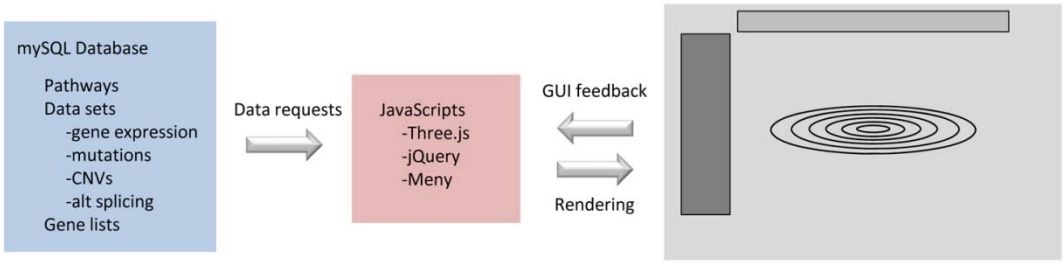
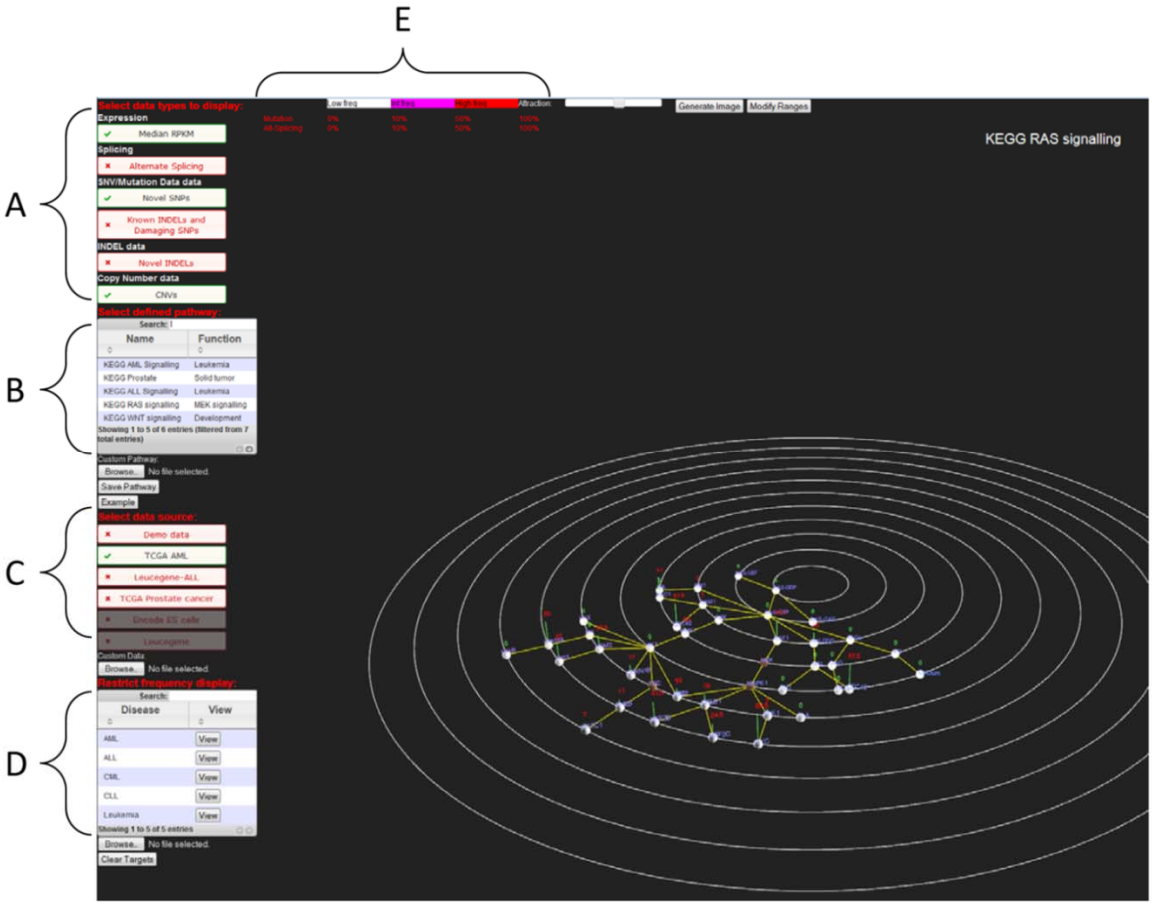
# Cascade – A novel tool for exploring multidimensional RNA-seq data

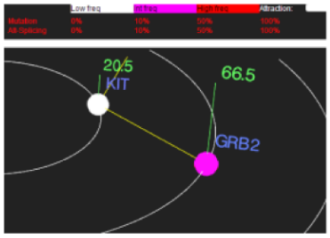
Cascade is a software tool which allows the representation and the exploration of multidimensional RNA-seq data, and specifically cancer genomics data. The object of the software is to reduce the dimensionality of RNA-seq data into a single intuitive view that builds on existing biological knowledge. Cascade consists of a my SQL database used to store experimental data and a main interface written in JavaScript which uses the three.js JavaScript 3D library for rendering of known biological pathways.



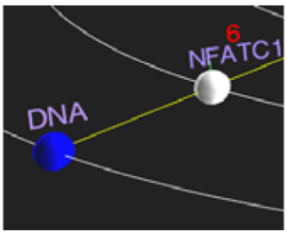

Known biological pathways are displayed as a series of connected nodes, where each node represents a gene. The dimensionality of the RNA-seq is achieved by using node colour to represent mutation frequencies, vertical bars with values attached to each node to show average gene expression (RPKM) values and shape and Z-axis position changes to highlight alternative splicing and copy number variations respectively. The thresholds used for colour/shape changes for nodes are user tunable through a menu option on the main screen. Users interact with Cascade using a space saving menu on the left-hand side to select features of the RNA-seq data to be displayed (A), pre-defined or custom biological pathways to view (B), specific datasets to use for visualization of features selected (C). Additionally, predefined or custom disease gene lists (D) can be used to restrict the colouring thresholds defined by the used (E). The “modify ranges” button (top centre) allows users alter the thresholds required for node colour changes (mutations) or ring appearance (splicing).



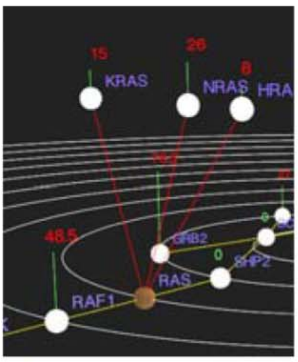
# Examples of node colouring and elements in Cascade



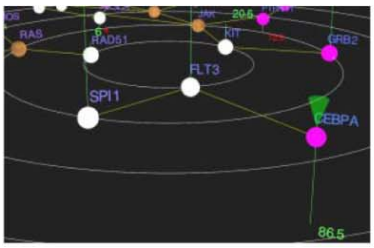

A gene (GRB2) with a mutation frequency between 10 and 50% in selected data set and an average RPKM value of 66.5.




Non-genic elements represented as blue nodes. Genes with outliers in dataset for gene expression have average RPKM values shown in red




Families of genes represented as brown nodes; Members can be expanded by clicking on parent node




Copy number changes shown by green (loss) or red (gain) cones with node below or above plane respectively

# Gene specific information

Clicking on specific nodes will bring up a vertically tabbed menu with information for: Generic description of gene retrieved dynamically from GeneCards, and sample specific information for gene expression (with mutated samples shown in red), mutations, alternative splicing and copy number variations. Samples can be sorted by column values within tabs.