



SALÃO DE INICIAÇÃO CIENTÍFICA
XXVIII SIC

paz no plural



Evento	Salão UFRGS 2016: SIC - XXVIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2016
Local	Campus do Vale - UFRGS
Título	Análise Transcritômica de <i>Cryptococcus gattii</i>
Autor	RODRIGO SILVA ARAUJO STREIT
Orientador	CHARLEY CHRISTIAN STAATS

Análise Transcritômica de *Cryptococcus gattii*

Rodrigo Silva Araujo Streit
Prof. Dr. Charley Christian Staats
Universidade Federal do Rio Grande do Sul

A levedura basidiomicética *Cryptococcus gattii* é um dos agentes etiológicos da criptococose, doença que pode acometer pele, pulmões e sistema nervoso central. As sequências genômica e transcritômica da linhagem R265 de *C. gattii* foram disponibilizadas pelo BroadInstitute (*Cryptococcus gattii* Sequencing Project, Broad Institute of Harvard and MIT) a partir 2009, passando por suas últimas atualizações em 2015. Apesar de sua recente atualização, os métodos empregados para a obtenção de modelos gênicos mantiveram-se os mesmos ao longo das versões, utilizando softwares de anotação automática como Augustus, FGESH e outros preditores *de novo*, metodologia à qual a literatura vem demonstrado haver erros inerentes e limitações em suas predições, especialmente em organismos onde é encontrada justaposição de genes. Ao passo que erros vêm sendo constatados na metodologia clássica, novos métodos de análise que utilizam dados de sequenciamento de última geração vêm sendo desenvolvidos e têm mostrando extrema eficiência e robustez na predição transcricional, de forma que têm sido constantes as correções e complementações de informações prévias de diversos organismos. Assim, o presente estudo tem por objetivo a análise e correção do transcrito da linhagem R265 de *C. gattii*, baseando-se em metodologias de RNA-seq. Para tanto, foram utilizadas 4 bibliotecas de *reads*, provenientes de sequenciamento de RNA de diferentes condições de cultivo de *C. gattii* (YNB acrescido do quelante de cobre BCS por 2h, YNB acrescido do quelante de zinco TPEN por 2h, lavado bronquioalveolar de camundongos infectados por 10 dias) através da plataforma IlluminaHiSeq. Inicialmente, foi realizado o alinhamento dos *reads* contra a sequência genômica disponibilizada pelo Broad Institute. Dos 160411087 *reads* presentes no agrupamento das 4 bibliotecas, houve o alinhamento de 139232388 (86.8%) *reads* contra o genoma. Os dados do alinhamento foram então utilizados para uma predição inicial de transcritos pelo software Cufflinks. Em seguida, os transcritos gerados foram submetidos ao software CodingQuarry, que gerou uma nova versão de transcritos a partir da combinação dos dados de RNAseq com algoritmos de predição automática. Atualmente, está sendo feita a revisão e correção manual dos modelos gerados, utilizando o software IGV, tendo por base os resultados do alinhamento e observando as regras de delimitação e estruturação de transcritos. Dos 6456 modelos que se fazem presentes na anotação do Broad Institute, 2468 já foram revisados, sendo que destes, 1001 (40.5%) apresentaram erros na anotação e 191 (7.7%) foram excluídos das análises por dados do alinhamento insatisfatórios para sua região, sendo aceitos os modelos prévios como corretos para esses casos. Por fim, para a confirmação dos resultados obtidos *in silico*, alguns transcritos serão selecionados aleatoriamente para a confirmação de predição através de RT-PCR.