NEPTUNE tranSMART

Quick Start Guide

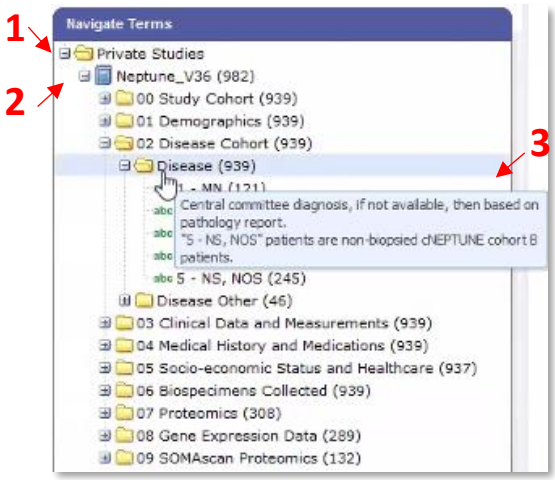
The NEPTUNE tranSMART instance can be accessed at https://neptune-transmart.med.umich.edu/

The application works best in Chrome, Firefox, or Safari. It cannot be used from a mobile device. The first time you log in (and once every 90 days after that), you will be asked to attest to a Data Use Statement.

Overview

To explore the data available, click on the plus sign (+) next to "Private Studies" in the **Navigate Terms** section (**1**). Then click on the plus sign next to the latest "NEPTUNE" version to expand the terms tree (**2**). The tree provides additional useful information:

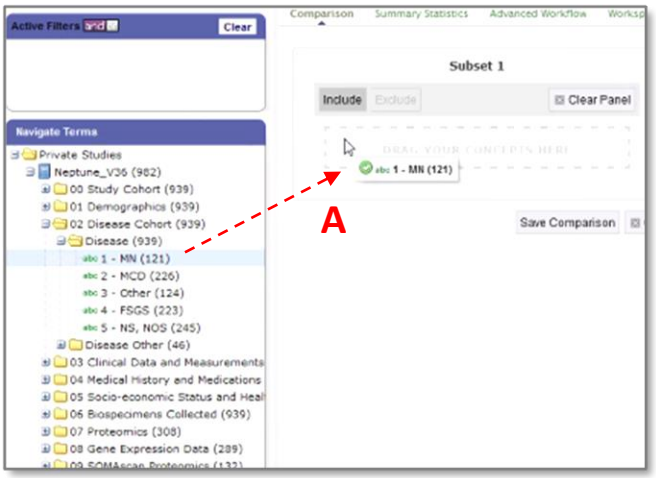
- The numbers in parenthesis after the terms are the number of samples that have that data.
- The **123** or **abc** before a term indicates if the data is numeric or categorical in nature.
- Hovering over a node, or right-clicking on it, will bring up a definition pop-up for some terms (**3**).



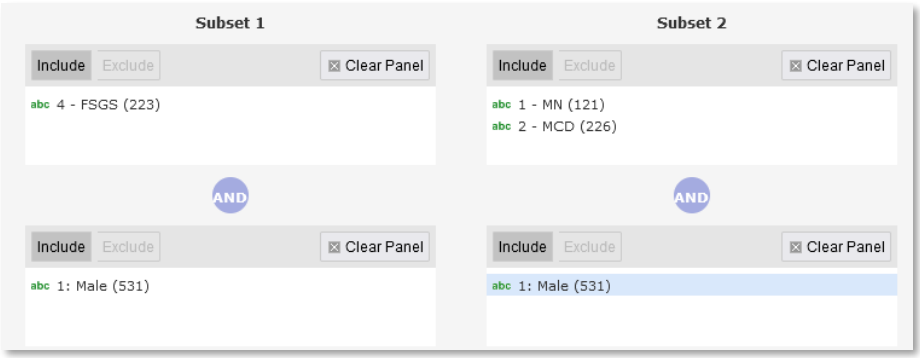
Selecting Data

The first step is to set the **Subset**(s) of data you wish to explore. Drag and drop terms from the tree into the **Subset** boxes on the right (**A**).

