

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
CURSO DE GRADUAÇÃO EM BIOMEDICINA

RODRIGO SILVA ARAUJO STREIT

**ANÁLISE DE FRAGMENTOS DE RNA DERIVADOS DE tRNA EM *Cryptococcus*
*gattii***

Porto Alegre

2017

RODRIGO SILVA ARAUJO STREIT

**ANÁLISE DE FRAGMENTOS DE RNA DERIVADOS DE tRNA EM *Cryptococcus*
*gattii***

Trabalho de Conclusão de Curso de Graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel em Biomedicina.

Área de habilitação: Biologia Molecular

Orientador: Prof. Dr. Charley Christian Staats

Porto Alegre

2017

CIP - Catalogação na Publicação

Streit, Rodrigo Silva Araujo
Análise de fragmentos de RNA derivados de tRNA em
Cryptococcus gattii / Rodrigo Silva Araujo Streit. --
2017.
66 f.
Orientador: Charley Christian Staats.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Ciências Básicas da Saúde, Curso de Biomedicina,
Porto Alegre, BR-RS, 2017.

1. Regulação da expressão gênica. 2. Cryptococcus.
3. pequenos RNAs. 4. fragmentos de tRNA. 5. RNA de
interferência. I. Staats, Charley Christian, orient.
II. Título.

RODRIGO SILVA ARAUJO STREIT

ANÁLISE DE FRAGMENTOS DE RNA DERIVADOS DE tRNA EM *Cryptococcus gattii*

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel(a) em Biomedicina.

Aprovado em: 19 de dezembro de 2017.

BANCA EXAMINADORA



Dr. Rodrigo Ligabue-Braun - UFRGS



Dra. Francine Melise dos Santos



Prof. Dr. Charley Christian Staats - UFRGS

AGRADECIMENTOS

Aos meus pais, Jorge e Maria Lúcia, por todo o amor e incentivo, me apoiando nas minhas decisões, me guiando pelo melhor caminho e, acima de tudo, me ajudando nos momentos mais difíceis. Seria impossível ter chegado até aqui sem o suporte e auxílio de vocês.

Aos meus avós, Fernando, Tatiana e Sônia, que sempre me incentivaram e me animaram com meu sonho, e ao meu primo, Pedro, que sempre foi um amigo inigualável e um irmão.

Ao meu orientador, Dr. Charley Staats, por me acolher no laboratório, por todo o conhecimento transmitido e pela disponibilidade para sanar todas as minhas dúvidas. Como um aluno sem conhecimento prévio de informática, sem sua paciência em ensinar e confiança no meu potencial, eu jamais teria desenvolvido um projeto como esse.

Aos professores Dra. Livia Kmetzsch, Dra. Marilene Henning Vainstein e Dr. Augusto Schrank pelo acolhimento no laboratório.

Aos meus amigos e colegas de laboratório Ane, Camila, Uriel e Patrícia, por não apenas me ajudarem nas minhas dificuldades e me transmitirem seus conhecimentos, como também por tornarem mais leve e divertido o meu dia-a-dia.

Aos demais colegas dos laboratórios 220, por todo o auxílio e por todos os bons momentos que vocês me proporcionaram.

Aos meus amigos e colegas formandos, por todas as risadas e bons momentos compartilhados ao longo da graduação. Muito obrigado especialmente à Maria, à Ana Paula, à Cristina, ao Natã e à Caroline Canani, por estarem sempre presentes e por aguentarem minhas reclamações constantes com relação a todas as provas da graduação.

Aos meus amigos Costa, Tex, Emerson, Mellanie e Gabriel Rosa, pelo companheirismo, amizade e por saber que posso contar com vocês sempre, nos bons e maus momentos.

RESUMO

Mecanismos de regulação da expressão gênica constituem um ponto chave de células e organismos, sendo fundamentais em diversos aspectos da funcionalidade celular, como quando expostos a estímulos e variações no ambiente. Neste contexto, moléculas de RNA não codificante também desempenham importantes funções. Recentemente, uma nova classe de pequenos RNAs com atividade regulatória pós-transcricional foi descrita em diversas espécies, sendo denominados fragmentos de RNA derivados de tRNAs (tRFs). Embora tRFs possuam uma função de regulação extremamente similar aos microRNAs, existem indícios de que tRFs exercem sua função independentemente da maquinaria celular utilizada pelos microRNAs na via de RNA de interferência (RNAi). As leveduras basidiomicéticas *Cryptococcus gattii* e *Cryptococcus neoformans* são os agentes etiológicos da criptococose, doença sistêmica que acomete principalmente pulmões e sistema nervoso central. É conhecido que algumas linhagens de *C. gattii* não possuem alguns genes envolvidos na via de RNAi, enquanto que *C. neoformans*, evolutivamente próximo a *C. gattii*, é proeficiente para esta via. Dessa forma, *C. gattii* apresenta um grande potencial como modelo de avaliação da independência dos tRFs da maquinaria de RNAi, o que poderia indicar a existência de uma via alternativa de regulação pós-transcricional nesta levedura. Assim, esse projeto teve por objetivo avaliar a presença de tRFs em *C. gattii*. Empregando metodologias *in silico* baseadas em dados de alinhamento de bibliotecas de pequenos RNAs, foram identificados 34 tRFs únicos para a linhagem R265 de *C. gattii*, originados de 23 das 53 sequências únicas de tRNAs preditas no genoma. Dentre estes, estão inclusos tRFs de uma possível nova classe, originada da região 5' *leader* do transcrito de tRNA primário. Por meio de ensaios de qRT-PCR, foram confirmadas as predições de tRFs representativos, além da identificação de expressão diferencial. Apesar da quantidade de tRNAs que geram tRFs ser extremamente baixa nesta levedura, quando comparada com outros organismos, a identificação de tRFs em uma linhagem de *C. gattii* que apresenta a inviabilidade da via de RNAi, e especialmente a evidência de expressão diferencial, sugerem a existência de regulação por meio de tRFs de forma independente da via de RNAi.

Palavras-Chave: *Cryptococcus*, fragmentos de tRNA, pequenos RNAs, regulação da expressão gênica, RNA de interferência, tRNAs

ABSTRACT

Gene expression regulation mechanisms are a key point for cells and organisms and essential to several aspects of cellular functionality, as those necessary to adapt to stimuli and variations on the environment. In this context, noncoding RNA molecules also have an important role. Recently, a new class of small RNAs with post-transcriptional regulatory activity was described in various species, known as tRNA-derived RNA fragments (tRFs). Although tRFs possess a regulation function extremely similar to microRNAs, there are indications that tRFs exert their functions independently of cellular machinery used by microRNAs in RNA interference pathway (RNAi). The basidiomycetous yeasts *Cryptococcus gattii* and *Cryptococcus neoformans* are the etiologic agents of cryptococcosis, a systemic disease which affects mainly the lungs and central nervous system. Some strains of *C. gattii* do not possess genes involved in RNAi pathway, while the evolutionarily close *C. neoformans* is proficient in such pathway. Therefore, *C. gattii* displays a great potential as a model for evaluation of tRFs RNAi machinery independent activity, which could point to an alternative post-transcriptional regulation pathway existence. Therefore, the aim of this work was to evaluate the presence of tRFs in *C. gattii* R265. Using *in silico* methodologies based on alignment data from small RNAs libraries, we identified 34 unique tRFs in *C. gattii* R265 strain, originated from 23 out of the 53 unique sequences of genome predicted tRNAs. Among these are tRFs of a possible new class, derived from the 5' leader region of primary tRNA transcript. With qRT-PCR assays, we were able to confirm the prediction of representative tRFs, as well as identifying differential expression of some tRFs. Although the number of tRF-producing tRNAs identified was lower compared to other organisms, the tRFs identification in a RNAi pathway-disabled strain of *C. gattii*, and especially the evidence of differential expression, suggest the existence of a RNAi pathway-independent regulation by the tRFs.

Keywords: *Cryptococcus*, gene expression regulation, RNA interference, small RNAs, tRNAs, tRNA fragments

LISTA DE FIGURAS

Figura 1 – Biogênese e funcionamentos dos microRNAs pela via de RNAi.....	17
Figura 2 – Produção e classificação dos principais tipos de tRFs.....	19
Figura 3 – <i>Pipeline</i> para a análise de tRF.....	35
Figura 4 – Representação do alinhamento de tRNAs produtores de tRFs.....	35
Figura 5 – Localização dos tRFs-In identificados na estrutura de seu tRNA de origem.....	37
Figura 6 – Localização dos tRFs-Ld identificados na estrutura de seu tRNA de origem.....	37
Figura 7 – Localização dos tRFs-5 identificados na estrutura de seu tRNA de origem.....	38
Figura 8 – Localização dos tRFs-3 identificados na estrutura de seu tRNA de origem.....	38
Figura 9 – Análise dos níveis relativos de tRFs em <i>C. gattii</i> R265 nas condições de cultivo controle, privação de zinco e privação de ferro.....	39

LISTA DE TABELAS

Tabela 1 – tRFs identificados empregando análises <i>in silico</i>	36
--	----

LISTA DE ABREVIATURAS E SIGLAS

μM	micromolar
$^{\circ}\text{C}$	graus Celsius
Ago	proteínas Argonauta da subfamília AGO
BPDS	Ácido bisulfônico de batofenantrolina
cDNA	DNA complementar
HIV	Vírus da imunodeficiência humana, do inglês, <i>human immunodeficiency virus</i>
lncRNA	RNA longo não-codificante, do inglês, <i>long non-coding RNA</i>
min	minutos
mL	mililitros
mM	milimolar
ncRNA	RNA não-codificante
PCR	Reação em cadeia da polimerase
PBS	Tampão fosfato salino
pH	potencial hidrogeniônico
piRNA	RNA de interação com PIWI
qRT-PCR	PCR quantitativo em tempo real associado à transcrição reversa
RISC	Complexo de silenciamento induzido por RNA
RNAi	RNA de interferência
rpm	Rotações por minuto
sRNA	pequeno RNA
siRNA	pequeno RNA de interferência
TPEN	N, N, N', N' – Tetraquis-2-pirimidil-etil-etiledodiamino
tRF	fragmento de RNA derivado de tRNA
tRF-In	tRF-interno
tRF-Ld	tRF- <i>leader</i>
YPD	Meio extrato de levedura/peptona/dextrose
YNB	Meio extrato de levedura/nitrogênio/base

SUMÁRIO

1	INTRODUÇÃO GERAL	12
1.1	Regulação da expressão gênica	12
1.2	RNAs não-codificantes.....	13
1.3	Via de RNA de interferência.....	15
1.4	Fragmentos de RNA derivados de tRNA.....	18
1.5	<i>Cryptococcus gattii</i> e <i>Cryptococcus neoformans</i>	20
2	JUSTIFICATIVA	22
3	OBJETIVOS	23
4	ARTIGO CIENTÍFICO	24
4.1	Abstract.....	26
4.2	Background	27
4.3	Métodos	30
4.4	Resultados	34
4.5	Discussão.....	41
4.6	Conclusão.....	45
4.7	Referências.....	46
4.8	Arquivos Suplementares	51
5	PERSPECTIVAS	54
6	REFERÊNCIAS BIBLIOGRÁFICAS	55
7	ANEXO A – NORMAS DE PUBLICAÇÃO NA REVISTA BMC Genomics	59

1 INTRODUÇÃO GERAL

1.1 Regulação da expressão gênica

A capacidade de adaptação ao ambiente permite aos organismos a sobrevivência e perpetuação da espécie, podendo ser observada desde organismos de menor complexidade, como bactérias, a sobreviver a estresses abióticos e bióticos, até na diferenciação celular em tempos e locais precisos durante o desenvolvimento de organismos multicelulares complexos. Essa adaptabilidade da célula ou organismo ocorre através de uma série de mecanismos celulares, os quais alteram a disponibilidade de produtos gênicos (RNA e proteínas) de acordo com os sinais extra- e intracelulares. Esses mecanismos são denominados como “regulação da expressão gênica”. Essa regulação em eucariotos ocorre em diversos pontos do processo de expressão gênica, como na transcrição gênica, no processamento e estabilidade do transcrito de RNA, na tradução do RNA mensageiro em proteínas bem como na regulação do funcionamento das proteínas produzidas, auxiliando no funcionamento da célula (CAUSTON et al., 2001; DAY; TUTE, 1998; LÓPEZ-MAURY; MARGUERAT; BÄHLER, 2008).

Uma das primeiras e mais bem elucidadas formas de regulação já descritas é o controle da transcrição (JACOB; MONOD, 1961), sendo também a forma de regulação comum a todos os organismos, apesar de diferir em grau de complexidade entre os mesmos. O controle da transcrição é a base de toda a regulação da expressão gênica, uma vez que é responsável por induzir ou restringir a transcrição de um precursor de RNA. Em eucariotos, a regulação da transcrição se dá por meio de dois mecanismos interligados: os fatores de transcrição e o remodelamento da cromatina (LEE; YOUNG, 2013). Os fatores de transcrição são proteínas com capacidade de ligação ao DNA, que se ligam a sequências específicas de regiões promotoras, reforçadoras e silenciadoras, e que causam, direta ou indiretamente (por meio de cofatores) a indução ou repressão do funcionamento ou acoplagem da maquinaria de transcrição à região promotora (FARNHAM, 2009). Já o remodelamento da cromatina por meio de histonas, as quais se tratam de proteínas de ligação ao DNA com função empacotadora, leva à formação de regiões inacessíveis aos fatores de transcrição, denominada heterocromatina, de forma a promover a repressão de genes localizados dentro da mesma. Somado a isso, as histonas alteram espacialmente a organização da cromatina, aproximando ou afastando regiões reguladoras tanto em *cis* (mesmo cromossomo) quanto em *trans* (cromossomos diferentes) de

determinados genes, de forma a alterar a transcrição dos mesmos (GONZALEZ-SANDOVAL; GASSER, 2016; GREWAL; MOAZED, 2003).

Outra forma de regulação da expressão é o controle pós-transcricional, que engloba mecanismos de processamento e estabilização da molécula de RNA recém transcrita. O principal mecanismo de processamento de RNA é o *splicing*, executado pelo complexo ribonucleoproteico denominado spliceossomo, que faz a remoção de *introns* e recombinação de *exons* do pré-mRNA, afim de gerar um mRNA maduro (BURGE et al., 1999), podendo também, por meio de *splicing* alternativo, gerar diferentes combinações de *exons* a partir de um mesmo pré-mRNA (STAMM et al., 2005). Já o mecanismo de regulação da estabilidade de moléculas de RNA mais bem descrito é o controle do decaimento de mRNAs, sendo regido pela ligação de diversas proteínas nas regiões 5' e 3' não traduzidas (UTR) dos mRNAs, podendo induzir ou inibir a degradação dessas moléculas por ribonucleases (SCHOENBERG; MAQUAT, 2012).

Por muito tempo, essas formas de regulação foram tratadas como a simples interação entre proteínas e os ácidos nucleicos que são seus substratos (DAY; TUIE, 1998), mas com o avanço das metodologias de biologia molecular e computacional, assim como o advento das tecnologias de sequenciamento, novos componentes das vias de regulação foram identificados, classificados como RNAs não-codificantes (ncRNAs), e que não apenas fazem parte de vias já descritas, como também são componentes principais de vias até então não descritas (EDDY, 2001).

1.2 RNAs não-codificantes

A existência de RNAs cujo objetivo final não é a produção de uma cadeia peptídica a partir de sua sequência de bases é reconhecida há décadas, como nos casos dos RNAs ribossômicos (rRNAs), RNAs de transferência (tRNAs), os pequenos RNAs nucleares (snRNAs) que compõe o spliceossomo e os pequenos RNAs nucleolares (snoRNAs), responsáveis pela maturação de rRNAs no nucléolo. Apesar disso, apenas recentemente os ncRNAs começaram a receber maior enfoque, devido às novas tecnologias de sequenciamento que possibilitaram a descoberta de novas classes de RNAs com função regulatória. Adicionalmente, tais tecnologias revelaram que grande parte do genoma humana é transcrito na forma desses RNAs funcionais não codificantes (DJEBALI et al., 2012). Assim, diversas classes de ncRNAs já foram identificadas, com funções que variam desde a regulação na

expressão gênica até a atuação como elementos estruturais (GONZALEZ-SANDOVAL; GASSER, 2016; MATERA; TERNS; TERNS, 2007; MERCER; DINGER; MATTICK, 2009). Apesar da grande variedade de ncRNAs, os mesmos podem ser divididos em dois grandes grupos de acordo com seu tamanho: os RNAs longos não-codificantes (lncRNAs) e os pequenos RNAs (sRNAs).

