



Evento	Salão UFRGS 2014: SIC - XXVI SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2014
Local	Porto Alegre
Título	Uma nova visão sobre o DNA
Autor	CAUÊ PETRY DA SILVA STEQUES
Orientador	DAVID RENATO CARRETA DOMINGUEZ

Introdução:

A entropia fornece uma previsão quantitativa da desordem. O sentido natural é de aumento de desordem. Pode-se calcular a entropia (S) utilizando probabilidade, para isto utiliza-se a formula $S = -\sum P(x) \log \left(\frac{1}{P(x)} \right)$, onde P é a probabilidade.

No método utilizado, as probabilidades utilizadas foram as de aparição cada base nitrogenada do Vírus. Ainda é possível comparar a diferença entre dois vírus. Para isto utilizamos a Divergência de Kullback-Leibler (que serve não só para vírus) que é dada por $D_{KL} = -\sum P(x) \log \frac{Q(x)}{P(x)}$, onde P(x) e Q(x) são distribuições de probabilidade. Vale lembrar que a divergência de Kullback-Leibler não é simétrica, ou seja, $D_{KL}(P, Q) \neq D_{KL}(Q, P)$, ainda se $P = Q$ resulta em uma divergência nula.

Método:

Primeiramente se fez um programa na linguagem C a fim de calcular as probabilidades de cada base nitrogenada de um vírus para assim calcular primeiramente a entropia. Após fazermos algumas análises baseadas na entropia, adicionamos ao programa a possibilidade de calcular a Divergência de Kullback-Leibler. Após realizar o programa, escolhemos um vírus a partir de um banco de dados de genoma, com a intenção que houvesse a referencia da localidade dos vírus, tal que um mesmo vírus tivesse cepas em diferentes localidades, assim poder-se-ia comparar as divergências com a distância real das origens das amostras.

Utilizamos o programa desenvolvido no genoma das diversas amostras do vírus. As amostras pertenciam à Taiwan (Hualian), Sri Lanka (desconhecida), China (Guanxi), Taiwan (desconhecida), China (Hainan) e Singapura (desconhecida). Analisamos primeiramente em função da entropia. Também repetimos o método com outras amostras. Comparamos não somente a um mesmo vírus em diferentes localizações, mas também Fungos com Vírus e Bactérias.

Após isto, desenvolveu-se um programa para calcular a divergência dos aminoácidos (triplets) de cada vírus, este programa ainda está em desenvolvimento.

Conclusões:

Com a análise que fizemos em relação às divergências das bases nitrogenadas podemos concluir que parece ter uma relação entre a Distancia de Kullback-Leibler e a distância, mas em alguns casos divergiu um pouco. Quanto à relação Fungo, Vírus, Bactéria, pelas distâncias de Kullback-Leibler podemos dizer que o vírus é o que está mais distante de todos. O vírus está mais próximo à bactéria do que ao fungo, e a bactéria está mais próxima ao fungo do que ao vírus. As divergências entre as diferentes bactérias analisadas em alguns casos foram maiores do que entre espécies, então talvez os resultados não sejam tão bons. Para ter uma certeza maior devemos calcular a divergência levando em consideração não só as bases nitrogenadas de cada amostra, mas também os triplets.