

Manual for the Blueprint Progenitors Web Application

<http://blueprint.haem.cam.ac.uk>

A companion to the paper:

“Transcriptional diversity during lineage commitment of human blood progenitors”.

1. Why use the web application:

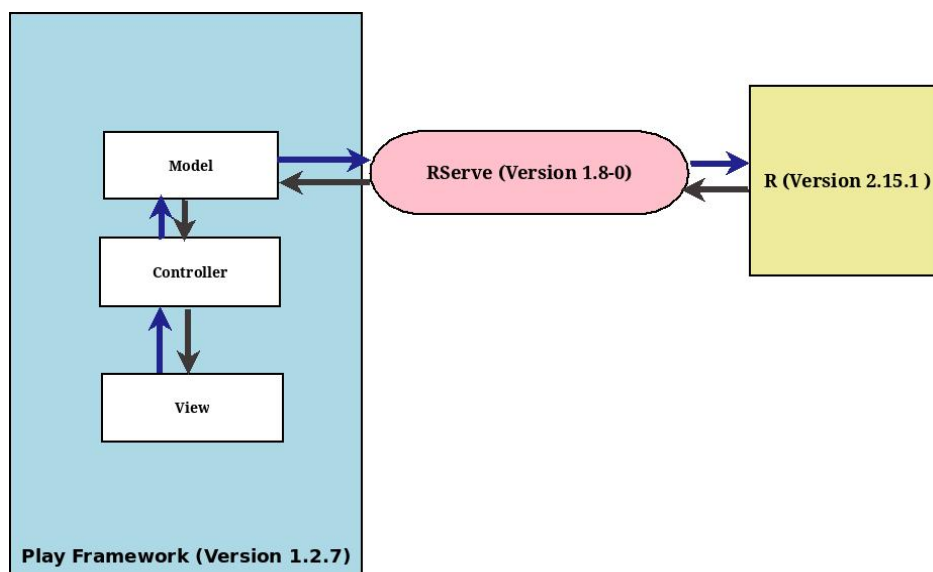
The application provides three different options for the visualisation of the relative expression of all genes and transcripts in the Ensembl v.70 annotation.

Input: The application accepts HGNC (e.g. ITGA2B) and Ensembl (e.g. ENSG00000005961) gene names, as well as, Ensembl transcript identifiers (e.g. ENST00000262407).

Output: The application outputs plots in jpeg format that are downloadable.

2. Work flow and packages:

The web application was built with the open source Play Framework v.1.2.7 (www.playframework.com/), which uses a Model-View-Controller architecture. The view component receives input from the user and sends it through the controller and model to R (<http://cran.r-project.org/>) via RServe v.1.8-0 (<http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/Urbanek.pdf>).

