

Manual de uso do Sistema BDGF

*Empresa Brasileira de Pesquisa Agropecuária
Embrapa Informática Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento*

DOCUMENTOS 163

Manual de uso do Sistema BDGF

Fábio Danilo Vieira

Autor

*Embrapa Informática Agropecuária
Campinas, SP
2019*

Exemplares desta publicação podem ser adquiridos na:

Embrapa Informática Agropecuária

Av. Dr. André Tosello, 209 - Cidade Universitária
Campinas, SP, Brasil
CEP. 13083-886
Fone: (19) 3211-5700
www.embrapa.br
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações
da Unidade Responsável

Presidente

Stanley Robson de Medeiros Oliveira

Secretário-Executivo

Carla Cristiane Osawa

Membros

Adriana Farah Gonzalez; Carla Geovana do Nascimento Macário; Jayme Garcia Arnal Barbedo; Kleber Xavier Sampaio de Souza; Luiz Antonio Falaguasta Barbosa; Magda Cruciol; Paula Regina Kuser Falcão; Ricardo Augusto Dante; Sônia Ternes

Suplentes

Goran Nescic

Michel Eduardo Beleza Yamagishi

Supervisão editorial

Kleber Xavier Sampaio de Souza

Revisão de texto

Nadir Rodrigues Pereira

Normalização bibliográfica

Carla Cristiane Osawa

Projeto gráfico da coleção

Carlos Eduardo Felice Barbeiro

Editoração eletrônica

Felipe Prado Jaconi sob supervisão de Magda Cruciol

Foto da capa

Pexels

1ª edição

Versão digital (2019)

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte,
constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Informática Agropecuária

Vieira, Fábio Danilo,

Manual de uso do Sistema BDGF / Fábio Danilo Vieira. - Campinas : Embrapa Informática Agropecuária, 2019.

PDF (30 p.) : il. -- (Documentos / Embrapa Informática Agropecuária, ISSN 1677-9274 ; 163).

1. Genótipo. 2. Fenótipo. 3. Banco de Dados de Genótipos e Fenótipos. 4. Sistema BDGF. I. Título. II. Embrapa Informática Agropecuária. III. Série.

CDD (21. ed.) 005.15

Autor

Fábio Danilo Vieira

Tecnólogo em Processamento de dados, mestre em Engenharia Agrícola, analista da Embrapa Informática Agropecuária, Campinas, SP.

Apresentação

Analisar grandes volumes e variedades de dados é, atualmente, um grande desafio para muitas áreas da ciência. Uma dessas áreas é a biologia, que constantemente precisa se reinventar para analisar a grande quantidade de informações geradas por sequenciadores automáticos de DNA, capazes de gerar dados genômicos em grande escala.

Um desafio tão grande quanto estudar esses dados é armazená-los da forma correta. Para integrar esse volume de dados distintos, é primordial utilizar uma estrutura de armazenamento robusta, como um Sistema Gerenciador de Banco de Dados (SGBD). Assim, uma questão importante a considerar é o equilíbrio entre normalização e desempenho durante o estágio de modelagem do banco de dados, pois isso terá um impacto direto na usabilidade e na experiência do usuário.

Buscando resolver esse problema de armazenamento eficiente e consultas rápidas num grande volume de dados, este trabalho apresenta o sistema Banco de Dados de Genótipos e Fenótipos (BDGF). Seu modelo de dados permite a implementação do tipo *JavaScript Object Notation* (JSON) em campos de tabelas relacionadas a fenótipos e do tipo texto em campos da tabela genótipos. O BDGF foi projetado para dar suporte a projetos de criação de animais da Embrapa, mas pode ser facilmente ajustado para armazenar dados de diversas fontes, tais como dados de plantas. Além disso, o sistema inclui políticas de acesso e segurança para fenótipos, genótipos e pedigree dos animais.

Neste documento apresentamos um manual de uso das principais funcionalidades do sistema BDGF, seja para usuários atuais ou para pessoas interessadas em como trabalhar com esse *software web* para armazenamento e recuperação de dados fenotípicos e genotípicos de animais.

Silvia Maria Fonseca Silveira Massruhá

Chefe-geral

Embrapa Informática Agropecuária

Sumário

Introdução	9
Como utilizar o sistema BDGF	9
Considerações Finais	29
Referências	29

Introdução

Nos últimos anos, a utilização de tecnologias de genotipagem em larga escala de marcadores moleculares do tipo *Single Nucleotide Polymorphisms* (SNP) para estimar o perfil genômico de animais permitiu o desenvolvimento tanto de estudos de associação genótipo-fenótipo em escala genômica (do inglês *genome-wide association studies* (GWAS)) quanto a introdução da tecnologia de seleção genômica em programas de melhoramento genético. Entretanto, essa situação implica a necessidade de armazenamento de grande volume de dados de genotipagem, fenotipagem e pedigree de um grande número de animais, uma tendência que possivelmente aumentará nos próximos anos, dada a diminuição dos custos para produção dos dados experimentais (Caetano, 2009).

Devido ao volume de dados (centenas de milhares de animais fenotipados e, possivelmente, genotipados para centenas de milhares de marcadores do tipo SNPs), uma questão importante a se considerar no desenvolvimento de uma solução computacional é a adequabilidade da modelagem à aplicação desejada, pois esta terá impacto direto nos tempos das consultas nos bancos de dados em que essa informação estará armazenada.

Visando o armazenamento e a consulta eficiente desse alto volume de dados, o sistema Banco de Dados de Genótipos e Fenótipos (BDGF) foi desenvolvido utilizando um modelo de dados inicialmente proposto por Higa e Oliveira (2015), com o principal objetivo de fornecer armazenamento adequado e consultas eficientes aos dados de genótipos e fenótipos num banco de dados relacional (do inglês *Relational Database Management System* (RDBMS)), neste caso o PostgreSQL.

O sistema é uma nova versão que surgiu a partir do sistema BDGF para Aves (Vieira, 2015), desenvolvido em 2015, e foi construído de forma que outros projetos de melhoramento animal da Embrapa possam adotar o mesmo modelo de banco de dados e também uma mesma linha de desenvolvimento. Além disso, o sistema inclui políticas de acesso e segurança aos dados de fenótipos e genótipos dos animais. Visando melhorar o desempenho na consulta aos dados, devido ao elevado número de registros, foi utilizada a tecnologia *JavaScript Object Notation* (JSON) para armazenamento dos genótipos e fenótipos. O sistema foi desenvolvido utilizando-se de ferramentas de software livre para ambiente web, tais como Java¹, Primefaces², Hibernate³, PostgreSQL⁴ e WildFly⁵.

Como utilizar o sistema BDGF

O usuário deve acessar o sistema BDGF por meio de um navegador Web tradicional, como o Firefox, por exemplo. Para exemplificar, utilizaremos como endereço a máquina de testes do sistema. O endereço disponível para os empregados da Embrapa poderem acessar é <http://www.lmb.cnptia.embrapa.br/bdgf>.

Ao acessar o endereço, a página de login da Figura 1 será exibida:

1 Disponível em: <<https://www.oracle.com/br/java/>>.

2 Disponível em: <<https://www.primefaces.org/>>.

3 Disponível em: <<http://hibernate.org/>>.

4 Disponível em: <<https://www.postgresql.org>>.

5 Disponível em: <<http://wildfly.org/downloads/>>.



Figura 1. Tela de login do sistema BDGF.

Para acessar as funcionalidades do sistema e do módulo, o usuário deve utilizar seu usuário e senha cadastrados no Lightweight Directory Access Protocol (LDAP) da Embrapa. Caso esteja cadastrado, o usuário será redirecionado para a tela da Figura 2.

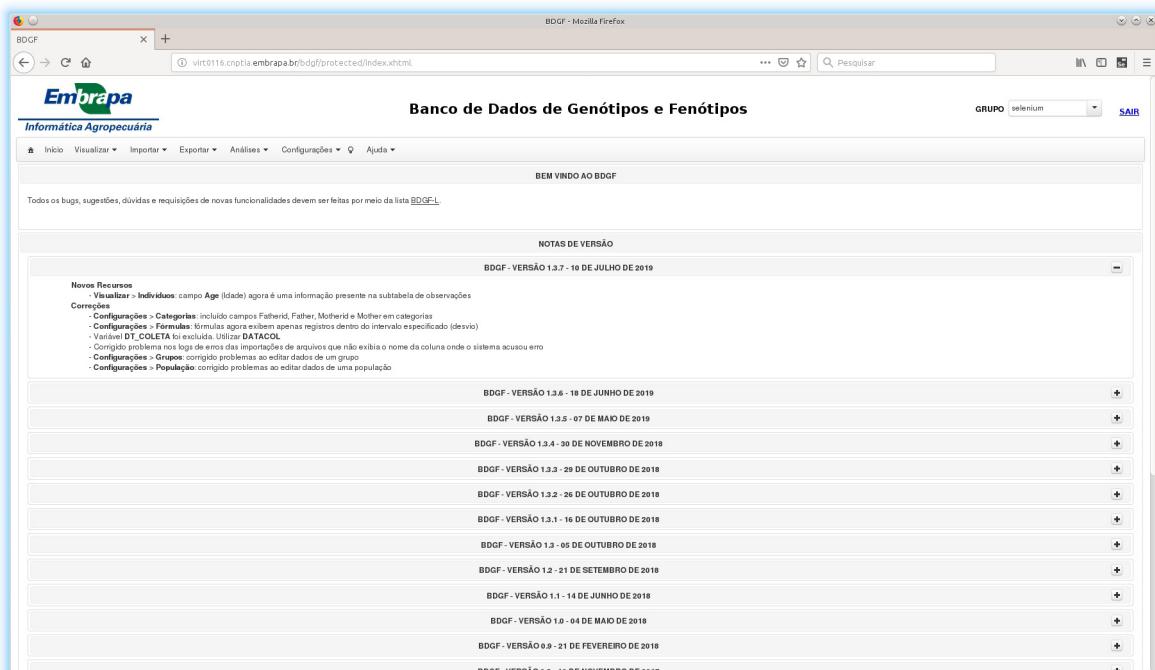


Figura 2. Tela principal do sistema BDGF.

