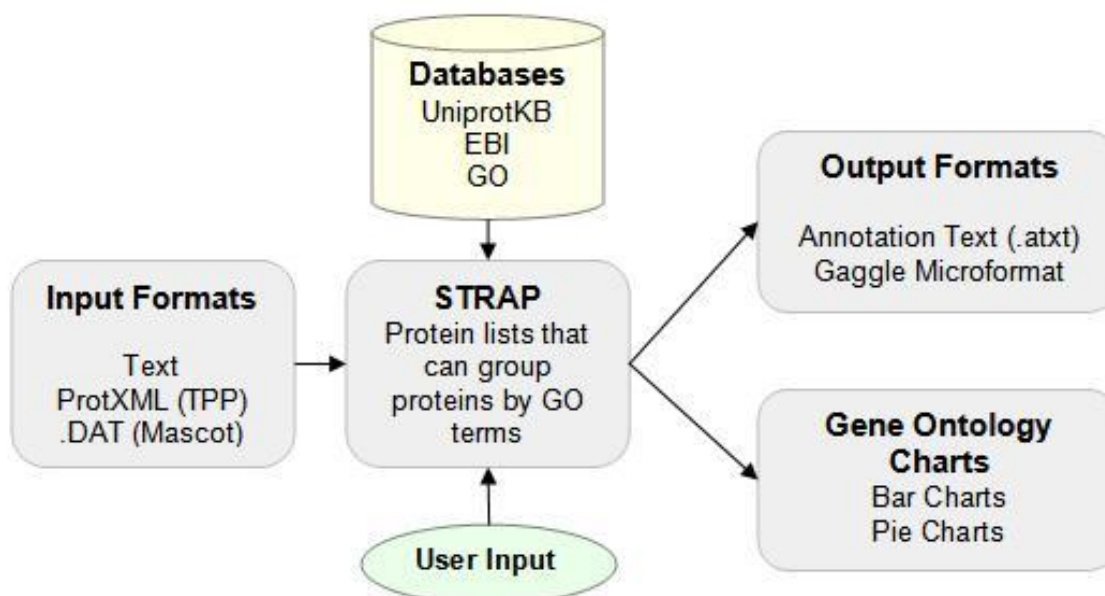# STRAP Tutorial

## A Guide to Using the Software Tool for Rapid Annotation of Proteins

STRAP saves you time by automatically annotating a protein list with information that helps you meaningfully interpret your mass spectrometry data. This tutorial will teach you how to use it.



### Installation

STRAP runs on Windows XP or higher and requires the Microsoft .NET Framework. The STRAP program and sample files can be downloaded as zip files (*STRAP.zip* and *STRAP-Sample-Files.zip*) from CPC Tools. There are directions on the Download Page for installing STRAP and for uncompressing the directory *STRAP Sample Files*. Installation of STRAP should automatically install the latest version of the .NET Framework.

### Opening and Saving Protein Lists

The starting point for any analysis with STRAP is a protein list in a variety of forms, usually originating from search results on raw spectra from mass spectrometry

experiments. Sample files are provided in the *STRAP Sample Files* directory for exploring the capabilities of STRAP. The sample data set originates from proteins that were overexpressed in each of three different B-cell proliferative states from a lymphomagenesis biomarker study.

STRAP can open protXML files generated from the Trans-Proteomic Pipeline (TPP), DAT files from Mascot searches, and plain-text files containing one UniProtKB protein name per line. To open the sample data, go to File->Open, select "Protein List (*.txt)" as the file type, and select set1.txt, set2.txt, and set3.txt.

