

TGAC Browser: visualisation solutions for big data in the genomic era

Anil S. Thanki¹, Xingdong Bian¹, Robert P. Davey¹

1. The Genome Analysis Centre, Norwich Research Park, Norwich NR4 7UH, UK

Email: Anil.Thanki@tgac.ac.uk

Demo: <http://tgac-browser.tgac.ac.uk>

Source Code: <https://github.com/tgac/tgacbrowser>

License: GPL v3

We present the TGAC Browser with novel rendering, annotation and analysis capabilities designed to overcome the shortcomings in available approaches. TGAC Browser, being a web-based client, utilises JavaScript libraries to provide a fast and intuitive genome browsing experience. We focus on harnessing Internet architectures as well as localised HPC hardware, concentrating on improved, more productive interfaces and analytical capabilities.

- **User-friendly:** Live data searching, track modification, and drag and drop selection; actions that are seamlessly powered by modern web browsers
- **Responsiveness:** Client-side rendering and caching, based on JSON fragments generated by server logic, helps decrease the server load and improves user experience
 - TGAC Browser visualises genomic data in different ways, based on the type and amount of data, which is more informative to the user and memory efficient.
- **Analysis Integration:** The ability to carry out heavyweight analysis tasks, using tools such as BLAST, via a dedicated extensible daemon
- **Annotation:** Users can edit annotations which can be persisted on the server, reloaded, and shared at a later date
- **Off-the-shelf Installation:** The only prerequisites are a web application container, such as Jetty or Tomcat, and a standard Ensembl database to host sequence features
- **Extensible:** Adaptable modular design to enable interfacing with other databases, e.g. GMOD
- **Data format:** TGAC Browser processes and visualises data directly from the Ensembl core schema as well as next-generation sequencing (NGS) data output, i.e. BAM/SAM, BigWig/wig, GFF, and VCF.