Os lncRNAs são um grupo heterogêneo de moléculas de RNA que apresentam um comprimento de 200 nucleotídeos ou mais (KAPRANOV et al., 2007), abrangendo uma grande quantidade de moléculas com as mais variadas características. A maioria dos lncRNAs apresentam muitas similaridades com os mRNAs, como a adição de 5'-cap, poliadenilação e *splicing*, podendo muitas vezes ser apenas distinguidos pela ausência de fases de leitura produtoras de proteínas, salvo exceções de lncRNAs que apresentam processamentos únicos, como os lncRNA predecessores de microRNAs e RNAs circulares (QUINN; CHANG, 2016). Apesar dessa similaridade com os mRNAs em termos de processamento, os lncRNAs podem ser transcritos de virtualmente qualquer porção do genoma, podendo ser originados de regiões intergênicas ou dentro de genes codificadores de proteínas, tanto da fita anti-senso quanto da fita senso, de regiões intrônicas ou exônicas, podendo ser até mesmo provenientes de uma ativação bidirecional do promotor de um gene codificador de proteínas (MERCER; DINGER; MATTICK, 2009). Da mesma forma que o termo lncRNA compreende diversas moléculas de tamanhos e biogênese diferentes, as funções exercidas pelos lncRNAs são variadas, atuando como guias para proteínas de remodelamento de cromatina com domínio de associação a RNA, na ativação e repressão da transcrição de genes em *cis* e *trans*, na regulação da estabilidade de mRNAs no citoplasma e no controle dos níveis de microRNAs por meio de sequestro dos mesmos (FATICA; BOZZONI, 2014).

Os sRNAs, por sua vez, são moléculas de RNAs cujas extensões não ultrapassam 200 nucleotídeos, apesar de grande parte das moléculas descritas neste grupo possuírem menos da metade desse comprimento (KAPRANOV et al., 2007). Diferentemente dos lncRNAs, os sRNAs não costumam apresentar características em comum entre suas diferentes classes, e, usualmente, serem extremamente distintas entre si em termos de conservação de sequência ou de estrutura. Apesar disso, sua biogênese tende a ser de duas formas: (i) pela transcrição de um gene próprio do sRNA, podendo esse ser policistrônico, ou (ii) a partir do processamento de outras classes de RNAs (MATERA; TERNS; TERNS, 2007; MATTICK; MAKUNIN, 2006). Os sRNAs em geral atuam associados a outras proteínas ou complexos ribonucleoproteicos, possuindo atividade estrutural, catalítica ou como um guia baseado em sequência de ácidos nucleicos, atuando em uma grande diversidade de fenômenos celulares (MATTICK;

MAKUNIN, 2006). Desde a identificação de rRNAs e tRNAs, os primeiros sRNAs a serem estabelecidos como ncRNAs funcionais, uma grande variedade de classes de sRNAs vem sendo constantemente relatada. Ainda assim, a literatura de sRNAs é composta majoritariamente por estudos de microRNAs, classe cuja associação com a via de RNA de interferência (RNAi) e uma ampla atuação na regulação da expressão gênica em eucariotos superiores (STEFANI; SLACK, 2008) lhe concedeu posição de destaque entre os ncRNAs.

Os microRNAs são sRNAs com tamanho variado entre 21 a 25 nucleotídeos, na sua maioria, e são a classe dominante de sRNAs reguladores (HE; HANNON, 2004). A grande maioria dos microRNAs são produzidos a partir de genes próprios. Entretanto, é possível também encontrar genes de microRNAs organizados em unidades policistônicas. Isto é majoritariamente observado em microRNAs que regulam genes envolvidos em processos biológicos similares (BARTEL, 2004). Também ocorre a produção de microRNAs a partir do processamento de mRNAs e lncRNAs, sendo produzidos tanto a partir de regiões exônicas quanto intrônicas (KIM; KIM, 2007; SLEZAK-PROCHAZKA et al., 2013). Independente da sua origem, os transcritos iniciais de microRNA, denominados pri-microRNAs, tem em comum a presença de uma estrutura na forma de *hairpin*, característica da qual depende todo o processamento do microRNA e entrada na via de RNAi (HE; HANNON, 2004). Essa inserção na via de RNAi é extremamente necessária, uma vez que os microRNAs exercem a sua função de reguladores negativos a nível pós-transcricional por meio da associação ao complexo de silenciamento induzido por RNA (RISC), o principal efetor da via de RNAi (CARTHEW; SONTHEIMER, 2009). Assim, a biogênese e funcionamento dos microRNAs apresentam-se sobrepostos ao funcionamento da via de RNAi.

1.3 Via de RNA de interferência

Inicialmente identificada em *Caenorhabditis elegans*, a partir de estudos com o microRNA *lin4* (LEE; FEINBAUM; AMBROS, 1993), a via de RNA de interferência tem se destacado como uma via de suma importância para a regulação da expressão gênica em eucariotos. Além de sua ocorrência extremamente difusa ao longo de diversas classes de eucariotos (SHABALINA; KOONIN, 2008), é proposto que, por meio de sua associação com microRNAs, a via de RNAi seja responsável pela regulação a nível pós transcricional de mais de 60% dos mRNAs em humanos (FRIEDMAN et al., 2009). Apesar de que a principal forma de atuação e de representação da via de RNAi seja a nível pós-transcricional por meio dos

microRNAs (WILSON; DOUDNA, 2013), ela também pode exercer funcionalidade por meio de associação com outros sRNAs, como com pequenos RNAs de interferência (siRNAs) (CARTHEW; SONTHEIMER, 2009) ou com RNAs de interação com Piwi (piRNAs) (THOMSON; LIN, 2009).

A via de RNAi associada aos microRNAs em mamíferos se inicia com a transcrição dos genes de microRNAs pela RNA polimerase II ou pela excisão de *introns* ou *exons* de mRNAs que contenham microRNAs, produzindo pri-microRNAs. Os pri-microRNAs são fitas-simples de RNA que, por meio de pareamento interno na fita, formam um *hairpin* com extremidades em fita-simples. Ainda no núcleo, o pri-microRNA passa pela primeira etapa de processamento, executada pelo complexo microprocessador, formando pelas proteínas Drosha e DGCR8. As proteínas DGCR8 se ligam as extremidades de fita-simples, guiando o complexo, e então a RNase Drosha faz a clivagem do pri-microRNA a 11 nucleotídeos de distância da divisão entre *hairpin* e fita-simples. Essa clivagem produz um *hairpin* de 60 a 70 nucleotídeos com uma extremidade 5' fosfato e uma extremidade 3' com usualmente 2 nucleotídeos sobressalentes, denominado pré-microRNA. Após o transporte para o citoplasma via Exportina-5, ocorre a segunda etapa de processamento, feita pela RNase Dicer. Por meio do reconhecimento da estrutura de *hairpin* e dos 2 nucleotídeos sobressalentes da extremidade 3' do pré-microRNA, Dicer faz a clivagem da porção do *hairpin* oposta às extremidades 5' e 3', transformando o *hairpin* de fita-simples em um RNA de fita-dupla (dsRNA), levando a formação de um duplex de microRNAs maduros. Da mesma forma que a clivagem por Drosha, a clivagem pela Dicer faz com que ambas as fitas formadas possuam uma extremidade 5' fosfato e uma extremidade 3' com nucleotídeos sobressalentes (WAHID et al., 2010; WILSON; DOUDNA, 2013). Estas informações estão ilustradas na Figura 1.

Após todo o processo de produção dos microRNAs, para que os mesmos possam efetuar sua função de repressão, uma das fitas do duplex deve ser carregada em proteínas Argonata da subfamília AGO (Ago). As Ago são as proteínas principais do RISC, contendo 4 domínios funcionais conservados: o domínio N-terminal, responsável pela separação do duplex de microRNAs e carregamento de um dos microRNAs em Ago, o domínio PAZ, responsável por se ligar à extremidade 3' do sRNA carregado, o domínio MID, responsável por se ligar à extremidade 5' do sRNA carregado, e o domínio PIWI, extremamente semelhante a uma RNase H e que possui atividade de endonuclease. Na maioria dos casos, para um dado duplex de microRNA, a mesma fita é a carregada na proteína argonata, sendo denominada fita guia, enquanto que a outra fita é liberada e degradada, sendo denominada fita passageira. Essa seleção se faz pela estabilidade termodinâmica do duplex de microRNAs, sendo que a fita que apresenta

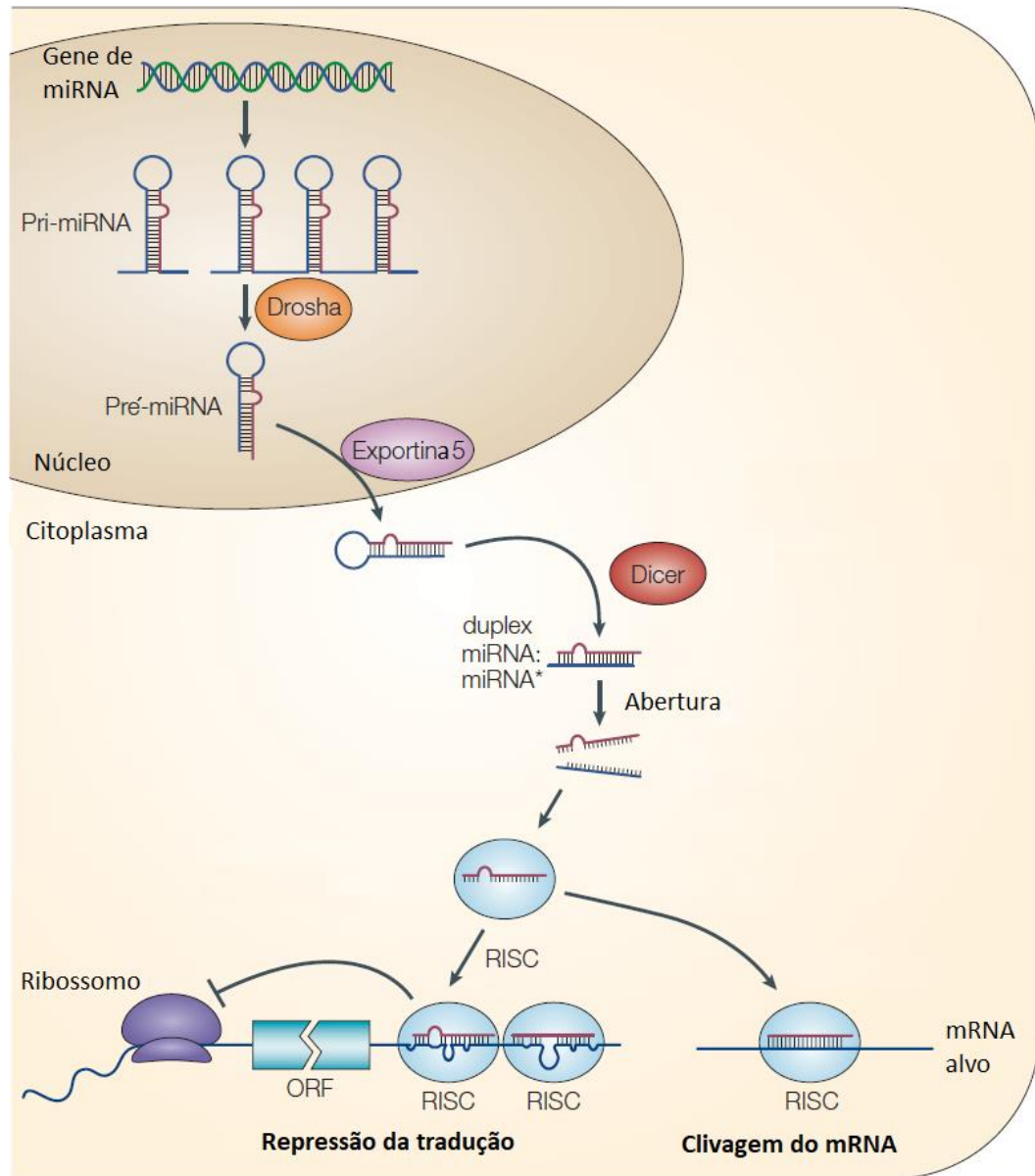


Figura 1: Biogênese e funcionamento dos microRNAs pela via de RNAi. A transcrição dos genes de microRNAs leva a produção de transcritos com estrutura de *hairpin* denominados pri-microRNAs. Ainda no núcleo, esses transcritos sofrem uma clivagem inicial de suas extremidades pela proteína Drosha, gerando os pré-microRNA, que é transportado para o citoplasma. No citoplasma, o pré-microRNA sofre uma nova clivagem pela proteína Dicer, produzindo um duplex de microRNAs, do qual uma das fitas será carregada no RISC e irá levar à repressão da tradução ou à clivagem do mRNA alvo. Adaptado de (HE; HANNON, 2004).

em sua porção 5' uma região de pareamento mais instável possui maior facilidade de interação com o domínio N-terminal de Ago, aumentando sua frequência de carregamento. Com a formação do RISC e com o carregamento do microRNA, o complexo irá identificar o mRNA alvo do microRNA e, a partir do pareamento entre o mRNA e a região *seed* do microRNA, que corresponde aos nucleotídeos 2 até 8 no sentido de 5' a 3', realizar o silenciamento desse mRNA. As regiões alvo das *seeds* de microRNAs geralmente se localizam na região 3' UTR dos mRNAs, e a extensão da complementariedade de bases entre o *seed* e seu alvo define a

forma pela qual o RISC irá gerar a repressão desse mRNA. Caso haja uma complementariedade perfeita entre o *seed* e a região alvo, Ago executa a clivagem da região alvo do mRNA por meio do seu domínio PIWI, levando à degradação do mesmo. Já no caso de a complementariedade ser suficiente para o reconhecimento do alvo, porém insuficiente para iniciar o processo de clivagem, ou no caso de a argonauta presente no RISC não possua um domínio PIWI funcional, ocorre a inibição da tradução do mRNA por meio de mecanismos não endonucleolíticos, como pela inibição da circularização do mRNA, competição de proteínas do RISC com proteínas necessárias para a tradução e por liberação prematura dos ribossomos, sendo que essa inibição pode ser seguida pela deadenilação e degradação do mRNA (BARTEL, 2004; WAHID et al., 2010; WILSON; DOUDNA, 2013). Estas informações estão ilustradas na Figura 1.

Com a crescente quantidade de sRNAs identificados, novas vias alternativas à via canônica de produção de microRNAs vêm sendo identificadas, com novas classes de sRNAs e classes já bem estabelecidas, mas até então não relacionadas com a via de RNAi, sendo encontradas em associação com proteínas Ago (DAUGAARD; HANSEN, 2017). Apesar de que a associação entre sRNAs e proteínas Ago não necessariamente signifique o funcionamento dos sRNAs por meio da via de RNAi, o potencial de novas formas de regulação leva à necessidade de estudos mais aprofundados sobre a associação desses sRNAs com a via.

