***ADGO 2.0* User’s Guide**

ADGO 2.0 is a newly developed web server that extends previous tentative server ADGO1 in analysis methods and coverage. This new version provides biological interpretations for both microarray data and lists of genes (proteins) using composite annotations. Main new features include

A. It provides multiple gene-set analysis methods as well as gene list analyses

B. Users can upload their own gene sets (e.g. predicted miRNA targets) to generate and analyze new composite sets.

C. It incorporates union and subtracted sets as well as intersection sets

D. It screens redundant composite annotation sets when generating and prioritizing those composite annotations.

**How to Use**

***Front Page***

1. Choose a species among the eight organisms supported
2. Upload data or paste the data into the panel. Then click ‘Upload’ to move to the next step.
   1. Microarray dataset or gene (protein) list is to be uploaded.
   2. Three file formats are acceptable. The first two are the formats of a microarray dataset for gene set analysis and the last format is for the gene list analysis. The three acceptable file formats are shown as follows:
      1. **[Array id – single value]**: For each gene (or probe set) ID, a single value such as fold change or t-statistic value is assigned.

C:\Users\Dougu Nam\Desktop\ip1.tif

* + 1. **[Array ID – multiple values]**: For each gene (or probe set) ID, multiple values for the two sample groups compared are assigned. The number of samples in each group should be specified later.
    2. **[Array ID]**: Only a list of gene (or protein) IDs without values.

C:\Users\Dougu Nam\Desktop\lkt.tif

***Input Options***

C:\Users\Dougu Nam\Desktop\ip.tif

1. **Array Type:** Both dual and single channel IDs are supported for the gene set analysis, but only dual channel IDs are supported for the gene list analysis. The dual channel IDs include ‘gene symbol’, ‘Ensemble’, ‘Refseq’, ‘Entrez’, ‘Uniprot’, and ‘Systematic name (yeast)’, and the single channel IDs include the probe set IDs of Affymetrix, Illumina and Agilent microarray chips.
2. **Take log 2:** If the values in the input data are needed to take a log transformation, click this option.
3. **Analysis Method**: ADGO 2.0 provides four methods for a gene set analysis of microarray data and two methods for a gene list analysis. Depending on the input file format (i, ii, and iii), only the applicable methods are displayed. For type i, Z-test2 and gene permutation methods5 are available. For type ii, all the four gene set analysis methods (Z-test, gene permutation, sample permutation5, and GSEA3) are provided. For type iii, the two gene list analysis methods (Fisher’s exact test and hypergeometric method) are available.
4. **Annotation Gene Sets**: Choose annotation gene sets by which you want to interpret your data. ADGO 2.0 currently provides from four to seven types of annotation categories depending on the species chosen (For Homo sapiens, three categories in the Gene Ontology, KEGG, chromosomal location, cis-regulatory motifs (TFBS), and OMIM gene sets are available). In addition, the user can upload their own collection of gene sets and then ADGO 2.0 generates *ad hoc* composite gene sets between the user’s gene set and the built-in gene sets chosen to interpret the input data.
5. **Gene Set Size** limits the number of members in an annotation gene set
6. **Type of Composite Gene sets** indicates different types of (composite) gene sets to be considered in your analysis such as single, intersection, difference, or union between a pair of gene sets.
7. **Filtering Composite Sets:** Composite gene sets are generated by taking intersection, unions, and differences of two gene sets. Each composite set is compared with existing single gene sets and will be screened if the new composite set is overlapped with any single set over the specified percentage.
8. **Filtering Significant Composite Sets:** provides two options to screen significant composite sets. In the ‘Strong Type’ option, a significant composite set is removed if an individual gene set used to generate the composite set has smaller p-value than the composite set. In the ‘Weak Type’ option, we display a significant composite set only if it has a smaller p-value than the individual gene sets.
9. For type ii format, an additional option will appear about sizes of the two sample groups.
10. Press the ‘**Analysis**’ button to obtain the result.

***Analysis Results***

The analysis results are also downloadable as text files.

**References**

1.Dougu Nam et al. ADGO: analysis of differentially expressed gene sets using composite GO annotation, *Bioinformatics*, **22**, 2249-2253 (2006)

2. S. -Y. Kim and D. J. Volsky, PAGE: parametric analysis of gene set enrichment, *BMC*

*Bioinformatics*, **6**: 144 (2005).

3. A. Subramanian *et al*. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, *PNAS*, **102** (43) 15545-15550 (2005).

4. L. Tian et al. Discovering statistically significant pathway in expression profiling studies, *PNAS*, **102** (38), 13544-13549 (2005).