

**UNIVERSIDADE DE SÃO PAULO**  
**FACULDADE DE CIÊNCIAS FARMACÊUTICAS**  
**Curso de Graduação em Farmácia-Bioquímica**

**Análise comparativa das vias de sinalização envolvidas na infecção  
por Malária**

**Isabelle Franco Moscardini**

Trabalho de Conclusão do Curso de  
Farmácia-Bioquímica da Faculdade de  
Ciências Farmacêuticas da  
Universidade de São Paulo.

Orientador(a):  
Prof. Dr. Helder I. Nakaya

São Paulo

2018

## SUMÁRIO

	Pág.
Lista de Abreviaturas .....	3
RESUMO .....	4
1. INTRODUÇÃO.....	5
2. OBJETIVOS.....	8
3. MATERIAIS E MÉTODOS.....	9
4. RESULTADOS.....	15
5. DISCUSSÃO.....	24
6. CONCLUSÃO.....	33
7. BIBLIOGRAFIA.....	34
8. ANEXOS.....	38

## LISTA DE ABREVIATURAS

GEO	<i>Gene Expression Omnibus</i>
PCA	Análise de Componentes Principais
DEGs	Genes Diferencialmente Expressos
CSBL	Laboratório de Biologia de Sistemas Computacionais
UP	Genes Positivamente Regulados
DOWN	Genes Negativamente Regulados
IFN	Interferon
TLR	Toll-Like Receptors

## RESUMO

MOSCARDINI, IF. **Análise Comparativa das vias de sinalização envolvidas na infecção por Malária**. 2018. Trabalho de Conclusão de Curso de Farmácia-Bioquímica – Faculdade de Ciências Farmacêuticas – Universidade de São Paulo, São Paulo, 2018.

Palavras-chave: *Plasmodium falciparum*, Transcriptômica, expressão gênica diferencial

**INTRODUÇÃO:** A malária é uma doença potencialmente fatal, causada por parasitas do gênero *Plasmodium*. Mesmo após o progresso obtido pela Organização Mundial de Saúde no combate à malária até o final da década de 70, ela atualmente é um dos mais sérios problemas mundiais de saúde pública. A infecção por *Plasmodium* leva a diversas mudanças no organismo, muitas estão relacionadas a expressão gênica do indivíduo. Essas mudanças podem ser melhor estudadas com o auxílio dos métodos como o Microarray e de ferramentas de bioinformática.

**OBJETIVO** Identificar as alterações transcricionais na infecção por malária a partir de três estudos diferentes para obter um melhor entendimento da reação do organismo frente à infecção, além de identificar os principais genes e vias envolvidas na infecção.

**MATERIAIS E MÉTODOS:** Esse trabalho analisou dados de transcriptômica de três estudos já publicados, realizamos uma meta-análise de cinco comparações entre diferentes grupos de infectados com grupos controle. Para o pré-processamento destes dados, utilizamos alguns pacotes desenvolvidos para o Software R. Para a análise destes dados, foram utilizadas diversas ferramentas de bioinformática.

**RESULTADOS:** O presente estudo encontrou os genes diferencialmente expressos consistentes entre as cinco comparações realizadas. A partir deste resultado, encontramos as vias biológicas enriquecidas durante a infecção, assim como os dez principais genes, tanto positivamente quanto negativamente regulados. Encontramos uma forte assinatura para o sistema imune entre as vias reguladas positivamente, principalmente vias relacionadas a citocinas pró inflamatórias e aos receptores do tipo *Toll-Like*, bem descritos na literatura associada ao sistema imune inato e infecções por malária. Entre as vias reguladas negativamente, encontramos o ciclo celular e o reparo de DNA, havendo poucos estudos sobre a perturbação destas vias na infecção por malária.

**CONCLUSÃO:** O presente trabalho identificou diversos genes consistentemente perturbados principais genes durante a infecção por malária, assim como as vias enriquecidas, relacionando a diversos aspectos já descritos em literatura, mas também identificando algumas diferenças pontuais e indicando possíveis caminhos a serem melhor estudados.

## 1. INTRODUÇÃO

A malária é uma doença potencialmente fatal, causada por parasitas do gênero *Plasmodium*. Em 2016, os dados da Organização Pan-Americana da saúde sugeriram que cerca de 3,2 bilhões de pessoas ao redor do mundo correm risco de serem infectadas pela doença. Nas Américas, aproximadamente 128 milhões de pessoas vivem em áreas de risco de contrair a doença (OPAS - WHO 2016).

Mesmo após o progresso obtido pela Organização Mundial de Saúde no combate à malária até o final da década de 70, ela atualmente é um dos mais sérios problemas mundiais de saúde pública. Mesmo com o avanço, encontra-se barreiras para sua total erradicação, sendo uma das principais dificuldades a crescente resistência aos antimaláricos e aos repelentes contra os mosquitos do gênero *Anopheles*, vetores da doença (World Health Organization, 2000–2010 e Nájera et al, 2011).

Os dados do *World Malaria Report* de 2017 estimaram 216 milhões de casos de malária em 2016, com o número de mortes atingindo os 445.000. Se comparado com o ano de 2015, em 2016 houve um aumento de cerca de 5 milhões de casos.

Em uma escala global, a região que concentra maior número de casos de malária é o continente africano, sobretudo a região subsaariana (Tauil, 2006). Alguns dos estudos de expressão gênica escolhidos para embasar este projeto provêm do oeste deste continente.

No continente americano, o Brasil é o país que apresenta o maior número de casos, com cerca de 300.000 casos anuais (Oliveira et al; 2010). Dessa forma, estudos e políticas públicas envolvendo a malária devem receber atenção uma vez que, além comprometer a saúde da população, essa doença causa danos ao nível social e econômico, representando um grande impacto para a receita de países endêmicos e uma carga importante nos serviços de saúde (Cutler et al., 2010).

Dado o cenário de danos em diversos níveis causados pela doença e da resistência crescente aos tratamentos atuais, observa-se o aumento de estudos relacionados ao tema, assim como o aumento de investimento por parte de fundações como a *Bill & Melinda Gates*, que tem voltado seus esforços para a erradicação da malária. Até o momento, a fundação investiu quase US \$ 2 bilhões em subsídios para combater a doença, além de defender o financiamento sustentado e crescente por parte de governos doadores e dos países endêmicos (The Bill & Melinda Gates foundation website). Segundo a *World Health Organization* o total de fundos investidos no controle e eliminação da malária chegou a 2,7 bilhões de dólares em 2016, incluindo contribuições dos governos de países endêmicos.

O parasita mais prevalente no continente africano é o *Plasmodium falciparum*, enquanto que o *Plasmodium vivax* é mais comum fora da região subsaariana (WHO, 2017). A infecção por *Plasmodium falciparum* é responsável pela maior parte das infecções com sintomas mais severos, podendo levar a quadros de anemia, trombocitopenia, hipoglicemia e acidose metabólica (Hu, 2016).

O complexo ciclo de vida do *Plasmodium*, ilustrado na figura 1, depende da expressão de inúmeras proteínas especializadas do hospedeiro, as quais determinam sua sobrevivência tanto no ambiente intracelular quanto extracelular, assim como a invasão de vários tipos celulares e a evasão das respostas imunológicas (Florens et al; 2002 e Doolan et al; 2008).

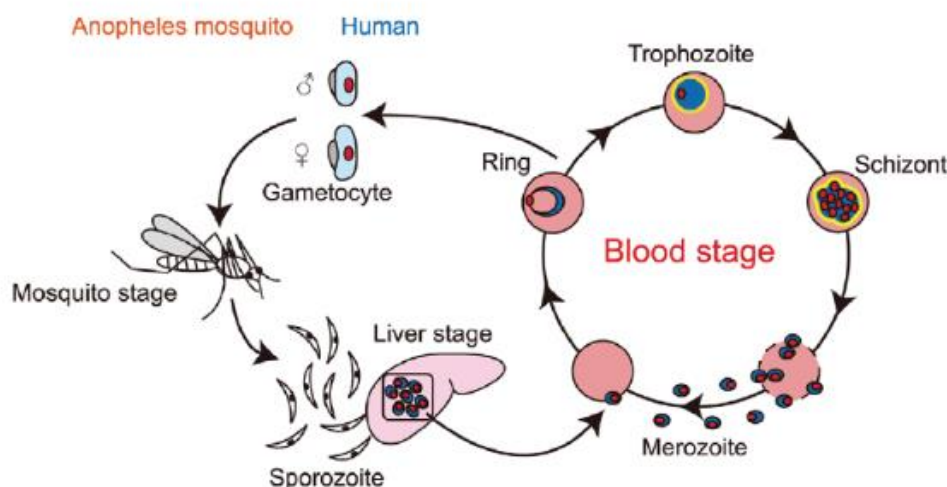


Figura 1: ilustração do ciclo de vida do *P. falciparum*. A infecção em humanos se inicia com a inoculação de esporozoítos presentes nas glândulas salivares de mosquitos *Anopheles* fêmeas. Os esporozoítos infectam as células hepáticas, onde se transformam e se reproduzem assexuadamente, liberando milhares de merozoítos, que ao infectar eritrócitos, iniciam o estágio sanguíneo, produzindo gametas ou entrando em uma nova fase de reprodução assexuada, liberando novos merozoítos (Enomoto et al, 2012 e Haldar, 2007).

A infecção por *Plasmodium* leva a diversas mudanças no organismo. Muitas destas estão relacionadas a alterações na síntese proteica, dependendo da expressão gênica do indivíduo. Essas mudanças ocasionadas nos organismos infectados podem ser avaliadas através de análises de expressão gênica, como por exemplo, análises de Microarray. As tecnologias de expressão gênica permitem obter um panorama da atividade transcricional em diferentes tecidos ou populações de células. Estes perfis são então comparados para identificar alterações de expressão gênica associadas a uma condição de interesse. Esses experimentos fornecem *insights* sobre os processos celulares normais, bem como a patogênese da doença estudada. (Matthew et al, 2015)

*Microarray* possibilita uma análise comparativa entre os diferentes genes expressados presentes no ser humano, sendo possível observar quais tiveram sua expressão aumentada ou diminuída após a infecção por malária. Os estudos utilizados no presente projeto comparam de forma similar, mas com algumas

particularidades, a expressão gênica entre pessoas infectadas por malária e pessoas saudáveis a partir de análises realizadas por esta técnica.

