Performing GO ANOVA

Preparing a data set for analysis requires importing the data, normalizing the data as appropriate for standard gene expression analysis, and inserting columns containing the experimental variables. Checkout our tutorial page for more details about preparing data. It is not necessary to perform a differential analysis of gene expression before GO ANOVA.

For the sake of example, the following walkthrough will consider an experiment that has been imported which includes two different tissues, brain tissue and heart tissue, extracted from a small set of patients.

The GO ANOVA function is available in the Gene Expression, microRNA Expression, RNA-Seq, and miRNA-Seq workflows.

- Select the Gene Expression workflow (or any of the other ones) from the Workflows drop-down on the upper right of the spread sheet
- Go to the **Biological Interpretation** section of the workflow
- Select Gene Set Analysis (Figure 1) and then Gene Set ANOVA
- Workflows
 Gene Expression

 Gene Expression
 X

 Import
 X

 QA/QC
 Analysis

 Visualization
 X

 Biological Interpretation
 Gene Set Analysis

 Gene Set Analysis
 Pathway Analysis

 Genomic Integration
 MicroRNA Integration

Figure 1. GO ANOVA dialog can be invoked via Gene Set Analysis option of the workflow

For this example analysis, the model was kept easy to interpret by including *Subject* and *Tissue* as the only ANOVA factors. Additionally, Tissue was added to the Disruption Factor(s). Including *Subject* controled for person to person variation, and including *Tissue* allowed the analysis of differential expression and of functional category disruption between tissue types. For the sake of simplicity and minimizing run time, the term *Subject* was not added to the *Disruption Factor(s)* box. Including it would have helped correct for subject specific gene expression patterns, though the results were largely unaffected in this case.

Performing GO ANOVA analysis on very large GO categories can take quite a bit of time. More importantly, very large categories may have too large a scope to be useful. To speed the operation and analyze only smaller GO categories, specify 20 genes as the maximum size for an analyzed GO category.

For the sample dataset, the GO ANOVA dialog setup should appear as in Figure 2 below.

| GO ANOVA of Spreadsheet 1 | | | X |
|--|---|---|-----------------------------------|
| While ANOVA finds differentially expressed genes, GO | ANOVA will identify differentially expressed functional g | oups of genes based on Gene Ontology de | signations. |
| Experimental Factor(s) | ANOVA Factor(s) | | Disruption Factor(s) (optional) 🕜 |
| 5. Type 6. Tissue 7. Subject | Add Factor > ? 6. Tissue 7. Subject | Add Factor > | 6. Tissue |
| 9. Scan Date | Add Interaction > 🕜 | < Remove Factor | |
| | < Remove Factor | | |
| Save Model Load Model | Contrasts | Advanced | |
| Exclude Skipped Probe Sets | | | |
| Exclude probe sets with Max v signal | < 3 | | |
| only exclude probe sets with differential expression p-value(s) > 0.05 | | | |
| Specify Output File C:/UserGuides/GOANOVA/ | GOANOVA | | Browse |
| | | | OK Cancel |
| | | | |

Copyright © 2018 by Partek Incorporated. All Rights Reserved. Reproduction of this material without express written consent from Partek Incorporated is strictly prohibited.

Figure 2. GO ANOVA configured for the user guide data set. Two factors added to the model

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.