All protein information is retrieved from the UniProt database, but additional gene ontology (GO) annotation can be imported from the EBI database by clicking Edit->Add Annotations from EBI.
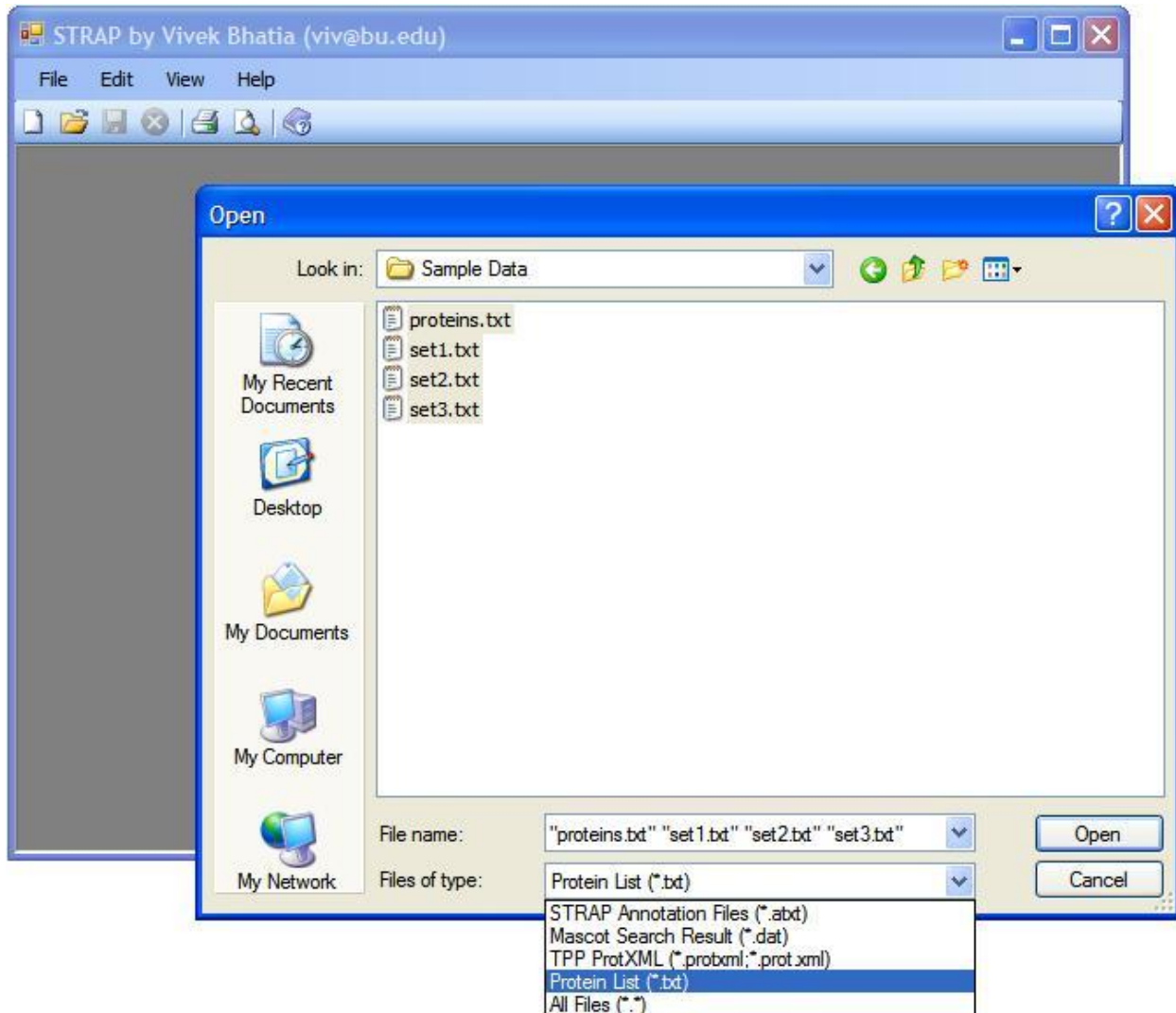
After pressing the Open button, STRAP will open three tabs and retrieve information about the proteins in your lists over several minutes.



Since it takes time for STRAP to annotate all proteins, it's important that you save these lists so that you don't waste time downloading annotations every time you want to look at a data set. Once the protein annotation table is completely filled, you can save it as a STRAP Annotation Text File (*.atxt).

If you need to close the currently open tab, right click on it, and click "Close Active Tab." You can also click the "X" button in the toolbar to achieve the same thing.

## The Protein Annotation Table

The protein annotation table has many categories. Each category can be sorted if you click on its header at the top of the window. The columns in STRAP's annotation table include columns for general protein information (e.g., gene name) as well as for GO terms. The GO categories used in STRAP are a simplification of those released by the Gene Ontology Consortium.

