela de saída da análise de expressão gênica diferencial realizada com o edgeR. Nesta tabela são fornecidos dados da diferença da expressão entre os grupos veículo (controle) e CTL em LogFC. Esta unidade se refere ao logaritmo na base dois (2) da diferença de expressão entre os grupos representada por *fold change* (FC) da expressão do grupo controle. A unidade “*fold change*” é calculada da seguinte forma:

$$\text{Fold change} = (\text{expressão média no grupo CTL}) / (\text{expressão média no grupo controle})$$

Posteriormente, a análise de vias metabólicas foi realizada com Pathview versão 1.34.0 (Luo; Brouwer, 2013), dentro do ambiente Galaxy. Nesta operação, unificamos as leituras do Salmon pelo nome dos genes no qual elas mapearam e rodamos o edgeR para obter os valores de LogFC. A tabela de saída do edgeR foi preparada para que mantivesse apenas as colunas “Gene Symbol” e “LogFC”. Esta tabela foi então usada na análise de vias por Pathview, segundo os parâmetros listados na tabela 5. As seguintes vias metabólicas foram analisadas: “00010 - Glycolysis/Gluconeogenesis”; “00020 - Citrate cycle (TCA cycle)”; “00190 - Oxidative phosphorylation” e “00480 - Glutathione metabolism”.

Tabela 5 – Parâmetros de entrada para a análise de vias *Pathview*

ETAPA	ENTRADA
Número de vias para plotar	Uma
Código da via no KEGG	00010
Espécie a se usar	Rato
Prover um arquivo de gene	DGE_CTL-VH_ANOTADO_LogFOLD_CHANGE_Grouped.tabular
O arquivo tem cabeçalho?	Sim
Formato para dos dados gênicos	“Gene Symbol”
Prover um arquivo de composto?	Falso
Integrar e representar no mesmo gráfico múltiplos estados de genes ou dados de compostos verdadeiros?	Sim
Os dados de genes e compostos estão verdadeiramente pareados?	Sim
Saída para o caminho	Sim
Plotar na mesma camada?	Sim
Mapear o gene NULL ou os dados reais compostos para a via metabólica?	Sim

O fluxograma de trabalho que seguimos para os dois fluxos de trabalhos seguiu da análise de qualidade, passou pelos pré-processamentos distintos, mapeamentos, análise de correlação, e chegou à árvore *neighbor joining*. A partir daí, seguimos apenas com o fluxo 2 pela análise de expressão diferencial e pela análise de vias *Pathview* (figura 2).

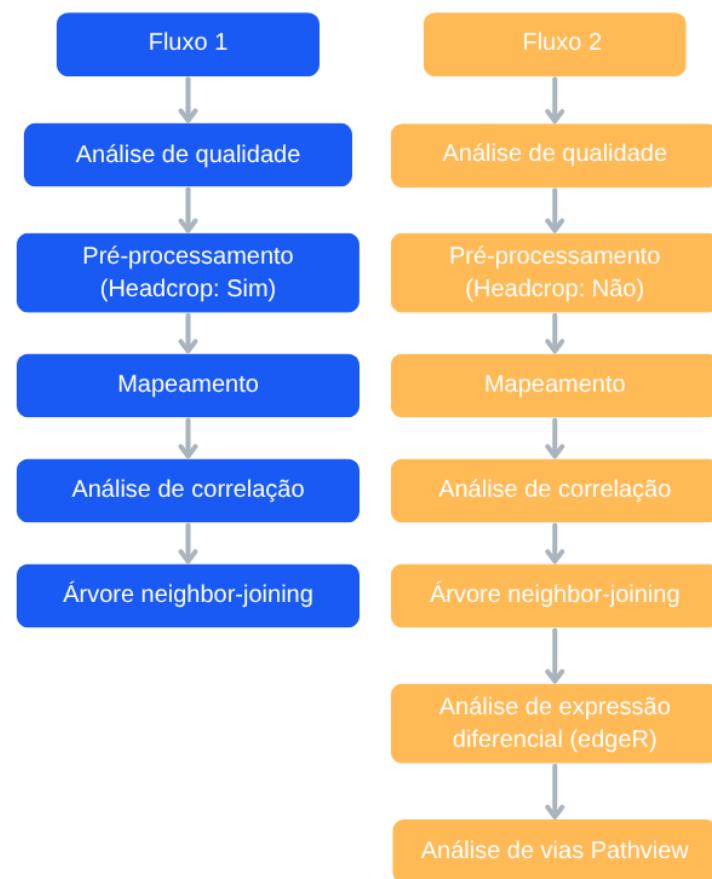


Figura 2 – Fluxograma de trabalho para os dois fluxos de pré-processamento

## 4. RESULTADOS

### 4.1. Análise de qualidade dos dados brutos

As amostras, tanto do grupo CTL quanto do grupo controle, apresentaram um total de sequências variando entre 3.284.215, referente à amostra CTL55 *forward* e 7.059.616 *reads*, referente à amostra VH61 *forward*, comprimento da sequência de 48 bp, e conteúdo de guanina-citosina (GC) entre 48 e 50%. A mediana do *Phread* oscilou entre 34 e 40. A análise de qualidade dos dados brutos apontou a presença dos adaptadores *PolyA*, *PolyG*, *Nextera Transposase Sequence* e *Illumina Universal Adapter*. Todas as amostras falharam no parâmetro “Conteúdo de bases das sequências”. Como é possível notar na figura 3, houve uma grande variação no conteúdo de bases nitrogenadas até aproximadamente a 15<sup>a</sup> posição das sequências das leituras.

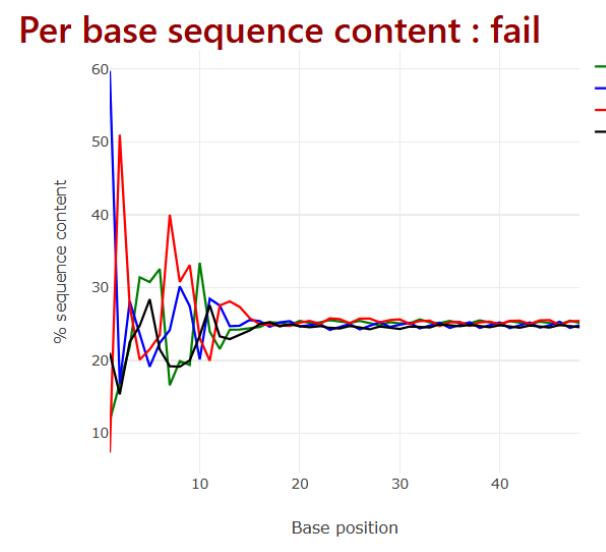


Figura 3 – Falha do conteúdo de sequência por base.

### 4.2. Análise de qualidade dos dados trimados

Como resultado do pré-processamento seguindo o fluxo 1, todas as amostras foram aprovadas em estatísticas básicas (tabela 6), qualidade por base da sequência, qualidade média por leitura, conteúdo de bases G e C em cada

sequência, proporção de bases N em cada posição da sequência, nível de duplicação das leituras, sequências super-representadas e presença de sequências de adaptadores nas leituras. No parâmetro distribuição de comprimento de sequência, todas as amostras receberam um aviso, como evidenciado pela figura 4.

