BMC Bioinformatics

POSTER PRESENTATION

Open Access

SNPlotz: a generic genome plot tool to aid the SNP association studies

Zhi-Liang Hu*, Rex Fernando, Dorian Garrick, James M Reecy

Background

High throughput SNP genotyping has empowered livestock researchers to perform genome wide association studies to discover genes responsible for complex traits or economically important quantitative traits. The analysis of large numbers of genotypes is not only a challenge for statistical data analysis but also presents a tedious job for users to interpret and comprehend the large number of statistics generated. A simple and fast approach is to plot the phenotypic association values of each SNP against their genome locations. The benefit is twofold: (1) certain genome locations may quickly draw
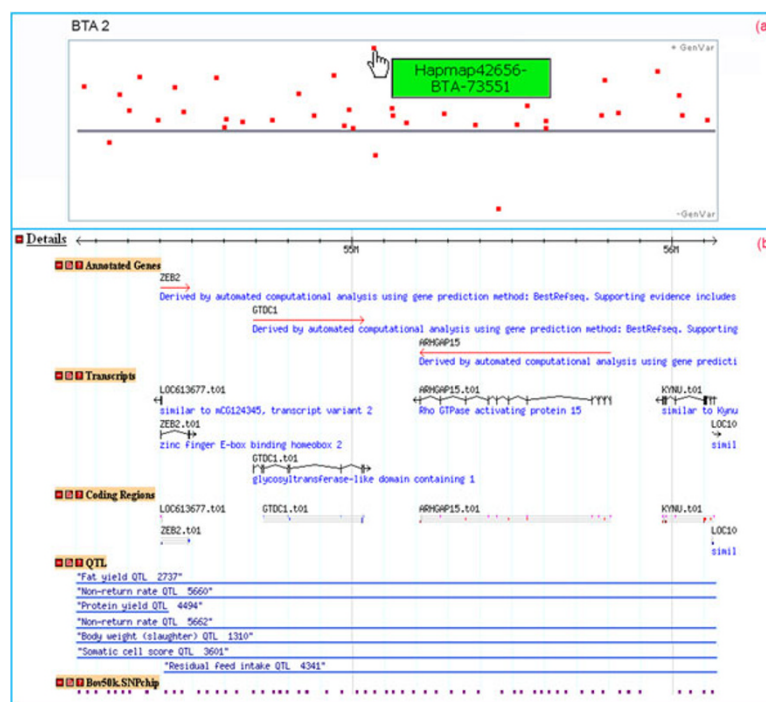
Figure 1 Sample view of SNPLOTz output showing (a) SNP phenotypic value plot on bovine chromosome 2; (b) Interested data point can be linked to GBrowse for alignment with various structural genomics features, such as annotated transcripts and previously mapped QTL etc.

* Correspondence: zhu@iastate.edu
Department of Animal Science and Center for Integrated Animal Genomics,
Iowa State University, Ames, IA 50011, USA

attention if many closely located SNPs are found to show higher degree of associations to traits. (2) It is feasible to locate causal candidate genes when the genome plot is aligned with genomic features such as transcripts, genes, or mapped QTL.

Materials and methods

We have developed an interactive genome plotting tool, SNPLOTz, for SNP association studies to achieve these goals. The software can plot any values associated with individual SNPs, such as estimates of their effects, with respect to their genome locations. The input data used to prime the design of the tool was from cattle 50K and pig 60K SNP chip association studies. The phenotype association analysis data was from Gensel (a software package by Garrick et al.; unpublished data). The tool is designed to take SNP IDs along with any kind of phenotypic association data, find the genomic locations of all SNPs preloaded in the backend MySQL database, and plot them in a two-dimensional graph. The output includes a whole genome plot and individual chromosome plots (Figure 1a). The whole genome plot can give user a quick overall impression, while the chromosome plots are easy for relative location comparisons. Furthermore, the tool also enables dynamic link of each data point to GBrowse [1] to visualize the SNP locations in comparison against other types of genome features such as annotated genes, curated QTL, etc. (Figure 1b). Users can upload as many as 6 data files; each can have as many as 8 columns of phenotype data; either Gensel serial IDs or Illumina SNP IDs may be used. Currently the tool is under further improvement to allow users a private area to store own data for re-use, and to couple with graphic capability of drawing correlation coefficients between any two SNP points for reference. A plan is also on the way to create a Java equivalent program to suit diverse needs of users.

Reference
1. Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S: **The generic genome browser: a building block for a model organism system database.** *Genome Res* 2002, 12:1599-610.