A tela principal do sistema é composta de um menu contendo as opções “Visualizar”, “Importar”, “Exportar”, “Análises”, “Configurações” e “Ajuda”, além da caixa de seleção da linguagem a ser utilizada (inglês ou português) no rodapé da página, e do link “Sair” no canto superior direito. Clicando em “Início”, o usuário será direcionado para essa mesma tela. Clicando em “Sair”, irá para a tela de Login.

A primeira funcionalidade que será mostrada será a opção de consulta de dados de indivíduos, que é o item visualizar indivíduos do menu “Visualizar”. Clicando nessa opção, a seguinte tela (Figura 3) será exibida com os animais cadastrados no sistema.

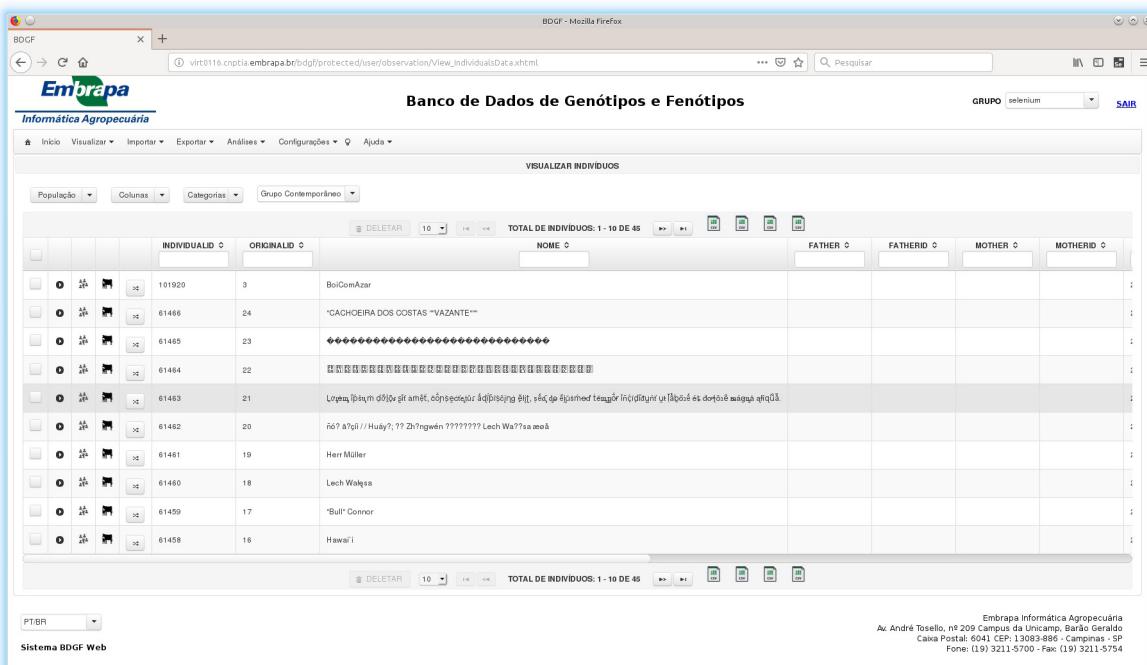


Figura 3. Tela com indivíduos cadastrados no sistema.

Na tela da Figura 3, o usuário encontra diversas informações sobre o indivíduo, tais como: código identificador do indivíduo (INDIVIDUALID), nome original (ORIGINALID), pai (FATHER), mãe (MOTHER), data de inserção na população (INSERTDATE), população (POPULACAONOME) e outras informações contidas nas variáveis JSON relativas ao tipo do indivíduo (gado de corte, ave etc.). Contudo, cabe ressaltar que essas variáveis dos fenótipos relacionados às espécies consideradas pelo sistema devem ser previamente registradas. Essas variáveis devem ser importadas do Sistema de Experimentos da Embrapa (SIExp) (Apolinário et al., 2016), em que foram definidas para a espécie com a qual o usuário trabalhará no seu grupo de usuários (ex: aves, gado de corte, suínos etc.).

Ao selecionar um ou mais indivíduos, é possível excluir esse(s) indivíduo(s). Para tanto, basta clicar no seguinte botão acima da listagem (Figura 4):



Figura 4. Botão para exclusão de indivíduos selecionados.

Também é possível baixar arquivos, do tipo CSV, dos animais (indivíduos), observações, pedigree e genótipos (arquivo PED) desses indivíduos. Para isso, basta clicar num dos botões da Figura 5.

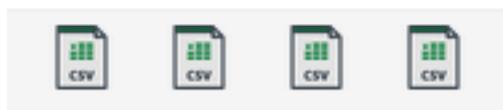


Figura 5. Botões para download de arquivos CSV de animais, observações, pedigree e genótipos.

Além desses botões, existem outras quatro funcionalidades na tela de visualização de indivíduos. São quatro ícones (Figura 6) que ficam ao lado de cada indivíduo na listagem exibida. O primeiro deles expande a linha e exibe os fenótipos (se existirem) cadastrados para aquele indivíduo. O segundo, e talvez mais o importante, abre uma janela exibindo o pedigree daquele indivíduo, ou seja, a sua árvore genealógica, com pai, mãe, avós etc. O terceiro ícone exibe uma janela com a progênie do animal, ou seja, seus filhos (se houver). O quarto botão abre uma janela onde é possível selecionar a nova população daquele indivíduo.

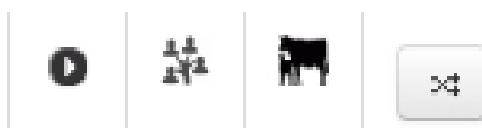


Figura 6. Botões para visualizar fenótipos, pedigree, progênie e alterar população atual do indivíduo, respectivamente.

Outra funcionalidade importante é a opção de visualização de observações (fenótipos) de indivíduos, que é o item “Observações” do menu “Visualizar”. Clicando nessa opção, a seguinte tela (Figura 7) será exibida com os arquivos de fenótipos importados para o sistema.

ID	Data de inserção	Nome	Descrição
15	2019-05-22 04:22:58,589	teste_formulas_observations.csv	teste formulas
14	2019-05-07 16:28:53,038	teste_formulas_observations.csv	teste formulas

Figura 7. Tela para visualização de observações (fenótipos) dos indivíduos.

Nessa tela de visualização de observações também é possível excluir os fenótipos selecionados por meio de um botão acima da lista. Além disso, é possível abrir a tela de visualização de indivíduos clicando no ícone ao lado de cada indivíduo, a qual irá mostrar somente os animais cadastrados naquele arquivo de fenótipo.

Na tela de importação de indivíduos (Figura 8), é possível o usuário selecionar o arquivo texto (CSV), que deve possuir formato específico (campos ORIGINALID E INDIVIDUALID). Caso o indivíduo esteja sendo cadastrado pela primeira vez, também será necessário especificar a população (campo POPULACAONOME) à qual ele pertence.

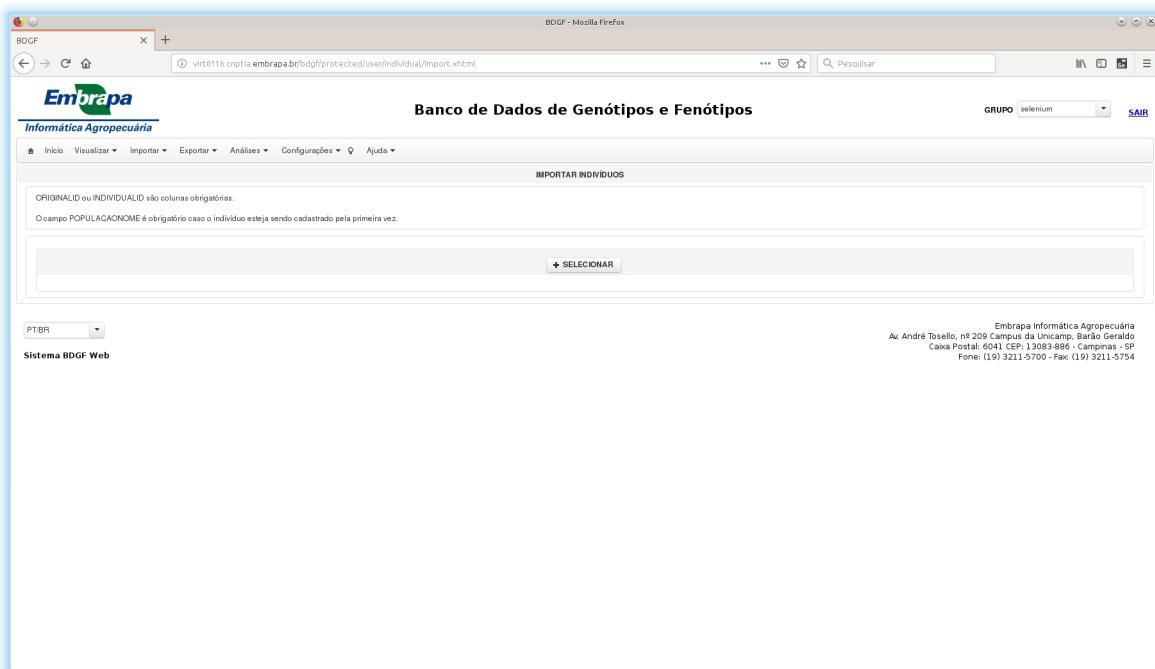


Figura 8. Tela para importação de indivíduos.

Para selecionar o arquivo, basta clicar no botão “Selecionar”. Após isso, se toda a análise de consistência estiver correta, deve-se clicar no botão “Submit” para armazenar os dados no banco.

Na tela de importação de observações (Figura 9) é possível o usuário selecionar o arquivo texto (CSV), que deve possuir formato específico (campos ORIGINALID, POPULACAONOME E DATACOL). O campo DATACOL deve conter as variáveis permitidas para aquele organismo no formato JSON (ou seja, chave:valor).