Clicking on the link in the Uniprot ID column will take you to the protein's entry in the UniProtKB database. For cases in which you'd like to create a custom FASTA database containing only proteins in a particular data set, you can export a FASTA file by clicking on File->Export to FASTA.

The sort feature can also be used to group proteins with similar GO terms:
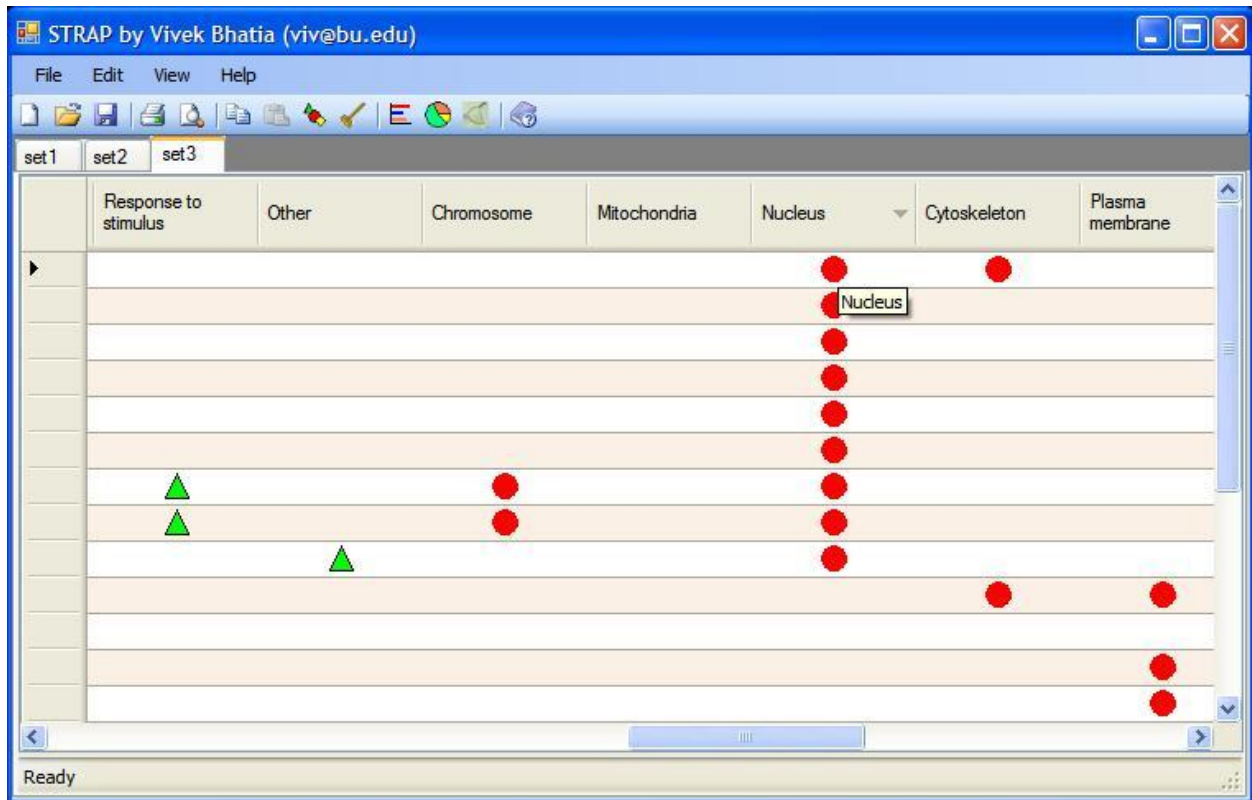


