Creating a pluggable visualisation toolsuite with BlueGenes Tool API

Yo Yehudi [1], Adrián Bazaga [1,2], Akshat Bhargava [1], Daniela Butano [1], Matthew Chadwick [1], Justin Clark-Casey [1], Sergio Contrino [1], Joshua Heimbach [1], Vivek Krishnakumar, Rachel Lyne [1], K.H. Reierskog [1], Julie Sullivan [1], Gos Micklem [1]
[1] Department of Genetics, University of Cambridge, Cambridge, United Kingdom, [2] STORM Therapeutics Ltd, Cambridge, United Kingdom

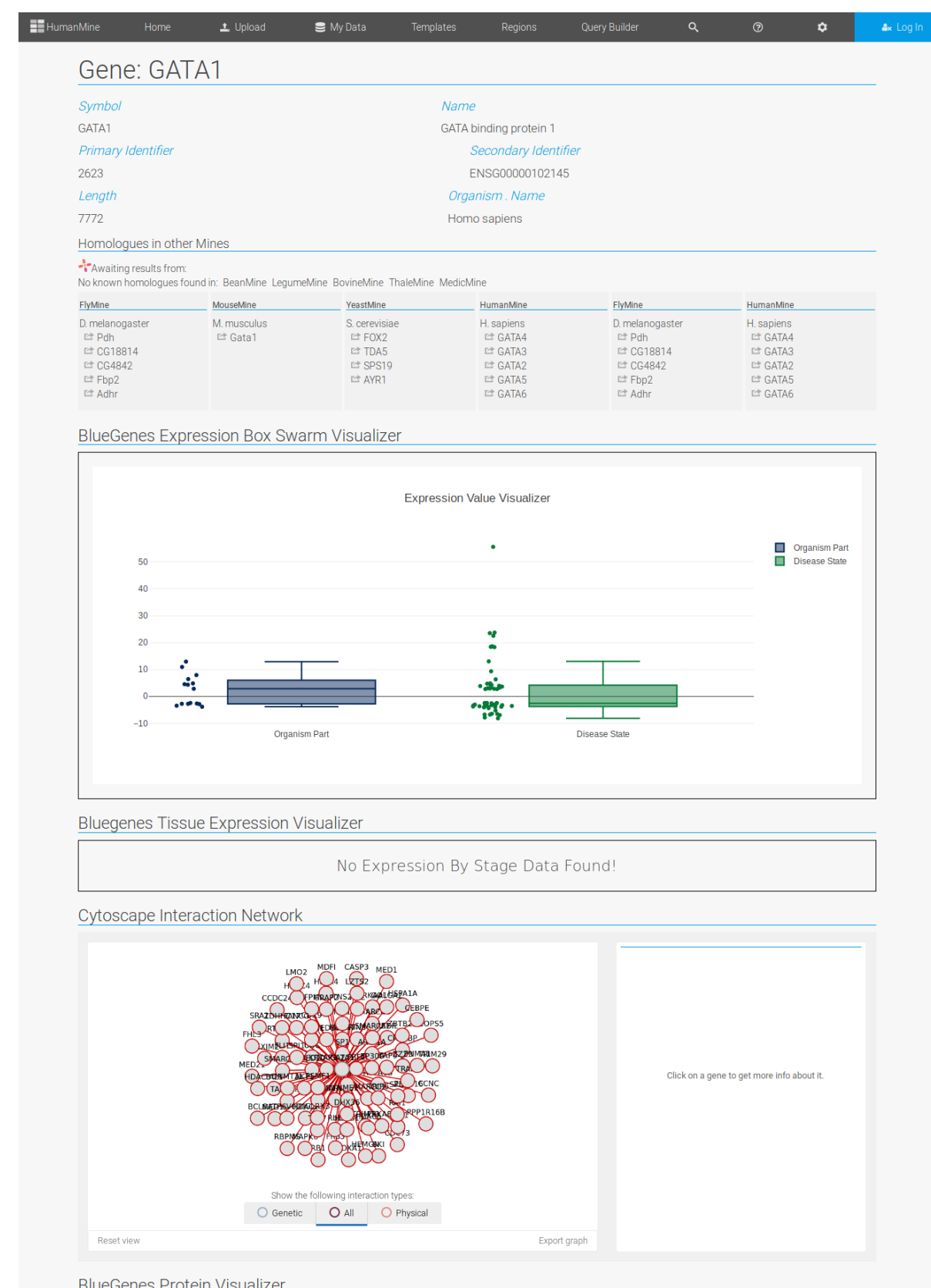GitHub: intermine   Twitter: @intermineorg   http://www.intermine.org
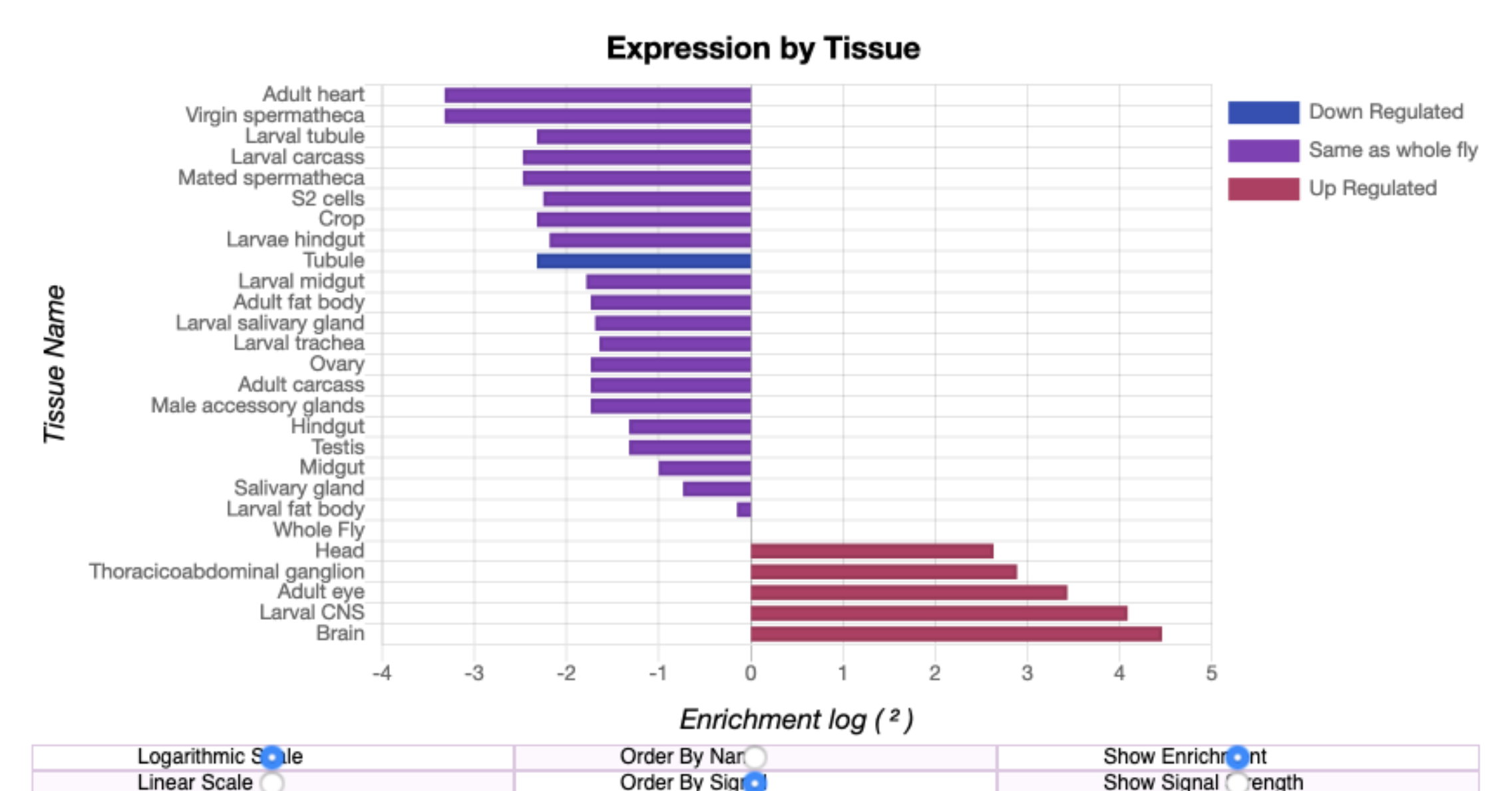
BlueGenes & InterMine

