UNIVERSIDADE DE MOGI DAS CRUZES
Alexandre Santos Simeone

PROSPECÇÃO E CARACTERIZAÇÃO *IN SILICO* DE RNAS LONGOS NÃO CODIFICANTES (lncRNAs) EM DADOS PÚBLICOS DE TRANSCRIPTOMAS DE FUNGOS PATOGÊNICOS HUMANOS

Orientador: Prof. Dr. Fabiano Bezerra Menegidio

Mogi das Cruzes, SP
2023

UNIVERSIDADE DE MOGI DAS CRUZES
Alexandre Santos Simeone

PROSPECÇÃO E CARACTERIZAÇÃO *IN SILICO* DE RNAS LONGOS NÃO CODIFICANTES (lncRNAs) EM DADOS PÚBLICOS DE TRANSCRIPTOMAS DE FUNGOS PATOGÊNICOS HUMANOS

Dissertação apresentada ao Programa de Pós-Graduação da Universidade de Mogi das Cruzes (UMC) como parte dos requisitos para a obtenção do grau de Mestre em Biotecnologia.

Área de Concentração: Biotecnologia aplicada a Recursos Naturais e Agronegócios (Genômica e Bioinformática).

Orientador: Prof. Dr. Fabiano Bezerra Menegidio

Mogi das Cruzes, SP
2023

FINANCIAMENTO



Ficha catalográfica
Universidade de Mogi das Cruzes - Biblioteca Central

S589p

Simeone, Alexandre Santos
Prospecção e caracterização in silico de RNAS longos não codificantes (incRNAs) em dados públicos de transcriptomas de fungos patogênicos humanos / Alexandre Santos Simeone. -- Mogi das Cruzes, SP: Universidade de Mogi das Cruzes, 2023.
120 f. : il.

Dissertação (Mestrado em Biotecnologia) -- Universidade de Mogi das Cruzes, 2023.
Orientação: Prof. Dr. Fabiano Bezerra Menegidio.

1. RNA longo não codificante 2. Transcriptoma/genética 3. Fungos I. Menegidio, Fabiano Bezerra II. Universidade de Mogi das Cruzes III. Título

CDD 579.5

Elaborada por Maria Cristina Alves Silva Xavier – CRB-8/10077

ATAS

ATA DA SESSÃO PÚBLICA DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO EM BIOTECNOLOGIA DA UNIVERSIDADE DE MOGI DAS CRUZES

Às quatorze horas do dia trinta de janeiro de dois mil e vinte e três, pela Universidade de Mogi das Cruzes, realizou-se a defesa de dissertação intitulada "PROSPECÇÃO E CARACTERIZAÇÃO *IN SILICO* DE RNAs LONGOS NÃO CODIFICANTES (lncRNAs) EM DADOS PÚBLICOS DE TRANSCRIPTOMAS DE FUNGOS PATOGÊNICOS HUMANOS", para obtenção do grau de mestre pelo(a) candidato(a) **Alexandre Santos Simeone**. Tendo sido o número de créditos alcançados pelo(a) mesmo(a) no total de 56 (cinquenta e seis), a saber: 32 unidades de crédito em disciplinas de pós-graduação e 24 unidades de crédito no preparo da dissertação, o(a) aluno(a) perfaz assim os requisitos para obtenção do grau de mestre. A comissão examinadora estava constituída dos senhores professores doutores Fabiano Bezerra Menegidio, e Mauricio Ferreira Marcondes Machado, da Universidade de Mogi das Cruzes, e Karine Frehner Kavalco, da Universidade Federal de Viçosa. Sob a presidência do primeiro, como orientador da dissertação. A sessão pública da defesa da dissertação foi aberta pelo senhor presidente da comissão que apresentou o(a) candidato(a). Em seguida o(a) candidato(a) realizou uma apresentação oral da dissertação. Ao final da apresentação da dissertação, seguiram-se as arguições pelos membros da comissão examinadora. A seguir a comissão, em sessão secreta, conforme julgamento discriminado por cada membro, considerou o(a) candidato(a)

Aprovado (aprovado(a) /reprovado(a)) por Unanimidade (unanimidade/maioria)

Mogi das Cruzes, 30 de janeiro de 2023.

Comissão Examinadora	Julgamento
Profª Drª Karine Frehner Kavalco	Aprovado (aprovado(a) / reprovado(a))
Prof. Dr. Mauricio Ferreira Marcondes Machado	Aprovado (aprovado(a) / reprovado(a))
Prof. Dr. Fabiano Bezerra Menegidio	Aprovado (aprovado(a) / reprovado(a))

gov.br Documento assinado digitalmente
FABIANO BEZERRA MENEGIDIO
Data: 09/02/2023 11:54:53-0300
Verifique em https://verificador.iti.br

DEDICATÓRIA

Para Cristina, Sávio, Maria, Luiz Enrique, Marcely, Lucas e Diego. Por testemunharem o que ninguém neste mestrado viu, me oferecer uma oportunidade e ter acreditado em mim.

AGRADECIMENTOS

Não há outra forma de iniciar esta seção sem agradecer primeiramente ao meu orientador, Professor Doutor Fabiano Bezerra Menegidio. Pela extraordinária paciência, oportunidade, ter me protegido, me apresentado e introduzido a área da Bioinformática, o mundo Linux e por tantas semelhanças que possuímos. Também agradeço ao Doutor David Aciole Barbosa, cuja importância foi fugaz para a realização deste trabalho, seus resultados e, sem a menor dúvida, é uma das pessoas mais inteligentes e dinâmicas e com quem pude conhecer e conviver. Obrigado por sempre ter algo a ensinar, a somar e por sempre me inspirarem. Meus caros, foi uma honra iniciar com vocês o que hoje é o LaBiOmicS, carregarei para sempre e com todo carinho esta vivência.

Às minhas amigas e companheiras de laboratório, Pâmela Correa, Rafaela Campos, Ana Humberto e Yara de Lima por terem feito com que a jornada se tornasse mais leve, divertida e encerrassem o silêncio e a solidão que eu sentia no início de todo o trabalho, coisas pelas quais eu não consigo me acostumar. Nunca se esqueçam: vocês merecem e têm o que é preciso para estar aqui e chegar aonde chegaram. Fico incomensuravelmente feliz em presenciar sua trajetória e sei o quanto vocês ainda irão contribuir com o laboratório e onde mais queiram estar. Pâmela, aliás, a Duda mal nasceu e já tem motivos de sobra para se orgulhar. Também não posso deixar de agradecer ao novo amigo e camarada Gledison Teixeira, que, assim como eu e nosso orientador, resolveu se reinventar e adentrar ao mundo da Ciência, buscando uma jornada menos atribulada e ainda mais gratificante. Querido, obrigado pela paciência, disposição e companhia. Você tem um futuro ainda mais brilhante neste lugar.

Aos pesquisadores chefes do Núcleo Integrado de Biotecnologia, Luiz Roberto Nunes e Daniela Leite Jabes, muito obrigado por permitirem que eu estivesse presente em seu laboratório e pudesse ter a oportunidade de ampliar meu campo de trabalho e conhecimento dentro da Biologia. Mesmo sabendo de onde vim, só posso agradecer por tudo que aprendi, mostrar que ainda tenho muito o que melhorar e tudo o que a Biotecnologia pode proporcionar.

Aos integrantes do LAGOOA, Leticia, Maíra, Michelle, Kenneth e Danilo, obrigado por sempre terem sido um ponto de apoio, descontração e gastronomia durante o gradual retorno à rotina presencial, por ser uma ponte de contato entre nossos laboratórios e por sempre lembrarem que antes de cientistas, somos seres humanos.

Aos meus novos amigos do Núcleo de Pesquisas Tecnológicas, Matheus e Luan, pela sempre divertidíssima convivência, disposição em ajudar e terem me salvado no trabalho do prof. Dr. Ricardo. Quero levar sempre vocês nesta jornada da vida e foi um prazer ser cobaia de seus experimentos.

Para os Drs. Alisson Chaves e Francisco Calaça, pelo entusiasmo e carinho sempre presentes quando surgiam dúvidas a respeito da micologia clínica e ambiental.

À minha terapeuta Silvana Vasconcelos, minha prima Dra. Fernanda Romeiro Dias e minha tia, Sueli Romeiro um agradecimento especial. Pela importância em me auxiliar a olhar o mundo sobre uma perspectiva completamente diferente da qual fui condicionado, como retirar de meus ombros as cargas que não são minhas, que não estou só e o quão importante a vida é. Assim como para você, que está lendo esta dissertação.

Para todos os meus alunos dos colégios, Gal. Humberto de Souza Mello, Giordano, Santa Bárbara e Beka, por sempre me instigarem, me incentivarem e sempre me ensinar, tal qual, tento fazer todos os dias com vocês. Obrigado pela paciência e perdão por nestes 2 anos não poder estar integralmente com vocês.

Aos meus amigos, Marcos Vinícius ("Marcão"), Teo, Anna Pessotti, Patrícia Tavares ("Paty"), Karine Leopoldina, Luiz e Leonardo Szajnbok e Caio Prendini por estarem aqui desde muito antes da biologia e foram fundamentais para que eu pudesse retornar do momento mais sombrio de minha vida. Não há um único dia que eu não agradeça por continuar aqui para abraçá-los mais uma vez e continuar criando mais recordações de amor, carinho, diversão e parceria. Também preciso agradecer à Bruna Correia Lima e sua família, por durante 3 anos terem oferecido seu lar e apoio durante o mestrado. Sou grato por somente tudo de positivo que aconteceu. À Maria do Carmo, Gustavo e Luiz Enrique, minha família. Sinto falta do que éramos, me devaneio imaginando o que poderíamos ter sido, mas a cada dia busco aceitar e assimilar o que nos tornamos. Embora nossa separação, o que nunca mudou foi meu amor e gratidão por vocês. Sem sua presença e apoio, nada teria dado certo e Gu, um agradecimento especial a você por ter me cedido seu computador com capacidade muito superior ao meu em diversos momentos que precisei me livrar de enrascadas. Obrigado por absolutamente tudo. Eu amo vocês.

À Marcely Silva, quem aceitou ser minha companheira de vida. Tanto em comum e tantas diferenças, até mesmo naquilo que me afligiu por 2 anos – algo que você também passou. Desde que chegou, tudo se tornou melhor, mais leve e ainda mais belo. Obrigado pela presença, disposição, paciência, cumplicidade, parceria, apoio, pelo seu lar, por ter me

transformado também em um “pai pet”, pelas “bruxarias” e por tanta descontração quando estamos juntos. Eu te amo.

À Universidade de Mogi das Cruzes por proporcionar toda a estrutura necessária para realização da minha pesquisa e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela bolsa de auxílio durante o período vigente de meu mestrado.

Meus sinceros e eternos agradecimentos!

“Seu pai me disse uma vez... nunca vou me esquecer...
Que a pesquisa científica consome a vida de uma pessoa. Ele me disse: ‘No meu leito de morte, quero me lembrar de vocês, não de um laboratório frio e asséptico.’”.

— Matéria Escura.

RESUMO

O presente projeto tem como objetivo a prospecção e caracterização *in silico* de RNAs longos não codificantes (lncRNAs) em dados públicos de transcriptomas de 15 espécies de fungos patogênicos humanos. A compreensão da regulação gênica em fungos patogênicos é crucial para o desenvolvimento de terapias eficazes, e os lncRNAs são uma classe emergente de moléculas reguladoras cuja função é ainda pouco compreendida nestes organismos. Para alcançar este objetivo, os dados de transcriptoma foram obtidos através *do Transcriptome Shotgun Assembly Sequence Database* (TSA) e submetidos a diversas ferramentas de bioinformática para avaliar o potencial codificador dos transcritos e anotá-los funcionalmente. Este projeto resultou na identificação de 29.746 lncRNAs em 15 espécies de fungos patogênicos humanos, muitos dos quais ainda não haviam sido caracterizados previamente. Estes dados fornecem uma base valiosa para futuras investigações sobre a regulação gênica em fungos patogênicos e podem ser utilizados para o desenvolvimento de novas terapias para infecções fúngicas. Além disso, a disponibilidade de dados públicos foi essencial para a realização deste projeto e contribuiu para a construção de novos conhecimentos na área da micologia clínica.

Palavras-chave: *lncRNA; RNASeq; transcriptoma; fungos patogênicos; bioinformática*.

ABSTRACT

The present project aims to prospect and in silico characterize long noncoding RNAs (lncRNAs) in public transcriptome data of 15 species of human pathogenic fungi. Understanding gene regulation in pathogenic fungi is crucial for the development of effective therapies, and lncRNAs are an emerging class of regulatory molecules whose function is still poorly understood in these organisms. To achieve this objective, transcriptome data was obtained from the Transcriptome Shotgun Assembly Sequence Database (TSA) and submitted to various bioinformatics tools to evaluate the coding potential of transcripts and functionally annotate them. This project resulted in the identification of 29746 lncRNAs in 15 species of human pathogenic fungi, many of which had not been previously characterized. These data provide a valuable foundation for future investigations into gene regulation in pathogenic fungi and can be used for the development of new therapies for fungal infections. Additionally, the availability of public data was essential for the completion of this project and contributed to the building of new knowledge in the field of clinical mycology.

Keywords: *lncRNA; RNASeq; transcriptome; pathogenic fungi; bioinformatics.*

LISTA DE ILUSTRAÇÕES

Figura 1	Diversidade do Reino Fungi	21
Figura 2	Esquematização dos tipos de ncRNAs	38
Figura 3	Fluxo de trabalho para prospecção de lncRNAs	61
Figura 4	Comparativo de CP com base no programa CPAT	69
Figura 5	Comparativo de CP com base no programa CPC2	72
Figura 6	Comparativo de CP com base no programa RNASamba	75
Figura 7	Distribuição dos transcritos como não-codificantes	77
Figura 8	Distribuição dos transcritos como codificantes	78
Figura 9	Distribuição de transcritos anotados no Rfam	81
Figura 10	Proporções dos transcritos anotados funcionalmente	83

LISTA DE TABELAS

Tabela 1	Métodos *in silico* para prospecção de lncRNA	54
Tabela 2	Transcriptomas disponíveis na base de dados TSA	63
Tabela 3	Metadados de sequenciamento e montagem dos transcriptomas	64
Tabela 4	Informações biomoleculares das espécies de fungos patogênicos	66
Tabela 5	Comparativo de CP com base no programa CPAT	69
Tabela 6	Comparativo de CP com base no programa CPC2	71
Tabela 7	Comparativo de CP com base no programa RNASamba	73
Tabela 8	Distribuição dos transcritos com base em seu potencial codificante	76
Tabela 9	Distribuição das anotações funcionais do EnTAP	82

LISTA DE ABREVIATURAS E SIGLAS

ABI	*Application Binary Interface*
ACM	*Association for Computing Machinery*
Adh	*Alcohol dehydrogenase*
API	*Application Programming Interface*
bd	*Banco de dados*
ADAMTS9-aS2	*ADAMTS9 antisense RNA 2*
AIDS	*Síndrome da Imunodeficiência Adquirida*
ANRIL	*Antisense RNA in the INK4 locus*
CCAT2	*Colon cancer associated transcript 2*
CDKN1A	*Cyclin dependent kinase inhibitor 1A*
CFLAR-AS1	*CFLAR antisense RNA 1*
CLI	*Command Line Interface*
COLDAIR	*Cold assisted intronic noncoding RNA*
CP	*Potencial Codificante*
CPAT	*Cdoing Potential Assessment Tool*
CPC2	*Coding Potential Calculator 2*
CM	*Modelo de Covariância*
DHFR	*Dihidrofolato redutase*
eggNOG	*Evolutionary genealogy of genes: Non-supervised Orthologous Groups*
elncRNAs	*Enhancer-like lncRNAs*
ENCODE	*Encyclopedia of DNA Elements*
EnTAP	*Eukaryotic Non-Model Transcriptome Annotation Pipeline*
FDA	*Food and Drug Administration*
FEELnc	*FlExible Extraction of LncRNAS*
FLC	*Flowering locus c*
FN	*False Negative*
FP	*False Positive*
GAS5	*Growth arrest specific 5*
GAS6-A61	*GAS6 antisense RNA 1*
H3K27me3	*Processo de trimetilação da lisina em histonas*
HCC	*Carcinoma hepatocelular*

HIV	*Human Imunodeficiency Virus*
HOTAIR	*HOX transcript antisense RNA*
HOTTIP	*HOXA distal transcript antisense RNA*
HOX	*Homeobox*
INFERNAL	*Inference of RNA Alignments*
InterPro	*The InterPro Consortium*
IRES	*Internal ribosome entry site*
KHPS1	*SPHK1a antisense transcript*
LINC00152	*Cytoskeleton regulator RNA*
LINE	*Long Interspersed Elements*
lncRNA	*Long noncoding RNA*
LTR	*Long Terminal Repeats*
m5C	*5-metilcitidina*
MALAT1	*Metastasis associated lung adenocarcinoma transcript 1*
mascRNAs	*MALAT1-associated small cytoplasmic RNA*
MIR31HG	*MIR31 host gene*
miRNAs	*MicroRNAs*
MLL	*Mixed-lineage leukaemia*
NCBI	*National Center for Biotechnology Information*
ncRNA	*Noncoding RNA*
ncRNA-a	*NcRNAs ativadores*
NF-YA	*Nuclear transcription factor Y subunit alpha*
nt	*Nucleotídeos*
ORF	*Open Reading Frame*
OSF	*Open Science Framework*
OVA	*Open Virtual Appliance*
PaaS	*Plataform-as-a-Service*
pb	*Pares de bases*
PC3	*Prostate cancer antigen 3*
PcG	*Polycomb group*
piRNAs	*Piwi-interacting RNAs*
plncRNAs	*Promoter-associated lncRNAs*
POU3F3	*POU class 3 homeobox 3*
PRC2	*Polycomb Repressive Complex 2*

PSA	*Próstata-específico*
PTEN	*Phosphatase and Tensin Homolog*
PTENP1	*Phosphatase and Tensin Homolog pseudogene 1*
RBH-BLAST	*Reciprocal Best HITS – BLAST*
Rme1	*Repressor de IME1*
RNAm	*RNA mensageiro*
RNAP	*RNA Polimerase*
SINE	*Short Interspersed Nuclear Elements*
siRNAs	*Small interference RNAs*
SMART	*Simple Modular Architecture Research Tool*
sncRNAs	*Pequenos RNAs não codificantes*
snoRNAs	*Small nucleolar RNAs*
snRNAs	*Small nuclear RNAs*
SPHK1	*Sphingosine kinase 1*
spRNA Pol IV	*Single-plypeptide nuclear RNA polymerase IV*
SUT	*Transcrição estável não anotada*
TN	*True Negative*
TP	*True Positive*
TSA	*Transcriptome Shotgun Assembly Sequence Database*
TUG1	*Taurine up-regulated 1*
TxG	*Trithorax*
UTR	*Untranslated Region*
vlincRNAs	*Very long intergenic noncoding RNAs*
WDR5	*WD repeat-containing protein 5*
XIC	*X inactivation center*
XIST	*X inactive specific transcript*
XUT	*Transcrição instável Xrn1*
ZEB2	*Zinc finger E-box binding homeobox 2*
ZEB2NAT	*ZEB2 antisense RNA1*

SUMÁRIO

CAPÍTULO 1. INTRODUÇÃO 20
1.1. REINO FUNGI 20
1.2.1. ESPÉCIES DE FUNGOS PATOGÊNICOS HUMANOS 25
1.2.1.1. ALTERNARIA MG1 25
1.2.1.2. ARTHRINIUM PHAEOSPERMUM 25
1.2.1.3. AUREOBASIDIUM MELANOGENUM 26
1.2.1.4. CANDIDA AFRICANA E CANDIDA ALBICANS 27
1.2.1.5. GEOTRICHUM CANDIDUM 28
1.2.1.6. IRPEX LACTEUS 29
1.2.1.7. MUCOR IRREGULARIS 30
1.2.1.8. PARACOCCIDIOIDES BRASILIENSIS 31
1.2.1.9. PAPILIOTREMA LAURENTII 32
1.2.1.10. PENICILLIUM JANTHINELLUM 32
1.2.1.11. PNEUMOCYSTIS JIROVECII 34
1.2.1.12. RHIZOPUS ARRHIZUS 35
1.2.1.13. YARROWIA SP 35
CAPÍTULO 2. CIÊNCIAS ÔMICAS 37
2.1. TRANSCRIPTOMA E RNA-SEQ 37
2.2. TRANSCRITOS 37
2.2.1. TRANSCRITOS CODIFICANTES DE PROTEÍNAS 38
2.2.2. TRANSCRITOS NÃO-CODIFICANTES DE PROTEÍNAS 39
2.2.2.1. RNAS PEQUENOS NÃO-CODIFICANTES DE PROTEÍNAS 39
2.2.2.2. RNAS LONGOS NÃO-CODIFICANTES DE PROTEÍNAS (LNCRNA) 40
2.2.2.2.1. ESTRUTURA DOS LNCRNAS 40
2.2.2.2.2. INTERAÇÕES MOLECULARES 41
2.2.2.2.2. MODIFICAÇÕES QUÍMICAS 42
2.2.2.2.3. BIOGÊNESE 43
2.2.2.2.4. LOCALIZAÇÃO INTRA E EXTRACELULAR 43
2.2.2.2.5. PADRÕES DE EXPRESSÃO 44
2.2.2.2.6. REGULADORES PRÉ-TRANSCRICIONAIS OU EPIGENÉTICOS 45
2.2.2.2.7. REGULADORES PÓS-TRANSCRICIONAIS – NÚCLEO CELULAR 47
2.2.2.2.8. REGULADORES PÓS-TRANSCRICIONAIS – CITOPLASMA 47

2.2.2.2.9. CLASSIFICAÇÃO 48
2.2.2.2.9.1. CLASSIFICAÇÃO BASEADA EM COMPRIMENTO 48
2.2.2.2.9.2. CLASSIFICAÇÃO BASEADA EM COMPRIMENTO 49
2.2.2.2.9.3. CLASSIFICAÇÃO EM LOCALIZAÇÃO INTRACELULAR 49
2.2.2.2.9.4. CLASSIFICAÇÃO EM ELEMENTOS REPETITIVOS 49
2.2.2.2.9.5. CLASSIFICAÇÃO FUNCIONAL 50
2.2.2.2.10. ARQUÉTIPOS DE AÇÃO DE MOLECULAR 50
2.2.2.2.10.1. ARQUÉTIPO SINALIZADOR 50
2.2.2.2.10.2. ARQUÉTIPO ISCA 51
2.2.2.2.10.3. ARQUÉTIPO GUIA 51
2.2.2.2.10.4. ARQUÉTIPO ANDAIME 52
2.2.2.2.11. MÉTODOS DE IDENTIFICAÇÃO IN SILICO 54
2.2.2.2.12. LNCRNAS ENVOLVIDOS EM DOENÇAS 55
2.2.2.2.13. LNCRNAS EM FUNGOS 57
2.2.2.2.14. LNCRNAS EM FUNGOS PATOGÊNICOS HUMANOS 58
3. OBJETIVO 59
3.1. OBJETIVOS ESPECÍFICOS 59
4. METODOLOGIA 60
4.1. OBTENÇÃO E IDENTIFICAÇÃO DE DADOS DE TRANSCRIPTOMA 60
4.2. FLUXO DE TRABALHO DE BIOINFORMÁTICA 61
4.2.1. ANÁLISE DE POTENCIAL DE CODIFICAÇÃO 61
4.2.2. ANOTAÇÃO FUNCIONAL DOS TRANSCRITOS CODIFICANTES DE PROTEÍNA 62
4.2.3. IDENTIFICAÇÃO DOS NCRNAS PUTATIVOS 63
4.2.4. DEFINIÇÃO DE LNCRNA PUTATIVO PARA ESSE TRABALHO 63
4.2.5. IDENTIFICAÇÃO DE CONSERVAÇÃO 63
5. RESULTADOS 64
5.1. OBTENÇÃO E IDENTIFICAÇÃO DE DADOS DE TRANSCRIPTOMA 64
5.2. ANÁLISE DE POTENCIAL DE CODIFICAÇÃO 67
5.3. ANOTAÇÃO FUNCIONAL DOS TRANSCRITOS NÃO-CODIFICANTES DE PROTEÍNA 79
6. DISCUSSÃO 93
7. CONCLUSÃO 96
8. PERSPECTIVAS FUTURAS 97

REFERÊNCIAS 98

CAPÍTULO 1. INTRODUÇÃO

1.1. Reino Fungi

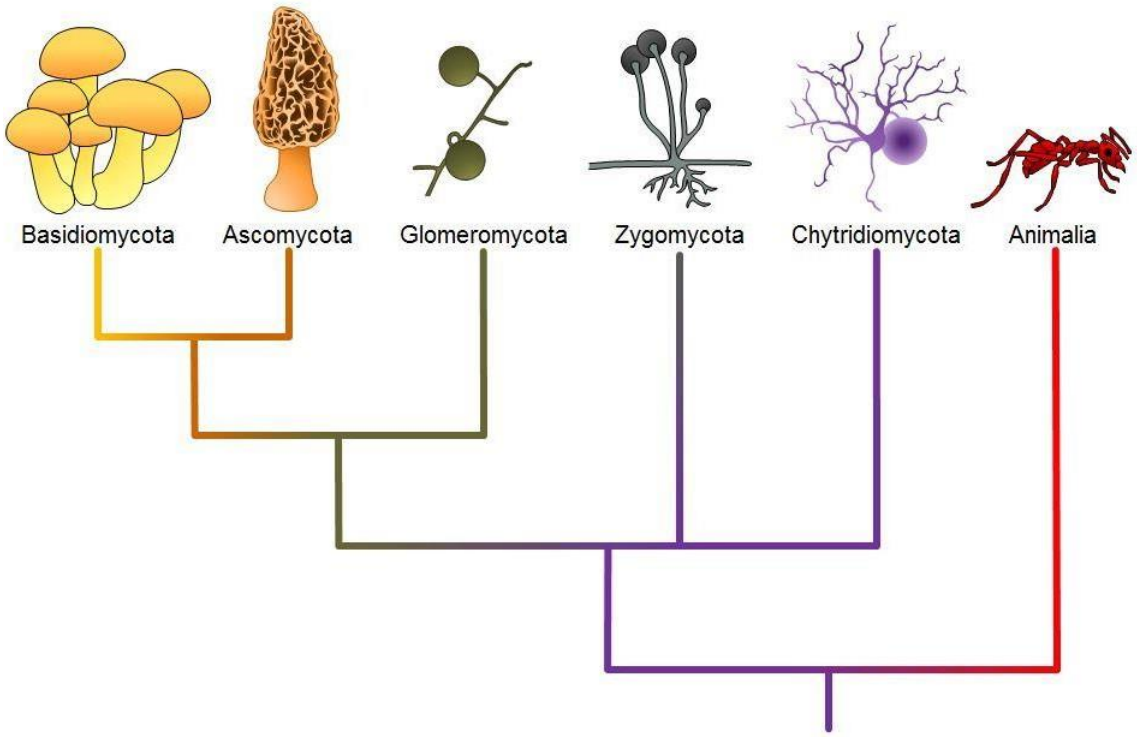
O Reino Fungi compreende todos os organismos conhecidos como fungos, os quais estão fora da dicotomia taxonômica clássica entre animais e vegetais, surgindo como um Terceiro Reino (NARANJO-ORTIZ e GALBADÓN, 2019). Eles possuem características como a presença de quitina na sua parede celular, produção de glicogênio como reserva energética e digestão extracorpórea, através da produção de enzimas, como pectinases, hidrolases e peroxidases, que exercem um papel importante na decomposição de matéria orgânica, como ligninas, celuloses e hemiceluloses. Estes processos contribuem significativamente para a ciclagem de nutrientes nos ecossistemas em que os fungos habitam (ALEXOPOULOS et al., 1952; LOGUERCIO-LEITE, 2010; KÜES, 2015; CALAÇA, 2016).

Podem ser encontrados em diversas formas, incluindo unicelular, não-filamentosa e microscópica - conhecida como levedura. Também podem apresentar formas filamentosas, chamadas bolores e mofos, e formas pluricelulares, filamentosas e macroscópicas, conhecidas popularmente como cogumelos (FORTUNA, 2020). As leveduras possuem estruturas morfológicas simples, geralmente encontradas em formas esféricas ou ovais. Já os bolores possuem uma estrutura denominada talo, constituído por longos filamentos de células chamadas hifas, que podem conter septos (TORTORA, 2017). Os cogumelos, são extremamente diversificados nas formas que podem ser encontrados. A ordem Agaricales apresenta a forma mais comum pela qual são conhecidos – assemelhando-se a um “guarda-chuva” e possuem estruturas denominadas chapéu ou píleo, lamelas, estipe, véu e anel (CLOSS e PUTZKE, 2017).

De acordo com Konopka et al. (2019) e Fisher e Garner. (2020), o Reino Fungi possui uma ampla diversidade de espécies, estimada em aproximadamente 6 milhões de espécies. Dentre essas, uma parcela significativa é relacionada a organismos vertebrados, com cerca de 625 espécies identificadas. Além disso, cerca de 200 espécies de fungos são consideradas como associadas ao ser humano, podendo ser caracterizadas por relações simbióticas ou comensais, principalmente como componentes do micobioma humano, ou

através de doenças infecciosas causadas por espécies patogênicas, sendo estas de interesse para a saúde humana. A diversidade do Reino Fungi é apresentada na Figura 1.

Figura 1. Diversidade do Reino Fungi. Representação dos cinco filos do Reino Fungi: Chytridiomycota (Chytrids), os Zygomycota (fungos conjugados), os Ascomycota (fungos de saco), os Basidiomycota (club fungi) e o filo Glomeromycota.



Fonte: Figura disponível em: OpenEd CUNY

1.2. Fungos patogênicos humanos

As doenças fúngicas que acometem os seres humanos compreendem uma variedade de categorias, que variam de acordo com os tecidos infectados pelos organismos. Essas categorias incluem micoses superficiais, cutâneas, subcutâneas, endêmicas ou oportunistas, algumas das quais podem progredir para quadros de infecções sistêmicas (MURRAY et al., 2017; SEGRELLES-CALVO et al., 2020; SAMADDAR et al., 2021). Além disso, o aumento do número de pacientes com imunossupressão severa é outro fator contribuinte para a crescente incidência de infecções fúngicas (TIRADO-SÁNCHEZ et al., 2020; DIAKOMOPOULOS et al., 2021). Embora a atual disponibilidade de fármacos antifúngicos, também há evidências do aumento de espécies resistentes ao tratamento por

este grupo medicamentoso, especialmente entre as que acometem pacientes que recebem tratamentos longos, ou que estão recebendo medicamentos de forma profilática, onde, em ambos os casos, pode haver disbiose na microbiota humana, propiciando um ambiente favorável para espécies ou cepas com esta característica resistentes (BERTO et al., 2018; MISHRA et al., 2022).

Para que um fungo seja capaz de infectar o homem e o acometa em alguma doença decorrente de sua infecção, é necessário que quatro critérios principais estejam presentes nestes organismos (KÖHLER et al., 2015): (i) Sua capacidade para se desenvolver em altas temperaturas, que sejam iguais ou acima de 37ºC (visto que é a temperatura corpórea do hospedeiro); (ii) Transpor as barreiras teciduais ou adentrar o organismo através de esporos ou conídios transportados pelo ar que atingem diretamente as cavidades aéreas dos pulmões e seios nasais de seu hospedeiro, para atingir os tecidos-alvo e parasitá-los; (iii) Ter habilidade para digerir e absorver componentes dos tecidos humanos; (iv) Deve ser capaz de resistir ao sistema imune humano.

Entretanto, os organismos patogênicos também podem utilizar vesículas extracelulares como mecanismo para causar modificações no sistema imunológico de seus hospedeiros, contribuindo para a infecção (RODRIGUES et al., 2007; MATOS et al.; 2018; BIELSKA e MAY, 2019).

Atualmente, está estabelecido que os fungos patogênicos humanos surgiram de forma independente em diferentes linhagens dentro do Reino Fungi, o que demonstra a versatilidade e adaptabilidade desses organismos para colonizar novos habitats, incluindo o homem (ROKAS, 2022). Além disso, é importante observar que muitos dos fungos patogênicos humanos, como as espécies dos gêneros *Aspergillus*, *Candida*, *Cryptococcus*, *Histoplasma*, *Coccidioides* e *Paracoccidioides*, também são capazes de infectar outros animais, incluindo mamíferos e outros vertebrados (SEYEDMOUSAVI et al., 2018). Essas infecções podem ser resultado da aquisição direta de esporos fúngicos do ambiente, mas também são conhecidos surtos zoonóticos como a transmissão de gatos domésticos para humanos do fungo *Sporothrix brasiliensis* (RODRIGUES et al., 2016).