If you'd like to manually change a protein's annotation, you can do so by double-clicking the corresponding GO category cell to toggle its value:
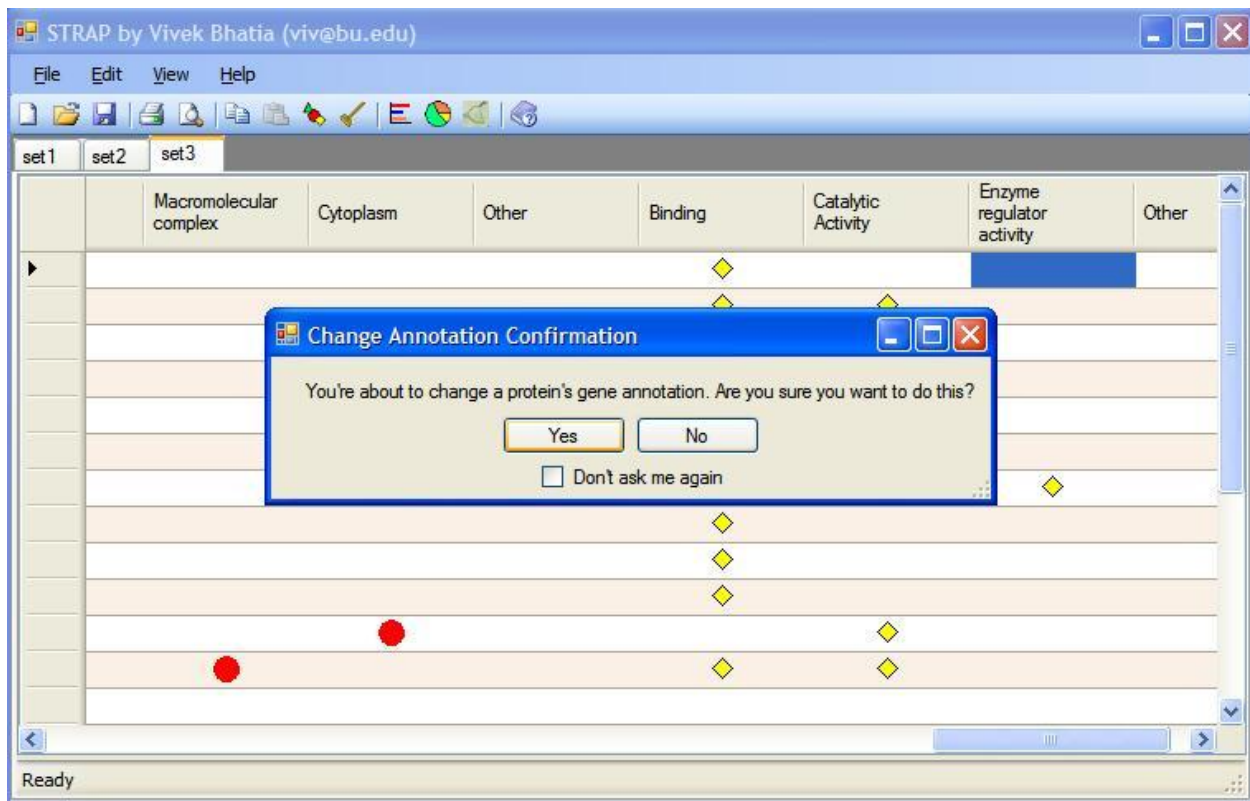
If you're curious about the lineage of a particular protein's GO terms, you can investigate further through the Gene Ontology Term Browser. This browser can be accessed by right-clicking on any protein and pressing "Browse GO terms..." to display a window such as this:
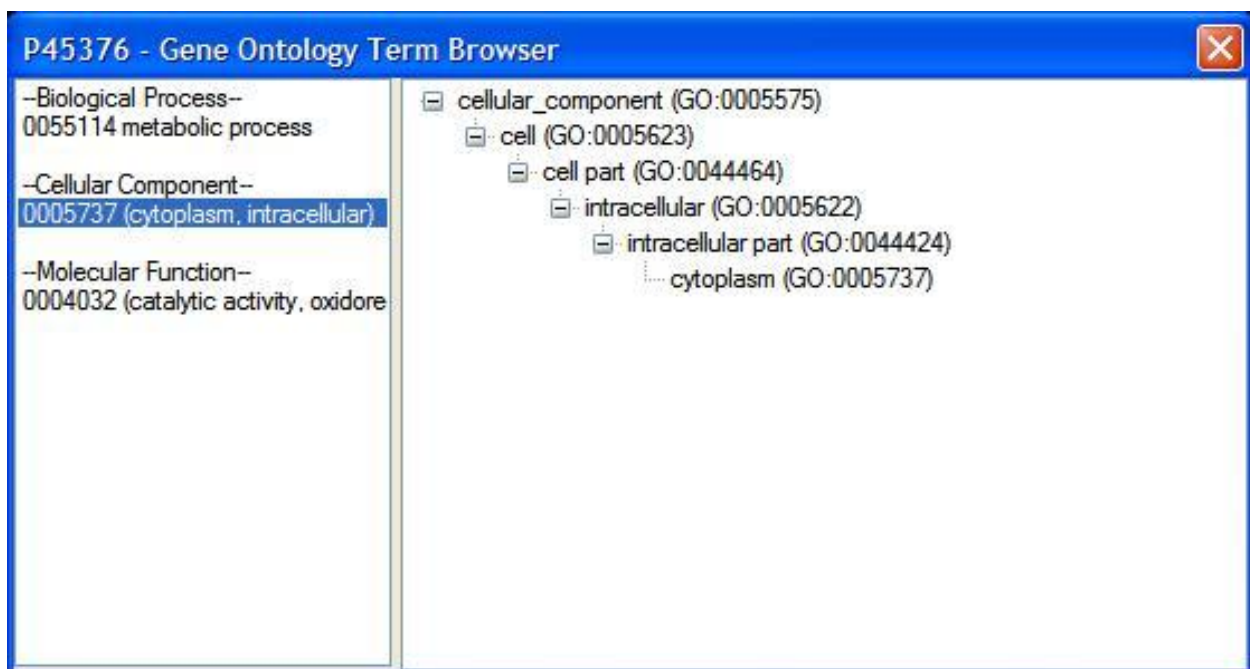
## Charts for GO Terms