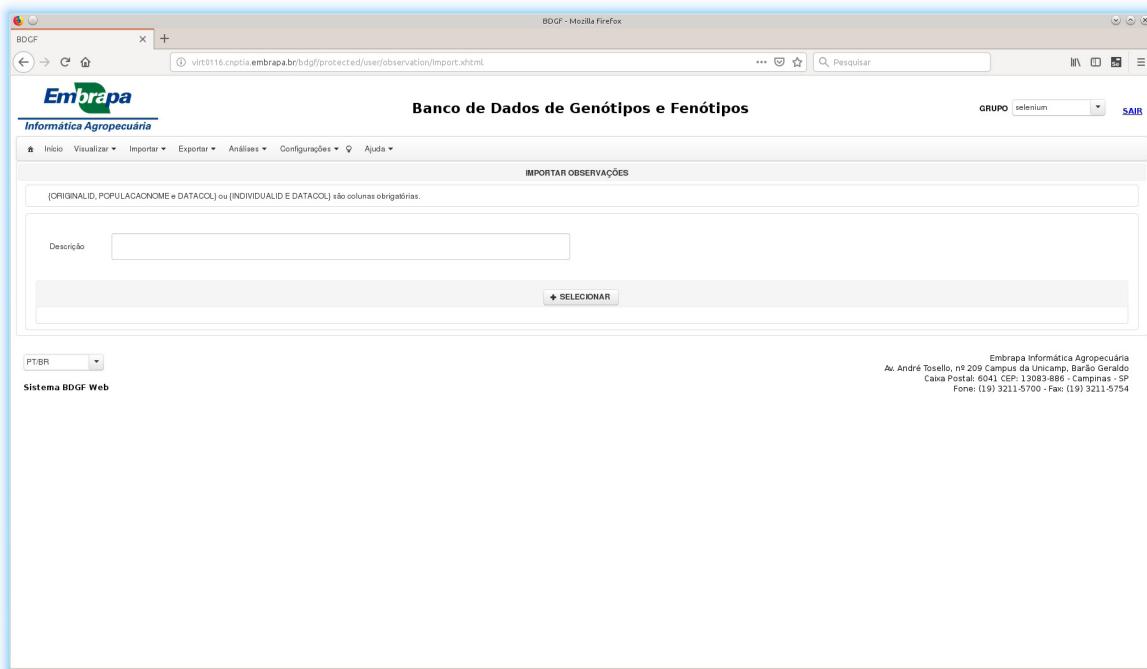


Figura 9. Tela para importação de observações (fenótipos) dos indivíduos.

Para selecionar o arquivo, deve-se clicar no botão “Selecionar”. Feito isso, se toda a análise de consistência estiver dentro do esperado, o próximo passo é clicar no botão “Submit” para armazenar os dados no banco.

A tela de importação de pedigree (Figura 10) permite que o usuário selecione o arquivo texto (CSV), que deve possuir formato específico (campos INDIVIDUALID, FATHERID E MOTHERID para versão com três colunas e campos ORIGINALID, POPULACAOID, ORIGINALIDPAI, POPULACAOIDPAI, ORIGINALIDMAE, POPULACAOIDMAE para versão com seis colunas e com id's das populações e, por fim, os campos ORIGINALID, POPULACAONOME, ORIGINALIDPAI, POPULACAONOMEPAI, ORIGINALIDMAE, POPULACAONOMEMAE para versão com seis colunas e nomes das populações).

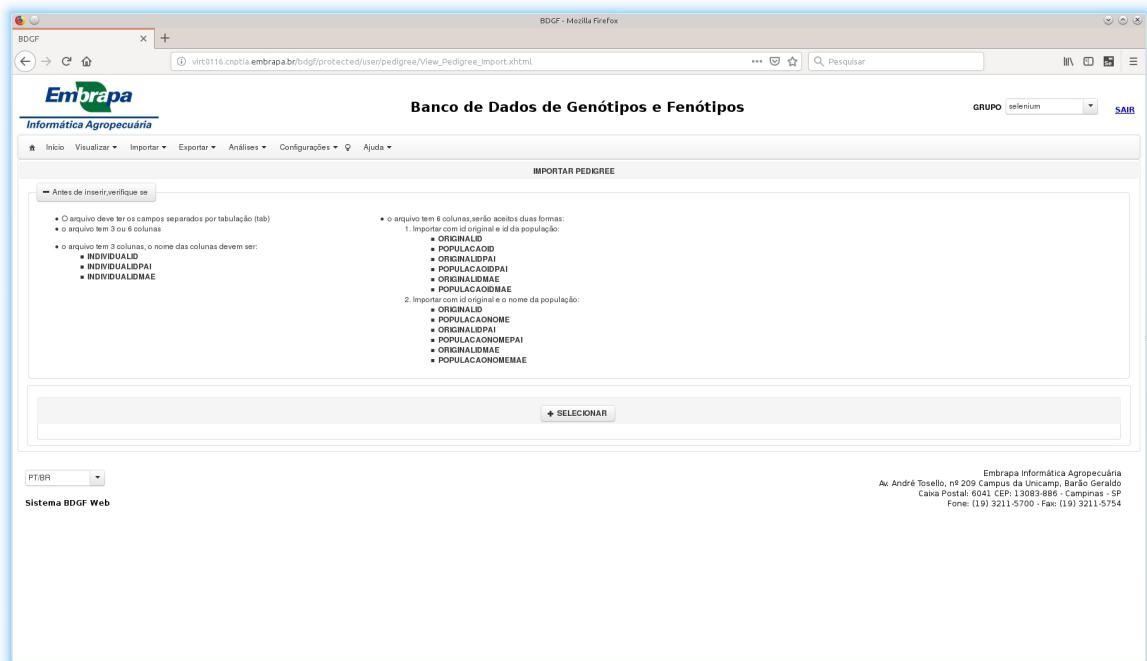


Figura 10. Tela para importação de pedigree dos indivíduos.

Para selecionar o arquivo, deve-se clicar no botão “Selecionar” e uma caixa de diálogo será exibida. Se toda a análise de consistência estiver correta, basta clicar no botão “Submit” para armazenar os dados de pedigree no banco de dados.

Também é possível importar dados de arquivos MAP (Figura 11), comumente utilizados por ferramentas de bioinformática, como Plink (Purcell et al., 2007), por exemplo. O Plink realiza análises em dados genotípicos e fenotípicos em busca de uma associação entre ambos, e precisa, em sua funcionalidade básica, de dois arquivos de entrada: um com extensão PED e outro com extensão MAP. O arquivo MAP, que mapeia os valores da coluna 7 em diante do arquivo PED, possui o seguinte formato de colunas (Tabela 1).

Coluna	Descrição
1	Cromossomo do SNP (1-22, X, Y ou 0 se não localizado)
2	Identificação do SNP
3	Distância Genética
4	Posição do Par de Nucleotídeos

Tabela 1. Colunas do arquivo MAP.

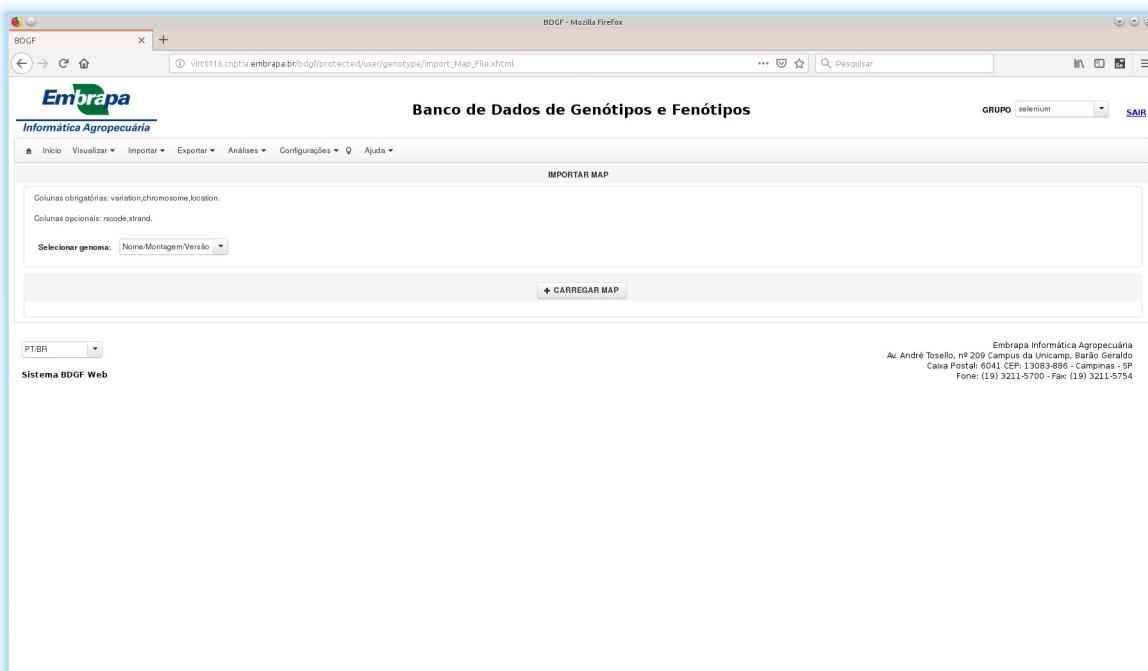


Figura 11. Tela para importação de arquivos MAP.

Para selecionar o arquivo, deve-se clicar no botão “Carregar MAP” e selecionar o genoma ao qual pertence aquele arquivo MAP. Após o upload do arquivo, se toda a análise de consistência estiver correta, o sistema armazenará os dados no banco.

Para a importação do arquivo PED, também utilizado por ferramentas de bioinformática, como Plink, utiliza-se a tela da Figura 12. O arquivo PED possui colunas separadas por espaços ou pelo caractere <TAB>, identificadas a seguir na Tabela 2.