Background:
InterMine is an open source data warehouse, designed to make it easy to merge disparate data sources and query them both programatically and/or via a user interface. Dozens of InterMines spanning a broad range of organisms can be found at registry.intermine.org

BlueGenes is a new user interface for InterMine databases, written in Clojure and ClojureScript. You can try it at bluegenes.apps.intermine.org

Clojure and ClojureScript are flexible, allowing rapid prototyping, succinct functional code, and one language can compile to the Java Virtual Machine (Clojure) while identical or near-identical code can also be compiled to JavaScript for the browser (ClojureScript), saving time and effort.
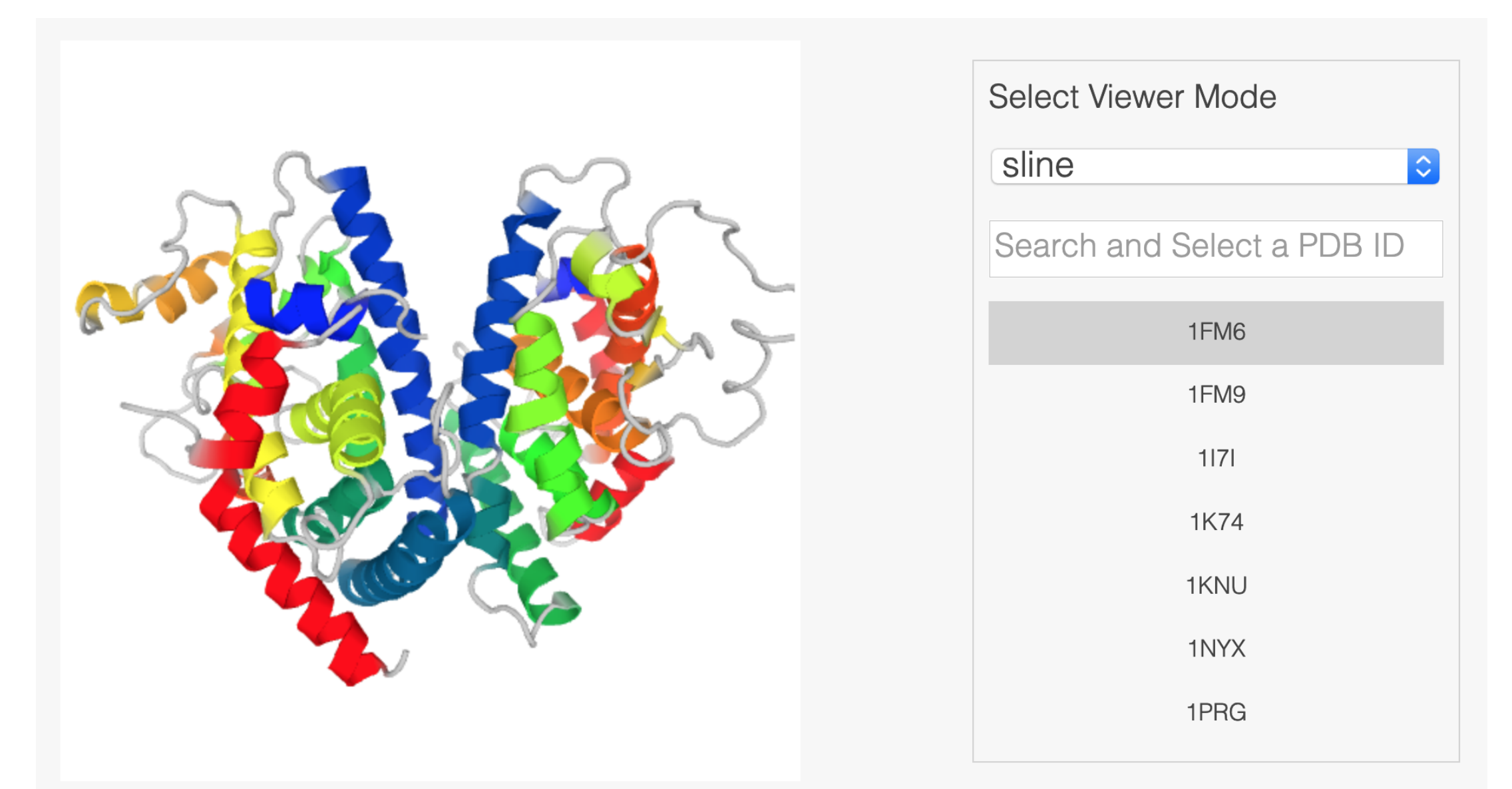

The BlueGenes Tool API

The Challenge:
Whilst Clojure has plenty of advantages, not everyone knows Clojure, and people may wish to:

 - **Re-use** others' open source libraries and JavaScript browser-based tools, such as BioJS components.
 - Write these tools in a **familiar language** (JavaScript, or perhaps something that compiles to JS, such as CoffeeScript or TypeScript)
 - People who already run their own InterMine may wish to **retain their existing visualisation tools** in the new UI.

