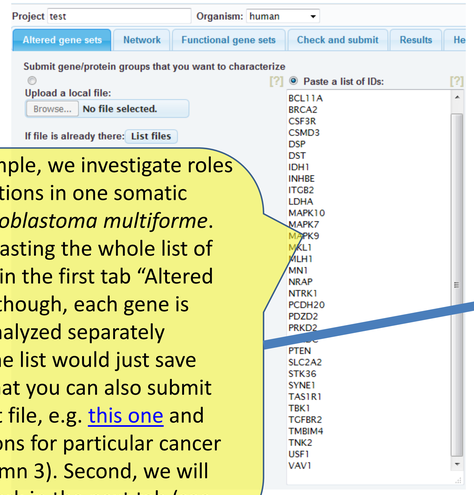
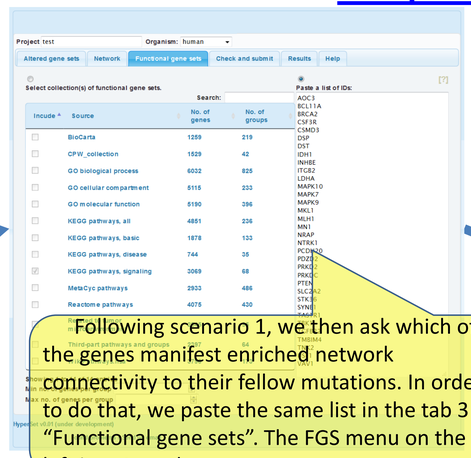
How to evaluate driver roles of candidate disease genes?

https://www.evinet.org



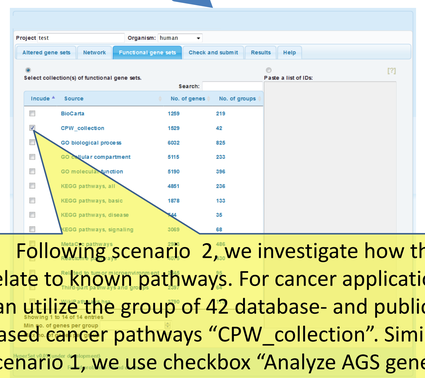
As an example, we investigate roles of point mutations in one somatic genome of *glioblastoma multiforme*. We start by pasting the whole list of 34 mutations in the first tab "Altered gene sets". Although, each gene is going to be analyzed separately (submitting the list would just save time). Note that you can also submit them as a text file, e.g. this one and select mutations for particular cancer genome (column 3). Second, we will select a network in the next tab (see details in tutorial "How to begin?").



Following scenario 1, we then ask which of the genes manifest enriched network connectivity to their fellow mutations. In order to do that, we paste the same list in the tab 3 "Functional gene sets". The FGS menu on the left is not used.

The idea of network-based functional analysis is in assuming that multiple genes involved in the disease would appear connected in the global network. Hence we should formulate gene-specific requests so that the answers confirm or reject respective hypotheses. The analysis is most applicable to cancer mutations (because of their abundance), but can well be applied to other diseases where candidate genes were derived in GWAS etc.

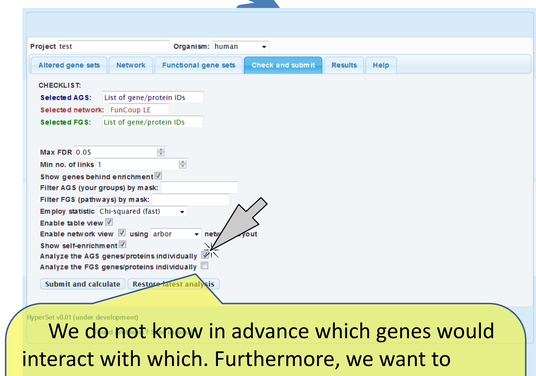
Example applications can be found in Merid et al., 2014 (cancer driver genes) and Bennet et al., 2011, Hong et al., 2010, Reynolds et al., 2010 (Alzheimer's disease).



