

Conclusões

Neste trabalho, foi criada uma ferramenta para o provimento de dados ao BLAST de maneira eficiente, realizando um gerenciamento de *buffer*, controle dos dados e escalonamento dos processos, que leva em consideração características específicas do BLAST no acesso ao banco. A ferramenta foi denominada BioProvider. Seu desenvolvimento foi feito de maneira que funcionasse com diferentes versões do BLAST e fosse, ao mesmo tempo, facilmente extensível para prover futuramente soluções de bancos de dados a outras ferramentas de Bioinformática.

Na primeira etapa do trabalho, foram analisados os comportamentos de diferentes versões e implementações do BLAST, sobretudo a versão mais recente do NCBI BLAST e a versão do WU-BLAST com código aberto. Foram estudados também os formatos dos arquivos do banco de dados para criar estratégias de fornecimento dos arquivos de maneira genérica, de modo a atender às versões de BLAST estudadas.

O BioProvider foi implementado de maneira não-intrusiva no código do BLAST, o que foi possível devido à utilização de um *driver* para realizar a comunicação com os processos BLAST. Além de não ser necessário modificar o código do BLAST para usar a ferramenta, esta funciona com as 2 versões do BLAST estudadas. Para tornar a ferramenta mais eficiente, foram criadas estratégias de atendimento e escalonamento dos processos BLAST.

Nos testes realizados, foi possível observar diversas situações nas quais houve uma melhora de desempenho do BLAST utilizando-se o BioProvider. Foi possível verificar também a influência de alguns fatores no desempenho do BLAST com o BioProvider.

5.1.

Trabalhos Futuros

Em trabalhos futuros, o BioProvider pode ser estendido em diversos pontos. Estes são listados a seguir:

- Inclusão do fornecimento ao BLAST de arquivos de bancos de dados de nucleotídeos. Poderão ser usadas estratégias de gerenciamento de *buffer* e escalonamento de processos semelhantes às já criadas.
- Criação e análise de desempenho de outras estratégias de atendimento de processos e gerenciamento de *buffer* a serem usadas com o BLAST. Uma possibilidade é a criação de múltiplos anéis em memória, aos quais devem ser alocados processos com velocidades semelhantes, de modo a não permitir que processos lentos atrasem a execução de processos rápidos durante a leitura do anel.
- Utilização de outras técnicas de fornecimento do banco de dados ao BLAST, como a criação dinâmica do arquivo de índices a ser fornecido a cada processo. Isto tornaria desnecessária a divisão do banco de dados em blocos e a criação de diferentes instâncias dos arquivos de índices e anotações durante o pré-processamento do banco.
- Utilização do BioProvider para prover outras soluções de bancos de dados ao BLAST. Um trabalho a ser feito é a inclusão de técnicas de compactação e descompactação de arquivos ao BioProvider que possibilitem o armazenamento de dados em formatos diferentes dos exigidos pelos processos, traduzindo-os em tempo de execução. Deste modo, poderão ser estudadas técnicas que incluam ao mesmo tempo gerenciamento de *buffer* e compactação para tornar o BLAST mais eficiente. Isto possibilitaria também a continuação de alguns dos estudos feitos em Rosa et al. (2006).
- Extensão do BioProvider para prover soluções de bancos de dados para outros aplicativos de Bioinformática. Uma das extensões consiste em fornecer à ferramenta FASTA técnicas de gerência de *buffer* semelhantes às já estudadas. Por ser também uma ferramenta de comparação de

biossequências, a FASTA possui características semelhantes ao BLAST e poderá se beneficiar de técnicas de gerência de *buffer* parecidas. Outros aplicativos também podem se beneficiar do uso do BioProvider para diferentes finalidades, aproveitando-se das características de transparência que este fornece aos aplicativos no fornecimento de dados.