iReport: HTML Reporting in Galaxy

Center for Translational Molecular Medicine

Saskia Hiltemann (1,2), Youri Hoogstrate (1,2), Hailiang Mei (3), Peter van der Spek (2), Guido Jenster (1), Andrew Stubbs (2)

1)Dept. of Urology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands; 2) Dept. of Bioinformatics, Erasmus MC, Rotterdam, The Netherlands
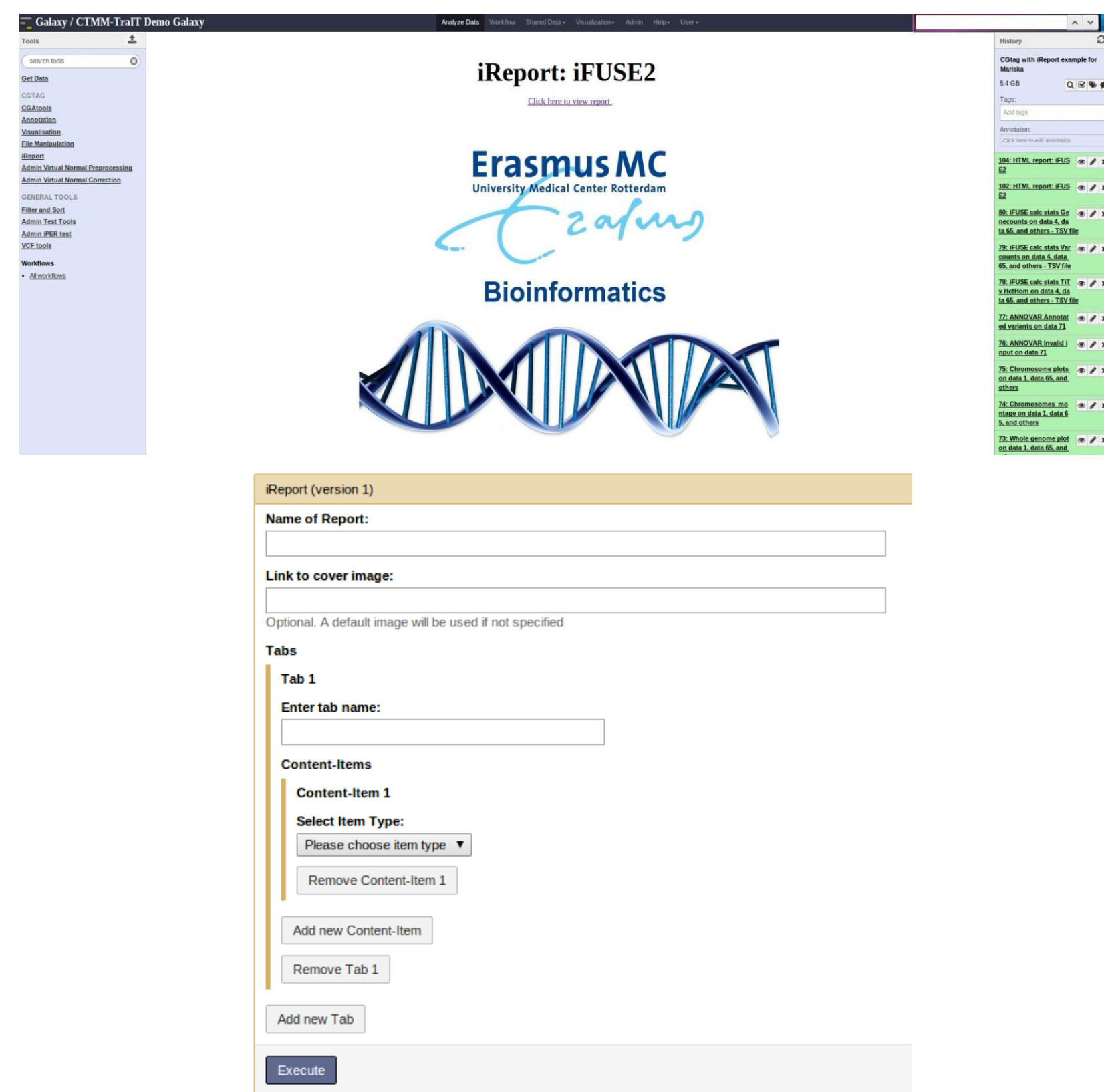3) Leiden University Medical Center (LUMC), Leiden, The Netherlands

iReport

TraIT Galaxy server: galaxy.ctmm-trait.nl
Example iReport (published history): tinyurl.com/llrzz9w