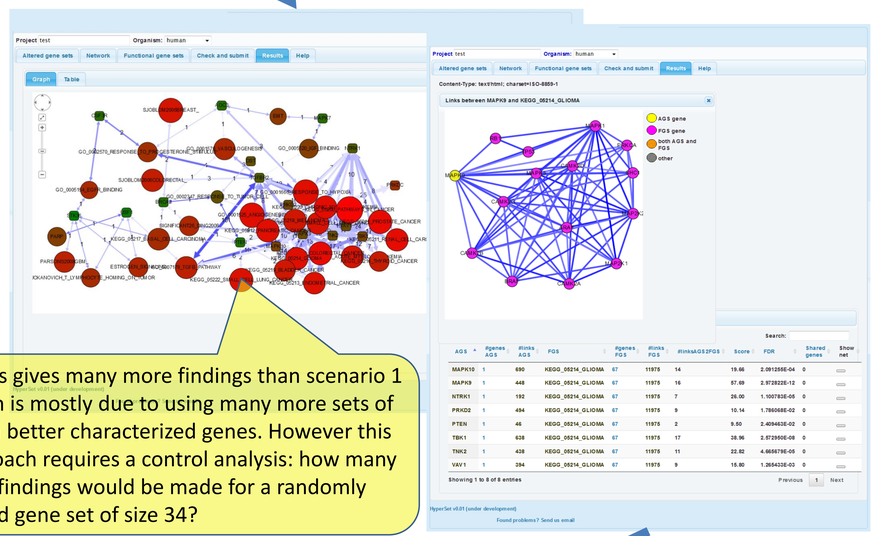
Following scenario 2, we investigate how the genes relate to known pathways. For cancer applications one can utilize the group of 42 database- and publication-based cancer pathways "CPW_collection". Similarly to scenario 1, we use checkbox "Analyze AGS genes / proteins individually" at the next tab.

There are two major modes in the analysis that can be run separately or in parallel. We can evaluate connections between a given candidate gene and a group of genes that ***together*** are likely to be implicated in the disease. Such a group can be either:

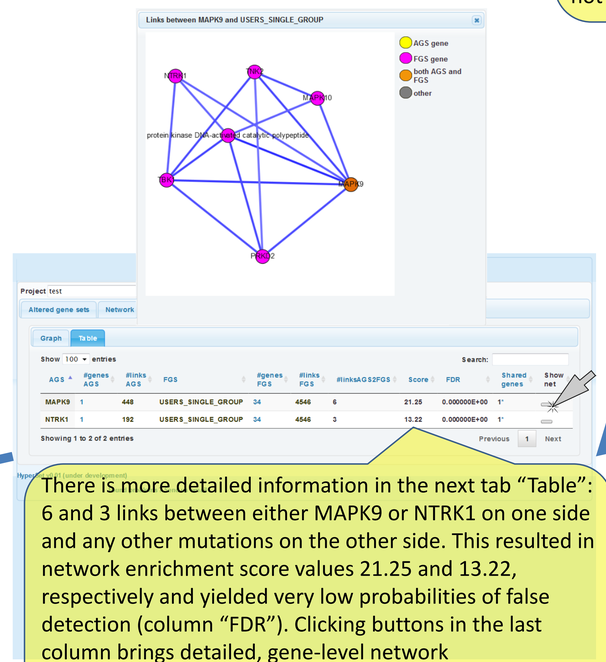
1) a set of altered genes discovered experimentally (typically the whole set of mutations, genetic variants, or differentially methylated or expressed genes), or
2) a curated gene set with well characterized functional role in the disease.



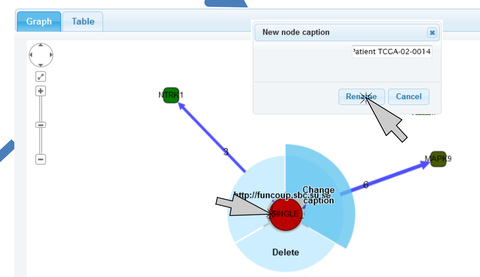
We do not know in advance which genes would interact with which. Furthermore, we want to evaluate individual genes' relations to ***the whole*** set of mutations. For these reasons, we check the box "Analyze AGS genes/proteins individually" and ***do not check*** the other box (or vice versa - it does not matter in this case).



This gives many more findings than scenario 1 which is mostly due to using many more sets of much better characterized genes. However this approach requires a control analysis: how many such findings would be made for a randomly picked gene set of size 34?



There is more detailed information in the next tab "Table": 6 and 3 links between either MAPK9 or NTRK1 on one side and any other mutations on the other side. This resulted in network enrichment score values 21.25 and 13.22, respectively and yielded very low probabilities of false detection (column "FDR"). Clicking buttons in the last column brings detailed, gene-level network representations of each line.



We see that only two genes appear significantly connected to the fellow mutations (in the given network). For presentation purposes one can rename nodes here.

Thus, stronger evidence was accumulated in the both scenarios for MAPK9 and NTRK1 and a few other genes scored high against multiple cancer pathways (e.g. TNK2, TGFBR2). See another analysis of the same mutation list using other parameters.

However a full-scale statistical framework should be implemented for a systematic analysis of multiple cancer genes and genomes. Examples can be found in Merid et al., 2014 and might require additional information from the same samples (methylation, copy number events etc.). Required software in R and/or perl can be downloaded from https://www.evinet.org