Os fungos desempenham um papel fundamental como decompositores de matéria orgânica na natureza, crescendo em diversos substratos e interagindo com uma ampla variedade de organismos. Conforme apresentado por Tekaia e Latge (2005), a maioria dos patógenos fúngicos são oportunistas, o que sugere que sua capacidade de infectar humanos pode ser compreendida ao analisar seus ambientes naturais. O solo é um nicho ecológico

comum onde muitos fungos patogênicos oportunistas podem ser encontrados. Segundo Casadevall (2006), um grama de solo pode abrigar bilhões de organismos microbianos de milhares de espécies pertencentes a dezenas de grupos taxonômicos, os quais também contribuem para a sobrevivência dos fungos na natureza. Um recente estudo ecológico evolutivo de mais de 1.200 espécies de fungos revelou uma associação significativa entre a capacidade de sobreviver em vários tipos diferentes de condições como termotolerância, a capacidade destes organismos em resistir a altas temperaturas (MUNIZ, 2015) e osmotolerância, a capacidade de se adaptar às alterações na disponibilidade de soluto em seu meio (COSTA, 2017) e patogenicidade oportunista (GOSTINČAR et al., 2018).

Sua biogeografia e fisiologia também demonstram ser fatores significativos para infectar humanos. Por exemplo, os fungos ascomicetos tendem a ser mais amplamente distribuídos e termotolerantes do que os fungos basidiomicetos (ROBERT et al., 2015; EGIDI et al., 2019), o que pode sugerir hipóteses a respeito do porquê existem mais linhagens de patógenos humanos oportunistas em Ascomicetos do que em Basidiomicetos. Uma abordagem para se compreender essa questão é o estudo das diferenças entre patógenos e fungos não patogênicos (CHATURVEDI e DE HOOG, 2020). Essas diferenças podem ser decorrentes de variabilidade e abundância (EGIDI et al., 2019, YAMAMOTO et al., 2012), capacidade de crescimento em temperatura corporal humana (ROBERT et al., 2015), adaptabilidade a diferentes níveis de oxigênio (CHUNG e LEE, 2020), preferência por reprodução sexuada ou assexuada (MAGWENE et al., 2011) e resposta à predação (NOVOHRADSKÁ et al., 2017).

Fatores evolutivos também demonstram estar presentes. As leveduras do gênero *Pneumocystis sp*. não podem sobreviver fora de um hospedeiro mamífero e se tornam patogênicas quando há uma condição de enfraquecimento de seu sistema imunológico (CUSHION, 2010; HAN et al., 2021). O mesmo ocorre com espécies do gênero *Candida sp.*, nas quais, quando não causam quadros clínicos de candidíase são comensais (OPULENTE et al., 2019, BENSASSON et al., 2019; ROMO e KUMAMOTO, 2020) o que pode sugerir um evento de coevolução.

No entanto, também há fungos que não são dependentes de seus hospedeiros para sobreviverem (TAYLOR e BERBEE. 2001; FISHER et al., 2001). A espécie *Aspergillus fumigatus* cresce melhor à temperatura do corpo humano e é muito mais tolerante ao estresse oxidativo associado à disponibilidade de nutrientes e oxigênio do que outra espécie filogeneticamente próxima do gênero, *A. fischeri* (MEAD et al., 2019). Diferenças na

capacidade de crescer à temperatura do corpo humano também são observadas entre organismos do complexo de espécies patogênicas de *Cryptococcus sp.*, que inclui os principais patógenos *Cryptococcus neoformans* e *C. gatti*, bem como seus parentes não patogênicos intimamente relacionados, como *Cryptococcus amylolentus* (FINDLEY et al., 2009; BLOOM et al, 2019).

Estas doenças fúngicas invasivas acarretam um grande impacto na saúde pública, com taxas de mortalidade de 30% a 90%, dependendo do patógeno e da população de pacientes. Aproximadamente 90% das mortes decorrentes de infecções fúngicas são causadas por espécies dos gêneros *Aspergillus*, *Cryptococcus*, *Candida* ou *Pneumocystis*, além de outros grupos, como *Coccidioides* e *Histoplasma*, os quais podem infectar mesmo hospedeiros imunocompetentes em regiões de endemismo para estes fungos. As espécies de *Cryptococcus* também são reconhecidas como uma das micoses mais significativas como ameaça à saúde humana. Ocasionando mais de 223.100 casos de meningite criptocócica anuais e levando a óbito mais de 181.000 pacientes portadores de HIV (Vírus da Imunodeficiência Humana), de acordo com a Organização Mundial da Saúde (PERFECT e BICANIC et al., 2015).

Outro patógeno fúngico emergente que ganhou atenção internacional é a levedura *Candida auris*, conhecida por ser resistente à diversos agentes antifúngicos e estar presente em graves infecções oportunistas em hospitais (FISHER et al., 2020). Inicialmente relatada em casos de infecções de ouvido (SATOH et al., 2009), foi logo depois detectada como causadora de surtos de infecções em outras regiões do corpo humano em diversas regiões do mundo e, recentemente, um caso de infecção envolvendo *C. auris* foi identificado no Brasil (ANVISA, 2021).

Contudo, em aspectos epidemiológicos, de 1990 a 2010, as dermatoses fúngicas foram a quarta causa mais prevalente de sequelas em humanos e sua prevalência aumentou 42,3% entre as pessoas que vivem com alguma deficiência por doenças não transmissíveis no mesmo período (VOS et al. 2012). Estudos epidemiológicos afirmam que cerca de 1,2 bilhão de pessoas ao redor do mundo sofrem de alguma doença fúngica (DENNING e BROMLEY, 2015). Contudo, devido alto número de indivíduos acometidos, trabalhos mais recentes (CHAVES et al., 2021) questionam esta estimativa e apontam que são casos que correspondem somente a dermatomicoses e ignoram doenças fúngicas sistêmicas. Assim, demonstrando a necessidade de se levantar dados de maior confiabilidade a respeito da quantidade de indivíduos infectados também por este tipo de micose ao redor do mundo.

Nos tópicos seguintes serão apresentadas os gêneros e espécies de fungos patogênicos para seres humanos que foram analisados neste trabalho e algumas características de sua biologia e virulência.

1.2.1. Espécies de fungos patogênicos humanos

1.2.1.1. Alternaria MG1

O gênero *Alternaria* é composto por fungos filamentosos pertencentes à classe Dothideomycetes, ordem Pleosporales e família Pleosporaceae. Este gênero é conhecido por ser patogênico em diversas plantas, além de ser um importante agente causador de alergias respiratórias em humanos. Dentre as espécies de *Alternaria*, a *Alternaria MG1* é considerada uma das principais causadoras de alergias respiratórias, sendo encontrada frequentemente em ambientes urbanos e rurais.

A *Alternaria MG1* é um fungo saprófito, ou seja, se alimenta de matéria orgânica em decomposição. No entanto, devido à sua capacidade de produzir metabólitos secundários alergênicos, esta espécie tem sido associada a diversos casos de rinites e asma (CHEN et al., 2018). Estudos também sugerem que a exposição prolongada à Alternaria MG1 pode aumentar o risco de desenvolvimento de asma persistente (BRUNEKREEF et al., 2002).

Além de sua capacidade de causar alergias respiratórias, a *Alternaria MG1* também é conhecida por ser um agente fitopatogênico. Esta espécie é responsável por causar manchas foliares em diversas plantas, incluindo frutas, hortaliças e culturas de grãos (FARR & ROSSMAN, 2019). A produção de enzimas hidrolíticas, como celulases e pectinases, permite que a *Alternaria MG1* invada as células vegetais, causando a deterioração da planta (GARCÍA-JIMÉNEZ et al., 2015).

1.2.1.2. ARTHRINIUM PHAEOSPERMUM

O gênero de fungo patogênico humano *Arthrinium* é composto por espécies com características morfológicas semelhantes, tendo como principal característica a presença de hifas septadas e ramificadas. Dentre as espécies desse gênero, a *Arthrinium phaeospermum*

é uma das mais relevantes devido ao seu potencial patogênico.

A. phaeospermum é um fungo filamentoso que apresenta hifas septadas e ramificadas, possuindo características similares a outras espécies do gênero. Esse fungo é comumente encontrado em ambientes úmidos, como em solos e em plantas (SABELLA et al., 2013).

Esse fungo tem sido associado a vários casos de infecção humana, principalmente em pacientes imunocomprometidos (KUMAR et al., 2008). As infecções causadas pela espécie são caracterizadas por lesões cutâneas, como feridas e bolhas, e podem se espalhar para outros órgãos, como os pulmões e o sistema nervoso central (KUMAR et al., 2008). Além disso, a *A. phaeospermum* é capaz de produzir enzimas, como a celulasana e a protease, que podem contribuir para sua virulência e sua capacidade de invadir tecidos humanos (KUMAR et al., 2008).

Outro estudo realizado por GUPTA et al. (2018) demonstrou que a *A. phaeospermum* é capaz de produzir metabólitos secundários com atividade antifúngica contra outras espécies de fungos. Isso sugere que essa espécie possa ter potencial como fonte de compostos com aplicação terapêutica para o tratamento de infecções fúngicas. Por fim, outros estudos têm demonstrado a capacidade de *A. phaeospermum* de se desenvolver em diferentes tipos de tecidos humanos, incluindo pele, tecido nervoso e pulmão (GERAMISHOAR et al., 2019; HE et al., 2022). Isso indica que esse fungo pode causar infecções sistêmicas, o que aumenta a importância de se compreender sua patogenicidade e se desenvolver estratégias de prevenção e tratamento.

1.2.1.3. AUREOBASIDIUM MELANOGENUM

O gênero *Aureobasidium* é composto por fungos patogênicos que possuem a capacidade de causar infecções em seres humanos, animais e plantas. Dentre as espécies desse gênero, *Aureobasidium melanogenum* é a mais frequentemente relacionada às infecções humanas.

Esses fungos são encontrados em ambientes variados, incluindo solo, água e superfícies de edifícios. *A. melanogenum* é comumente encontrada em materiais de construção, como tintas, argamassas e papel de parede (CHEN et al., 2016). Além disso, essa espécie também pode ser encontrada em ambientes hospitalares, como em unidades de terapia intensiva e em pacientes imunocomprometidos (ABBES et al., 2016).

A. melanogenum é conhecida por causar infecções invasivas em pacientes imunocomprometidos, como aqueles com AIDS ou pacientes submetidos a transplante de células-tronco (CHEN et al., 2016). Essas infecções podem se manifestar como pneumonias, septicemias e infecções cutâneas (BROOKS et al., 1984). Além disso, essa espécie também é associada ao desenvolvimento de aspergilomas, uma condição em que o fungo cresce em uma cavidade pulmonar (CHEN et al., 2016).

A capacidade da *A. melanogenum* de causar infecções invasivas em pacientes imunocomprometidos é atribuída à sua capacidade de produzir melanina, um pigmento que confere a esses fungos resistência aos mecanismos de defesa do hospedeiro (CHEN et al., 2016). Além disso, essa espécie também possui uma série de enzimas hidrolíticas que lhe permitem degradar componentes da parede celular do hospedeiro, o que contribui para sua virulência (BROOKS et al., 1984).

1.2.1.4. CANDIDA AFRICANA E *CANDIDA ALBICANS*

O gênero *Candida* é um dos principais patógenos fúngicos humanos, tendo a capacidade de causar infecções em vários órgãos e tecidos do corpo. Entre as espécies mais comuns estão a *Candida albicans* e a *Candida africana*, sendo estas as principais responsáveis por infecções em pacientes imunocomprometidos.

A *Candida albicans* é a espécie mais prevalente entre os pacientes com infecções fúngicas, tendo sido relatado que cerca de 80% das infecções fúngicas são causadas por esta espécie (KURTZMAN e ROBNETT, 1998). Esta espécie é normalmente presente na flora intestinal e vaginal, mas pode causar infecções invasivas em pacientes imunocomprometidos, como aqueles com AIDS, pacientes oncológicos e pacientes submetidos a tratamento com antibióticos (KENNEDY et al., 2010).

Já a *Candida africana* é uma espécie relativamente recente, tendo sido descrita pela primeira vez em 2009 (TIETZ et al., 2001). Esta espécie tem sido associada a infecções invasivas em pacientes imunocomprometidos, como aqueles com AIDS e pacientes oncológicos (TIETZ et al., 2001; KENNEDY et al., 2010). Além disso, a *Candida africana* tem mostrado resistência aos antifúngicos azólicos, o que a torna uma ameaça crescente para o tratamento de infecções fúngicas KENNEDY et al., 2010).

As infecções causadas por *Candida* podem variar desde infecções de pele e mucosas até infecções invasivas, como sepse e meningite. O diagnóstico e tratamento dessas infecções requer a identificação precisa da espécie envolvida e a avaliação da sensibilidade aos antifúngicos (KENNEDY et al., 2010).

1.2.1.5. GEOTRICHUM CANDIDUM

O gênero *Geotrichum* é composto por fungos filamentosos pertencentes à classe de Ascomycetes. Dentre as espécies deste gênero, *Geotrichum candidum* é uma das mais estudadas devido ao seu potencial patogênico em humanos. Esta espécie é comumente encontrada em ambientes naturais, como solo e água, e pode ser isolada de diversos tipos de alimentos, como leite e queijos (KÖHLER et al., 2016).

G. candidum é capaz de causar infecções em diferentes órgãos e sistemas do corpo humano, como pulmões, ouvidos, pele e trato gastrointestinal (KEENE et al., 2019). A forma mais comum de infecção é através da inalação de esporos presentes no ar, que podem causar asma e bronquite (KÖHLER et al., 2016). Além disso, *G. candidum* também pode causar infecções cutâneas, como dermatite seborreica e onicomicose (KEENE et al., 2019).

A capacidade de *G. candidum* de causar infecções em humanos é atribuída à sua capacidade de aderir e colonizar tecidos, bem como à produção de enzimas hidrolíticas (KÖHLER et al., 2016). Ainda assim, é importante ressaltar que a maioria das infecções causadas por esta espécie são consideradas fúngicas oportunistas, ou seja, ocorrem em indivíduos com sistemas imunológicos debilitados (KEENE et al., 2019).

Além de sua capacidade patogênica, *G. candidum* também é conhecida por sua importância na indústria alimentícia. Esta espécie é utilizada como fermento em diversos tipos de queijos, como o *brie* e o *camembert*, para conferir sabor e textura característicos a esses produtos (KÖHLER et al., 2016). Além disso, a *G. candidum* também é utilizada como fermento em outros produtos alimentícios, como pães e vinhos (KEENE et al., 2019).

Apesar de sua importância na indústria alimentícia, a *G. candidum* também pode ser considerada um contaminante em alimentos. A presença desta espécie em produtos alimentícios pode indicar falta de higiene durante o processamento, e a ingestão de alimentos contaminados pode levar a problemas de saúde (KÖHLER et al., 2016).

A identificação precisa da *G. candidum* é importante tanto para o controle de sua presença em alimentos quanto para o diagnóstico e tratamento de infecções fúngicas. A identificação pode ser realizada através de técnicas de microbiologia clássica, como cultivo em meios seletivos e testes de caracterização bioquímica, ou através de técnicas moleculares, como PCR e sequenciamento de DNA (KEENE et al., 2019).

1.2.1.6. IRPEX LACTEUS

O gênero *Irpex* é uma das principais fontes de fungos patogênicos para os seres humanos. Dentro desse gênero, a espécie *Irpex lacteus* é uma das mais estudadas e frequentemente encontradas em infecções fúngicas humanas. Esses fungos são encontrados comumente em ambientes naturais, como florestas e solos, e podem ser disseminados através do ar ou através de contato direto com materiais contaminados.

Irpex lacteus é conhecido por causar infecções fúngicas invasivas em pacientes imunocomprometidos, como aqueles com câncer, HIV ou pacientes em tratamento com terapia imunossupressora (ENOCH et al., 2006). Além disso, esses fungos também podem causar infecções cutâneas e subcutâneas em indivíduos saudáveis (MARTINS et al., 2017).

Os sinais e sintomas de infecção por *Irpex lacteus* variam de acordo com a localização da infecção, mas geralmente incluem vermelhidão, inchaço, dor e descamação na área afetada (SKIADA et al., 2020). Além disso, esses fungos também podem causar sintomas gerais, como febre e fraqueza.

O diagnóstico de infecções por *Irpex lacteus* geralmente é feito através da cultura de tecidos ou secreções da área afetada, seguida de identificação do fungo através de técnicas de microscopia e testes bioquímicos (GARCIA-HERMOSO et al., 2018). O tratamento geralmente inclui o uso de antifúngicos, como anfotericina B ou itraconazol, mas pode ser complicado pela resistência a esses medicamentos (AL-HATMI et al., 2019).

1.2.1.7. MUCOR IRREGULARIS

Os fungos do gênero *Mucor* são considerados como patógenos oportunistas em humanos, o que significa que eles podem causar infecções em indivíduos com sistemas imunológicos comprometidos. Dentre as várias espécies desse gênero, a *Mucor irregularis* é uma das mais frequentemente associadas a infecções humanas (SKIADA et al., 2020).

A *Mucor irregularis* é um fungo filamentoso que pertence à ordem *Mucorales* e à família *Mucoraceae*. Ele é amplamente distribuído na natureza e pode ser encontrado em solos, plantas e alimentos (WANG et al., 2018). Esse fungo é caracterizado por sua capacidade de formar colônias fúngicas filamentosas, com estruturas reprodutivas denominadas esporangióforos que contêm esporângios (RAUTHAN e SHARMA et al., 2021).

As infecções causadas pela *M. irregularis* são conhecidas como mucormicoses, que são consideradas como doenças invasivas e potencialmente fatais (KWON-CHUNG et al., 1992). Essas infecções podem afetar diversos órgãos e tecidos, incluindo o nariz, o cérebro, os olhos e os pulmões (GARCIA-HERMOSO et al., 2018). As principais vias de entrada para a infecção são a via respiratória e a via cutânea (GARCIA-HERMOSO et al., 2018).

Os indivíduos mais suscetíveis às infecções causadas pela *M. irregularis* são aqueles com sistemas imunológicos comprometidos, como pacientes com diabetes, câncer, transplantes de órgãos sólidos e pacientes em terapia imunossupressora (KWON-CHUNG et al., 2017). O diagnóstico dessas infecções é baseado na história clínica, exame físico e resultados de exames laboratoriais, como a cultura fúngica e a microscopia de esfregaços (KWON-CHUNG et al., 1992).

O tratamento das infecções causadas pela *M. irregularis* é baseado na administração de antifúngicos, como a amfotericina B, associada à terapia de suporte, como a desbridamento cirúrgico e a terapia intensiva (KWON-CHUNG et al., 1992). A prevenção dessas infecções inclui o controle dos fatores de risco, como o controle glicêmico em pacientes com diabetes e a evitação de feridas cutâneas em pacientes imunossuprimidos (KWON-CHUNG et al., 1992). Além disso, é importante destacar que as infecções causadas pela *M. irregularis* são mais comuns em pacientes com doenças crônicas, como o diabetes, e em pacientes submetidos a tratamentos intensivos, como a terapia intensiva e a quimioterapia (KWON-CHUNG et al., 1992). A incidência dessas

infecções tem aumentado nos últimos anos devido ao aumento da população de pacientes imunossuprimidos, incluindo aqueles com câncer e aqueles que receberam transplantes de órgãos sólidos (SKIADA et al., 2020).

Apesar de ser uma espécie pouco comum, é importante estar atento às infecções causadas pela *M. irregularis* devido à sua alta mortalidade. A identificação precoce da infecção e o tratamento adequado são fundamentais para melhorar o prognóstico dos pacientes (RAUTHAN e SHARMA et al., 2021). Além disso, a prevenção dessas infecções é crucial para evitar a ocorrência de casos graves e fatais (KWON-CHUNG et al., 1992).

1.2.1.8. PARACOCCIDIOIDES BRASILIENSIS

O gênero *Paracoccidioides* é composto por fungos dimórficos, ou seja, os indivíduos desse gênero possuem duas formas morfológicas distintas: uma forma filamentosa, que se caracteriza por se desenvolver em meios líquidos, e uma forma leveduriforme, que é caracterizada por ser unicelular e se desenvolver em meios sólidos (DURLACHER et al., 2017). Dentre as espécies desse gênero, *Paracoccidioides brasiliensis* é a principal responsável por causar a paracoccidioidomicose, uma doença fúngica que afeta principalmente os indivíduos do sexo masculino e que é endêmica em regiões como América Latina (RESTREPO et al., 2016).

A paracoccidioidomicose é caracterizada por ser uma doença granulomatosa crônica, que pode se manifestar de forma aguda ou subaguda (RESTREPO et al., 2016). Os sintomas incluem tosse, expectoração, dor torácica, fadiga, perda de peso e, em casos avançados, acometimento de órgãos como os pulmões, ossos e gânglios linfáticos (DURLACHER et al., 2017). A infecção pelo fungo ocorre através da inalação do ar contendo esporos fúngicos, que se desenvolvem na forma leveduriforme após a exposição às condições adequadas (RESTREPO et al., 2016).

O diagnóstico da paracoccidioidomicose é feito através de exames como a cultivação do fungo, a microscopia direta e o teste sorológico (DURLACHER et al., 2017). O tratamento é baseado em terapia antifúngica, sendo a itraconazol a droga de escolha (RESTREPO et al., 2016). A prevenção da doença inclui medidas como a redução da exposição ao fungo e a vacinação (DURLACHER et al., 2017).

1.2.1.9. PAPILIOTREMA LAURENTII

O gênero de fungo patogênico humano *Papiliotrema* é composto por espécies que podem causar infecções fúngicas sistêmicas e cutâneas em seres humanos. Dentre as espécies pertencentes a este gênero, *Papiliotrema laurentii* é a mais estudada e relacionada à doença humana.

Papiliotrema laurentii é um fungo filamentoso que foi descrito pela primeira vez por Laurent em 1948 como uma espécie nova do gênero *Papiliotrema* (LAURENT, 1948). Ele é caracterizado pela presença de septos filamentosos e células de ascoides com ascos tetrasporangiados (LONDERO et al., 2019). A espécie é comumente encontrada em solos úmidos e em vegetação de florestas tropicais (LONDERO et al., 2019).

A infecção causada por *P. laurentii* é rara e geralmente se apresenta como uma infecção sistêmica fúngica, como meningite criptocócica (LONDERO et al., 2019; CASADEVALL et al., 2002). Além disso, também foi relatado que *P. laurentii* pode causar infecções cutâneas, como piodermite (CASADEVALL et al., 2002; LONDERO et al., 2019). Os sintomas da infecção variam de acordo com a localização da infecção, mas geralmente incluem febre, cefaleia, confusão, alterações neurológicas e sinais de meningite (CASADEVALL et al., 2002; LONDERO et al., 2019).

O diagnóstico da infecção por *P. laurentii* é baseado em exames clínicos e laboratoriais. O exame microscópico é usado para identificar o fungo e o exame de cultura é usado para confirmar a identificação (CASADEVALL et al., 2002; LONDERO et al., 2019). O tratamento da infecção é geralmente feito com antifúngicos, como anfotericina B e fluconazol (CASADEVALL et al., 2002; LONDERO et al., 2019).

É importante ressaltar que, devido à raridade da infecção causada por *P. laurentii*, ainda há muito a ser estudado sobre este fungo e sua relação com a doença humana. Portanto, é fundamental continuar a realizar pesquisas e estudos para entender melhor a patogenicidade e epidemiologia desta espécie.

1.2.1.10. PENICILLIUM JANTHINELLUM

O gênero *Penicillium* é um dos mais comuns e diversificados entre os fungos filamentosos. Ele é composto por cerca de 300 espécies, com diferentes características morfológicas e biologia. Dentre elas, a espécie *Penicillium janthinellum* é uma das mais

relevantes do ponto de vista médico e farmacêutico, pois é um dos principais produtores de penicilinas, um grupo de antibióticos amplamente utilizado no tratamento de infecções bacterianas.

De acordo com Frisvad e Samson (2016), *P. janthinellum* é uma espécie de dimorfismo opcional, ou seja, é capaz de apresentar as formas assexuadas (micélio) e sexuadas (esporos) de acordo com as condições ambientais. Isso pode ser observado pela coloração do micélio, que varia de branco a azulado-verde, dependendo da presença de oxigênio. Além disso, esta espécie é conhecida por ser produtora de penicilinas e outros compostos de interesse farmacêutico (FRISVAD & SAMSON, 2016).

Bentley et al. (2011) relatam que *P. janthinellum* foi originalmente descrita por Saccardo em 1887, e é uma espécie cosmopolita, podendo ser encontrada em diferentes ambientes, como solos, alimentos e até mesmo no ar. É comum encontrar esta espécie em cultivos de cogumelos, onde é considerada uma praga devido à sua capacidade de produzir enzimas líticas que causam danos às células dos cogumelos (BENTLEY et al., 2011).

De acordo com Klotz et al. (2016), *P. janthinellum* é um dos principais produtores de penicilina G, um dos antibióticos mais utilizados na história da medicina. A produção de penicilina G é controlada por um conjunto de genes reguladores, que são ativados quando as condições ambientais são favoráveis para a produção de penicilina. Esta espécie também é capaz de produzir outros compostos de interesse, como citrinina, xantofilas e ácido cicloexenoico (KLOTZ et al., 2016).

Além disso, é importante mencionar que *P. janthinellum* também tem sido estudado como um potencial agente biotecnológico, devido à sua capacidade de produzir compostos de interesse industrial, como enzimas e pigmentos (KLOTZ et al., 2016; FRISVAD & SAMSON, 2016). Estudos também têm mostrado que *P. janthinellum* tem atividade antifúngica e antioxidantes, o que pode ter aplicações na área de saúde humana e animal (MAHLO et al., 2013; DOMINGUETE et al., 2021).

De acordo com Chen et al. (2016), *P. janthinellum* é um fungo comumente encontrado em ambientes hospitalares, como ar-condicionado, ventilação e superfícies de contato. Ele é capaz de produzir enzimas líticas e proteínas que podem causar danos aos tecidos humanos, além de ser um agente de infecção em pacientes imunocomprometidos. Os autores relataram um caso de óbito em um paciente idoso com LES devido a uma infecção fúngica disseminada causada por *P. janthinellum*. O paciente apresentou sintomas como falta de ar, dor torácica e insuficiência renal aguda (SANTOS et al., 2018), e foi

diagnosticado com um quadro de sepse fúngica disseminada. A análise microbiológica revelou a presença de *P. janthinellum* em amostras de sangue e tecido pulmonar (CHEN et al., 2016).

1.2.1.11. PNEUMOCYSTIS JIROVECII

O gênero *Pneumocystis* é composto por fungos unicelulares obrigatórios que habitam o trato respiratório superior dos seres humanos e animais. Dentre as espécies desse gênero, *Pneumocystis jirovecii* é a mais importante em relação à patologia humana. Essa espécie é o agente etiológico da pneumocistose, uma infecção pulmonar oportunista que afeta principalmente indivíduos imunodeprimidos, como pacientes com AIDS (Síndrome da Imunodeficiência Adquirida) ou submetidos a tratamentos imunossupressores (MOLINA et al., 2020).

Pneumocystis jirovecii é um fungo unicelular que apresenta características morfológicas e fisiológicas atípicas para fungos. Ele é composto por organismos unicelulares que se apresentam como cistos ou formas trofozóicas, sendo que as células trofozóicas são raramente encontradas no tecido pulmonar (WALKER et al., 2011). A forma cística é responsável pela transmissão da infecção, já que é resistente ao ambiente externo e pode permanecer viável por longos períodos (MOLINA et al., 2020).

Pneumocystis jirovecii é uma espécie que se reproduz de forma assexuada, através de divisão celular binária. Apesar de não possuir parede celular, essa espécie apresenta características semelhantes a fungos, como a capacidade de formar septos celulares e a presença de filamentos citoplasmáticos (WALKER et al., 2011).

A infecção por *Pneumocystis jirovecii* é caracterizada por sintomas respiratórios, como tosse, dispneia e hipoxemia (MOLINA et al., 2020). A gravidade da doença varia de acordo com o estado imunológico do indivíduo, sendo geralmente assintomática em indivíduos saudáveis e potencialmente fatal em pacientes imunodeprimidos (WALKER et al., 2011). O tratamento da pneumocistose é realizado com a administração de sulfonamidas associadas a trimetoprima, como a pentamidina ou a atovaquona (MOLINA et al., 2020).

1.2.1.12. RHIZOPUS ARRHIZUS

O gênero *Rhizopus* é um dos principais patógenos fúngicos humanos, sendo responsável por diversas infecções oportunistas em indivíduos imunodeprimidos. Dentre as espécies deste gênero, *Rhizopus arrhizus* é a mais frequentemente isolada em casos de infecção.

R. arrhizus é um fungo filamentoso que apresenta hifas septadas, com presença de esporangióforos e esporangiósporos. É comumente encontrado em ambientes úmidos, como frutas e vegetais em decomposição, e pode ser isolado a partir de diversos tecidos humanos, como o nariz, boca e trato respiratório (CHEN et al., 2016).

A infecção por *R. arrhizus* é mais comum em indivíduos imunocomprometidos, como pacientes com diabetes mellitus, pacientes em terapia imunossupressora ou aqueles com doenças hematológicas malignas (CHAO et al., 2022). A infecção pode se manifestar de diversas formas, desde casos leves, como rinofaringite, até casos graves, como meningite fúngica ou abscessos pulmonares (WICKES, 2013).

Além disso, *R. arrhizus* também é conhecida por sua capacidade de produzir enzimas líticas, como a protease, que podem contribuir para a patogenicidade da infecção (HASSAN e VOIGT et al., 2019). A resistência aos antifúngicos também é uma preocupação crescente, com estudos demonstrando a presença de mecanismos de resistência aos antifúngicos azólicos e poliene (JEONG et al., 2019).

1.2.1.13. YARROWIA SP.

O gênero *Yarrowia* é composto por fungos que pertencem à classe Ascomycetes, ordem Saccharomycetales e família Dipodascaceae. Este gênero é amplamente distribuído em diversos ambientes naturais, incluindo solos, água e alimentos. Entretanto, algumas espécies do gênero *Yarrowia* têm sido relatadas como patogênicas para seres humanos.

A especie *Yarrowia sp.* é um dos principais agentes patogênicos humanos deste gênero. Ela é conhecida por causar infecções em diferentes órgãos e tecidos, incluindo o sistema nervoso central, ossos, pulmões, pele e unhas (ABBES et al., 2016; KWON-CHUNG et al., 1992). Estas infecções são geralmente associadas a pacientes com comprometimento imunológico, como aqueles com HIV ou pacientes submetidos a transplante de órgãos (ABBES et al., 2016).

Além disso, a especie *Yarrowia sp.* também tem sido relacionada a infecções fúngicas disseminadas, como a candidemia (SANTOS et al., 2018). A candidemia é uma infecção sistêmica causada pela espécie *Candida*, que pode ser fatal se não for tratada adequadamente (SANTOS et al., 2018). A especie *Yarrowia sp.* tem sido relatada como um importante agente etiológico da candidemia, principalmente em pacientes com comprometimento imunológico (SANTOS et al., 2018). Também tem sido relacionada a outras infecções fúngicas, como a aspergilose (KWON-CHUNG et al., 1992) e a criptococose (ABBES et al., 2016). A aspergilose é uma infecção causada pela espécie *Aspergillus*, que pode afetar os pulmões e outros órgãos (KWON-CHUNG et al., 1992). Já a criptococose é uma infecção causada pela espécie *Cryptococcus*, que pode afetar o sistema nervoso central e outros. órgãos (ABBES et al., 2018).

CAPÍTULO 2. CIÊNCIAS ÔMICAS

2.1. Transcriptoma e RNA-Seq

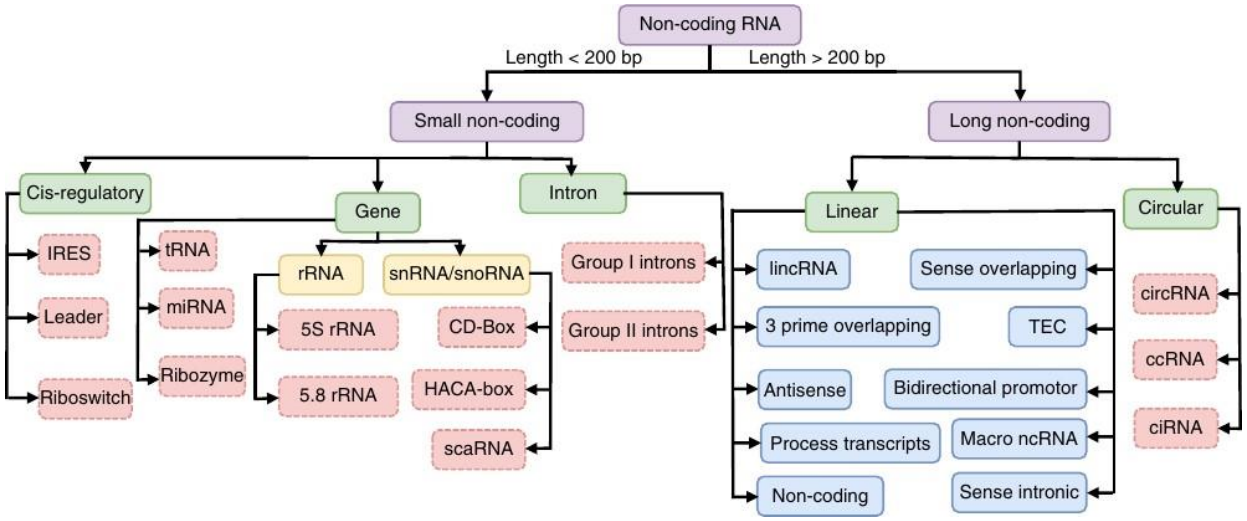
O que conhecemos como RNA-seq descreve uma coleção de métodos experimentais e computacionais para determinar a identidade e abundância de sequências de RNA em amostras biológicas. Assim, é identificada a ordem de cada nucleotídeo de adenosina, citosina, guanina e uracila presente em uma molécula de RNA, e desta forma definimos o conjunto total destes elementos como transcriptoma (WANG et al., 2009). Os métodos experimentais envolvem isolamento de RNA de células, tecidos ou amostras de animais inteiros, preparação de bibliotecas baseadas em amostras e o sequenciamento dos ácidos nucleicos (KORPELAINEN et al., 2014). Como resultado, esta tecnologia e suas metodologias proporcionaram a descoberta de novos transcritos, modelos de genes e novas classes de RNAs não codificantes (WASHIETL et al., 2011).