1.4 Fragmentos de RNA derivados de tRNA

Os fragmentos de RNA derivados de tRNA (tRFs) são uma classe recém descrita de sRNAs produzidos a partir do processamento de transcritos de tRNA, possuindo de 13 a 30 nucleotídeos e atuando na regulação da expressão gênica. Esses fragmentos podem ser produzidos a partir tanto de tRNAs maduros quanto de transcritos de tRNAs primários, e atualmente são classificados de acordo com a região da qual se originam no tRNA: (i) tRF-5 quando produzidos a partir da extremidade 5' do tRNA maduro, (ii) tRF-In (tRF-interno) quando produzidos a partir de regiões internas do tRNA maduro que não incluam uma extremidade, (iii) tRF-3 quando produzidos a partir da extremidade 3' do tRNA maduro e contenham a sequência CCA em sua extremidade, e (iv) tRF-1 quando produzidos a partir do 3' *trailer* do transcrito de tRNA primário (Figura 2) (GEBETSBERGER; POLACEK, 2013; KEAM; HUTVAGNER, 2015; ZHENG et al., 2016).

Por muito tempo, estudos de sRNAs identificaram a presença desses fragmentos, mas por previamente serem considerados produtos da degradação de tRNAs ou artefatos de

sequenciamento, esses fragmentos eram excluídos de qualquer análise. Porém, uma avaliação mais aprofundada de características desses fragmentos levou à caracterização de uma produção ativa de tais moléculas: (i) os tRFs possuem uma produção extremamente precisa, com grandes quantidades de fragmentos provenientes de clivagem na mesma posição do tRNA, indicando que não são produzidos aleatoriamente; (ii) os níveis de produção de tRFs não se correlacionam com os níveis de expressão dos seus tRNAs progenitores, indicando que existem taxas de produção específicas para produção de tRFs, não sendo apenas a degradação dos tRNAs; (iii) os níveis de produção de tRFs são afetados por estímulos estressores, indicando que os mesmos são produzidos como uma forma de resposta ao ambiente e, possivelmente, possuem papel regulatório; (iv) tRFs provenientes de diferentes porções de um mesmo tRNA podem não apresentar níveis de produção semelhantes, indicando que são unidades independentes e produzidas ativamente (BAKOWSKA-ŻYWICKA et al., 2016; GEBETSBERGER; POLACEK, 2013; KUMAR; KUSCU; DUTTA, 2016).

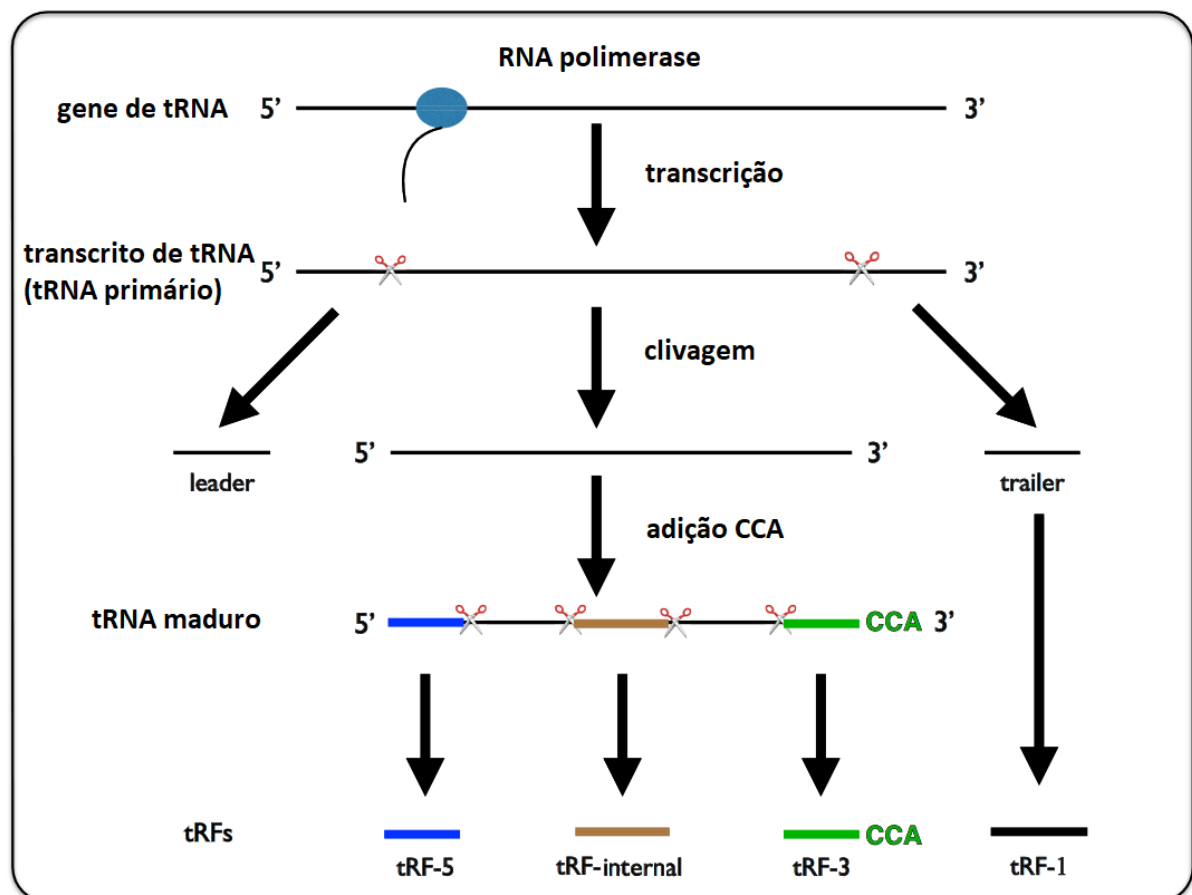


Figura 2: Produção e classificação dos principais tipos de tRFs. A produção dos tRFs pode ocorrer tanto a partir do transcrito de tRNA primário quanto do tRNA maduro. Quando produzido a partir do tRNA maduro, o tRF pode ser classificado como tRF-5 (azul) quando originado da extremidade 5', tRF-In (marrom) quando originado de uma porção interna ou tRF-3 (verde) quando originado da extremidade 3'. Já quando é originado do trailer da região 3' do tRNA primário, o tRF é classificado como tRF-1 (preto). Adaptado de (ZHENG et al., 2016).

Apesar de já terem sido demonstradas as funções de alguns tRFs (KUMAR; KUSCU; DUTTA, 2016), pouco se sabe sobre sua biogênese e forma de atuação, sendo que grande parte das evidências apontam para a associação dos tRFs com a via de RNAi. Devido ao seu tamanho compatível com o dos sRNAs utilizados pelas argonautas e a identificação em diversos estudos a associação de tRFs com argonautas, os tRFs se apresentam como fortes candidatos a uma via não-canônica de RNAi (KUMAR et al., 2014). Foi demonstrado que tRFs-3 associados a argonautas são capazes de se ligar a regiões 3' UTR de mRNAs e produzir a repressão da expressão, utilizando o mesmo mecanismo de *seed* que microRNAs (KUMAR et al., 2014), assim como alguns tRFs-5. Na mesma linha da via de RNAi, a constatação de que tRFs-5 se concentram no núcleo e são capazes de reprimir a expressão de genes alvo também aponta para uma atuação semelhante aos siRNAs e piRNAs (KUMAR et al., 2014; ZHANG et al., 2016).

Mesmo com esses indícios, ainda é questionável a atribuição do funcionamento dos tRFs à via de RNAi. Apesar da associação de tRFs-5 e tRFs-3 com argonauta e seus papéis semelhantes aos exercidos por sRNAs da via de RNAi, os tRFs-1 não apresentam associação com argonautas. Além disso, a produção de tRFs independe das proteínas Dicer, diferindo de outros sRNAs utilizados pela via (KUMAR et al., 2014). Por fim, tRFs se fazem presentes também em organismos evolutivamente prévios ao desenvolvimento da via de RNAi ou em que houveram deleções de proteínas da via, sendo que alguns mecanismos independentes de proteínas argonauta já foram identificados (BAKOWSKA-ŻYWICKA et al., 2016; GEBETSBERGER et al., 2012). Outras classes de sRNAs cujas características e função se assemelham as dos microRNAs já foram identificadas em bactérias, atuando por meio de mecanismos semelhante à via de RNAi, apesar de não-relacionados (DESNOYERS; BOUCHARD; MASSÉ, 2013; GOTTESMAN, 2005). Assim, torna-se necessário avaliar a extensão da dependência da via de RNAi por parte dos tRFs, visto que os mesmos podem possuir uma via de atuação própria, ancestral à via de RNAi, sendo a utilização de argonautas para exercer funcionalidade uma exceção ao invés de uma regra.

1.5 *Cryptococcus gattii* e *Cryptococcus neoformans*

C. gattii é uma levedura basidiomicética que, juntamente com *C. neoformans*, são os agentes etiológicos da criptococose, doença fúngica fatal de maior prevalência no mundo (KWON-CHUNG et al., 2014; SCHMIEDEL; ZIMMERLI, 2016). A criptococose afeta especialmente imunocomprometidos, sendo que seu quadro mais grave, a meningite

criptococócica, é responsável por cerca de 15% das mortes relacionadas à infecção por HIV no mundo, com uma taxa de mortalidade em torno de 95% (RAJASINGHAM et al., 2017). A infecção ocorre a partir da inalação de esporos das leveduras, presentes em ambientes contaminados por excretas de aves (*C. neoformans*) ou em cascas de árvores (*C. gattii*), e normalmente evolui para meningite criptococócica, apesar de ser uma doença de caráter sistêmico (KWON-CHUNG et al., 2014; LIN; HEITMAN, 2006). *C. neoformans* é responsável por cerca de 80% dos casos de criptococose, mas apesar da menor frequência de casos de criptococose causados por *C. gattii*, o mesmo possui a capacidade de infectar indivíduos imunocompetentes, sendo responsável por diversos surtos de criptococose na América do Norte (BYRNES III et al., 2009a, 2009b; KIDD et al., 2004; KWON-CHUNG et al., 2014).

Há diversas evidências que *C. neoformans* apresenta uma via de RNAi completa e funcional, enquanto a linhagem R265 de *C. gattii* apresenta deleções nos genes que codificam as proteínas Argonauta e RNA polimerase dependente de RNA, de forma a inviabilizar o funcionamento da via nessa linhagem. Além disso, em *C. neoformans*, foram identificados fragmentos de tRNAs associados a proteínas argonauta, apesar de que os mesmos não foram explorados (FERETZAKI et al., 2016; JIANG et al., 2012; WANG et al., 2010). De forma interessante, a perda da funcionalidade da via de RNAi na linhagem R265 de *C. gattii* não levou a uma perda da virulência, uma vez que essa linhagem é um isolado hipervirulento do surto da ilha de Vancouver (KIDD et al., 2004), sendo provável que algum mecanismo alternativo de controle pós-transcricional tenha suplantado a ausência da regulação por microRNAs.

2 JUSTIFICATIVA

Em vista das informações fornecidas, *C. gattii* apresenta um bom modelo para estudos de tRFs e comparação com organismos que apresentam a via de RNAi conservada, com potencial para a avaliação do grau de dependência dos tRFs para com a via de RNAi, assim como para a identificação de possíveis novas vias envolvendo sRNAs capazes de atuar como microRNAs, incluindo os tRFs.

3 OBJETIVOS

2.1 Objetivo geral

Avaliar a existência de tRFs em *Cryptococcus* spp.

2.2 Objetivos específicos

- Construir uma *pipeline* para a descrição de *in silico* tRFs em *Cryptococcus* spp.
- Descrever os tRFs em *C. neoformans* e *C. gattii*.
- Avaliar o grau de conservação das sequências e da expressão dos tRFs entre *C. neoformans* e *C. gattii*.
- Confirmar a presença de tRFs em *C. gattii* por meio de ensaios *in vitro*.

4 ARTIGO CIENTÍFICO

O artigo intitulado “Análise de fragmentos de RNA derivados de tRNA em *Cryptococcus gattii*” foi formatado conforme as normas para publicação junto ao periódico *BMC Genomics*.

1 **Análise de Fragmentos de RNA Derivados de tRNAs em *Cryptococcus gattii***

2 ***Analysis of tRNA-derived RNA fragments in Cryptococcus gattii***

3

4 Rodrigo Silva Araujo Streit¹, Ane Wichine Acosta Garcia^{2,3}, Livia Kmetzsch^{2,3}, Augusto
5 Schrank^{2,3}, Marilene Henning Vainstein^{2,3}, Charley Christian Staats^{2,3}

6

7 ¹Acadêmico do curso de Biomedicina, Instituto de Ciências Básicas da Saúde, Universidade
8 Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

9 ²Programa de Pós-Graduação em Biologia Celular e Molecular, Centro de Biotecnologia,
10 Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

11 ³Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade
12 Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

13

14 *Autor de correspondência (C. C. Staats)

15 Centro de Biotecnologia

16 Universidade Federal do Rio Grande do Sul (UFRGS)

17 91501-970, Porto Alegre, RS, Brazil.

18 staats@cbiot.ufrgs.br

19 Telefone: +55 (51) 3308-6091

20

21

22

23

24

25

26 4.1 Abstract

27 *Background:* Fragmentos de RNA derivados de tRNA são pequenos RNAs de função
28 regulatória presentes em uma vasta gama de organismos. Existem indícios de sua associação
29 com a via de RNA de interferência, atuando de forma semelhante a microRNAs, siRNAs e
30 piRNAs. Entretanto, a presença desses fragmentos em organismos com a ausência da via de
31 RNA de interferência levanta a hipótese de uma via alternativa de regulação da expressão
32 gênica para os mesmos. Assim, o presente estudo tem por objetivo analisar a presença desses
33 fragmentos em uma linhagem de *C. gattii* que apresenta a via de RNA de interferência inativa.
34 *Resultados:* Empregando análises *in silico* utilizando dados de RNA-seq da linhagem R265 de
35 *C. gattii*, foram identificados 34 fragmentos de RNA derivados tRNA, oriundos de 23 dos 53
36 tRNAs únicos. Por meio de ensaios de qRT-PCR, foram confirmadas as predições de tRFs,
37 sendo também detectada a expressão diferencial de tRFs.
38 *Conclusões:* Apesar da identificação de fragmentos de RNA derivados de tRNA ser pequena
39 em comparação com outros organismos, a presença desses fragmentos em um organismo sem
40 uma via de RNA de interferência funcional, e especialmente a evidência de expressão
41 diferencial, sugerem a existência de mecanismos alternativos para o funcionamento de alguns
42 fragmentos de tRNA.

43

44 **Palavras-chave:** RNA não-codificante, pequeno RNA, RNAs regulatórios, tRNAs, tRFs

45

46

47

48

49

50

51 4.2 Background

52

53 Ao longo dos últimos anos, os sRNAs (pequenos RNAs), classe de RNAs não codificantes com
54 menos de 200 nucleotídeos [1], vêm se destacando como importantes peças na regulação da
55 expressão gênica. Dentre as classes já caracterizadas, os microRNAs recebem um lugar de
56 destaque pela sua ampla atividade na regulação pós-transcricional em eucariotos superiores
57 [2,3], atuando por meio da via de RNAi (RNA de interferência) [4]. A via de RNAi também
58 exerce suas funções utilizando outros sRNAs bem estudados, como os siRNAs (pequenos
59 RNAs de interferência) [5] e piRNAs (RNAs de interação com PIWI) [6], tanto em nível pós-
60 transcricional quanto em nível transcricional, e novas classes de sRNAs regulatórios associados
61 a essa via são constantemente identificadas [7].