Observamos hoje um grande número de estudos gerando dados de expressão gênica tanto em relação à malária como também outras doenças. Entretanto, a análise de comparação entre esses estudos ainda não é amplamente realizada. Apesar disso, esta análise é de grande importância, pois a partir dela pode-se avaliar de forma mais consistente os genes envolvidos em cada patologia, sendo possível o conhecimento das principais vias biológicas envolvidas na infecção.

Conhecendo minuciosamente o mecanismo fisiopatológico, vias biológicas e genes envolvidos, criam-se as bases para novas pesquisas, incluindo estudos ao nível de fármacos para o tratamento da malária, cuja necessidade se dá pelo rápido crescimento da resistência aos fármacos já existentes (Hastings et al, 2005).

Desta forma, o presente estudo teve como objetivo realizar a análise funcional de três estudos de expressão gênica diferencial utilizando diversas ferramentas no intuito de gerar maiores conhecimentos sobre os detalhes da fisiopatologia e das redes biológicas e moleculares envolvidas na infecção por malária.

## **2. OBJETIVO(S)**

### **Objetivo geral:**

Identificar as alterações transcricionais na infecção por malária a partir de três estudos diferentes para obter um melhor entendimento da reação do organismo frente à infecção.

### **Objetivos específicos**

- Identificar os genes diferencialmente expressos entre indivíduos infectados com malária e indivíduos saudáveis
- Identificar genes consistentemente alterados em todas as comparações
- Identificar processos biológicos relacionados à infecção por malária



### 3. MATERIAIS E MÉTODOS

#### 3.1. Levantamento de Estudos e critérios de inclusão e exclusão

O levantamento de estudos de expressão gênica de infecção de malária foi realizado no banco de dados *Gene Expression Omnibus* (GEO), um repositório público internacional que cura, armazena e distribui dados de *microarray*, sequenciamentos de última geração (RNA-Seq) e outras formas de dados de transcriptoma (Clough, 2016).

#### 3.2. Pré Processamento

##### 3.2.1. Escolha dos estudos

Neste trabalho, foram utilizados três estudos de *microarray* contendo dados de expressão gênica de indivíduos infectados por *Plasmodium falciparum*. O estudo GSE5418 (Ockenhouse e Vahey, 2006) contém 71 amostras de células mononucleares periféricas do sangue (PBMC), tanto de pacientes saudáveis, quanto de pacientes naturalmente infectados com malária, além de voluntários desafiados com mosquitos infectados por *P. falciparum*.

Já o estudo GSE34404 (Idaghdour Y et al, 2012) agrega o total de 155 amostras de sangue de crianças do oeste da África, contendo o perfil de expressão gênica e o genótipo de cada indivíduo. Este estudo possui 3 grupos principais: controle, sintomáticos de baixa parasitemia e sintomáticos de alta parasitemia. Todos os indivíduos foram testados com o teste rápido para diagnóstico de malária. Para a realização das análises, os indivíduos foram divididos em apenas dois grupos, o grupo controle (indivíduos cujo o resultado do teste foi negativo – “Negatif”) e o grupo infectado por malária (indivíduos cujo o resultado do teste foi positivo – “Positif”).

O estudo GSE35858 (Popper S; 2016) foi realizado entre 37 indivíduos vietnamitas, incluindo indivíduos admitidos no hospital com infecção por *Plasmodium falciparum* que apresentaram ou não complicações, assim como indivíduos controle, ou seja, não infectados. Para nossa análise, foram removidas as amostras pertencentes ao grupo “*complicated malaria*”, devido às variáveis clínicas não estarem corretamente distribuídas entre os lotes experimentais (GenePix 4000A e GenePix 4000B), o que poderia ser um fator de confusão durante a análise e, possivelmente, diminuir a especificidade durante a detecção de genes diferencialmente expressos entre os grupos.

De forma geral, realizou-se comparações entre indivíduos infectados e grupo controle, como demonstrado na tabela 1. A segunda coluna da tabela 1 (Estudo) contém o identificador (ID) dos estudos utilizados.

Comparação	Estudo	Grupo Perturbado (GP)	Num. amostras (GP)	Grupo Controle (GC)	Num. amostras (GC)
Malaria X Controle	GSE34404	Positif	93	Negatif	62
Malária Não Complicada X Controle	GSE35858	Uncomplicated Malaria	9	Healthy	8
Malária Aguda X Controle	GSE5418	Acute Malaria	15	Baseline Uninfected	22
Malária Experimental X Controle	GSE5418	Experimental Malaria	22	Baseline Uninfected	22
Malária Tratada X Controle	GSE5418	Treated Malaria	12	Baseline Uninfected	22

*Tabela 1: Comparações realizadas para cada estudo GSE. Para cada comparação, as colunas 3 e 5 identificam quais grupos de cada estudo GSE foram utilizados na análise. As colunas 4 e 6 identificam o número de amostras do respectivo grupo.*

Plataforma	Empresa	Título/Nome	Num. de Probes	Estudo
GPL15240	Stanford Functional Genomics Facility	SMD Print_1162 lc-51	38784	GSE35858
GPL10558	Illumina Inc	Illumina HumanHT-12 V4.0 expression beadchip	48107	GSE34404
GPL96	Affymetrix	Affymetrix Human Genome U133A Array	22283	GSE5418

*Tabela 2: Tecnologia empregada nas plataformas de Microarray utilizadas em cada estudo GSE*

### 3.2.2. Aquisição de dados e Avaliação de Qualidade

Para evitar que a presença de ruído nos estudos influenciasse negativamente as análises, foi realizado um pré-processamento. Para isso, foram desenvolvidos *scripts* na linguagem de programação R, em que foram utilizados pacotes do repositório *Bioconductor* (HUBER et al; 2015), como *GEOquery* (Davis; 2007), *affy* (Gautier; 2004) e *arrayQualityMetrics* (Kauffman; 2009)

O *Bioconductor* é um repositório *open source* que disponibiliza ferramentas para uma análise compreensiva de dados genômicos em larga escala. Para avaliar a qualidade das amostras, foi utilizando o pacote *arrayQualityMetrics*, que fornece um relatório de diagnóstico para dados de *microarray* (e.g. distância entre amostras, *boxplots*, gráficos de densidade, *PCA*, entre outros). Além disso, o pacote identifica possíveis *outliers*, amostras com valores aberrantes em relação à população observada (Kauffmann et al, 2009). Todas as amostras que falharam em pelo menos dois testes foram caracterizadas como *outliers* e removidas.

Avaliou-se a qualidade de todos os estudos. Após a avaliação, todos os estudos tiveram, ao menos, uma amostra removida.

### 3.2.3. Análise de Componentes Principais

A análise de componentes principais (PCA) consiste em uma técnica de redução de dimensionalidade, em que as coordenadas das amostras dos estudos são reformuladas em outro sistema de eixo, denominados de componentes principais. É um método versátil e conveniente para análise de dados de alta dimensionalidade, capaz de fornecer uma visão geral de dados multivariados, o que permite melhor explorar a variabilidade existente nos dados e detectar amostras outliers (Bro, R. and Smilde, A.K, 2014). Esta análise faz parte do relatório gerado pelo *arrayQualityMetrics* e foi aplicado à todos os estudos.

### 3.2.4. Normalização dos dados

Após o controle de qualidade, os dados brutos foram transformados em escala logarítmica e normalizados por quantil. Para os estudos GSE35858 e GSE34404,

a normalização foi feita com a função *normalize.quantiles* do pacote *preprocessCore* (Bolstad, 2018), que consiste em uma coleção de funções para pré-processamento de dados de *microarray*.

Já para o estudo GSE5418, utilizou-se a função *rma* da biblioteca *affy*, uma biblioteca dedicada ao pré-processamento de tecnologias Affymetrix. A função *rma* consiste em três passos, sendo o primeiro deles a correção de fundo, caracterizada pela remoção de possíveis artefatos técnicos e ruídos, como por exemplo, a interferência luminosa de sondas situadas nas proximidades. O segundo passo é a normalização em si, removendo efeitos específicos de cada *microarray* e tornando-os comparáveis. Por fim, a função *rma* combina as intensidades de cada amostra para que as medidas finais representem os níveis de expressão.

Após a etapa de normalização, a análise de qualidade dos estudos foi realizada novamente para uma segunda verificação da presença de *outliers*. Nesta etapa não foram encontradas amostras *outliers* em nenhum dos três estudos.

### 3.3. Análise de genes diferencialmente expressos

Na busca de genes relacionados à infecção por malária foram analisados, primeiramente, os genes diferencialmente expressos (DEGs) para cada uma das comparações listadas anteriormente na Tabela 1. Para tal, utilizamos o pacote *limma* do R, desenvolvido para análises de diferença de expressão em desenhos experimentais complexos, através de modelos lineares. (Ritchie, M.E *et al*; 2015). Dessa forma, obteve-se uma tabela de cada comparação contendo valores de *fold-change*, que representa o quanto um dado gene está aumentado ou diminuído em relação ao controle, e seus respectivos p-valores. Para definir se um gene estava diferencialmente expresso, utilizamos um limiar de *fold-change* absoluto maior que 1.25 e p-valor ajustado menor que 0.01.

Sendo um dos objetivos do presente estudo encontrar dados de expressão consistentes, foi realizada uma busca de genes diferencialmente expressos em todas as cinco comparações realizadas. Para tal utilizamos uma ferramenta em fase de desenvolvimento no Laboratório de Biologia de Sistemas Computacionais (CSBL) – o *Metavolcano*. Esta ferramenta está sendo desenvolvida na linguagem R, e permite combinar, de uma forma estatisticamente coerente, diversos resultados de análises de diferenças de expressão, tais como resultados obtidos com o *limma* ou GEO2R.