Tabela 6 – Parâmetros estatísticos básicos das amostras trimadas nos fluxos 1 e 2

amostra (SRR)	sentido	Fluxo 1 (com <i>headcrop</i> )			Fluxo 2 (sem <i>headcrop</i> )		
		número total de sequências (reads)	comprimento das sequências (bp)	conteúdo de GC (%)	número total de sequências (reads)	comprimento das sequências (bp)	conteúdo de GC (%)
SRR2078355	forward	3.259.707	31-38	49	3.259.707	34-48	48
	reverse	3.259.707	33-38	49	3.259.707	35-48	48
SRR2078356	forward	5.476.900	28-38	50	5.476.900	35-48	50
	reverse	5.476.900	33-38	50	5.476.900	35-48	50
SRR2078648	forward	5.637.280	35-38	51	5.637.280	35-48	50
	reverse	5.637.280	31-38	51	5.637.280	35-48	50
SRR2078649	forward	5.247.788	35-38	50	5.247.788	35-48	50
	reverse	5.247.788	29-38	50	5.247.788	35-48	50
SRR2078651	forward	5.043.494	35-38	50	5.043.494	35-48	50
	reverse	5.043.494	30-38	50	5.043.494	35-48	50
SRR2078654	forward	5.759.174	34-38	51	5.759.175	35-48	50
	reverse	5.759.174	35-38	51	5.759.175	24-48	50
SRR2078659	forward	5.309.868	30-38	50	5.309.868	35-48	49
	reverse	5.309.868	30-38	50	5.309.868	35-48	49
SRR2078661	forward	7.010.518	35-38	50	7.010.518	33-48	50
	reverse	7.010.518	31-38	50	7.010.518	33-48	50

# Report

Wed Oct 1 01:40:45 2025

forward

## Summary

### Basic Statistics

- Per base sequence quality
- Per sequence quality scores
- Per base sequence content
- Per sequence GC content
- Per base N content
- Sequence Length Distribution
- Sequence Duplication Levels
- Overrepresented sequences
- Adapter Content

## Basic Statistics: pass

Measure	Value
Filename	forward
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	5476900
Sequences Flagged As Poor Quality	0
Sequence length	28 - 38
%GC:	50

Figura 4 – Representação da análise de qualidade das amostras trimadas no fluxo 1.

Com relação ao fluxo 2, a qualidade de todas as amostras após a trimagem foi aprovada na maioria dos parâmetros avaliados, com exceção da reprovação no conteúdo por base das sequências e do aviso para o parâmetro distribuição do tamanho das leituras (figura 5). As estatísticas básicas das amostras trimadas para o fluxo 2 estão apresentadas na tabela 6.

# Report

Thu Sep 11 01:10:24 2025

forward

## Summary

- ### Basic Statistics
- Per base sequence quality
  - Per sequence quality scores
  - Per base sequence content
  - Per sequence GC content
  - Per base N content
  - Sequence Length Distribution
  - Sequence Duplication Levels
  - Overrepresented sequences
  - Adapter Content

## Basic Statistics: pass

Measure	Value
Filename	forward
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	3259707
Sequences Flagged As Poor Quality	0
Sequence length	34 - 48
%GC:	48

Figura 5 – Representação da análise de qualidade das amostras trimadas no fluxo 2

#### 4.3. Similaridade da expressão gênica entre os transcriptomas

A árvore *neighbor-joining* construída a partir da distância das expressões gênicas entre os transcriptomas analisados do fluxo 1 mostra as amostras de CTL pouco relacionadas, bem como não há um agrupamento claro das amostras de veículo (figura 6).

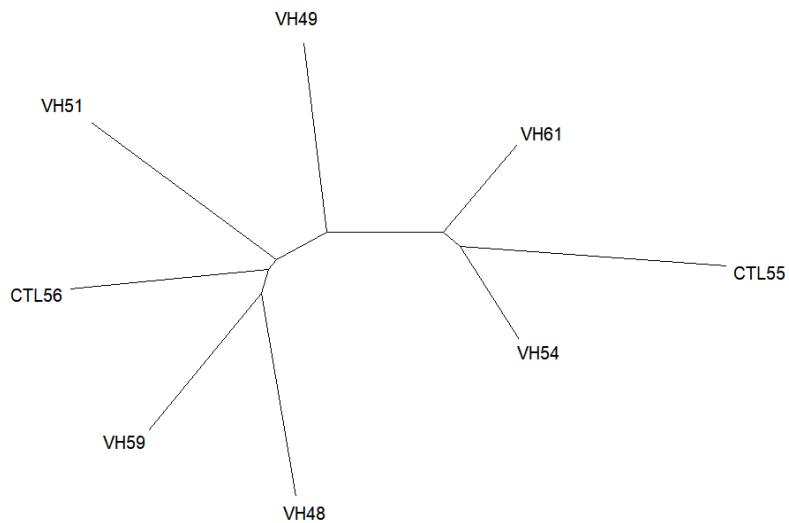


Figura 6 – Representação da similaridade de expressão gênica entre as amostras analisadas pelo fluxo 1.

A árvore *neighbor-joining* construída a partir da distância das expressões gênicas entre os transcriptomas analisados do fluxo 2 evidencia as amostras de CTL mais proximamente relacionadas, assim como o grupo controle forma um agrupamento melhor definido (figura 7).

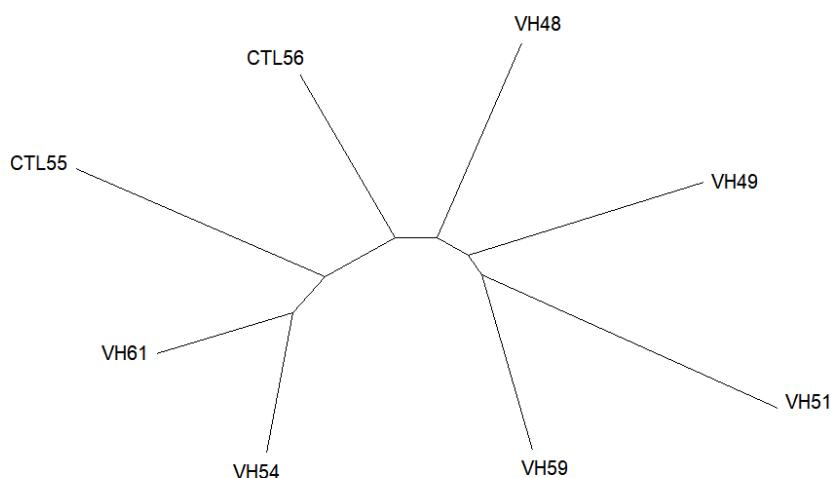


Figura 7 – Árvore filogenética *neighbor-joining* das amostras do fluxo 2

#### 4.3.1. Análise de genes diferencialmente expressos

Os resultados da análise de genes diferencialmente expressos (DEG) no fluxo 2 mostram a regulação negativa de três genes (tabela 7). O gene que codifica para a proteína de ligação a nucleotídeos de guanina, o Gnb1 foi o gene com maior diferença na expressão em relação à condição controle. Gnb1 apresentou expressão de cerca de vinte oito mil vezes menor nas células tratadas com CTL em relação à condição controle (tabela 7). O segundo gene mais diferencialmente expresso foi o gene Pkp4 (placofilina 4) e o terceiro gene foi PTPRF Trio (proteína de domínio triplo funcional), aproximadamente 6500 e 6200 vezes menos expresso no CTL em relação ao controle, respectivamente.

Tabela 7 – Resultados da análise de genes diferencialmente expressos no fluxo 2

Gene ID	Símbolo	Fold Change	FDR
NM_008142.5	Gnb1	-28229,37	0,004637935974
NM_026361.2	Pkp4	-6491,42	0,04019869298
NM_001081302.1	PTPRF (Trio)	-6176,94	0,04092260684

#### 4.4. Análise dos genes mais diferencialmente expressos no CTL em relação ao controle

Os dez genes mais diferencialmente expressos com expressão positiva no CTL, dada em *fold change* em relação ao controle estão apresentados na tabela 8. Para nenhum destes genes, entretanto, foi encontrada significância na diferença de expressão (tabela 8). Dentre os genes com maior diferença de expressão destaca-se o Map2 (proteína 2 associada a microtúbulos), com uma expressão mais de onze mil vezes maior no tratamento com CTL em relação ao veículo ( $F = 19,70$ ;  $FDR = 1$ ;  $P = 0,009471$ ). Nesta tabela também é possível notar que o gene DCX (*doublecortin*) teve uma expressão cerca de oito mil vezes maior no grupo CTL em relação ao controle ( $F = 22,07$ ;  $FDR = 1$ ;  $P = 0,004358$ ). Quando observamos os genes mais diferencialmente expressos com regulação negativa (tabela 9), encontramos o Gnb1, com expressão cerca de vinte oito mil vezes menor nas células tratadas com CTL em relação à condição controle ( $F = 198,25$ ;  $FDR = 0,0046$ ;  $P = 0,0000000342$ ).