62 Fragmentos de RNA derivados de tRNA (tRFs) são sRNAs de 13 a 30 nucleotídeos produzidos
63 a partir do processamento de tRNAs e que possuem atividade regulatória [8]. A classificação
64 proposta para tRFs é feita de acordo com a região do tRNA da qual se originam, sendo que a
65 maior parte dos tRFs até então identificados são originados das extremidades 5' e 3' ou regiões
66 internas de tRNAs maduros (tRF-5, tRF-3 e tRF-interno, respectivamente), assim como do 3'
67 *trailer* de transcritos de tRNA primário (tRF-1) [9,10]. A existência de fragmentos de tRNA já
68 era evidenciada em estudos nos quais eram considerados produtos de degradação [11,12], mas
69 apenas após a identificação da atividade regulatória de um tRF-1 na proliferação celular em
70 células LNCaP e C4-2 [13], os tRFs passaram a ser estudados como moléculas funcionais.
71 Desde então, a presença de tRFs tem sido relatada em uma ampla gama de organismos, sendo
72 que sua função principal é atribuída a seu funcionamento em conjunto com a via de RNAi [14].
73 Tanto tRFs-5 quanto tRFs-3 apresentam associação com argonautas, sendo demonstrada a
74 capacidade de tRFs-3 de atuar no silenciamento a nível pós-transcricional em um

75 comportamento microRNA-like. Já tRFs-5 aparentam ter uma função mais relacionada a
76 siRNAs e piRNAs, concentrando-se no núcleo e atuando na regulação da transcrição [8,15].
77 Apesar da atribuição da funcionalidade dos tRFs à via de RNAi, poucos tRFs-5 e tRFs-3 tem
78 sua função demonstrada por meio dessa via. Adicionalmente, tRFs-1 não apresentam
79 associação com argonautas [14] e não há evidências da forma de atuação dos tRFs-In (tRF-
80 interno). A presença de tRFs em organismos evolutivamente anteriores às proteínas da via de
81 RNAi já foi demonstrada, como em *Escherichia coli* e *Haloferax volcanii* [16,17], junto com a
82 existência de mecanismos alternativos para o funcionamento dos mesmos. Análises *in vitro* em
83 *Saccharomyces cerevisiae*, organismo que apresenta a perda de proteínas da via de RNAi,
84 indicaram a presença em grande número de tRFs, assim como sua responsividade a estímulos
85 estressores, reforçando a possibilidade da atuação de tRFs de maneira independente à via de
86 RNAi [18]. Já foi demonstrado que outras classes de sRNAs presentes em bactérias apresentam
87 mecanismos de funcionamento extremamente semelhantes a microRNAs, atuando por meio do
88 pareamento com mRNAs e recrutamento de chaperonas Hfq para o silenciamento do alvo [19].
89 Assim, a extensão da dependência da via de RNAi por parte dos tRFs precisa ser estudada, uma
90 vez que vias alternativas podem ser tão relevantes quanto a via de RNAi para o funcionamento
91 dos tRFs.

92 As leveduras *C. gattii* e *C. neoformans* são os agentes etiológicos da criptococose humana,
93 doença fúngica fatal de maior prevalência no mundo [20,21]. Há evidências que demonstraram
94 a ausência dos genes das proteínas argonauta e RNA-polimerase dependente de RNA,
95 proteínas-chave da atividade da via de RNAi, na linhagem R265 de *C. gattii*, enquanto *C.*
96 *neoformans* apresenta uma via de RNAi e regulação por microRNAs funcional [22–24]. De
97 forma interessante, o estudo da funcionalidade da via de RNAi em *C. neoformans* identificou a
98 presença de fragmentos de tRNA [23], apesar de que esse fato não foi explorado pelos autores.
99 Devido a essa proximidade com um organismo que apresenta via de RNAi funcional, o estudo

100 de tRFs na linhagem R265 de *C. gattii* apresenta um grande potencial na identificação da
101 influência dessa via no funcionamento dos tRFs. Assim, o presente estudo tem por objetivo
102 avaliar por meio de ensaios *in silico* a presença de tRFs em *C. gattii*, afim de avaliar o impacto
103 da perda da via de RNAi em termos de conservação.

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125 **4.3 Métodos**

126

127 *Dados de RNA-seq*

128

129 Para as análises de identificação e expressão diferencial de tRFs em *C. gattii*, foram utilizadas
130 bibliotecas de pequenos RNAs provenientes de cultivos da linhagem R265 em condição
131 controle (YNB), privação de ferro (BPDS) e privação de zinco (TPEN), utilizadas por Ferrarezi
132 e colaboradores [25] e presentes no *Sequence Read Archive* com o código de acesso
133 SRR5209175. Para as análises de identificação de tRFs em *C. neoformans*, foi utilizada a
134 biblioteca de pequenos RNAs proveniente do cultivo da linhagem H99, disponibilizada no GEO
135 *database* [26] (código de acesso GSE43363).

136

137 *Identificação de tRFs in silico*

138

139 Com o intuito de identificar tRFs em *C. gattii* e *C. neoformans*, foi desenvolvida uma *pipeline*
140 de análise baseada no servidor online *tRF2Cancer* [10] que permitisse a identificação de tRFs
141 em qualquer organismo a partir de dados de RNA-seq.

142 Foram preditos os genes de tRNA para a linhagem R265 de *C. gattii* [27] empregando os
143 programas *tRNAscan-SE* [28] (versão 2.0) e *Infernal* [29] (versão 1.1.7), e pelo servidor online
144 *RNA-space* [30], sendo os resultados comparados e aceitos os genes com maior suporte entre as
145 diferentes ferramentas. Os genes de tRNA da linhagem H99 de *C. neoformans* foram obtidos a
146 partir do banco de dados online tRNAdb [31]. A predição do tRNA sem seus íntrons foi
147 utilizada como tRNA maduro, enquanto que a predição do tRNA acrescido de seus íntrons e de
148 100 nucleotídeos a jusante e a montante foi utilizada como tRNA primário. tRNAs de sequência
149 iguais foram unidos, mantendo apenas sequências únicas.

150 Cada biblioteca de *reads* foi separada pelo tamanho dos *reads*, e então alinhada contra as
151 sequências dos tRNAs maduros e primários por meio do software *Bowtie2* (versão 2.2.6) [32],
152 sendo excluídos alinhamentos com *mismatches* ou alinhamento compartilhado. Foram então
153 executadas as contagens de *reads* alinhados contra cada tRNA, utilizando o software *Samtools*
154 (versão 0.1.19) [33], e de *reads* alinhados contra cada nucleotídeo de cada tRNA, utilizando o
155 software *Bedtools* (versão 2.17.0) [34]. Os dados das contagens foram carregados em um código
156 em linguagem *python* desenvolvido utilizando os modelos estatísticos do servidor tRF2Cancer
157 [10] para a identificação de tRFs (Arquivo Suplementar 1). Regiões cujas contagens de *reads*
158 se apresentaram desviantes de uma distribuição aleatória ($p < 0.01$) foram reportadas pelo
159 código, sendo classificadas como tRFs caso a região desviante apresente a mesma extensão que
160 o os *reads* do alinhamento. Quando tRFs de diferentes tamanhos foram identificados em uma
161 mesma região, variando a extensão de apenas uma das extremidades, o tRF de maior extensão
162 foi aceito para a anotação do tRF, visto que, para todos os casos, esse foi o tRF de maior
163 expressão. Os tRFs foram então anotados e classificados de acordo com sua região de origem.
164 Para a visualização dos alinhamentos foi utilizado o software *IGV* [35] e para a produção das
165 representações de origens dos tRFs foi utilizado o software *VARNA* [36].

166

167 *Linhagem celular e condições de cultivo*

168

169 Células da linhagem R265 de *C. gattii* foram pré-cultivadas em meio YPD líquido por 18 horas
170 a 200 rpm e a uma temperatura de 30°C. As células foram então centrifugadas (5000g por 5
171 min) e lavadas duas vezes em tampão PBS (NaCl 137 mM; KCl 2,7 mM; Na₂HPO₄ 10 mM;
172 KH₂PO₄ 1,8 mM; pH 7,4). As células foram suspensas em meio YNB (*Yeast Nitrogen Base*,
173 sem sulfato de amônio e aminoácidos) e foi realizada a determinação da densidade de células
174 com o auxílio de uma câmara de Neubauer. As células quantificadas foram então inoculadas

175 em uma concentração aproximada de 10^7 células em 100mL de meio YNB para cada uma das
176 condições de cultivo. Para a condição controle, foi utilizado apenas meio YNB; para condição
177 de privação de zinco, foi utilizado meio YNB acrescido de $10 \mu\text{M}$ do quelante de zinco TPEN
178 (N, N, N', N' – Tetraquis-2-pirimidil-etil-etiledodiamino); e para a condição de privação de
179 ferro, foi utilizado meio YNB acrescido de 1 mM do quelante de ferro BPDS (ácido bisulfônico
180 de batofenantrolina). Cada condição de cultivo foi executada em triplicata. Os cultivos foram
181 mantidos por 18 horas a 200 rpm e a uma temperatura de 30°C . Após o cultivo, as células foram
182 recuperadas por meio de centrifugação (5000g por 5 min), o sobrenadante foi descartado, as
183 células foram imediatamente congeladas em nitrogênio líquido e armazenadas em ultrafreezer
184 (-80°C) até sua liofilização.

185

186 *Extração e quantificação de RNA*

187

188 Após sua liofilização, as células foram maceradas e foi feita a extração do RNA utilizando o
189 reagente Trizol[®] (Invitrogen), seguindo as instruções do fabricante. A integridade do RNA foi
190 avaliada por meio de eletroforese em gel de agarose 1% e a quantificação do RNA foi realizada
191 através de análise fluorométrica no sistema Qubit (Life Technologies), conforme as
192 recomendações do fabricante.

193

194 *qRT-PCR*

195

196 Para confirmar a predição e a expressão dos tRFs em *C. gattii* nas condições de cultivo, foi
197 realizada a síntese de cDNA e posterior PCR em tempo real para 3 tRFs, 2 provenientes da
198 região 5' *leader* de tRNAs ProAGG, sendo eles tRF-Ld(26) ProAGG e tRF-Ld(25), e 1
199 proveniente da região interna de tRNAs GluCTC, tRF-In GluCTC (Tabela 1), tendo como

200 normalizador o snRNA U6. A síntese do cDNA dos tRFs e de U6 foi realizada por *stem-loop*
201 RT-PCR[37] a partir de amostras de RNA tratadas com Dnase (Promega). O ensaio de qRT-
202 PCR foi desenvolvido em um sistema StepOne Real-Time PCR (Applied Biosystems),
203 utilizando as configurações de ciclagem térmica com uma fase inicial de 95°C por 10 min,
204 seguida de 50 ciclos de 95°C por 15 s, 50°C por 15 s e 72°C por 60 s. As reações de PCR foram
205 feitas com um volume final de 20 µL, contendo 2 µL de cDNA (2,5 ng/µL), 2 µL SYBR Green
206 (1:1000) (Invitrogen), 0,1 µL dNTP (5mM), 2 µL PCR buffer 10x, 1,2 µL MgCl₂, 0,5 U
207 Platinum Taq DNA Polymerase (Invitrogen), e 0,2 µL de cada primer (5 pmol). As curvas de
208 *melting* foram analisadas ao fim das reações afim de confirmar a presença de apenas 1 produto
209 de PCR. Cada amostra de cDNA foi analisada em triplicata. A expressão relativa foi
210 determinada pelo método $2^{-\Delta CT}$. As análises estatísticas foram realizadas por *one-way ANOVA*
211 *test*, seguido de *Dunnet's post test*, utilizando o software *GraphPadPrism6*. O mesmo software
212 foi utilizado para a obtenção dos gráficos dos resultados.

213

214

215

216

217

218

219

220

221

222

223

224

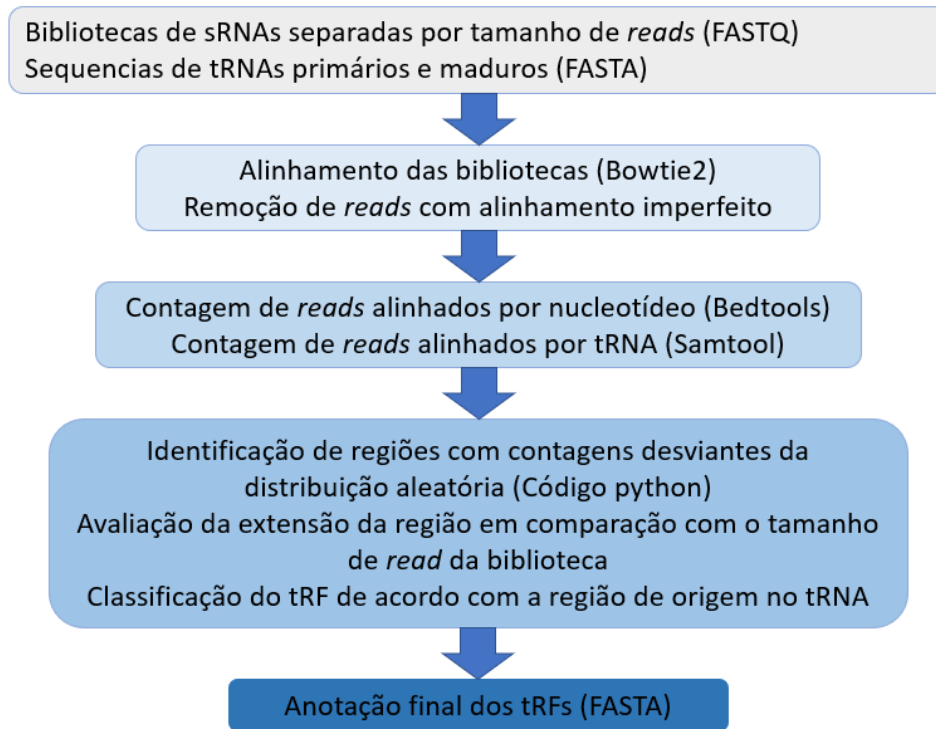
225 4.4 Resultados

226

227 Para avaliar a presença de tRFs em *C. gattii*, foi empregada uma *pipeline* de análises baseada
228 no servidor online tRF2Cancer. Por meio dessa *pipeline* (Figura 3), foi possível identificar 34
229 tRFs, produzidos a partir de 23 das 53 sequências únicas de tRNA (Tabela 1, Figura 4). Apesar
230 dos tRFs identificados apresentarem um padrão de processamento extremamente preciso, na
231 qual *reads* do mesmo tamanho para uma dada região produtora de tRF se encontravam
232 alinhados nas mesmas posições, para todos as regiões produtoras de tRFs identificadas, na
233 exceção da região produtora de tRF-5' do tRNA de ProAGG, existem formas minoritárias de
234 tRFs identificadas por meio da *pipeline*. Essas formas minoritárias variam em poucos
235 nucleotídeos da forma majoritária da região, que em todos os casos foi o tRF de maior extensão,
236 sendo que essa variação ocorre exclusivamente em uma das extremidades, podendo indicar
237 processamento do tRF. Também foi observado que a contagem de *reads* e tamanho de tRFs
238 provenientes de um mesmo tRNA não são iguais, demonstrando processamentos individuais.
239 De forma interessante, apesar do número de tRFs identificados provenientes das extremidades
240 5', 3' e regiões internas de tRNAs maduros ser extremamente similar, não foram identificados
241 tRFs provenientes da região do 3' *trailer* do transcrito de tRNA primário (Figuras 5, 7 e 8).
242 Além disso, foram identificados 3 tRFs provenientes exclusivamente da região 5' *leader* do
243 transcrito de tRNA primário de ProAGG, possuindo sua extremidade 3' precisamente no último
244 nucleotídeo a montante da sequência do tRNA maduro (Figura 6). Avaliando os tRFs com
245 relação à sua extensão, foi visto que os mesmos, em sua maioria, se encontram dentro da
246 distribuição de 13 a 30 nucleotídeos, com picos em 18, 22 e 29 nucleotídeos.