Coluna	Descrição
1	Identificação da Família ou Grupo dos Indivíduos
2	Identificação do Indivíduo
3	Identificação do Pai
4	Identificação da Mãe
5	Sexo do Indivíduo; (1=macho; 2=fêmea; other=desconhecido)
6	Fenótipo do Indivíduo; (-9=sem informação; 0=não afetado; 1=afetado)
7 em diante	Genótipos do Indivíduo. (ex: AA, AG, GG etc.)

Tabela 2. Colunas do arquivo PED.

As primeiras 6 colunas são obrigatorias, sendo que, da 7 em diante, os valores são os pares de alelo dos genótipos pesquisados, os quais estarão numa quantidade de acordo com os marcadores SNPs mapeados no arquivo MAP.

The screenshot shows a Mozilla Firefox browser window with the BDGF (Banco de Dados de Genótipos e Fenótipos) application. The title bar says 'BDGF - Mozilla Firefox'. The main page header is 'Embrapa Informática Agropecuária' and the sub-header is 'Banco de Dados de Genótipos e Fenótipos'. A dropdown menu 'GRUPO' is set to 'selenium'. A 'SAIR' (Logout) link is visible. The main content area is titled 'PED' and contains a form for importing a PED file. The form has fields for 'Importar com:' (radio buttons for 'Individual' or 'Original'), 'Conjunto de marcadores:' (dropdown menu 'Selecionar marcador'), 'Descrição:' (text input), 'Licença:' (text input), 'Data da genotipagem:' (text input), and a 'CARREGAR PED' (button). Below the form is a note: '* campos obrigatórios' and 'obs: Arquivos ped maiores que 1GB podem demorar alguns minutos para serem processados!'. At the bottom left is a language dropdown 'PT-BR' and at the bottom right is contact information for Embrapa: Av. André Tosello, nº 209 Campus da Unicamp, Barão Geraldo, Caixa Postal: 6041 CEP: 13083-886 - Campinas - SP, Fone: (19) 3211-5754.

Figura 12. Tela para importação de arquivo PED.

Para selecionar o arquivo, deve-se clicar no botão “Carregar PED”, selecionar o conjunto de marcadores importados de um MAP e preencher os campos obrigatórios “Descrição”, “Licença” e “Data de genotipagem”. Preenchidos esses campos, se toda a análise de consistência estiver correta, o sistema armazenará automaticamente os dados no banco.

O caminho inverso também é possível, ou seja, exportar (ou fazer o download) os arquivos MAP salvos no banco de dados (Figura 13). Além disso, é possível excluir um desses arquivos salvos por meio do botão “Deletar”.

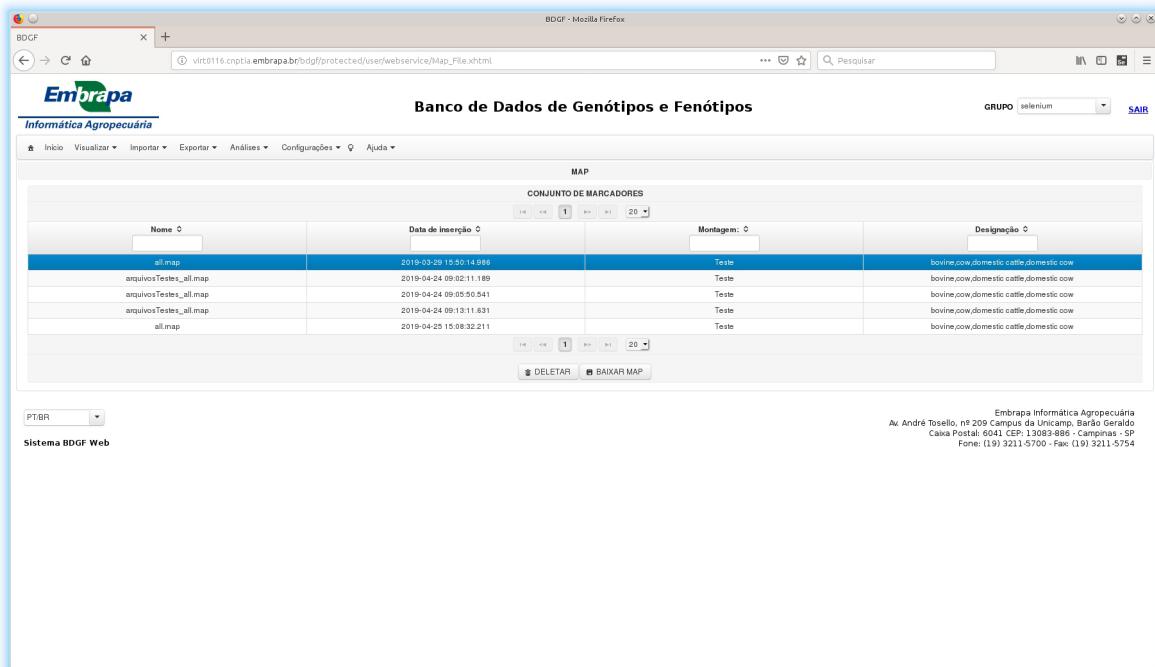


Figura 13. Tela para excluir ou exportar (baixar) arquivos MAP para computador local.

Para fazer o download de um dos arquivos MAP, basta selecionar uma das linhas e clicar no botão “Baixar MAP”, e um arquivo MAP será salvo em seu computador.

Também é possível exportar (baixar) arquivos PED salvos no banco BDGF para o computador local (Figura 14), além da possibilidade de excluir um desses arquivos. Para realizar o download de um dos arquivos PED, basta selecionar uma das linhas e clicar no botão “Baixar PED”, e um arquivo do tipo PED será salvo em seu computador. Dependendo da quantidade de dados salvo, o download poderá demorar alguns minutos. Nesta tela também é possível baixar o arquivo MAP relativo àquele PED clicando no botão “Baixar MAP”.



Figura 14. Tela para excluir ou exportar (baixar) arquivos PED para computador local.

A próxima funcionalidade é o item “Verificar indivíduos”, no menu “Análises”. Na tela da Figura 15 é possível verificar arquivos do tipo CSV de indivíduos e observações antes de fazer sua importação para o sistema BDGF. Se houver inconsistências em relação ao banco, ou mesmo se aparecerem erros no arquivo, uma tabela com erros e avisos é exibida para o usuário.



Figura 15. Tela para verificação de arquivos CSV de indivíduos e observações.

Para selecionar o arquivo a ser verificado, deve-se clicar no botão “Carregar Arquivo” e selecionar o arquivo CSV de indivíduos ou observações. Após a análise, o sistema informará ao usuário se encontrou inconsistências ou se os dados estão consistentes.

A segunda funcionalidade do menu “Análises” é uma das mais importantes do sistema BDGF, que é o item “Encontrar indivíduos duplicados” (Figura 16). Por meio dela, é possível buscar na base de dados por animais que foram importados e que são possivelmente duplicados. Isso pode acontecer, pois um mesmo indivíduo (animal) pode aparecer em planilhas de diversas fazendas (por motivos variados, como a venda ou troca de rebanhos por parte dos criadores). E quando essas planilhas são importadas, esses indivíduos ganham um novo id do sistema, porém, com os mesmos dados fenotípicos.

The screenshot shows the BDGF application running in Mozilla Firefox. The main title is "Banco de Dados de Genótipos e Fenótipos". The interface includes a navigation bar with links like Início, Visualizar, Importar, Exportar, Análises, Configurações, and Ajuda. A dropdown menu "GRUPO" is set to "selenium". A search bar at the top right contains the text "Pesquisar" and a dropdown "GRUPO" set to "selenium". Below the search bar is a button labeled "SAIR". The main content area is titled "DUPLICACÕES" and displays a table of individuals. The table has columns: INDIVIDUALID, ORIGINALID, NOME, FATHER, MOTHER, START DATE, and others. A total of 45 individuals are listed. A button "Procurar duplicações" is located at the top right of the table area.

Figura 16. Tela para procura de indivíduos duplicados no banco de dados.

Para descobrir quais os possíveis animais duplicados, basta clicar no botão que aparece no ínicio de cada linha dos indivíduos indicados pelo sistema como potenciais repetidos. Ao clicar nesse botão, a seguinte tela (Figura 17) será exibida:

The screenshot shows the BDGF application running in Mozilla Firefox. The main title is "Banco de Dados de Genótipos e Fenótipos". The interface includes a navigation bar with links like Início, Visualizar, Importar, Exportar, Análises, Configurações, and Ajuda. A dropdown menu "GRUPO" is set to "selenium". A search bar at the top right contains the text "Pesquisar" and a dropdown "GRUPO" set to "selenium". Below the search bar is a button labeled "SAIR". The main content area is titled "POSSÍVEIS DUPLICADOS" and displays two rows of individual data. The first row is for "Herr Müller" and the second for "Lech Wa?sa wa?". A message at the bottom says "Clique no '+' para adicionar indivíduos duplicados manualmente" (Click '+' to add manually duplicated individuals). A button "Associar" is located at the bottom center of the table area.

Figura 17. Tela para associação de indivíduos duplicados em um animal somente.

Nesta tela, o usuário seleciona o(s) animal(is) na tabela inferior que devem ser associados (mesclados) com o animal selecionado na tela anterior e que aparece na tabela superior. Após a seleção do(s) indivíduo(s), basta clicar no botão “Associar” que todos esses indivíduos assumem um único id, tendo todos os dados de pedigree e observações redirecionados automaticamente para esse único animal.

Partindo para o menu “Configurações” do sistema, pode-se encontrar funcionalidades ligadas ao gerenciamento de dados de outras tabelas do banco de dados do sistema, tais como grupos de usuários, categorias de variáveis fenotípicas, grupos contemporâneos (Figura 18), genomas, populações de animais e fórmulas.