### **3.4. Análise de redes**

#### **3.4.1. *Network Analyst***

Após encontrar os genes diferencialmente expressos presentes nas cinco comparações, foram realizadas análises de redes com o objetivo de encontrar vias gênicas possivelmente enriquecidas na presença de infecção por malária. Para tanto, utilizou-se o *NetworkAnalyst*, uma ferramenta *web* para análises de dados ômicos. Ao combinar procedimentos estatísticos bem estabelecidos com técnicas de visualização de dados de última geração, o *NetworkAnalyst* permite que grandes conjuntos de dados complexos de expressão gênica sejam analisados para identificar características, padrões, funções e conexões importantes, levando à geração de novos recursos biológicos (Xia J *et al*, 2015).

Listas de DEGs separados por sua direção (UP – positivamente regulados, DOWN – negativamente regulados) foram fornecidas à ferramenta, e criou-se uma rede de genes combinados com genes membros de redes de proteína-proteína (PPI), provenientes do banco de dados *IMEX Interactome* (Orchard *et al*, 2012). Para a construção da rede, o *NetworkAnalyst* recomenda que se mantenha a rede entre 200 e 2000 nós (genes), uma vez que redes maiores são difíceis de navegar e interpretar, enquanto redes menores não oferecem compreensão a um nível de sistema. Dessa forma, para os DEGs UP escolhemos a *Minimum Network* que é adequada para simplificar uma rede densa a fim de estudar os principais

conectivos e, para os DEGs DOWN, por haver menos genes nesta condição, utilizamos a *First order Network*.

### 3.4.2. GeneMANIA

GeneMANIA (<http://genemania.org>) é uma ferramenta autônoma que realiza previsões da função de genes de forma rápida e eficiente. O GeneMANIA utiliza um algoritmo (Mostafavi et al., 2008) que tem a capacidade de derivar previsões a partir de uma combinação de fontes de dados potencialmente heterogêneas. Isso permite que os usuários construam uma rede de interação funcional gene-gene composta a partir de uma lista de genes. A rede resultante inclui os genes mais relacionados à lista original, assim como suas anotações funcionais do Gene Ontology (Ashburner et al; 2000). O banco de dados do GeneMANIA que possui mais de 1800 redes, contendo mais de 500 milhões de interações que abrangem diversos organismos, entre eles o *H. sapiens*, organismo de estudo do presente projeto (Montejo et al, 2014). Através do GeneMANIA os dados foram preparados para serem analisados no Gephi.

### 3.4.3. Gephi

Apesar das análises de enriquecimento de vias biológicas terem sido realizadas na ferramenta *NetworkAnalyst*, a partir do banco de dados do Reactome, os grafos das redes enriquecidas foram construídos através do *software* Gephi. O Gephi é um *software* para análise de grafos e redes que utiliza um mecanismo de renderização 3D para exibir redes massivas em tempo real e utilizando conjuntos de dados complexos para produzir resultados para explorações visuais (Bastian et al, 2009).

## 3.5. Análise de Genes pelo GeneCards

GeneCards é uma ferramenta que permite a visualização e interação com o amplo banco de dados de genes, doenças, variantes, proteínas, células e vias biológicas humanas (Stelzer et al; 2016). Esta ferramenta foi utilizada para entender as

funções e vias relacionadas dos principais genes identificados como positivamente e negativamente regulados.

## 4. RESULTADOS

### 4.1. Controle de Qualidade

O resultado do controle de qualidade dos estudos, realizado através do pacote *arrayQualityMetrics*, é composto de diversos gráficos de diagnóstico, como os de PCA, ilustrados nas as figuras 1, 2 e 3. Esses gráficos mostram a variabilidade existente entre as amostras de cada estudo, de forma que, quanto mais próximos dois pontos (amostras) estiverem entre si, menor a variabilidade existente entre eles. Partindo deste princípio, uma amostra que se encontra muito distante das outras é um possível *outlier*. A partir deste e de outros resultados obtidos com o controle de qualidade, foram escolhidas as amostras a serem excluídas da análise.

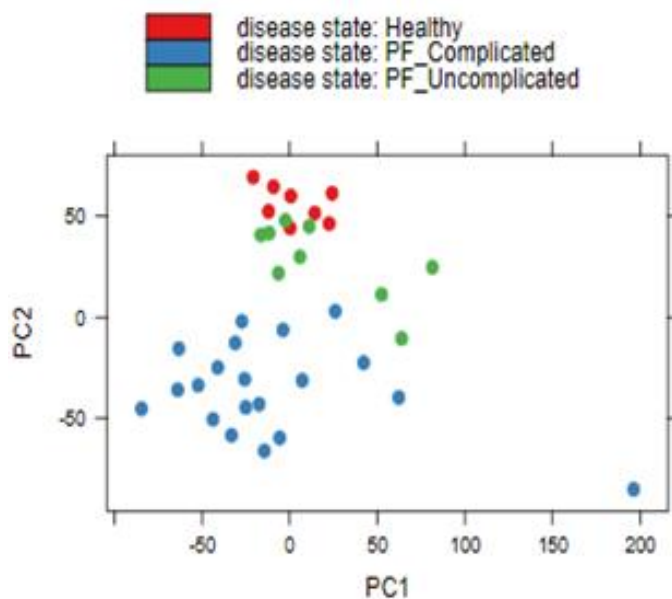


Figura 2: Gráfico PCA do estudo GSE35858. Os pontos representam as amostras do estudo. Em vermelho temos os indivíduos do grupo controle, em azul os indivíduos com malária complicada e em verde, com malária não complicada.

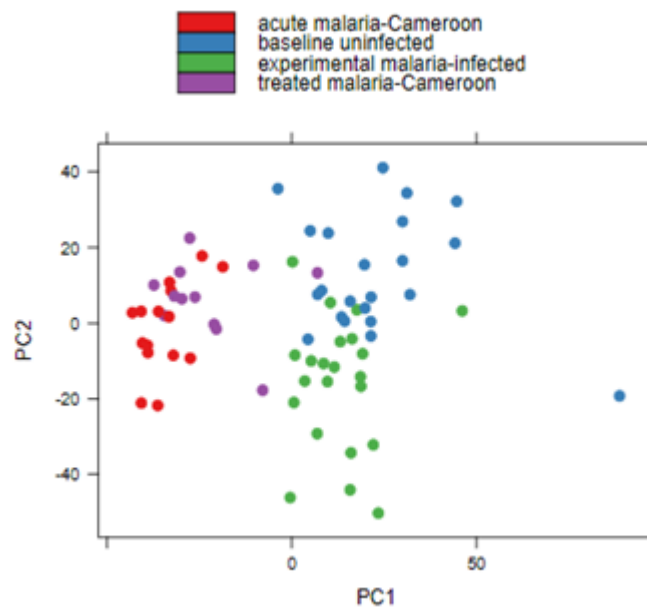


Figura 3: Gráfico PCA do estudo GSE5418. Os pontos no mapa representam as amostras do estudo, classificadas segundo a cor. Os pontos azuis representam os indivíduos controle, os verdes os indivíduos com malária experimental, os vermelhos representam os indivíduos com malária aguda e os roxos, indivíduos convalescentes

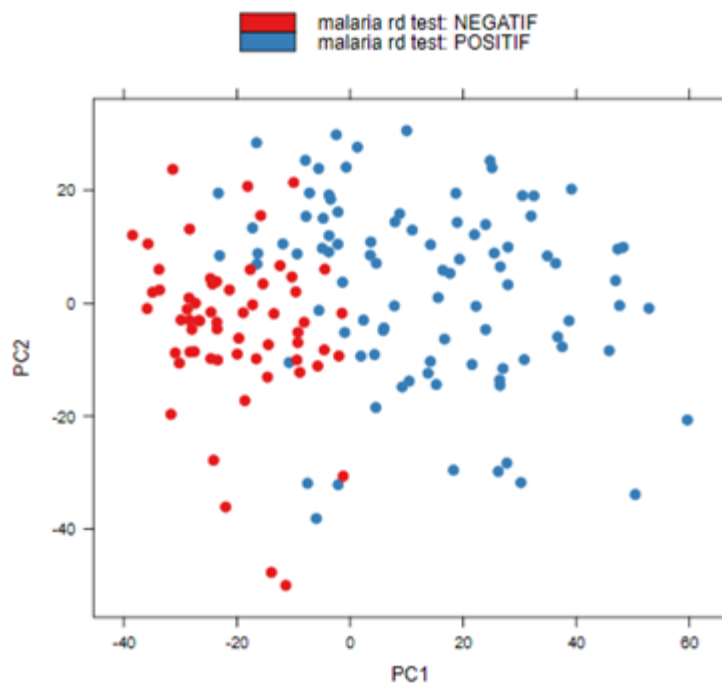


Figura 4: Gráfico PCA do estudo GSE34404. Os pontos no mapa representam as amostras, sendo que os pontos em vermelho representam os indivíduos controle e, em azul, os indivíduos infectados com malária.



## 4.2. Normalização

Após a normalização constuiu-se graficos da densidade dos valores de expressão (figura 4) para observar a qualidade da normalização e garantir a comparabilidade entre as amostras.