You can create pie charts and bar charts to quickly interpret the biology of data sets via their associated GO information. Pie charts illustrate the variety of Biological Process, Cellular Component, and Molecular Function GO terms in *one* data set, whereas bar charts illustrate the differences in associated GO terms between *several* data sets.

To create a GO pie chart, first select the tab that contains the fully annotated data set of interest. Then, click the pie button on the toolbar. The resulting pie chart can be saved as an image by right-clicking on the pie chart itself and clicking "Save Image As...".
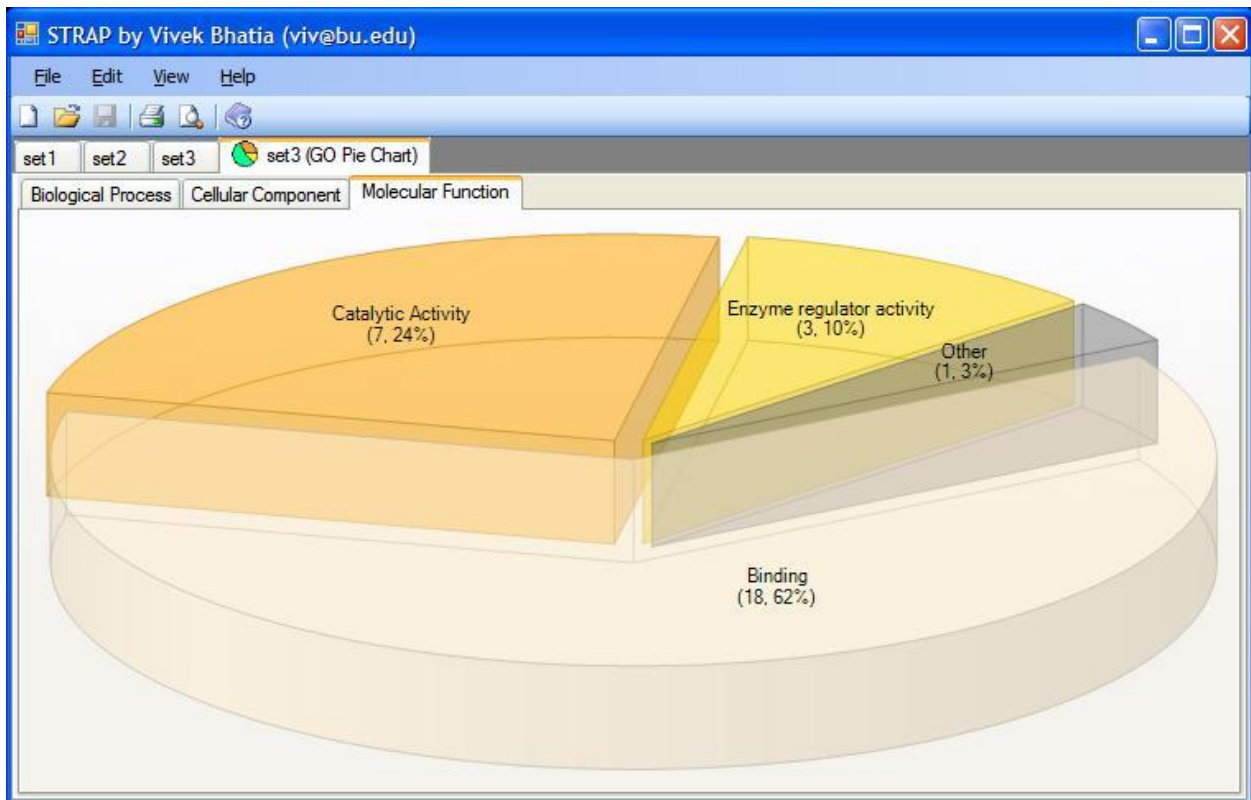