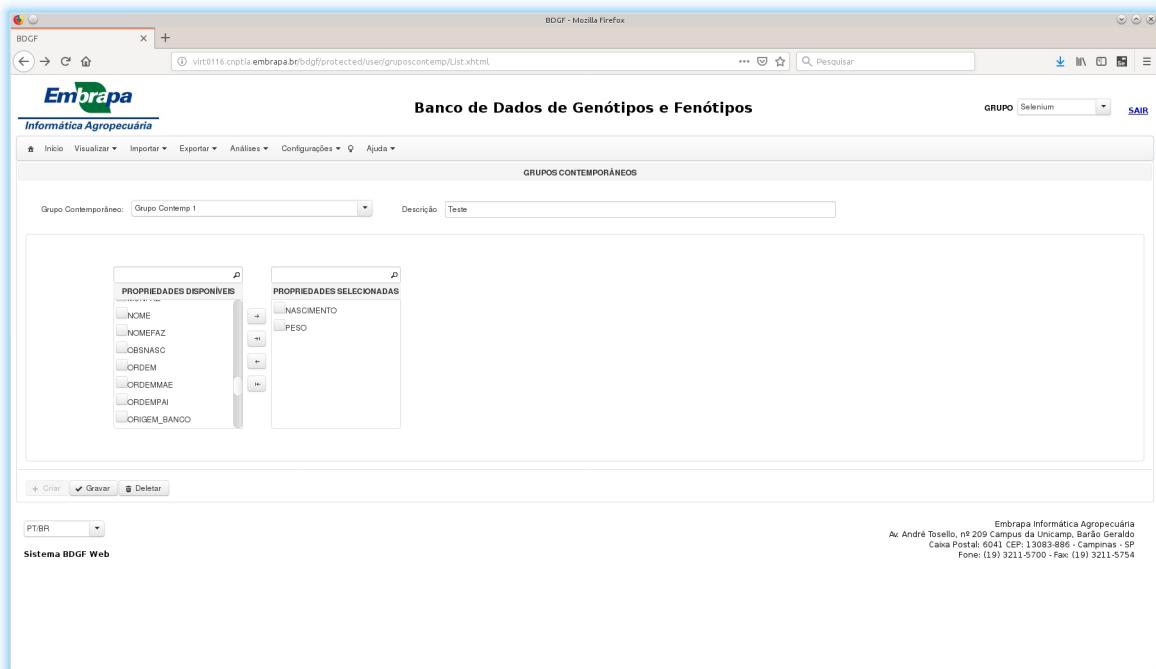


Figura 18. Tela para inserção, edição e remoção de categoria de grupos contemporâneos.

O primeiro item do menu “Configurações” trata do gerenciamento dos dados de grupos de usuários. Ao se clicar neste item, a seguinte tela da Figura 19 é exibida:

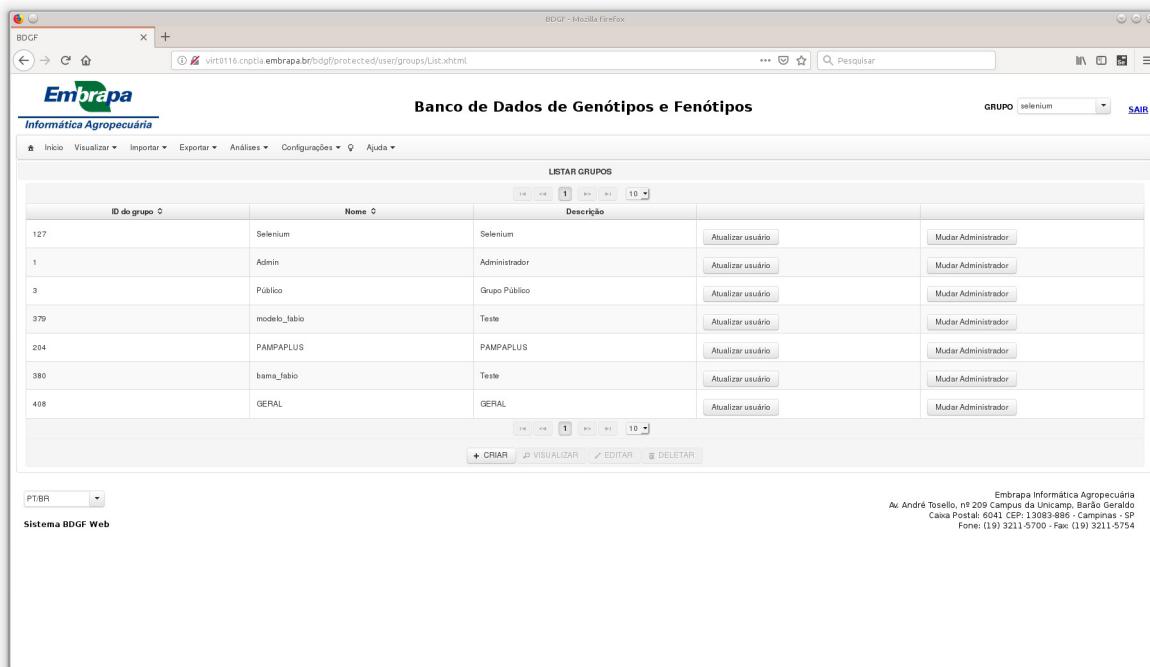


Figura 19. Tela para inserção, edição, remoção e visualização de grupos.

A tela da Figura 19 possibilita ao usuário inserir, editar, excluir e visualizar detalhes dos dados de um grupo de usuários. Para isso, deve-se utilizar os botões “Criar”, “Editar”, “Deletar” e “Visualizar”, respectivamente. Janelas de diálogo serão abertas para cada um dos procedimentos, as quais guiarão o usuário quanto aos campos de preenchimento obrigatório (no caso de criação e edição de dados).

Além dessas funcionalidades, nesta tela é possível adicionar e retirar usuários do grupo por meio do botão “Atualizar usuário”, presente em cada linha da tabela de grupos. Ao se clicar neste botão, a seguinte tela de diálogo da Figura 20 é exibida:

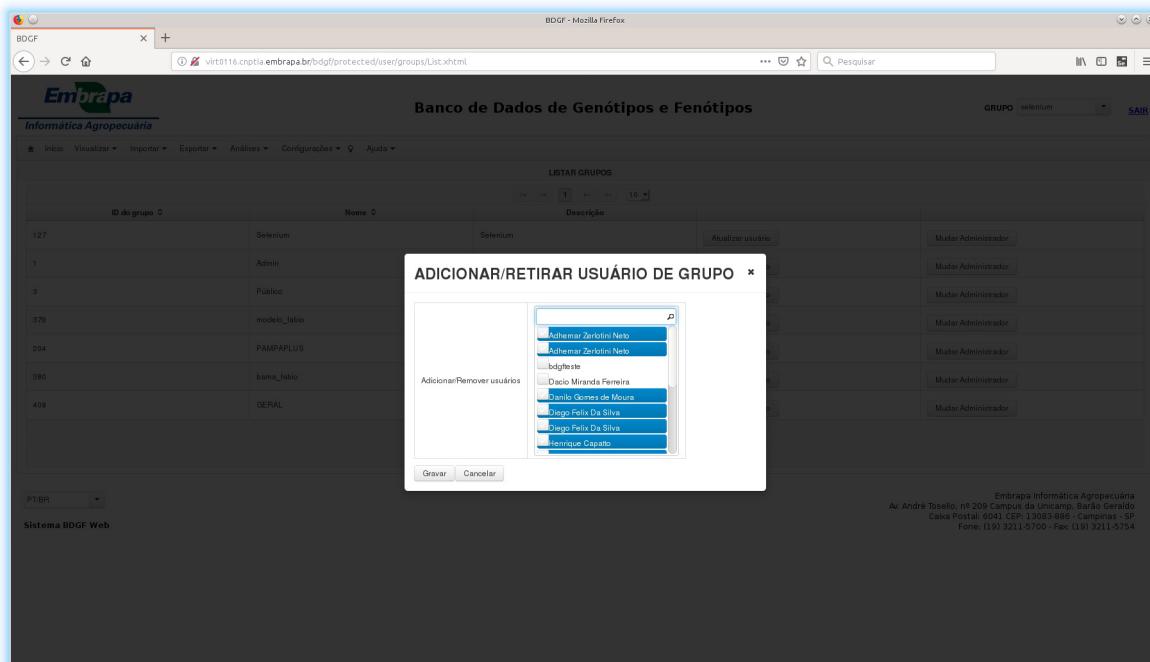


Figura 20. Tela para adicionar e retirar usuários de um grupo.

Nessa tela é possível adicionar e/ou retirar usuários do grupo apenas marcando ou desmarcando as caixas de seleção. Após a escolha de quais usuários comporão o grupo, basta clicar no botão “Salvar” que a composição do grupo estará atualizada.

Outra função encontrada na tela de gerenciamento de grupos é alterar o administrador de um grupo. Clicando no botão “Mudar administrador” a janela da Figura 21 é chamada:

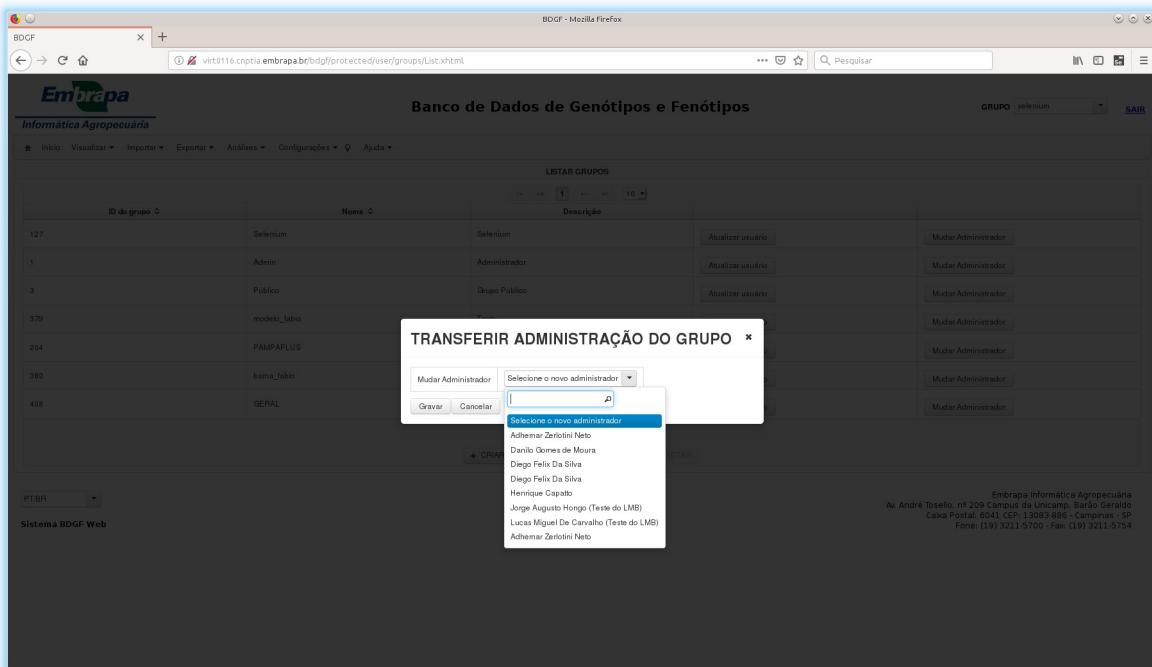


Figura 21. Tela para mudança de administrador do grupo de usuários.