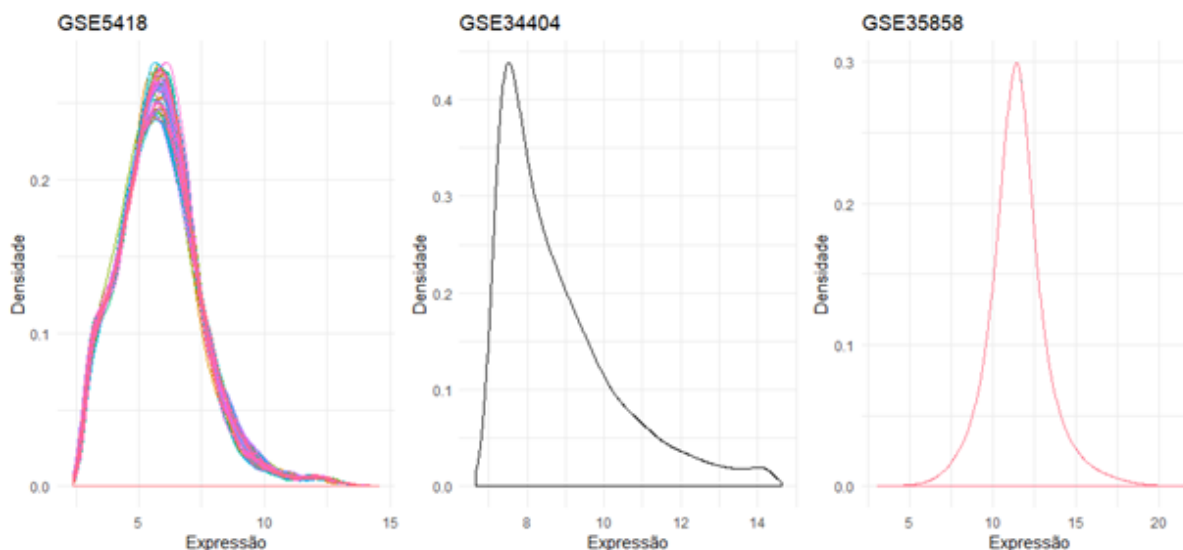


Figura 5: Resultado da Normalização dos dados dos três estudos utilizados

## 4.3. Genes Diferencialmente Expressos

Após a realizar a análise de diferença de expressão através do pacote *limma*, identificou-se os DEGs escolhendo-se valores de p-valor ajustado menores que 0.01 e de *fold-change* maiores que 1.25 para identificar os genes positivamente regulados (UP) e os valores de p-valor ajustado menores que 0.01 e de *fold-change* menores que - 1.25 para os genes negativamente regulados na presença de infecção por malária (DOWN).

## 4.4. Identificação de DEGs Consistentes

Para encontrar os DEGs consistentes, ou seja, que apareceram em todas as comparações, foi realizada a análise por Metavolcano, em que não há um estabelecimento prévio de limites de valores de p-valor e *fold-change* para cada comparação. Neste caso, pode-se ter uma visão geral da diferença de expressão nos estudos, não se excluindo possíveis DEGs relevantes e que poderiam não ser

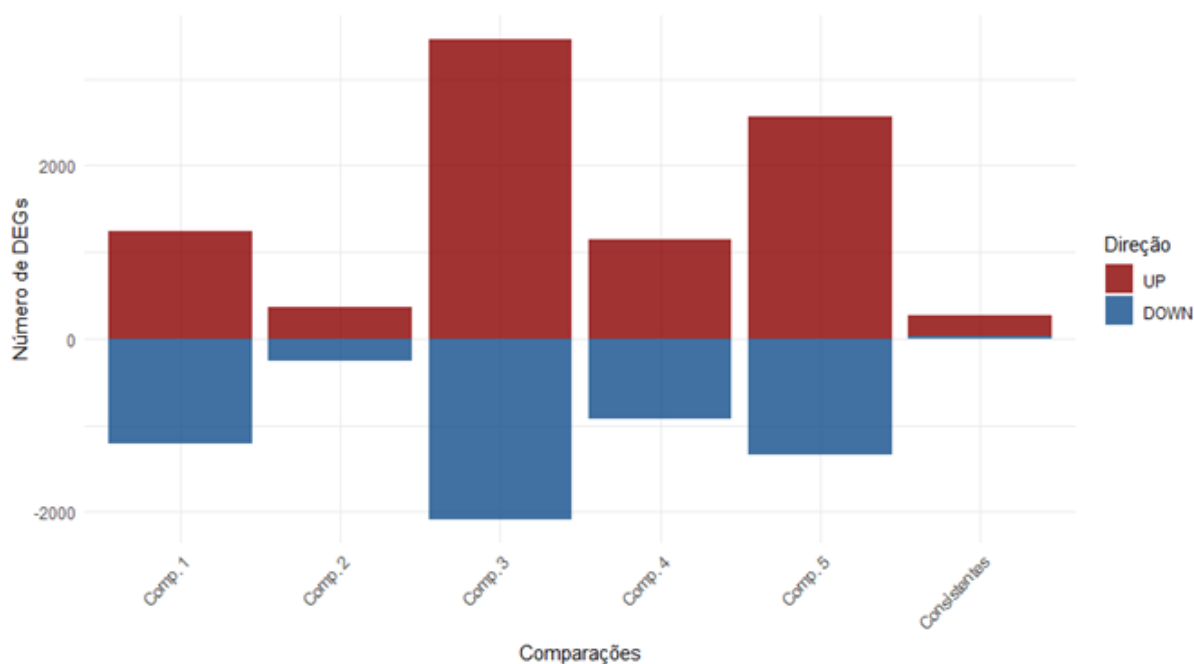
identificados em algum dos estudos por conta dos limites previamente estabelecidos.

Estes resultados estão demonstrados através da Tabela 3 e Figura 2. Na tabela é possível identificar o número de DEGs em cada comparação, relacionando às comparações da tabela 1. A partir da Figura 2 é possível visualizar essa diferença no número de DEGs encontrados em cada uma das comparações.

Através da análise por Metavolcano obtivemos como resultado uma tabela cujas colunas representam, para cada gene identificado, o seu respectivo p-valor e *fold-change*, além de definir o dircon para cada gene. O dircon é uma métrica que representa a quantidade de vezes que um gene foi identificado como diferencialmente expresso na análise, levando em consideração a sua direção. Em nosso caso, como foram realizadas cinco comparações, o dircon variou de -5 a 5, sendo que o primeiro representa um gene que esteve negativamente regulado nas cinco comparações e o segundo representa um gene que tenha sido identificado como positivamente regulado nas cinco comparações. Os genes que possuíam o Dircon igual a 5 ou -5 foram classificados como DEGs consistentes.

	<b>Comparações</b>	<b>UP</b>	<b>DOWN</b>
<b>Comp. 1</b>	Malária Aguda X Controle	3454	2090
<b>Comp. 2</b>	Malária Experimental X Controle	1144	920
<b>Comp. 3</b>	Malária Não Complicada X Controle	366	253
<b>Comp. 4</b>	Malária X Controle	1241	1210
<b>Comp. 5</b>	Malária Tratada X Controle	2561	1341
<b>-</b>	Consistentes	231	33

*Tabela 3: Número de DEGs UP e DOWN para cada comparação e número total de DEGs UP e DOWN consistentes - Genes identificados como DEGs em todas as comparações.*



*Figura 6: Representação do número de genes UP e DOWN encontrados em cada comparação e, por último, à direita, o número total de DEGs consistentes de acordo com a direção*

Além de encontrar os DEGs consistentes, esta análise permitiu encontrar os principais genes consistentemente envolvidos na infecção por malária. Uma vez selecionados apenas os DEGs consistentes, selecionamos tanto para um valor de dircon igual a 5 quanto para -5 os primeiros dez genes com o melhor score.

O score representa a multiplicação do log do p-valor por -1 e em seguida multiplicado pelo valor do *fold-change*. Dessa forma, obtém-se um valor positivo que, levando em consideração ambas as medidas, representa os genes mais provavelmente relacionados à infecção por malária, como descrito nas tabelas 4 e 5 e na figura 7 abaixo.

Positivamente Regulados		
Gene_ID	Fold-Change	P-valor
CNIH4	0.7040949	9,62E-30
SOCS3	0.7944026	5,75E-14
TYMS	0.6425880	2,73E-17
VAMP3	0.5355982	9,97E-21
JAK2	0.7138037	1,10E-11
IL10RB	0.4943331	1,20E-18
LITAF	0.6093132	3,59E-09
TFE3	0.3564524	9,52E-19
AHCY	0.3882118	4,08E-15
PFKP	0.4076223	1,67E-13

Tabela 4: Top10 genes regulados positivamente e seus respectivos fold-change e p-valores

Negativamente Regulados		
Gene_ID	Fold-Change	P-valor
LRRN3	-1.09991574	2,62E-12
PAFAH1B1	-0.34685176	1,10E-23
NUPL2	-0.30309847	1,45E-21
TCL1A	-0.86534384	0,003126
TUG1	-0.27977026	2,67E-16
NDE1	-0.37454709	5,44E-10
LTB	-0.50908911	1,71E-05
EXOSC2	-0.25817981	1,57E-14
IL11RA	-0.49398744	0,015242
AMT	-0.25200133	1,88E-09

Tabela 5: Top10 genes regulados negativamente e seus respectivos fold-change e p-valores

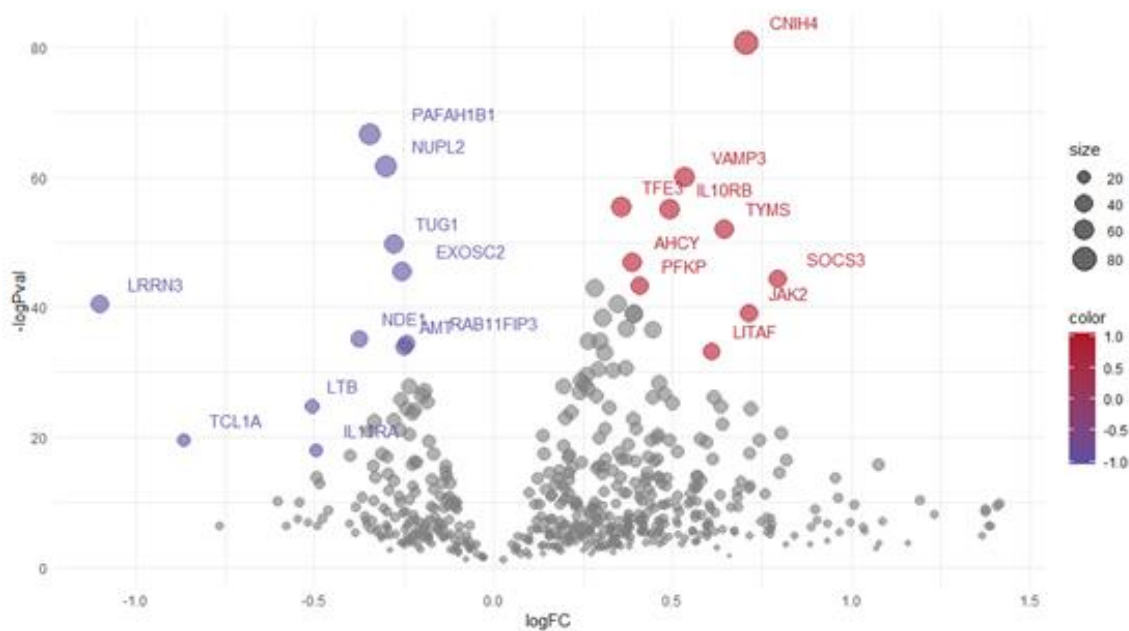


Figura 7: Volcano Plot ilustrando os 'Top 10' genes positivamente e negativamente regulados. Estes genes representam os genes consistentes que possivelmente estão mais relacionados à infecção por malária

#### 4.5. Análise de redes

Após a realização da análise de rede com a utilização da ferramenta *Gephi*, obtivemos como resultado uma rede de conexões entre genes regulados positivamente e uma rede de conexões entre genes regulados negativamente, como nas figuras 8 e 9 abaixo.