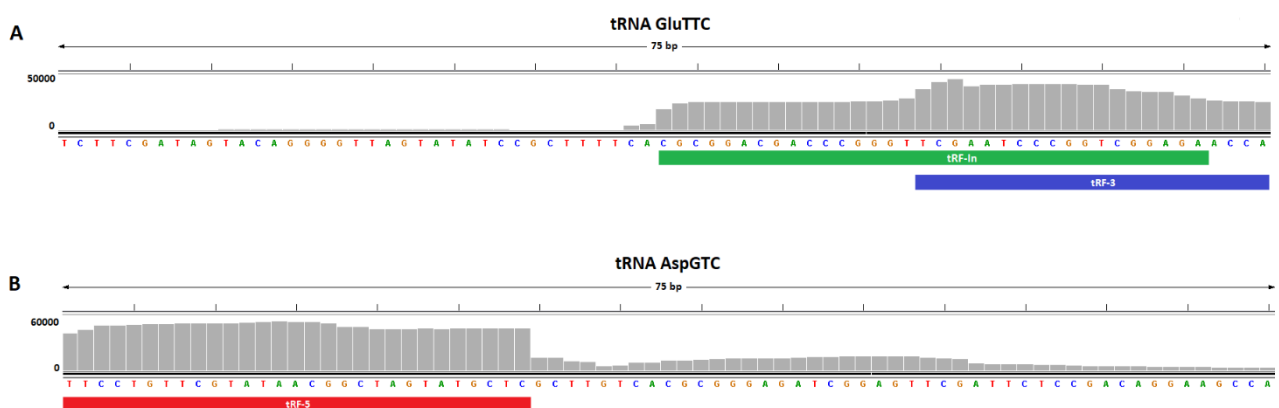
247 Após a identificação de tRFs em *C. gattii*, a mesma *pipeline* foi executada em *C. neoformans*,
248 a fim de que fossem identificados tRFs para comparação e avaliação da conservação.
249 Entretanto, devido a problemas com os dados de RNA-seq depositados na base de dados que

250 impediram a obtenção dos *reads* limpos de adaptadores, não foi possível executar as análises
 251 em *C. neoformans*.



252 **Figura 3: Pipeline para a análise de tRFs.**

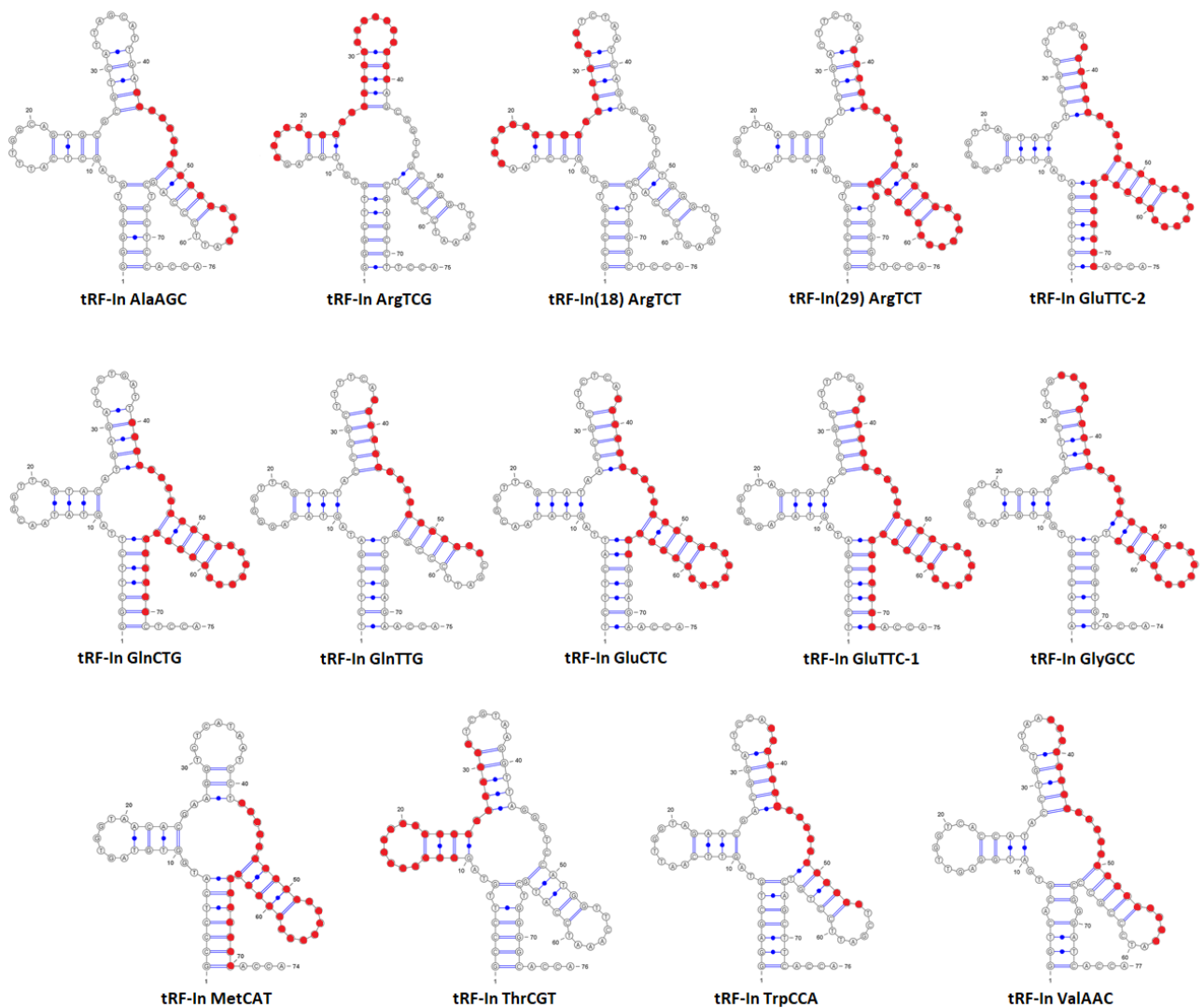
253



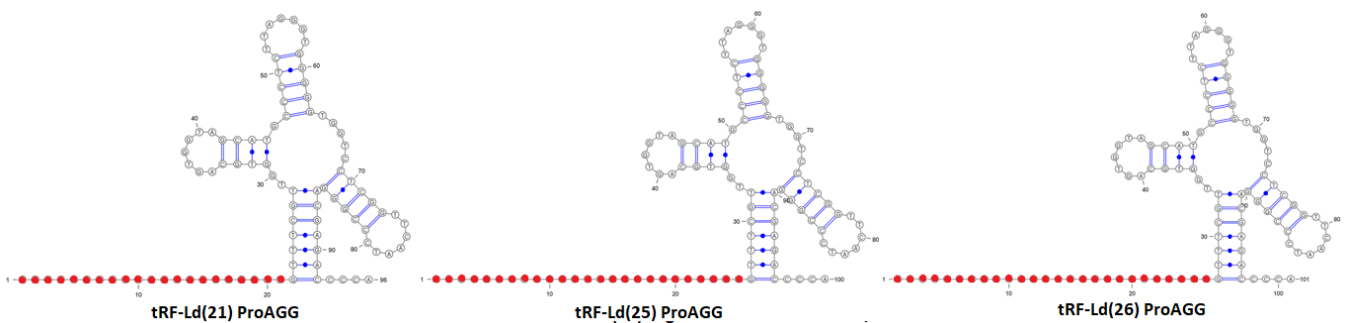
254 **Figura 4: Representação do alinhamento de tRNAs produtores de tRFs.** Representação dos
 255 alinhamentos de *reads* da biblioteca controle contra os tRNAs (A) GluTTC e (B) AspGTC, com
 256 a cobertura de *reads* em cinza e os tRFs destacados em vermelho (tRF-5), verde (tRF-In) e azul
 257 (tRF-5). Visualização pelo software IGV.

258 **Tabela 1: tRFs identificados por meio das análises *in silico*.** Listagem dos 34 tRFs
 259 identificados nas análises das 3 bibliotecas de *reads* da linhagem R265 de *C. gattii*, contendo o
 260 tRNA de origem, a sequência, a classe e o tamanho do tRF.

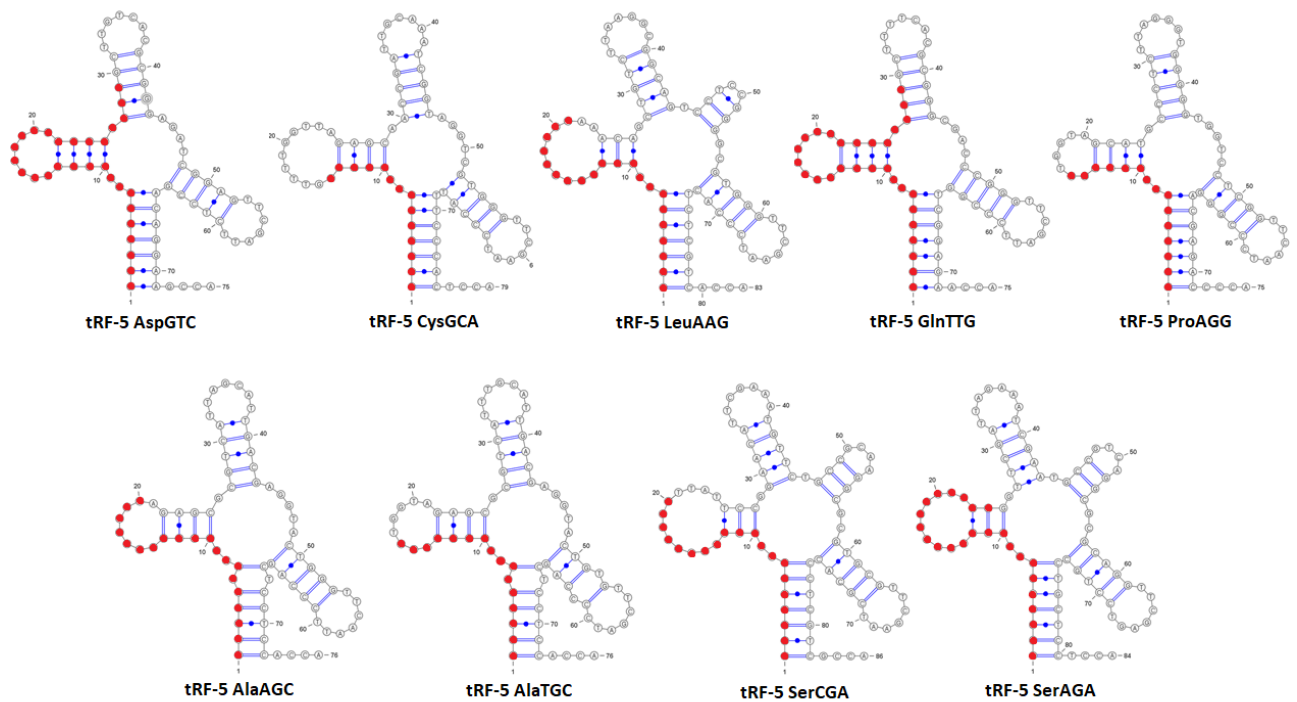
tRNA	Sequência	Classe	Tamanho
MetCAT	CAGCTCTGAGTTTCGATTCTCAGTGAGGGC	Interno	29
MetCAT	GTTCGATTCTCAGTGAGGGCACCA	3'	24
GlyGCC	CCATCGATGCGGCAAGGGTTCGACTCCCT	Interno	29
ValAAC	CACACAGGAGGTCGGCGTTTCG	Interno	22
ThrTGT	CTCACAGCCGTGGCACCA	3'	18
ThrCGT	CTCAGGGGTTAGAGTATGGCCT	Interno	22
GlnCTG	CTCGTGACCTGGGTTTCGACTCCCAGAGAGGC	Interno	31
AspGTC	TTCCTGTTCGTATAACGGCTAGTATGCTC	5'	29
GluTTC-1	CGCGGGCGACCCGGGTTTCGATTCCCGGTCGGAGA	Interno	34
GluTTC-1	TCGATTCCCGGTCGGAGAACCA	3'	22
GluTTC-2	CGCGGACGACCCGGGTTTCGAATCCCGGTCGGAGA	Interno	34
GluTTC-2	TCGAATCCCGGTCGGAGAACCA	3'	22
GlnTTG	GGCGACATCGTCTAGTGGTTAGGACGTTT	5'	29
GlnTTG	TCCGAACGACCTGGGTT	Interno	17
ProAGG	GTTTCGTTGGTGCAG	5'	15
ProAGG	AACAGTTGGGCTTCGGTGGCATAGCC	Leader	26
ProAGG	ATCAGTGGTCGTCGGTGGCATAGCC	Leader	25
ProAGG	TTGGGCTTCGGTGGCATAGCC	Leader	21
SerAGA	GGAGTAGTGCGACAGTGGTACTC	5'	24
GluCTC	CGCGGTAGACCAGGGTTCAACTCCCTGCT	Interno	29
AlaTGC	GGGGGTGTAGCTCACT	5'	16
LysTTT	CCCCGTCTGAAGGCTCCA	3'	17
ArgCCT	TCCCCACCGTTGGTTCCA	3'	18
ArgTCT	TGGTTAAGGCTTCTGACT	Interno	18
ArgTCT	CAGAGGATTGTGGGTTTCGAGTCCCACCTT	Interno	29
ArgTCG	ATCCCGTCGAGCCTTCCA	3'	18
ArgTCG	TGGTAGCATTCTAGACTTCGAATCT	Interno	25
CysGCA	GTGGGAGTAGCTCA	5'	14
TrpCCA	AATCCGTCGGTCGCAGGT	Interno	18
LeuAAG	GGCGAGATGTGTGAGTGGTCT	5'	21
SerCGA	GGCGAGGTCGGATAGCCCGG	5'	20
SerCGA	CGCACCCCTCGTCGCCA	3'	16
AlaAGC	GGGGGTGTAGCTCATTGTC	5'	20
AlaAGC	CGAGGTAAGGTTTCA	Interno	16



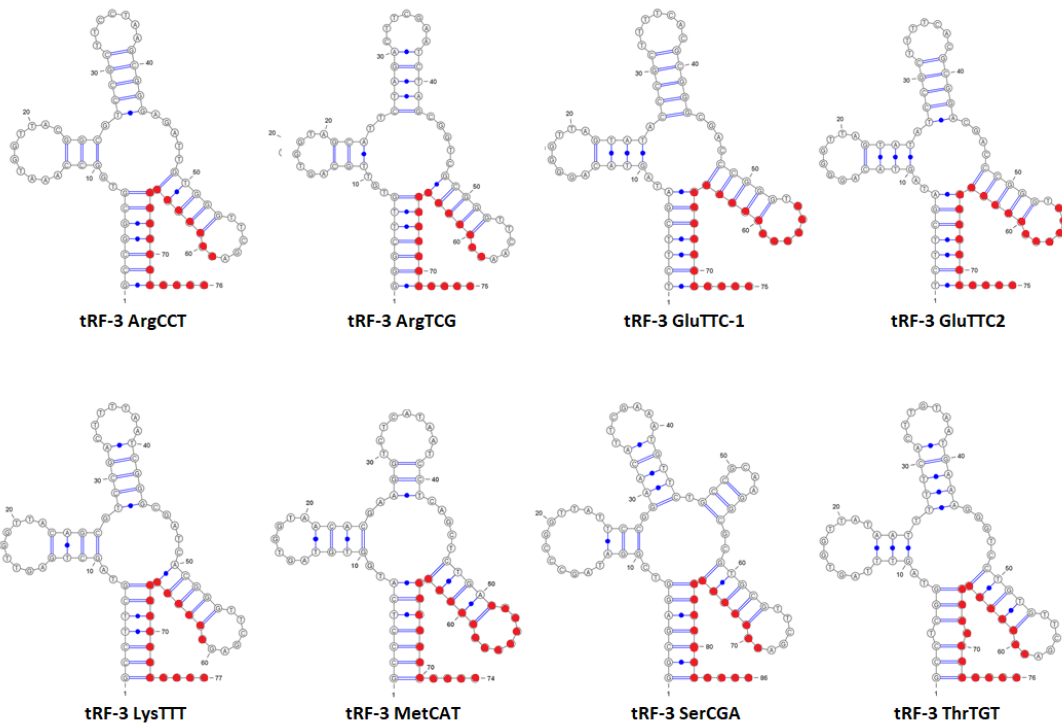
261 **Figura 5: Localização dos tRFs-In identificados na estrutura de seu tRNA de origem.**
 262 Representação das estruturas dos tRNAs produtores dos 14 tRFs-In identificados, geradas pelo
 263 software VARNA, destacando em vermelho as regiões de origem de tRFs.