Ao longo das últimas décadas, esta tecnologia tornou-se uma ferramenta indispensável para a análise de expressão gênica diferencial e *splicing* diferencial de RNAms (STARK et al., 2019). No entanto, à medida que as tecnologias de sequenciamento de última geração se desenvolveram, seus métodos de análise também se aprimoraram para explorar diversos aspectos diferentes da biologia da molécula de RNA, como a expressão gênica de célula única (*Single Cell Sequencing*) (ANDREWS et al., 2021), os produtos de sua tradução (TAVARES-FERREIRA et al., 2022), sua estrutura molecular (XU et al., 2021) e sendo aplicadas também na medicina humana, auxiliando na identificação de agentes patogênicos e na abordagem mais eficiente para seu tratamento após o diagnóstico (WANG et al., 2022).

2.2. Transcritos

Os transcritos de RNA são moléculas de RNA que são produzidas a partir da transcrição do DNA. A transcrição é o processo pelo qual uma sequência específica de nucleotídeos do DNA é copiada para produzir uma molécula complementar de RNA (ALBERTS et al., 2010). Existem diferentes tipos de transcritos de RNA, incluindo o RNA mensageiro (RNAm), o RNA ribossômico (rRNA) e o RNA não codificador (ncRNA) (MATTICK, 2004). A principais classes de transcritos estão representadas na Figura 2.

Figura 2. Esquematização dos tipos de ncRNAs. Classificação dos diferentes transcritos em codificantes e não-codificantes de proteínas.



Fonte: AMIN et al., 2019

2.2.1. Transcritos codificantes de proteínas

Antes que ocorra a codificação de um transcrito em uma proteína, é necessário que ocorra o processo de transcrição. Em organismos eucariotos, isto é feito pelas enzimas RNA polimerase I, II ou III. O RNAm neste grupo de organismos é sintetizado pela RNA polimerase II a qual necessita de uma série de outras proteínas auxiliares, denominadas de fatores gerais de transcrição, para iniciar a transcrição sobre um DNA molde purificado (ALBERTS et al., 2010). Desta forma, para que um transcrito seja codificador de proteínas, é necessário que alguns fatores estejam presentes:

- Fases de leitura abertas (ORFs ou *Open Reading Frames*) (XU et al, 2021);
- Semelhança de sequência com proteínas ou domínios de proteínas conhecidos, modelos estatísticos de sequência de codificação de proteínas, taxas de substituição sinônimas versus não sinônimas (FRITH et al., 2006);
- Apresentarem quantidades maiores de guanina-citosina (GC) em regiões exônicas (OLIVER e MARÍN, 1996) sendo este, um critério para recentes ferramentas computacionais de predição de codificação proteica (XU et al., 2021).

2.2.2. Transcritos não-codificantes de proteínas

Por volta da década de 1940, os ácidos nucleicos já estavam consolidados dentro da comunidade científica, mas ainda havia um ceticismo dentre os cientistas sobre a capacidade destas moléculas em armazenar e, principalmente, transmitir informações sobre os seres vivos (DAHM, 2008). Este cenário começou a mudar quando pesquisadores descobriram - ao utilizar fungos como modelo de estudo - que a informação e conformação genética consistiam, principalmente, em instruções para a produção (hoje, codificação) de proteínas. Sendo estas versáteis macromoléculas que realizam a maioria das funções celulares, iniciando o que foi conhecido como Dogma Central da Biologia Molecular (DNA – RNA – Proteína) (PRAY, 2008). Desta forma, difundindo pelos anos seguintes o conceito que todo gene expressaria uma proteína responsável por alguma função em um organismo.

No entanto, a partir dos anos 1950 foram descritos RNAs que não possuíam uma função codificadora, mas eram capazes de exercer algum papel biológico, chamados de RNAs não codificantes (PALAZZO e LEE, 2015). Esses RNAs não despertavam a atenção dos cientistas até meados da década de 1970, quando ainda se acreditava que todo o aspecto funcional do genoma humano era representado pelas proteínas que ele codifica e alguns poucos RNAs não codificantes (como os rRNAs e tRNAs) (OHNO, 1972). Este conceito permaneceu difundido na comunidade científica até o início dos anos 80, quando foram descobertas novas classes de ncRNAs, os quais, possuíam importantes funções relacionadas a atividades enzimáticas, como a Ribonuclease P, uma precursora do tRNA (STARK et al., 1978).

2.2.2.1. RNAs pequenos não-codificantes de proteínas

Os sncRNAs (pequenos RNAs não codificantes) são transcritos de comprimento inferior a 200 pares de base que desempenham papéis importantes na regulação da expressão gênica e em processos biológicos, incluindo vias metabólicas, supressão de infecções virais e estágios do desenvolvimento (KAPRANOV et al., 2007; HOLLEY e TOPKARA, 2011; ROMANO et al., 2017, BATISTA DOS SANTOS, 2022). Devido à sua variedade de locais e mecanismos de ação, eles podem ser classificados em diferentes grupos. Entre eles, os *microRNAs* (miRNAs), *small nuclear RNAs* (snRNAs), *small nucleolar RNAs* (snoRNAs), *small interference RNAs* (siRNAs) e *piwi-interacting RNAs* (piRNAs), que interagem com as

proteínas PIWI (*P-element-induced wimpy testis*) (BROSNAN et al., 2007; GOMES et al., 2013). O grupo dos miRNAs é amplamente estudado, com cerca de 20 a 22 nucleotídeos reconhecidos como essenciais em praticamente todos os processos e desenvolvimento de doenças (CHEN et al., 2017; BARTEL, 2018). Sua ação consiste em mediar o silenciamento de genes pós-transcrição, inativando mRNAs e diminuindo sua tradução ou degradando-os (AALTO e PASQUINELLI, 2012; MENDELL e OLSON 2012; BARTEL, 2018).

2.2.2.2. RNAs longos não-codificantes de proteínas (lncRNA)

Na década de 1980, os cientistas utilizavam técnicas de hibridização diferencial de cDNA para clonar e estudar genes com padrões específicos de expressão tecidual e temporal. Inicialmente, os esforços se concentraram em genes que codificam proteínas conhecidas, mas posteriormente foi adotada uma abordagem a posteriori, sem levar em conta a potencial capacidade de codificação do RNA. Através desta abordagem, foi descoberto o primeiro gene não codificador de proteínas denominado H19 (*H19 imprinted maternally expressed transcript*) mas classificado como RNAm (JARROUX et al., 2017 *apud* PACHNIS et al., 1984).

O H19, foi inicialmente observado em *Mus musculus*, mas que foi caracterizado também em Homo sapiens através da demonstração de que seu gene não apresentava uma ORF conservada e que o transcrito não se associaria com a maquinaria de tradução (BRANNAN et al., 1990). Pouco tempo após esta descoberta, também foi sugerido que o lncRNA XIST (*X inactive specific transcript*) em humanos atuaria como um RNA funcional, com provável papel estrutural no núcleo, uma vez que seu gene também não codificava nenhuma proteína (BROWN et al., 1992).

2.2.2.2.1. <u>Estrutura dos lncRNAs</u>

Uma das principais características dos lncRNAs é a sua tendência de se dobrar em estruturas secundárias de maior ordem e termodinamicamente estáveis. Sua estrutura secundária e terciária são altamente conservadas e desempenham um papel importante no mecanismo de ação do RNA (NOVIKOVA, 2013). É sugerido pela comunidade científica que, de forma geral, os lncRNAs são organizados em um nível de complexidade estrutural intermediário, maior do que o observado em um RNAm comum, mas menor do que o

encontrado em ribossomos (ANDRIC et al., 2021).

Os lncRNAs também possuem a capacidade de formar ligações de hidrogênio, o que contribui para a formação de estruturas secundárias, incluindo duplas-hélices, grampos, loops, protuberâncias e pseudonós. Um dos elementos encontrados em lncRNAs é o chamado double stem-loop, que geralmente está associado à remodelagem da cromatina, uma função desempenhada por diversos transcritos desta classe (QIAN et al., 2019). Estruturas repetitivas também podem ser encontradas em lncRNAs, resultando na criação de domínios funcionais derivados de inserções e remodelagens de elementos repetitivos, como, por exemplo, no caso do lncRNA XIST, que contém uma região composta por estes elementos capazes de se dobrar em duas estruturas conhecidas como *stem-loop* (WUTZ et al., 2002).

2.2.2.2.2. Interações moleculares

Os lncRNAs são capazes de interagir com uma ampla variedade de moléculas, gerando diversos complexos RNA-proteína, RNA-DNA e RNA-RNA. As proteínas tendem a interagir com RNAs onde se formam estruturas secundárias complexas, em combinação com a abundância de proteínas ligadoras de RNA, o que proporciona uma ampla interface de comunicação entre o transcriptoma e o proteoma (ENTEZARI et al., 2022). Um dos exemplos mais conhecidos dessa interação é a associação entre o lncRNA HOTAIR (*HOX transcript antisense RNA*) e o PRC2 (*Polycomb Repressive Complex 2*), uma histona metiltransferase que atua no silenciamento gênico, além de repressão da transcrição independentemente do complexo proteico gerado (PORTOSO et al., 2017).

Os lncRNAs também são capazes de interagir com diferentes regiões do seu sítio de transcrição, sugerindo sua capacidade de regular padrões de expressão gênica tanto de forma local (cis) quanto distante (trans). Exemplos dessa habilidade incluem o lncRNA PAUPAR (*PAX6 upstream antisense RNA*), que é capaz de interagir com mais de 2.800 sítios localizados em múltiplos cromossomos, modulando sua atividade de transcrição em trans (VANCE et al., 2014) e o lncRNA KHPS1 (*SPHK1a antisense transcript*), que é capaz de se ancorar ao promotor do gene SPHK1 (*Sphingosine kinase 1*) e formar um tríplex entre RNA-DNA para induzir sua transcrição (POSTEPSKA-IGIELSKA et al., 2015). A associação genômica entre lncRNAs e cromatina pode ocorrer através do pareamento direto de bases entre RNA e DNA ou pelo pareamento de bases do lncRNA com sequências de RNA nascentes no locus que contém genes em processo de transcrição, que podem agir como guias

para o lncRNA até seu alvo em um genoma (PEREIRA, 2017).

Os lncRNAs também são conhecidos por seu papel na regulação da recomposição éxon-éxon. A formação de uma estrutura secundária conjunta entre duas moléculas de RNA, envolvendo um pareamento de bases intra e intermoleculares, permite que lncRNAs ocultem sítios de splicing e bloqueiem a montagem de um spliceossomo. Isso exige uma extensa complementaridade com a molécula de pre-RNAm. Além disso, a formação de um complexo lncRNA-RNAm pode aumentar a expressão do RNAm, encobrindo sítios de ligação de miRNAs. Em alguns casos, a ligação do lncRNA tende a suprimir a tradução do RNAm, como é o caso do 7SL RNA, que se liga na região 3' UTR do RNAm TP53 (ABDELMOHSEN et al., 2014). Outros lncRNAs também contêm repetições do tipo Alu, que são capazes de marcar diversos RNAm para serem degradados através da formação de sítios de ligação para uma proteína indutora de decaimento (GONG e MAQUAT, 2011).

2.2.2.2.2. Modificações químicas

Os lncRNAs apresentam modificações químicas internas, como é o caso da N6-metiladenosina (m6A), que é uma das modificações químicas mais prevalentes, junto com a N7-metilguanosina (m7G). A posição dessas modificações na molécula de lncRNA pode afetar a estabilidade e as estruturas secundárias dos RNAs. Essas alterações químicas podem ter importantes consequências para o processamento e função dos lncRNAs, como é o caso do XIST que requer m6A para exercer uma repressão transcricional eficiente (PATIL et al., 2016).

Também podem ser suscetíveis à modificações químicas internas 5-metilcitidina (m5C), pseudouridina (ψ) e uridilação (“cauda U”). A modificação m6A, por exemplo, é capaz de estabilizar ou desestabilizar as estruturas secundárias dos lncRNAs, dependendo de sua posição na molécula, e pode ter importantes consequências para o processamento e função dessa classe de RNAs, como no caso do lncRNA XIST que requer m6A para exercer uma repressão transcricional eficiente (PATIL et al., 2016). A pseudouridina é a modificação química mais abundante em ncRNAs, sendo propostos cerca de 2 mil sítios únicos para ela em lncRNAs (SHAFIK et al., 2015). Já a uridilação consiste na adição de uridinas na extremidade 3’ da molécula, em um processo que também já foi demonstrado em RNAm como importante para a degradação de alguns transcritos que não sofreram o processo de poliadenilação (PEREIRA, 2017).

2.2.2.2.3. Biogênese

O processo transcricional dos lncRNAs possui muitas características que também estão presentes em diversos RNAm: o *capping* da extremidade 5', a estrutura multiexônica, a poliadenilação da extremidade 3' e os mecanismos de *splicing* normal e alternativo. Da mesma forma como os RNAs codificantes de proteínas, a maioria dos lncRNAs anotados é transcrita pela RNA polimerase II, mas alguns também podem surgir pela RNAP-I ou RNAP-III e pela spRNA Pol IV *(single-plypeptide nuclear RNA polymerase IV*). Alguns também podem passar por formas alternativas de processamento, como splicing não sequencial de íntrons em um processo denominado *back-splicing* e *capping* de ambas as extremidades feito por snoRNAs (*small nucleolar RNAs*) (PEREIRA, 2017). Em alguns casos, os transcritos podem sofrer uma clivagem pós-transcricional que forma uma hélice tripla na extremidade 3', que atua protegendo a molécula madura de clivagem nucleotídica. Também podem ser transcritos tanto a partir de diversos loci do genoma mitocondrial, quanto de diversos loci genômicos em regiões intergênicas e nas orientações de senso e antisenso de genes codificantes de proteína. Com sítios de início de transcrição localizados em íntrons, éxons e regiões promotoras 3' e 5' de UTRs (*Untranslated Regions*). Em células humanas, existem evidências de clivagem pós-transcricional, nos quais, os lncRNAs pode agir como substrato para a geração de small RNAs ("RNAs pequenos") e como precursores de miRNAs, piRNAs, siRNAs e mascRNAs (MALAT1-associated small cytoplasmic RNA) e snoRNAs (WILUSZ, et al., 2012).

2.2.2.2.4. Localização intra e extracelular

A localização subcelular dos lncRNAs está diretamente relacionada com suas funções biológicas. Eles podem ser divididos em três grupos principais: (i) lncRNAs inteiramente nucleares localizados em cis; (ii) lncRNAs principalmente nucleares que atuam em trans; e (iii) lncRNAs que se localizam e funcionam amplamente no citoplasma. O primeiro grupo é composto por moléculas que possuem funções nucleares e se acumulam em seus sítios de transcrição em cis. No entanto, eles também podem ser subdivididos em transcritos que atuam tanto em cis quanto em trans. Os lncRNAs que atuam em cis podem ser moduladores ou modificadores da cromatina e formar tríplex RNA-DNA. Os que atuam em trans podem formar contatos transcromossômicos e modulam a arquitetura nuclear

(HACISULEYMAN et al., 2014).

O grupo dos lncRNAs principalmente nucleares é formado por transcritos que também se localizam no núcleo, mas que precisam ser realocados de seus sítios de transcrição para que possam exercer alguma influência na expressão gênica; que pode ocorrer tanto de modo global, como gene-específica. Isso ocorre devido ao seu acúmulo em corpos nucleares específicos, como ocorre com MALAT1 (*metastasis associated lung adenocarcinoma transcript 1*) (TRIPATHI et al., 2010). O terceiro e último grupo, consiste em lncRNAS, que precisam ser exportados para o citoplasma e, desta forma, consigam executar seus efeitos na expressão gênica, os quais, podem ocorrer de múltiplas formas, como pela interferência em modificações pós-traducionais, resultando em traduções de sinais aberrantes, além de estarem presentes em processos como a localização de proteínas, sua tradução e estabilização de mRNAs (LIN et al., 2016; LEE et al., 2016).

Já para o meio externo da célula, o transporte para locais distantes através de vesículas extracelulares, podendo ser detectados em sobrenadantes de culturas celulares, como em fluidos corporais como sangue e urina também são descritos para lncRNAs (ARITA et al., 2013; LORENZEN et al., 2015). Deste modo, podem ser secretados sob três modos: (i) empacotamento de vesículas envoltas por membranas, como exossomos e microvesículas, que podem ser secretadas sem que sejam degradadas por RNAses; (ii) liberação na circulação por células tumorais primárias e circulantes, células adjacentes ao tumor, células imunes e células sanguíneas; (iii) encapsulamento em HDL, em corpos apoptóticos ou associarem-se a complexos proteicos (SHI et al., 2016).

2.2.2.2.5. Padrões de expressão

Os lncRNAs, de modo geral são expressos comumente para a manutenção de funções celulares básicas e que apresentam uma alta especificidade, tanto tecidual como em estágio ontogênico de um organismo. Em moscas-de-fruta (*Drosophila spp.*) foi demonstrado que a expressão de lncRNAs é altamente restrita em determinados períodos quando comparada aos genes que codificam proteínas (CHEN et al., 2016). De modo semelhante, um padrão espaço-temporal é visto em diversos outros organismos de outros grupos zoológicos como primatas, nos quais os lncRNAs apresentam uma fundamental importância durante seu desenvolvimento através de sua ação conjunta com diversos RNAm (LI et al., 2015). Além disso, a expressão desta classe de transcritos não-codificantes é altamente dinâmica; onde

diversos lncRNAs podem apresentar um tempo de meia-vida mais curto em relação aos codificantes de proteína. Também existem os que podem apresentar um tempo maior de meia-vida, estando envolvidos em funções de genes de manutenção (TANI et al., 2012).

Seu nível transcricional também tende a ser baixo, o que corresponde com sua alta especificidade tecidual. Entretanto, estes níveis baixos também podem ser um reflexo de alterações em seus padrões de expressão, como no caso de alguns tumores, onde foi observado que há baixa expressão do lncRNA GAS6-A61 (*GAS6 antisense RNA 1*) em comparação a um tecido normal, está inversamente correlacionada à metástase em um linfonodo - o que também mostra o seu potencial para utilização em diagnóstico de doenças (HAN et al., 2013).

2.2.2.2.6. Reguladores pré-transcricionais ou epigenéticos

Quando ocorre o empacotamento de uma molécula de DNA eucariótico na cromatina consequências importantes acontecem para o processo de transcrição. Seus fatores de transcrição e outros componentes da maquinaria transcricional são dependentes de seu acesso para a fita de DNA, para que possam utilizá-la como um modelo de síntese de RNAs. Assim, a estrutura da cromatina pode ser considerada um aspecto crítico da expressão gênica. Dentro do núcleo, os lncRNAs podem regular a expressão gênica a nível pré-transcricional através da associação e recrutamento de complexos remodeladores da cromatina. Em mamíferos, as fêmeas possuem duas cópias do cromossomo X (XX), enquanto, nos machos, há somente uma (XY) (KHALIL et al., 2009). Esse desequilíbrio seria capaz de acarretar diversos problemas no momento da expressão de diversos genes associados a este cromossomo. Porém, o mecanismo de compensação de dose age para tornar igual a expressão dos genes encontrados em cromossomos sexuais para ambos os sexos. Logo, nas fêmeas, este mecanismo surge mais previamente durante seu desenvolvimento e ocorre principalmente, através do silenciamento do segundo cromossomo X (BROCKDORFF e TURNER, 2015). Neste sentido, o lncRNA XIST (*X inactive specific transcript*) possui uma atuação primordial neste silenciamento e nesta inativação cromossômica. Sua origem se dá no locus XIC (*X inactivation center*) e se acumula de modo cis ao longo de todo o cromossomo em que será transcrito para realizar o silenciamento. Ele age através da interação com o complexo PRC2 *(Polycomb Repressive Complex 2*), que é responsável pela trimetilação de histonas e pelo estabelecimento de estado transcricional reprimido dos genes (CERASE et al., 2015).

Contudo, o lncRNA HOTAIR (*Hox transcript antisense RNA*) é transcrito na fita oposta do cluster gênico HOXC (Homeobox C) e pode interagir com PRC2; mesmo atuando em trans através do recrutamento deste complexo para sua região promotora do gene HOXD (Homeobox D) e silenciando a expressão gênica (SARMA et al., 2014; RINN et al., 2007). Logo, os diferentes lncRNAs XIST e HOTAIR são capazes de interagir e recrutar o PRC2 em seu loci e em contextos genômicos diferentes, porém mantendo em comum sua capacidade de repressão da expressão de genes via processo de trimetilação da lisina em histonas (H3K27me3) (BROCKDORFF, 2013).

Existem lncRNAs que são localizados principalmente dentro do núcleo das células e podem atuar na regulação da expressão gênica a nível de transcrição, o que implica em mudanças na organização espacial de um determinado genoma e do nível da transcrição de genes que são codificantes de proteínas (VANCE e PONTING, 2014). Considerando isto, é possível também subdividir os lncRNAs em plncRNAs (*promoter-associated lncRNAs*) ou elncRNAs (*enhancer-like lncRNAs*), nos quais, ambos podem sofrer processo de poliadenilação e receberem o cap. Entretanto, ainda são indistinguíveis com relação ao seu tamanho em quantidade de nucleotídeos, número de éxons e orientação em relação ao gene codificante de proteínas (MARQUES et al., 2013).

Os plncRNAs são comumente descritos como transcritos com ~250-500 nt e provém de regiões proximais do sítio de início de transcrição gênica (TAFT et al., 2009). Sua atuação consiste no aumento ou diminuição do nível de expressão gênica, como ocorre com o gene DHFR (dihidrofolato redutase), que possui dois promotores de tamanhos maior e menor. O início do processo de transcrição no promotor menor irá gerar um lncRNA que interage tanto com o promotor maior quanto com seu fator de transcrição ligante e pode reduzir os níveis de expressão do gene DHFR (MARTIANOV et al., 2007), no caso, o TFIIB. Também há o lncRNA Evf-2 (*DLX6-AS1 – DLX6 antisense RNA 1*), que age como um plncRNA e atua em cis como um coativador da proteína Dlx-2 e forma um complexo estável que intensifica a atividade transcricional dos genes adjacentes no locus Dlx5/6 (FENG et al., 2006).

Já os elncRNAs, são lncRNAs derivados de regiões com atividades conhecidas como enhancer (“acentuadoras”) e tendem a ser mais restritos a um tecido específico. Não possuem abundância e tendem a ser menos conservados sob um ponto de vista evolutivo (MARQUES et al., 2013; RINN e CHANG, 2012). Um exemplo deste grupo é o conhecido como HOTTIP (*HOXA distal transcript antisense RNA*), que é transcrito a partir da região 5’ final do cluster gênico HOXA (*Homeobox A*) através de um looping cromossômico que o deixará

permanecendo em proximidade espacial com os genes presentes. Isto permite que interaja diretamente com a proteína WDR5 (*WD repeat-containing protein 5*) e recrutamento do complexo MLL (*Mixed-Lineage Leukaemia*), que é responsável por realizar a metilação de histonas na região promotora H3K4me3 e manutenção da ativação transcricional local (HAN et al., 2018).

2.2.2.2.7. Reguladores pós-transcricionais – núcleo celular

A nível pós-transcricional, os lncRNAs também podem exercer uma função regulatória através de diversos mecanismos. No núcleo das células, estes processos incluem sua atuação afetando três processos principais: (i) a estabilidade nuclear; (ii) o splicing de RNAm; (iii) a distribuição celular (VILLEGAS e ZAPHIROPOULOS, 2015; ANGRAND et al., 2015). O gene ZEB2 (*zinc finger E-box binding homeobox 2*) é transcrito em uma fita senso do DNA, na região do braço longo do cromossomo 2. Em sua fita antisenso se origina um lncRNA chamado ZEB2NAT (*ZEB2 antisense RNA1*), que pode esconder o sítio de splicing de um íntron na região 5' UTR do RNAm presente no ZEB2 devido ao pareamento por complementaridade. Esta interação não permite a ligação do spliceossomo, resultando na manutenção do íntron 5' UTR que possui um sítio interno de entrada ribossômica (IRES). Deste modo, ocorre a ligação da maquinária de tradução ao IRES resultando em uma maior eficiência da tradução da proteína Zeb2 (HUNG e CHANG, 2010).

2.2.2.2.8. Reguladores pós-transcricionais – citoplasma

No citoplasma, a regulação mediada pelos lncRNAs pode agir principalmente na estabilidade e degradação do RNAm ou também modulando o processo de tradução. Neste sentido, os lncRNAs podem atuar como captadores de miRNAs quando determinados RNAm e lncRNAs possuem sítios de ligação de miRNAs semelhantes. Assim, podendo ocorrer o pareamento de regiões complementares do lncRNA e do miRNA. Os lncRNAs que possuírem sítios de ligação de miRNAs em sua sequência, sequestram esses miRNAs e evitam que encontrem seus RNAm-alvos (ANGRAND et al., 2015). O gene supressor tumoral PTEN (*Phosphatase and Tensin Homolog*) apresenta um pseudogene localizado em outro cromossomo que possui uma sequência de nucleotídeos muito similar. Se acreditava que este pseudogene era biologicamente inativo, mas foi observado que ele pode ser transcrito e dá

origem a um lncRNA denominado PTENP1 (*Phosphatase and Tensin Homolog pseudogene 1*) e por conta de sua semelhança também foram observadas regiões conhecidas como seed conservadas de ligações de miRNAs na região 3' UTR do lncRNA PTENP1 para famílias dos miRNAs miR-17, miR-21, miR-214, miR-19 e miR-26. Logo, este lncRNA pode atuar como uma isca para atrair miRNAs que teriam como target o RNAm do gene PTEN (POLISENO et al., 2010).

2.2.2.2.9. Classificação

A definição sobre como classificar um lncRNA ainda gera uma discussão dentro da comunidade científica. Principalmente pela dificuldade em distingui-lo de outras classes de RNAs não codificantes como RNAms, após a descoberta de que alguns RNAS não codificantes longos são traduzidos em peptídeos menores (BÁNFAI et al., 2012) e alguns transcritos que se acreditava não serem codificantes apresentarem função de tradução de proteína (DINGER et al., 2011).

Porém, algumas características desta classe de RNAs permitem uma distinção mais clara sobre como diferenciá-lo (GUTTMAN et al., 2009; ULITSKY, 2016). Dentre elas, duas se destacam como características básicas dos lncRNAs: (i) a sequência de nucleotídeos do transcrito, que obrigatoriamente deve ser igual ou maior que 200 pares de base; (ii) não serem codificantes de proteínas. Também devem ter uma ORF ausente; o que pode sugerir que estes RNAs exerçam outras funções no meio (DINGER et al., 2008), como em fatores de regulação metabólica (MONGELLI et al., 2019).

2.2.2.2.9.1. Classificação baseada em comprimento

Dentre os métodos de classificação desta classe de ncRNAs, esta é uma das mais simples. Consistindo basicamente no comprimento estimado dos transcritos que não possuem capacidade de codificar uma proteína. No caso dos lncRNAs, um limite de 200 nucleotídeos seria o suficiente para que pudessem separá-los de outros small ncRNAs (ncRNAs $\geq$ 200 nt), dos longos ncRNAs (>200 nt) (KAPRANOV et al., 2007). Para isso, uma terceira categoria havia sido sugerida, denominada vlincRNAs (*very long intergenic noncoding RNAs*), que se correspondem a lncRNAs que apresentam uma sequência de nucleotídeos que pode variar entre 50 kb (kilobases) até aproximadamente 1Mb (Megabases) (KAPRANOV et al., 2010;

ST LAURENT et al., 2013). Também houve outras subclassificações propostas, como pequenos lncRNAs (200-950 nt), médios lncRNAs (950-4.800 nt) e longos lncRNAs (>4.800 nt). Com base no que foi observado no projeto ENCODE e considerando estas classificações, *H. sapiens* possui, em maioria, pequenos lncRNAs (58%) (MA et al., 2013).

2.2.2.2.9.2. Classificação baseada em comprimento

Por conta de suas funções biológicas serem definidas em grande parte por sua conformação tridimensional, foi proposto um modelo de classificação com base na estrutura secundária dos lncRNAs. Deste modo, poderiam ser classificados como tipo I (altamente estruturados, com motifs conformacionais complexos e com subdomínios); tipo ii (RNAs pouco estruturados, com múltiplas organizações stem-loop) e tipo iii (não estruturados e desordenados) (SANBONMATSU et al., 2016).

2.2.2.2.9.3. Classificação em localização intracelular

Alguns RNAs que estão associados à cromatina agem como componentes integrais de sua estrutura e podem apresentar um potencial regulatório de sua expressão gênica proximal (MONDAL et al., 2010); isso permitiu que os lncRNAs também fossem classificados conforme os compartimentos que eram encontrados no meio intracelular. Parte desta classificação foi proporcionada também pelo consórcio ENCODE (*Encyclopedia of DNA Elements*), o qual pode identificar milhares de lncRNAs presentes no núcleo e no citoplasma em 15 linhagens de células e, também, possibilitou a caracterização da composição de sua fração subnuclear: composta pelo nucléolo, nucleoplasma e cromatina. (DJEBALI et al., 2012; ENCODE, 2012).

2.2.2.2.9.4. Classificação em elementos repetitivos

Alguns RNAs que estão associados à cromatina agem como componentes integrais de sua estrutura e podem apresentar um potencial regulatório de sua expressão gênica proximal (MONDAL et al., 2010); isso permitiu que os lncRNAs também fossem classificados conforme os compartimentos que eram encontrados no meio intracelular. Parte desta classificação foi proporcionada também pelo consórcio ENCODE (*Encyclopedia of*

DNA Elements), o qual pode identificar milhares de lncRNAs presentes no núcleo e no citoplasma em 15 linhagens de células e, também, possibilitou a caracterização da composição de sua fração subnuclear: composta pelo nucléolo, nucleoplasma e cromatina. (DJEBALI et al., 2012; ENCODE, 2012).

2.2.2.2.9.5. Classificação funcional

Embora não codifiquem proteínas, os lncRNAs podem exercer diversas funções no meio onde estão localizados e alguns podem ser subclassificados com base nesta característica. Os ncRNA-a ("ncRNAs ativadores"), apresentam funções semelhantes aos RNAs potenciadores, mas somente regulam positivamente genes que estão próximos de sua localização (ST LAURENT et al., 2015). Também existem aqueles que são classificados como ceRNAs (*competing endogenous RNAs*) que são lncRNAs que apresentam similaridade com sequências de transcritos codificantes de proteínas e podem atuar como competidoras por moléculas regulatórias (TAY et al., 2014). Os lncRNAs também podem agir como moléculas precursoras de outros RNAs mais curtos, como ocorre com o lncRNA H19, que precede o miRNA miR-675 (KENIRY et al., 2011). Alguns lncRNAs também podem codificar alguns peptídeos de menor tamanho e se associarem com ribossomos, porém, isto não implica em uma síntese proteica, mas em sua regulação durante a tradução (VAN HEESCH et al., 2014; ST. LAURENT et al., 2015).

2.2.2.2.10. Arquétipos de ação de molecular

2.2.2.2.10.1. Arquétipo Sinalizador

Como seu nome sugere, os lncRNAs que pertencem à esta classe podem agir como sinalizadores moleculares, devido sua expressão ocorrer em locais e momentos específicos. Alguns podem agir realizando uma rápida regulação – o que se trata de função exercida comumente por proteínas e faz com que determinado processo seja executado de modo mais rápido por conta dos lncRNAs não necessitarem da tradução proteica. Outros lncRNAs que se tornam somente subprodutos do processo de transcrição também podem apresentar papeis de regulação por conta de sua iniciação, alongamento ou término de sua sequência (WANG e CHANG, 2011).