To create a GO bar chart, first make sure you have at least two data sets open and fully annotated. Then, click on the bar chart icon in the toolbar and configure the bar chart to your liking in this dialog box:
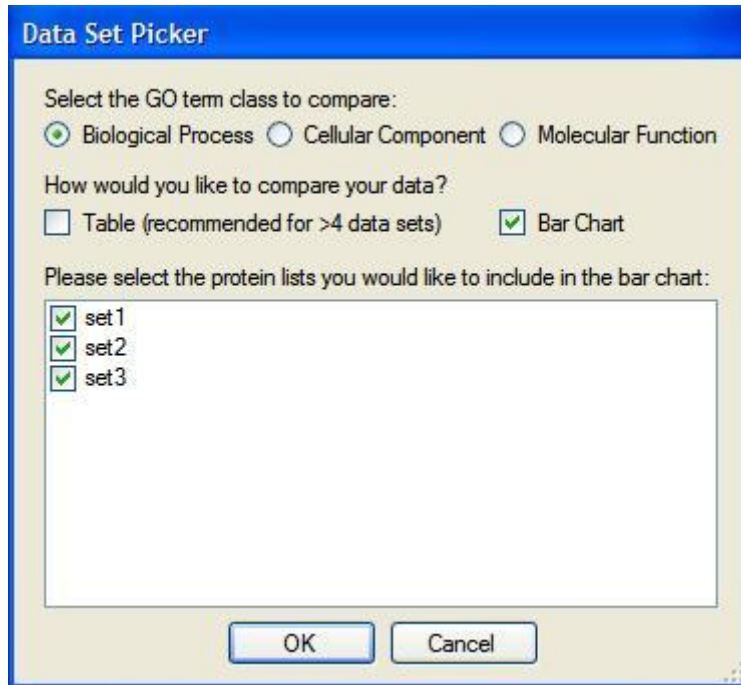
Press *OK* to see the resulting graph and/or table. To adjust the *y*-axis of the bar graph to see all possible GO terms, right click anywhere on the bar graph and select *Set Scale to Default*.

For comparisons that involve more than 4 files, it may be more useful to see the comparison in tabular form. If you have more than 4 files, check the "Table" checkbox to see your comparison in tabular form.

