

COMPARAÇÃO VISUAL DE MAPAS E DADOS GENÔMICOS VIA INCORPORAÇÃO DE TECNOLOGIAS TABLEAU, POWER BI E JBROWSE

A. Selli*¹, L.T. Andrietta², R.V. Ventura²

¹Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo / ²Departamento de Nutrição e Produção Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo

*alana.selli@usp.br

Objetivos

A seleção genômica faz o uso de marcadores moleculares para a identificação de animais geneticamente superiores, assumindo-se que estes estejam em desequilíbrio de ligação com QTLs (*Quantitative Trait Loci*), importantes regiões genômicas associadas a características de interesse. Ferramentas de *Business Intelligence* (BI), como o Power BI e Tableau, permitem a integração de fontes de dados distintas e a visualização de dados de forma interativa, possibilitando o uso do JBrowse, visualizador de dados genômicos, de forma mais objetiva. O objetivo deste estudo foi avaliar diferentes painéis de genotipagem em relação à posição, número de marcadores e proximidade destes em relação a QTLs previamente identificados pela literatura, via implementação de duas ferramentas BI, bem como analisar o uso da ferramenta web JBrowse para a visualização do conteúdo alélico oriundo do genoma de bovinos.

Métodos e Procedimentos

Foram utilizados 25 painéis comerciais que caracterizam o genoma bovino, reportados pelo programa USDA NRSP-8. Os *softwares* Power BI e Tableau foram utilizados para gerar visualizações e avaliar a composição de marcadores dos diferentes painéis. Posteriormente, o pacote R GALLO (FONSECA et al., 2020) foi utilizado para a extração de QTLs provenientes do Animal QTL database, que por sua vez foram inseridos no programa Tableau para contrastar a posição destes com marcadores previamente identificados. 5 marcadores foram selecionados por estarem

próximos a um QTL, e a partir destes um arquivo em formato VCF (*Variant Call Format*) foi simulado e inserido no programa JBrowse.

Resultados

Ambos Power BI e Tableau mostraram-se eficientes na visualização dos dados genômicos, permitindo a identificação de grandes diferenças nos painéis comerciais. A análise do banco de QTLs mostrou que metade (52,83%) dos QTLs identificados estão relacionados à produção de leite. Para a visualização de conteúdo alélico de indivíduos no JBrowse, foi necessária a conversão de um arquivo em formato de genotipagem comercial para a extensão VCF, já que o programa não permite a inserção de arquivos genotípicos comumente utilizados no melhoramento animal.

Conclusões

As ferramentas BI mostraram-se eficientes para a integração e visualização de dados genômicos, apresentando vantagens práticas em relação a ferramentas já existentes para a visualização deste tipo de dados, por serem de fácil uso e não exigirem arquivos em formatos específicos.

Referências Bibliográficas

- BUELS, R. et al. JBrowse: A dynamic web platform for genome visualization and analysis. *Genome Biology*, v. 17, n. 1, p. 1–12, 2016.
- FONSECA, P. A. S. et al. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for

positional candidate loci. **GigaScience**, v. 9, n. 12, p. 1–9, 30 dez. 2020.

VISUAL COMPARISON OF MAPS AND GENOMIC DATA VIA INCORPORATION OF THE TABLEAU, POWER BI AND JBROWSE TECHNOLOGIES

A. Selli*¹, L.T. Andrietta², R.V. Ventura²

¹Faculty of Animal Science and Food Engineering, University of São Paulo / ²Department of Animal Nutrition and Production, College of Veterinary and Animal Science, University of São Paulo

*alana.selli@usp.br

Objectives

The genomic selection makes use of molecular markers for the identification of genetically superior animals, assuming these markers are in linkage disequilibrium with QTL (Quantitative Trait Loci), important genomic regions associated with traits of interest. Business Intelligence (BI) tools, such as Power BI and Tableau, allow the integration of distinct data sources and the visualization of data in an interactive way, making possible the use of JBrowse, a genomic data visualization tool, in an objective way. The objective of this study was to evaluate different genotyping maps regarding the position, number of markers, and the proximity of these in relation to QTL previously identified in the literature, as well as to analyze the use of the JBrowse web tool for the visualization of the allelic content of the bovine genome.

Materials and Methods

25 maps which characterize the bovine genome were used, reported by the USDA NRSP-8 program. The Power BI and Tableau softwares were used to create visualization and evaluate the composition of markers in the different maps. Afterwards, the R package GALLO (FONSECA et al., 2020) was used for the extraction of QTL from the Animal QTL database, which were inserted in Tableau to contrast the position of these with the previously identified markers. 5 markers were selected for being close to a QTL, and from these a VCF (Variant Call Format) file was simulated and inserted in the JBrowse software.

Results

Both Power BI and Tableau were efficient in the visualization of genomic data, allowing the identification of important differences between the maps. The analysis of the QTL database revealed that half (52,83%) of the QTLs identified were related to milk production. For the visualization of allelic content from cattle in JBrowse, it was necessary to convert a file from the commercial genotyping format to the VCF format, since the software does not allow for the insertion of genotypic files commonly used in animal breeding.

Conclusions

The BI tools were efficient in the integration and visualization of genomic data, showing advantages over tools designed specifically for the visualization of genomic data, since they are easy to use and do not require specific file formats.

References

- BUELS, R. et al. JBrowse: A dynamic web platform for genome visualization and analysis. **Genome Biology**, v. 17, n. 1, p. 1–12, 2016.
- FONSECA, P. A. S. et al. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. **GigaScience**, v. 9, n. 12, p. 1–9, 30 dez. 2020.