Figura 8: Análise de redes reguladas positivamente pelo Gephi. A região central da rede foi ampliada a fim de facilitar a observação dos nós (genes). Alguns nós não participavam de redes e foram excluídos. O tamanho dos nós é proporcional ao valor de *eigenvector centrality*, que mensura a importância do nó na rede. As cores representam a estrutura da comunidade representada, segundo o *Louvain modularity method*. A espessura das arestas está representada segundo o peso.

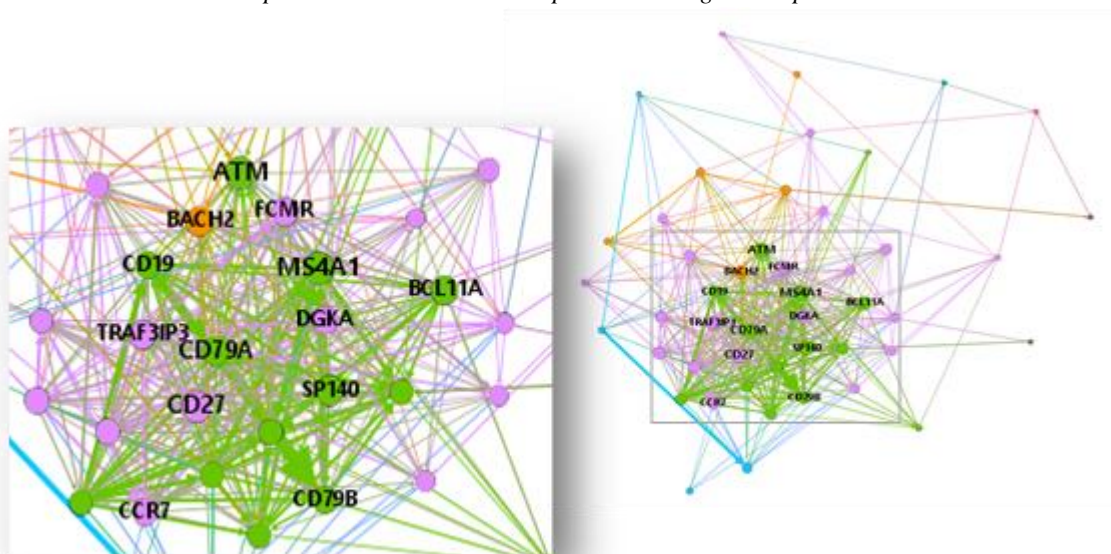


Figura 9: Análise de redes reguladas negativamente pelo Gephi. A região central da rede foi ampliada a fim de facilitar a observação dos nós (genes). Alguns nós não participavam de redes e foram excluídos. Os parâmetros de cor, espessura das arestas e tamanho dos nós seguiu o mesmo raciocínio das vias reguladas positivamente.

É importante ressaltar que a construção de rede realizada pelo Gephi é direcional, ou seja, nos casos aplicáveis, observam-se as setas indicando a direção da interação entre os nós. Outro ponto relevante é o tamanho do nó, calculado com base na medida da centralidade de autovetor (*eigenvector centrality*) (Smith e Timberlake, 2001). Esta medida, proposta por Bonacich (1972), considera no cálculo da centralidade de determinado elemento a importância dos elementos a ele conectados.

As cores representam as comunidades segundo o método *Louvain*, um método simples e eficiente para identificar comunidades em grandes redes. O método revela hierarquias de comunidades e permite ampliar comunidades para descobrir sub-estruturas altamente conexas. (Blondel, Vincent D, 2008).

## **4.6. Análise de Enriquecimento**

### **4.6.1 Reactome**

Através da ferramenta *NetworkAnalyst* foi realizado o enriquecimento de vias biológicas utilizando o banco de dados *Reactome* e, como resultado, obtivemos uma tabela com as principais vias enriquecidas, o número total de genes já descritos como pertencentes ou potencialmente pertencentes a esta via e o número de genes encontrados na análise que estão relacionados a esta via. Além destes três dados, a tabela também informa os respectivos p-valores e p-valores ajustados (FDR). A partir desta tabela, obtivemos as Figuras 10 e 11, que ilustram as 20 principais vias enriquecidas, tanto para os genes regulados negativamente quanto para os genes regulados positivamente.

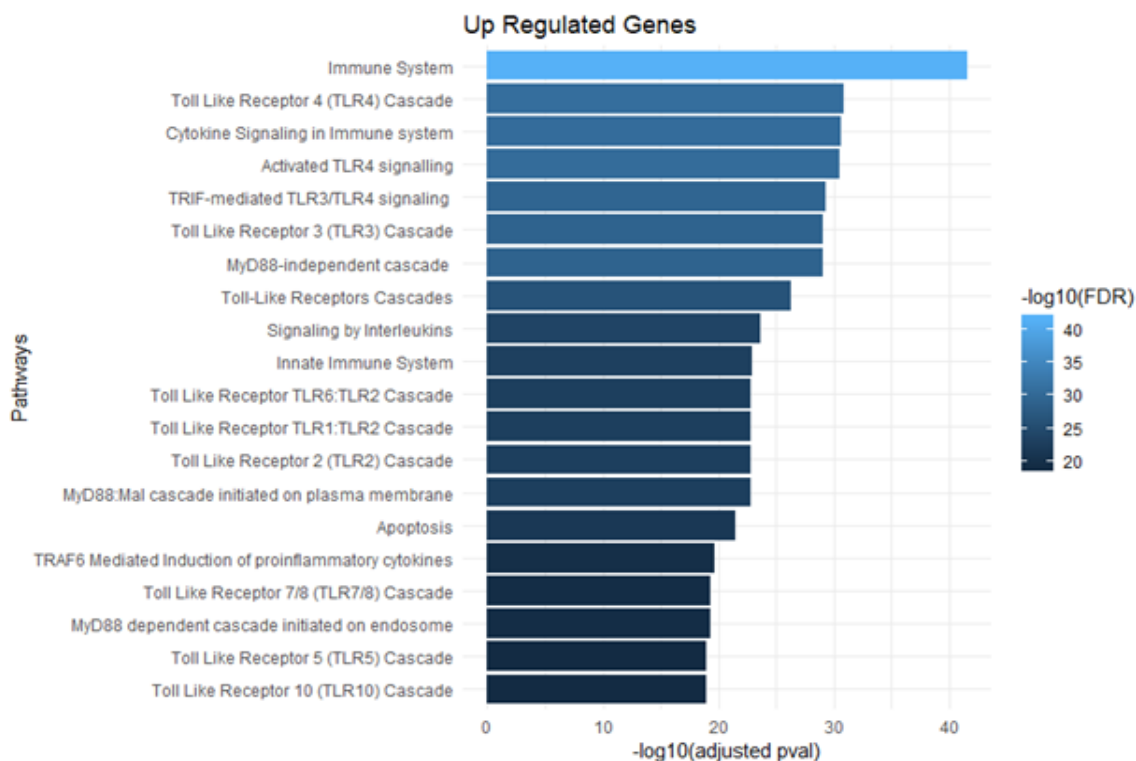


Figura 10: Resultado do enriquecimento das vias reguladas positivamente. A imagem ilustra as 20 principais vias enriquecidas segundo o P-valor ajustado (FDR)

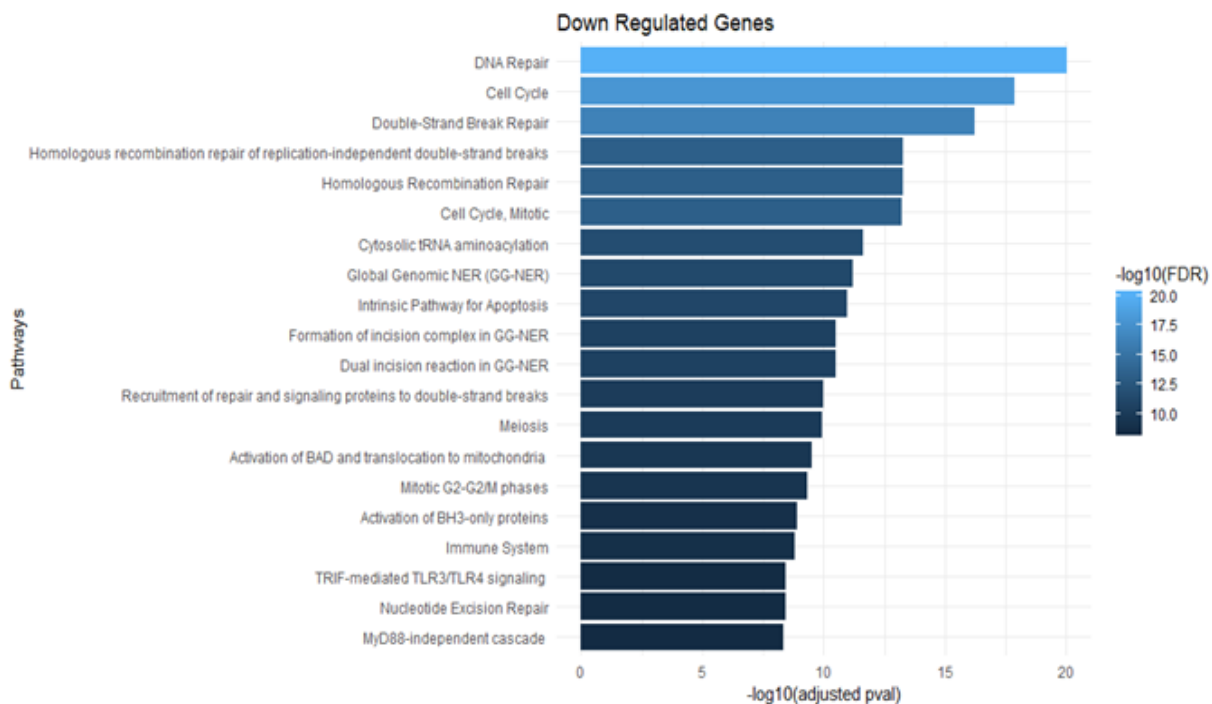


Figura 11: Resultado do enriquecimento das vias reguladas negativamente. A imagem ilustra as 20 principais vias enriquecidas segundo o P-valor ajustado (FDR)

## 5. DISCUSSÃO

Neste estudo identificamos genes consistentemente perturbados na infecção por malária, e para entender os processos biológicos envolvidos, realizou-se o enriquecimento de vias para os DEGs consistentes e em seguida buscamos as funções biológicas de alguns dos DEGs consistentes mais perturbados na plataforma Genecards (Stelzer et al, 2016).