Basic structure

Cover page
User-defined title and cover image (url)

Tab-based report
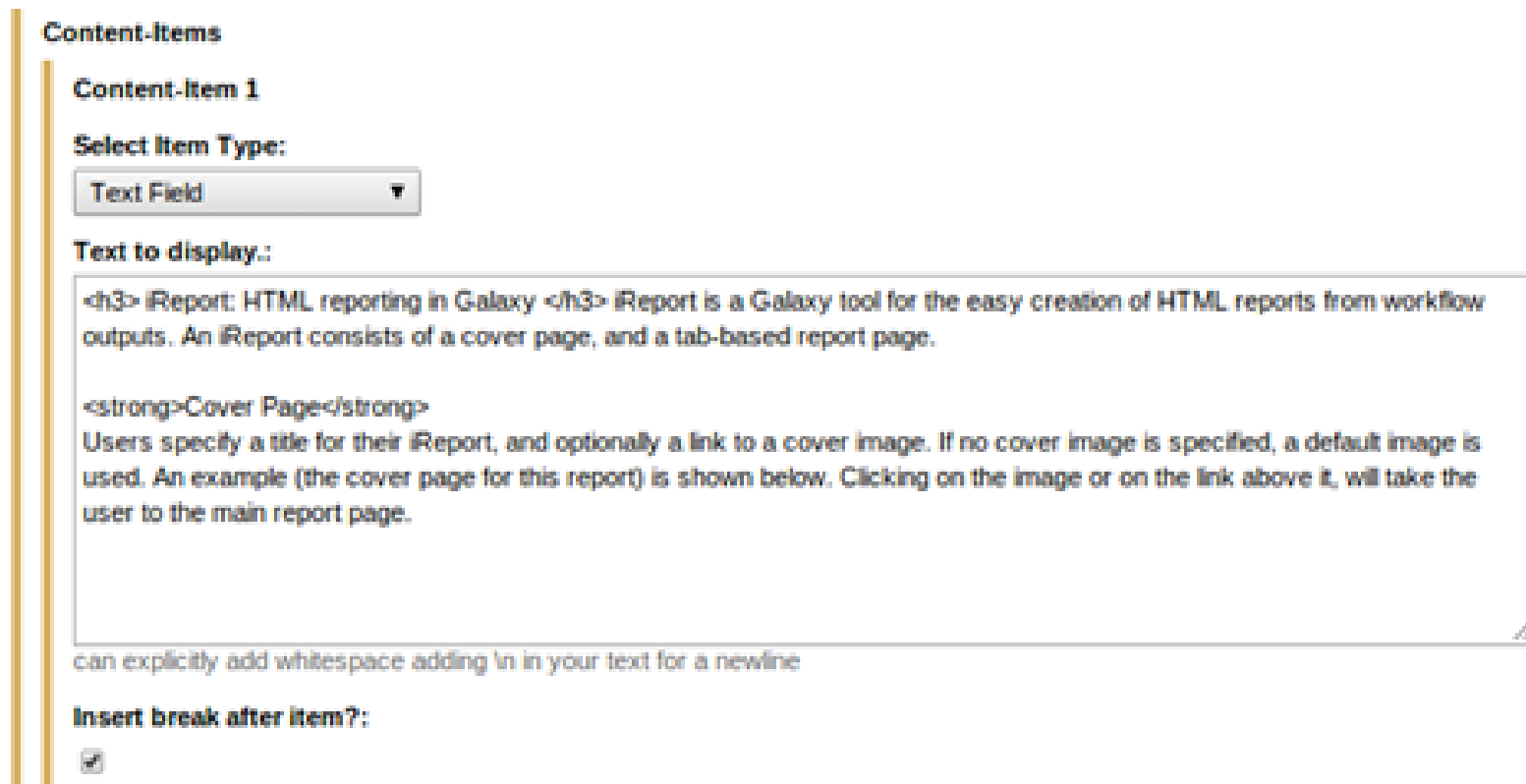Add any number of tabs, then add any number of *content items* to each tab.