O lncRNA PANDA (*P21 associated ncRNA DNA damage activated*) que está localizado no sítio de início da transcrição do gene CDKN1A (*cyclin dependent kinase inhibitor 1A*), no qual, é ativado por meio de algum dano ao DNA de maneira da proteína p53, irá interagir com o fator de transcrição NF-YA (*nuclear transcription factor Y subunit alpha*) e ativará um processo de apoptose; o que habilita a parada do ciclo celular por meio da limitação da expressão de genes pró-apoptóticos (HUNG et al., 2013).

2.2.2.2.10.2. Arquétipo Isca

Esta categoria corresponde a lncRNAs que se ligam e captam outras moléculas que podem ser, geralmente, proteínas ou outros RNAs e inibem a sua ação. Em síntese, este arquétipo age como um regulador negativo de um efetor funcional, ou seja, consegue diminuir a atividade de determinada molécula efetora. Assim, em teoria, este processo de nocaute de um lncRNA deste grupo conseguiria mimetizar algum efeito de ganho de função de sua respectiva molécula que exerceria alguma função (HUNG et al., 2013).

Um exemplo de um lncRNA que pertence a este grupo é o GAS5 (*growth arrest specific 5*). Um transcrito que possui ~600nt e que é sintetizado pela RNA polimerase II. Este lncRNA interage com glicocorticoides que se ligam e ativam o receptor GR, que é um fator de transcrição que, quando se conecta ao seu ligante, age como um ativador transcricional e se liga a elementos que respondem a glicocorticoides (GRE), localizados na região promotora de genes-alvo. Quando o GAS5 interage com GR, ele irá o reprimir através da formação de um *motif* de RNA que irá mimetizar seu *motif* de DNA equivalente. Esta repressão ocorrerá devido a competição e ligação direta desse lncRNA ao domínio do GR que se liga à molécula de DNA, o que impedirá que se associe com seu GRE (KINO et al., 2010).

2.2.2.2.10.3. Arquétipo Guia

Os lncRNAs desta classe também são conhecidos como “guias”, os quais também exercem a ligação e recrutamento de riboucleoproteínas e as redirecionam para locais específicos em seus respectivos genes-alvo. Após este encaminhamento, ocorrerão alterações na expressão gênica em modo cis e em modo trans. As moléculas efetoras com ação catalítica que são conduzidas por lncRNAs podem ser complexos que ativam proteínas do grupo TxG (Trithorax), complexos repressores como proteínas do grupo PcG (*Polycomb group*) ou um

conjunto de combinações de fatores de transcrição. No caso da ocorrência de algum nocaute de um lncRNA do tipo guia, as possíveis predições seriam de alguma interferência na localização da molécula efetora ou até mesmo da perda de função de seu próprio efetor (HUNG et al., 2013).

Na brassicácea *Arabdophis thaliana*, quando submetida ao frio do inverno na região onde ocorre, alguns processos de metilação denominados H3K27me3 irão ocorrer em sua cromatina do repressor floral FLC (*flowering locus c*), resultando em sua repressão epigenética. Este processo é mediado por um lncRNA denominado COLDAIR (*cold assisted intronic noncoding RNA*), que interage fisicamente com um componente do complexo repressivo PRC2 (*Polycomb Repressive Complex 2*) e o redireciona para o locus onde está localizado o FLC. Então, o lncRNA COLDAIR, o qual é transcrito a partir de um íntron que está localizado no mesmo sentido de seu gene-alvo; assim, ele atuará como um guia para que afete o silenciamento no locus onde está o FLC e resultando o efeito biológico denominado vernalização (HUNG et al., 2013) – que consiste na indução das plantas a florescerem através da exposição a temperaturas baixas não congelantes. A descoberta desta interação entre lncRNAs e plantas pode sugerir que a relação entre ncRNA e PRC2 seja um mecanismo evolutivamente conservado de repressão gênica (HEO e SUNG et al., 2011).

2.2.2.2.10.4. Arquétipo Andaime

Era comum acreditar que as proteínas eram os elementos bioquímicos responsáveis por formarem alicerces moleculares, nos quais outras moléculas poderiam se acoplar para que ocorressem a montagem de outros complexos funcionais (GOOD et al., 2011). Porém, os lncRNAs que pertencem ao grupo dos andaimes apresentam uma função semelhante – em que são plataformas centrais para que outros componentes moleculares sejam anexados. Este arquétipo apresenta uma grande complexidade funcional, por conta de seus lncRNAs possuírem domínios diferentes que podem interagir com efetores distintos. Deste modo, um transcrito desta classe pode se ligar a diversos outros efetores de modo simultâneo e, assim, unificar, ao mesmo tempo e local, efetores com atividade de estímulo ou repressão transcricional. Uma possível previsão para ocorrer um processo de nocaute deste arquétipo pode incluir interferência na localidade de sua molécula efetora ou até mesmo pela perda de função do próprio efetor, uma vez que ocorreria sua dissociação do complexo lncRNA e seus efetores (WANG e CHANG, 2011).

O processo de repressão transcricional por meio da interação entre lncRNA e complexos modificadores de cromatina também podem ser encontradas no locus INK4b/ARF/INK4a. A expressão deste locus de supressão tumoral é dependente da metilação direcionada por proteínas do PcG. Deste mesmo locus irá se originar o lncRNA ANRIL (*antisense RNA in the INK4 locus*), devido à atuação de fatores de transcrição. As proteínas que pertencem aos complexos PRC1 e PRC2 (*Polycomb Repressive Complex* 1 e 2) apresentam domínios de ligação de RNAs em suas subunidades CBX7 e SUZ12, respectivamente. Logo, este lncRNA pode interagir com a subunidade SUZ12 de PRC2 para que ocorra a indução da metilação de H3K27 e o silenciamento de seu locus. Além disso, ANRIL também é capaz de se ligar à CBX7 em PRC1, o que proporciona a manutenção de seu silenciamento (ZHU et al., 2011). Assim, o lncRNA ANRIL, agindo como um andaime, recruta múltiplos conjuntos de complexos modificadores de cromatina para modular dinamicamente sua atividade transcricional (WANG e CHANG, 2011; KUGEL e GOODRICH, 2012).

Embora existam estes arquétipos que buscam contribuir para a compreensão, organização e definição sobre o que são os lncRNAs, muitos destes ncRNAs apresentam combinações destes arquétipos que podem desempenhar diversas funções biológicas. Considerando isto, por apresentarem uma resposta à estímulos endógenos ou exógenos, também possuem sítios para a ligação de complexos modificadores de cromatina e expressão gênica. Como por exemplo, o lncRNA previamente citado HOTAIR ocorre por um fenômeno múltiplo - regulatório. Por ser transcrito preferencialmente em células localizadas em regiões posteriores e distais, apresentará uma ação de especificação anatômica. Também por ser capaz de se ligar aos complexos PRC2 e LSD1, age como um “esqueleto molecular” que fornecerá o alicerce para que ocorram a acoplagem de outras moléculas que formarão outros complexos. Além disso, por também conseguir direcionar PRC2 para seu locus adequado, também pode atuar como um arquétipo guide (WANG e CHANG, 2011).

2.2.2.2.11. Métodos de identificação *in silico*

Distinguir lncRNAs por abordagens computacionais é um desafio devido às similaridades que os lncRNAs apresentam aos transcritos codificantes de proteínas. Na última década, diversos trabalhos coletaram lncRNAs oriundos de bancos de dados biológicos como RefSeq, Ensembl Genomes, GENCODE entre outros, com a finalidade de construir datasets específicos de lncRNAs. Para tornar o processo de predição de lncRNAs mais fácil, algumas classes para classificá-los foram estabelecidas: (i) recursos relacionados aos códons, (ii) recursos relacionados às regiões de leitura abertas (ORFs), (iii) características relacionadas à quantidade de guanina-citosina (GC), (iv) recursos relacionados às sequências de codificação, (v) recursos relacionados ao processo de transcrição, (vi) recursos relacionados à suas características estruturais. Muitos métodos computacionais foram desenvolvidos para prever lncRNAs, e eles utilizam diferentes recursos e diferentes técnicas de aprendizado de máquina para construir modelos de predição (Tabela 1). Desta forma, as ferramentas computacionais que utilizam estas classes para classificar os lncRNAs podem utilizar três abordagens específicas: (i) métodos de classificação binários, (ii) métodos baseados em aprendizagem profunda, (iii) aprendizado de máquinas baseado em conjuntos (XU et al., 2021).

Dentre os programas que fornecem esta predição, estão os programas:

1. *Coding Potential Assessment Tool* (CPAT): ferramenta que não considera o alinhamento de sequências, mas quatro características: cobertura e tamanho das regiões de leitura abertas (ORFs), pontuação no teste *Fickett* e pontuação de uso de hexâmeros (WANG et al., 2013);
2. *Coding Potential Calculator* (CPC2): utiliza uma abordagem de aprendizado de máquinas denominada Máquina de Vetores de Suporte (SVM) para classificar as sequências não codificantes com base em sua pontuação no teste *Fickett*, o comprimento e integridade de suas ORFs e seu ponto isoelétrico (pI). Enquanto a pontuação do teste *Fickett* é derivada da frequência de nucleotídeos do transcrito completo, as outras três características são calculadas com base na ORF putativa mais longa identificada *in silico* para que a ferramenta classifique as sequências como codificantes ou não (KANG et al., 2017);
3. RNAsamba: ao contrário das ferramentas anteriores, utiliza abordagens de *Aprendizagem Profunda*, na qual, estabelece sua classificação por meio de redes neurais artificiais se baseando na identificação de padrões presentes nas sequências e nas ORFs. Assim,

possibilitando a identificação de padrões que diferenciam os transcritos quanto sua capacidade de codificar uma proteína. Dentre estes padrões, estão a frequência de *k-mers*, aminoácidos e o comprimento das ORFs existentes (CAMARGO et al., 2020).

Tabela 1. Métodos *in silico* para prospecção de lncRNA.

FERRAMENTA	PARÂMETROS	CLASSE	MODELO
CPC2	Teste *Fickett* Comprimento de ORFs Integridade das ORFs Ponto Isoelétrico	Aprendizagem de Máquinas	Máquina de Vetores de Suporte
CPAT	Tamanho de ORFs Cobertura de ORFs Teste *Fickett* Viés do uso de Hexâmeros	Aprendizagem de Máquinas	Regressão Logística
RNAsamba	Proteínas putativas Frequência de *k-mers* Frequência de aminoácidos Comprimento de ORFs	Aprendizagem Profunda	Rede Neural Recorrente

2.2.2.2.12. lncRNAs envolvidos em doenças

Os lncRNAs são moléculas que estão envolvidas em processos biológicos cruciais como a proliferação celular, diferenciação, regulação epigenética, remodelagem da cromatina, *splicing*, tradução e diversos fatores do metabolismo. Devido sua atuação em diversos contextos biomoleculares, quando ocorrem alterações ou desregulagem destes ncRNAs, eles podem, também, estar associados ao surgimento e desenvolvimento de diversas doenças complexas, como desordens neurológicas, diabetes e doenças cardiovasculares (PEREIRA, 2017).

Não obstante, o câncer é uma patologia onde os lncRNAs são amplamente estudados; uma vez que podem ser causados também por alterações genéticas que influenciam nas vias de expressão gênica que são responsáveis pela manutenção da homeostase do meio celular. Por serem tecido-específicos, os lncRNAs, quando correlacionados com conjuntos distintos de genes que influenciam em funções que determinam o fenótipo transformado das células do câncer. Diversos destes lncRNAs são transcricionalmente por oncogenes ou fatores de supressão tumoral, enquanto, outros transcritos do grupo exercem estas mesmas funções. O

lncRNA MIR31HG (*MIR31 host gene*) é altamente expresso em adenocarcinomas ductais pancreáticos, estando, também, envolvido na indução da proliferação e invasão das células tumorais (YANG et al., 2016). Do mesmo modo, o lncRNA CCAT2 (*colon cancer associated transcript 2*) em carcinoma hepatocelular (HCC), que, além de promover a migração e proliferação celular, também está envolvido durante a inibição do processo de apoptose das células HCCs de modo *in vitro* (ZHOU et al., 2016). Porém o lncRNA ADAMTS9-aS2 (*ADAMTS9 antisense RNA 2*) participa da supressão tumoral inibindo a migração celular (YAO et al., 2014) e, similarmente, o ncRNA TUG1 (*taurine up-regulated 1*) possui ação na indução de apoptose em células do glioma (LI et al., 2015). Estes fatores os tornaram um interessante objeto de pesquisa como biomarcadores, tanto para a detecção e classificação de diferentes tipos de subclasses de tumores, quanto para realizar a predição de respostas aos seus tratamentos.

Embora ainda seja uma grande dificuldade utilizá-los desta maneira, existem moléculas que apresentam um grande potencial para esta aplicação: como o previamente citado H19 no câncer de mama (ZHANG et al., 2016), o MALAT1 durante a metástase de tumor gástrico (HEMASHETTAR et al., 2011) e os LINC00152 (*cytoskeleton regulator RNA*), CFLAR-AS1 (*CFLAR antisense RNA 1*) e POU3F3 (*POU class 3 homeobox 3*) em carcinomas esofágicos (HU et al., 2016).

Já em cânceres de próstata, o chamado antígeno prostático específico (PSA) é muito utilizado como biomolécula para rastreamento de neoplasias, entretanto, possui a limitação de ser expressa somente no tecido da próstata e não necessariamente em um tumor nesta glândula. Neste contexto, o lncRNA PC3 (*prostate cancer antigen 3*) surge como um potencial candidato a um biomarcador, uma vez que não é expresso em todos os tecidos normais ou malignos que não tenham sua origem na próstata e sendo sua expressão significativamente maior em pacientes que apresentam este tipo de tumor em comparação a outros pacientes que apresentam outras disfunções nesta região (BUSSEMAKERS et al., 1999; KOK et al., 2002). Este lncRNA também pode ser detectado na urina e é um dos biomarcadores mais específicos para esta patologia, levando à criação do teste PROGENSA PCA3: o primeiro teste de diagnóstico molecular em amostras de urina aprovado pelo órgão estadunidense FDA (*Food and Drug Administration*) (QI et al., 2016). Entretanto, a área terapêutica em tumores tem atingido um grande avanço no desenvolvimento de oligonucleotídeos aprimorados, que são capazes de alvejar a célula tumoral in vivo de modo mais efetivo, estável e menos tóxico. Estas moléculas têm a semelhança de serem mediadoras

a degradação de lncRNAs específicos e bloquear a atividade dos lncRNAs pela ligação a elas ou pela ligação dos lncRNAs com proteínas específicas. Também existe a abordagem para aproveitar a alta expressão específica de alguns transcritos em células tumorais para diminuir os riscos de tecidos normais serem afetados pelo tratamento (PEREIRA, 2017).

2.2.2.2.13. lncRNAs em fungos

Em fungos, a maior parte dos lncRNAs é descrita por agir em cis, sendo o modo trans raramente descrito. Um exemplo ocorre em *S. cerevisiae*, em que uma via de mecanismo trans foi declarada para a transcrição estável não anotada (SUT) SUT457, que interage fisicamente com 12 genes essenciais para organização telomérica e manutenção da homeostase (KYRIAKOU et, al., 2022). Além disso, recentemente, um lncRNA denominado HAX1 foi identificado como um transativador da expressão de celulase na espécie *Trichoderma reesei* (TILL et al., 2018).

Curiosamente, estes lncRNAs causam um impacto positivo em seus genes-alvo, assim como, genes trans que causam um efeito repressor também são encontrados: membros da classe da transcrição instável Xrn1 (XUTs), como TY1, frequentemente mediam o silenciamento de trans-genes alvo localizados (BERRETTA et al., 2009). Evidenciando, assim, as importantes funções de regulação metabólica dessa classe de RNAs também presentes nestes organismos.

Ainda em *S. cerevisiae*, os lncRNAs IRT1, RME2 e RME3, possuem uma atuação muito importante na regulação da meiose. Os três são semelhantes quanto à sua função e mecanismos regulatórios. o IRT1 inibe a expressão do sentido localizado à jusante e parcialmente sobreposto ao gene indutor da meiose 1, IME1. Evitando assim, a diferenciação errônea de células germinativas e esporulação em células haploides. Em células cultivadas em estado haploide, o fator de transcrição repressivo da meiose Rme1 (repressor de IME1) induz a produção do IRT1 que age como mediador da deposição de marcas de cromatina repressivas por recrutando a histona metiltransferase Set2 e a histona desacetilase Set3, que leva a um bloco de iniciação da transcrição ação de IME1. Além disso, IRT1 impede a ligação do gene Pog1, responsável por iniciar processos de transcrição da levedura (VAN WERVEN et al., 2012).

2.2.2.2.14. lncRNAs em fungos patogênicos humanos

A presença desta classe de RNAs não codificantes em fungos patogênicos humanos ainda permanece pouco explorada pela comunidade científica. Há somente um número bastante limitado de trabalhos que buscam descobrir a influência de lncRNAs na virulência destes organismos durante seu processo infeccioso. O primeiro estudo a trazer evidências sobre sua influência em sua patogenicidade foi feito por Chacko e colaboradores (2015), no qual descreveu o lncRNA RZE1, que está relacionado com o desenvolvimento das hifas no fungo patogênico *Cryptococcus neoformans*. Mais recentemente, o lncRNA denominado DINOR foi confirmado como atuante em fatores de virulência durante a infecção pelo patógeno multirresistente *Candida auris* (GAO et al., 2021) e a quantificação desta classe de transcritos não-codificantes presentes em outras espécies do gênero foi realizada em um estudo de Hovhannisyan e Galbadón (2021), porém, sem a descrição de quais são os lncRNAs presentes no grupo.

Desta forma, ainda são necessários mais estudos que tenham por objetivo ampliar o limitado conhecimento disponível deste grupo de RNAs não codificantes de proteínas que estão presentes no Reino Fungi.

3. OBJETIVO

O presente trabalho teve como objetivo primário realizar a prospecção e caracterização *in silico* de RNAs longos não-codificantes de proteínas (lncRNAs) de diferentes espécies de fungos patogênicos humanos através de transcriptomas depositados no banco de dados *Transcriptome Shotgun Assembly Sequence Database* (TSA), disponível no *National Center for Biotechnology Information* (NCBI).

3.1. Objetivos específicos

- Identificar dados de transcriptoma de diferentes espécies de fungos patogênicos humanos depositados no TSA;
- Desenvolver um fluxo de trabalho reprodutível para a prospecção e caracterização dos lncRNAs;
- Identificar e classificar os transcritos codificantes de proteínas;
- Identificar e classificar os transcritos não-codificantes de proteínas;
- Anotar funcionalmente os transcritos codificantes de proteínas presentes nos transcriptomas;
- Anotar funcionalmente os ncRNAs presentes nos transcriptomas;
- Realizar a prospecção de lncRNAs nos transcriptomas identificados na base de dados TSA;

4. METODOLOGIA

4.1. Obtenção e identificação de dados de transcriptoma

A metodologia utilizada para obtenção das sequências brutas dos fungos patogênicos para humanos foi composta de duas etapas principais: *download* de dados e seleção de espécies.

Etapa 1: *Download* de dados

- Foi utilizado o repositório *Transcriptome Shotgun Assembly Sequence Database* (TSA; https://www.ncbi.nlm.nih.gov/genbank/tsa/) mantido pelo *National Center for Biotechnology Information* (NCBI; https://www.ncbi.nlm.nih.gov/) para obtenção das sequências brutas dos fungos.
- Foi realizada uma busca utilizando o termo "*fungi*" para obtenção das sequências. Além disso, foi realizado o filtro "*taxonomic groups; Fungi*" e selecionado os projetos do tipo "TSA".

Etapa 2: Seleção de espécies

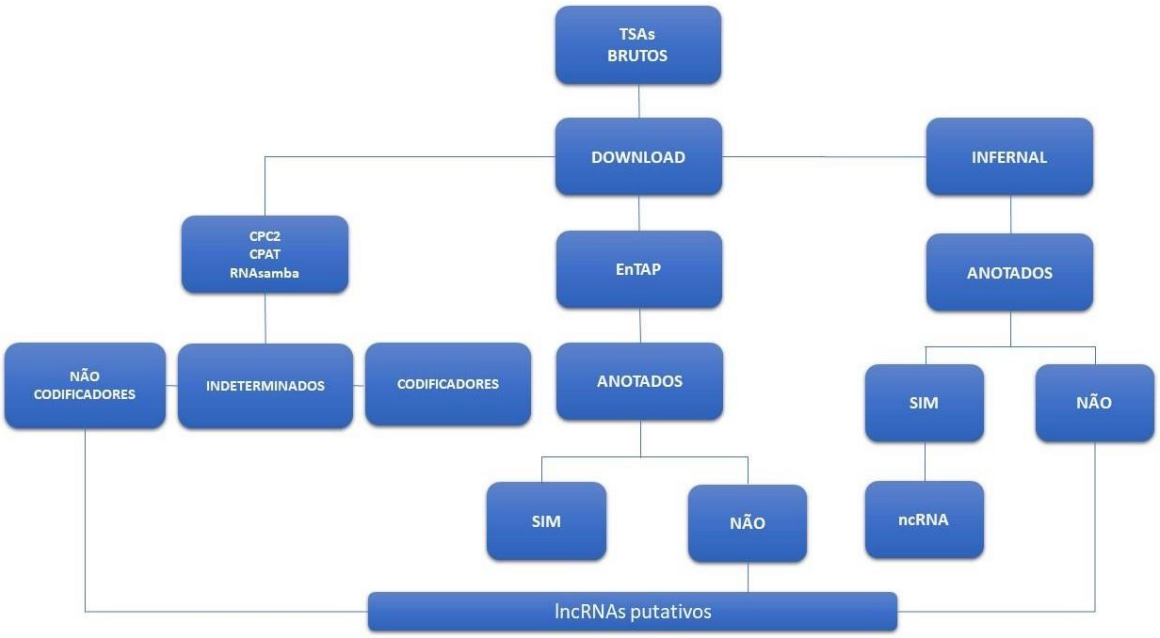
- Para selecionar quais espécies seriam utilizadas para as análises, foi realizada uma filtragem baseada em evidências na literatura de fungos com patogenicidade confirmada para humanos.
- Os critérios de seleção foram baseados em estudos prévios que comprovaram a patogenicidade das espécies selecionadas para humanos.

Essa metodologia garantiu a obtenção de sequências confiáveis e relevantes para as análises, uma vez que foram selecionadas apenas espécies com patogenicidade comprovada para humanos.

4.2. Fluxo de trabalho de bioinformática

Um fluxo de trabalho de bioinformática foi desenvolvido para prever lncRNAs e ncRNAs nos transcriptomas identificados. Este fluxo é descrito na Figura 3.

Figura 3. Fluxo de trabalho para prospecção de lncRNAs. A metodologia possibilita a identificação de sequências codificadoras e não-codificadoras de proteínas, com a possibilidade de posterior anotação funcional.



4.2.1. Análise de potencial de codificação

No presente estudo, foi realizada a predição do potencial de codificação proteica dos transcritos através da utilização de três ferramentas computacionais específicas: *Coding Potential Assessment Tool* (KANG et al., 2017), *Coding Potential Calculator* (WANG et al., 2013) e RNAsamba (CAMARGO et al., 2020). Estas ferramentas foram acessadas através de suas respectivas interfaces gráficas online, e forneceram como principal resultado o potencial codificante de cada sequência.

Os transcritos foram submetidos às ferramentas, e os resultados foram organizados baseando-se em três categorias: (i) aqueles transcritos que foram anotados como "não codificantes" em unanimidade, ou seja, com 100% de concordância pelas três ferramentas; (ii)

as sequências que foram anotados como “codificantes” em unanimidade, ou seja, com 100% de concordância pelas três ferramentas; (iii) as sequências que não apresentaram concordância entre os três programas e foram classificadas como “indeterminadas”.

Cada uma das ferramentas utilizadas analisou as sequências presentes nos transcriptomas através de diferentes algoritmos de análise (Tabela 1) para classificar cada uma como sendo ou não capaz de codificar uma proteína. Por definição, uma das características dos lncRNAs é o seu comprimento em nucleotídeos, maior ou igual a 200 pares de base. Dado que para o depósito de novos transcriptomas ao banco de dados TSA do NCBI é necessário que todo transcrito tenha pelo menos 200nt de tamanho, todos os transcritos utilizados neste trabalho possuiam o tamanho necessário para se enquadrar nesta categoria. Assim, toda sequência classificada pelos programas de cálculo de potencial codificador como “não codificante”, foi considerada inicalmente como um possível lncRNA.

4.2.2. Anotação funcional dos transcritos codificantes de proteína

A pipeline de anotação funcional foi realizada utilizando a ferramenta *Eukaryotic Non-Model Transcriptome Annotation Pipeline* (EnTAP) (HART et al., 2019). Essa ferramenta baseia-se em princípios de ontologia e similaridade de genes para realizar a anotação, utilizando bancos de dados pré-configurados através de um *script* desenvolvido no Laboratório de Bioinformática e Ciências Ômicas (LaBiOmics) para execução em uma interface CLI (*Command Line Interface*) Linux. Durante esse processo foram utilizados todos os transcritos contidos no transcriptoma dos fungos patogênicos selecionados para esse trabalho.

Durante sua execução, o fluxo de trabalho da ferramenta EnTAP incluiu uma etapa de busca por similaridade contra um banco de dados de proteínas no formato DIAMOND (BUCHFINK et al., 2014), utilizando o banco eggNOG (JENSEN et al., 2007) e um conjunto de dados customizado contendo nucleotídeos de organismos do Reino Fungi depositados no NCBI.

Outra etapa importante foi a identificação de grupos de genes ortólogos, utilizando as famílias de genes presentes no eggNOG, domínios proteicos depositados no SMART (*Simple Modular Architecture Research Tool*) (LETUNIC, 2015) e Pfam (FINN et al., 2014), termos do *Gene Ontology* (ASHBURNER et al., 2000) e vias metabólicas do KEGG (KANEHISA et al., 2000).

4.2.3. Identificação dos ncRNAs putativos

Para identificar os transcritos correspondentes aos ncRNAs presentes nos transcriptomas de fungos patogênicos estudados, utilizamos o módulo *cmscan* do programa INFERNAL (NAWROCKI et al., 2009). Este módulo é especializado na busca de sequências em coleções de modelos de covariância (CMs). Durante o processo, cada sequência contida nos transcriptomas de referência foi utilizada para pesquisar o banco de dados de destino de CMs do Rfam 11.0.2 2.09.2013 (GRIFFITHS-JONES et al., 2003) e gerar uma lista de classificação de saída dos CMs com as correspondências mais significativas para os transcrtitos de entrada.

4.2.4. Definição de lncRNA putativo para esse trabalho

Durante o presente trabalho, utilizamos a definição apresentada por Menegidio et al. (2021) e Barbosa et al. (2022) para a classificação dos lncRNAs putativos. Nesse sentido, tratamos como lncRNA putativo todos aqueles transcritos que:

1. Foram classificados como não codificantes de proteínas pelas três ferramentas de CP;
2. Não tiveram uma anotação funcional durante a análise pelo EnTAP;
3. Não tiveram uma anotação funcional durante a análise pelo INFERNAL.

4.2.5. Identificação de conservação

Para identificar lncRNAs conservados, alinhamos as sequências de lncRNA putativos de *Candida africana* e *Candida albicans* contra os lncRNAs apresentados por Hovhannisyan e Gabaldón (2021). Também foram alinhados os lncRNAs de *Paracoccidioides brasiliensis* Pb18 contra os lncRNAs apresentados por Menegidio et al. (2021). A presente análise teve como base inicialmente o alinhamento utilizando o programa *minimap2* (LI, 2018), além das ferramentas *Samtools* (LI et al., 2009) e *Bedtools* (QUINLAN, 2014). Após o processo de alinhamento, os arquivos gerados foram filtrados com o parâmetro "*Samtools view -q 20*" para remover aqueles com uma pontuação de alinhamento MAPQ menor que 20, removendo assim quaisquer alinhamentos de baixa qualidade.

5. RESULTADOS

5.1. Obtenção e identificação de dados de transcriptoma

Para o presente estudo, foi realizada uma busca utilizando critérios específicos no banco de dados de TSA do NCBI, resultando na obtenção de um total de 187 transcriptomas brutos de fungos. Após a análise desses dados, foram selecionados 19 transcritos pertencentes a 15 espécies diferentes, conforme mostrado na Tabela 2. Essa seleção foi baseada em critérios previamente descritos no tópico da Metodologia, levando em consideração fatores como a disponibilidade de sequenciamento para diferentes cepas. Os transcritos que não foram selecionados para este estudo não foram utilizados nas análises posteriores.

Tabela 2. Transcriptomas disponíveis na base de dados TSA. Lista dos transcriptomas utilizados no presente trabalho com evidência de patogenicidade em humanos baseada em literatura.

ESPÉCIE	TSA ID	ASSEMBLY ID	REFERÊNCIA
Alternaria sp. MG1	GEMY01	MG1	Pastor e Guarro, 2008
Arthrinium phaeospermum	GHWG01	M.B.Ellis	Li et al., 2020
Aureobasidium melanogenum	GEEO01	CBS 110374	Wang et al., 2019
Aureobasidium melanogenum	GIXY01	CBS 110374	Wang et al., 2019
Candida africana	GEVV02	NA	Gharehbolagh et al., 2020
Candida albicans	GEVW02	SC5314	Nobile e Johnson et al., 2015
Geotrichum candidum	GGKD01	LMA-244	Keene et al., 2019
Irpex lacteus	GBRE01	F17	Chaudhary, 2015
Mucor irregularis	GFBC01	B50	Zhang et al., 2013
Mucor irregularis	GHMU01	B50	Zhang et al., 2013
Papiliotrema laurentii RY1	GFWE01	RY1	Asadzadeh et al., 2020
Paracoccidioides brasiliensis Pb18	GHTY01	Pb18	Menegidio et al., 2021
Penicillium janthinellum	GBSP01	NCIM1366	Li et al., 2020
Penicillium janthinellum	GBSQ01	NCIM1366	Li et al., 2020
Penicillium janthinellum	GIPK01	NCIM1366	Li et al., 2020
Pneumocystis jirovecii	HAAA01	RU7	Calderón et al., 2014
Rhizopus oryzae	GDUK01	97-1192	Li et al., 2006
Yarrowia sp. JCM 30694	IAAH01	JCM_30694	Trabelski et al.,2015
Yarrowia sp. JCM 30695	IAAI01	JCM 30695	Trabelski et al.,2015

De acordo com os metadados disponíveis na base de dados TSA, foram obtidas 19 sequências de montagem de transcriptomas. Destas, 16 foram montadas utilizando uma abordagem de novo, ou seja, sem referência a uma sequência conhecida. Para a realização destas montagens, foram utilizados quatro programas diferentes, são eles: *Newbler* (MARGULIES et al., 2005) *SOAP* (LI et al., 2008), *Trinity* (GRABHERR et al., 2011) e

RSEM (LI e DEWEY, 2011). Já as outras três sequências foram montadas utilizando a abordagem por referência, ou seja, baseando-se em uma sequência conhecida. Neste caso, foram utilizados dois programas diferentes: MIRA (CHEVREUX et al., 2004) e a suíte de programas *Tuxedo* (TRAPNELL et al., 2012) (Tabela 3).

Tabela 3. Metadados de sequenciamento e montagem dos transcriptomas. Código de identificação de TSA de cada espécie, seus respectivos métodos de sequenciamento e montagens dos transcriptomas.

TSA	SEQUENCIAMENTO	MONTAGEM
GBRE01	Illumina Hi-Seq	*De novo*
GBSP01	Illumina	*De novo*
GBSQ01	Illumina	*De novo*
GDUK01	Illumina HiSeq	*De novo*
GEEO01	Illumina HiSeq	*De novo*
GEMY01	Illumina Hi-Seq	*De novo*
GEVV02	Illumina Hi-Seq	*De novo*
GEVW02	Illumina HiSeq 2000	*De novo*
GFBC01	Illumina Hi-Seq	*De novo*
GFWE01	Illumina NextSeq 500	*De novo*
GGKD01	IonTorrent	Referência
GHMU01	Illumina Hi-Seq	*De novo*
GHTY01	Illumina MiSeq	Referência
GHWG01	Illumina Hi-Seq	*De novo*
GIPK01	Illumina HiSeq	*De novo*
GIXY01	Illumina NovaSeq	*De novo*
HAAA01	Illumina	Referência
IAAH01	Illumina HiSeq - 2500	*De novo*
IAAI01	Illumina HiSeq - 2500	*De novo*

A partir dos dados disponíveis no NCBI, foi possível a identificação de informações biomoleculares das 15 espécies obtidas. Na Tabela 4 estão descritas informações sobre a presença de um genoma de referência disponível, a presença de genomas mitocondriais, a quantidade de genes, ncRNAs e lncRNAs. A curadoria dos metadados das espécies presentes neste trabalho também foi enriquecida através das informações disponíveis no banco de dados ENSEMBL.