As vias mais enriquecidas para os genes que estavam negativamente regulados indicaram uma perturbação no processo de divisão celular e mecanismos de reparo de DNA, mais especificamente reparo de quebra da dupla fita. Isso está em total acordo com os 10 genes negativamente regulados mais perturbados, no qual cinco deles participam diretamente de mecanismos do ciclo celular. Por exemplo, os genes PAFAH1B1 e NDE1 (Chen et al, 2017 e Wynne et al, 2018) estão relacionados com as diversas funcionalidades dos microtúbulos, entre eles a duplicação do centrôssoma, o acoplamento do núcleo ao centrôssoma e a formação e função do fuso mitótico. Ou seja, possuem o papel de regular a biogênese e a manutenção do ciclo celular, da mitose e das organelas.

O gene NUPL2 é responsável pela exportação de mRNA do núcleo para o citoplasma, estando envolvido em vias como o ciclo celular e mitose. O gene TCL1A está relacionado com o processo de proliferação celular, estabilizando o potencial da membrana mitocondrial e promovendo a sobrevivência celular, através da fosforilação e ativação de AKT1, AKT2 e AKT3 (Hsi et al, 2014). O gene TUG1 é traduzido em um RNA não codificador, que age na regulação epigenética da transcrição. Esse RNA promove a proliferação celular e, assim como o gene TCL1A, está regulado positivamente em células tumorais (Zhang et al, 2018).

Além dos genes relacionados ao ciclo celular, encontramos entre os principais genes regulados negativamente o gene EXOSC2 que codifica um componente não catalítico do complexo exossomo de RNA que participa de inúmeros eventos



de processamento e degradação de RNA celular, especificamente mRNAs intrinsecamente instáveis. (Morton et al, 2018) O gene EXOSC8, envolvido na mesma via, aparece entre os genes consistentes negativamente regulados.

Também foram identificados dois genes dentre com outras funções além do controle do ciclo celular e reparo de DNA, mas relacionados à atividade de citocinas: o LTB, relacionado à atividade de NF-kappaB e o IL11RA codifica o receptor da citocina IL-11, membro dos receptores de citocinas hematopoiéticas.

Por fim, o gene AMT codifica uma proteína envolvida na via de Clivagem da glicina, um processo relacionado com a via do metabolismo do Folato – Vitamina B12 - Metionina. Esta via já foi descrita como alterada na presença de infecção por malária, sendo uma das possíveis razões, a utilização de metionina pelo parasita (Brabin et al, 2003).

Um estudo (Achtman et al, 2012) analisou dados de Microarray de camundongos e demonstrou uma regulação positiva da via de reparo de DNA nos animais tratados com a substância anti-inflamatória IDR-1018, se comparado aos tratados com apenas solução salina. O mesmo estudo identificou que camundongos tratados com IDR-1018 obtiveram melhores resultados frente a infecção, uma vez que cerca de 44% deles sobreviveram, enquanto todos os tratados apenas com a solução salina morreram após desenvolverem malária cerebral. Uma hipótese possível seria o envolvimento do reparo de DNA, encontrado negativamente regulado em nosso estudo, com a inflamação, de forma que a atividade de uma substância anti-inflamatória pudesse reverter o sentido desta via e levar a uma possível proteção frente a infecção pelo *Plasmodium berghei*.

Encontrou-se poucos estudos relacionando o ciclo celular e a infecção por malária na literatura. O estudo realizado por Asghar et al, em 2017, envolvendo ambos os temas, sugere o aumento da expressão de CDKN2A - inibidor de kinase dependente de ciclina 2<sup>a</sup>, já descrito como um inibidor do ciclo celular – na fase

aguda da infecção. O estudo realizado em humanos indica que durante a infecção aguda por *P. falciparum* a dinâmica do envelhecimento celular é afetada, associando este fato a uma possível diminuição dos telômeros em células sanguíneas periféricas.

Outros estudos associam o tamanho dos telômeros a infecções, relacionando o fato observado ao estresse celular e danos oxidativos resultantes tanto da ativação do sistema imune inato quanto adaptativo (Ilmonen and Kotrschal, 2008). Espera-se que as células imunológicas sejam particularmente vulneráveis ao encurtamento dos telômeros e à senescência acelerada, devido à sua rápida proliferação e alta produção de espécies reativas de oxigênio pelos fagócitos durante as respostas inflamatórias (Vallejo et al 2004).

As vias reguladas positivamente estão relacionadas principalmente a modulação do sistema imune durante a infecção. A assinatura para os Toll-Like Receptors (TLR) mais forte parece estar relacionada ao TLR 4, mas entre as vinte primeiras vias positivamente reguladas, notamos também assinaturas de TLR 1, TLR2, TLR 3, TLR5, TLR6, TLR7, TLR8 e TLR10, indicando que a infecção por malária parece ter influência sobre diversos receptores

Os TLRs são uma família de proteínas transmembrana que reconhecem padrões moleculares associados a patógenos. São extremamente importantes em respostas imunes inatas. Há cerca de dez TLRs funcionais já descritos em humanos (Corr e O'Neill, 2009), dentre eles, a ativação ou mudanças na expressão de TLR2, TLR4 e TLR9 parecem ter uma influência importante no desfecho da doença (Esposito et al., 2012 e Franklin et al., 2009).

Certos componentes imunologicamente ativos derivados de parasitas, foram descritas como mediadores das respostas inflamatórias via TLR2 e, em menor grau, via TLR4 (Krishnegowda et al; 2005).

---

Além disso, já foi descrito que alguns TLR são influenciados pela Hemozoína, subproduto do metabolismo do parasita, que estimula o TLR9 (Parroche et al., 2007), e, sua associação com o fibrinogênio hospedeiro, também ativa TLR4 em monócitos humanos (Barrera et al., 2011).

Outra via que se mostrou positivamente regulada foi a sinalização por Myd88, esta via já foi descrita como relacionada a infecção por malária, por estar intimamente ligada à atividade das citocinas pró-inflamatórias. A ativação de TLR é mediada por uma série de proteínas adaptadoras, entre elas a MyD88, o que leva a expressão de genes pró-inflamatórios (Pasare et al; 2005). Na malária aguda, a associação de TLR9 com MyD88 é essencial para o início de uma resposta imune mediada pela interleucina 12 (IL-12) / interferon gama (IFN- $\gamma$ ) o que eleva ainda mais a expressão de citocinas pró-inflamatórias (Franklin et al; 2009).

No estudo de Barboza e colaboradores, em 2014, o aumento de citocinas pró-inflamatórias foi dependente da expressão de MyD88, sugerindo que a ausência dessa proteína pode ter um efeito negativo sobre a resposta imune. Em outro estudo, Franklin e colaboradores mostraram que o MyD88 é essencial para induzir a produção de TNF- $\alpha$  e IFN- $\gamma$  durante a infecção por *P. chabaudi* (Poovassery et al; 2009). Além disso, esses autores mostraram que, na ausência da expressão do MyD88, os camundongos exibiam uma patologia atenuada.

Afim de melhor entendermos os mecanismos subjacentes às vias reguladas positivamente, analisou-se os 10 dos genes mais perturbados entre os genes regulados positivamente através da plataforma GeneCards (Stelzer et al; 2016). Estes genes apresentam vias relacionadas, principalmente, a atividade do sistema imune durante a infecção. O gene IL10RB, por exemplo, codifica uma proteína pertencente à família de receptores de citocinas, que é uma cadeia acessória essencial para a ativação do complexo receptor da interleucina 10 (IL-10).

---

Os genes JAK2 e SOCS3 estão envolvidos com a sinalização de citocinas. Enquanto o gene JAK2 codifica uma proteína quinase necessária para a resposta ao interferon gama, ativando a transdução de sinal pela fosforilação de proteínas STAT, o gene SOCS3 faz parte da via de feedback negativo desta sinalização, codificando uma proteína capaz de se ligar à proteína JAK2 quinase, inibindo a sua atividade. O gene SOCS3 tem sua ativação induzida justamente por citocinas como IL-10 e interferon (Trilling et al, 2014)

Ainda em relação a sinalização de citocinas o gene LITAF codifica o fator TNF alfa induzido por lipopolissacarídeos. A transcrição deste gene é induzida pelo supressor tumoral p53, sendo também implicado na via apoptótica induzida por p53. Além da inflamação, o LITAF tem sido identificado como um potencial gene supressor de tumor, devido ao fato de sua expressão poder ser induzida pela p53 (Polyak et al; 1997)

O gene LITAF também foi identificado como um parceiro funcional de STAT6, ambos são ativados por LPS e associam-se ao receptor toll-like-2/4 para formar um complexo dependente de MyD88, e é fosforilado por p38- $\alpha$ . O LITAF fosforilado e o STAT6 (B) consequentemente interagem para formar um complexo proteico antes da translocação para o núcleo, onde o LITAF se liga especificamente à sequência promotora, ativando assim a expressão de genes a jusante, como TNF- $\alpha$  e IL-6 (Tang et al; 2005). LITAF é altamente expresso em macrófagos em vários tecidos inflamatórios agudos (Locksley et al; 2001)

Os genes da família VAMP também foram identificados como consistente, entre eles VAMP3, VAMP5 e VAMP7, sendo que o VAMP3 apareceu entre os top10 genes regulados positivamente. Esses genes são membros da família da proteína de membrana associada a vesículas, entre suas vias relacionadas está a apresentação de antígenos mediados pelo MHC de classe I. (Nair-Gupta et al, 2014). Na literatura há muitos dados sobre a atividade deste gene no parasita

*Plasmodium falciparum* durante a infecção, mas pouco se sabe sobre sua ativação nos seres humanos.