264 **Figura 6: Localização dos tRFs-Ld identificados na estrutura de seu tRNA de origem.**
 265 Representação das estruturas dos tRNAs produtores dos 3 tRFs-Ld identificados, geradas pelo
 266 software VARNA, destacando em vermelho as regiões de origem de tRFs.

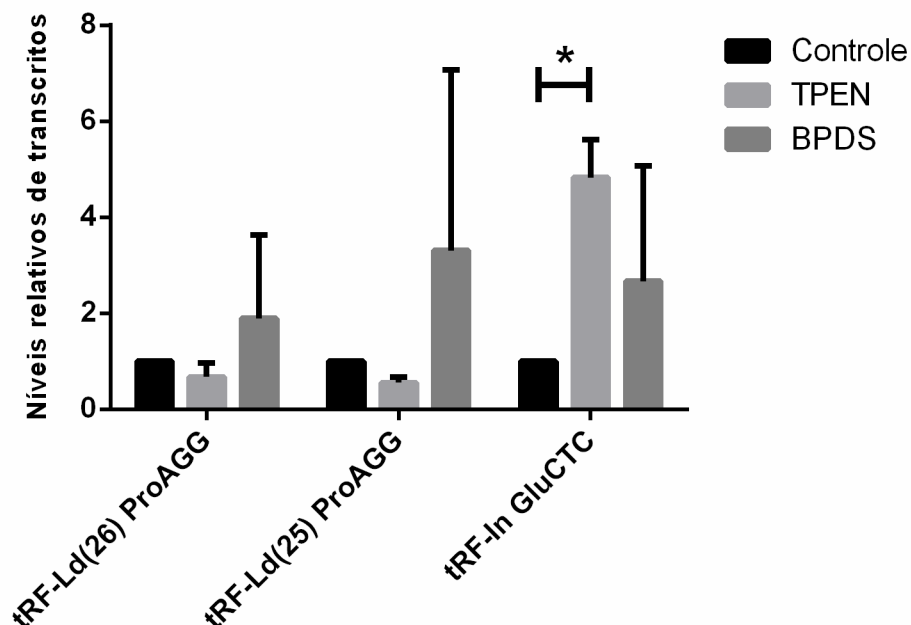


267 **Figura 7: Localização dos tRFs-5 identificados na estrutura de seu tRNA de origem.**
 268 Representação das estruturas dos tRNAs produtores dos 9 tRFs-5 identificados, geradas pelo
 269 software VARNA, destacando em vermelho as regiões de origem de tRFs.



270 **Figura 8: Localização dos tRFs-3 identificados na estrutura de seu tRNA de origem.**
 271 Representação das estruturas dos tRNAs produtores dos 8 tRFs-3 identificados, geradas pelo
 272 software VARNA, destacando em vermelho as regiões de origem de tRFs.

273 A fim de que fossem confirmadas as predições dos tRFs, foi executado um ensaio de qRT-PCR
 274 associado à síntese de cDNA por *Stem-loop* para dois tRF-Ld do tRNA ProAGG e para o tRF-
 275 In do tRNA GluCTC. Esses tRFs foram selecionados para que houvesse a confirmação das
 276 predições de tRFs produzidos a partir tanto de tRNAs maduros como de tRNAs primários,
 277 sendo selecionados devido ao tamanho apropriado para o funcionamento dos *primers*, assim
 278 como por sua alta contagem de *reads* nas análises *in silico*. Para essa análise, foram utilizadas
 279 as três condições de cultivo para assegurar a detecção dos tRFs, visto que os mesmos
 280 apresentaram variação das contagens de *reads* nas diferentes bibliotecas. A partir dessa análise,
 281 foi possível identificar a expressão dos 3 tRFs, sendo que para o tRF do tRNA GluCTC foi visto
 282 um aumento significativo da expressão na condição de privação de zinco, sendo quatro vezes
 283 maior quando comparada à condição controle (Figura 9). A expressão diferencial do tRF em
 284 uma condição de estresse específica reforça a possibilidade de que os tRFs possam estar
 285 exercendo sua função de regulação na ausência de via de RNAi funcional.



286 **Figura 9: Análise dos níveis relativos de tRFs em *C. gattii* R265 nas diferentes condições**
 287 **de cultivo.** Gráfico do resultado da avaliação de expressão dos tRFs tRF-Ld(26) ProAGG, tRF-
 288 Ld(25) ProAGG e tRF-In GluCTC por ensaio de qRT-PCR nas condições controle, privação de

289 zinco (TPEN) e privação de ferro (BPDS). As barras representam as médias e desvios padrão
290 obtidos a partir de triplicatas técnicas e biológicas. Expressão relativa baseada em $2^{-\Delta CT}$ a partir
291 de dados normalizados por snRNA U6. Análise estatística por *one way ANOVA test*: (*) $p <$
292 0.05.

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314 4.5 Discussão

315 Apesar da disponibilidade de dados de sequenciamento de pequenos RNAs seja crescente
316 devido aos estudos de microRNAs, a descoberta dos tRFs é recente e poucas informações com
317 relação aos seus padrões de processamento se encontram na literatura, de forma que até então,
318 a metodologia de análise *in silico* de tRFs em organismos que não *Homo sapiens* era feita de
319 forma subjetiva, valendo-se apenas da contagem de *reads* e região de alinhamento dos
320 fragmentos como parâmetro de identificação [14,38,39]. Esse tipo de análise pode superestimar
321 a presença de tRFs, visto que não há um critério fixo que divide os tRFs de produtos de
322 degradação de tRNAs, assim como pode limitar sua identificação, uma vez que a procura por
323 tRFs ocorre apenas em posições do tRNA bem estabelecidas para a produção de tais moléculas
324 e ocasiona dificuldade na identificação de classes menos convencionais. Utilizando como base
325 os modelos estatísticos do servidor *tRF2Cancer* [10], nós fomos capazes de desenvolver uma
326 *pipeline* de análise de tRFs com suporte estatístico passível de utilização em qualquer
327 organismo. Por meio dessa *pipeline*, fomos capazes de identificar tRFs para 22 dos 39 tRNAs
328 preditos para a linhagem R265 de *C. gattii*. Apesar da identificação de tRFs, essa quantidade
329 de tRNAs produtores de tRFs é mais baixa comparado a outros organismos, como no camarão
330 *Triops cancriformis* [39], no qual 39 dos 45 tRNAs produzem tRFs, e na planta modelo
331 *Arabidopsis thaliana* [38], no qual 37 de 48 tRNAs produzem tRFs, indicando a dependência
332 da via de RNAi por alguns tRFs. Como já demonstrado, a estabilidade de microRNAs depende
333 de sua associação com proteínas argonauta [40], impedindo a atuação de nucleases
334 citoplasmáticas sobre os microRNAs, e a ausência de argonautas leva a perda da capacidade de
335 detecção de microRNAs por degradação dos mesmos. Assim, visto que existem indícios de
336 associação de tRFs com proteínas argonauta e que os mesmos partilham diversas características
337 com microRNAs, é possível que o mesmo mecanismo de proteção se faça presente, de forma
338 que tRFs que não apresentem associação com outras moléculas em vias de regulação

339 alternativas à via de RNAi, na ausência de argonautas, sejam rapidamente degradados e
340 incapazes de serem detectados. Da mesma forma, ainda é possível que as proteínas argonauta
341 estejam ligadas à produção de tRFs, assim como participam do processamento de alguns
342 microRNAs [41], visto que ainda não é completamente elucidada a forma de produção dos
343 tRFs.

344 É necessário, porém, ressaltar as limitações dessa comparação. Primeiramente, os organismos
345 comparados são evolutivamente distantes de *C. gattii*, além de que eucariotos superiores
346 apresentam maiores quantidades de RNAs não-codificantes quando comparados com
347 organismos de complexidade inferior [42], de forma que a diferença observada pode ocorrer,
348 em parte, devido a uma ocorrência naturalmente mais elevada de tRFs nessas classes de
349 organismos ao invés de devido à perda da funcionalidade da via de RNAi. Além disso, como
350 previamente citado, a maioria dos estudos de tRFs é executada sem apoio estatístico ou
351 restrições de classificação claras, de forma que a menor exigência dessa metodologia em
352 comparação com a utilizada no presente estudo pode ser responsável por parte dessa quantidade
353 reduzida de tRFs identificados. Dessa forma, seria de extremo interesse a análise de organismos
354 mais próximos por meio dessa *pipeline* afim de tornar mais precisa a avaliação do impacto da
355 perda da via de RNAi. Ainda assim, a identificação de tRFs invariavelmente fortalece a
356 hipótese de que os mesmos são capazes de atuar por meio de mecanismos alternativos à via de
357 RNAi.

358 Com relação às características dos tRFs identificados, também foi possível verificar diferenças
359 com os padrões da literatura. Grande parte dos tRFs identificados eram pertencentes às classes
360 tRF-3 e tRF-5, como o observado em organismos com via de RNAi funcional [14], mas a classe
361 que apresentou maior frequência foi a de tRF-In. Ainda que essa classe foi a mais recentemente
362 identificada e, devido a isso, a grande maioria dos estudos de tRFs não apresenta informações
363 com relação à presença dos mesmos, o fato de que os mesmos não foram identificados nesses

364 estudos sugere que a frequência desses tRFs seja extremamente inferior a das demais classes
365 nos organismos analisados. Assim, é possível considerar que essa alteração na dominância de
366 classes de tRFs esteja também associada à perda da via de RNAi, especialmente pelo fato de
367 que é demonstrado que tRFs-3 e tRFs-5 apresentam associação com argonauta. Além disso, não
368 foram identificados tRFs-1, mesmo que essa classe de tRF não tenha associação relevante com
369 argonautas. Apesar de que a ausência de argonautas não explique a falta de tRFs-1, foi
370 demonstrado que essa classe de tRF é extremamente variável em quantidade entre diferentes
371 organismos que apresentam ou não argonautas [14], devido a fatores ainda não identificados, o
372 que poderia explicar a ausência dos mesmos em *C. gattii*. De forma interessante, foram
373 identificados tRFs provenientes do 5' *leader* do transcrito de tRNA primário. Recentemente,
374 houve a identificação de tRFs originados a partir de regiões 5' *leader* em células tumorais [43],
375 entretanto, não havia nenhuma outra descrição na literatura de uma classe de tRFs produzida a
376 partir exclusivamente da região 5' *leader* até o presente momento, de forma que o achado
377 reforça a existência dessa nova classe, aqui denominada tRFs-Ld (tRFs-*leader*). Esses tRFs, de
378 forma similar aos tRFs-1, apresentam sua extremidade 3' no como o ponto de excisão do 5'
379 *leader* durante a maturação do tRNA, sugerindo que a provável forma de produção dos mesmos
380 seja por meio da atuação da RNase P [44], assim como a RNase Z é responsável pela produção
381 de tRFs-1 [8,15].

382 A identificação desses tRFs não usuais, assim como a grande quantidade de tRFs-I identificados
383 nas mais diversas porções dos tRNAs, ressalta a importância da adoção de uma metodologia de
384 identificação de tRFs que permita a busca com algum grau de confiança estatística em regiões
385 do tRNA que não apenas as principais descritas na literatura, uma vez que tRFs aparentam
386 serem produzidos a partir de virtualmente qualquer região do transcrito de tRNA.

387 Embora o padrão de distribuição de classes de tRFs tenha sido alterado com relação a
388 organismos que apresentam via de RNAi funcional, a distribuição de tamanho dos tRFs foi

389 extremamente similar. Quase a totalidade dos tRFs se apresentaram dentro do intervalo de 14 a
390 30 nucleotídeos, sendo possível observar um enriquecimento de tRFs nos tamanhos de 18 e 22
391 nucleotídeos, assim como é visto em outros organismos [14], sugerindo que o tamanho do tRF
392 não determina se o mesmo vai ou não atuar por meio da via de RNAi. Também foi identificado
393 um enriquecimento de tRFs no tamanho de 29 nucleotídeos, sendo sua maioria tRFs-I, o que
394 explica a ausência de detecção desse pico em estudos prévios.

395 As análises de qRT-PCR foram capazes não apenas de confirmar a predição dos tRFs por nossa
396 *pipeline*, como também identificaram a expressão diferencial de tRFs. A identificação de
397 eventos de expressão diferencial de tRFs em condições de estresse específicas reforça a ideia
398 de que os mesmos ainda são capazes de exercer suas funções mesmo na ausência da via de
399 RNAi. A avaliação de expressão, porém, foi executada apenas em 3 tRFs, e com relação a 2
400 condições estressoras, de forma que a real extensão dos eventos de expressão diferencial de
401 tRFs é incerta. Infelizmente, avaliações de expressão diferencial *in silico* não poderiam ser
402 executadas afim de avaliar os demais tRFs utilizando nossas bibliotecas de RNA-seq, devido a
403 limitações de nossos dados. Nossas bibliotecas foram geradas sem replicatas, de forma que não
404 apenas a grande maioria dos softwares de análises de expressão diferencial não conseguem
405 executar suas análises devido ao passo de normalização por replicatas, como também, mesmo
406 os softwares que o fazem, devido a perda da capacidade de avaliar a variação biológica das
407 amostras, não apresentam resultados confiáveis [45,46]. Assim, se tornam necessárias
408 avaliações posteriores para a identificação de eventos de expressão diferencial de tRFs em *C.*
409 *gattii*.

410

411

412

413

414 **4.6 Conclusão**

415

416 Com os resultados obtidos nesse estudo, foi observada a presença de tRFs e expressão
417 diferencial dos mesmos na ausência de funcionamento da via de RNA de interferência. Apesar
418 da identificação e expressão diferencial, os padrões e quantidade de tRFs identificados não são
419 iguais ao observado em organismos que apresentam essa via funcional. Dessa forma,
420 concluimos que, apesar de que alguns tRFs aparentam necessitar de uma via de RNA de
421 interferência funcional e associação a argonautas para o seu funcionamento, existem tRFs que
422 apresentam independência da via e sugerem a existência de formas alternativas de regulação da
423 expressão por sRNAs.

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439 4.7 Referências

- 440 1. Kapranov P, Cheng J, Dike S, Nix DA, Dutttagupta R, Willingham AT, et al. RNA maps
441 reveal new RNA classes and a possible function for pervasive transcription. *Science* (80-.).
442 American Association for the Advancement of Science; 2007;316:1484–8.
- 443 2. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev.*
444 *Genet.* Nature Publishing Group; 2004;5:522–31.
- 445 3. Wahid F, Shehzad A, Khan T, Kim YY. MicroRNAs: synthesis, mechanism, function, and
446 recent clinical trials. *Biochim. Biophys. Acta (BBA)-Molecular Cell Res.* Elsevier;
447 2010;1803:1231–43.
- 448 4. Wilson RC, Doudna JA. Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.*
449 *Annual Reviews*; 2013;42:217–39.
- 450 5. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. *Cell.*
451 Elsevier; 2009;136:642–55.
- 452 6. Thomson T, Lin H. The biogenesis and function of PIWI proteins and piRNAs: progress and
453 prospect. *Annu. Rev. Cell Dev. Annual Reviews*; 2009;25:355–76.
- 454 7. Daugaard I, Hansen TB. Biogenesis and Function of Ago-Associated RNAs. *Trends Genet.*
455 Elsevier; 2017;
- 456 8. Gebetsberger J, Polacek N. Slicing tRNAs to boost functional ncRNA diversity. *RNA Biol.*
457 Taylor & Francis; 2013;10:1798–806.
- 458 9. Keam SP, Hutvagner G. tRNA-derived fragments (tRFs): emerging new roles for an ancient
459 RNA in the regulation of gene expression. *Life. Multidisciplinary Digital Publishing Institute*;
460 2015;5:1638–51.
- 461 10. Zheng L-L, Xu W-L, Liu S, Sun W-J, Li J-H, Wu J, et al. tRF2Cancer: A web server to
462 detect tRNA-derived small RNA fragments (tRFs) and their expression in multiple cancers.
463 *Nucleic Acids Res.* Oxford University Press; 2016;44:W185–93.