Content items can be one of:
- Text
- Image
- Table
- PDF
- Link

Content Items

Text
Specify a text file from history, or enter text directly into a text field.
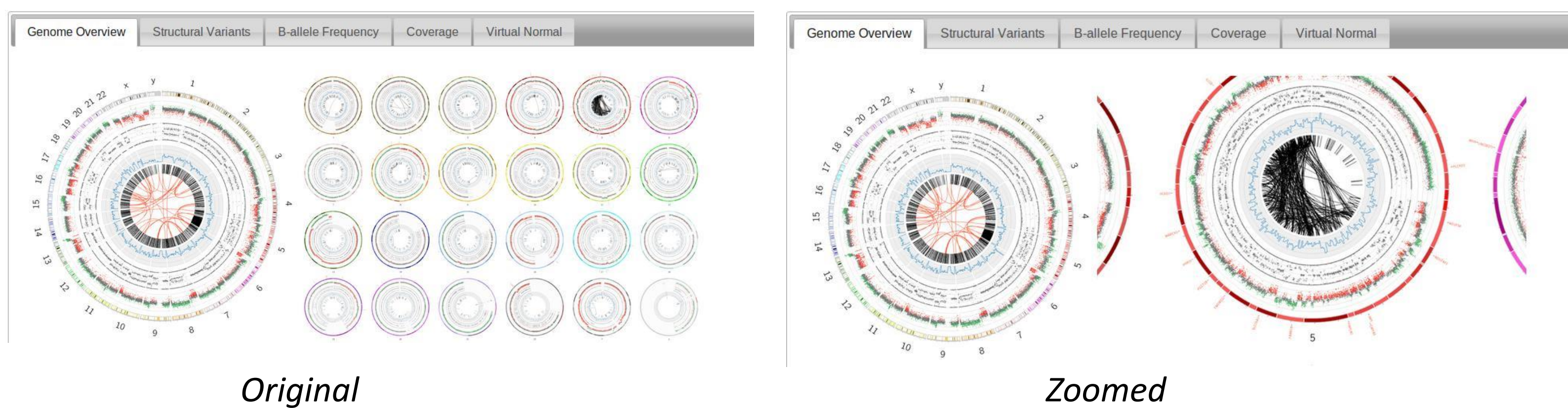

Input is sanitized and printed verbatim, with the exception of a few html tags (h1-h6, strong, em)

\n and \t interpreted as newline and tab respectively to allow for some control over whitespace

Image
Specify image file from history, and an image width (0 for original size).
If image is scaled down, a jQuery *zoom-on-mouseover* effect is added to the image

Original   *Zoomed*

Table
Specify tab delimited file from history
jQuery library *datatables* (http://datatables.net) used to create paginated, searchable, and sortable tables (if desired)

Can also generate weblinks from column values by specifying url prefix and suffix
(e.g. from column with gene names, create links to gene cards by prepending fixed url to all cells)

PDF
Specify PDF file from history.

