



**Universidade:
presente!**

UFRGS
PROPEAQ



XXXI SIC

21. 25. OUTUBRO • CAMPUS DO VALE

Evento	Salão UFRGS 2019: SIC - XXXI SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2019
Local	Campus do Vale - UFRGS
Título	Influência do índice de confiança, pipeline de clusterização e classificadores em análises de metabarcoding
Autor	LAURA BEZERRA COUTINHO
Orientador	AFONSO LUIS BARTH

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

HOSPITAL DE CLÍNICAS DE PORTO ALEGRE

INFLUÊNCIA DO ÍNDICE DE CONFIANÇA, *PIPELINE* DE CLUSTERIZAÇÃO E CLASSIFICADORES EM ANÁLISES DE *METABARCODING*

Autora: Laura Bezerra Coutinho

Orientador: Afonso Luís Barth

As complexas comunidades microbiológicas presentes nos mais diversos ambientes - inclusive o corpo humano - são o que chamamos de microbiota e a compreensão de sua composição e funcionamento é alvo de grande interesse científico, tendo em vista a gama de interações e funções nas quais ela está envolvida. O *metabarcoding* a partir do 16S rDNA é, hoje, um método amplamente aplicado em estudos de composição da microbiota. Nesse contexto, as análises bioinformáticas dos dados gerados enfrentam o grande desafio da garantia de qualidade e reprodutibilidade dos resultados. *Pipelines* de análise padronizadas são uma alternativa para essa questão, no entanto, o uso de parâmetros apropriados podem gerar impacto direto nos resultados obtidos. O objetivo deste trabalho foi, portanto, elucidar a influência de parâmetros de análise nos resultados de análises de dados de 16S rDNA referentes à microbiota intestinal humana, utilizando como referência o *pipeline* do BMP - *Brazilian Microbiome Project*. Para isso foram comparados os resultados de diferentes combinações entre índices de confiabilidade, classificadores e *pipelines* para clusterização. Nossos resultados indicam diferenças claras entre a aplicação de diferentes parâmetros ao *pipeline*, gerando diferentes efeitos na quantidade de taxa identificados, de OTUs classificadas e na precisão da classificação. Resultados de combinações de classificadores e índices de confiança apresentam variações entre os dois *pipelines* de clusterização, no USEARCH havendo pouca diferenciação com a alteração dos classificadores e no VSEARCH apresentando maiores disparidades - com destaque para o Mothur, cujos resultados de número de taxa identificados e OTUs classificadas foram acima dos demais, não respondendo inclusive ao aumento do índice de confiança. Destacamos ainda o papel da remoção de sequências quiméricas na qualidade dos resultados. Com isso, salientamos a importância da inclusão em estudos de microbiota de detalhes dos parâmetros e métodos aplicados, garantindo a validação, qualidade e reprodutibilidade do resultado. Compreender os *pipelines* aplicados e os efeitos de seus parâmetros é essencial. Testar diferentes *pipelines* e variações dos parâmetros é recomendável.