Basta o usuário selecionar o novo administrador do grupo e clicar no botão “Gravar” para concretizar a alteração. Vale lembrar que o usuário responsável por essa alteração deve ser o administrador atual do grupo a ser modificado.

A tela da Figura 22 tem a função de adicionar ou modificar categorias de variáveis (características) fenotípicas e fórmulas baseadas nessas variáveis utilizadas por cada espécie de indivíduo cadastrado (ex: bovino de corte). A criação de categorias permite a exibição apenas do conjunto de variáveis selecionadas na tela de visualização de indivíduos.

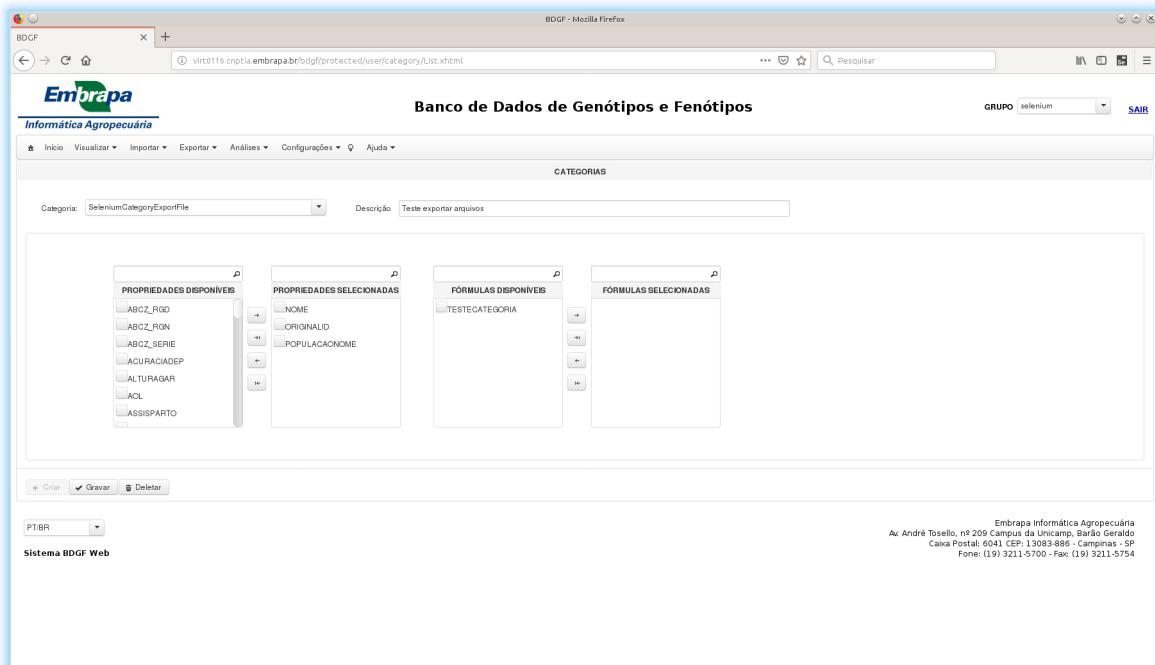


Figura 22. Tela para inserção, edição e remoção de categoria de variáveis fenotípicas.

Nesta tela, o usuário deve selecionar (ou criar) uma categoria e nela inserir as variáveis selecionadas no quadro “Propriedades disponíveis”. Também é possível adicionar a uma categoria as fórmulas criadas pelos usuários do grupo atual. Basta selecioná-las no quadro “Fórmulas disponíveis” e enviá-las para o quadro “Fórmulas selecionadas”. Por fim, basta clicar no botão “Gravar” para efetivar as alterações. A criação de fórmulas será vista adiante neste documento.

A tela da Figura 23 permite ao usuário inserir, editar e excluir dados de um genoma de um organismo já cadastrado anteriormente. Os dados do genoma informam qual o chip utilizado para genotipagem, a versão e a qual organismo se refere essa genotipagem (ex: Bos indicus, Bos taurus etc.). Para tanto, basta utilizar os botões “Novo”, “Editar” e “Deletar”, respectivamente. Antes da importação do arquivo MAP, será necessário adicionar os genomas que serão vinculados a estes arquivos MAPs, pois, para cada importação de arquivo MAP, deve-se indicar de qual organismo provêm os dados.

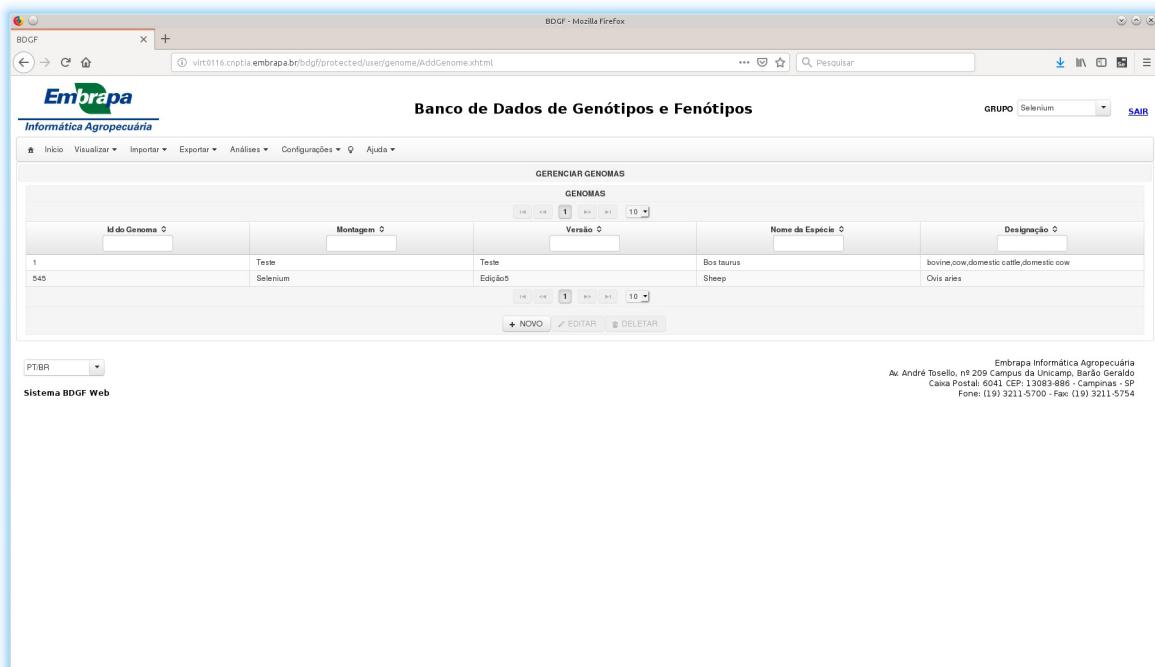


Figura 23. Tela para inserção, edição e remoção de genomas.

Serão abertas janelas de diálogo correspondentes para cada um dos procedimentos, as quais guiarão o usuário quanto aos campos de preenchimento obrigatório (no caso de criação e edição de dados). Um exemplo pode ser visto na Figura 24 a seguir. Nesse caso, serão três campos obrigatórios a serem preenchidos: “Espécie”, onde deverá ser selecionada dentre a lista fornecida pelo sistema uma espécie para o genoma; nome da “Montagem” e “Versão” da montagem.

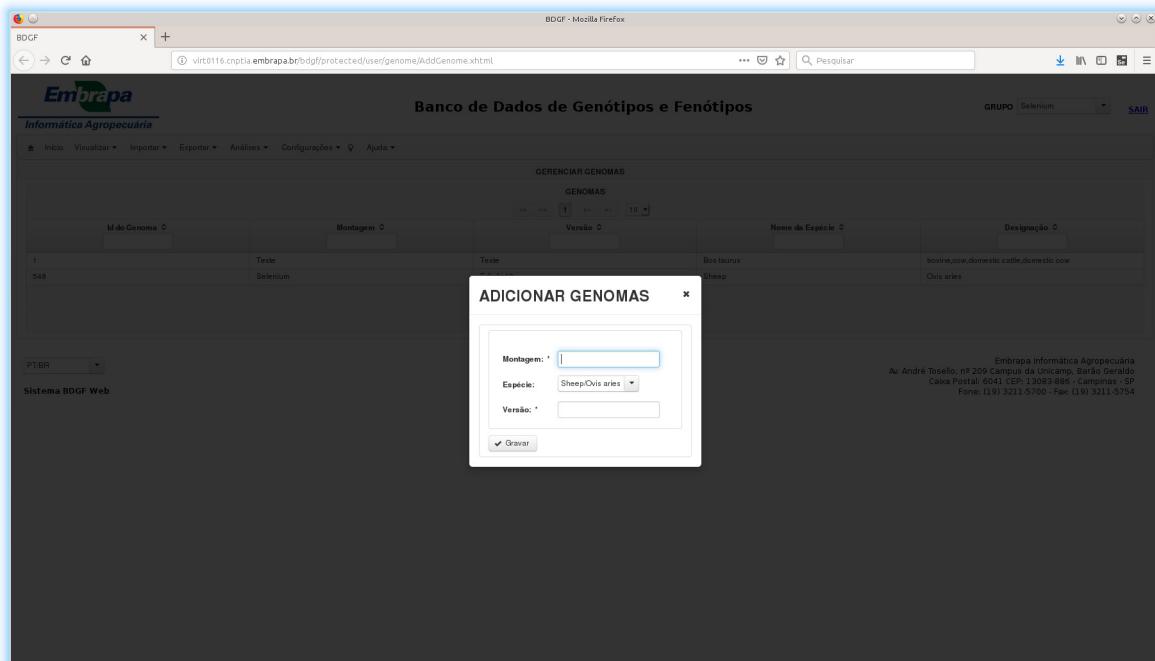


Figura 24. Janela de diálogo para criação de genomas.

