Genômica Comparativa: do Alinhamento de Seqüências ao Enovelamento Proteíco

Ana Tereza Ribeiro de Vasconcelos

Laboratorio Nacional de Computação Científica, Dep de Matemática Aplicada e Computacional, Laboratório de Bioinformática, Petrópolis, Rio de Janeiro.

Diversos genomas completos encontram-se hoje disponíveis em bancos de dados de acesso público. Embora os procedimentos de seqüenciamento de genes e de genomas tenham tornado-se razoavelmente rotineiros, o mesmo não pode ser dito de sua anotação e interpretação. Mecanismos de mineração de dados estão sendo intensamente aplicados sobre os dados obtidos e têm por objetivo compreender e organizar a crescente quantidade de informações biológicas hoje disponível. Um dos mecanismos é a genômica comparativa onde muitos aspectos devem ser observados como por exemplo similaridade entre seqüências, localização gênica, comprimento e número de regiões codificantes, quantidade e perfis de regiões não-codificantes, regiões altamente. Embora a comparação entre seqüências seja uma prática comum em bioinformática, observa-se a existência de um generoso grau de degeneração informacional entre seqüências de nucleotídeos (e consequentemente na seqüência de aminoácidos de suas proteínas), que não necessariamente altera a função biológica de seu gene e produtos. Encontram-se na literatura exemplos de proteínas com baixo grau de similaridade seqüencial, que entretanto apresentam estrutura tridimensional protéica conservada assim como similaridade funcional. Assume-se portanto a premissa de que na natureza, a estrutura protéica é mais conservada que a informação de sua seqüência de aminoácidos. Portanto, o conhecimento da estrutura tridimensional de proteínas é relevante na determinação de funções biológicas de proteínas em organismos distintos, assim como no estudo de relações evolutivas entre estes.