There are three major types of GO terms: biological process, cellular component, and molecular function. Comparisons between data sets can only be made within one of these major GO categories at a time. These images can be further edited or saved via the context menu as well. Here is an example of a comparison of molecular function GO terms between the three protein lists in the sample data:
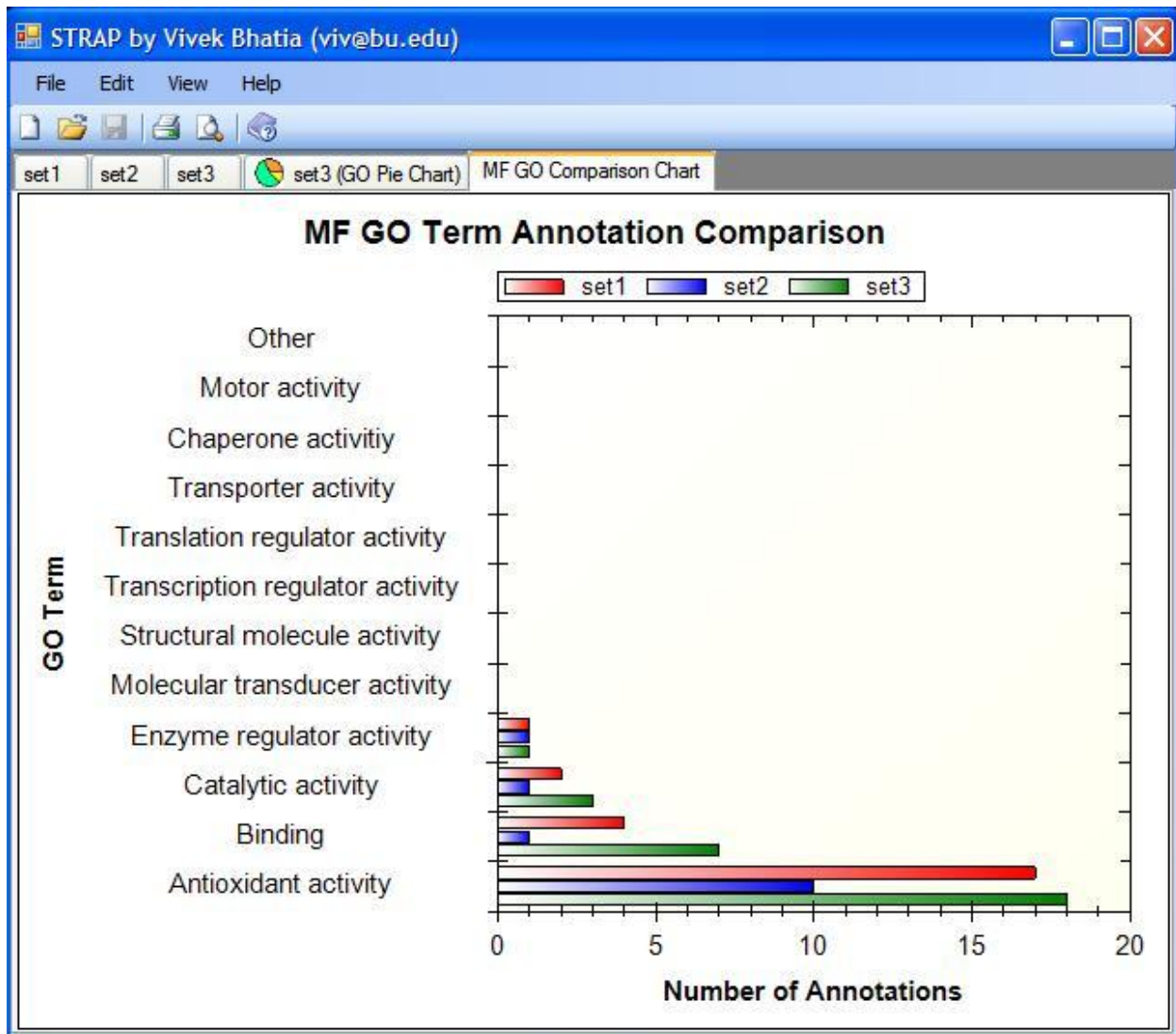
## Sending Data to the Gaggle

The Institute for Systems Biology has created a tool called the Gaggle which connects to many other data analysis programs (most of which are gene-oriented). STRAP can export genes associated with each protein to the Gaggle via Firegoose. To do this, open an annotated protein list and click the goose icon in the STRAP toolbar. This will open up an HTML page in Firefox, and the data contained in the HTML page can be broadcast to other Geese (e.g., Cytoscape).

## Conclusion

This tutorial gives you an idea of what STRAP can do and how it can fit into your data processing workflow. In summary, STRAP can:

- Annotate a list of proteins given a DAT, protXML, or plain text file

- Group proteins in a list by their gene ontology (GO) terms
- Modify GO annotations of proteins
- Use a GO pie chart to interpret a protein list
- Use a GO bar chart to compare protein lists
- Export proteins in lists to tools in the Gaggle

If you have any questions or suggestions, please Contact Us.