A tela da Figura 25 permite ao usuário inserir, editar, excluir e visualizar detalhes dos dados de uma população, a qual deverá estar vinculada a um grupo de usuários e um organismo (ex: bovino de corte). Para isso, deve-se utilizar os botões “Criar”, “Visualizar”, “Editar” e “Deletar”, respectivamente.

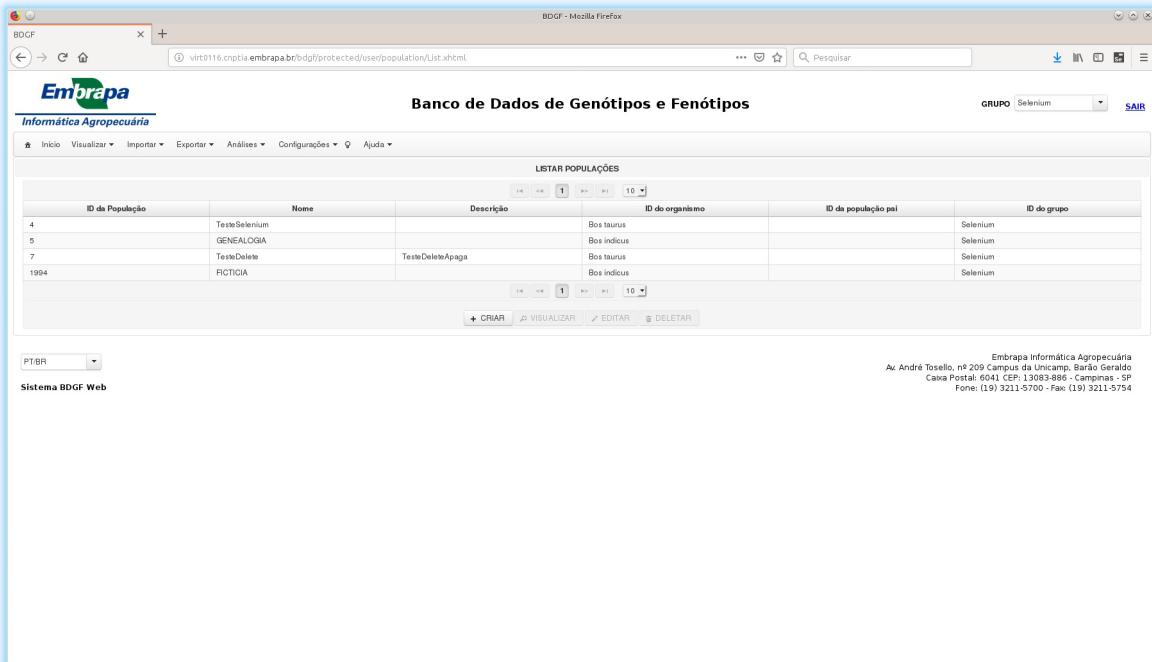


Figura 25. Tela para inserção, edição, remoção e visualização de populações.

Também serão abertas janelas de diálogo para cada um dos procedimentos, como da Figura 26. Os campos “Nome” e “ID do organismo” são de preenchimento obrigatório. Após finalizar o preenchimento, basta clicar no botão “Gravar”.

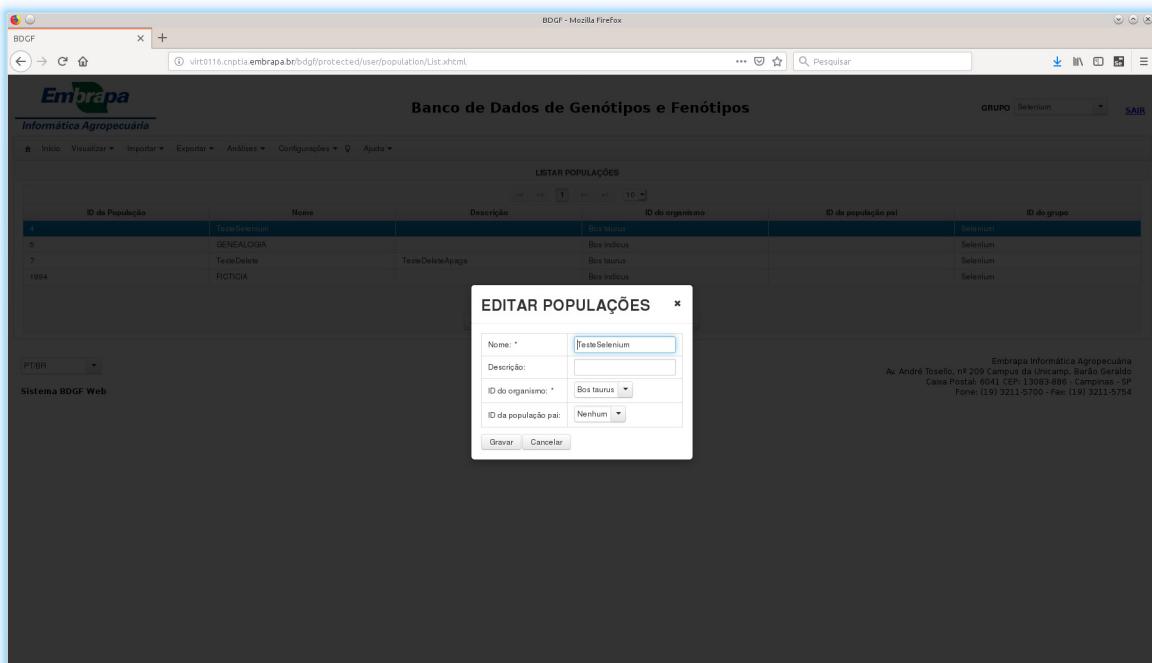


Figura 26. Janela de diálogo para edição de dados de população.

No sistema BDGF também é possível criar fórmulas (Figura 27), que são utilizadas para realizar cálculos estatísticos em cima dos dados salvos no sistema. Esses dados são os fenótipos observados para cada animal, tais como peso atual, peso ao nascimento, altura etc. Para essas variáveis é possível calcularmos a média, mediana, valores mínimo, máximo e mais próximo, utilizando o desvio-padrão e o número máximo de dias entre a coleta e a data de nascimento do indivíduo.

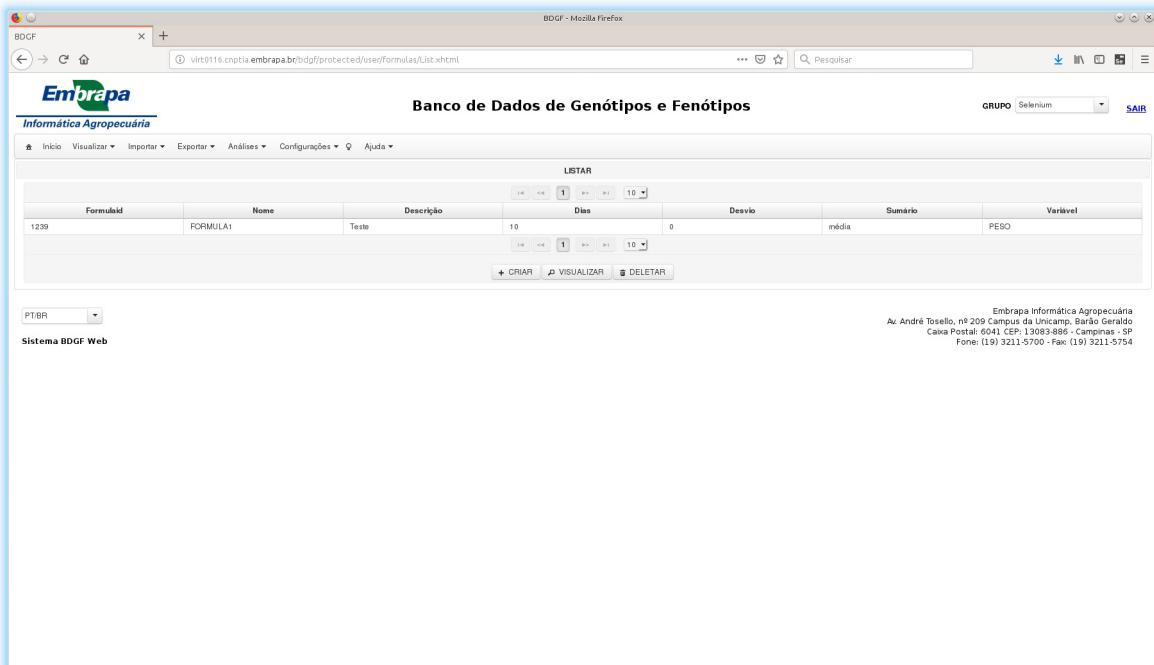


Figura 27. Tela para inserção, visualização e deleção de fórmulas do sistema.