Genes relacionados ao metabolismo também foram encontrados entre os principais genes regulados positivamente. O gene PFKP codifica a isoforma específica de plaquetas da enzima fosfofrutoquinase, possuindo um papel importante na regulação da glicólise (Mor et al., 2011). Esta via já foi descrita como ligada à infecção por malária, em 2016 o estudo de microarray realizado por Wan-Chung comparou pacientes infectados com não infectados e identificou a via glicolítica como positivamente regulada. Entre os genes regulados positivamente o estudo identificou o gene ALDOA e PGAM1, que também apareceram no presente estudo como consistentes positivamente regulados. Para Wan-Chung os resultados apoiam uma relação da acidose metabólica e hipoglicemia (comumente observados na infecção por malária) com uma regulação positiva geral dos genes que codificam uma série de enzimas glicolíticas, transportadores de glicose e H<sup>+</sup> - ATPases.

Nosso estudo também identificou os genes SLC25A46, SLC2A3, SLC35B1, SLC38A2, SLC3A2 e SLC7A5 como consistentes. Estes genes estão relacionados ao transporte de substâncias, entre elas a glicose, sendo que todos, com exceção dos genes SLC25A46 e SLC38A2, também foram descritos como positivamente regulados no estudo de Hu, WC.

O gene TYMS também foi identificado nesta categoria, este gene codifica a enzima Timidilato sintetase, cuja função mantém o pool de dTMP (monofosfato de timidina-5-primo) crítico para replicação e reparo de DNA. As interações proteína-proteína deste gene incluem interações com o gene AMT, encontrado no presente estudo como um dos principais genes negativamente regulados.



Também relacionado ao metabolismo da metionina, identificamos o gene AHCY entre os principais genes regulados positivamente. Responsável por regular a concentração de S-adenosylhomocysteine, importante para reações de transmetilação, processo que faz parte da biotransformação da metionina,

Como já era esperado os principais genes regulados positivamente estão relacionados a forte indução do sistema imune frente a infecção por *Plasmodium falciparum*. Mas além destes processos, encontramos também alterações em vias do metabolismo, tanto de açúcares como de aminoácidos, vias que também já haviam sido identificadas como relacionadas à malária e que, o presente estudo contribui para o melhor conhecimento dos possíveis genes envolvidos nestas vias.

Além disso, realizamos uma comparação dos resultados encontrados neste trabalho com os resultados encontrados outros estudos de transcriptoma na infecção por malária, incluindo os estudos utilizados como base, com exceção do estudo GSE35858, que não possui nenhuma publicação vinculada.

No estudo GSE34404, os autores identificaram DEGs entre pacientes infectados com malária e controles não infectados. O enriquecimento de vias para tais genes mostrou uma forte assinatura de imunidade inata induzida (regulação positiva de genes induzíveis por IFN, módulos associados a neutrófilos e marcadores de fagocitose). No entanto, encontrou-se vários processos negativamente regulados do sistema imune adaptativo, tais como os genes do MHC, células T, células B e vias de sinalização de células T citotóxicas.

O estudo GSE5418 identificou diversos genes relacionados ao *Interferon gama* regulados positivamente após a infecção por malária, o que é consistente com o modelo no qual a infecção por malária induz principalmente uma resposta imune do tipo TH1. Estes genes, além de *interferon gama*, incluíram proteína de ligação ao guanilato induzível por *interferon 1* (GBP1), proteína 2 de ligação a guanilato induzível por *interferon* (GBP2), Janus quinase1 (JAK1) e transportador com processamento de antígeno 1 (TAP1).

O presente trabalho também encontrou uma forte assinatura de regulação positiva da via de Interferon. Como particularidades, o estudo GSE5418 encontrou o gene JAK1 positivamente regulado, já esse presente estudo encontrou um outro gene da mesma família, denominado JAK2 (fold change 0,7138 e p-valor 0,0011). Entre as diferenças já descritas entre os dois genes, a principal está relacionada à via de sinalização, em que o JAK2 está relacionado com IFN $\gamma$  (Griffiths et al; 2005) e o JAK1 com o IFN $\alpha$  /  $\beta$  e IFN $\gamma$  (Ho et al; 1998).

Outros componentes importantes da resposta imune TH $\alpha$  $\beta$  incluem regulação positiva de FAS, BCL6, TNFRSF10A e MIP1 $\alpha$ , resultado também observado na pesquisa de Griffiths em 2005, que estudou os leucócitos do sangue periférico de crianças quenianas com malária aguda por *P. falciparum*. Em nosso estudo, os genes BCL6 e FAS foram encontrados em todas as comparações positivamente regulados, com p-valor significativo e *fold-change* elevado. O gene TNFRSF10A não foi classificado como consistente, porém, o gene TNFRSF4, membro da superfamília dos receptores de TNF, foi encontrado de forma consistente e com o valor de *fold-change* elevado, além de um p-valor significativo.

Portanto, o estudo encontrou diversos resultados já descritos em literatura, mas também complementou estas informações com novas informações de outros genes que podem estar envolvidos na infecção. Além disso, o presente estudo encontrou assinaturas para vias ainda pouco descritas como relacionadas com a infecção por malária, como por exemplo, o ciclo celular e o reparo de DNA.



## **6. CONCLUSÃO (ÕES)**

O presente trabalho cumpriu com os objetivos propostos de encontrar os principais genes e vias relacionadas com a infecção por malária. Dentre os resultados, o estudo encontrou diversos aspectos já descritos em literatura, o que é de suma importância para a confiabilidade dos resultados, mas também identificou algumas diferenças pontuais e indicou possíveis caminhos a serem melhor estudados.

É importante ressaltar que, falando em números totais, o presente estudo realizou uma comparação de 139 indivíduos infectados com 92 indivíduos controles, dados obtidos de três estudos e cinco comparações diferentes. Dentre os indivíduos infectados encontramos voluntários de partes diferentes do continente africano, Vietnã e Estados Unidos e que possuíam quadros clínicos muito diversificados: infecções de alta parasitemia, infecções de baixa parasitemia, infecções experimentais, indivíduos tratados e malária não complicada. Dessa forma, nossos resultados compreendem processos que se apresentam consistentes, relevando muitos aspectos relacionados a possíveis variações genéticas.

## 7. BIBLIOGRAFIA

Achtman AH, Pilat S, Law CW, Lynn DJ, Janot L, Mayer ML, Ma S, Kindrachuk J, Finlay BB, Brinkman FS, Smyth GK, Hancock RE, Schofield L. Effective adjunctive therapy by an innate defense regulatory peptide in a preclinical model of severe malaria. **Science Translational Medicine**, 2012 23;4(135):135ra64

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. **Nature Genetics**. 2000; 25(1):25-9

Asghar M, Yman V, Homann MV, Sondén K, Hammar U, Hasselquist D, Färnert A. Cellular aging dynamics after acute malaria infection: A 12-month longitudinal study. **Aging Cell**; 2018

Bastian M., Heymann S., Jacomy M. Gephi: an open source software for exploring and manipulating networks. **International AAAI Conference on Weblogs and Social Media**. 2009

Blondel, Vincent D; Guillaume, Jean-Loup; Lambiotte, Renaud; Lefebvre, Etienne "Fast unfolding of communities in large networks". **Journal of Statistical Mechanics: Theory and Experiment**. 2008; (10): P10008

Bonacich, P. Factoring and Weighting Approaches to Status Scores and Clique Identification. **Journal of Mathematical Sociology**. 1972; 2, 113-120.

Bolstad B. *preprocessCore*: A collection of pre-processing functions. **R package version 1.42.0** 2018

Brabin BJ, Fletcher KA, Brown N. Do disturbances within the folate pathway contribute to low birth weight in malaria? **Trends in Parasitology**. 2003; 19(1):39-43

Bro, R. and Smilde, A.K. Principal component analysis. **Analytical Methods**. 2014; 6, 2812

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. **BMC Bioinformatics**. 2013; 128(14).

Chen J, Cai Z, Zhang L, Yin Y, Chen X, Chen C, Zhang Y, Zhai S, Long X, Liu X, Wang X. Lis1 Regulates Germinal Center B Cell Antigen Acquisition and Affinity Maturation. **Journal Immunology**. 2017; 1;198(11):4304-4311

Clough E, Barrett T. The Gene Expression Omnibus Database. **Methods in Molecular Biology**. 2016; 1418:93-110.

Corr SC, O'Neill LA. Genetic variation in Toll-like receptor signalling and the risk of inflammatory and immune diseases. **Journal of Innate Immunity**. 2009; 1(4):350-7

Cutler, David; Winnie Fung; Michael Kremer; Monica Singhal e Tom Vogl. "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." **American Economic Journal: Applied Economics**, 2010, 2(2): 72-94.

Davis S, Meltzer P. "GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor." **Bioinformatics**. 2007; 14, 1846–1847.

Doolan DL, Mu Y, Unal B, Sundaresh S, Hirst S, Valdez C, et al. Profiling humoral immune responses to *P. falciparum* infection with protein microarrays. **Proteomics**. 2008; 8(22):4680-94.

