

# Aplicação Java Web para armazenamento de dados experimentais no projeto Rede Genômica Animal

Leonardo Barros Almeida de Ana<sup>1</sup>  
Roberto Hiroshi Higa<sup>2</sup>

O projeto “Rede Genômica Animal” faz parte da carteira de projetos do Macroprograma 1 do SEG - Sistema Embrapa de Gestão e tem por objetivo prospectar genes relacionados a manifestações de interesse econômico para pecuária brasileira, por meio de análise de expressão gênica e genotipagem por *Single Nucleotide Polymorphism* (SNP).

Logo no início do projeto, identificou-se a necessidade do desenvolvimento de uma aplicação para armazenamento tanto dos dados brutos, produzidos pelos experimentos expressão gênica e genotipagem por SNP, quanto dos resultados de análises desses dados.

Uma das maneiras de lidar com esse tipo de necessidade em projetos de pesquisa na área genômica consiste em adotar esquemas de banco de dados genéricos com tabelas preparadas para o armazenamento de informações genômicas. Uma das primeiras iniciativas nessa área é o esquema genérico CHADO (EMMERT et al., 2007), que foi desenvolvido pelo projeto GMOD (GMOD, 2010) e tem por objetivo desenvolver ferramentas para criar e administrar bancos de dados genômicos.

---

<sup>1</sup> Universidade Estadual de Campinas; [leonardo@cnptia.embrapa.br](mailto:leonardo@cnptia.embrapa.br)

<sup>2</sup> Embrapa Informática Agropecuária; [roberto@cnptia.embrapa.br](mailto:roberto@cnptia.embrapa.br)

O objetivo deste trabalho é desenvolver uma aplicação web para banco de dados genômicos, a partir do módulo Mage do GMOD para armazenagem dos dados experimentais brutos produzidos no âmbito do projeto “Rede Genômica Animal” e as suas respectivas análises.

O servidor de bancos de dados adotado foi o PostgreSQL 8.4.2 (PostgreSQL, 2010a) , sendo que para facilitar a manipulação das tabelas, via scripts SQL, é utilizado o cliente PGAdmin III (PostgreSQL, 2010b).

Na modelagem de dados, inicialmente, considerou-se o módulo MAGE do CHADO e, então, foram realizadas modificações no modelo de dados para compatibilizá-lo com as necessidades do projeto “Rede Genômica Animal”. O software DIA (The Gnome Project, 2010) foi utilizado para documentar o novo modelo de dados. A modelagem conceitual, lógica e física do banco de dados, realizada a partir do módulo Mage do CHADO resultou em um modelo consensual composto por 34 tabelas. A plataforma de desenvolvimento adotada consiste da linguagem de programação Java (ORACLE CORPORATION, 2010a), utilizada a partir do ambiente de desenvolvimento integrado - IDE Netbeans (ORACLE CORPORATION, 2010b).

Inicialmente, foi desenvolvida uma aplicação local, utilizando o Framework Java Hibernate (HIBERNATE, 2010) para mapeamento, comunicação e persistência dos objetos contidos na base de dados PostgreSql e os componentes do Java Swing para construção da interface com o usuário.

A versão local da aplicação contempla 4 tipos de usuário: *root*, administrador, analista de dados/resultados e comum. Apenas o usuário *root* possui privilégios para cadastrar novos projetos e atribuir a função de administrador a um usuário. Ele é o responsável por iniciar e encerrar um projeto dentro da aplicação. O usuário comum possui permissão apenas para consultar dados cadastrais dos projetos, tais como título, administrador, descrição, data de início e finalização do projeto, se este estiver encerrado. Já o usuário analista de dados e resultados é o responsável pelo resultado das análises realizadas a partir dos dados brutos. Seus privilégios compreendem a realização de *downloads* de

dados brutos e análises, e o cadastro de novo resultados de análise. Por fim, o usuário Administrador poderá realizar todas as funções dos demais usuários, cadastrar novos usuários, atribuir-lhes funções e fazer o *upload* de dados brutos.

Atualmente, encontra-se em desenvolvimento a versão web da aplicação, utilizando as tecnologias JSP – Java Server Pages e JSF - Java Server Faces e o servidor de aplicações Tomcat (APACHE SOFTWARE FOUNDATION, 2010).

## Referências

- APACHE SOFTWARE FOUNDATION. 2010. Disponível em: <<http://tomcat.apache.org>>. Acesso em: 14 jul. 2010.
- GMOD. 2010. Disponível em: <<http://gmod.org>>. Acesso em: 19 de mar. 2010.
- THE GNOME PROJECT. 2010. Disponível em: <<http://live.gnome.org/Dia>>. Acesso em: 4 de maio 2010.
- HIBERNATE. 2010. Disponível em: <<http://www.hibernate.org>>. Acesso em: 9 jun. 2010.
- MUNGALL, C.; EMMERT, D. B.; J. A Chado case study: ontology-based modular schema for representing genome-associated biological information. **Bioinformatics**, v. 23, n. 13, 337-346, 2007.
- ORACLE CORPORATION. 2010a. Disponível em: <http://www.java.com>. Acesso em: 30 mar. 2010.
- ORACLE CORPORATION. 2010b. Disponível em: <<http://netbeans.org>>. Acesso em: 5 abr. 2010.
- POSTGRESQL. 2010a. Disponível em: <<http://www.postgresql.org>>. Acesso em: 29 abr. 2010.
- POSTGRESQL. **pgAdmin**. 2010b. Disponível em: <<http://www.pgadmin.org>> Acesso em: 29 abr. 2010.