Tabela 8 – Dez genes mais diferencialmente expressos com expressão positiva no CTL dada em *fold change* em relação ao controle

<b>Gene ID</b>	<b>Símbolo</b>	<b>Fold Change</b>	<b>F</b>	<b>FDR</b>	<b>P</b>
XM_006495756.4	Map2	11030,09	19,7098	1	0,009471
XM_036162881.1	Dclk1	10752,55	19,7680	1	0,009230
XM_036160951.1	Apc	9644,48	20,9459	1	0,006552
NM_001401929.1	Ids	9209,67	112,3847	0,3435	0,000022
NM_001410010.1	Dcx	8318,06	22,0763	1	0,004358
NM_001426499.1	Amph	7686,11	22,8907	1	0,003162
NM_010906.4	Nfix	7033,92	22,8980	1	0,002717
NR_182008.1	Snap25	6904,04	23,1119	1	0,002460
NM_001243047.1	Dlg2	6798,95	71,9650	0,6231	0,000088
XM_006529075.4	Hdlbp	6769,31	23,2348	1	0,002275

Tabela 9 – Dez genes mais diferencialmente expressos com expressão negativa no CTL dada em *fold change* em relação ao controle

<b>Gene ID</b>	<b>Símbolo</b>	<b>Fold Change</b>	<b>F</b>	<b>FDR</b>	<b>P</b>
NM_008142.5	Gnb1	-28229,37	198,25	0,0046	0,0000000342
XM_006528759.1	Tmsb4x	-16391,05	6,82	1	0,037622
NM_008478.3	L1cam	-15809,66	7,23	1	0,033354
XM_006511268.4	Tmod2	-13597,49	13,26	1	0,007557

NM_001159634.1	Prrc2b	-12028,14	5,64	1	0,051951
NM_001146123.1	Psap	-11764,96	3,78	1	0,096930
NM_175836.2	Sptbn1	-11747,44	2,82	1	0,142450
XM_036155218.1	Rpl4	-11091,43	6,01	1	0,045984
XM_006522883.5	Cxadr	-10542,70	3,88	1	0,093303
XM_017321377.3	Chl1	-9606,35	9,68	1	0,015401

#### 4.5. Análise de genes envolvidos no mecanismo de ação do CTL

Dentre os genes de interesse listados, os genes Eno 2, Sdhc, Ndufs1 e Ndufs6 apresentam maior expressão diferencial no grupo clorotalonil em relação ao controle. O gene associado ao mecanismo de ação do CTL com maior expressão dada em *fold change* positiva em relação ao controle é Eno2 (enolase 2), que está 3.769,4 vezes mais expresso no grupo CTL ( $F = 16,29$ ;  $FDR = 1$ ;  $P = 0,005624867687$ ). O gene Sdhc (complexo da succinato desidrogenase, subunidade C) está cerca de 700 vezes mais expresso no grupo clorotalonil ( $F = 8,96$ ;  $FDR = 1$ ;  $P = 0,02159$ ) enquanto o Ndufs1 (subunidade central S1 da ubiquinona oxidorredutase;  $F = 8,64861$ ;  $FDR = 1$ ;  $P = 0,02325$ ) e o Ndufs6 (subunidade central S6 da NADH:ubiquinona oxidorredutase;  $F = 8,46239$ ;  $FDR = 1$ ;  $P = 0,02429$ ) estão cerca de 600 e 450 vezes mais expressos, respectivamente, em relação ao controle (tabela 11).

Tabela 11 – Dez genes relacionados ao mecanismo de ação do CTL mais expressos no grupo CTL

Gene ID	Símbolo (nome) do gene	Fold change	F	FDR	P
NM_001409687.1	Eno2 (enolase 2)	3769,42	16,29449	1	0,00562
XM_006496955.3	Sdhc (complexo da succinato desidrogenase, subunidade C)	694,73	8,96967	1	0,02159
XM_036164425.1	Ndufs1 (subunidade central S1 da ubiquinona oxidorredutase)	597,36	8,64861	1	0,02325
NM_001423716.1	Ndufs6 (subunidade central S6 da NADH:ubiquinona oxidorredutase)	452,79	8,46239	1	0,02429
NR_187114.1	Ndufs6 (subunidade central S6 da NADH:ubiquinona oxidorredutase)	219,79	8,06190	1	0,02675

NM_001378832.1	Slc37a4 (transportador de soluto família 37 [transportador de glicose-6-fosfato])	218,46	8,21644	1	0,02576
NM_001378833.1	Slc37a4 (transportador de soluto família 37 [transportador de glicose-6-fosfato])	195,95	8,15445	1	0,02615
NM_001361905.1	Ogdh (oxoglutarato [alfa-cetoglutarato] desidrogenase [lipoamida])	49,17	7,74586	1	0,02892
XM_030245650.1	Ogdh (oxoglutarato [alfa-cetoglutarato] desidrogenase [lipoamida])	39,56	3,13463		0,10292
XM_036164417.1	Ndufs1 (subunidade central S1 da ubiquinona oxidorredutase)	37,76	10,50521		0,01549

O gene associado ao mecanismo de ação do CTL com maior diferença de expressão dada em *fold change* negativa em relação ao controle é Ldha, 4.584,85 vezes menos expresso no CTL ( $F = 3,12794$  ; FDR = 1 ; P = 0,11947; tabela 12). Outros genes mais inibidos no grupo CTL são o Pkm, o Ogdh e o Ndufs1, cerca de 4.600, 2.100 e 1.100 vezes, respectivamente, mais expressos no controle.

Tabela 12 – Dez genes relacionados ao mecanismo de ação do CTL menos expressos no grupo CTL

Gene ID	Símbolo (nome) do gene	Fold-change	F	FDR	P
NM_001136069.2	Ldha (lactato desidrogenase subunidade a)	-4584,85	3,12794	1	0,11947
NM_001405491.1	Pkm (piruvato kinase)	-2094,01	3,98663	1	0,07222
XM_030245651.1	Ogdh (oxoglutarato [alfa-cetoglutarato] desidrogenase [lipoamida])	-1694,53	16,22300	1	0,00215
NM_001378867.1	Pkm (piruvato kinase)	-1210,69	2,92026	1	0,11658
NM_001160038.1	Ndufs1 (subunidade central S1 da ubiquinona oxidorredutase)	-1127,67	4,08526	1	0,06928
XM_030241961.1	Sdhc (complexo da succinato desidrogenase, subunidade C)	-577,63	3,21327	1	0,10163
NR_153798.1	Ndufab1 (subunidade AB1 da NADH:ubiquinona oxidorredutase)	-507,76	5,04554	1	0,04709
NM_00136	Ndufs7 (subunidade central S7 da	-368,30	2,72038	1	0,12840

4694.1	NADH:ubiquinona oxidorredutase)				
XR_00395 0460.1	Pfkp (fosfofrutocinase)	-325,67	2,67584	1	0,13123
NM_01027 4.3	Gpd2 (glicerol fosfato desidrogenase 2)	-240,47	2,52100	1	0,14173

A análise da via glicolítica pelo Pathview mostra que os genes que codificam para as enzimas GCK (glicoquinase, E.C. 2.7.1.2), ADPGK (a glicoquinase dependente de ADP, E.C. 2.7.1.147), PGK (fosfoglicerato quinase, E.C. 2.7.2.3) e Eno (enolase, E.C. 4.2.1.11) foram inibidos transcripcionalmente (em verde na figura 8). Dentro das enzimas com regulação transcripcional positiva, indicada pela cor vermelha na figura 8, estão a HK (hexokinase, E.C. 2.7.1.1), a GPI (glicose-6-fosfato isomerase, E.C. 5.3.1.9), a PFK (6-fosfofrutokinase, E.C. 2.7.1.11), a ALDO (frutose-bisfosfato aldolase, classe I, E.C. 4.1.2.13), a GAPDH (gliceraldeído 3-fosfato desidrogenase, E.C. 1.2.1.12) e a PK (piruvato kinase, E.C. 2.7.1.40).

