

FERRAMENTAS PARA VISUALIZAÇÃO E ANÁLISE DE DADOS DE METAGENOMAS

Luciano Antonio Digiampietri, Vivian Mayumi Yamassaki Pereira, Camila Izidio Costa

Escola de Artes, Ciências e Humanidades, USP
{digiampietri, vivian.pereira, camila.izidio.costa}@usp.br

Objetivos

Este trabalho tem como objetivo apresentar um estudo e o desenvolvimento de ferramentas para analisar e comparar dados de metagenomas. Metagenômica é uma forma de se estudar comunidades microbianas em nichos ecológicos específicos, por exemplo, o estômago de animais, o fundo do mar, as nuvens e o interior de cavernas. Neste projeto, estamos inicialmente analisando dados de metagenomas da compostagem que ocorre no Zoológico Municipal de São Paulo e as ferramentas desenvolvidas visam a comparar a evolução microbacteriana da compostagem ao longo do tempo, de forma a identificar correlações entre o crescimento (ou decaimento) da variedade e quantidade de espécies, bem como identificar genes ou conjuntos de genes de interesse.

Métodos/Procedimentos

A metodologia foi iniciada com a identificação, junto aos especialistas do domínio, das principais dificuldades quanto à análise e comparação de dados de metagenomas. Com uma lista de desafios em mãos, o passo seguinte consistiu da identificação de soluções possíveis na literatura e de ferramentas que pudessem auxiliar na resolução destes desafios. Com base neste estudo, identificou-se que a comparação entre a quantidade e diversidade de espécies em metagenomas é uma tarefa importante que ainda apresenta alguns desafios ligados a quantidade de dados e precisão dos resultados. A partir daí foram desenvolvidas e estão sendo testadas ferramentas para comparar os dados de metagenomas.

Resultados

Até o momento, foram desenvolvidas duas ferramentas para a comparação de dados

metagenômicos: uma para a organização e contagem das sequências de DNA coletadas de acordo com a classificação taxonômica utilizando dados de duas ferramentas diferentes (uma mais precisa, porém que só consegue identificar a taxonomia de uma pequena quantidade dos dados e outra mais abrangente). A Figura 1 apresenta a distribuição das classes mais abundantes em 5 coletas realizadas na mesma composteira.

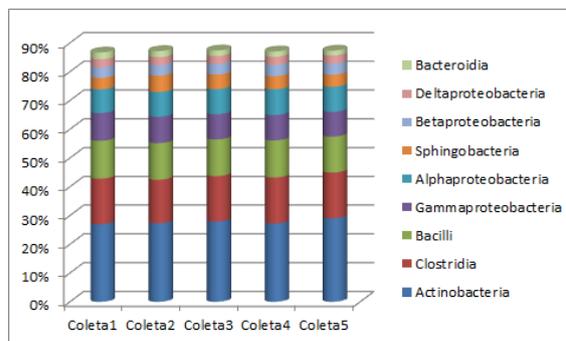


Figura 1: Distribuição das Classes nas Coletas

A segunda ferramenta desenvolvida correlaciona os dados taxonômicos para identificar relações positivas e negativas entre as frequências das espécies identificadas.

Conclusões

Este projeto está desenvolvendo ferramentas para auxiliar na análise de dados de metagenomas focando na distribuição taxonômica das prováveis espécies com base no DNA sequenciado. Em particular, na comparação entre diferentes coletas de um mesmo ambiente e na correlação entre as frequências das espécies encontradas.

Agradecimentos

Agradecemos ao Programa de Educação Tutorial (MEC/SESu), a FAPESP e a Pró-Reitoria de Pesquisa da USP que financiaram esse trabalho.