Depending on browser settings this will be shown in page itself or a link will appear to download or view file in new window
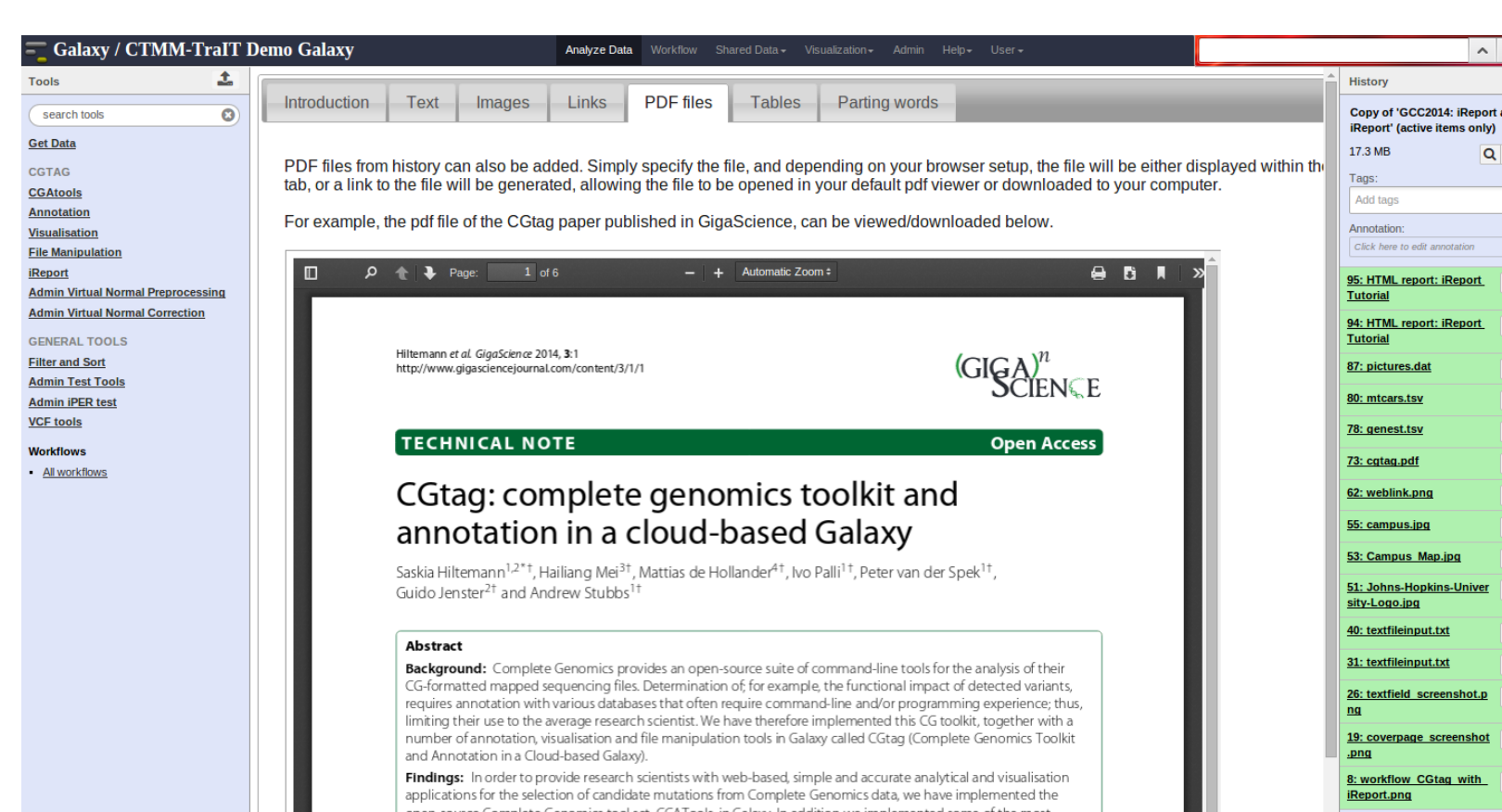
Links
Links to web addresses, files from history, or all files in an archive in the history can also be created

The CTMM TraIT Project

TraIT will develop a long-lasting IT infrastructure for translational medicine that will facilitate the collection, storage, analysis, archiving, sharing and securing of the data generated in the CTMM operational translational research projects.

The project builds on existing expertise to create an IT infrastructure that will help to accelerate the translational research in the Dutch Life Sciences and Health sector.