- 464 11. Speer J, Gehrke CW, Kuo KC, Waalkes TP, Borek E. tRNA breakdown products as markers
465 for cancer. *Cancer*. Wiley Online Library; 1979;44:2120–3.
- 466 12. Borek E, Baliga BS, Gehrke CW, Kuo CW, Belman S, Troll W, et al. High turnover rate of
467 transfer RNA in tumor tissue. *Cancer Res. AACR*; 1977;37:3362–6.
- 468 13. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA
469 fragments (tRFs). *Genes Dev. Cold Spring Harbor Lab*; 2009;23:2639–49.
- 470 14. Kumar P, Anaya J, Mudunuri SB, Dutta A. Meta-analysis of tRNA derived RNA fragments
471 reveals that they are evolutionarily conserved and associate with AGO proteins to recognize
472 specific RNA targets. *BMC Biol. BioMed Central*; 2014;12:78.
- 473 15. Kumar P, Kuscu C, Dutta A. Biogenesis and function of transfer RNA-related fragments
474 (tRFs). *Trends Biochem. Sci. Elsevier*; 2016;41:679–89.
- 475 16. Gebetsberger J, Zywicki M, Künzi A, Polacek N. tRNA-derived fragments target the
476 ribosome and function as regulatory non-coding RNA in *Haloferax volcanii*. *Archaea. Hindawi*
477 *Publishing Corporation*; 2012;2012.
- 478 17. Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JT, et al. A 3' external
479 transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent
480 transcriptional noise. *Mol. Cell. Elsevier*; 2015;58:393–405.
- 481 18. Bąkowska-Żywicka K, Mleczko AM, Kasprzyk M, Machtel P, Żywicki M, Twardowski T.
482 The widespread occurrence of tRNA-derived fragments in *Saccharomyces cerevisiae*. *FEBS*
483 *Open Bio. Wiley Online Library*; 2016;6:1186–200.
- 484 19. Gottesman S. Micros for microbes: non-coding regulatory RNAs in bacteria. *TRENDS*
485 *Genet. Elsevier*; 2005;21:399–404.
- 486 20. Kwon-Chung KJ, Fraser JA, Doering TL, Wang ZA, Janbon G, Idnurm A, et al.
487 *Cryptococcus neoformans* and *Cryptococcus gattii*, the etiologic agents of cryptococcosis. *Cold*
488 *Spring Harb. Perspect. Med. Cold Spring Harbor Laboratory Press*; 2014;4:a019760.

- 489 21. Schmiedel Y, Zimmerli S. Common invasive fungal diseases: an overview of invasive
490 candidiasis, aspergillosis, cryptococcosis, and *Pneumocystis pneumonia*. *Swiss Med. Wkly.*
491 *EMH Schweizerischer Ärzteverlag*; 2016;146.
- 492 22. Feretzaki M, Billmyre RB, Clancey SA, Wang X, Heitman J. Gene network polymorphism
493 illuminates loss and retention of novel RNAi silencing components in the *Cryptococcus*
494 pathogenic species complex. *PLoS Genet. Public Library of Science*; 2016;12:e1005868.
- 495 23. Jiang N, Yang Y, Janbon G, Pan J, Zhu X. Identification and functional demonstration of
496 miRNAs in the fungus *Cryptococcus neoformans*. *PLoS One. Public Library of Science*;
497 2012;7:e52734.
- 498 24. Wang X, Hsueh Y-P, Li W, Floyd A, Skalsky R, Heitman J. Sex-induced silencing defends
499 the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev. Cold Spring Harbor Lab*;
500 2010;24:2566–82.
- 501 25. Ferrareze PAG, Streit RSA, dos Santos FM, Schrank A, Kmetzsch L, Vainstein MH, et al.
502 sRNAs as possible regulators of retrotransposon activity in *Cryptococcus gattii* VGII. *BMC*
503 *Genomics. BioMed Central*; 2017;18:294.
- 504 26. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI
505 GEO: archive for functional genomics data sets—update. *Nucleic Acids Res. Oxford University*
506 *Press*; 2012;41:D991–5.
- 507 27. Farrer RA, Desjardins CA, Sakthikumar S, Gujja S, Saif S, Zeng Q, et al. Genome evolution
508 and innovation across the four major lineages of *Cryptococcus gattii*. *MBio. Am Soc Microbiol*;
509 2015;6:e00868-15.
- 510 28. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA
511 genes in genomic sequence. *Nucleic Acids Res. Oxford Univ Press*; 1997;25:955–64.
- 512 29. Nawrocki EP, Kolbe DL, Eddy SR. Infernal 1.0: inference of RNA alignments.
513 *Bioinformatics. Oxford University Press*; 2009;25:1335–7.

- 514 30. Cros M-J, De Monte A, Mariette J, Bardou P, Grenier-Boley B, Gautheret D, et al.
515 RNAspace. org: An integrated environment for the prediction, annotation, and analysis of
516 ncRNA. *Rna*. Cold Spring Harbor Lab; 2011;17:1947–56.
- 517 31. Chan PP, Lowe TM. GtRNADB 2.0: an expanded database of transfer RNA genes identified
518 in complete and draft genomes. *Nucleic Acids Res*. Oxford University Press; 2016;44:D184–
519 9.
- 520 32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*.
521 *Nature Research*; 2012;9:357–9.
- 522 33. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence
523 alignment/map format and SAMtools. *Bioinformatics*. Oxford University Press; 2009;25:2078–
524 9.
- 525 34. Quinlan AR. BEDTools: the Swiss-army tool for genome feature analysis. *Curr. Protoc*.
526 *Bioinforma*. Wiley Online Library; 2014;11–2.
- 527 35. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-
528 performance genomics data visualization and exploration. *Brief. Bioinform*. Oxford University
529 Press; 2013;14:178–92.
- 530 36. Darty K, Denise A, Ponty Y. VARNA: Interactive drawing and editing of the RNA
531 secondary structure. *Bioinformatics*. Oxford Univ Press; 2009;btp250.
- 532 37. Czimmerer Z, Hulvely J, Simandi Z, Varallyay E, Havelda Z, Szabo E, et al. A versatile
533 method to design stem-loop primer-based quantitative PCR assays for detecting small
534 regulatory RNA molecules. *PLoS One*. Public Library of Science; 2013;8:e55168.
- 535 38. Loss-Morais G, Waterhouse PM, Margis R. Description of plant tRNA-derived RNA
536 fragments (tRFs) associated with argonaute and identification of their putative targets. *Biol*.
537 *Direct*. BioMed Central; 2013;8:6.
- 538 39. Hirose Y, Ikeda KT, Noro E, Hiraoka K, Tomita M, Kanai A. Precise mapping and

- 539 dynamics of tRNA-derived fragments (tRFs) in the development of *Triops cancriformis*
540 (tadpole shrimp). *BMC Genet. BioMed Central*; 2015;16:83.
- 541 40. Winter J, Diederichs S. Argonaute proteins regulate microRNA stability: Increased
542 microRNA abundance by Argonaute proteins is due to microRNA stabilization. *RNA Biol.*
543 *Taylor & Francis*; 2011;8:1149–57.
- 544 41. Cheloufi S, Dos Santos CO, Chong MMW, Hannon GJ. A dicer-independent miRNA
545 biogenesis pathway that requires Ago catalysis. *Nature. Nature Research*; 2010;465:584–9.
- 546 42. Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and
547 eukaryotic complexity. *Bioessays. Wiley Online Library*; 2007;29:288–99.
- 548 43. Olvedy M, Scaravilli M, Hoogstrate Y, Visakorpi T, Jenster G, Martens ES. A
549 comprehensive repertoire of tRNA-derived fragments in prostate cancer. *Oncotarget. Impact*
550 *Journals, LLC*; 2016;7:24766.
- 551 44. Frank DN, Pace NR. Ribonuclease P: unity and diversity in a tRNA processing ribozyme.
552 *Annual Reviews* 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA; 1998.
- 553 45. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
554 expression analysis of digital gene expression data. *Bioinformatics. Oxford University Press*;
555 2010;26:139–40.
- 556 46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
557 RNA-seq data with DESeq2. *Genome Biol. BioMed Central*; 2014;15:550.
- 558
- 559
- 560
- 561
- 562
- 563

564 **4.8 Arquivos Suplementares**

565

566 **Arquivo Suplementar 1: Código em linguagem python para a detecção de tRFs**

567

```
568 import os
```

```
569 from decimal import Decimal
```

570

```
571 fasta_file = raw_input("Enter tRNA's fasta file location: ")
```

```
572 samtools_file = raw_input("Enter total coverage file location: ")
```

```
573 bedtools_file = raw_input("Enter nucleotide coverage file location: ")
```

```
574 read_len = float(input("Read length: "))
```

```
575 read_len_int = int(read_len)
```

```
576 fasta = {}
```

```
577 outputname = "output_" + str(read_len_int) + ".txt"
```

```
578 output = open(outputname, 'w')
```

579

```
580 with open(fasta_file) as file_one:
```

```
581     for line in file_one:
```

```
582         line = line.strip()
```

```
583         if not line:
```

```
584             continue
```

```
585         if line.startswith(">"):
```

```
586             active_sequence_name = line[1:]
```

```
587             if active_sequence_name not in fasta:
```

```
588                 fasta[active_sequence_name] = []
```

```
589             continue
```

```
590             sequence = line
```



```
618         with open("temporary_file.txt", 'r') as count:
619             lines = count.readlines()
620             firstbase = 0
621             while firstbase < int(seqLen):
622                 basecount = lines[firstbase]
623                 firstbase += 1
624                 Psum = 0
625                 K = int(basecount)
626                 N = int(trnacov)
627                 R = int(read_len)
628                 L = int(seqLen)
629                 Q = float(Pr)
630                 while K <= N:
631                     result = 1
632                     for i in range(1, K+1):
633                         result = result * (N-i+1) / i
634                     P = Decimal(result)*(Decimal(Q)**(Decimal(K)))*((1-
635 Decimal(Q))**(Decimal(N)-Decimal(K)))
636                     Psum += P
637                     K += 1
638                 Pc = Psum*100
639                 if Pc < 1:
640                     print >>output, firstbase, int(basecount)
641                 else :
642                     continue
643
644 os.remove("temporary_file.txt")
645
```

5 PERSPECTIVAS

- Avaliar a presença de tRFs em *Cryptococcus neoformans* para a comparação com *Cryptococcus gattii*
- Avaliar o papel de tRFs de *C. gattii* no processo de regulação da expressão gênica empregando ensaios funcionais.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- BAKOWSKA-ŻYWICKA, Kamilla et al. The widespread occurrence of tRNA-derived fragments in *Saccharomyces cerevisiae*. **FEBS open bio**, [s. l.], v. 6, n. 12, p. 1186–1200, 2016.
- BARTEL, David P. MicroRNAs: genomics, biogenesis, mechanism, and function. **cell**, [s. l.], v. 116, n. 2, p. 281–297, 2004.
- BURGE, Christopher B. et al. Splicing of precursors to mRNAs by the spliceosomes. In: 2354 NUCLEIC ACIDS RESEARCH 1999, **Anais...** : Citeseer, 1999.
- BYRNES III, Edmond J. et al. Molecular evidence that the range of the Vancouver Island outbreak of *Cryptococcus gattii* infection has expanded into the Pacific Northwest in the United States. **The Journal of infectious diseases**, [s. l.], v. 199, n. 7, p. 1081–1086, 2009. a.
- BYRNES III, Edmond J. et al. First reported case of *Cryptococcus gattii* in the Southeastern USA: implications for travel-associated acquisition of an emerging pathogen. **PLoS One**, [s. l.], v. 4, n. 6, p. e5851, 2009. b.
- CARTHEW, Richard W.; SONTHEIMER, Erik J. Origins and mechanisms of miRNAs and siRNAs. **Cell**, [s. l.], v. 136, n. 4, p. 642–655, 2009.
- CAUSTON, Helen C. et al. Remodeling of yeast genome expression in response to environmental changes. **Molecular biology of the cell**, [s. l.], v. 12, n. 2, p. 323–337, 2001.
- DAUGAARD, Iben; HANSEN, Thomas Birkballe. Biogenesis and Function of Ago-Associated RNAs. **Trends in Genetics**, [s. l.], 2017.
- DAY, D. A.; TUIITE, Mick F. Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. **Journal of Endocrinology**, [s. l.], v. 157, n. 3, p. 361–371, 1998.
- DESNOYERS, Guillaume; BOUCHARD, Marie-Pier; MASSÉ, Eric. New insights into small RNA-dependent translational regulation in prokaryotes. **TRENDS in Genetics**, [s. l.], v. 29, n. 2, p. 92–98, 2013.
- DJEBALI, Sarah et al. Landscape of transcription in human cells. **Nature**, [s. l.], v. 489, n. 7414, p. 101–108, 2012.
- EDDY, Sean R. Non-coding RNA genes and the modern RNA world. **Nature Reviews Genetics**, [s. l.], v. 2, n. 12, p. 919–929, 2001.
- FARNHAM, Peggy J. Insights from genomic profiling of transcription factors. **Nature Reviews Genetics**, [s. l.], v. 10, n. 9, p. 605–616, 2009.
- FATICA, Alessandro; BOZZONI, Irene. Long non-coding RNAs: new players in cell differentiation and development. **Nature Reviews Genetics**, [s. l.], v. 15, n. 1, p. 7–21, 2014.
- FERETZAKI, Marianna et al. Gene network polymorphism illuminates loss and retention of

- novel RNAi silencing components in the *Cryptococcus* pathogenic species complex. **PLoS genetics**, [s. l.], v. 12, n. 3, p. e1005868, 2016.
- FRIEDMAN, Robin C. et al. Most mammalian mRNAs are conserved targets of microRNAs. **Genome research**, [s. l.], v. 19, n. 1, p. 92–105, 2009.
- GEBETSBERGER, Jennifer et al. tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Haloferax volcanii*. **Archaea**, [s. l.], v. 2012, 2012.
- GEBETSBERGER, Jennifer; POLACEK, Norbert. Slicing tRNAs to boost functional ncRNA diversity. **RNA biology**, [s. l.], v. 10, n. 12, p. 1798–1806, 2013.
- GONZALEZ-SANDOVAL, Adriana; GASSER, Susan M. On TADs and LADs: spatial control over gene expression. **Trends in Genetics**, [s. l.], v. 32, n. 8, p. 485–495, 2016.
- GOTTESMAN, Susan. Micros for microbes: non-coding regulatory RNAs in bacteria. **TRENDS in Genetics**, [s. l.], v. 21, n. 7, p. 399–404, 2005.
- GREWAL, Shiv I. S.; MOAZED, Danesh. Heterochromatin and epigenetic control of gene expression. **science**, [s. l.], v. 301, n. 5634, p. 798–802, 2003.
- HE, Lin; HANNON, Gregory J. MicroRNAs: small RNAs with a big role in gene regulation. **Nature Reviews Genetics**, [s. l.], v. 5, n. 7, p. 522–531, 2004.
- JACOB, François; MONOD, Jacques. Genetic regulatory mechanisms in the synthesis of proteins. **Journal of molecular biology**, [s. l.], v. 3, n. 3, p. 318–356, 1961.
- JIANG, Nan et al. Identification and functional demonstration of miRNAs in the fungus *Cryptococcus neoformans*. **PloS one**, [s. l.], v. 7, n. 12, p. e52734, 2012.
- KAPRANOV, Philipp et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. **Science**, [s. l.], v. 316, n. 5830, p. 1484–1488, 2007.
- KEAM, Simon P.; HUTVAGNER, Gyorgy. tRNA-derived fragments (tRFs): emerging new roles for an ancient RNA in the regulation of gene expression. **Life**, [s. l.], v. 5, n. 4, p. 1638–1651, 2015.
- KIDD, S. E. et al. A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 101, n. 49, p. 17258–17263, 2004.
- KIM, Young-Kook; KIM, V.Narry. Processing of intronic microRNAs. **The EMBO journal**, [s. l.], v. 26, n. 3, p. 775–783, 2007.
- KUMAR, Pankaj et al. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. **BMC biology**, [s. l.], v. 12, n. 1, p. 78, 2014.
- KUMAR, Pankaj; KUSCU, Canan; DUTTA, Anindya. Biogenesis and function of transfer