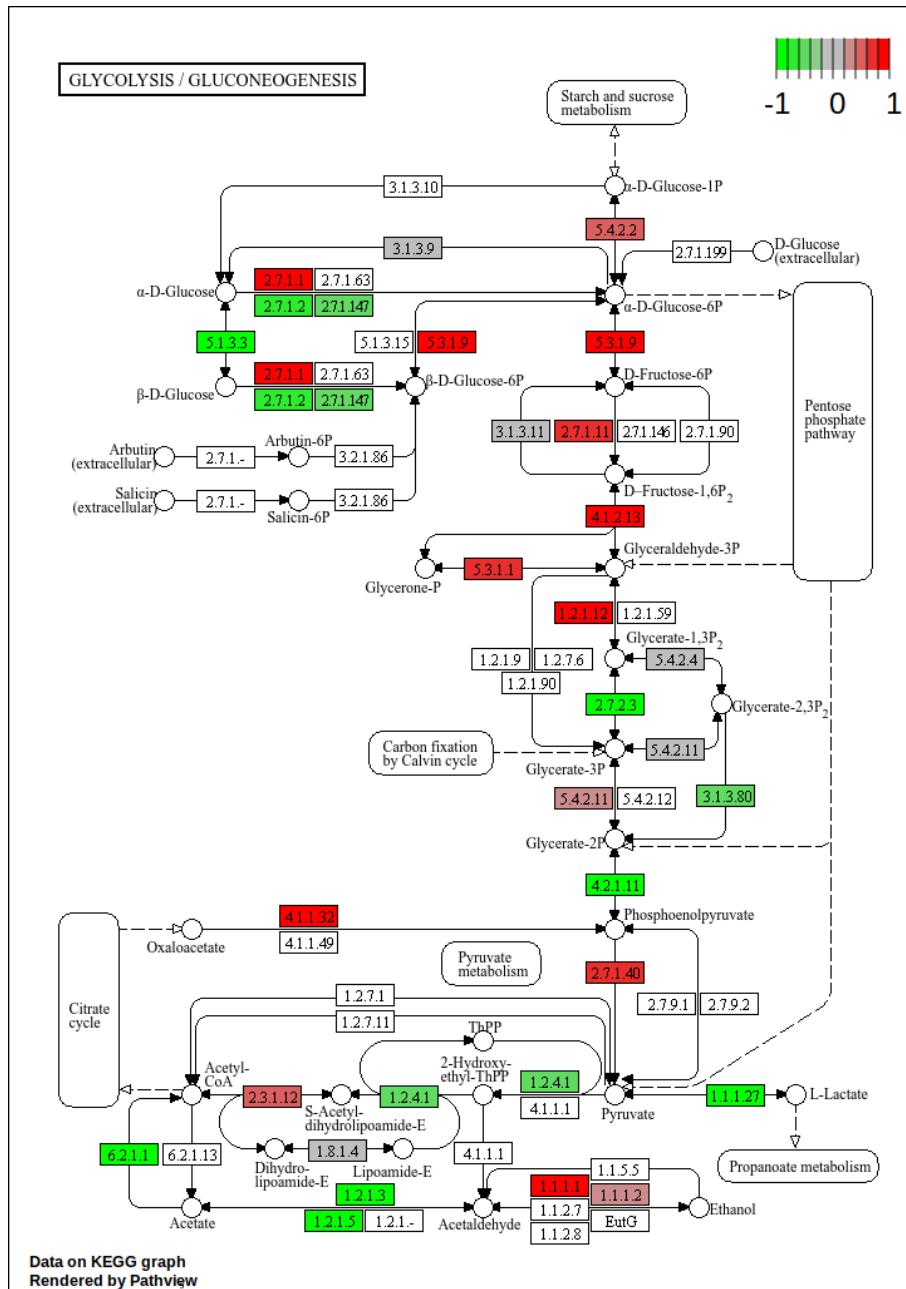


Figura 8 – Análise de via da rota glicolítica

A análise de vias referente ao ciclo do ácido tricarboxílico mostra quatro enzimas com transcrição regulada positivamente em resposta ao tratamento com CTL (em vermelho na figura 9). São elas: a ACLY (ATP citrato [pro-S]-liase, E.C. 2.3.3.8), a DLST (componente E2 da 2-oxoglutarato desidrogenase [succiniltransferase de diidrolipoamida], E.C. 2.3.1.61), a SDH (succinato desidrogenase, E.C. 1.3.5.1) e a MDH1 (malato desidrogenase, E.C. 1.1.1.37). Dentre as enzimas com regulação transcripcional negativa (em verde na figura 9)

estão: a IDH1 (isocitrato desidrogenase, E.C. 1.1.1.42), a OGDH (componente E1 da 2-oxoglutarato desidrogenase, E.C. 1.2.4.2), a sucD (subunidade alfa da succinil-CoA sintetase, E.C. 6.2.1.5), a LSC1 (subunidade alfa da succinil-CoA sintetase, E.C. 6.2.1.4) e a fumA (fumarato hidratase, classe I, E.C. 4.2.1.2).

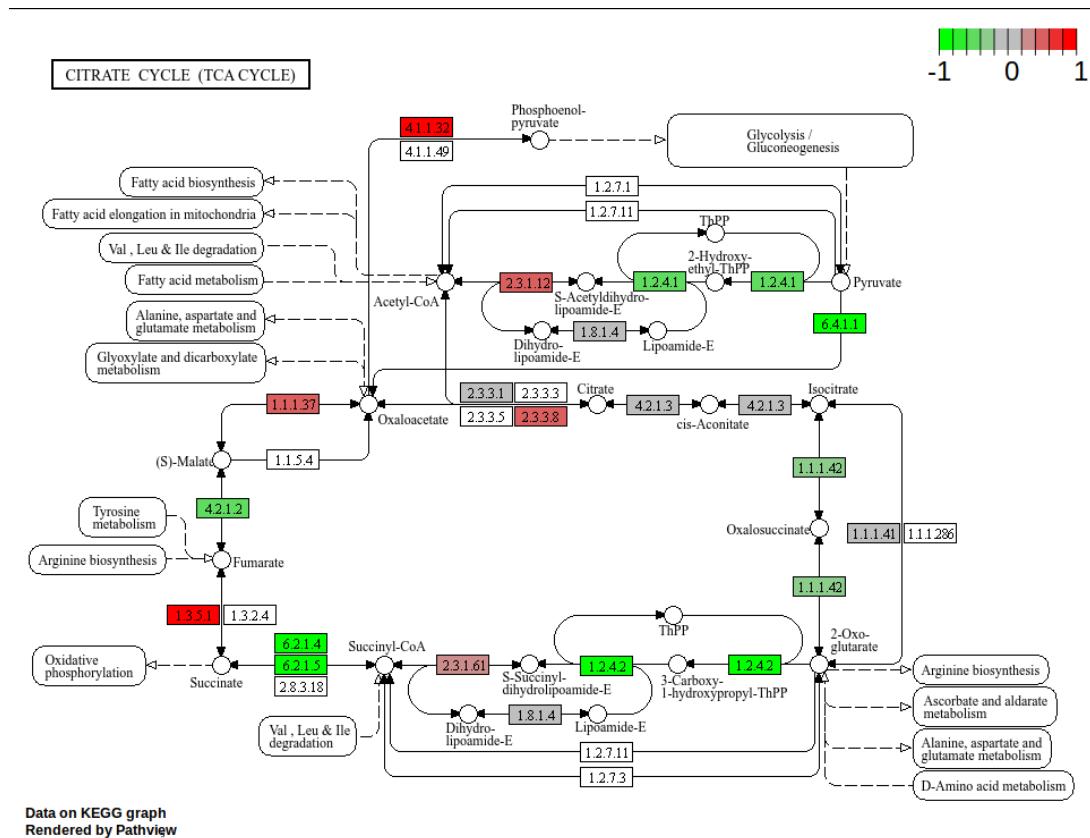


Figura 9 – Rota do ciclo do ácido tricarboxílico (TCA)