An example: iFUSE2

CGtag Galaxy workflow for downstream analysis of Complete Genomics Data (Hiltemann et al, *GigaScience*, 2013) generates many outputs.

iFUSE (Hiltemann et al, *Bioinformatics*, 2012) is a web-based tool for exploration, visualisation and prioritization of candidate fusion genes, and has since been ported into Galaxy.

iFUSE2 combines these two pipelines with newly incorporated copy number variation methods (OncoFUSE) and presents the results in an iReport

Overview page containing Circos plots and tables with summary statistics
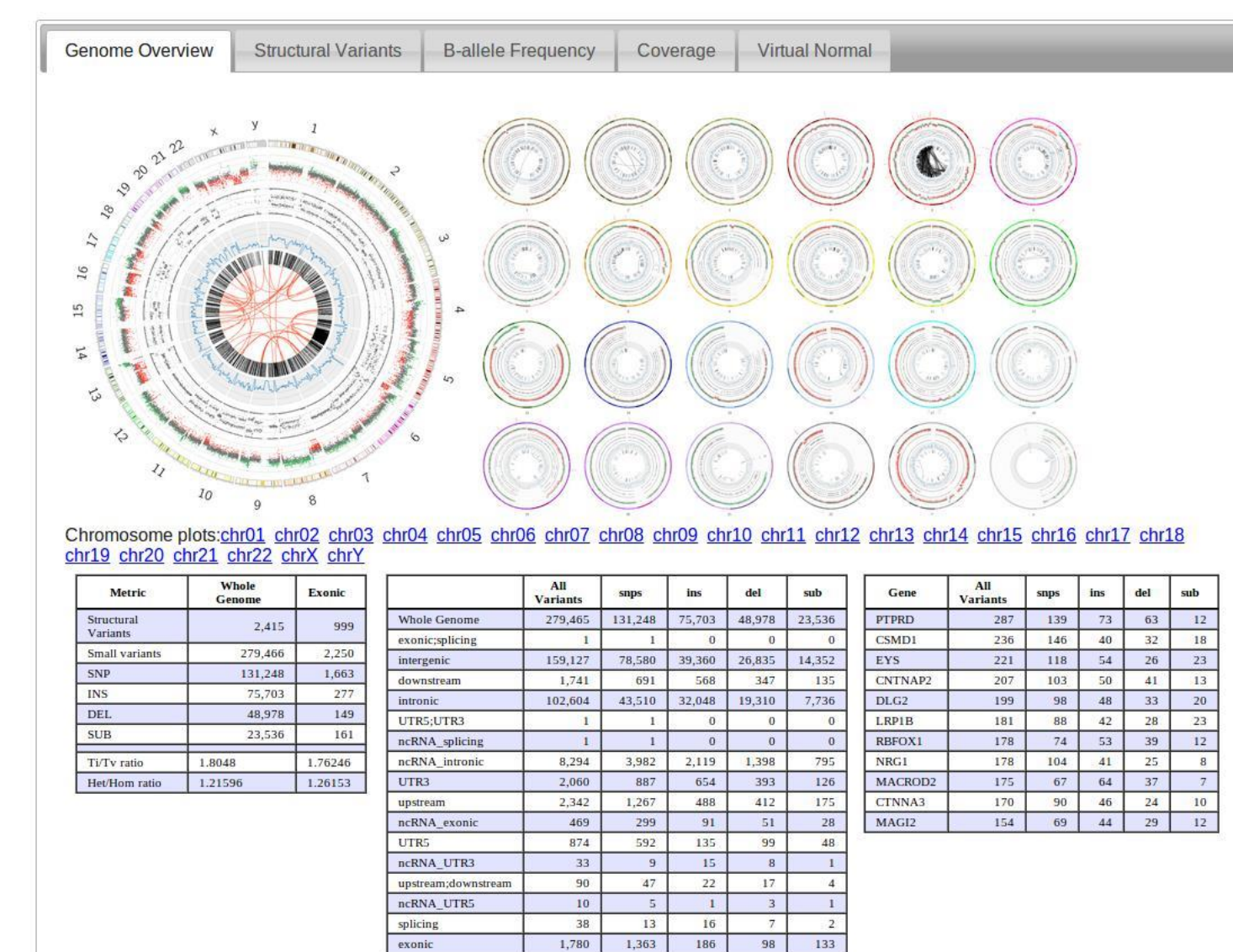