Our Solution:
ClojureScript is completely compatible with JavaScript applications, so we created a specification that allows developers to easily make their JavaScript-based tools BlueGenes compatible. It has a clearly defined file structure and method signature that all BlueGenes compatible-tools must provide, including the object type (e. g. Gene, Protein), and unique identifier (e.g. the InterMine ID for BRCA1).

Building Your Own Tool

While the specification is clear (and strict), this isn't a barrier to creating your own tool, or making an existing JavaScript tool BlueGenes compatible. We provide a Yeoman generator, which will step a prospective tool creator through a short series of questions and auto-generate most of the boilerplate files, create stub methods, and set up a test suite and node server task. With the help of this generator, a standards-compliant JavaScript visualisation tool may only need a few lines of custom code to become BlueGenes compatible.

https://github.com/intermine/generator-bluegenes-tool/

Examples

Tissue expression graph for the *D. melanogaster* **ey gene**
Implemented by Akshat Bhargava as part of Google Summer of Code 2019. Data from FlyAtlas [1].

https://github.com/intermine/bluegenes-tissue-expression-visualizer

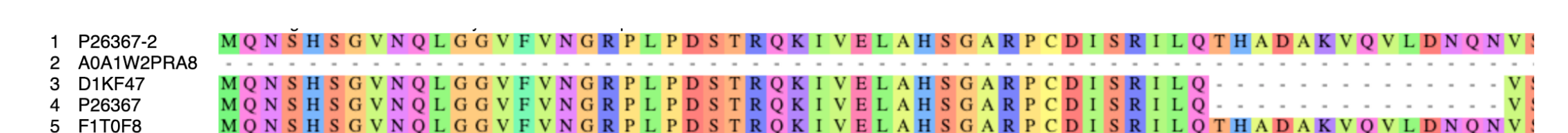**Protein 3D structure visualiser for the gene** *PPARG* **[2] in** *H. sapiens*
This tool uses InterMine to fetch all proteins associated with a given gene, then fetches protein stucture information directly from PDB [3]. The tool itself is a thin data-fetching layer over bio-pv [4], a BioJS compatible protein visualisation library.

https://github.com/intermine/bluegenes-protein-visualizer

BioJS multiple sequence alignment viewer for *H. sapiens* **PAX6 proteins**
This tool is a thin wrapper around a fork of MSAViewer [5], a BioJS compatible multiple sequence alignment tool. The sequence is truncated to fit on this poster. Data from UniProt [6].

https://github.com/intermine/bluegenes-msa-viewer

With thanks to...

References
[1] Robinson SW, Herzyk P, Dow JA, Leader DP. FlyAtlas: database of gene expression in the tissues of Drosophila melanogaster. Nucleic Acids Res. 2013 Jan;41(Database issue):D744-50. doi: 10.1093/nar/gks1141. Epub 2012 Nov 29. PubMed PMID: 23203866; PubMed Central PMCID: PMC3531048. flyatlas.org
[2] Gampe RT Jr, Montana VG, Lambert MH, Miller AB, Bledsoe RK, Milburn MV,Kliewer SA, Willson TM, Xu HE. Asymmetry in the PPARgamma/RXRalpha crystalstructure reveals the molecular basis of heterodimerization among nuclearreceptors. Mol Cell. 2000 Mar;5(3):545-55.
[3] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne.(2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242. rcsb.org
[4] Marco Biasini, bio-pv. (2015, July 21). pv: v1.8.1 (Version v1.8.1). Zenodo. http://doi.org/10.5281/zenodo.20980
[5] Guy Yachdav et al. "MSAViewer: interactive JavaScript visualization of multiple sequence alignments." Bioinformatics (2016)
[6] UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47: D506-515 (2019)