Com relação à análise de via da fosforilação oxidativa, o conjunto de alterações de expressão gênica revelou expressão negativa (em verde na figura 10) da subunidade A da NADH-quinona oxidoreduktase (E.C. 7.1.1.2) do complexo I da cadeia transportadora de elétrons. As enzimas com regulação transcripcional positiva (em vermelho na figura 10) foram a succinato desidrogenase (SDH , E.C. 1.3.5.1), a citocromo c oxidase (E.C. 7.1.1.9), a subunidade a transportadora de H<sup>+</sup> do tipo F (ATPF0A, E.C. 7.1.2.2) e a subunidade alfa da ATPase de troca H<sup>+</sup>/K<sup>+</sup> (ATP4A E.C. 7.2.2.19), constituintes da unidade F1 da ATP sintase.

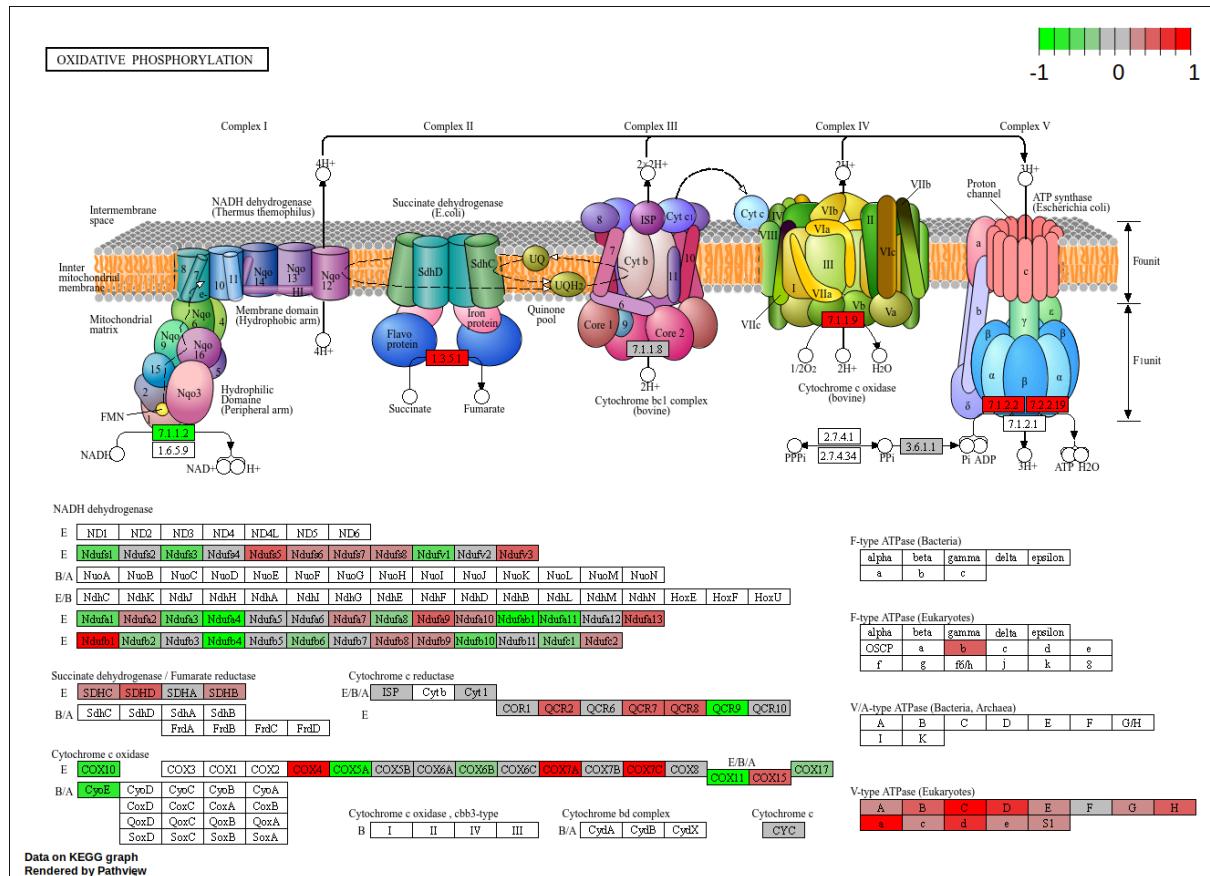


Figura 10 – Rota da fosforilação oxidativa

Com relação ao metabolismo de glutationa, a análise de via aponta a regulação positiva do gene que codifica para a subunidade catalítica da enzima glutamato cisteína ligase (GCL, E.C. 6.3.2.2) nas culturas de células expostas ao CTL em relação às culturas de células expostas ao veículo (em vermelho na figura 11). Além disso, o tratamento com CTL causou a regulação negativa (em verde na figura 11) dos genes codificantes para as enzimas 5-oxiprolinase (E.C. 3.5.2.9), glutationa sintetase (GS, E.C. 6.3.2.3), bem como para GGT (gama-glutamiltranspeptidase/glutationa hidrolase, E.C. 3.4.19.13).

