

Flapjack

High Performance Graphical Genotyping

Iain Milne, Micha Bayer, Paul Shaw, Linda Cardle, David Marshall
 SCRI, Invergowrie, Dundee DD2 5DA, Scotland.

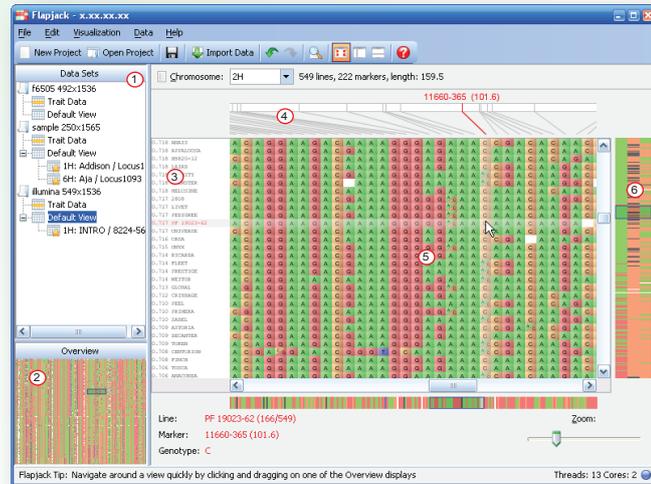


<http://bioinf.scri.ac.uk/flapjack>

Introduction

New software tools for graphical genotyping and haplotype visualization are required that can routinely handle the large data volumes generated by the high throughput Illumina SNP platform and comparable genotyping technologies. Flapjack is a new visualization tool to facilitate analysis of these data types. Its visualizations are rendered in real-time allowing for rapid navigation and comparisons between lines, markers and chromosomes.

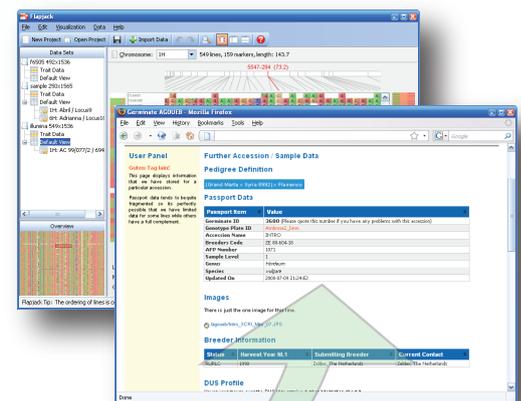
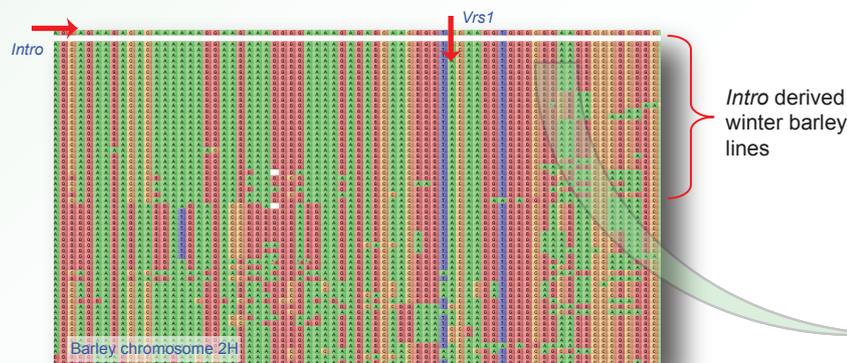
Flapjack is available for Windows, OS X, Linux and Solaris.



- 1 Navigation of multiple simultaneous data sets
- 2 Whole chromosome overview
- 3 Phenotypic traits heat-map
- 4 SNP distribution on genetic map
- 5 Primary genotype display area
- 6 Dynamic whole-marker (and its neighbours) display area

Sorting to identify common haplotypes

Sorting relative to a given line for a given chromosome region can be used to identify co-inherited haplotypes. In the figure below, we have sorted by the chromosomal region around the *Vrs1* locus to show the chromosome 2H haplotypes in winter barley varieties that have inherited the deficiens allele at the *Vrs1* locus from the cultivar *Intro*.



Germinate (see poster 786) displaying additional information on *Intro*. Flapjack can query any database that supports a simple key=value URL query, allowing for comprehensive data on line and markers to be called up at any time.

Colour models

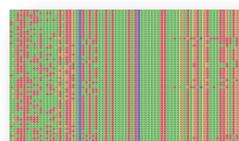
Flapjack provides a number of alternative graphical genotype views with individual alleles coloured by state, frequency or similarity to a given standard line.



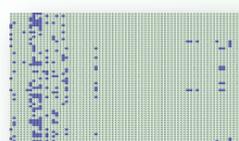
Homozygous and heterozygous representation



Instantly highlight all heterozygotes



Colouring by line (or marker) similarity displays the selected line in a darker colour, then colours each allele on every other line based on whether it is the same (red) or not (green) as the comparison line.



The nucleotide colour scheme assumes that the data contains alleles that are of the form A, C, G, or T. Each base is assigned its own colour, and these colours apply to both homozygous and heterozygous alleles.

Colouring by allele frequency will use a specified frequency threshold to display rare alleles and haplotypes within a marker in one of two colours - those below the threshold (blue), and those above it (green).