- RNA-related fragments (tRFs). **Trends in biochemical sciences**, [s. l.], v. 41, n. 8, p. 679–689, 2016.
- KWON-CHUNG, Kyung J. et al. Cryptococcus neoformans and Cryptococcus gattii, the etiologic agents of cryptococcosis. **Cold Spring Harbor perspectives in medicine**, [s. l.], v. 4, n. 7, p. a019760, 2014.
- LEE, Rosalind C.; FEINBAUM, Rhonda L.; AMBROS, Victor. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, [s. l.], v. 75, n. 5, p. 843–854, 1993.
- LEE, Tong Ihn; YOUNG, Richard A. Transcriptional regulation and its misregulation in disease. **Cell**, [s. l.], v. 152, n. 6, p. 1237–1251, 2013.
- LIN, Xiaorong; HEITMAN, Joseph. The biology of the *Cryptococcus neoformans* species complex. **Annu. Rev. Microbiol.**, [s. l.], v. 60, p. 69–105, 2006.
- LÓPEZ-MAURY, Luis; MARGUERAT, Samuel; BÄHLER, Jürg. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. **Nature Reviews Genetics**, [s. l.], v. 9, n. 8, p. 583–593, 2008.
- MATERA, A.Gregory; TERNS, Rebecca M.; TERNS, Michael P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. **Nature reviews Molecular cell biology**, [s. l.], v. 8, n. 3, p. 209–220, 2007.
- MATTICK, John S.; MAKUNIN, Igor V. Non-coding RNA. **Human molecular genetics**, [s. l.], v. 15, n. suppl_1, p. R17–R29, 2006.
- MERCER, Tim R.; DINGER, Marcel E.; MATTICK, John S. Long non-coding RNAs: insights into functions. **Nature Reviews Genetics**, [s. l.], v. 10, n. 3, p. 155–159, 2009.
- QUINN, Jeffrey J.; CHANG, Howard Y. Unique features of long non-coding RNA biogenesis and function. **Nature Reviews Genetics**, [s. l.], v. 17, n. 1, p. 47–62, 2016.
- RAJASINGHAM, Radha et al. Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. **The Lancet Infectious Diseases**, [s. l.], 2017.
- SCHMIEDEL, Yvonne; ZIMMERLI, Stephan. Common invasive fungal diseases: an overview of invasive candidiasis, aspergillosis, cryptococcosis, and *Pneumocystis pneumonia*. **Swiss medical weekly**, [s. l.], v. 146, 2016.
- SCHOENBERG, Daniel R.; MAQUAT, Lynne E. Regulation of cytoplasmic mRNA decay. **Nature Reviews Genetics**, [s. l.], v. 13, n. 4, p. 246–259, 2012.
- SHABALINA, Svetlana A.; KOONIN, Eugene V. Origins and evolution of eukaryotic RNA interference. **Trends in ecology & evolution**, [s. l.], v. 23, n. 10, p. 578–587, 2008.
- SLEZAK-PROCHAZKA, Izabella et al. Cellular localization and processing of primary

- transcripts of exonic microRNAs. **PloS one**, [s. l.], v. 8, n. 9, p. e76647, 2013.
- STAMM, Stefan et al. Function of alternative splicing. **Gene**, [s. l.], v. 344, p. 1–20, 2005.
- STEFANI, Giovanni; SLACK, Frank J. Small non-coding RNAs in animal development. **Nature reviews Molecular cell biology**, [s. l.], v. 9, n. 3, p. 219–230, 2008.
- THOMSON, Travis; LIN, Haifan. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. **Annual Review of Cell and Developmental**, [s. l.], v. 25, p. 355–376, 2009.
- WAHID, Fazli et al. MicroRNAs: synthesis, mechanism, function, and recent clinical trials. **Biochimica et Biophysica Acta (BBA)-Molecular Cell Research**, [s. l.], v. 1803, n. 11, p. 1231–1243, 2010.
- WANG, Xuying et al. Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. **Genes & development**, [s. l.], v. 24, n. 22, p. 2566–2582, 2010.
- WILSON, Ross C.; DOUDNA, Jennifer A. Molecular mechanisms of RNA interference. **Annual review of biophysics**, [s. l.], v. 42, p. 217–239, 2013.
- ZHANG, Xue et al. IL-4 inhibits the biogenesis of an epigenetically suppressive PIWI-interacting RNA to upregulate CD1a molecules on monocytes/dendritic cells. **The Journal of Immunology**, [s. l.], v. 196, n. 4, p. 1591–1603, 2016.
- ZHENG, Ling-Ling et al. tRF2Cancer: A web server to detect tRNA-derived small RNA fragments (tRFs) and their expression in multiple cancers. **Nucleic acids research**, [s. l.], v. 44, n. W1, p. W185–W193, 2016.

7 ANEXO A – NORMAS DE PUBLICAÇÃO NA REVISTA BMC Genomics

Criteria

Research articles should report on original primary research, but may report on systematic reviews of published research provided they adhere to the appropriate reporting guidelines which are detailed in our [editorial policies](#). Please note that non-commissioned pooled analyses of selected published research will not be considered.

BMC Genomics strongly encourages that all datasets on which the conclusions of the paper rely should be available to readers. We encourage authors to ensure that their datasets are either deposited in publicly available repositories (where available and appropriate) or presented in the main manuscript or additional supporting files whenever possible. Please see Springer Nature's [information on recommended repositories](#). Where a widely established research community expectation for data archiving in public repositories exists, submission to a community-endorsed, public repository is mandatory. A list of data where deposition is required, with the appropriate repositories, can be found on the Editorial [Policies Page](#).

Preparing your manuscript

The information below details the section headings that you should include in your manuscript and what information should be within each section.

Please note that your manuscript must include a 'Declarations' section including all of the subheadings (please see below for more information).

Title page

The title page should:

- present a title that includes, if appropriate, the study design
- list the full names, institutional addresses and email addresses for all authors
 - if a collaboration group should be listed as an author, please list the Group name as an author. If you would like the names of the individual members of the Group to be searchable through their individual PubMed records, please include this information in the “Acknowledgements” section in accordance with the instructions below
- indicate the corresponding author

Abstract

The Abstract should not exceed 350 words. Please minimize the use of abbreviations and do not cite references in the abstract. The abstract must include the following separate sections:

- **Background:** the context and purpose of the study
- **Results:** the main findings
- **Conclusions:** a brief summary and potential implications

Keywords

Three to ten keywords representing the main content of the article.

Background

The Background section should explain the background to the study, its aims, a summary of the existing literature and why this study was necessary.

Results

This should include the findings of the study including, if appropriate, results of statistical analysis which must be included either in the text or as tables and figures.

Discussion

For research articles this section should discuss the implications of the findings in context of existing research and highlight limitations of the study. For study protocols and methodology manuscripts this section should include a discussion of any practical or operational issues involved in performing the study and any issues not covered in other sections.

Conclusions

This should state clearly the main conclusions and provide an explanation of the importance and relevance of the study to the field.

Methods (can also be placed after Background)

The methods section should include:

- the aim, design and setting of the study
- the characteristics of participants or description of materials
- a clear description of all processes, interventions and comparisons. Generic names should generally be used. When proprietary brands are used in research, include the brand names in parentheses
- the type of statistical analysis used, including a power calculation if appropriate

List of abbreviations

If abbreviations are used in the text they should be defined in the text at first use, and a list of abbreviations can be provided.

Declarations

All manuscripts must contain the following sections under the heading 'Declarations':

- Ethics approval and consent to participate
- Consent for publication
- Availability of data and material
- Competing interests

- Funding
- Authors' contributions
- Acknowledgements
- Authors' information (optional)

Please see below for details on the information to be included in these sections.

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Ethics approval and consent to participate

Manuscripts reporting studies involving human participants, human data or human tissue must:

- include a statement on ethics approval and consent (even where the need for approval was waived)
- include the name of the ethics committee that approved the study and the committee's reference number if appropriate

Studies involving animals must include a statement on ethics approval.

See our [editorial policies](#) for more information.

If your manuscript does not report on or involve the use of any animal or human data or tissue, please state “Not applicable” in this section.

Consent for publication

If your manuscript contains any individual person's data in any form (including any individual details, images or videos), consent for publication must be obtained from that person, or in the case of children, their parent or legal guardian. All presentations of case reports must have consent for publication.

You can use your institutional consent form or our [consent form](#) if you prefer. You should not send the form to us on submission, but we may request to see a copy at any stage (including after publication).

See our [editorial policies](#) for more information on consent for publication.

If your manuscript does not contain data from any individual person, please state “Not applicable” in this section.

Availability of data and materials

All manuscripts must include an 'Availability of data and materials' statement. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. By data we mean the minimal dataset that would be necessary to interpret, replicate and build upon the findings reported in

the article. We recognise it is not always possible to share research data publicly, for instance when individual privacy could be compromised, and in such instances data availability should still be stated in the manuscript along with any conditions for access.

Data availability statements can take one of the following forms (or a combination of more than one if required for multiple datasets):

- The datasets generated and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]
- The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
- All data generated or analysed during this study are included in this published article [and its supplementary information files].
- The datasets generated and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.
- Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.
- The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name].
- Not applicable. If your manuscript does not contain any data, please state 'Not applicable' in this section.

More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available [here](#).

BioMed Central also requires that authors cite any publicly available data on which the conclusions of the paper rely in the manuscript. Data citations should include a persistent identifier (such as a DOI) and should ideally be included in the reference list. Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style. Dataset identifiers including DOIs should be expressed as full URLs. For example:

Hao Z, AghaKouchak A, Nakhjiri N, Farahmand A. Global integrated drought monitoring and prediction system (GIDMaPS) data sets. figshare. 2014.
<http://dx.doi.org/10.6084/m9.figshare.853801>

With the corresponding text in the Availability of data and materials statement:

The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS].^[Reference number]

Competing interests

All financial and non-financial competing interests must be declared in this section.

See our [editorial policies](#) for a full explanation of competing interests. If you are unsure whether you or any of your co-authors have a competing interest please contact the editorial office.

Please use the authors initials to refer to each authors' competing interests in this section.

If you do not have any competing interests, please state "The authors declare that they have no competing interests" in this section.

Funding

All sources of funding for the research reported should be declared. The role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript should be declared.

Authors' contributions

The individual contributions of authors to the manuscript should be specified in this section. Guidance and criteria for authorship can be found in our [editorial policies](#).

Please use initials to refer to each author's contribution in this section, for example: "FC analyzed and interpreted the patient data regarding the hematological disease and the transplant. RH performed the histological examination of the kidney, and was a major contributor in writing the manuscript. All authors read and approved the final manuscript."

Acknowledgements

Please acknowledge anyone who contributed towards the article who does not meet the criteria for authorship including anyone who provided professional writing services or materials.

Authors should obtain permission to acknowledge from all those mentioned in the Acknowledgements section.

See our [editorial policies](#) for a full explanation of acknowledgements and authorship criteria.

If you do not have anyone to acknowledge, please write "Not applicable" in this section.

Group authorship (for manuscripts involving a collaboration group): if you would like the names of the individual members of a collaboration Group to be searchable through their individual PubMed records, please ensure that the title of the collaboration Group is included on the title page and in the submission system and also include collaborating author names as the last paragraph of the "Acknowledgements" section. Please add authors in the format First Name, Middle initial(s) (optional), Last Name. You can add institution or country information for each author if you wish, but this should be consistent across all authors.

Please note that individual names may not be present in the PubMed record at the time a published article is initially included in PubMed as it takes PubMed additional time to code this information.

Authors' information

This section is optional.

You may choose to use this section to include any relevant information about the author(s) that may aid the reader's interpretation of the article, and understand the standpoint of the author(s). This may include details about the authors' qualifications, current positions they hold at institutions or societies, or any other relevant background information. Please refer to authors using their initials. Note this section should not be used to describe any competing interests.

Endnotes

Endnotes should be designated within the text using a superscript lowercase letter and all notes (along with their corresponding letter) should be included in the Endnotes section. Please format this section in a paragraph rather than a list.

References

All references, including URLs, must be numbered consecutively, in square brackets, in the order in which they are cited in the text, followed by any in tables or legends. The reference numbers must be finalized and the reference list fully formatted before submission.

Examples of the BioMed Central reference style are shown below. Please ensure that the reference style is followed precisely.

See our editorial policies for author guidance on good citation practice.

Web links and URLs: All web links and URLs, including links to the authors' own websites, should be given a reference number and included in the reference list rather than within the text of the manuscript. They should be provided in full, including both the title of the site and the URL, as well as the date the site was accessed, in the following format: The Mouse Tumor Biology Database. <http://tumor.informatics.jax.org/mtbwi/index.do>. Accessed 20 May 2013. If an author or group of authors can clearly be associated with a web link (e.g. for blogs) they should be included in the reference.

Example reference style:

Article within a journal

Smith JJ. The world of science. *Am J Sci.* 1999;36:234-5.

Article within a journal (no page numbers)

Rohrmann S, Overvad K, Bueno-de-Mesquita HB, Jakobsen MU, Egeberg R, Tjønneland A, et al. Meat consumption and mortality - results from the European Prospective Investigation into Cancer and Nutrition. *BMC Med.* 2013;11:63.

Article within a journal by DOI

Slifka MK, Whitton JL. Clinical implications of dysregulated cytokine production. *Dig J Mol Med*. 2000; doi:10.1007/s801090000086.

Article within a journal supplement

Frumin AM, Nussbaum J, Esposito M. Functional asplenia: demonstration of splenic activity by bone marrow scan. *Blood* 1979;59 Suppl 1:26-32.

Book chapter, or an article within a book

Wyllie AH, Kerr JFR, Currie AR. Cell death: the significance of apoptosis. In: Bourne GH, Danielli JF, Jeon KW, editors. *International review of cytology*. London: Academic; 1980. p. 251-306.

OnlineFirst chapter in a series (without a volume designation but with a DOI)

Saito Y, Hyuga H. Rate equation approaches to amplification of enantiomeric excess and chiral symmetry breaking. *Top Curr Chem*. 2007. doi:10.1007/128_2006_108.

Complete book, authored

Blenkinsopp A, Paxton P. *Symptoms in the pharmacy: a guide to the management of common illness*. 3rd ed. Oxford: Blackwell Science; 1998.

Online document

Doe J. Title of subordinate document. In: *The dictionary of substances and their effects*. Royal Society of Chemistry. 1999. [http://www.rsc.org/dose/title of subordinate document](http://www.rsc.org/dose/title%20of%20subordinate%20document). Accessed 15 Jan 1999.

Online database

Healthwise Knowledgebase. *US Pharmacopeia*, Rockville. 1998. <http://www.healthwise.org>. Accessed 21 Sept 1998.

Supplementary material/private homepage

Doe J. Title of supplementary material. 2000. <http://www.privatehomepage.com>. Accessed 22 Feb 2000.

University site

Doe, J: Title of preprint. <http://www.uni-heidelberg.de/mydata.html> (1999). Accessed 25 Dec 1999.

FTP site

Doe, J: Trivial HTTP, RFC2169. <ftp://ftp.isi.edu/in-notes/rfc2169.txt> (1999). Accessed 12 Nov 1999.

Organization site

ISSN International Centre: The ISSN register. <http://www.issn.org> (2006). Accessed 20 Feb 2007.

Dataset with persistent identifier

Zheng L-Y, Guo X-S, He B, Sun L-J, Peng Y, Dong S-S, et al. Genome data from sweet and grain sorghum (*Sorghum bicolor*). *GigaScience Database*. 2011. <http://dx.doi.org/10.5524/100012>.

Figures, tables additional files

See [General formatting guidelines](#) for information on how to format figures, tables and additional files.