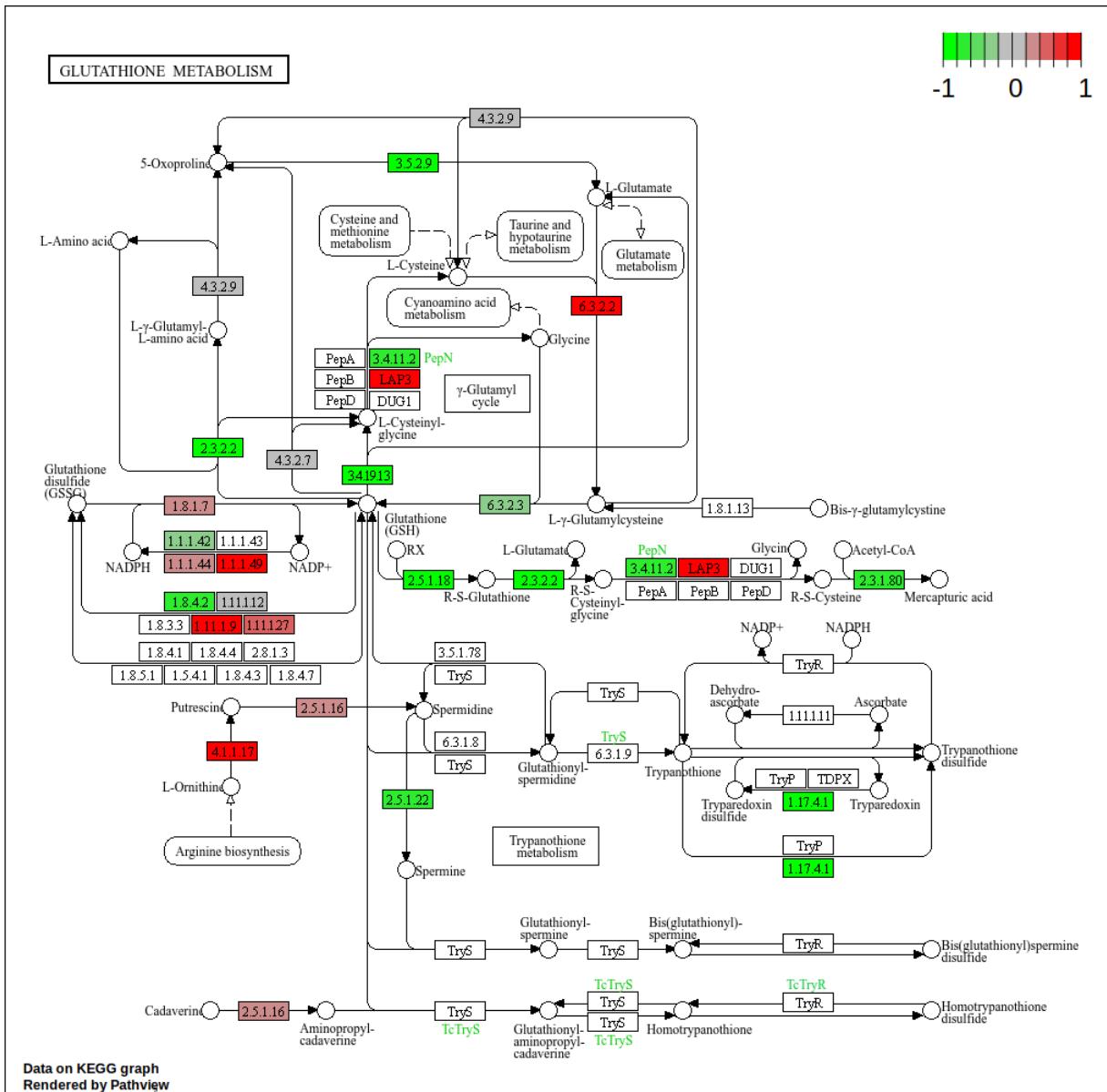


Figura 11 – Rota do metabolismo da glutatona

## 5. DISCUSSÃO

Nosso estudo mostra o efeito do pré-processamento e do desenho experimental nas análises de RNA-Seq, bem como o efeito do CTL em cultura primária de neurônios de camundongo.

### 5.1. Efeito do pré-processamento

#### 5.1.1. Comparação entre o fluxo 1 e 2

A análise de qualidade das amostras de RNA-Seq selecionadas mostrou que elas estavam homogêneas e apresentavam boa qualidade geral. Entretanto, as amostras falharam no conteúdo da sequência por base, variando até a 15<sup>a</sup> base das *reads*. Sobre o sequenciamento da nova geração (*next generation sequencing, NGS*) da Illumina, Korpelainen (2015) afirma “foi demonstrado que o priming usando hexâmeros aleatórios induz vieses na composição de nucleotídeos no início das leituras de RNA-seq”. Dessa forma, a falta de compatibilidade no início das sequências das leituras pode ser um reflexo do próprio método de NGS utilizado na obtenção dos transcriptomas. Tendo em vista esta variabilidade no início das sequências, resolvemos seguir dois fluxos de análise: o fluxo 1, no qual a trimagem excluiu as 10 primeiras bases das leituras, e o fluxo 2, no qual estas 10 primeiras bases foram mantidas.

A similaridade entre os transcriptomas foi quantificada por análise de correlação da expressão gênica entre todas as amostras para cada um dos fluxos. Para visualizar a relação entre os transcriptomas, construímos árvores *neighbor-joining* baseadas nas distâncias entre as expressões gênicas, calculadas a partir dos coeficientes de correlação. Nestas árvores, esperávamos a formação de dois grupos claros: o grupo CTL e o grupo controle.

A árvore para o fluxo 1 não agrupou as amostras de veículo e nem as de CTL de forma clara. As amostras de clorotalonil apareceram de lados opostos da árvore, sugerindo pouca similaridade entre os transcriptomas. Uma provável explicação está relacionada à trimagem realizada no fluxo 1. Após a trimagem, o comprimento de sequência das amostras do fluxo 1 oscilou entre 28 e 38 pares de base. Molinari *et al.* (2018) afirmam que a taxa de duplicatas não-naturais em *reads* de 50 bp são maiores do que em *reads* com 100 bp ou mais. Leituras muito pequenas têm pouco

valor por serem muito ambíguas para serem informativas (Bolger; Lohse; Usadel, 2014). Esses dados corroboram com a ideia de que *reads* menores tendem a apresentar maior redundância e menor precisão no mapeamento. Em sequências curtas, como o caso da nossa (48 bp), pode ser mais vantajoso manter essas bases iniciais da leitura na trimagem para que as leituras tenham comprimento suficiente para serem informativas (Korpelainen, 2015). Logo, a trimagem das dez primeiras bases das *reads* pode ter afetado o mapeamento e diminuído o conteúdo de informação.

A árvore para o fluxo 2, por outro lado, mostrou maior proximidade entre os transcriptomas das amostras CTL. Apesar disso, os transcriptomas de amostras controle no fluxo 2 não formaram um grupo único. Uma possível explicação para a variabilidade entre as amostras controle é o fato do bioprojeto incluir 55 amostras de veículo e não ser claro a respeito de quais amostras de RNA-Seq foram obtidas como controle no experimento em que as células foram expostas ao CTL. Além disso, neste bioprojeto foram geradas somente duas amostras de RNA-Seq tratadas com CTL. Apesar destas limitações, com este desenho experimental esperamos encontrar poucos genes com expressão diferencial, mas indicação robusta do efeito do CTL na expressão destes genes em modelo de culturas primárias de neurônios de camundongo. Para tanto, a escolha por seguir o fluxo 2 para análise e interpretação de dados apresentou-se como mais adequada para o nosso trabalho.

### 5.1.2. Análise de genes diferencialmente expressos (DEG)

A partir da visualização da similaridade entre transcriptomas na forma de árvores filogenéticas, decidimos aprofundar as análises no fluxo 2. Neste fluxo de trabalho, três genes foram diferencialmente expressos, todos os quais foram regulados negativamente no tratamento com CTL. Estes genes são: Gnb1, Pkp4 e PTPRF (Trio).

A Subunidade beta-1 da proteína de ligação a nucleotídeos de guanina (Gnb1) age principalmente com a sinalização celular por meio de proteínas heterotriméricas e dos receptores acoplados à proteína G (GPCRs). É um ator central na transdução de sinais de membrana, funcionando como a subunidade regulatória G $\beta$  que, ao se desconectar da subunidade G $\alpha$ , desencadeia uma vasta rede de respostas celulares, controlando enzimas e canais iônicos essenciais para a homeostase e o desenvolvimento (Petrovski *et al.*, 2016). Variantes patogênicas da

Gnb1 foram associadas com a encefalopatia e com a síndrome de West (Nasvytis; Čiauškaitė; Jurkevičienė, 2024; Hemati *et al.*, 2018), além de ter sido apontada com papel pró-cancerígeno para os cânceres de pulmão e mama (Zhang *et al.*, 2023). As evidências sugerem que há relação do CTL com a maioria dos tumores quando a sua regulação [CTL] é positiva. Ainda, segundo estudos de carcinogenicidade, o CTL foi enquadrado como provável carcinogênico para humanos (USEPA, 1999).