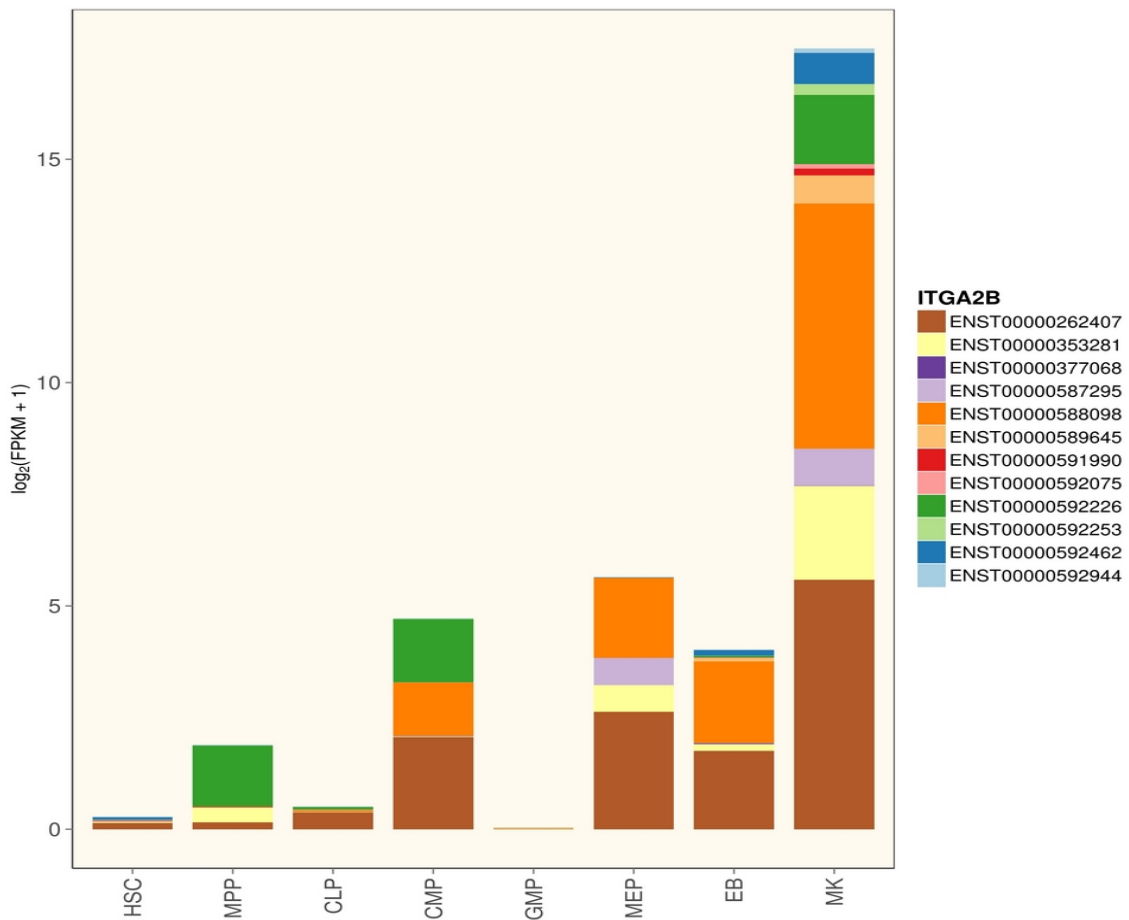


3. How to use the website:

The website contains two main sections: data and tools. The data page contains the supplementary tables to our paper. The tool page provides three different visualisation methods for genes and transcripts: bar plots, river plots and heat maps. The application queries an R data frame with the gene and transcript expression data, and uses the `RColorBrewer`, `reshape`, `gplots` and `ggplot2` R packages for visualizing the data.

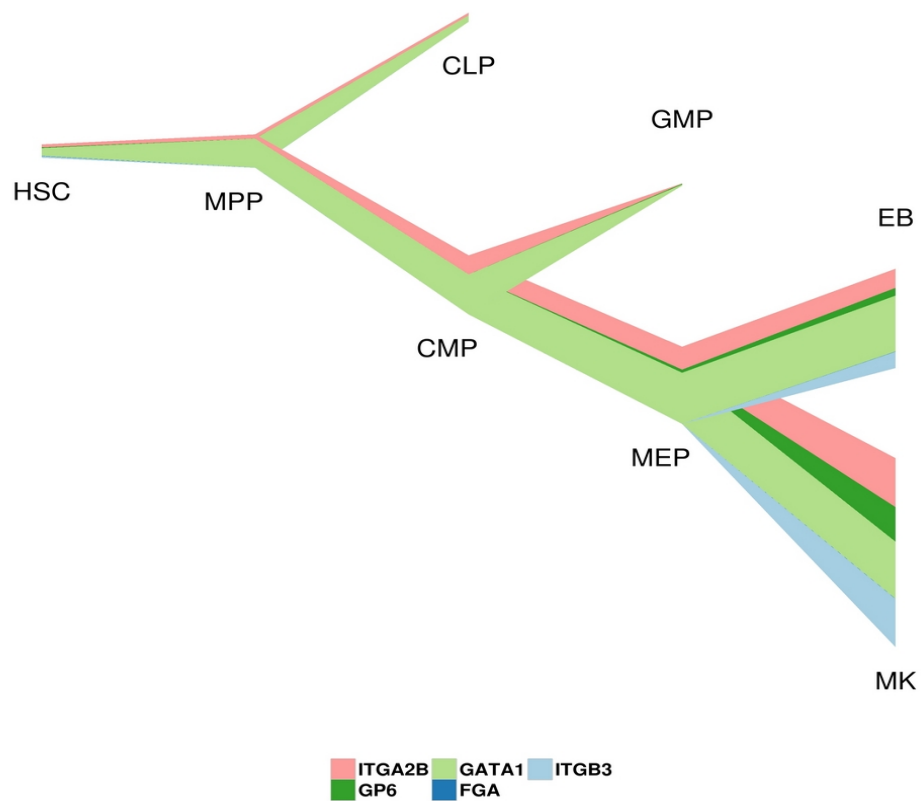
3.1. Bar plots

Bar plots are available for a single gene or transcript.



3.2. River plots

A river plot shows the relative expression across eight cell types. River plots are available for a maximum of 12 genes or transcripts.



3.3. Heat maps

A heat map shows the relative expression across eight cell types in a matrix of coloured cells. Heat maps are available for a list of genes or transcripts. For a single gene, a heat map of all its transcripts will be displayed. For multiple genes, the annotation type determines if the expression of those genes or of their transcripts is displayed.

Heat maps can be visualized under different settings by selecting 'Cluster heat map' or 'Scale heat map'. The clustering option groups rows (with genes/transcripts) according to their expression values using the `heatmap.2` function in the `gplots` R package. The scaling option scales and centres the expression values. If a single gene name is input but the transcript button is selected, the heat map will display the expression of all transcripts of that gene.