Tabs for Structural Variants (fusion gene candidates), Small Variants, Copy Number Variations (e.g. OncoFUSE), SNP array data, and more
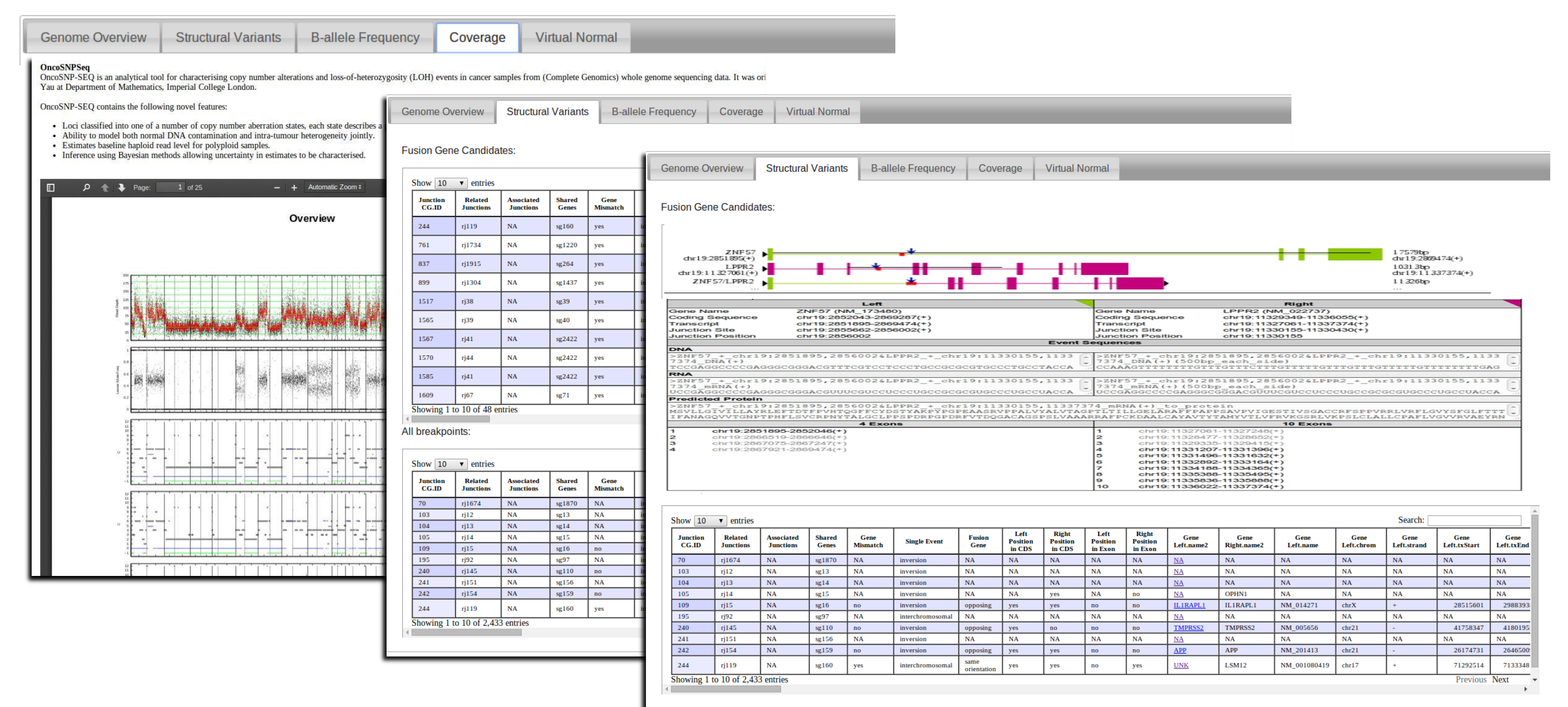
Tools used: CGATools, Circos, GNUplot, R, OncoFUSE, OncoSNPSeq