A proteína tirosina fosfatase tipo F ou proteína de domínio triplo funcional (PTPRF Trio) é uma fosfatase importante que desempenha funções na regulação do crescimento celular, diferenciação (Chagnon; Uetani; Tremblay, 2004) e transformação oncogênica (Tian *et al.*, 2018). Ainda segundo Tian *et al* (2018), a baixa expressão de PTPRF está relacionada com estágios avançados de câncer gástrico, além da baixa taxa de sobrevivência dos pacientes. Ademais, um estudo demonstrou uma associação entre a superexpressão do RNA longo não codificante lncRNA TCONS\_00004099 e a menor expressão do PTPRF, o que sugere uma relação indireta entre o lncRNA TCONS\_00004099 e a PTPRF na regulação do desenvolvimento de gliomas (Wang *et al.*, 2021).

A placofilina 4 (Pkp4) desempenha papel como reguladora da atividade de proteínas da família Rho – família de pequenas proteínas G de sinalização que regulam, dentre suas funções, muitos aspectos da dinâmica da actina intracelular (Mott; Owen, 2010) durante a citocinese (The Uniprot Consortium, 2025). Este dado aponta uma provável relação entre a exposição ao CTL e alterações na dinâmica da mobilidade do citoesqueleto.

#### 5.1.3. Análise da expressão de genes mais e menos expressos no grupo clorotalonil em relação ao controle

Dois genes envolvidos na organização do citoesqueleto foram mais expressos nos neurônios expostos à CTL em relação ao controle: Map2 e o DCX. A Map2, proteína 2 associada a microtúbulos, é uma proteína específica para neurônios localizada no compartimento somatodendrítico e é conhecida pela sua propriedade de polimerização de microtúbulos (DeGiosio *et al.*, 2024). A DCX, que codifica para a proteína de migração neuronal *doublecortin*, exerce papel central na migração de células neuronais até o córtex, ajuda na estabilização e agrupamento dos microtúbulos em neurônios migratórios (Gleeson *et al.*, 1999) e mutações nesse gene são a causa mais comum de lisencefalia ligada ao cromossomo X em

humanos (Kato; Dobyns, 2003). Em conjunto, estes dados sugerem que o CTL afeta o citoesqueleto dos neurônios de camundongo em cultura de célula primária.

#### 5.1.4. Análise da expressão de genes de interesse

Embora o mecanismo de ação do clorotalonil não seja completamente compreendido, a literatura atribui seus efeitos antifúngicos à inibição da glicólise e da respiração celular, através da sua reação com enzimas de grupos tiol funcionais dessas vias, como a GAPDH – gliceraldeído-3-fosfato desidrogenase (Tillman *et al.*, 1973; Long; Siegel, 1975). Assim sendo, analisamos individualmente os genes que participam destas vias, bem como analisamos as vias de forma integrada.

A análise de vias da rota glicolítica mostra uma regulação complexa, com genes mais e menos expressos. A GAPDH foi positivamente transcrita, contrariando o esperado para a exposição ao CTL. A superexpressão de GAPDH pode ser um indicador de que a célula está passando por estresse energético e, por isso, tentando compensar através de uma maior produção de energia. Junto com o dado que aponta a inibição transcricional do complexo da piruvato desidrogenase, a análise de vias sugere que o CTL afeta a obtenção de energia a partir de glicose.

Com relação ao metabolismo mitocondrial, a análise de vias mostra perfil consistente com inibição do ciclo de Krebs e do complexo I da cadeia transportadora de elétrons nas células expostas ao CTL. Contudo, a análise de vias mostra os complexos II, IV e V ativados transcricionalmente. Esta resposta dual de ativação e inibição de vias e complexos mitocondriais pode sugerir desbalanço metabólico mitocondrial e produção de espécies reativas de oxigênio (ERO). Embora não haja dados sobre o efeito do CTL na inibição do complexo I da cadeia respiratória em modelo de sistema nervoso, o estudo de Wang *et al* (2024) apontou que a exposição de óocitos de camundongos ao CTL causou danos à função mitocondrial, levando a uma diminuição do ATP e um aumento na quantidade de ERO. Estudos mostraram que os intermediários de metabolização ditiol e tritiol do CTL inibiram o complexo II da respiração mitocondrial (Hartwig; Mak Comission, 2025). A observação da ativação transcricional da SDH do complexo II, contrariando o que é esperado segundo a literatura, sugere que a célula ativou um mecanismo compensatório. Quando a atividade da SDH é prejudicada, a célula pode responder aumentando a atividade transcricional deste complexo na tentativa de restabelecer sua função e preservar seu fluxo metabólico (Hartwig; Mak Commission, 2025).

A análise das vias do metabolismo de glutationa evidencia um desbalanço na sua síntese. Nota-se regulação positiva do gene que codifica a subunidade catalítica da glutamato-cisteína ligase (GCL), enquanto a glutationa sintetase está diminuída nas culturas expostas ao CTL em comparação ao veículo, sugerindo um desbalanço da via biossintética. Além disso, a inibição da transcrição de genes envolvidos no ciclo gama-glutamil indica um possível declínio do conteúdo intracelular de glutationa reduzida (GSH). Esse padrão é compatível com o mecanismo conhecido de biotransformação do CTL. A molécula reage com tióis de baixo peso molecular, sendo conjugada à GSH como primeiro passo do seu metabolismo em mamíferos, processo catalisado pela glutationa-S-transferase, GST, (USEPA, 1999). A conjugação intensiva somada ao desbalanço na via de síntese, contribui para a depleção de GSH, que desempenha papel central, junto à GST, na defesa antioxidante celular e no metabolismo de xenobióticos. Níveis reduzidos de glutationa diminuem a capacidade de desintoxicação de xenobióticos e favorecem o aumento do estresse oxidativo (Gallagher; Canada; Di Giulio, 1992).

## 6. CONCLUSÃO

Neste trabalho, o efeito do clorotalonil em culturas celulares do córtex pré-frontal foi determinado através de dados públicos de sequenciamento de RNA. Foram encontradas somente duas amostras de RNA-Seq para CTL no banco de dados. Adicionalmente, todas amostras de RNA-Seq deste bioprojeto variaram no conteúdo das quinze primeiras bases. Apesar destas limitações, o fluxo de trabalho 2 foi robusto o suficiente para agrupar os transcriptomas das amostras de CTL e identificar o efeito do CTL na repressão dos genes Gnb1, Pkp4 e PTPRF (Trio). Este é o primeiro trabalho que mostra que o CTL afeta genes envolvidos no citoesqueleto e sinalização celular. Além disso, os efeitos do CTL na expressão destes genes pode estar associada à carcinogenicidade deste pesticida. Com relação aos mecanismos de ação, diferentemente do esperado, a análise de vias da glicólise não mostrou inibição em enzimas-chave, como a PFK e nem da GAPDH, descrita como um alvo do CTL. Entretanto, pode-se verificar que há um desbalanço entre o ciclo de Krebs e a cadeia transportadora de elétrons, o que pode levar à produção de espécies reativas de oxigênio. Conforme o esperado, a análise da via do metabolismo da glutationa sugere comprometimento da defesa antioxidante celular. Em conjunto, este trabalho mostra que as análises de transcriptoma servem tanto para levantar hipóteses e eleger genes candidatos, quanto para entender as bases moleculares dos danos de agentes tóxicos.