Na criação de uma fórmula, um diálogo com formulário será aberto, com campos relacionados ao cálculo das fórmulas. Um exemplo pode ser visto na Figura 28 seguinte:

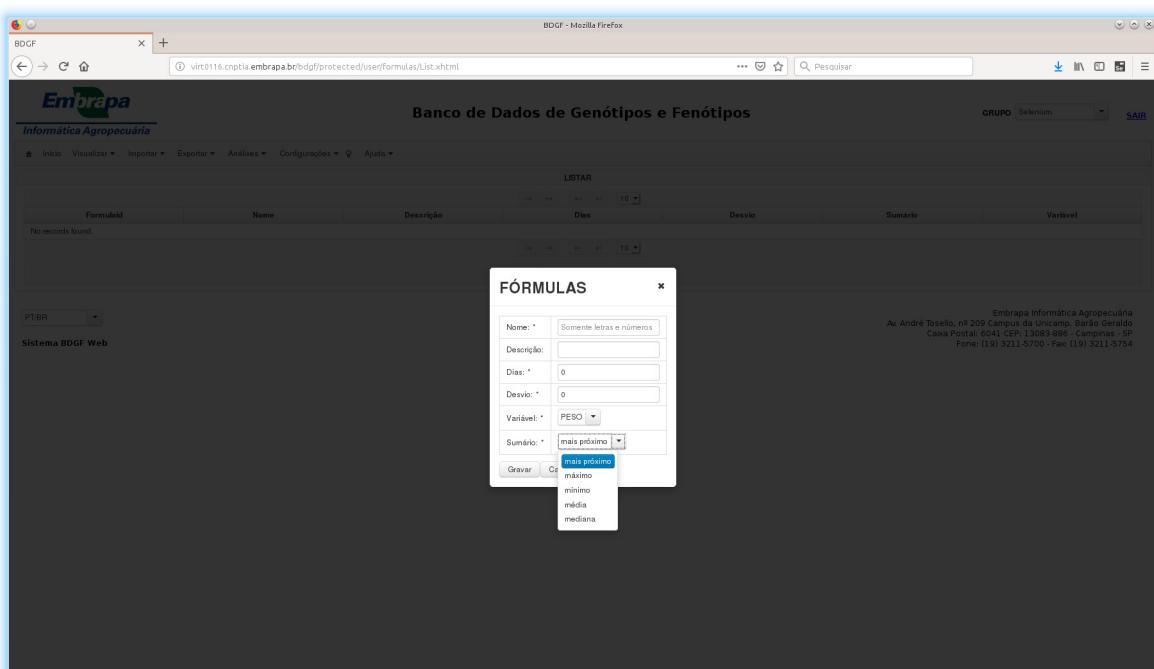


Figura 28. Janela de diálogo para inserção dos dados de uma fórmula.

Basta preencher todos os campos e, ao finalizar, clicar em “Gravar” para efetivar as informações no sistema. Assim que uma fórmula é criada, os cálculos serão aplicados à variável escolhida no formulário preenchido, e para ver os resultados deve-se acessar a página de visualização de indivíduos (animais). É possível excluir uma fórmula que não seja mais necessária. Para tanto, basta selecionar a fórmula a ser excluída na tabela e, em seguida, clicar no botão “Deletar”.

A tela da Figura 29 ilustra a funcionalidade de Logs do sistema, que realiza o monitoramento do sistema, e exibe todas as operações realizadas no software, tais como: importar arquivos, inserir, editar e excluir registros. Junto a isso, somam-se informações indicando a data e o horário do evento, assim como o usuário que realizou a operação no sistema.

LogId	Pesquisador	Operação
2846	bdgfteste	OPERAÇÃO: EXPORTAR TIPO ARQUIVO: PEDIGREE INFORMAÇÕES ADICIONAIS: Populações selecionadas [FICTIONIA] / Quantidade de indivíduos selecionados [8]
2845	bdgfteste	OPERAÇÃO: IMPORTAR TIPO ARQUIVO: PEDIGREE INFORMAÇÕES ADICIONAIS: pedigree.csv
2844	bdgfteste	OPERAÇÃO: EXPORTAR TIPO ARQUIVO: PEDIGREE INFORMAÇÕES ADICIONAIS: Populações selecionadas [FICTIONIA] / Quantidade de indivíduos selecionados [8]
2843	bdgfteste	OPERAÇÃO: IMPORTAR TIPO ARQUIVO: OBSERVAÇÕES INFORMAÇÕES ADICIONAIS: observations.csv
2842	bdgfteste	OPERAÇÃO: EXPORTAR TIPO ARQUIVO: INDIVÍDUOS INFORMAÇÕES ADICIONAIS: Populações selecionadas [FICTIONIA] / Quantidade de indivíduos selecionados [8]
2841	bdgfteste	OPERAÇÃO: IMPORTAR TIPO ARQUIVO: PEDIGREE INFORMAÇÕES ADICIONAIS: pedigree.csv
2840	bdgfteste	OPERAÇÃO: IMPORTAR TIPO ARQUIVO: INDIVÍDUOS INFORMAÇÕES ADICIONAIS: pedigree.csv
2839	bdgfteste	OPERAÇÃO: EXCLUIR TIPO ARQUIVO: INDIVÍDUOS INFORMAÇÕES ADICIONAIS: Populações selecionadas [FICTIONIA],GENEALOGIA,SELENIUM,BASE,SELENIUM.EDIT,SELENIUM_GRUPO_01,SELENIUM_GRUPO_02,SeleniumTeste,TESTE,TESTI
2838	bdgfteste	OPERAÇÃO: IMPORTAR TIPO ARQUIVO: PEDIGREE INFORMAÇÕES ADICIONAIS: pedigreeSQL.csv
2837	bdgfteste	OPERAÇÃO: EXCLUIR TIPO ARQUIVO: MAP INFORMAÇÕES ADICIONAIS: teste_telenium.map

Figura 29. Tela para visualização de logs do sistema.

Na tela de gerenciamento de logs é possível, ainda, filtrar as informações pelo usuário, bastando, para isso, clicar no nome de cada um deles, que estão acima da lista de logs registrados do sistema.

A tela seguinte (Figura 30) exibe as variáveis importadas do SIExp. A importação dessas variáveis ocorre por meio da tela da Figura 31 (podendo ser feita apenas pelos usuários do grupo “Admin”), e que foram definidas para a espécie na qual o usuário trabalhará no seu grupo de usuários (ex: bovinos, mamona etc.).

Variável	Tipo	Descrição
UFFAZ	STRING	Estado da fazenda Estado da fazenda
ESC_MUCOSA	NUMBER	Mucosa Escreva visual, nota 1 a 9, onde 9 é desclassificante Valores Permitidos
		1 bronze claro 2 bronze 3 bronze escuro 4 cinza claro 5 cinza 6 cinza escuro 7 rosada ou despigmentada – desclassificante 8 manchas de despigmentação - desclassificante 9 onça - desclassificante

Figura 30. Tela para visualização das variáveis fenotípicas aceitas para um organismo importadas do SIExp.

O sistema fará a importação das variáveis por meio de um arquivo JSON com uma estrutura de variáveis pré-definidas para cada espécie já existente no SIExp, ou seja, quais valores serão válidos para cada variável e o tipo de cada uma delas (string, number, enumeration etc.)

PT-BR

Sistema BDGF Web

Embrapa Informática Agropecuária
Av. André Tosello, nº 209 Campus da Unicamp, Barão Geraldo
Caixa Postal: 6041 CEP: 13083-886 - Campinas - SP
Fone: (19) 3211-9700 - Fax: (19) 3211-9754

Figura 31. Tela para importação de variáveis do sistema SIExp (apenas usuário Admin).

Por fim, a tela da Figura 32 permite ao usuário inserir, editar e excluir dados de um organismo (ou espécie). Para tanto, basta utilizar os botões “Novo”, “Editar” e “Deletar”, respectivamente. Novas janelas serão abertas para cada procedimento.

ID do organismo	Nome	Designação
9940	Sheep	Ovis aries
3988	Ricinus communis	Mamona, Castor bean
9031	Gallus gallus	bantam/chicken
9913	Bos taurus	bovine,cow,domestic,cattle,domestic cow
9915	Bos indicus	zebu, zebu cattle
9000	Teste	Teste

Figura 32. Tela que usuário administrador utiliza para adicionar organismos (espécies) permitidas.

No próprio sistema BDGF existe um menu com ajuda on-line, que também guiará o usuário a resolver grande parte de suas dúvidas.

Considerações Finais

O sistema BDGF foi projetado para apoiar programas de melhoramento animal da Embrapa, mas pode ser facilmente ajustado para armazenar dados de diversas fontes, como dados de plantas. Ele permite o armazenamento e acesso rápido aos dados de pedigree, fenótipos e genótipos, utilizados para realização de avaliações genéticas. Além disso, o sistema inclui políticas de acesso e segurança para fenótipos, genótipos e pedigree dos animais. Como trabalhos futuros, já estão em curso a integração do BDGF com sistemas de avaliação genética e a construção de APIs de fornecimento de relatórios, para que outros sistemas possam obter informações da base de dados BDGF.

Referências

APOLINÁRIO, D. R. de F.; QUEIROS, L. R.; VACARI, I.; CRUZ, S. A. B. da. **SIExp - Sistema de Informação de Experimentos da Embrapa**. Versão v. 1.7.6. Campinas: Embrapa Informática Agropecuária, 2016.

CAETANO, A. R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v. 38, p. 64-71, July 2009. Número especial. DOI: <http://dx.doi.org/10.1590/S1516-35982009001300008>.

HIGA, R. H.; OLIVEIRA, G. B. de. **Banco de Dados de Genótipos e Fenótipos (BDGF) para suporte a estudos de associação genômica ampla e seleção genômica em programas de melhoramento animal**. Campinas: Embrapa Informática Agropecuária, 2015. 30 p. (Embrapa Informática Agropecuária. Documentos, 133). Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/138127/1/Doc133.pdf>>. Acesso em: 22 jul. 2019.

PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M. A. R.; BENDER, D.; MALLER, J.; SKLAR, P.; BAKKER, P. I. W. de; DALY, M. J.; SHAM, P. C. PLINK: a toolset for whole-genome association and population-based linkage analysis. **The American Journal of Human Genetics**, v. 81, n. 3, p. 559-575, Sept. 2007. DOI: <https://doi.org/10.1086/519795>.

VIEIRA, F. D. **Sistema BDGF de Aves**. Versão 1.0. Campinas: Embrapa Informática Agropecuária, 2015. 1 CD-ROM.