Tabela 4. Informações biomoleculares das espécies de fungos patogênicos.

ESPÉCIE	Genoma	Mitogenoma	Genes	ncRNA	lcnRNA
Alternaria sp. MG1	NCBI - ASM357452v1	Sim	13.440	NA	NA
Arthrinium phaeospermum	NCBI - ASM650353v1	Não	NA	NA	NA
Aureobasidium melanogenum - GEEO01	NCBI – GCA_000721775.1 ENSEMBL - GCA_000721775	Não	10.582	ENSEMBL - 305	ENSEMBL - 3
A. melanogenum - GIXY01	NCBI - GCA_000721775.1 ENSEMBL - GCA_000721775	Não	10.582	ENSEMBL - 305	ENSEMBL - 3
Candida africana	NA	Não	NA	NA	NA
Candida albicans	NCBI - ASM18296v3 ENSEMBL - Cand_albi_SC5314_V4	Sim	6.478	NCBI - 215	273
Geotrichum candidum	NCBI - LMA-244_clib	Sim	5.396	NCBI - 114	22
Irpex lacteus	NCBI – ASM198639v2	Não	NA	NA	NA
Mucor irregularis - GHMU01	NCBI - GCA_000587855.1	Não	NA	NA	NA
M. irregularis - GFBC01	NCBI - B50	Não	NA	NA	NA
Papiliotrema laurentii RY1	NCBI - ASM1292261v1	Não	NA	NA	NA
Paracoccidioides brasiliensis Pb18	NCBI - Paracocci_br_Pb18_V2	Sim	8.654	NCBI - 132 / ENSEMBL - 286	132
Penicillium janthinellum - GBSP01	NCBI - ASM236980v1	Não	NA	NA	NA
P. janthinellum - GBSQ01	NCBI - ASM236980v1	Não	NA	NA	NA
P. janthinellum - GIPK01	NCBI - ASM236980v1	Não	NA	NA	NA
Pneumocystis jirovecii	NCBI - GCA_001477535.1 ENSEMBL - Pneu_jiro_RU7_V2	Não	3.861	102	102
Rhizopus oryzae	NCBI - MucRam97-1192-1.0	Sim	NA	NA	NA
Yarrowia sp. JCM 30694	NCBI - JCM_30695_assembly_v001 / ENSEMBL - ASM176148v1	Sim	NA	NA	NA
Yarrowia sp. JCM 30695	NCBI - JCM_30695_assembly_v001 / ENSEMBL - ASM176148v1	Sim	NA	NA	NA

5.2. Análise de potencial de codificação

A predição e classificação dos transcritos que codificam ou não para proteína foram realizadas utilizando três ferramentas de bioinformáticas: CPAT, CPC2 e RNASamba (KANG et al., 2017; WANG et al., 2013; CAMARGO et al., 2020). Estas ferramentas são amplamente utilizadas na comunidade científica para identificar transcritos codificantes e não-codificantes. Os transcritos foram submetidos à análise de potencial codificante (CP) utilizando essas ferramentas para 14 espécies diferentes. Após a execução dos programas, os transcritos foram classificados em três categorias: codificantes, não-codificantes e indeterminados. A categoria codificantes refere-se aos transcritos que codificam proteína, a categoria não-codificantes corresponde aos transcritos que não codificam proteínas e a categoria indeterminado representa os transcritos cuja classificação não pôde ser determinada por meio da concordância de resultados entre os três programas e sendo calculada por meio da soma da quantidade de total de transcritos codificantes e não codificantes e o resultado desta operação foi subtraído pelo total de transcritos presentes nos transcriptomas brutos de cada espécie.

O programa CPAT foi utilizado para analisar a proporção de sequências codificantes e não-codificantes nos transcriptomas de diversas espécies de fungos:

- A espécie *Alternaria* MG1 apresentou 7.192 sequências anotadas como codificantes (38,84%) e 11.326 (61,16%) não-codificantes de um total de 18.518 transcritos em seu transcriptoma.
- Já a espécie *Athrinium phaeospermum* possui 12.861 transcritos, dos quais 8.470 (65,86%) são codificantes e 4931 (38,34%) são não-codificantes.
- A cepa GEEO01 de *Aureobasidium melanogenum*, de 20.507 sequências, apresentou 9.695 (47,28%) sequências codificantes e 10.812 (52,72%) sequências não-codificantes.
- A cepa GIXY01 de *A. melanogenum* apresentou 19.673 (38,37%) sequências codificantes e 19.131 (38,29%) sequências não-codificantes de um total de 38.809 sequências.

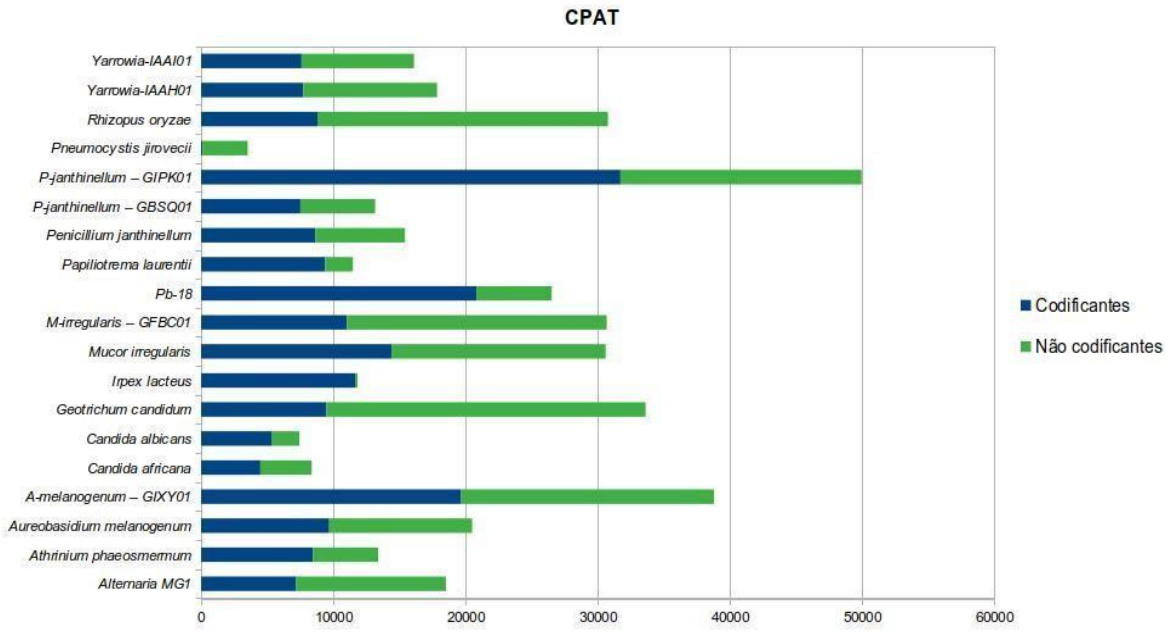
- A espécie *Candida africana* apresentou 4.481 (53,55%) de transcritos codificantes e 3.887 (46,45%) não-codificantes, enquanto a espécie *C. albicans* apresentou 5.319 (17,38%) transcritos codificantes e 2.118 (6,92%) não-codificantes.
- *Geotrichum candidum* apresentou 9.486 (31,9%) de sequências codificantes e 24.147 (71,80%) não-codificantes em um transcriptoma de 33.633 sequências.
- *Irpex lacteus*, o único macrofungo desta pesquisa, apresentou 11645 (72,31%) sequências codificantes e 195 (1,21%) de sequências não-codificantes em um transcriptoma de 11840 sequências.
- A cepa GHMU01 de *Mucor irregularis* apresentou 14.405 (47,06%) de transcritos codificantes e 19.652 (74,10%), não-codificantes, enquanto a cepa GFBC01 apresentou 11.039 (41,63%) são codificantes e 19.652 (74,10%) não-codificantes.
- Em *Paracoccidioides brasiliensis* – Pb18, com um transcriptoma de 11.482 sequências, 9.383 (81,72%) foram anotadas como codificantes e 2.099 (18,28%) como não-codificantes.
- Em *Papiliotrema laurentii*, das 26.520 sequências de seu transcriptoma, 20.827 (78,53%) foram classificadas como codificantes e 5.693 (21,47%) como não-codificantes.
- Em *Penicillium janthinellum*, a cepa GBSP01 apresentou 8.649 (56,06%) transcritos codificantes e 6.778 (43,94%) não-codificantes, a cepa GBSQ01 apresentou 7.510 (57,03%) transcritos codificantes e 5.658 (42,97%) não-codificantes, e a cepa GIPK01 apresentou 31.728 (63,49%) transcritos codificantes e 18.245 (36,51%) não-codificantes.
- *Pneumocystis jirovecii* apresentou 111 (3,15%) transcritos codificantes e 3.413 (96,85%) transcritos não-codificantes em um transcriptoma de 3.524 sequências.
- *Rhizopus oryzae* apresentou 8.803 (28,76%) transcritos codificantes e 21.798 (71,24%) transcritos não-codificantes em um transcriptoma de 30.601 sequências.
- Em *Yarrowia sp.* com a cepa IAAH01 apresentou 7.721 transcritos codificantes (43,25%) e 10.133 transcritos não-codificantes (56,75%).
- No isolado IAAI01 de *Yarrowia sp.* foram obtidos 7.599 transcritos codificantes (47,18%) e 8.506 transcritos não-codificantes (52.82%).

Esses dados são resumidos na Tabela 5 e Figura 4.

Tabela 5. Comparativo de CP com base no programa CPAT.

ESPÉCIE	Transcritos	Codificantes (%)	Não codificantes (%)
Alternaria sp. MG1	18.518	38.84	61.16
Arthrinium phaeospermum	12.861	65.86	38.34
Aureobasidium melanogenum	20.507	47.28	52.72
Aureobasidium melanogenum	38.809	39.37	38.29
Candida africana	7.442	53.55	46.45
Candida albicans	8.370	17.38	6.92
Geotrichum candidum	33.633	31.90	71.80
Irpex lacteus	11.840	72.31	1.21
Mucor irregularis	30.608	47.06	52.94
Mucor irregularis	30.691	41.63	74.10
Papiliotrema laurentii RY1	11.482	78.53	21.47
Paracoccidioides brasiliensis Pb18	26.520	81.72	18.28
Penicillium janthinellum	15.427	56.06	43.94
Penicillium janthinellum	13.184	57.03	42.97
Penicillium janthinellum	49.885	63.49	36.51
Pneumocystis jirovecii	3.524	3.15	96.85
Rhizopus oryzae	30.601	28.77	71.82
Yarrowia sp. JCM 30694	17.854	43.25	56.75
Yarrowia sp. JCM 30695	16.105	47.18	52.82

Figura 4. Comparativo de CP com base no programa CPAT.



O programa CPC2 também foi utilizado para analisar a proporção de sequências codificantes e não-codificantes:

- A espécie *Alternaria* MG1 apresentou 5.329 sequências codificantes, representando 28,78% do total, e 13.189 sequências não-codificantes, representando 71,22% do total.
- *Athrinium phaeospermum* apresentou 7.824 sequências codificantes (60,84%) e 5.037 sequências não-codificantes (39,16%).
- Em *Aureobasidium melanogenum* com GEEO01, 7.987 sequências foram anotadas como codificantes (38,95%) e 12.520 como não-codificantes (61,05%).
- Já *A. melanogenum* no isolado GIXY01 apresentou 17.922 sequências codificantes (35,86%) e 20.887 sequências não-codificantes (41,80%).
- *Candida africana* obteve 4.330 transcritos codificantes (51,74%) e 4.037 transcritos não-codificantes (48,24%). *C. albicans* possui 5253 transcritos codificantes (17,17%) e 2.184 transcritos não-codificantes (7,14%).
- *Geotrichum candidum* possui 4.012 sequências codificantes (22,47%) e 29.621 sequências não-codificantes (88,07%).
- *Irpex lacteus* obteve 11.391 sequências codificantes (70,73%) e 449 sequências não-codificantes (2,79%).
- *Mucor irregularis* em GHMU01 apresentou 15.346 transcritos codificantes (50,14%) e 15.262 transcritos não-codificantes (49,86%).
- Já a cepa GFBC01 apresentou 9.224 transcritos codificantes (34,78%) e 21.467 transcritos não-codificantes (80,95%).
- Em *Paracoccidioides brasiliensis* – Pb18, 18.736 sequências foram anotadas como codificantes (70,65%) e 7.757 sequências como não-codificantes (29,25%).
- *Papiliotrema laurentii* apresentou 9.535 sequências codificantes (83,04%) e 1.947 sequências não-codificantes (16,96%).
- *Penicillium janthinellum*, com a cepa GBSP01 apresentou 7.946 transcritos codificantes (51,51%) e 7.481 transcritos não-codificantes (48,49%).
- A cepa GBSQ01 de *P. janthinellum* constou 7.163 transcritos codificantes (54,40%) e 6.005 transcritos não-codificantes (45,60%) e a cepa GIPK01, 29.636 transcritos codificantes (59,30%) e 20.610 transcritos não-codificantes (41,24%).

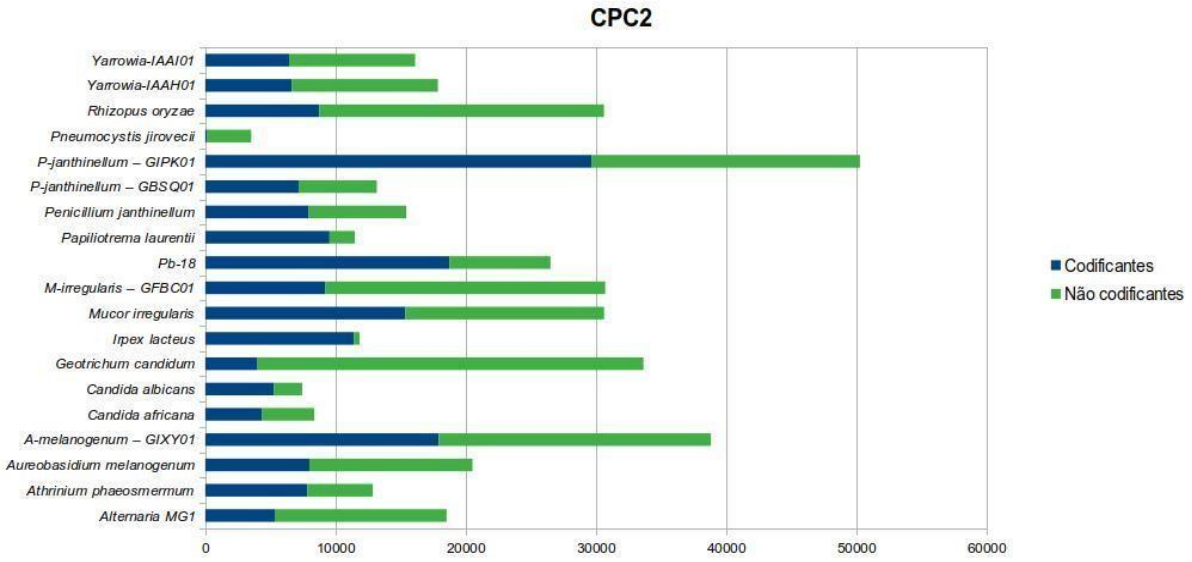
- *Pneumocystis jirovecii* apresentou 160 transcritos codificantes (4,54%) e 3.364 transcritos não-codificantes (95,46%).
- *Rhizopus oryzae* apresentou 8.758 transcritos codificantes (28,62%) e 21.843 transcritos não-codificantes (71,38%).
- Em *Yarrowia sp.* com a cepa IAAH01 apresentou 6.658 transcritos codificantes (37,29%) e 11.196 transcritos não-codificantes (62,71%).
- Já a cepa IAAI01 de *Yarrowia sp.* apresentou 6.433 transcritos codificantes (39,94%) e 9.672 transcritos não-codificantes (60,06%).

Essas informações estão resumidas na Tabela 6 e Figura 5.

Tabela 6. Comparativo de potencial codificador com base no programa CPC2.

ESPÉCIE	Transcritos	Codificantes (%)	Não codificantes (%)
Alternaria sp. MG1	18.518	28.78	71.22
Arthrinium phaeospermum	12.861	60.84	39.16
Aureobasidium melanogenum - GEEO01	20.507	38.95	61.05
Aureobasidium melanogenum - GIPK01	38.809	35.86	41.80
Candida africana	7.442	51.74	48.24
Candida albicans	8.370	17.17	7.14
Geotrichum candidum	33.633	22.47	88.07
Irpex lacteus	11.840	70.73	2.79
Mucor irregularis	30.608	50.14	49.86
Mucor irregularis	30.691	34.78	80.95
Papiliotrema laurentii RY1	11.482	70.65	29.25
Paracoccidioides brasiliensis Pb18	26.520	83.04	16.96
Penicillium janthinellum- GBSP01	15.427	51.51	48.49
Penicillium janthinellum - GBSQ01	13.184	54.40	45.60
Penicillium janthinellum - GIXY01	49.885	59.30	41.24
Pneumocystis jirovecii	3.524	4.54	95.46
Rhizopus oryzae	30.601	28.62	71.38
Yarrowia sp. JCM 30694	17.854	37.29	62.71
Yarrowia sp. JCM 30695	16.105	39.94	60.06

Figura 5. Comparativo de CP com base no programa CPC2.



Por fim, o programa RNAsamba também foi utilizado para analisar as sequências de RNA das espécies de fungos:

- A espécie *Alternaria* MG1 apresentou 14.113 (76.21%) sequências codificantes e 4.405 (23.79%) não-codificantes.
- A espécie *Athrinium phaeospermum* possuiu 10.925 (84.95%) sequências codificantes e 1.936 (15.05%) não-codificantes.
- A espécie *Aureobasidium melanogenum*, na cepa GEEO01, apresentou 16.440 (80.17%) sequências codificantes e 4.067 (19.83%) não-codificantes.
- Já na cepa GIXY01 de A. *melanogeum*, foram identificadas 32.268 (64.57%) sequências codificantes e 6.541 (13.09%) não-codificantes.
- A espécie *Candida africana* apresentou 6.275 (74.99%) de transcritos codificantes e 2.093 (25.1%) não-codificantes.
- *C. albicans* possuiu 6.334 (20.70%) transcritos codificantes e 1.103 (3.60%) não-codificantes.
- *Geotrichum candidum* possuiu 24.035 (71.46%) de sequências codificantes e 9.598 (53.76%) não-codificantes.
- A espécie *Irpex lacteus* obteve 11.826 (73.43%) sequências codificantes e 14 (0.09%) de sequências não-codificantes, sendo esta, a espécie com a menor quantidade de transcritos não codificantes entre as três ferramentas utilizadas.

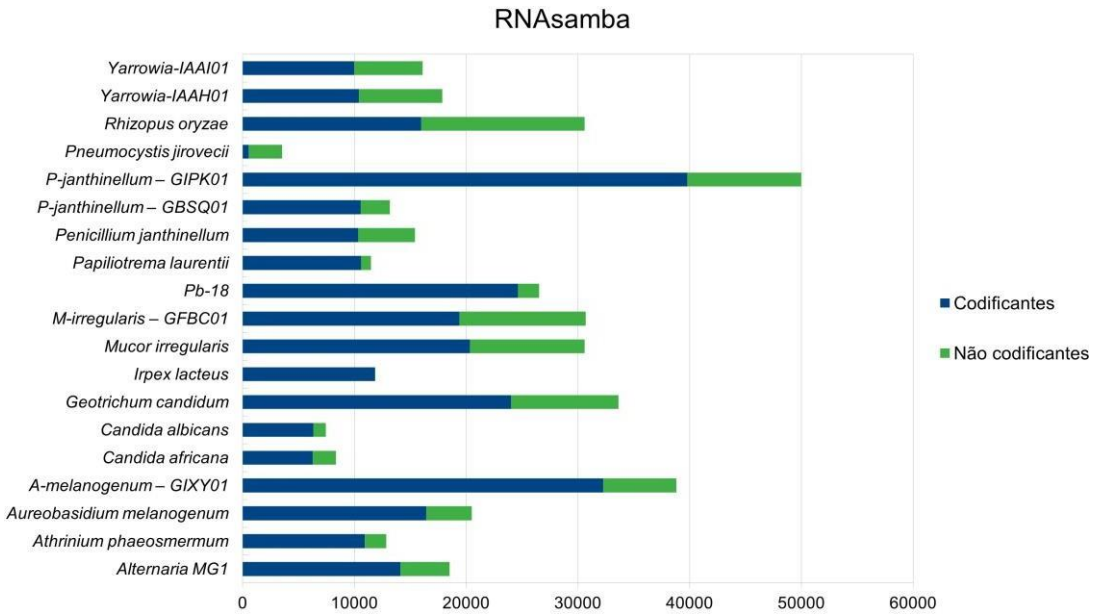
- A espécie *Mucor irregularis*, na cepa GHMU01, apresentou 20.352 (66.49%) de transcritos codificantes e 10.256 (33.51%) não-codificantes.
- Já na cepa GFBC01 de *M. irregularis*, 19.394 (73.13%) foram identificadas como codificantes e 11.297 (42.60%) não-codificantes.
- Em *Paracoccidioides brasiliensis* – Pb18, foram identificadas 24.637 (92.90%) sequências codificantes e 1.883 (7.1%) não-codificantes.
- A espécie *Papiliotrema laurentii* apresentou 10.606 (92.37%) sequências codificantes e 876 (7.6%) não-codificantes.
- A espécie *Penicillium janthinellum*, na cepa GBSP01, apresentou 10.330 (66.96%) sequências codificantes e 5.079 (32.92%) não-codificantes.
- Na cepa GBSQ01 de *P. janthinellum*, foram identificadas 2.605 (80.22%) sequências codificantes e 2.605 (19.78%) não-codificantes.
- E na cepa GIPK01 de *P. janthinellum*, foram identificadas 39.778 (79.60%) sequências codificantes e 10.195 (20.40%) não-codificantes.
- A espécie *Pneumocystis jirovecii* apresentou 544 (15.44%) sequências codificantes e 2.980 (84.56%) não-codificantes.
- A espécie *Rhizopus oryzae* apresentou 15.965 (52.17%) sequências codificantes e 14.636 (47.83%) não-codificantes.
- A espécie *Yarrowia sp.*, na cepa IAAH01, apresentou 10.419 (58.36%) sequências codificantes e 7.435 (41.64%) não-codificantes.
- Enquanto o isolado IAAI01 de *Yarrowia sp.*, obteve 9.994 (60.06%) sequências codificantes e 6.111 (37.94%) não-codificantes.

Essas informações estão resumidas na Tabela 7 e Figura 6.

Tabela 7. Comparativo de CP com base no programa RNASamba.

ESPÉCIE	Transcritos	Codificantes (%)	Não codificantes (%)
ESPÉCIE	**Transcritos**	28.78	71.22
Alternaria sp. MG1	18.518	60.84	39.16
Arthrinium phaeospermum	12.861	38.95	61.05
Aureobasidium melanogenum - GEEO01	20.507	35.86	41.80
Aureobasidium melanogenum - GIPK01	38.809	51.74	48.24
Candida africana	7.442	17.17	7.14
Candida albicans	8.370	22.47	88.07
Geotrichum candidum	33.633	70.73	2.79
Irpex lacteus	11.840	50.14	49.86
Mucor irregularis	30.608	34.78	80.95
Mucor irregularis	30.691	70.65	29.25
Papiliotrema laurentii RY1	11.482	83.04	16.96
Paracoccidioides brasiliensis Pb18	26.520	51.51	48.49
Penicillium janthinellum- GBSP01	15.427	54.40	45.60
Penicillium janthinellum - GBSQ01	13.184	59.30	41.24
Penicillium janthinellum - GIXY01	49.885	4.54	95.46
Pneumocystis jirovecii	3.524	28.62	71.38
Rhizopus oryzae	30.601	37.29	62.71
Yarrowia sp. JCM 30694	17.854	39.94	60.06

Figura 6. Comparativo de CP com base no programa RNASamba**.**



Por fim, avaliamos a sobreposição na classificação de sequências identificadas como codificadoras e não codificadoras de proteínas pelas três ferramentas de CP. Identificamos nesse processo 183.494 sequências genéticas que foram classificadas como indeterminadas, devido à falta de concordância entre as anotações dos três programas utilizados. Nesse sentido, a classificação demostrou que a espécie *Alternaria* MG1 apresentou o maior número de sequências indeterminadas, com 9.533 (51,48%) sequências. Já *Athrinium phaeospermum* apresentou 4.173 (32,45%), *Aureobasidium melanogenum* 14.971 (73%), e a cepa GIXY01 de 16.781 (66,84%). *Candida africana* obteve 4.974 (40,25%) e *C. albicans* 3.369 (72,58%). *Geotrichum candidum* apresentou 24.410 (0,47%), *Irpex lacteus* 56 (66,19%), *Mucor irregularis* com a cepa GHMU01 20.258 (37,86%), cepa GFBC01 11.620 (23,6%), *Paracoccidioides brasiliensis* - Pb18 3.533 (30,77%), *Papiliotrema laurentii* 6.260 (23,6%), *Penicillium janthinellum* com a cepa GBSP01 9.425 (61,09%), cepa GBSQ01 7.154 (54,26%), cepa GIPK01 30.374 (60,89%), *Pneumocystis jirovecii* 488 (13,85%), *Rhizopus oryzae* 8.427 (27,54%), *Yarrowia sp.* IAAH01 3.937 (22,05%), e *Yarrowia sp.* IAAI01 3.751 (23,29%). Essas informações são resumidas na Tabela 8 e nas Figuras 7 e 8.

Tabela 8. Distribuição dos transcritos com base em seu potencial codificante. Distribuição com base na concordância na classificação das três ferramentas de potencial de codificação.

ESPÉCIE	Codificantes (%)	Não codificantes (%)	Indeterminado (%)
Alternaria sp. MG1	4.802 (25.93%)	4.183 (22.59%)	9.533 (51.48%)
Arthrinium phaeospermum	7.166 (55.71%)	1.522 (11,83%)	5.827 (45,31%)
Aureobasidium melanogenum - GEEO01	3.536 (17,24%)	2.000 (9,75%)	14.971 (73%)
Aureobasidium melanogenum - GIXY01	16.083 (41,44%)	5.945 (15,32%)	16.781 (43,24%)
Candida africana	2.388 (32,09%)	1.006 (13,52%)	4.974 (63.84%)
Candida albicans	3.760 (44,92%)	314 (3,75%)	3.368 (40,24%)
Geotrichum candidum	1.960 (5,82%)	7.263 (21,59%)	24.410 (72,58%)
Irpex lacteus	0	0	0
Mucor irregularis -GFBC01	7.913 (25,85%)	11.158 (36,45%)	4.967(16,22%)
Mucor irregularis	3.698 (12,5%)	2.946 (9,6%)	24.047 (78,35%)
Papiliotrema laurentii RY1	5.945 (51,77%)	1.769 (15,40%)	1.886 (16,42%)
Paracoccidioides brasiliensis Pb18	7.775 (29,31%)	174 (0,66%)	3.533 (13,32%)
Penicillium janthinellum- GBSP01	4.115 (26,67%)	1.887 (12,23%)	9.425 (61,09%)
Penicillium janthinellum - GBSQ01	3.911 (29,66%)	1.267 (9,61%)	7.990 (60,60%)
Penicillium janthinellum - GIPK01	16.446 (32,96%)	2.631 (5,27%)	30.896 (61,93%)
Pneumocystis jirovecii	75 (2,13%)	2.961 (84,2%)	488 (13,85%)
Rhizopus oryzae	7.594 (24,82%)	15.480 (50,59%)	7.527 (24,60%)
Yarrowia sp. JCM 30694	7.646 (42,82%)	7.348 (41,16%)	2.860 (16,02%)
Yarrowia sp. JCM 30695	7.503 (46,59%)	6.004 (37,28%)	2.598 (16,13%)

Figura 7. Distribuição dos transcritos como não codificantes. Distribuição baseada na concordância na classificação das três ferramentas de potencial de codificação para as espécies: (a) *Alternaria* MG1; (b) *Aureobasidium melanogenum* – GEEO01, (c) *A. melanogenum* – GIXY01; (d) *Artrhinium phaeospermum*; (e) *Candida africana*; (f) *Candida albicans*; (g) *Geotrichum candidum*; (h) *Irpex lacteus*; (i) *Mucor irregularis* – GFBCO1; (j) *M. irregularis* - GHMU01; (k) *Paracoccidioides brasiliensis*; (l) *Penicillium janthinellum* GBSP01; (m) *P. janthinellum* - GBSQ01; (n) *P. janthinellum* - GIPK01; (o) *Pneumocystis jirovecii*; (p) *Papiliotrema laurentii*; (q) *Rhizopus oryzae*; (r) *Yarrowia* – IAAH01; (s) *Yarrowia* - IAAI01

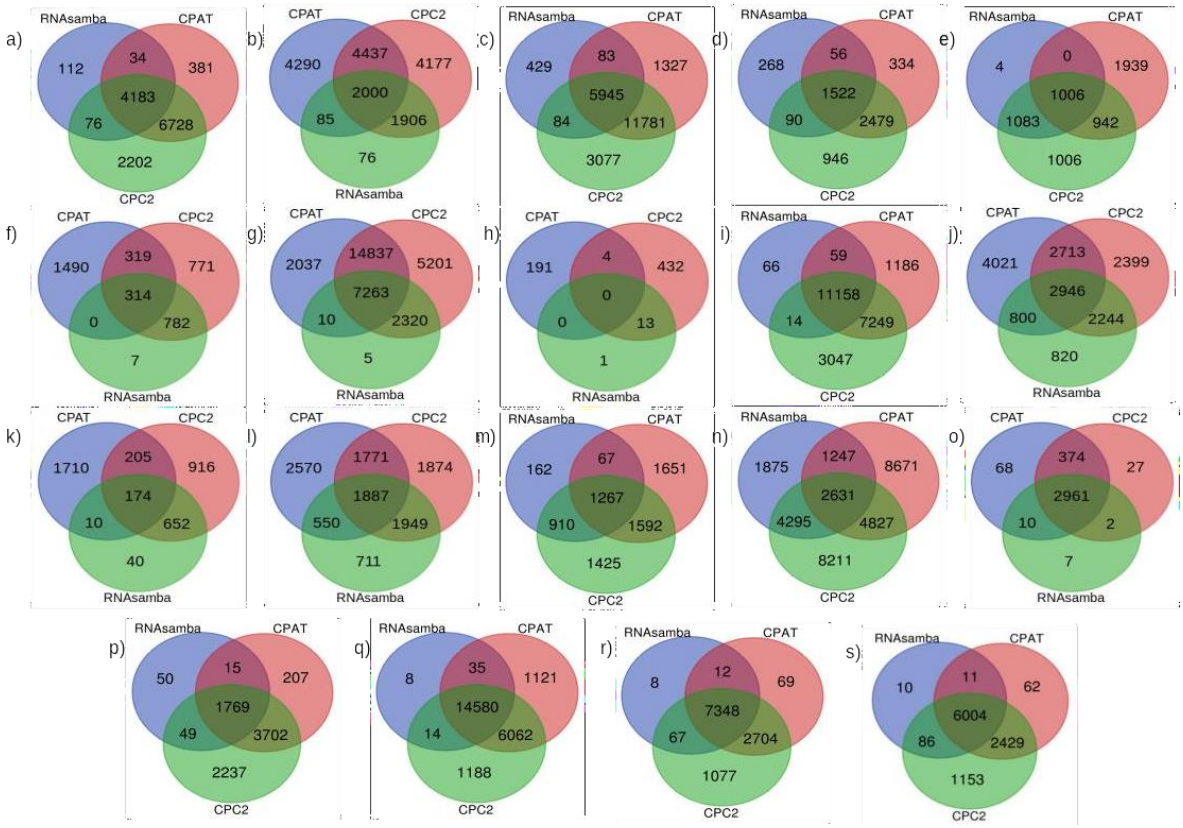
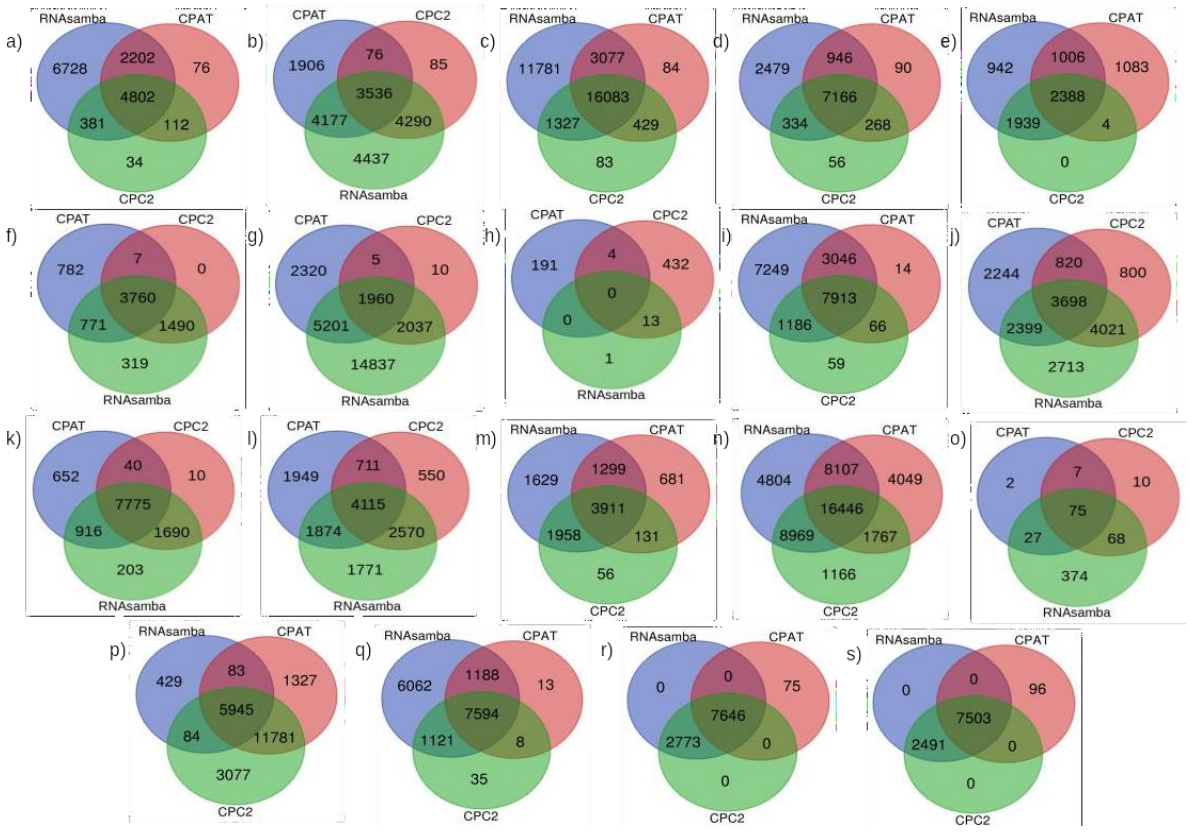