## REFERÊNCIAS

- ARVANITES, Anthony C.; BOERTH, Donald W. Modeling of the mechanism of nucleophilic aromatic substitution of fungicide chlorothalonil by glutathione. *Journal of Molecular Modeling*, v. 7, p. 245-256, 2001. DOI: 10.1007/s008940100032.
- BARRETO, Juliano da Silva; TAROUCO, Fabio de Melo; ROSA, Carlos Eduardo da. Chlorothalonil causes redox state change leading to oxidative stress generation in *Danio rerio*. *Aquatic Toxicology*, v. 225, p. 105527, 2020.
- BATUT, B. *et al.* RNA-Seq Data Analysis in Galaxy. In: *Methods in Molecular Biology*. [s.l.] Humana Press Inc., 2021. v. 2284 p. 367–392.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, v. 30, n. 15, p. 2114–2120, 2014. DOI: 10.1093/bioinformatics/btu170.
- BOMBARDI, Larissa Mies. Agrotóxicos e colonialismo químico. São Paulo: Elefante, 2023.
- BRANDINE, Guilherme de Sena; SMITH, Andrew D. Falco: high-speed FastQC emulation for quality control of sequencing data. *F1000Research*, v. 8, n. 1874, 2021.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde e Ambiente. Boletim Epidemiológico, v. 54, n. 12, 2023.
- BRITTON, N. F. *et al.* RNA-seq Data Analysis: A Practical Approach. New York: [s.n.].
- CARSON, Rachel. Primavera silenciosa. 2. ed. São Paulo: Gaia, 2010.
- CHAGNON, Mélanie J.; UETANI, Noriko; TREMBLAY, Michel L. Functional significance of the LAR receptor protein tyrosine phosphatase family in development and diseases. *Biochemistry and Cell Biology*, v. 82, n. 6, p. 664–675, 2004.
- CHEN, Yun; COLONNA, Marco. Microglia in Alzheimer's disease at single-cell level: Are there common patterns in humans and mice? *Journal of Experimental Medicine*, v. 218, n. 9, e20202717, 2021.
- CORRÊA, Marcia Leopoldina Montanari et al. Agrotóxicos, saúde e ambiente: ação estratégica e políticas públicas em territórios do agronegócio. *Revista de Políticas Públicas*, v. 24, n. 1, p. 11-27, 2020.
- FAO. Pesticides use and trade – 1990–2022. FAOSTAT Analytical Briefs, n. 89, 2024.

FRENCH, John E. *et al.* Diversity outbred mice identify population-based exposure thresholds and genetic factors that influence benzene-induced genotoxicity. Environmental Health Perspectives, v. 123, n. 3, p. 237-245, 2014.

GALLAGHER, Evan P.; CANADA, Andrew T.; DI GIULIO, Richard T. The protective role of glutathione in chlorothalonil-induced toxicity to channel catfish. Aquatic Toxicology, v. 23, n. 3-4, p. 155-168, 1992.

GLEESON, Joseph G. *et al.* Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons. Neuron, v. 23, n. 2, p. 257-271, 1999.

HARTWIG, A.; MAK COMMISSION. Chlorothalonil. MAK Collection for Occupational Health and Safety, 2019/2025.

HEMATI, Parisa *et al.* Refining the phenotype associated with GNB1 mutations. American Journal of Medical Genetics A, v. 176, n. 11, 2018.

HOFF, Mariana Leivas Müller *et al.* Transcriptome analysis identifies key metabolic changes in the hooded seal. PLoS ONE, v. 12, n. 1, 2017.

IBAMA. Os 10 ingredientes ativos mais vendidos – agrotóxicos químicos: 2023. 2024.

INSTITUTO NACIONAL DE CÂNCER (Brasil). Agrotóxico. Disponível em: <https://www.gov.br/lnca/pt-br/assuntos/causas-e-prevencao-do-cancer/exposicao-no-trabalho-e-no-ambiente/agrotoxico>. Acesso em: 20 out 2025.

KATO, Mitsuhiro; DOBYNS, William B. Lissencephaly and the molecular basis of neuronal migration. Human Molecular Genetics, v. 12, 2003.

KORPELAINEN, Eija *et al.* RNA-seq data analysis: a practical approach. Boca Raton: CRC Press, 2015.

KRENAK, Ailton. Ideias para adiar o fim do mundo. 2. ed. São Paulo: Companhia das Letras, 2020.

LIU, Ruijie *et al.* Modelling sample and observational level variability improves power in RNA-seq analyses. Nucleic Acids Research, v. 43, n. 15, 2015.

LONG, J. W.; SIEGEL, M. R. Mechanism of action and fate of the fungicide chlorothalonil in biological systems. Chemico-Biological Interactions, v. 10, 1975.

LUO, Weijun; BROUWER, Cory. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics, v. 29, 2013.

MOLINARI, M. D. C. *et al.* INFLUÊNCIA DO TAMANHO DE READS E PRESENÇA DE DUPLICATAS NÃO NATURAIS EM BIBLIOTECAS DE RNA-Seq DE SOJA SOB DEFÍCIT HÍDRICO In: CONGRESSO BRASILEIRO DE SOJA, 2018.

MOTT, Helen R.; OWEN, Darerca. Structure of Rho Family Targets. In: *Handbook of Cell Signaling*, 2010.

NASVYTIS, Matas; ČIAUŠKAITĖ, Julija; JURKEVIČIENĖ, Giedrė. GNB1 Encephalopathy. *Medicina (Kaunas)*, v. 60, 2024.

NORTOX S/A. Clorotalonil Nortox: rótulo e informações técnicas. Arapongas, PR: Nortox S/A, 2024. Registro MAPA n.º 00621.

PATRO, R. *et al.* Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, v. 14, 2017.

PEARSON, B. L. *et al.* Identification of chemicals that mimic transcriptional changes associated with autism, brain aging and neurodegeneration. *Nature Communications*, v. 7, 2016.

PETROVSKI, Slavé *et al.* Germline De Novo Mutations in GNB1 Cause Severe Neurodevelopmental Disability, Hypotonia, and Seizure. *American Journal of Human Genetics*, v. 98, 2016.

ROBINSON, Mark D.; McCARTHY, Davis J.; SMYTH, Gordon K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, v. 26, 2010.

SAMIH, Anahita; FERREIRA, Maurício Alexander de Moura; NIKOLOSKI, Zoran. Gene expression and protein abundance: Just how associated are these molecular traits? *Biotechnology Advances*, v. 86, 2026.

SRIVASTAVA, A. *et al.* Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biology*, v. 20, 2019.

TIAN, Xiang'an *et al.* PTPRF as a novel tumor suppressor through deactivation of ERK1/2 signaling in gastric adenocarcinoma. *OncoTargets and Therapy*, v. 11, 2018.

TILLMAN, R. W.; SIEGEL, M. R.; LONG, J. W. Mechanism of action and fate of the fungicide chlorothalonil (2,4,5,6-tetrachloroisophthalonitrile) in biological systems: I. Reactions with cells and subcellular components of *Saccharomyces pastorianus*. *Pesticide Biochemistry and Physiology*, v. 3, 1973.

THE GALAXY COMMUNITY. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Research*, v. 52, 2024.

THE UNIPROT CONSORTIUM. UniProt: UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*, v. 53, 2025.

UNITED STATES. Environmental Protection Agency. Reregistration Eligibility Decision: Chlorothalonil. 1999.

VAN GURP, Thomas P.; MCINTYRE, Lauren M.; VERHOEVEN, Koen J. F. Consistent Errors in First Strand cDNA Due to Random Hexamer Mispriming. *PLOS One*, v. 8, 2013.

VEBER, Bruno et al. Mutagenicity of the agriculture pesticide chlorothalonil assessed by somatic mutation and recombination test in *Drosophila melanogaster*. *Environmental and Molecular Mutagenesis*, v. 65, 2024.

VINCENT, P. G.; SISLER, H. Mechanism of antifungal action of 2,4,5,6-tetrachloroisophthalonitrile. *Physiologia Plantarum*, v. 21, 1968.

WAN, Nian-Feng et al. Pesticides have negative effects on non-target organisms. *Nature Communications*, v. 16, 2025.

WANG, Yuhao et al. LncRNA TCONS\_00004099-derived microRNA regulates oncogenesis through PTPRF in gliomas. *Annals of Translational Medicine*, v. 9, 2021.

ZHANG, Xin et al. GNB1 promotes hepatocellular carcinoma progression by targeting BAG2 to activate P38/MAPK signaling. *Cancer Science*, v. 114, 2023.