Enomoto, M., Kawazu, S., Kawai, S., Furuyama, W., Ikegami, T., Watanabe, J., & Mikoshiba, K. Blockage of Spontaneous Ca<sup>2+</sup> Oscillation Causes Cell Death in Intraerythrocytic Plasmodium falciparum. **PLoS ONE**. 2012, 7(7), e39499

Esposito S, Molteni CG, Zampiero A, Baggi E, Lavizzari A, Semino M, Daleno C, Groppo M, Scala A, Terranova L, Miozzo M, Pelucchi C, Principi N. Role of polymorphisms of toll-like receptor (TLR) 4, TLR9, toll-interleukin 1 receptor domain containing adaptor protein (TIRAP) and FCGR2A genes in malaria susceptibility and severity in Burundian children. **Malaria Journal**. 2012 12;11:196

Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, Matthews L, May B, Milacic M, Rothfels K, Shamovsky V, Webber M, Weiser J, Williams M, Wu G, Stein L, Hermjakob H, D'Eustachio P. **Nucleic Acids Research**. 2016; 44(D1):D481-7.

Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, et al. A proteomic view of the *Plasmodium falciparum* life cycle. **Nature**. 2002; 419(6906):520-6.

Franklin BS, Parroche P, Ataíde MA, Lauw F, Ropert C, de Oliveira RB, Pereira D, Tada MS, Nogueira P, da Silva LHP, Bjorkbacka H, Golenbock DT, Gazzinelli RT. Malaria primes the innate immune response due to interferon-gamma induced enhancement of toll-like receptor expression and function. **Proceeds of National. Academy of Science U. S. A**. 2009; 106:5789–5794

Gautier L, Cope L, Bolstad BM, Irizarry RA. "affyanalysis of Affymetrix GeneChip data at the probe level." **Bioinformatics**. 2004; 20(3), 307–315

Global report on Antimalarial Drug efficacy and Drug Resistance: **World Health Organization**. 2000–2010

Griffiths, M. J., M. J. Shafi, S. J. Popper, C. A. Hemingway, M. M. Kortok, A. Wathen, K. A. Rockett, R. Mott, M. Levin, C. R. Newton, K. Marsh, D. A. Relman, and D. P. Kwiatkowski. Genomewide analysis of the host response to malaria in Kenyan children. **Journal of Infection Disease**. 2005. 191:1599-1611

Haldar K, Mohandas N. Erythrocyte remodeling by malaria parasites. **Current Opinion in Hematology**. 2007; 14: 203–209

Hastings IM, Watkins WM. Intensity of malaria transmission and the evolution of drug resistance. **Acta Tropical**. 2005; 94(3):218-29

Hsi AC, Robirds DH, Luo J, Kreisel FH, Frater JL, Nguyen TT. T-cell prolymphocytic leukemia frequently shows cutaneous involvement and is associated with gains of MYC, loss of ATM, and TCL1A rearrangement. **American Journal of Surgery Pathology**. 2014;38(11):1468-83

Ho, M., T. Schollaardt, S. Snape, S. Looareesuwan, P. Suntharasamai, and N. J. White. Endogenous interleukin-10 modulates proinflammatory response in Plasmodium falciparum malaria. **Journal of Infection Disease**. 1998 178:520-525

Hu WC. Microarray analysis of PBMC after Plasmodium falciparum infection: Molecular insights into disease pathogenesis. **Asian Pacific Journal of Tropical Medicine**. 2016; 9(4):313-323

Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. Orchestrating high-throughput genomic analysis with Bioconductor. **Nature Methods**. 2015 Feb;12(2):115-21.

Idaghdour Y, Quinlan J, Goulet JP, Berghout J, Gbeha E, Bruat V, de Malliard T, Grenier JC, Gomez S, Gros P, Rahimy MC, Sanni A, Awadalla P. Evidence for additive and interaction effects of host genotype and infection in malaria. **Proceeds of the National Academy of Sciences USA**. 2012; 109(42):16786-93.

Ilmonen P, Kotrschal A, Penn DJ. Telomere attrition due to infection. **PLoS One**. 2008 14;3(5):e2143

J.A. Nájera , M. González-Silva , P.L. Alonso , Some lessons for the future from the Global Malaria Eradication Programme (1955–1969),. **PLoS Medicine**. 2011; 8.

Kanehisa M and Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes; **Nucleic Acids Research**. 2000; 28(1): 27–30.

Kauffmann A, Gentleman R, Huber W. “arrayQualityMetrics—a bioconductor package for quality assessment of microarray data.” **Bioinformatics**, 2009; 25(3), 415–6.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. **Nucleic Acids Research**. 2016

Montojo J, Zuberi k, Rodriguez H, Bader HD, Morris Q; GeneMANIA: Fast gene network construction and function prediction for Cytoscape; **F1000 Research**. 2014; 1,3:153

I. Mor, E.C. Cheung, K.H. Vousden Control of glycolysis through regulation of PFK1: old friends and recent additions. Cold Spring Harbor **Symposia on Quantitative Biology**. 2011; 76 pp. 211-216

Morton DJ, Kuiper EG, Jones SK, Leung SW, Corbett AH, Fasken MB. The RNA exosome and RNA exosome-linked disease. **RNA**. 2018; 24(2):127-142

Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. **Genome Biology**. 2008; 9 Suppl 1:S4

Nair-Gupta P, Baccarini A, Tung N, Seyffer F, Florey O, Huang Y, Banerjee M, Overholtzer M, Roche PA, Tampé R, Brown BD, Amsen D, Whiteheart SW, Blander JM. TLR signals induce phagosomal MHC-I delivery from the endosomal recycling compartment to allow cross presentation. **Cell**. 2014; 158(3):506-21

Nosten F., Brasseur P.; **Drugs**. 2002; 62:1315

Ockenhouse CF, Hu WC, Kester KE, Cummings JF, Stewart A, Heppner DG, Jedlicka AE, Scott AL, Wolfe ND, Vahey M, Burke DS. Common and divergent immune response signaling pathways discovered in peripheral blood mononuclear cell gene expression patterns in presymptomatic and clinically apparent malaria. **Infection and Immunity**. 2006; 74(10):5561-73

Oliveira-Ferreira J, Lacerda MV, Brasil P, Ladislau JL, Tauil PL, Daniel-Ribeiro CT. Malaria in Brazil: an overview. **Malaria Journal**. 2010; 9:115

Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bidwell S, Bridge A, Briganti L, Brinkman FS, Cesareni G, Chatr-aryamontri A, Chautard E, Chen C, Dumousseau M, Goll J, Hancock RE, Hannick LI, Jurisica I, Khadake J, Lynn DJ, Mahadevan U, Perfetto L, Raghunath A, Ricard-Blum S, Roechert B, Salwinski L, Stümpflen V, Tyers M, Uetz P, Xenarios I, Hermjakob H. Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. **Nature Methods**. 2012 Apr;9(4):345-50.

Organização Mundial da Saúde, Representação da Organização Pan-Americana da saúde no Brasil Website. 2016

Pasare C, Medzhitov R. Toll-like receptors: linking innate and adaptive immunity. **Advances in Experimental Medicine and Biology**. 2005; 560:11–18

Polyak K, Xia Y, Zweier JL, Kinzler KW and Vogelstein B: A model for p53-induced apoptosis. **Nature**. 1997; 389:300–305

Poovassery JS, Sarr D, Smith G, Nagy T, Moore JM. Malaria-induced murine pregnancy failure: distinct roles for IFN-gamma and TNF. **Journal of Immunology**. 2009

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. **Nucleic Acids Research**. 2015 Apr 20;43(7):e47.

Smith, D. A.; Timberlake, M. F. World city networks and hierarchies: 1977–1997, **American Behavioral Scientist**. 2001; 44(10), 1656-1678.

Stelzer G, Rosen R, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Iny Stein T, Nudel R, Lieder I, Mazor Y, Kaplan S, Dahary D, Warshawsky D, Guan - Golan Y, Kohn A, Rappaport N, Safran M, and Lancet D. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analysis , **Current Protocols in Bioinformatics**. 2016; 54:1.30.1 - 1.30.33

Tang X, Marciano DL, Leeman SE and Amar S: LPS induces the interaction of a transcription factor, LPS-induced TNF-alpha factor and STAT6 (B) with effects on multiple cytokines. **Proceeds of the National Academy of Sciences USA**. 2005; 102:5132–5137

Targett G., Drakeley, C., Jawara, M.; **The Journal of Infectious Diseases**. 2001; 183:1254

Tauil PL. Perspectivas de controle de doenças transmitidas por vetores no Brasil. **Revista da Sociedade Brasileira de Medicina Tropical**. 2006; 39(3):275-7.

The Bill & Melinda Gates Foundation Website

Trilling M, Le VT, Rashidi-Alavijeh J, Katschinski B, Scheller J, Rose-John S, Androsiac GE, Jonjic S, Poli V, Pfeffer K, Hengel H. "Activated" STAT proteins: a paradoxical consequence of inhibited JAK-STAT signaling in cytomegalovirus-infected cells. **Journal of Immunology**..2014;192(1):447-58

Vallejo AN, Weyand CM, Goronzy JJ. T-cell senescence: a culprit of immune abnormalities in chronic inflammation and persistent infection. **Trends in Molecular Medicine**. 2004 10: 119–124

WHO. Severe falciparum malaria. World Health Organization, Communicable Diseases Cluster. **Transactions of the Royal Society of Tropical Medicine and Hygiene**. 2000; 94 Suppl 1:S1-90. Review

Wynne CL, Vallee RB. Cdk1 phosphorylation of the dynein adapter Nde1 controls cargo binding from G2 to anaphase. **Journal of Cell Biology**. 2018; 3;217(9):3019-3029

World Health Organization. World Malaria Report. **Geneva: WHO**. 2017

Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. **Nature Protocols**. 2015; 10(6):823-44

Zhang Z, Wang X, Cao S, Han X, Wang Z, Zhao X, Liu X, Li G, Pan X, Lei D. The Long Noncoding RNA TUG1 Promotes Laryngeal Cancer Proliferation and Migration. **Cell Physiology Biochemistry**. 2018; 27;49(6):2511-2520

## 8. ANEXOS

Não Aplicável

28/09/2018




---

Data e assinatura do aluno(a)

28/09/2018




---

Data e assinatura do orientador(a)