GERENCIAMENTO, ANÁLISE E DISPONIBILIZAÇÃO DA INFORMAÇÃO DO PROJETO DE MELHORAMENTO DE TRIGO, POR FERRAMENTAS DE BIOINFORMÁTICA

Bueno, E. F.[1]; Nhani Junior, A.[2]*

O Projeto Melhoramento Genético de Trigo para o Brasil, conduzido pela Embrapa Trigo, tem como objetivo a criação de novas cultivares resistentes a estresses bióticos e abióticos, elencando como principais objetos de estudo a ferrugem da folha, a giberela, o excesso de alumínio do solo e a seca. Com o avanço das técnicas em biologia molecular, ocorreu a geração e o acúmulo de um grande volume e diversidade de informação em gramíneas. A utilização de ferramentas de bioinformática tornou-se deste modo essencial para a análise, o armazenamento e a disponibilização destes dados. Um grande número de sites disponibilizam pela internet estas informações, como por exemplo o NCBI-GenBank, Gramene, GrainGenes, entre outros. Entretanto, a compilação destas informações é laboriosa. O GMOD (Generic Model Organism Database) é um projeto de desenvolvimento de módulos com ferramentas “open source” para manipulação e gerenciamento de dados. O módulo CHADO é um banco de dados relacional utilizado para armazenamento de informações de seqüências nucleotídicas, protéicas e análises de similaridade; o GBROWSE possibilita a visualização dos dados e relações entre eles (inseridos no CHADO) e CMAP é usado para a geração de mapas genéticos. Por atenderem, com pouca ou nenhuma modificação, as demandas do projeto de melhoramento, as ferramentas do GMOD foram instaladas no Ubuntu Linux, usando um tutorial validado pelo CNPTIA. Com o objetivo de facilitar a visualização das relações que ocorrem entre trigo e arroz, seqüências gênicas e genômicas de arroz relacionadas a genes de resistência aos estresses citados foram inseridas no banco de dados. As seqüências gênicas de arroz foram utilizadas na busca de seqüências similares em trigo e na tribo Triticeae. Os arquivos de alinhamento gerados foram submetidos a pré-processamento através de ferramentas do GMOD e também armazenadas no banco de dados CHADO. Os próximos passos serão configurar o GBROWSE para que sejam exibidas as similaridades entre seqüências de trigo e arroz e desenvolver um script para normalizar os arquivos de inserção quanto a erros de identificação de códigos específicos (vocabulário controlado).

[1] Acadêmico do curso de Sistemas de Informação, Faculdade Meridional. Bolsista Embrapa Trigo.
[2] Pesquisador da Embrapa Trigo, *orientador.