TraIT Galaxy HPC CLOUD Architecture

The rapid evolution of NGS technologies together with decreasing cost are creating a challenge to store and analyze the vast amount of sequencing data that are generated by experimental biologists. Configuring suitable data analysis software and having access to readily available computation and storage are the two major bottlenecks faced by many research groups
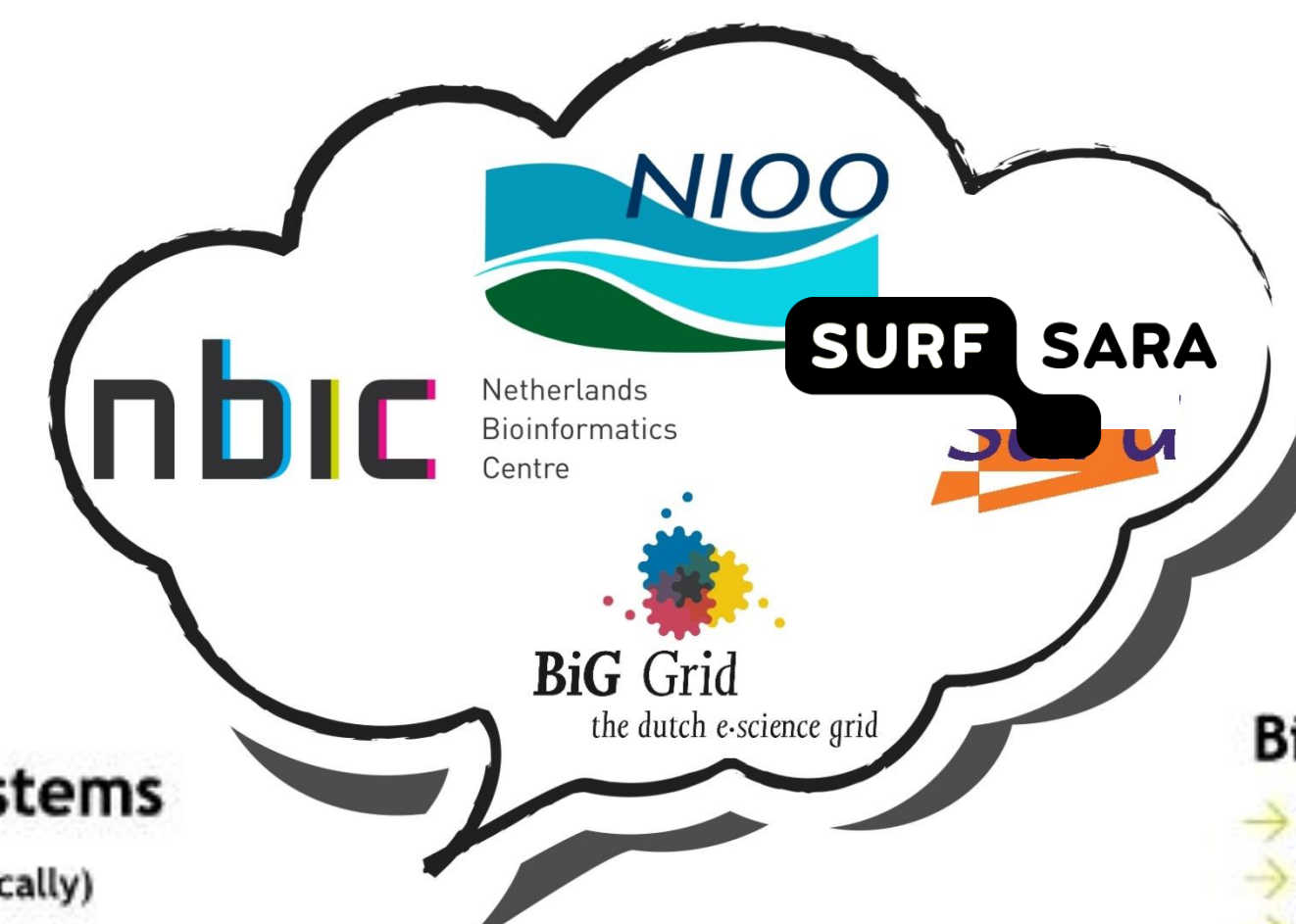
Advantages of Cloud Systems
- Rapid elasticity (scale up/down dynamically)
- Full administrational rights
- Perfect for project-based research
- Access to powerful compute systems

BiGGrid & SARA HPC Cloud (Calligo)
- 19 x Intel Xeon 32 core processors (608 cores)
- 19 x 256GB RAM (7.75 TB RAM)
- 400TB shared storage