Tutorial for the *C. elegans* Matrisome Annotator
http://ce-matrisome-annotator.permalink.cc/

1. On the input page, you can search or manually browse through *C. elegans* matrisome by clicking on [+] of the "Browse *C. elegans* matrisome" field.

2. Upload a CSV file or copy/paste your gene list in the "Upload your gene list" field. In order to be recognized, the CSV file should contain only one column and no header. Independent of how the gene list is submitted, a window will appear that displays how the gene list is parsed. Check that your dataset was imported correctly. Under "Select input options", choose the gene ID type used in your dataset (*see Note 1*). You have the option to name your analysis.

3. Click "Apply import settings" button. A table will appear displaying your input genes with their corresponding Gene name, Gene description, WormBase ID, and Entrez gene ID. This table is searchable. The total number of mapped genes is displayed at the bottom left of the table. The genes which could not be mapped to WormBase IDs are listed in the table below (*see Note 2*).

4. Click on the "Data analysis" tab on the left-side bar. This tab appears when at least one submitted gene was successfully mapped.

5. At the top of the "Data analysis" page you will see the number of identifiers submitted, how many were recognized, and how many were identified as being part of the *C. elegans* matrisome. Three tabs are provided to explore the matrisome signature of your gene list:

1) The occurrence page lists the affiliations of each gene to their respective matrisome divisions and categories. Each gene can be further examined at the bottom of the page by clicking on [+] for the table of interest. The "In matrisome (%)" column, indicates the distribution of the matrisome genes of your dataset across the different matrisome categories.

2) The Venn diagram is to scale and interactive and by hovering with the mouse over each area you can determine how many genes belong to each matrisome division.

3) The complete fraction page summarizes the distributions of both the published matrisome (719 genes) and the user-supplied list in a circular graph. The core section is split into the four matrisome divisions and each division is subdivided in its periphery into its corresponding categories.

6. The Download / Export tab provides the option to download the data tables (Excel, CSV), illustrations (PDF and PNG) as well as R RDS files. We recommend to download the Excel file (both checkboxes checked) to have the entire matrisome analysis as well as your uploaded gene list (mapped and unmapped) in one file.

Notes:

1. We recommend using the most stable identifier, which are in decreasing order: Ensembl ID / WormBase ID (e.g., WBGene00000149), Entrez ID (e.g. 180783) and gene name / symbol (e.g., *apl-1*).

2. Problems with genes that fail to map: this could be due to multiple reasons: (A) copy / pasting errors, (B) selecting the wrong gene identifier, but also (C) through potential gene annotation changes. Failing to map genes is usually the case with retired gene/sequences names (e.g. pqn-4 and T25B9.10). For instance, *C. elegans* official gene names and sequence names are regularly updated and thus are often not mapped anymore. If genes failed to map, go to http://www.wormbase.org/ and obtain the corresponding Ensembl ID / WormBase ID (e.g., WBGene00012016).