



NOVAS FERRAMENTAS PARA ANÁLISE DA ALFA DIVERSIDADE EM DADOS COMPOSICIONAIS

Susiane Paloma Haugg, discente de graduação, Universidade Federal do Pampa, Campus São Gabriel¹

Priscila Thiago Dobbler, discente de doutorado, Universidade Federal do Pampa, Campus São Gabriel

Luiz Fernando Würdig Roesch, docente, Universidade Federal do Pampa

susianehaugg.aluno@unipampa.edu.br¹

Em estudos de microbiomas, medidas de alfa diversidade são utilizadas para analisar e contrapor amostras. Em pesquisas onde se utiliza o gene marcador 16S, medidas de alfa diversidade geralmente utilizadas reconhecem as leituras de sequência como contagem de táxons, todavia essas contagens não são absolutas, uma vez que o método gera dados composicionais. O Índice de diversidade de Shannon e o Teste de Simpson são largamente utilizados nos dados de contagem de espécies, sendo estes calculados com base na frequência de cada táxon. Entretanto, converter as contagens em abundância relativa ocasiona em distorções dos dados. Uma alternativa para a retificação das análises de dados composicionais é a transformação dos dados através da razão logarítmica central de Aitchison (clr). Contudo, utilizar os dados convertidos em clr para calcular a alfa diversidade através dos índices mencionados seria incorreto, visto que a transformação compila dados negativos para taxa de baixa abundância e as medidas citadas não admitem valores negativos em suas fórmulas, além dos mesmos não poderem ser utilizados no cálculo de frequência de táxon. As pesquisas em microbiomas utilizando o gene marcador 16S aumentaram nos últimos anos, mas a ausência de uma medida de alfa diversidade padrão para dados composicionais impede um vasto uso dessa opção. Uma nova ferramenta para a análise de alfa diversidade foi desenvolvida baseada na Área abaixo da curva de abundância transformada e o objetivo principal deste trabalho foi testar a viabilidade de uso dessa ferramenta em estudos de microbiomas. Foram pré selecionados 10 conjuntos de dados empíricos envolvendo o microbioma de diferentes ambientes. Todos os conjuntos possuíam pelo menos dois tratamentos com repetições. O software estatístico R foi utilizado para realizar as análises. A alfa diversidade foi estimada através do Índice de diversidade de Shannon e também da nova ferramenta. O novo índice se mostrou capaz de produzir resultados similares aos que foram obtidos utilizando o índice de Shannon, ambos forneceram valores significativos ($p < 0,5$) para 5 dos conjuntos de dados testados. Ambos os índices testados permitiram a mesma conclusão biológica, no entanto um diferencial da nova ferramenta é a sua possibilidade de detectar a biosfera abundante e a biosfera rara. Por conseguinte, a nova ferramenta não se mostrou melhor que o índice de Shannon, bem como não se mostrou inferior. Uma perspectiva futura para a continuidade das análises é realizar os testes a partir de dados simulados, onde se conhecerá os valores que deverão ser obtidos.

Agradecimentos: FAPERGS, UNIPAMPA

Palavras-chave: Microbiomas; Bioinformática; Alfa Diversidade; Microrganismos.