Figura 8. Distribuição dos transcritos como codificantes. Distribuição baseada na concordância na classificação das três ferramentas de potencial de codificação para as espécies: (a) *Alternaria* MG1; (b) *Aureobasidium melanogenum* – GEEO01, (c) *A. melanogenum* – GIXY01; (d) *Artrhinium phaeospermum*; (e) *Candida africana*; (f) *Candida albicans*; (g) *Geotrichum candidum*; (h) *Irpex lacteus*; (i) *Mucor irregularis* – GFBCO1; (j) *M. irregularis* - GHMU01; (k) *Paracoccidioides brasiliensis*; (l) *Penicillium janthinellum* GBSP01; (m) *P. janthinellum* - GBSQ01; (n) *P. janthinellum* - GIPK01; (o) *Pneumocystis jirovecii*; (p) *Papiliotrema laurentii*; (q) *Rhizopus oryzae*; (r) *Yarrowia* – IAAH01; (s) *Yarrowia* - IAAI01



5.3. Anotação funcional dos transcritos não-codificantes de proteína

O módulo cmscan da ferramenta INFERNAL foi utilizado para a anotação de ncRNAs putativos com base em todos os transcritos presentes nos 19 transcriptomas das 15 espécies de fungos patogênicos analisados durante esse trabalho. Os dados de anotação são descritos abaixo:

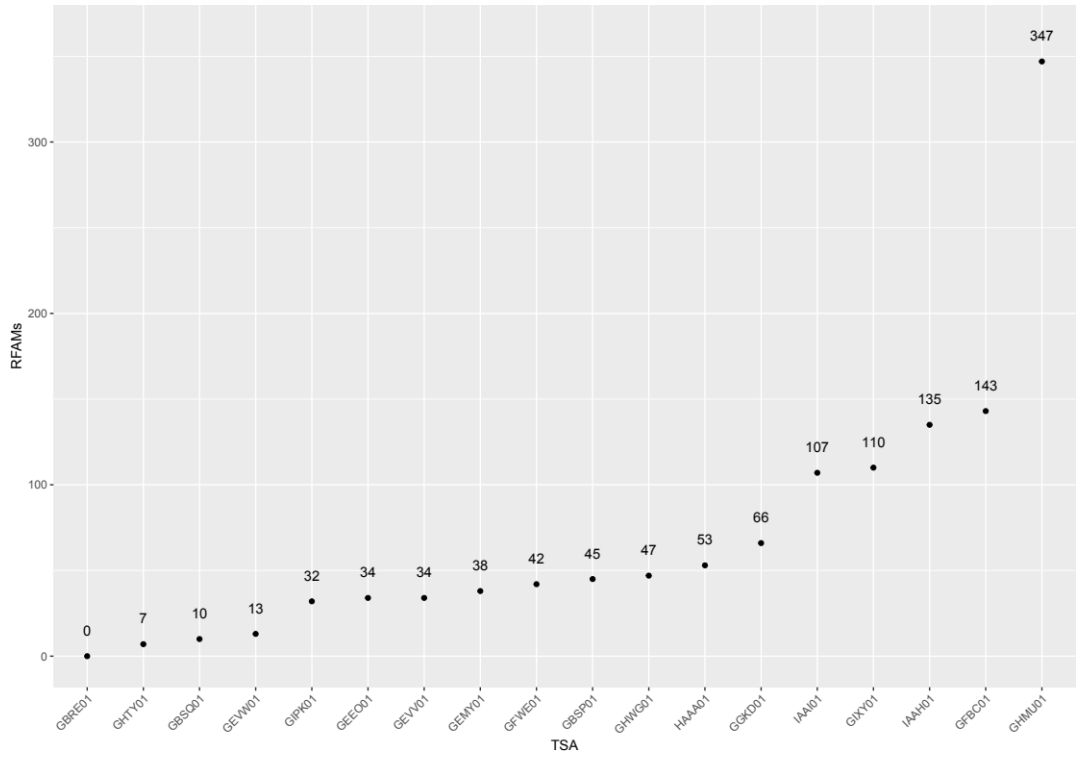
- A espécie A*lternaria* MG1 apresentou 24 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: SSU_rRNA_eukarya, SSU_rRNA_bacteria, SSU_rRNA_archaea, PK-G12rRNA, tRNA, tmRNA
- A espécie *A. phaeospermum* apresentou 37 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: flavi_FSE, sR52, sR53, tRNA, SSU_rRNA_bacteria, SSU_rRNA_archaea, SSU_rRNA_eukarya
- A cepa *A. melanogenum* - GIXY01 apresentou 29 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: Metazoa_SRP, tRNA, Afu_294, snoZ101, GcvB, SNORD14.
- A cepa *A. melanogenum* GIXY01 apresentou 80 sequências anotadas pelo INFERNAL atr'avés dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: sR52 ceN28 sR53, SNORD52, TP53TG1_1, tRNA, SNORD33, SNORD50, SSU_rRNA_eukarya, SSU_rRNA_archaea, SSU_rRNA_bacteria, CRISPR-DR2
- A espécie *C. africana* apresentou 28 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: snosnR66, tRNA, tRNA-Sec, SNORD79, snR79, STnc380, rliD, MIR473
- A espécie *C. albicans* apresentou 12 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: MIR821, MIR530, Afu_191 snosnR69, HCV_SLIV

- A espécie *G. candidum* apresentou 59 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: SSU_rRNA_eukarya, SSU_rRNA_archaea, SSU_rRNA_bacteria
- A espécie *I. lacteus* não apresentou sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013.
- A espécie *M. irregularis* apresentou 151 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: tRNA, MIR530, mir-268, MIR396 MIR2118, MIR1023, MIR408
- A espécie *M. irregularis* GFBC01 apresentou 129 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: U2, Metazoa_SRP, Fungi_SRP
- A espécie *P. laurentii* apresentou 23 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: mraW, SSU_rRNA_archaea, tRNA mir-253, SSU_rRNA_eukarya, SSU_rRNA_bacteria
- A espécie *P. brasiliensis* apresentou 5 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: SSU_rRNA_eukarya, snoU18, SSU_rRNA_archaea, 5_8S_rRNA, Trp_leader
- A espécie *P. janthinellum* apresentou 26 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: snR51, Afu_513, SNORD27, snoR118
- A cepa *P. janthinellum* GBSQ01 apresentou 8 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: Trp_leader, SNORD29 SSU_rRNA_bacteria, SSU_rRNA_archaea, SSU_rRNA_eukarya, mir-1224 tmRNA, snoZ13_snr52, tRNA, snoR38
- A cepa *P. janthinellum* GIPK01 apresentou 28 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as

sequências anotadas, destacam-se as famílias de RNAs: tRNA, SSU_rRNA_bacteria, SSU_rRNA_archaea, Afu_294, snR51

- A espécie *P. jirovecii* apresentou 49 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: snR75, mir-460, U6, sR17, tRNA, sR38, Afu_298, sR2
- A espécie *R. oryzae* apresentou 161 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: tRNA, SSU_rRNA_bacteria, SSU_rRNA_archaea, Afu_294, snR51
- A cepa *Yarrowia* – 30694 apresentou 110 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: tRNA, SnoMe18S-Um1356, 5S_rRNA
- A cepa *Yarrowia* – 30695 apresentou 86 sequências anotadas pelo INFERNAL através dos modelos de covariância (CMs) do RFAM 11.0.2 2.09.2013. Dentre as sequências anotadas, destacam-se as famílias de RNAs: SSU_rRNA_eukarya,, SSU_rRNA_archaea, SSU_rRNA_bacteria, tRNA, snosnR55

Figura 9. Distribuição de transcritos anotados no Rfam.



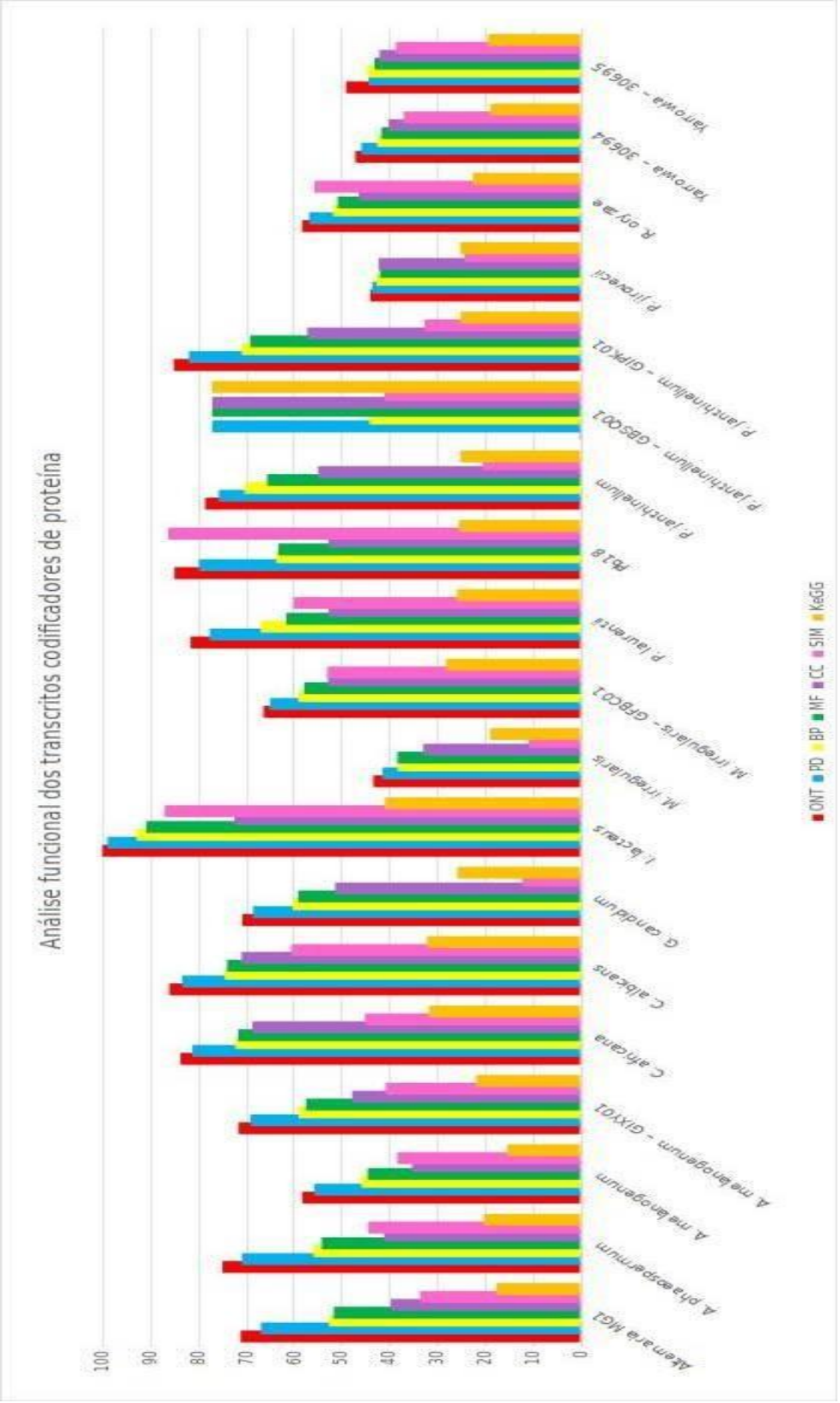
5.4. Anotação funcional dos transcritos codificantes de proteína

O programa EnTAP foi utilizado para a anotação funcional de todos os transcritos presentes nos 19 transcriptomas das 15 espécies de fungos patogênicos analisados durante esse trabalho. Os dados de anotação são apresentados na Figura 10 e na Tabela 9.

Tabela 9. Distribuição das anotações funcionais do EnTAP. A coluna **Espécie** compreende os fungos analisados, **SIM** corresponde à análise de similaridade, **ORT** predição de genes ortólogos, **KeGG** quantidade de vias metabólicas, **BP** processos biológicos, **CC** componentes celulares, **MF** funções moleculares, **PD** domínios proteicos, **COD** quantidade de transcritos codificantes, **NC** quantidade de transcritos não codificantes, **IND** quantidaade de transcritos indeterminados, **Infernal** quantidade de transcritos com anotações no módulo *cmscan* e **#seq** quantidade total de sequências em cada transcriptoma.

	SIM	**ORT**	**KeGG**	**BP**	**CC**	**MF**	**PD**	**COD**	**NC**	**IND**	**Infernal**	**#seq**
Alternaria MG1	6.197	13.149	3.247	9.726	7.334	9.545	12.369	4.802	4.183	9.533	24	18.518
A. phaeospermum	5.706	9.627	2.588	7.172	5.280	6.958	9.106	7.166	1.522	4.173	37	12.861
A. melanogenum GEEO01	7.843	11.919	3.133	9.378	7.215	9.116	11.408	3.536	2.000	14.971	29	20.507
A. melanogenum GIXY01	15.800	27.727	8.453	22.864	18.524	22.260	26.757	16.083	5.945	16.781	80	38.809
C. africana	3.768	7.000	2.646	6.035	5.741	5.987	6.789	2.388	1.006	4.974	28	8.368
C. albicans	4.497	6.397	2.385	5.532	5.276	5.498	6.197	3.759	314	3.369	12	7.442
G. candidum	4.084	23.775	8.633	20.254	17.238	19.845	23.019	1.960	7.263	24.410	59	33.633
I. lacteus	10.291	11.834	4.827	10.989	8.568	10.745	11.716	11.784	0	56	0	11.840
M. irregulares GHMU01	3.289	13.239	5.758	11.679	10.049	11.698	12.658	6.116	4.234	20.258	151	30.608
M. irregularis GFBC01	16.231	20.332	8.611	18.096	16.208	17.718	19.919	7.913	11.158	11.620	129	30.691
P. laurentii	15.910	21.620	6.821	17.725	13.948	16.301	20.559	18.491	1.769	6.260	23	26.520
P. brasiliensis	9.897	9.745	2.907	7.303	6.052	7.252	9.158	7.775	174	3.533	5	11.482
P. janthinellum GBSP01	3.145	12.104	3.864	10.828	8.461	10.100	11.664	4.115	1.887	9.425	26	15.427
P. janthinellum GBSQ01	5.397	15	10.147	5.805	10.140	10.140	10.140	4.652	1.362	7.154	8	13.168
P. janthinellum GIPK01	16.269	42.485	12.514	35.378	28.531	34.501	40.906	16.933	2.666	30.374	28	49.973
P. jirovecii	853	1.547	885	1.503	1.485	1.476	1.533	75	2.961	488	49	3.524
R. oryzae	17.003	17.795	6.880	15.817	14.174	15.517	17.352	7.594	14.580	8.427	161	30.601
Yarrowia – 30694	6.598	8.393	3.359	7.578	7.157	7.424	8.172	6.569	7.348	3.937	110	17.854
Yarrowia – 30695	6.212	7.881	3.100	7.174	6.763	6.925	7.125	6.350	6.004	3.751	86	16.105

Figura 10. Proporções dos transcritos anotados funcionalmente.



De um total de 397.931 transcritos pertencentes aos 19 transcriptomas das 15 espécies de fungos patogênicos, observamos a seguinte anotação funcional:

- A espécie *Alternaria* MG1 apresentou 6.197 sequências anotadas através da análise de similaridade e 13.149 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 3.246 genes em vias metabólicas, sendo 9.726 na classe de Processos Biológicos (BP), 7.334 na classe Componente Celular (CC) e 9.545 na classe Função Molecular (MF). Também foram observados 12.369 domínios proteicos. Esses dados se somam as 4.802 sequências codificadoras de proteína, 4.183 sequências não-codificadoras de proteínas, 9.533 sequências indeterminadas e 24 sequências anotadas como ncRNAs.
- A espécie *A. phaeospermum* apresentou 5.706 sequências anotadas através da análise de similaridade e 9.627 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 2.588 genes em vias metabólicas, sendo 7172 na classe de Processos Biológicos (BP), 5.280 na classe Componente Celular (CC) e 6.958 na classe Função Molecular (MF). Também foram observados 9.106 domínios proteicos. Esses dados se somam as 7.166 sequências codificadoras de proteína, 1.522 sequências não-codificadoras de proteínas, 4.173 sequências e 37 sequências anotadas como ncRNAs.
- A espécie *A. melanogenum* GEEO01 apresentou 7.843 sequências anotadas através da análise de similaridade e 11.919 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 3.133 genes em vias metabólicas, sendo 9.378 na classe de Processos Biológicos (BP), 9.378 na classe Componente Celular (CC) e 7215 na classe Função Molecular (MF). Também foram observados 11.408 domínios proteicos. Esses dados se somam as 3.536 sequências codificadoras de proteína, 2.000 sequências não-codificadoras de proteínas, 14.971 sequências indeterminadas e 29 sequências anotadas como ncRNAs.
- A espécie *A. melanogenum* GIXY01 apresentou 15.800 sequências anotadas através da análise de similaridade e 27.727 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 8.453 genes em vias metabólicas, sendo 22.864 na classe de Processos Biológicos (BP), 18.524 na classe Componente Celular (CC) e 22.260 na classe Função Molecular (MF). Também foram observados 26.757

domínios proteicos. Esses dados se somam as 16.083 sequências codificadoras de proteína, 5.945 sequências não-codificadoras de proteínas, 16.781 sequências indeterminadas e 80 sequências anotadas como ncRNAs.

- A espécie *C. africana* apresentou 3.768 sequências anotadas através da análise de similaridade e 7.000 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 2.646 genes em vias metabólicas, sendo 6.035 na classe de Processos Biológicos (BP), 6.035 na classe Componente Celular (CC) e 5.741 na classe Função Molecular (MF). Também foram observados 6.789 domínios proteicos. Esses dados se somam as 2.388 sequências codificadoras de proteína, 1.006 sequências não-codificadoras de proteínas, 4.974 sequências indeterminadas e 28 sequências anotadas como ncRNAs.
- A espécie *C. albicans* apresentou 4.497 sequências anotadas através da análise de similaridade e 6.397 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 2.385 genes em vias metabólicas, sendo 5.532 na classe de Processos Biológicos (BP), 5.276 na classe Componente Celular (CC) e 5.498 na classe Função Molecular (MF). Também foram observados 6.197 domínios proteicos. Esses dados se somam as 3.759 sequências codificadoras de proteína, 314 sequências não-codificadoras de proteínas, 3.369 sequências indeterminadas e 12 sequências anotadas como ncRNAs.
- A espécie *G. candidum* apresentou 4084 sequências anotadas através da análise de similaridade e 23.775 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 8.633 genes em vias metabólicas, sendo 20.254 na classe de Processos Biológicos (BP), 17.238 na classe Componente Celular (CC) e 19.845 na classe Função Molecular (MF). Também foram observados 23.019 domínios proteicos. Esses dados se somam as 1960 sequências codificadoras de proteína, 7.263 sequências não-codificadoras de proteínas, 24.410 sequências indeterminadas e 59 sequências anotadas como ncRNAs.
- A espécie *I. lacteus* apresentou 10.291 sequências anotadas através da análise de similaridade e 11834 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 4.827 genes em vias metabólicas, sendo 10.989 na classe de Processos Biológicos (BP), 8.568 na classe Componente Celular (CC) e 10.745 na classe Função Molecular (MF). Também foram observados 11.716

domínios proteicos. Esses dados se somam as 11.784 sequências codificadoras de proteína, 0 sequências não-codificadoras de proteínas, 56 sequências indeterminadas e 0 sequências anotadas como ncRNAs.

- A cepa *M. irregularis* GHMU01 apresentou 3.289 sequências anotadas através da análise de similaridade e 13.239 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 5.758 genes em vias metabólicas, sendo 11.679 na classe de Processos Biológicos (BP), 10.049 na classe Componente Celular (CC) e 11.698 na classe Função Molecular (MF). Também foram observados 12.658 domínios proteicos. Esses dados se somam as 6.116 sequências codificadoras de proteína, 4.234 sequências não-codificadoras de proteínas, 20.258 sequências indeterminadas e 151 sequências anotadas como ncRNAs.
- A cepa *M. irregularis* GFBC01 apresentou 16.231 sequências anotadas através da análise de similaridade e 20.332 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 8.611 genes em vias metabólicas, sendo 18.096 na classe de Processos Biológicos (BP), 16.208 na classe Componente Celular (CC) e 17.718 na classe Função Molecular (MF). Também foram observados 19.919 domínios proteicos. Esses dados se somam as 7.913 sequências codificadoras de proteína, 11158 sequências não-codificadoras de proteínas, 11.620 sequências indeterminadas e 129 sequências anotadas como ncRNAs.
- A espécie *P. laurentii* apresentou 21.620 sequências anotadas através da análise de similaridade e 15910 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 6.821 genes em vias metabólicas, sendo 17.725 na classe de Processos Biológicos (BP), 13.948 na classe Componente Celular (CC) e 16.301 na classe Função Molecular (MF). Também foram observados 20559 domínios proteicos. Esses dados se somam as 18.491 sequências codificadoras de proteína, 1.769 sequências não-codificadoras de proteínas, 6.260 sequências indeterminadas e 23 sequências anotadas como ncRNAs.
- A espécie *P. brasiliensis* apresentou 9.897 sequências anotadas através da análise de similaridade e 9.745 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 2.907 genes em vias metabólicas, sendo 7.303 na classe de Processos Biológicos (BP), 6.052 na classe Componente Celular (CC) e 7.252 na classe Função Molecular (MF). Também foram observados 9158 domínios proteicos.

Esses dados se somam as 7.775 sequências codificadoras de proteína, 174 sequências não-codificadoras de proteínas, 3533 sequências indeterminadas e 5 sequências anotadas como ncRNAs.

- A espécie *P. janthinellum* apresentou 3.145 sequências anotadas através da análise de similaridade e 12.104 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 3.864 genes em vias metabólicas, sendo 10.828 na classe de Processos Biológicos (BP), 8.461 na classe Componente Celular (CC) e 10.100 na classe Função Molecular (MF). Também foram observados 11.664 domínios proteicos. Esses dados se somam as 4.115 sequências codificadoras de proteína, 1.887 sequências não-codificadoras de proteínas, 9.425 sequências indeterminadas e 26 sequências anotadas como ncRNAs.
- A cepa *P. janthinellum* GBSQ01 apresentou 5.397 sequências anotadas através da análise de similaridade e 15 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 10.147 genes em vias metabólicas, sendo 5.805 na classe de Processos Biológicos (BP), 10.140 na classe Componente Celular (CC) e 10.140 na classe Função Molecular (MF). Também foram observados 10.140 domínios proteicos. Esses dados se somam as 4.652 sequências codificadoras de proteína, 1.362 sequências não-codificadoras de proteínas, 7.154 sequências indeterminadas e 8 sequências anotadas como ncRNAs.
- A cepa *P. janthinellum* GIPK01 apresentou 16.269 sequências anotadas através da análise de similaridade e 42.485 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 12.514 genes em vias metabólicas, sendo 35.378 na classe de Processos Biológicos (BP), 28.531 na classe Componente Celular (CC) e 34.501 na classe Função Molecular (MF). Também foram observados 40.906 domínios proteicos. Esses dados se somam as 16.933 sequências codificadoras de proteína, 2.666 sequências não-codificadoras de proteínas, 30.374 sequências indeterminadas e 28 sequências anotadas como ncRNAs.
- A espécie *P. jirovecii* apresentou 853 sequências anotadas através da análise de similaridade e 1.547 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 885 genes em vias metabólicas, sendo 1.503 na classe de Processos Biológicos (BP), 1.485 na classe Componente Celular (CC) e 1.485 na classe Função Molecular (MF). Também foram observados 1.533 domínios proteicos.

Esses dados se somam as 75 sequências codificadoras de proteína, 2.961 sequências não-codificadoras de proteínas, 3.751 sequências indeterminadas e 49 sequências anotadas como ncRNAs.

- A espécie *R. oryzae* apresentou 17.003 sequências anotadas através da análise de similaridade e 17.795 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 6.880 genes em vias metabólicas, sendo 15.817 na classe de Processos Biológicos (BP), 14174 na classe Componente Celular (CC) e 14.174 na classe Função Molecular (MF). Também foram observados 17.352 domínios proteicos. Esses dados se somam as 7.594 sequências codificadoras de proteína, 14.580 sequências não-codificadoras de proteínas, 8.427 sequências indeterminadas e 161 sequências anotadas como ncRNAs.
- A cepa *Yarrowia* – 30694 apresentou 6.598 sequências anotadas através da análise de similaridade e 8.393 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 3.359 genes em vias metabólicas, sendo 7.578 na classe de Processos Biológicos (BP), 7157 na classe Componente Celular (CC) e 7.424 na classe Função Molecular (MF). Também foram observados 8172 domínios proteicos. Esses dados se somam as 6.569 sequências codificadoras de proteína, 7348 sequências não-codificadoras de proteínas, 3.937 sequências indeterminadas e 110 sequências anotadas como ncRNAs.
- A cepa *Yarrowia* – 30695 apresentou 6.212 sequências anotadas através da análise de similaridade e 7.881 sequências anotadas através da predição de genes ortólogos. Além disso foram observados 3.100 genes em vias metabólicas, sendo 7.174 na classe de Processos Biológicos (BP), 6.763 na classe Componente Celular (CC) e 6.925 na classe Função Molecular (MF). Também foram observados 7125 domínios proteicos. Esses dados se somam as 6350 sequências codificadoras de proteína, 6004 sequências não-codificadoras de proteínas, 3.751 sequências indeterminadas e 86 sequências anotadas como ncRNAs.

5.5. Distribuição dos lncRNA putativos entre as espécies fúngicas

Conforme descrito na Metodologia desse trabalho, tratamos como lncRNA putativo todos aqueles transcritos que: (i) foram classificados como não codificantes de proteínas pelas três ferramentas de CP; (ii) não tiveram uma anotação funcional durante a análise pelo EnTAP; (iii) não tiveram uma anotação funcional durante a análise pelo INFERNAL. Abaixo apresentamos os lncRNAs putativos identificados durante a análise dos 19 transcriptomas e caracterizamos informações biomoleculares estruturais dessas sequências:

- A espécie *Alternaria* MG1 apresentou 2.119 lncRNA putativos com conteúdo GC% de 51.43%, comprimento mínimo de 201pb, médio de 459pb e máximo de 8.331pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A espécie *Arthrinium phaeospermum* apresentou 220 lncRNA putativos com conteúdo GC% de 52.10%, comprimento mínimo de 202pb, médio de 479pb e máximo de 9789pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A cepa *Aureobasidium melanogenum* (GEEO01) apresentou 609 lncRNA putativos com conteúdo GC% de 50,82%, comprimento mínimo de 201pb, médio de 394pb e máximo de 6.209pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 3 lncRNAs descritos até o momento.
- A cepa *Aureobasidium melanogenum* (GIXY01) apresentou 8 lncRNA putativos com conteúdo GC% de 49.54%, comprimento mínimo de 220pb, médio de 290pb e máximo de 635pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 3 lncRNAs descritos até o momento.
- A espécie *Candida africana* apresentou 141 lncRNA putativos com conteúdo GC% de 32.77%, comprimento mínimo de 201pb, médio de 366pb e máximo de 7.090pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.

- A espécie *Candida albicans* apresentou 37 lncRNA putativos com conteúdo GC% de 31.56%, comprimento mínimo de 204pb, médio de 416pb e máximo de 2.457pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 273 lncRNAs descritos até o momento.
- A espécie *Geotrichum candidum* apresentou 2.335 lncRNA putativos com conteúdo GC% de 48.26%, comprimento mínimo de 200pb, médio de 267pb e máximo de 1.246pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 22 lncRNAs descritos até o momento.
- A espécie Irpex lacteus apresentou 0 lncRNA putativos nesta etapa da análise
- A cepa *Mucor irregularis* (GHMU01) apresentou 1.849 lncRNA putativos com conteúdo GC% de 37.46%, comprimento mínimo de 202pb, médio de 931pb e máximo de 26.878pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A cepa *Mucor irregularis* (GFBC01) apresentou 3.995 lncRNA putativos com conteúdo GC% de 36.80%, comprimento mínimo de 351pb, médio de 523pb e máximo de 5.442pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A espécie *Papiliotrema laurentii* RY1 apresentou 1.136 lncRNA putativos com conteúdo GC% de 55.85%, comprimento mínimo de 240pb, médio de 398pb e máximo de 2192pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A espécie *Paracoccidioides brasiliensis* no isolado Pb18 apresentou 8 lncRNA putativos com conteúdo GC% de 45.17%, comprimento mínimo de 202pb, médio de 229pb e máximo de 490pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 132 lncRNAs descritos até o momento.
- A cepa *P. janthinellum* (GBSP01) apresentou 394 lncRNA putativos com conteúdo GC% de 50.33%, comprimento mínimo de 201pb, médio de 297pb e máximo de 12.258pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a

espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.

- A cepa *P. janthinellum* (GBSQ01) apresentou 271 lncRNA putativos com conteúdo GC% de 50.96%, comprimento mínimo de 202pb, médio de 979pb e máximo de 5.840pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A cepa *P. janthinellum* (GIPK01) apresentou 634 lncRNA putativos com conteúdo GC% de 51.63%, comprimento mínimo de 201pb, médio de 294pb e máximo de 11.723pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A espécie *Pneumocystis jirovecii* apresentou 1.365 lncRNA putativos com conteúdo GC% de 27.92%, comprimento mínimo de 205pb, médio de 968pb e máximo de 12.527pb. Conforme dados apresentados na Tabela 4, o banco de dados do Ensembl apresenta um total de 102 lncRNAs descritos até o momento.
- A espécie *Rhizopus oryzae* apresentou 6.906 lncRNA putativos com conteúdo GC% de 34.82%, comprimento mínimo de 201pb, médio de 288pb e máximo de 6.194pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A cepa *Yarrowia sp.* JCM 30694 apresentou 4.236 lncRNA putativos com conteúdo GC% de 46.51%, comprimento mínimo de 200pb, médio de 274pb e máximo de 2186pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.
- A cepa *Yarrowia sp.* JCM 30695 apresentou 3.188 lncRNA putativos com conteúdo GC% de 47.97%, comprimento mínimo de 200pb, médio de 275pb e máximo de 2.340pb. Conforme dados apresentados na Tabela 4, esses dados são inéditos para a espécie, já que inexistem informações sobre lncRNA nos bancos de dados do NCBI e do Ensembl.

Esses resultados demonstram que o presente trabalho proporcionou informações inéditas sobre lncRNAs para 13 fungos patogênicos dos 19 estudados. Além disso, o trabalho permitiu a prospecção de novos lncRNAs ainda não conhecidos para as 5 espécies de fungos que já apresentavam qualquer informação de lncRNA junto aos bancos de dados do NCBI ou ENSEMBL.

6. DISCUSSÃO

Nos últimos anos, diversos estudos tem mostrado que lncRNAs podem ser utilizados como marcadores de doença e alvos terapêuticos em várias condições (BADOWSKI et al., 2022), incluindo doenças infecciosas (WANG et al., 2022). Além disso, a prospecção de lncRNAs em fungos patogênicos tem sido um campo em rápido crescimento (SARAVANAKUMAR, 2022).

