

canEvolve Help Documents:

This section includes instruction on how to use canEvolve database functions. If you need further assistance, please do not hesitate to contact with canEvolve team at help {at} canevoive [dot] org

Help document version: 1

Date: June 4, 2012

Differential Gene Expression Function:

Differential gene expression analysis looks for genes whose expression changes between two or more different groups. With canEvolve DGE function, users can query genes between two classes.

Usage:

- 1) Open "Diff. Gene Expr." Section from left menu.
- 2) Select one of the cancer type from drop down menu and click "Get Studies" button.
- 3) You will see the dataset those have DGE analysis results on the "Studies" section. Please select the dataset that you are interested.
- 4) Enter the gene symbol(s) that you want to query or if you are interested with pathway please select MSigDB defined pathway from drop down list and click "Get Pathway Members" button. You should add 1 gene symbol per line. (canEvolve compatible with HUGO gene symbols.)
- 5) Once you complete your settings click "Get Results" button to get the results and heatmaps.

Differential miRNA Expression Function:

Differential miRNA expression analysis looks for miRNAs whose expression changes between two or more different groups. With this canEvolve function users can query miRNAs between two classes.

Usage:

- 1) Open "Diff. miRNA Expr." section from left menu.
- 2) Select one of the cancer type from drop down menu and click "Get Studies" button.
- 3) You will see the dataset those have differential miRNA expression analysis results on the "Studies" section. Please select the studies that you are interested.
- 4) Enter the miRNA symbol(s) that you want to query. (canEvolve compatible with miRbase symbols)
- 5) Once you complete your settings click "Get Results" button to get the results and heatmaps.

Regulatory Networks:

This canEvolve function has been developed based on ARACNE (<http://wiki.c2b2.columbia.edu/califanolab/index.php/Software/ARACNE>) tools. ARACNE, a novel algorithm, using microarray expression profiles, specifically

designed to scale up to the complexity of regulatory networks in mammalian cells, yet general enough to address a wider range of network deconvolution problems. With regulatory network function, users can visualize gene based regulatory networks.

Usage:

- 1) Open “Regulatory Network” section from left menu.
- 2) Select cancer type and click “Get Studies” button
- 3) Select a study from right list
- 4) Enter gene symbol (You can only use one gene symbol each query) that you are interested
- 5) Click “Show Aracne Network” button to visualize the regulatory network based on your gene symbol. (Visualization process may takes a minute)

WGCNA Networks Function:

The WGCNA network function has been developed based on WGCNA package. The authors of the study describe the WGCNA with following words, “Weighted correlation network analysis (WGCNA) can be used for finding clusters (modules) of highly correlated genes, for summarizing such clusters using the module eigengene or an intramodular hub gene, for relating modules to one another and to external sample traits (using eigengene network methodology), and for calculating module membership measures.” Here with canEvolve users can access the pre-analysed WGCNA results for different cancer datasets.

Usage:

- 1) Open “WGCNA Networks” section from left menu.
- 2) Select cancer type that you are interested and click “Get Studies” button.
- 3) Select a dataset from right list
- 4) Use “Show WGCNA Results” button to access analysis results and figures.

Survival Analysis Function:

Survival analysis is used to understand prognostic differences between classes. The canEvolve uses gene expression data with time and event information. The query output includes the both “Cox” and “log-rank” survival analysis results and K-M plots.

Usage:

- 1) Open “Survival analysis” section from left menu
- 2) Select a cancer type and click “Get Studies” button
- 3) Select a study that you are interested from right list
- 4) Enter a gene symbol (You can only use one gene symbol each query)
- 5) Click “Show Survival Results” button

Gene Set Enrichment Function:

For gene set enrichment analysis, canEvolve use GSEA tool from Broad Institute (<http://www.broadinstitute.org/gsea/index.jsp>). Broad Institute define GSEA as a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states.

Usage:

- 1) Open “Gene Set Enrichment” section from left menu.
- 2) Select a cancer type and click “Get Studies” button
- 3) Select a dataset from right menu.
- 4) Click “Show GSEA Results” button
- 5) The canEvolve dataset stores GSEA results based on 6 different collections. You may download the results, check the snapshots, or access enrichment results in a different format such as HTML or Excel file.

Copy Number Function:

It is known that copy number alterations has a role in the cancer progression. With this function, users can query the copy number of SNPs based on gene symbols. If the dataset includes 2 or more class users also see the SNP copy number based on class information.

Usage:

- 1) Open “Copy Number” section from left menu.
- 2) Select cancer type and click “Get Studies” button
- 3) Select a dataset from right menu
- 4) Enter gene symbol(s) that you are interested (Enter one gene symbol per line)
- 5) Click “Get Results” button

GemiNi Function:

GemiNi is a recently published method for integrating gene and miRNA expression profiles in cancer using feed-forward loops (FFLs) consisting of transcription factors (TFs), miRNAs and their common target genes. The dChip-GemiNI (Gene and miRNA Network-based Integration) method statistically ranks computationally predicted FFLs by their explanatory power to account for differential gene and miRNA expression between two biological conditions such as normal and cancer. GemiNI integrates not only gene and miRNA expression data but also computationally derived information about TF-target gene and miRNA-mRNA interactions. Please check our paper if you need more information about method (<http://www.ncbi.nlm.nih.gov/pubmed/22645320>)

Usage:

Please visit dChip-GemiNi web site.

Copy Number & Gene Expression Integration Function:

Copy number and gene expression data are extensively used in genetic research. However, the analysis methods specifically used for these data have their own limitations. Such limitations may be related with either the nature of the data or the methods applied. It is believed that integrative analysis methods used for both data types would be crucial to overcome the limitations. With this function, canEvolve users can query results based on gene symbol and to access the correlation value and FDR.

Usage:

- 1) Open “Copy number & Gene Expr.”
- 2) Select cancer type and click “Get Studies” button

- 3) Select a dataset from right menu
- 4) Enter a gene symbol(s) that you are interested
- 5) Click "Get Results" button to show the results. You can also show results about dataset that you select

Protein-Protein Interactions Function:

This section has been developed based on String database resources. (<http://string-db.org/>) We filtered the Homo sapiens data only.

Usage:

- 1) Enter gene symbol that you are interested (This section allows you to use only one gene symbol per query)
- 2) Set connection power (max: 999 - min:1) (We suggest that to start from higher connection power)
- 3) Click "Show Protein-Protein Interaction" button to visualize the interactions