Em um estudo recente, Hovhannisyan e Gabaldón (2021) apresentaram um perfil de lncRNAs amplo em leveduras patogênicas do gênero *Candida*. Os autores analisaram a presença de lncRNAs em diferentes espécies do clado CTG, incluindo *Candida albicans*, *Candida tropicalis*, *Candida parapsilosis* e *Candida auris*. O estudo mostrou variações significativas na quantidade de lncRNA antisenso entre as espécies estudadas, com *C. albicans* apresentando o maior número. No entanto, essa diferença não foi observada quando apenas os dados de interação hospedeiro-patógeno foram analisados. Fatores como diferenças biológicas, variedade de amostras e número de amostras analisadas foram identificados como influenciando essas variações. Além disso, foi observado que a transcrição antisenso é extensa em *C. albicans*, mas é menos comum em outras espécies. O estudo também analisou a distribuição dos lncRNA ao longo dos cromossomos e encontrou variações na distribuição entre as espécies. Um fator importante a ser observado é que esse estudo foi utilizado para a construção do banco de dados de validação utilizado no presente trabalho. Dos novos lncRNAs putativos identificados no presente trabalho, observou-se uma correspondência de 501 lncRNAs em *Candida africana* e 231 em *Candida albicans* para os 21366 lncRNAs apresentados por Hovhannisyan e Gabaldón (2021). Além disso, os autores apresentaram 7222 lncRNAs de *Candida albicans* anteriormente não conhecidos. Deve-se destacar também que os autores não apresentaram lncRNAs para *Candida africana*, sendo os apresentados nessa dissertação os primeiros para essa espécie. Por fim, é importante observar que Hovhannisyan e Gabaldón (2021) utilizaram 2.561 bibliotecas de RNA-Seq disponíveis publicamente no banco de dados *Sequence Read Archive* (SRA), sendo 2.177 para *C. albicans*, 129 para *C. parapsilosis*, 123 para *C. glabrata*, 86 para *C. auris* e 46 para *C. tropicalis*, demonstrando assim a importância de bibliotecas públicas em bases de dados biológicas para a construção de novos conhecimentos com base no conceito de *Open Science*. No entanto, os autores não disponibilizaram os dados gerados em seu estudo na base de dados Transcriptome *Shotgun Assembly Sequence Database* (TSA), o que excluiu a possibilidade de

compararmos o nosso processo de prospecção de lncRNAs nos dados gerados pela equipe. Essa comparação seria de extrema importância para avaliar a replicabilidade do nosso fluxo de trabalho e da metodologia desenvolvida por Hovhannisyan e Gabaldón (2021).

Gao et al. (2021) também expandiu o conhecimento atual sobre os lncRNAs em *Candida auris*. Em seu trabalho, os autores identificaram um lncRNA chamado DINOR, envolvido com morfogênese, danos no DNA e replicação do DNA. Da mesma forma, os autores demonstraram que esse lncRNA estaria envolvido na orquestração de múltiplas respostas ao estresse.

Em um trabalho recente do nosso laboratório, Menegidio et al. (2021) descreveram um amplo perfil transcriptômico do fungo patogênico *Paracoccidioides brasiliensis* (Pb18) com a ajuda de uma abordagem baseada em RNA-seq. Os resultados confirmaram que cerca de 75% dos genes anotados no genoma de *P. brasiliensis* são transcritos *in vivo* e que cerca de 19% deles podem exibir isomorfos alternativos. Além disso, os autores identificaram 627 transcritos que não correspondiam a nenhum gene mapeado anteriormente no genoma, representados por 114 transcritos codificantes (provavelmente derivados de genes codificantes de proteínas anteriormente não mapeados) e 513 ncRNAs, incluindo 203 lncRNAs. Semelhante ao realizado com *Candida*, comparamos os lncRNAs obtidos na presente dissertação com os 203 lncRNAs descritos por Menegidio et al. (2021). Identificamos apenas 33 dos lncRNA putativos apresentados nesse trabalho entre os 203 lncRNAs apresentados por Menegidio et al. (2021). Quando os lncRNAs putativos são comparados com os 11766 transcritos disponibilizado pelos autores, observamos uma correspondência para 156 dos transcritos. Esses resultados demonstram, assim como relatado por Hovhannisyan e Gabaldón (2021), que a metodologia utilizada pode ocasionar um problema de replicabilidade dos resultados ao se utilizar o mesmo conjunto de dados. Enquanto que no presente trabalho foram utilizadas as ferramentas CPAT, CPC2 e RNASamba para a prospecção dos lncRNA, Menegidio et al. (2021) realizou a prospecção de potencial de codificação através das ferramentas *Coding Potential Predictor* (CPPred) (TONG et al., 2019), *Coding Potential Calculator* (CPC2) (KANG et al., 2017), *Coding-Noncoding Index* (CNCI) (SUN et al., 2013), *Coding-Potential Assessment Tool* (CPAT) (WANG et al., 2013) e *FlExible Extraction of LncRNAS* (FeeLNC) (WUCHER et al, 2017). Como apresentado por Iwakiri et al. (2016), Gu et al. (2017), Pinkney et al. (2020) e Duan et al. (2021), os diferentes algoritmos utilizados nas ferramentas de predição de lncRNA podem gerar inconsistências na identificação das sequências.

Para os outros fungos patogênicos estudados não foram encontrados outros trabalhos que buscaram identificar lncRNAs em dados de transcriptoma ou genoma. Da mesma forma, com exceção do gênero *Penicillium*, também não foram encontrados trabalhos para os gêneros associados as outras espécies analisadas durante essa dissertação. Esse resultado demonstra que a prospecção de lncRNAs permanece como uma área de importância para os estudos biomoleculares em fungos e que os dados aqui apresentados são inovadores e de extrema importância para ampliar a caracterização desses organismos patogênicos.

7. CONCLUSÃO

Em conclusão, este estudo demonstrou a importância da análise *in silico* de dados de transcriptomas publicados para a prospecção e caracterização de novos lncRNAs em fungos patogênicos humanos. A utilização de dados públicos de sequenciamento permitiu a identificação de novos lncRNAs, que podem servir de alvos potenciais para o desenvolvimento de terapias contra estes patógenos. Além disso, os resultados obtidos também sugerem que esses lncRNAs podem desempenhar papéis importantes na regulação da expressão gênica e na patogenicidade dos fungos estudados. Desta forma, é fundamental a continuidade da deposição de dados de sequenciamento em bancos de dados biológicos públicos, para que possam ser utilizados para a descoberta de novos alvos terapêuticos e para a compreensão dos mecanismos subjacentes à patogenicidade desses microrganismos.

8. PERSPECTIVAS FUTURAS

Como apresentado na discussão da presente dissertação, os dados gerados nesse trabalho foram apenas os seus primeiros passos na prospecção de lncRNAs em fungos patogênicos humanos. O fluxo de trabalho aqui apresentado também pode ser reaproveitado para a prospecção desses mesmos elementos biomoleculares em fungos não-patogênicos ou de interesse veterinário, agrícola, industrial, ecológico e/ou biotecnológico. Existem perspectivas de melhorias e modificações a fazer, sobretudo na perspectiva de tornar a presente prospecção em um processo contínuo, com todas as etapas integradas numa só unidade.

Além disso, é importante agregar os dados apresentados nessa dissertação no genoma dos organismos estudados, realizando uma reanotação dos mesmos com o acréscimo dos dados gerados nesse trabalho. Essa etapa proporcionará um ganho nos dados disponíveis para a comunidade científica e permitirá um melhor aproveitamento de dados genômicos, transcriptômicos e proteômicos pela área da micologia.

Por fim, novas propostas como análises focadas na co-expressão de genes, estrutura prevista, loci genômicos, redes de interação de lncRNA-miRNA-ncRNA-proteínas, entre outros, são perspectivas para ampliar o presente objeto de estudo.

REFERÊNCIAS

AALTO, A. P.; PASQUINELLI, A. E. Small noncoding RNAs mount a silent revolution in gene expression. **Current Opinion in Cell Biology**, v. 24, n. 3, p. 333–340, 1 jun. 2012.

ABBES, S. et al. Analysis of virulence factors and*in vivo*biofilm-forming capacity of*Yarrowia lipolytica* isolated from patients with fungemia. **Medical Mycology**, v. 55, n. 2, p. 193–202, 19 jul. 2016.

ABDELMOHSEN, K. et al. 7SL RNA represses p53 translation by competing with HuR. **Nucleic Acids Research**, v. 42, n. 15, p. 10099–10111, 14 ago. 2014.

ACIOLE BARBOSA, D. et al. Transcriptomic Profiling and Microsatellite Identification in Cobia (Rachycentron canadum), Using High-Throughput RNA Sequencing. **Marine Biotechnology**, v. 24, n. 1, p. 255–262, 2 dez. 2021.

ALBERTS, B. et al. **Biologia Molecular da Célula**. [s.l.] Artmed Editora, 2010.

ALEXOPOULOS, C. J. Introductory Mycology. **Soil Science**, v. 74, n. 6, p. 481, 1 dez. 1952.

AL-HATMI, A. M. S. et al. Large-scale generation and analysis of filamentous fungal DNA barcodes boosts coverage for kingdom fungi and reveals thresholds for fungal species and higher taxon delimitation. **Studies in Mycology**, v. 92, p. 135–154, mar. 2019.

AMIN, N.; MCGRATH, A.; CHEN, Y.-P. P. Evaluation of deep learning in noncoding RNA classification. **Nature Machine Intelligence**, v. 1, n. 5, p. 246–256, 13 maio 2019.

ANDREWS, T. S. et al. Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. **Nature Protocols**, v. 16, n. 1, p. 1–9, 1 jan. 2021.

ANDRIC, V. et al. A scaffold lncRNA shapes the mitosis to meiosis switch. **Nature Communications**, v. 12, n. 1, p. 770, 3 fev. 2021.

ANGRAND, P.-O. et al. The role of long noncoding RNAs in genome formatting and expression. **Frontiers in Genetics**, v. 6, 29 abr. 2015.

ANVISA. **Identificação de caso de Candida auris no Brasil | Biblioteca Virtual em Saúde MS**. Disponível em: <https://bvsms.saude.gov.br/identificacao-de-caso-de-candida-auris-no-brasil/>. Acesso em: 18 fev. 2023.

ARITA, T. et al. Circulating long noncoding RNAs in plasma of patients with gastric cancer. **Anticancer Research**, v. 33, n. 8, p. 3185–3193, 1 ago. 2013.

ASADZADEH, M. et al. Molecular identification, genotypic heterogeneity and comparative pathogenicity of environmental isolates of *Papiliotrema laurentii*. **Journal of Medical Microbiology**, v. 69, n. 11, p. 1285–1292, 1 nov. 2020.

ASHBURNER, M. et al. Gene Ontology: tool for the unification of biology. **Nature Genetics**, v. 25, n. 1, p. 25–29, maio 2000.

BADOWSKI, C.; HE, B.; GARMIRE, L. X. Blood-derived lncRNAs as biomarkers for cancer diagnosis: the Good, the Bad and the Beauty. **npj Precision Oncology**, v. 6, n. 1, p. 1–18, 21 jun. 2022.

BARTEL, D. P. Metazoan MicroRNAs. **Cell**, v. 173, n. 1, p. 20–51, mar. 2018.

BARTEL, D. P. MicroRNAs: Target Recognition and Regulatory Functions. **Cell**, v. 136, n. 2, p. 215–233, jan. 2009.

BATISTA DOS SANTOS, S. E. et al. Identification and Characterization of Polymorphisms in piRNA Regions. **Current Issues in Molecular Biology**, v. 44, n. 2, p. 942–951, 1 fev. 2022.

BENSASSON, D. et al. Diverse Lineages of Candida albicans Live on Old Oaks. **Genetics**, v. 211, n. 1, p. 277–288, 1 jan. 2019.

BERRETTA, J.; MORILLON, A. Pervasive transcription constitutes a new level of eukaryotic genome regulation. **EMBO reports**, v. 10, n. 9, p. 973–982, 14 ago. 2009.

BERTO, C. et al. BASES DA RESISTÊNCIA ANTIFÚNGICA: UMA REVISÃO COMENTADA. **Revista Uningá**, v. 55, n. 3, p. 52–71, 20 set. 2018.

BIELSKA, E. & MAY, R. C. Quantifying donor-to-donor variation in macrophage responses to the human fungal pathogen *Cryptococcus neoformans*. **PLOS ONE**, v. 13, n. 3, p. e0194615, 29 mar. 2019.

BIELSKA, E.; MAY, R. C. Extracellular vesicles of human pathogenic fungi. **Current Opinion in Microbiology**, v. 52, p. 90–99, dez. 2019.

BLOOM, A. L. M. et al. Translational Regulation Promotes Oxidative Stress Resistance in the Human Fungal Pathogen *Cryptococcus neoformans*. **mBio**, v. 10, n. 6, 24 dez. 2019.

BONGOMIN, F. et al. Global and Multi-National Prevalence of Fungal Diseases—Estimate Precision. **Journal of Fungi**, v. 3, n. 4, p. 57, 18 out. 2017.

BRANNAN, C. I. et al. The product of the H19 gene may function as an RNA. **Molecular and Cellular Biology**, v. 10, n. 1, p. 28–36, jan. 1990.

BROOKS, A. M. V.; LAZARUS, M. G.; WEINER, J. M. Soft contact lens contamination by *Alternaria alternata*. **Medical Journal of Australia**, v. 140, n. 8, p. 490–491, abr. 1984.

BROSNAN, C. A.; VOINNET, O. The long and the short of noncoding RNAs. **Current Opinion in Cell Biology**, v. 21, n. 3, p. 416–425, jun. 2009.

BROWN, C. J. et al. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. **Cell**, v. 71, n. 3, p. 527–542, out. 1992.

BRUNEKREEF, B. et al. Associations between markers of respiratory morbidity in European children. **European Respiratory Journal**, v. 19, n. 3, p. 479–486, 1 mar. 2002.
BUCHFINK, B.; XIE, C.; HUSON, D. H. Fast and sensitive protein alignment using

DIAMOND. **Nature Methods**, v. 12, n. 1, p. 59–60, 17 nov. 2014.

BUZINA, W. et al. The Polypore Mushroom Irpex lacteus, a New Causative Agent of Fungal Infections. **Journal of Clinical Microbiology**, v. 43, n. 4, p. 2009–2011, 1 abr. 2005.

CALAÇA, F. J. S.; XAVIER-SANTOS, S. New records of coprophilous ascomycetes (Fungi: Ascomycota) from Brazil and Neotropical Region. **Check List**, v. 12, n. 6, p. 2009, 6 dez. 2016.

CALDERÓN, E. J. et al. High prevalence of *Pneumocystis jirovecii* colonization among HIV-positive patients in southern Brazil. **Medical Mycology**, v. 52, n. 8, p. 804–809, 6 out. 2014.

CAMACHO, C. et al. BLAST+: architecture and applications. **BMC Bioinformatics**, v. 10, n. 1, p. 421, 2009.

CAMARGO, A. P. et al. RNAsamba: neural network-based assessment of the protein-coding potential of RNA sequences. **NAR genomics and bioinformatics**, v. 2, n. 1, p. lqz024, 1 mar. 2020.
CANTANI, A.; CIASCHI, V. Epidemiology of alternaria alternata allergy: a prospective study in 6840 Italian asthmatic children. **European review for medical and pharmacological sciences**, v. 8, n. 6, p. 289–294, 1 nov. 2004.

CARTHEW, R. W.; SONTHEIMER, E. J. Origins and Mechanisms of miRNAs and siRNAs. **Cell**, v. 136, n. 4, p. 642–655, fev. 2009.

CASADEVALL, A.; NOSANCHUK, J. D. Impact of Melanin on Microbial Virulence and Clinical Resistance to Antimicrobial Compounds. **Antimicrobial Agents and Chemotherapy**, v. 50, n. 11, p. 3519–3528, nov. 2006.

CATARINE, A.; COSTA, T. **UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO CENTRO DE CIÊNCIAS DA SAÚDE PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA Resposta a estresses consecutivos em *Saccharomyces cerevisiae***. [s.l: s.n.]. Disponível em: <http://repositorio.ufes.br:8080/bitstream/10/7103/1/tese_10830_Disserta%C3%A7%C3%A3o_Ane%20Catarine%20Tosi%20Costa.pdf>. Acesso em: 13 fev. 2023.

CERASE, A. et al. Xist localization and function: new insights from multiple levels. **Genome Biology**, v. 16, n. 1, 15 ago. 2015.

ČERNOŠA, A. et al. Virulence Traits and Population Genomics of the Black Yeast Aureobasidium melanogenum. **Journal of Fungi**, v. 7, n. 8, p. 665, 16 ago. 2021.

CHACKO, N. et al. The lncRNA RZE1 Controls Cryptococcal Morphological Transition. **PLOS Genetics**, v. 11, n. 11, p. e1005692, 20 nov. 2015.

CHANDER, J. et al. Mucormycosis: Battle with the Deadly Enemy over a Five-Year Period in India. **Journal of Fungi**, v. 4, n. 2, p. 46, 1 jun. 2018.

CHAO, C.-M.; LAI, C.-C.; YU, W.-L. COVID-19 associated mucormycosis – An emerging threat. **Journal of Microbiology, Immunology and Infection**, jan. 2022.

CHATURVEDI, V.; DE HOOG, G. S. Onygenalean Fungi as Major Human and Animal Pathogens. **Mycopathologia**, v. 185, n. 1, p. 1–8, 30 jan. 2020.

CHAUDHARY, R. Isolation and identification of bioactive compounds from *Irpex lacteus* Wild Fleshy Fungi. **Journal of Pharmaceutical Sciences and Research**, v. 7, n. 7, p. 424, 2015.

CHAVES, A. F. A. et al. What is the elephant in the room when considering new therapies for fungal diseases? **Critical Reviews in Microbiology**, v. 47, n. 3, p. 275–289, 29 jan. 2021.

CHEN, B. et al. Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis. **Scientific Reports**, v. 6, n. 1, mar. 2016.

CHEN, J. et al. Evolutionary analysis across mammals reveals distinct classes of long noncoding RNAs. **Genome Biology**, v. 17, n. 1, 2 fev. 2016.

CHEN, W. et al. The Long Non-coding RNA Landscape of the Mouse Eye. **Investigative Ophthalmology & Visual Science**, v. 58, n. 14, p. 6308–6317, 1 dez. 2017.

CHEVREUX, B. Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. **Genome Research**, v. 14, n. 6, p. 1147–1159, 12 maio 2004.

CHUNG, H.; LEE, Y.-H. Hypoxia: A Double-Edged Sword During Fungal Pathogenesis? **Frontiers in Microbiology**, v. 11, 2020.

CLOSS, F. K.; PUTZKE, J. Fungos aquáticos (Oomycota, Chytridiomycota) ocorrentes em anfíbios anuros em dois remanescentes de Mata Mtlântica, localizados em Santa Cruz do Sul e Venâncio Aires, RS, Brasil. **Caderno de Pesquisa**, v. 29, n. 1, p. 01, 5 jan. 2017.

CUSHION, M. T.; STRINGER, J. R. Stealth and Opportunism: Alternative Lifestyles of Species in the Fungal Genus *Pneumocystis*. **Annual Review of Microbiology**, v. 64, n. 1, p. 431–452, 13 out. 2010.

DAHM, R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. **Human genetics**, v. 122, n. 6, p. 565–81, 2008.

DANG, H. X. et al. The Alternaria genomes database: a comprehensive resource for a fungal genus comprised of saprophytes, plant pathogens, and allergenic species. **BMC Genomics**, v. 16, n. 1, 25 mar. 2015.

DE HOOG, G. S. et al. Commentaries: Name Changes in Medically Important Fungi and Their Implications for Clinical Practice. **Journal of Clinical Microbiology**, v. 53, n. 4, p. 1056–1062, abr. 2015.

DENNING, D. W.; BROMLEY, M. J. How to bolster the antifungal pipeline. **Science**, v. 347, n. 6229, p. 1414–1416, 26 mar. 2015.

DIAKOMOPOULOS, A.; DALAMAGA, M.; PAPADAVID, E. Understanding the enigmatic association between mycosis fungoides and psoriasis: Report of two cases and review of the literature. **Metabolism Open**, v. 12, p. 100148, 1 dez. 2021.

DJEBALI, S. et al. Landscape of transcription in human cells. **Nature**, v. 489, n. 7414, p. 101–108, 1 set. 2012.

DOMINGUETE, L. C. B. Avaliação da produção de antioxidantes e inibidores de acetilcolinesterase por fungos filamentosos. **repositorio.ufmg.br**, 28 maio 2021.

DOWNS, SARA H. et al. Clinical Importance of *Alternaria* Exposure in Children. **American Journal of Respiratory and Critical Care Medicine**, v. 164, n. 3, p. 455–459, 1 ago. 2001.

DUAN, Y. et al. A systematic evaluation of bioinformatics tools for identification of long noncoding RNAs. **RNA**, p. rna.074724.120, 14 out. 2020.

DUAN, Z. et al. Comprehensive investigation of a dye-decolorizing peroxidase and a manganese peroxidase from Irpex lacteus F17, a lignin-degrading basidiomycete. **AMB Express**, v. 8, n. 1, 17 jul. 2018.

DURLACHER, R. R. [UNESP. Lesões orais e correlação com aspectos clínicos, epidemiológicos e laboratoriais da paracoccidioidomicose. **repositorio.unesp.br**, 17 out. 2017.

EGIDI, E. et al. A few Ascomycota taxa dominate soil fungal communities worldwide. **Nature Communications**, v. 10, n. 1, 30 maio 2019.

ELGART, M. L. Atlas of clinical fungi. **Journal of the American Academy of Dermatology**, v. 37, n. 2, p. 291, ago. 1997.

EL-GEBALI, S. et al. The Pfam protein families database in 2019. **Nucleic Acids Research**, v. 47, n. D1, p. D427–D432, 24 out. 2018.

ENOCH, D. A. Invasive fungal infections: a review of epidemiology and management options. **Journal of Medical Microbiology**, v. 55, n. 7, p. 809–818, 1 jul. 2006.

ENTEZARI, M. et al. LncRNA-miRNA axis in tumor progression and therapy response: An emphasis on molecular interactions and therapeutic interventions. **Biomedicine & Pharmacotherapy**, v. 154, p. 113609, 1 out. 2022.

FARR, D.F., & ROSSMAN, A.Y. **Fungal Databases, U.S. National Fungus Collections**. Disponível em: <https://nt.ars-grin.gov/fungaldatabases/>. Acesso em: 13 fev. 2023.

FENG, J. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. **Genes & Development**, v. 20, n. 11, p. 1470–1484, 1 jun. 2006.

FINDLEY, K. et al. Human Skin Fungal Diversity. **Nature**, v. 498, n. 7454, p. 367–370, 20 jun. 2013.

FINN, R. D. et al. Pfam: the protein families database. **Nucleic Acids Research**, v. 42, n. D1, p. D222–D230, 27 nov. 2013.

FISHER, M. C. et al. Biogeographic range expansion into South America by Coccidioides immitis mirrors New World patterns of human migration. **Proceedings of the National Academy of Sciences**, v. 98, n. 8, p. 4558–4562, 3 abr. 2001.

FISHER, M. C.; GARNER, T. W. J. Chytrid fungi and global amphibian declines. **Nature Reviews Microbiology**, 25 fev. 2020.

FISHER, M. C.; HUDOME, S. M. Nosocomial infections in the neonatal intensive care unit. **Current Opinion in Infectious Diseases**, v. 14, n. 3, p. 303, 1 jun. 2001.

FORTUNA, J. L. Apostila da Disciplina de Biologia dos Fungos. Teixeira de Freitas: Projeto Fungus Extremus, UNEB, Campus X. 2020, 27 p.

FRISVAD, J. C. & SAMSON, R. A. Aspergillus is monophyletic: Evidence from multiple gene phylogenies and extrolites profiles. **Studies in Mycology**, v. 85, p. 199–213, set. 2016.

FRITH, M. C. et al. Discrimination of Non-Protein-Coding Transcripts from Protein-Coding mRNA. **RNA Biology**, v. 3, n. 1, p. 40–48, jan. 2006.

GAO, J. et al. LncRNA DINOR is a virulence factor and global regulator of stress responses in Candida auris. **Nature Microbiology**, v. 6, n. 7, p. 842–851, 1 jul. 2021.

GARCIA-HERMOSO, D. et al. [Fungal keratitis: A 5-year monocentric retrospective study on Reunion Island]. **Journal francais d'ophtalmologie**, v. 41, n. 4, p. 321–325, 1 abr. 2018.

GARCÍA-JIMÉNEZ, J. et al. Detection of Grapevine Fungal Trunk Pathogens on Pruning Shears and Evaluation of Their Potential for Spread of Infection. **Plant Disease**, v. 99, n. 7, p. 976–981, jul. 2015.

GERAMISHOAR, M. et al. Laboratory and Epidemiology Communications First Case of Cerebral Phaeohyphomycosis Caused by Nattrassia mangiferae in Iran. **Jpn. J. Infect. Dis**, v. 57, 2004.

GHAREHBOLAGH, S. A. et al. Distribution, antifungal susceptibility pattern and intra-*Candida albicans* species complex prevalence of Candida africana: A systematic review and meta-analysis. **PLOS ONE**, v. 15, n. 8, p. e0237046, 20 ago. 2020.

GIARDINE, B. Galaxy: A platform for interactive large-scale genome analysis. **Genome Research**, v. 15, n. 10, p. 1451–1455, 16 set. 2005.

GIRARDI, L. S. et al. Aureobasidium pullulans septicemia. **Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America**, v. 16, n. 2, p. 338–339, 1 fev. 1993.

GOMES, A.; NOLASCO, S.; SOARES, H. Noncoding RNAs: Multi-Tasking Molecules in the Cell. **International Journal of Molecular Sciences**, v. 14, n. 8, p. 16010–16039, 31 jul. 2013.

GONG, C.; MAQUAT, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. **Nature**, v. 470, n. 7333, p. 284–288, fev. 2011.

GOOD, M. C.; ZALATAN, J. G.; LIM, W. A. Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. **Science**, v. 332, n. 6030, p. 680–686, 5 maio 2011.

GOSTINČAR, C. et al. Fungi between extremotolerance and opportunistic pathogenicity on humans. **Fungal Diversity**, v. 93, n. 1, p. 195–213, nov. 2018.

GRABHERR, M. G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nature Biotechnology, v. 29, n. 7, p. 644–652, 15 maio 2011.

GRIFFITHS-JONES, S. Rfam: an RNA family database. **Nucleic Acids Research**, v. 31, n. 1, p. 439–441, 1 jan. 2003.

GROENEWALD, M. et al. *Yarrowia lipolytica*: Safety assessment of an oleaginous yeast with a great industrial potential. **Critical Reviews in Microbiology**, v. 40, n. 3, p. 187–206, 14 mar. 2013.

GU, C. et al. Global network random walk for predicting potential human lncRNA-disease associations. **Scientific Reports**, v. 7, n. 1, p. 12442, 29 set. 2017.

GUPTA, M. K. et al. Endophytic Fungi: A Source of Potential Antifungal Compounds. **Journal of Fungi**, v. 4, n. 3, p. 77, 25 jun. 2018.

GUTTMAN, M. et al. Chromatin signature reveals over a thousand highly conserved large noncoding RNAs in mammals. **Nature**, v. 458, n. 7235, p. 223–227, 1 fev. 2009.

HACISULEYMAN, E. et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. **Nature Structural & Molecular Biology**, v. 21, n. 2, p. 198–206, 26 jan. 2014.

HAN, H. et al. Aggregation-Induced Emission-Based Platforms for the Treatment of Bacteria, Fungi, and Viruses. **Advanced Healthcare Materials**, v. 10, n. 24, p. 2100736, 30 jun. 2021.

HAN, L. et al. Low expression of long noncoding RNA GAS6-AS1 predicts a poor prognosis in patients with NSCLC. **Medical Oncology**, v. 30, n. 4, 24 ago. 2013.

HAN, Q. et al. Long noncoding RNA CRCMSL suppresses tumor invasive and metastasis in colorectal carcinoma through nucleocytoplasmic shuttling of HMGB2. **Oncogene**, v. 38, n. 16, p. 3019–3032, 1 abr. 2019.

HART, A. J. et al. EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. **Molecular Ecology Resources**, v. 20, n. 2, p. 591–604, 31 dez. 2019.

HASSAN, M. I. A.; VOIGT, K. Pathogenicity patterns of mucormycosis: epidemiology, interaction with immune cells and virulence factors. **Medical Mycology**, v. 57, n. Supplement_2, p. S245–S256, 28 fev. 2019.

HAWKES, M. et al. *Aureobasidium pullulans* infection: Fungemia in an infant and a review of human cases. **Diagnostic Microbiology and Infectious Disease**, v. 51, n. 3, p. 209–213, mar. 2005.

HE, Y. et al. Phaeohyphomycosis in China. **Frontiers in Cellular and Infection Microbiology**, v. 12, 13 jun. 2022.

HEMASHETTAR, B. M. et al. Chronic Rhinofacial Mucormycosis Caused by *Mucor irregularis* (*Rhizomucor variabilis*) in India. **Journal of Clinical Microbiology**, v. 49, n. 6, p. 2372–2375, jun. 2011.

HEO, J. B.; SUNG, S. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. **Science**, v. 331, n. 6013, p. 76–79, 2 dez. 2010.

HOLLEY, C. L.; TOPKARA, V. K. An Introduction to Small Noncoding RNAs: miRNA and snoRNA. **Cardiovascular Drugs and Therapy**, v. 25, n. 2, p. 151–159, abr. 2011.

HOUBRAKEN, J. et al. *Penicillium araracuarense* sp. nov., *Penicillium elleniae* sp. nov., *Penicillium penarojense* sp. nov., *Penicillium vanderhammenii* sp. nov. and *Penicillium wotroi* sp. nov., isolated from leaf litter. **INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY**, v. 61, n. 6, p. 1462–1475, 16 jul. 2010.

HOVHANNISYAN, H.; GABALDÓN, T. The Noncoding RNA landscape of *Candida* yeast pathogens. **Nature Communications**, v. 12, n. 1, dez. 2021.

HU, H.; JIE, H.-Y.; ZHENG, X.-X. Three Circulating LncRNA Predict Early Progress of Esophageal Squamous Cell Carcinoma. **Cellular Physiology and Biochemistry**, v. 40, n. 1-2, p. 117–125, 2016.

HUNG, T. et al. Extensive and coordinated transcription of noncoding RNAs within cell cycle promoters. **Nature genetics**, v. 43, n. 7, p. 621–629, 5 jun. 2011.

HUNG, T.; CHANG, H. Y. Long noncoding RNA in genome regulation. **RNA Biology**, v. 7, n. 5, p. 582–585, set. 2010.

HUTCHINS, A. P.; PEI, D. Transposable elements at the center of the crossroads between embryogenesis, embryonic stem cells, reprogramming, and long noncoding RNAs. **Science Bulletin**, v. 60, n. 20, p. 1722–1733, out. 2015.

IBANEZ PEREZ, R. et al. Peritonitis by *Aureobasidium pullulans* in continuous ambulatory peritoneal dialysis. **Nephrology Dialysis Transplantation**, v. 12, n. 7, p. 1544–1545, 1 jul. 1997.

INTRA, J.; SARTO, C.; BRAMBILLA, P. A rare case of cutaneous *Papiliotrema (Cryptococcus) laurentii* infection in a 23-year-old Caucasian woman affected by an

autoimmune thyroid disorder with hypothyroidism. **European Journal of Clinical Microbiology & Infectious Diseases**, v. 40, n. 3, p. 647–650, 6 out. 2020.

IRGA, N. et al. Hyperthyroidism Secondary to Disseminated Mucormycosis in a Child with Acute Lymphoblastic Leukemia: Case Report and a Review of Published Reports. **Mycopathologia**, v. 175, n. 1-2, p. 123–127, 25 set. 2012.

IWAKIRI, J.; HAMADA, M.; ASAI, K. Bioinformatics tools for lncRNA research. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, v. 1859, n. 1, p. 23–30, 1 jan. 2016.

JACQUES et al. DD3PCA3, a Very Sensitive and Specific Marker to Detect Prostate Tumors. **Cancer Research**, v. 62, n. 9, p. 2695–2698, maio 2002.

JARROUX, J.; MORILLON, A.; PINSKAYA, M. History, Discovery, and Classification of lncRNAs. **Advances in Experimental Medicine and Biology**, p. 1–46, 2017.

JENSEN, L. J. et al. eggNOG: automated construction and annotation of orthologous groups of genes. **Nucleic Acids Research**, v. 36, n. Database, p. D250–D254, 23 dez. 2007.

JEONG, W. et al. The epidemiology and clinical manifestations of mucormycosis: a systematic review and meta-analysis of case reports. **Clinical Microbiology and Infection**, v. 25, n. 1, p. 26–34, jan. 2019.

KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v. 28, n. 1, p. 27–30, 1 jan. 2000.

KANG, Y.-J. et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. **Nucleic Acids Research**, v. 45, n. W1, p. W12–W16, 18 maio 2017.

KANNAN, S. et al. Transposable Element Insertions in Long Intergenic Noncoding RNA Genes. **Frontiers in Bioengineering and Biotechnology**, v. 3, 9 jun. 2015.

KANO, Y. et al. First case report of fungal meningitis due to a polypore mushroom *Irpex lacteus*. **Neurological Sciences**, v. 42, n. 5, p. 2075–2078, 5 jan. 2021.

KAPRANOV, P. et al. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. **Science**, v. 316, n. 5830, p. 1484–1488, 8 jun. 2007.

KAPUSTA, A. et al. Transposable Elements Are Major Contributors to the Origin, Diversification, and Regulation of Vertebrate Long Noncoding RNAs. **PLoS Genetics**, v. 9, n. 4, p. e1003470, 25 abr. 2013.

KAPUSTA, A.; FESCHOTTE, C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. **Trends in Genetics**, v. 30, n. 10, p. 439–452, out. 2014.

KAUR, M.; AGGARWAL, N. K. Enzymatic activities of pathogenic species of *Alternaria*, isolated from *Parthenium*. **Indian Journal of Weed Science**, v. 49, n. 2, p. 207, 2017.

KEENE, S. et al. Cutaneous geotrichosis due to *Geotrichum candidum* in a burn patient. **Access Microbiology**, v. 1, n. 1, 1 mar. 2019.

KENNEDY, M. A.; SOBEL, J. D. Vulvovaginal Candidiasis Caused by Non-albicans *Candida* Species: New Insights. **Current Infectious Disease Reports**, v. 12, n. 6, p. 465–470, 4 set. 2010.

KHALIL, A. M. et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 28, p. 11667–72, 2009.

KHEDRI, S. et al. Iranian HIV/AIDS patients with oropharyngeal candidiasis: identification, prevalence and antifungal susceptibility of Candida species. **Letters in Applied Microbiology**, v. 67, n. 4, p. 392–399, 16 ago. 2018.

KINO, T. et al. Noncoding RNA Gas5 Is a Growth Arrest- and Starvation-Associated Repressor of the Glucocorticoid Receptor. **Science Signaling**, v. 3, n. 107, p. ra8–ra8, 2 fev. 2010.

KLOTZ, S. A. et al. Serum Amyloid P Component and Systemic Fungal Infection: Does It Protect the Host or Is It a Trojan Horse? **Open Forum Infectious Diseases**, v. 3, n. 3, p. ofw166, maio 2016.

KÖHLER, J. R. et al. Fungi that Infect Humans. **Microbiology Spectrum**, v. 5, n. 3, 9 jun. 2017.

KÖHLER, J. R.; CASADEVALL, A.; PERFECT, J. The Spectrum of Fungi That Infects Humans. **Cold Spring Harbor Perspectives in Medicine**, v. 5, n. 1, p. a019273–a019273, 3 nov. 2014.

KONOPKA, J. B. et al. One Health: Fungal Pathogens of Humans, Animals, and Plants. **American Society for Microbiology**: 18 out, 2019.

KORPELAINEN, E. et al. **RNA-seq Data Analysis: A Practical Approach**. [s.l.] CRC Press, 2014.

KÜES, U. Fungal enzymes for environmental management. **Current Opinion in Biotechnology**, v. 33, p. 268–278, jun. 2015.

KUGEL, J. F.; GOODRICH, J. A. Noncoding RNAs: key regulators of mammalian transcription. **Trends in Biochemical Sciences**, v. 37, n. 4, p. 144–151, 1 abr. 2012.

KUMAR, K. K.; HALLIKERI, K. Phaeohyphomycosis. **Indian Journal of Pathology and Microbiology**, v. 51, n. 4, p. 556, 2008.

KURTZMAN, C. P.; ROBNETT, C. J. Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5’ end of the large-subunit (26S) ribosomal DNA gene. **Journal of clinical microbiology**, v. 35, n. 5, p. 1216–1223, 1997.

KWON-CHUNG, K. J.; BENNETT, J. E. Medical mycology. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 34, p. 504–504, 1 dez. 1992.

KYRIAKOU, A. et al. Fungal Infections and Nail Psoriasis: An Update. **Journal of Fungi**, v. 8, n. 2, p. 154, 3 fev. 2022.

LANIGAN, G.; PAYNE, A.; COCKRUM, P. PRODUCTION OF TREMORGENIC TOXINS BY *PENICILLIUM JANTHINELLUM* BIOURGE: A POSSIBLE AETIOLOGICAL FACTOR IN RYEGRASS STAGGERS. **Australian Journal of Experimental Biology and Medical Science**, v. 57, n. 1, p. 31–37, fev. 1979.

LAURENT, G. Sur un nouveau genre de champignons ascomycètes: *Papiliotrema*. **Annales de Mycologie**, 46(1), 1-15. 1948.

LEE, J. W. et al. Biosynthesis of Novel Exopolymers by *Aureobasidium pullulans*. **Applied and Environmental Microbiology**, v. 65, n. 12, p. 5265–5271, dez. 1999.

LEE, S. et al. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. **Cell**, v. 164, n. 1-2, p. 69–80, jan. 2016.

LETUNIC, I.; DOERKS, T.; BORK, P. SMART: recent updates, new developments and status in 2015. **Nucleic Acids Research**, v. 43, n. D1, p. D257–D260, 9 out. 2014.

LI, B.; DEWEY, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 1, 4 ago. 2011.

LI, H. et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078–2079, 8 jun. 2009.

LI, H. New strategies to improve minimap2 alignment accuracy. **Bioinformatics**, v. 37, n. 23, p. 4572–4574, 8 out. 2021.

LI, J. et al. LncRNA TUG1 acts as a tumor suppressor in human glioma by promoting cell apoptosis. **Experimental Biology and Medicine**, v. 241, n. 6, p. 644–649, 1 mar. 2016.

LI, J. et al. *Pneumocystis carinii* Pneumonia in Patients With Connective Tissue Disease. **JCR: Journal of Clinical Rheumatology**, v. 12, n. 3, p. 114, 1 jun. 2006.

LI, R. et al. SOAP: short oligonucleotide alignment program. **Bioinformatics**, v. 24, n. 5, p. 713–714, 28 jan. 2008.

LI, S. et al. Whole-genome sequence of *Arthrinium phaeospermum*, a globally distributed pathogenic fungus. **Genomics**, v. 112, n. 1, p. 919–929, jan. 2020.

LI, X. et al. *Penicillium janthinellum* Pneumonia in an SLE Patient: A Case Study. **Infection and Drug Resistance**, v. Volume 13, p. 2745–2749, ago. 2020.

LIN, A. et al. The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer. **Nature Cell Biology**, v. 18, n. 2, p. 213–224, 11 jan. 2016.

LOBRITZ, R. W. Granulomatous Pulmonary Disease Secondary to *Alternaria*. **JAMA: The Journal of the American Medical Association**, v. 241, n. 6, p. 596, 9 fev. 1979.

LOGUERCIO-LEITE, C.; CAMPOS-SANTANA, M. DE. Mycodiversity of xylophilous basidiomycetes (Basidiomycota, Fungi) in Mondaí, Santa Catarina, Brazil II: A new addition. **Biotemas**, v. 23, n. 1, p. 13–18, 26 abr. 2010.

LOMBARD, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. D1, p. D490–D495, 21 nov. 2013.

LONDERO, M. R. et al. Catheter-related infection due to *Papiliotrema laurentii* in an oncologic patient: Case report and systematic review. **The Brazilian Journal of Infectious Diseases**, v. 23, n. 6, p. 451–461, 1 nov. 2019.

LORENZEN, J. M. et al. Long noncoding RNAs in urine are detectable and may enable early detection of acute T cell–mediated rejection of renal allografts. **Clinical chemistry**, v. 61, n. 12, p. 1505-1514, 2015.

LU X.L., LIU Z.H., SHEN Y.N. et al. Primary cutaneous zygomycosis caused by *Rhizomucor variabilis*: a new endemic zygomycosis? A case report and review of 6 cases reported from China. **Clin Infect Dis**. 2009; 49: e39–43.

LU X.L., NAJAFZADEH M.J., DOLATABADI S. et al. Taxonomy and epidemiology of *Mucor irregularis*, agent of chronic cutaneous mucormycosis. **Persoonia**. 2013; 30: 48–56.

MAHLO, S. M.; MCGAW, L. J.; ELOFF, J. N. Antifungal activity and cytotoxicity of isolated compounds from leaves of *Breonadia salicina*. **Journal of Ethnopharmacology**, v. 148, n. 3, p. 909–913, jul. 2013.

MARGULIES, M.et al. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v. 437, n. 7057, p. 376-380, 2005.

MARION et al. DD3: A New Prostate-specific Gene, Highly Overexpressed in Prostate Cancer1. **Cancer Research**, v. 59, n. 23, p. 5975–5979, dez. 1999.

MARQUES, A. C. et al. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. **Genome biology**, v. 14, n. 11, p. 1-14, 2013.

MARTIANOV, I. et al. Repression of the human dihydrofolate reductase gene by a noncoding interfering transcript. **Nature**, v. 445, n. 7128, p. 666-670, 2007.

MARTINS, C., DIAS, C., & MARTINS, A. (2017). *Irpex lacteus* infections: a clinical and epidemiological review. **Journal of clinical microbiology**, 55(3), 731-739.

MATOS, T. et al. Organization of Patient Management and Fungal Epidemiology in Cystic Fibrosis. **Mycopathologia**, v. 183, n. 1, p. 7–19, 3 nov. 2018.

MATTICK, J. S. RNA regulation: a new genetics? **Nature Reviews Genetics**, v. 5, n. 4, p. 316–323, abr. 2004.

MEAD, M. E. et al. Characterizing the Pathogenic, Genomic, and Chemical Traits of *Aspergillus fischeri*, a Close Relative of the Major Human Fungal Pathogen *Aspergillus fumigatus*. **mSphere**, v. 4, n. 1, 20 fev. 2019.

MENDELL, JOSHUA T.; OLSON, ERIC N. MicroRNAs in Stress Signaling and Human Disease. **Cell**, v. 148, n. 6, p. 1172–1187, mar. 2012.

MENEGIDIO, F. B. et al. Transcriptomic profiling identifies novel transcripts, isomorphs, and noncoding RNAs in *Paracoccidioides brasiliensis*. **Medical Mycology**, v. 59, n. 2, p. 197–200, 21 jul. 2020.

MILLAR, B. C. et al. Isolation of *Alternaria alternata* from an Emollient Cream: Implications for Public Health. **Mycopathologia**, v. 156, n. 4, p. 273–277, 2003.

MISHRA, P. et al. The Spectrum of Invasive Fungal Sinusitis in COVID-19 Patients: Experience from a Tertiary Care Referral Center in Northern India. **Journal of Fungi**, v. 8, n. 3, p. 223, 24 fev. 2022.

MISHRA, S. K.; WANG, H. Computational Analysis Predicts Hundreds of Coding lncRNAs in Zebrafish. **Biology**, v. 10, n. 5, p. 371, 1 maio 2021.

MOLINA, B. et al. [Infectious myositis]. **La Revue De Medecine Interne**, v. 41, n. 4, p. 241–249, 1 abr. 2020.

MONDAL, T. et al. Characterization of the RNA content of chromatin. **Genome Research**, v. 20, n. 7, p. 899–907, 19 abr. 2010.

MUNIZ, E. R. Termotolerância de *Metarhizium* spp. e efeito de sua formulação sobre a virulência para *Rhipicephalus microplus* expostos a condições estressantes de temperatura e umidade. **repositorio.bc.ufg.br**, 1 abr. 2015.

MURRAY, P. R.; ROSENTHAL, K. S.; PFALLER, M. A. **Microbiología médica**. [s.l.] Elsevier Health Sciences, 2017.

NARANJO-ORTIZ, M. A.; GABALDÓN, T. Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. **Biological Reviews**, v. 94, n. 6, p. 2101–2137, 29 out. 2019.

NAWROCKI, E. P.; KOLBE, D. L.; EDDY, S. R. Infernal 1.0: inference of RNA alignments. **Bioinformatics**, v. 25, n. 10, p. 1335–1337, 23 mar. 2009.

NECSULEA, A. et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. **Nature**, v. 505, n. 7485, p. 635-640, 2014.

NOBILE, C. J.; JOHNSON, A. D. *Candida albicans* Biofilms and Human Disease. **Annual Review of Microbiology**, v. 69, n. 1, p. 71–92, 15 out. 2015.

NOVIKOVA, I. V. et al. Rise of the RNA Machines: Exploring the Structure of Long noncoding RNAs. **Journal of Molecular Biology**, v. 425, n. 19, p. 3731–3746, out. 2013.

NOVOHRADSKÁ, S.; FERLING, I.; HILLMANN, F. Exploring Virulence Determinants of Filamentous Fungal Pathogens through Interactions with Soil Amoebae. **Frontiers in Cellular and Infection Microbiology**, v. 7, 5 dez. 2017.

ODDS, F. C. et al. Molecular Phylogenetics of *Candida albicans*. **Eukaryotic Cell**, v. 6, n. 6, p. 1041–1052, 6 abr. 2007.

OLIVER, J. L.; MARÍN, A. A relationship between GC content and coding-sequence length. **Journal of Molecular Evolution**, v. 43, n. 3, p. 216–223, set. 1996.

OPULENTE, D. A. et al. Pathogenic budding yeasts isolated outside of clinical settings. **FEMS Yeast Research**, v. 19, n. 3, 24 abr. 2019.

ØROM, U. A. et al. Long noncoding RNAs with Enhancer-like Function in Human Cells. **Cell**, v. 143, n. 1, p. 46–58, out. 2010.

PACHNIS, V.; BELAYEW, A.; TILGHMAN, S. M. Locus unlinked to alpha-fetoprotein under the control of the murine raf and Rif genes. **Proceedings of the National Academy of Sciences**, v. 81, n. 17, p. 5523–5527, 1 set. 1984.

PALAZZO, A. F.; LEE, E. S. Noncoding RNA: what is functional and what is junk? **Frontiers in Genetics**, v. 6, 26 jan. 2015.

PALMER, J.; STAJICH, J. **Funannotate: eukaryotic genome annotation pipeline**. Acesso em: 17 fev. 2023.

PANT, H. et al. Fungal-Specific Humoral Response in Eosinophilic Mucus Chronic Rhinosinusitis. **The Laryngoscope**, v. 115, n. 4, p. 601–606, abr. 2005.

PARKINSON, J.; BLAXTER, M. Expressed Sequence Tags: An Overview. **Methods in Molecular Biology**, p. 1–12, 2009.

PASTOR, F. J.; GUARRO, J. *Alternaria* infections: laboratory diagnosis and relevant clinical features. **Clinical Microbiology and Infection**, v. 14, n. 8, p. 734–746, ago. 2008.

PATIL, D. P. et al. m6A RNA methylation promotes XIST-mediated transcriptional repression. **Nature**, v. 537, n. 7620, p. 369–373, set. 2016.

PEÇANHA, P. M. et al. Paracoccidioidomycosis: What We Know and What Is New in Epidemiology, Diagnosis, and Treatment. **Journal of Fungi**, v. 8, n. 10, p. 1098, 18 out. 2022.

PEREIRA, T. C. Introdução ao universo dos noncoding RNAs. **repositorio.usp.br**, 2017.

PERFECT, J. R.; BICANIC, T. Cryptococcosis diagnosis and treatment: What do we know now. **Fungal Genetics and Biology**, v. 78, p. 49–54, maio 2015.

PERFECT, J. R.; CASADEVALL, A. Cryptococcosis. **Infectious Disease Clinics of North America**, v. 16, n. 4, p. 837–874, dez. 2002.

PINKNEY, H. R.; WRIGHT, B. M.; DIERMEIER, S. D. The lncRNA Toolkit: Databases and *In Silico* Tools for lncRNA Analysis. **Noncoding RNA**, v. 6, n. 4, p. 49, 16 dez. 2020.

POLISENO, L. et al. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. **Nature**, v. 465, n. 7301, p. 1033–1038, 1 jun. 2010.

PORTOSO, M. et al. PRC 2 is dispensable for *HOTAIR* -mediated transcriptional repression. **The EMBO Journal**, v. 36, n. 8, p. 981–994, 6 fev. 2017.

POSTEPSKA-IGIELSKA, A. et al. LncRNA Khps1 Regulates Expression of the Proto-oncogene SPHK1 via Triplex-Mediated Changes in Chromatin Structure. **Molecular Cell**, v. 60, n. 4, p. 626–636, nov. 2015.

PRAY, L. Discovery of DNA structure and function: Watson and Crick. **Nature Education**, v. 1, n. 1, p. 100, 2008.

QI, P.; ZHOU, X.; DU, X. Circulating long noncoding RNAs in cancer: current status and future perspectives. **Molecular Cancer**, v. 15, n. 1, 17 maio 2016.

QIAN, X. et al. Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. **Trends in Biochemical Sciences**, v. 44, n. 1, p. 33–52, jan. 2019.

QUINLAN, A. R.; HALL, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. **Bioinformatics**, v. 26, n. 6, p. 841–842, 28 jan. 2010.

QUINN, J. J.; CHANG, H. Y. Unique features of long noncoding RNA biogenesis and function. **Nature Reviews Genetics**, v. 17, n. 1, p. 47–62, 15 dez. 2015.

RAUTHAN, P.; SHARMA, D. C. Mucormycosis: Pathogenesis, Diagnosis, and Management. **Asian Journal of Pharmaceutical Research and Development**, v. 9, n. 3, p. 144–153, 15 jun. 2021.

RAWLINGS, N. D.; BARRETT, A. J.; BATEMAN, A. MEROPS: the peptidase database. **Nucleic Acids Research**, v. 38, n. suppl_1, p. D227–D233, 5 nov. 2009.

RESTREPO, A.; RESTREPO, M. C.; PÉREZ, M. Paracoccidioidomicosis: una revisión actualizada. **Revista de la Facultad de Medicina**, 64(2), 307-319. 2016.

RINN, J. L. et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. **Cell**, v. 129, n. 7, p. 1311–1323, jun. 2007.

RINN, J. L.; CHANG, H. Y. Genome Regulation by Long noncoding RNAs. **Annual Review of Biochemistry**, v. 81, n. 1, p. 145–166, 7 jul. 2012.

RODRIGUES, A. M.; DE HOOG, G. S.; DE CAMARGO, Z. P. *Sporothrix* Species Causing Outbreaks in Animals and Humans Driven by Animal–Animal Transmission. **PLoS Pathogens**, v. 12, n. 7, 14 jul. 2016.

RODRIGUES, M. L. et al. Vesicular Polysaccharide Export *in Cryptococcus neoformans* Is a Eukaryotic Solution to the Problem of Fungal Trans-Cell Wall Transport. **Eukaryotic Cell**, v. 6, n. 1, p. 48–59, 17 nov. 2006.

ROKAS, A. Evolution of the human pathogenic lifestyle in fungi. **Nature Microbiology**, v. 7, n. 5, p. 607–619, 1 maio 2022.

ROMANO, G. et al. Small noncoding RNA and cancer. **Carcinogenesis**, v. 38, n. 5, p. 485–491, 26 abr. 2017.

ROMO, J. A.; KUMAMOTO, C. A. On Commensalism of *Candida*. **Journal of Fungi**, v. 6, n. 1, 17 jan. 2020.

S, OHNO. So much "junk" DNA in our genome. In "Evolution of Genetic Systems". **Brookhaven Symposium in Biology**, v. 23, p. 366–370, 1972.

SABELLA, E. et al. Arthrinium phaeospermum isolated from Tuber borchii ascomata: the first evidence for a “Mycorrhization Helper Fungus”?. **Mycological Progress**, v. 14, n. 8, 19 jul. 2015.

SAMADDAR, A.; SHARMA, A. Emergomycosis, an Emerging Systemic Mycosis in Immunocompromised Patients: Current Trends and Future Prospects. **Frontiers in Medicine**, v. 8, 23 abr. 2021.

SANBONMATSU, K. Y. Towards structural classification of long noncoding RNAs. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, v. 1859, n. 1, p. 41–45, jan. 2016.

SANTOS, A. C. A. et al. Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. **Nature Communications**, v. 9, n. 1, 7 jun. 2018.

SARAVANAKUMAR, K. et al. Bioinformatics strategies for studying the molecular mechanisms of fungal extracellular vesicles with a focus on infection and immune responses. **Briefings in Bioinformatics**, v. 23, n. 4, 6 jul. 2022.

SARMA, K. et al. ATRX Directs Binding of PRC2 to Xist RNA and Polycomb Targets. **Cell**, v. 159, n. 4, p. 869–883, nov. 2014.

SATOH, K. et al. Isolation and identification of*Rhizomucor pusillus*from pleural zygomycosis in an immunocompetent patient. **Medical Mycology**, v. 47, n. 8, p. 869–873, dez. 2009.

SCHULER, G. D. et al. A Gene Map of the Human Genome. **Science**, v. 274, n. 5287, p. 540–546, 25 out. 1996.

SEGRELLES-CALVO, G.; DE S ARAÚJO, G. R.; FRASES, S. Systemic mycoses: a potential alert for complications in COVID-19 patients. **Future Microbiology**, v. 15, n. 14, p. 1405–1413, set. 2020.

SEPPEY, M.; MANNI, M.; ZDOBNOV, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. **Methods in Molecular Biology**, p. 227–245, 2019.

SEYEDMOUSAVI, S. et al. Fungal infections in animals: a patchwork of different situations. **Medical Mycology**, v. 56, n. suppl_1, p. S165–S187, 1 abr. 2018.

SHAFIK, A. et al. The emerging epitranscriptomics of long noncoding RNAs. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, v. 1859, n. 1, p. 59–70, jan. 2016.

SHARIFYNIA, S. et al. In vitro antifungal susceptibility profiles of *Candida albicans* complex isolated from patients with respiratory infections. **Acta Med Iran**, 54(6):376-381. 2016.

SHI, T.; GAO, G.; CAO, Y. Long noncoding RNAs as Novel Biomarkers Have a Promising Future in Cancer Diagnostics. **Disease Markers**, v. 2016, p. 1–10, 2016.

SKIADA, A.; PAVLEAS, I.; DROGARI-APIRANTHITOU, M. Epidemiology and Diagnosis of Mucormycosis: An Update. **Journal of Fungi**, v. 6, n. 4, p. 265, 2 nov. 2020.

ST LAURENT, G. et al. VlincRNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. **Genome Biology**, v. 14, n. 7, p. R73, 22 jul. 2013.

ST. LAURENT, G.; WAHLESTEDT, C.; KAPRANOV, P. The Landscape of long noncoding RNA classification. **Trends in Genetics**, v. 31, n. 5, p. 239–251, maio 2015.

STARK, B. C. et al. Ribonuclease P: an enzyme with an essential RNA component. **Proceedings of the National Academy of Sciences of the United States of America**, v. 75, n. 8, p. 3717–3721, 1 ago. 1978.

STARK, B. C. et al. Ribonuclease P: an enzyme with an essential RNA component. **Proceedings of the National Academy of Sciences of the United States of America**, v. 75, n. 8, p. 3717–3721, 1 ago. 1978.

STARK, P. C. et al. Fungal Levels in the Home and Allergic Rhinitis by 5 Years of Age. **Environmental Health Perspectives**, v. 113, n. 10, p. 1405–1409, out. 2005.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nature Reviews Genetics**, v. 20, n. 11, p. 631–656, 24 jul. 2019.

SUN, L. et al. Utilizing sequence intrinsic composition to classify protein-coding and long noncoding transcripts. **Nucleic Acids Research**, v. 41, n. 17, p. e166–e166, 27 jul. 2013.

SWIFT, A.; HEALE, R.; TWYCROSS, A. What are sensitivity and specificity? **Evidence Based Nursing**, v. 23, n. 1, p. ebnurs-2019-103225, 12 nov. 2019.

TAFT, R. J. et al. Evolution, biogenesis and function of promoter-associated RNAs. **Cell Cycle**, v. 8, n. 15, p. 2332–2338, ago. 2009.

TANI, H. et al. Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. **Genome Research**, v. 22, n. 5, p. 947–956, 1 maio 2012.

TARIQ, T. et al. Spontaneous fungal peritonitis: Micro-organisms, management and mortality in liver cirrhosis-A systematic review. **World Journal of Hepatology**, v. 11, n. 7, p. 596–606, 27 jul. 2019.

TAVARES-FERREIRA, D. et al. Spatial transcriptomics of dorsal root ganglia identifies molecular signatures of human nociceptors. **Science Translational Medicine**, v. 14, n. 632, 16 fev. 2022.

TAYLOR, J. W.; BERBEE, M. L. Fungal Molecular Evolution: Gene Trees and Geologic Time. **Systematics and Evolution**, p. 229–245, 2001.

TEKAIA, F.; LATGÉ, J.-P. *Aspergillus fumigatus*: saprophyte or pathogen? **Current Opinion in Microbiology**, v. 8, n. 4, p. 385–392, ago. 2005.

THE ENCODE PROJECT CONSORTIUM. An integrated encyclopedia of DNA elements in the human genome. **Nature**, v. 489, n. 7414, p. 57–74, set. 2012.

TIETZ, H. J. et al. Phenotypic and genotypic characterization of unusual vaginal isolates of *Candida albicans* from Africa. **Journal of clinical microbiology**, v. 33, n. 9, p. 2462–2465, 1995.

TIETZ, H.-J. et al. *Candida africana* sp. nov., a new human pathogen or a variant of *Candida albicans*? **Mycoses**, v. 44, n. 11-12, p. 437–445, dez. 2001.

TILL, P.; MACH, R. L.; MACH-AIGNER, A. R. A current view on long noncoding RNAs in yeast and filamentous fungi. **Applied Microbiology and Biotechnology**, v. 102, n. 17, p. 7319–7331, 4 jul. 2018.

TIRADO-SÁNCHEZ, A.; GONZÁLEZ, G. M.; BONIFAZ, A. Endemic mycoses: epidemiology and diagnostic strategies. **Expert Review of Anti-infective Therapy**, v. 18, n. 11, p. 1105–1117, 20 jul. 2020.

TONG, X.; LIU, S. CPPred: coding potential prediction based on the global description of RNA sequence. **Nucleic Acids Research**, v. 47, n. 8, p. e43–e43, 7 maio 2019.

TRABELSI, H. et al. Fungemia caused by *Yarrowia lipolytica*. **Mycopathologia**, v. 179, n. 5-6, p. 437–445, 1 jun. 2015.

TRAPNELL, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, v. 7, n. 3, p. 562–578, 1 mar. 2012.

TRIPATHI, V. et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. **Molecular Cell**, v. 39, n. 6, p. 925–938, 24 set. 2010.

ULITSKY, I. Evolution to the rescue: using comparative genomics to understand long noncoding RNAs. **Nature Reviews Genetics**, v. 17, n. 10, p. 601–614, 30 ago. 2016.

VAN HEESCH, S. et al. Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. **Genome Biology**, v. 15, n. 1, p. R6, 7 jan. 2014.

VAN WERVEN, FOLKERT J. et al. Transcription of Two Long Noncoding RNAs Mediates Mating-Type Control of Gametogenesis in Budding Yeast. **Cell**, v. 150, n. 6, p. 1170–1181, set. 2012.

VANCE, K. W. et al. The long noncoding RNA Paupar regulates the expression of both local and distal genes. **The EMBO Journal**, v. 33, n. 4, p. 296–311, 1 fev. 2014.

VANCE, K. W.; PONTING, C. P. Transcriptional regulatory functions of nuclear long noncoding RNAs. **Trends in Genetics**, v. 30, n. 8, p. 348–355, ago. 2014.

VARONA, S.; LAVÍN, J. L.; OGUIZA, J. A. Secretomes of medically important fungi reflect morphological and phylogenetic diversity. **Fungal Biology**, v. 124, n. 11, p. 915–923, 1 nov. 2020.

VILLEGAS, V.; ZAPHIROPOULOS, P. Neighboring Gene Regulation by Antisense Long noncoding RNAs. **International Journal of Molecular Sciences**, v. 16, n. 2, p. 3251–3266, 3 fev. 2015.

VOS, C. et al. Arbuscular mycorrhizal fungi reduce root-knot nematode penetration through altered root exudation of their host. **Plant and Soil**, v. 354, n. 1-2, p. 335–345, 29 nov. 2011.

WALKER, D. H.; MCGINNIS, M. R. Diseases Caused by Fungi. **Pathobiology of Human Disease**, p. 217–221, 2014.

WANG, KEVIN C.; CHANG, HOWARD Y. Molecular Mechanisms of Long noncoding RNAs. **Molecular Cell**, v. 43, n. 6, p. 904–914, set. 2011.

WANG, L. et al. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. **Nucleic Acids Research**, v. 41, n. 6, p. e74–e74, 17 jan. 2013.

WANG, L. F. et al. Infectious disease in an era of global change. **Nature Reviews Microbiology**, v. 20, n. 20, 13 out. 2022.

WANG, M. et al. Comparative pathogenicity of opportunistic black yeasts in *Aureobasidium*. **Mycoses**, 19 jul. 2019.

WANG, S.-B.; LI, R.-Y.; YU, J. Identification and susceptibility of *Rhizomucor spp*. isolated from patients with cutaneous zygomycosis in China. **Medical Mycology**, v. 49, n. 8, p. 799–805, 1 nov. 2011.

WASHIETL, S.; KELLIS, M.; GARBER, M. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. **Genome Research**, v. 24, n. 4, p. 616–628, 15 jan. 2014.

WICKES, B. L. Rhizopus oryzae: Genetic Secrets of an Emerging Human Pathogen. **Soil Biology**, p. 255–281, 2013.

WILUSZ, J. E. et al. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. **Genes & Development**, v. 26, n. 21, p. 2392–2407, 16 out. 2012.

WUCHER, V. et al. FEELnc: a tool for long noncoding RNA annotation and its application to the dog transcriptome. **Nucleic Acids Research**, v. 45, n. 8, p. e57, 5 maio 2017.

WUTZ, A.; RASMUSSEN, T. P.; JAENISCH, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. **Nature Genetics**, v. 30, n. 2, p. 167–174, 1 fev. 2002.

XIA, Z.-K.; WANG, W.-L.; YANG, R.-Y. Slowly Progressive Cutaneous, Rhinofacial, and Pulmonary Mucormycosis Caused by Mucor irregularis in an Immunocompetent Woman. **Clinical Infectious Diseases**, v. 56, n. 7, p. 993–995, 14 dez. 2012.

XU, X. et al. A systematic review of computational methods for predicting long noncoding RNAs. **Briefings in functional genomics**, v. 20, n. 3, p. 162-173, 2021.

YAMAMOTO, N. et al. Particle-size distributions and seasonal diversity of allergenic and pathogenic fungi in outdoor air. **The ISME Journal**, v. 6, n. 10, p. 1801–1811, 5 abr. 2012.

YANG, H. et al. Long noncoding RNA MIR31HG exhibits oncogenic property in pancreatic ductal adenocarcinoma and is negatively regulated by miR-193b. **Oncogene**, v. 35, n. 28, p. 3647–3657, 14 jul. 2016.

YAO, J. et al. A new tumor suppressor LncRNA ADAMTS9-AS2 is regulated by DNMT1 and inhibits migration of glioma cells. **Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine**, v. 35, n. 8, p. 7935–7944, 1 ago. 2014.

YAZDANPANAH, A.; KHAITHIR, T. M. N. Issues in Identifying Germ Tube Positive Yeasts by Conventional Methods. **Journal of Clinical Laboratory Analysis**, v. 28, n. 1, p. 1–9, 27 dez. 2013.

ZDOBNOV, E. M.; APWEILER, R. InterProScan - an integration platform for the signature-recognition methods in InterPro. **Bioinformatics**, v. 17, n. 9, p. 847–848, 1 set. 2001.

ZHANG, K. et al. Circulating lncRNA H19 in plasma as a novel biomarker for breast cancer. **Cancer Biomarkers: Section A of Disease Markers**, v. 17, n. 2, p. 187–194, 24 jun. 2016.

ZHANG, S.; LI, R.; YU, J. Drug Combinations against *Mucor irregularis In Vitro*. **Antimicrobial Agents and Chemotherapy**, v. 57, n. 7, p. 3395–3397, jul. 2013.

ZHENG R.Y.; CHEN C.Q. A non-thermophilic Rhizomucor causing human primary cutaneous mucormycosis. **Mycosystema**.; 4: 45–57. 1991.

ZHOU, N. et al. Long noncoding RNA CCAT2 functions as an oncogene in hepatocellular carcinoma, regulating cellular proliferation, migration and apoptosis. **Oncology letters**, v. 12, n. 1, p. 132-138. 2016.

ZHU, Y. X. et al. Multilocus sequence typing analysis of *Candida africana* from vulvovaginal candidiasis. **BMC infectious diseases**, v. 19, n. 1, p. 1-8. 2019.