- If an entire folder is dragged over, all samples with a value for that data point are included in the subset regardless of the value (e.g. dragging the folder "Disease" would include all patients with values of "MN", "MCD", "FSGS", "Other", or "NS, NOS").
- If only a single term is included, then only samples having that value are included (e.g. dragging "MN" would include only those patients assigned a value of MN).
- If a numeric term is selected, values can be further restricted by choosing an option from the pop-up menu. Select "No Value" to use all samples with any value.
- Terms in the same subset box are part of an "OR" statement.
- Subset boxes are joined together in "AND" statements.
- Two different subsets (**Subset 1** and **Subset 2**) can be defined and compared. Example **B** compares **Subset 1**: "FSGS" AND "Male" to **Subset 2**: "MN" OR "MCD" AND "Male".

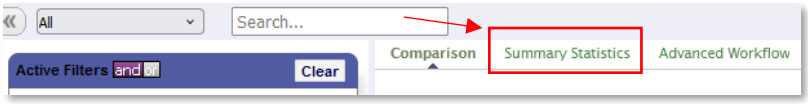


B



Analyzing Data

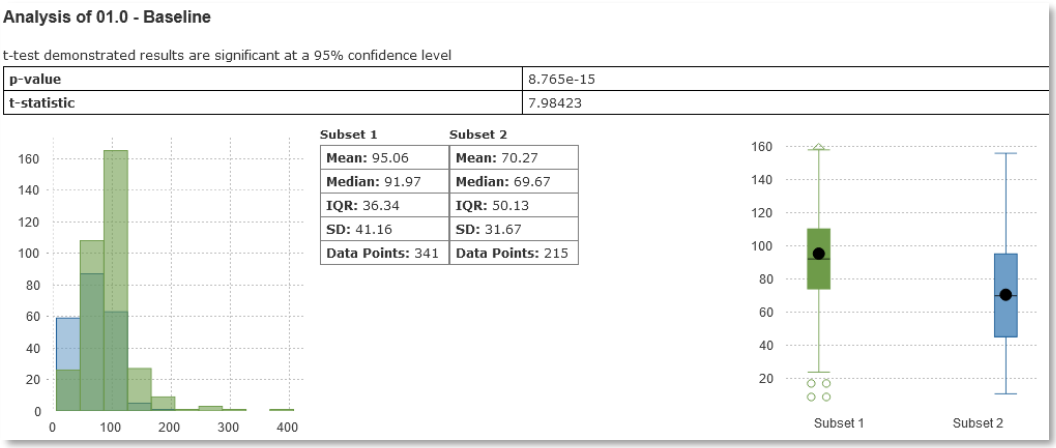
After setting a subset, clicking on the **Summary Statistics** tab will bring up an overview of the subset(s) by Age, Sex, and Race. Scroll through the results to see summaries and breakdowns of the terms used in the subset statements. For the subsets created, compare a concept across those Subsets by dragging an additional term from the tree directly onto the **Summary Statistics** results page.



Example 1: Compare eGFR at Baseline visit between Subset 1 of MN or MCD and Subset 2 of FSGS.

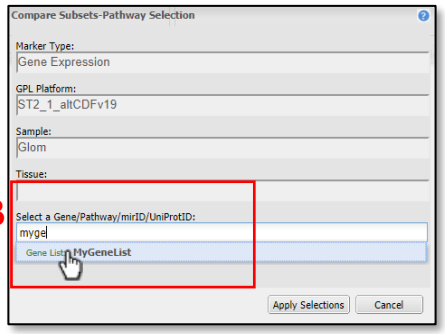
1. From the tree, drag and drop "1- MN" and "2 - MCD" into the first box of Subset 1 (to create the "OR" statement). Then drag and drop "4 - FSGS" into the first box of Subset 2. Click on Summary Statistics to see a comparison of the two subsets.

2. Drag and drop the baseline visit ("01.0 – Baseline") from the tree under "eGFR (mL/min/1.73 m2)" onto the Summary Statistics results page. A box plot comparing eGFR values in each subset will appear at the bottom of the page:



NEPTUNEtmSupport@umich.edu

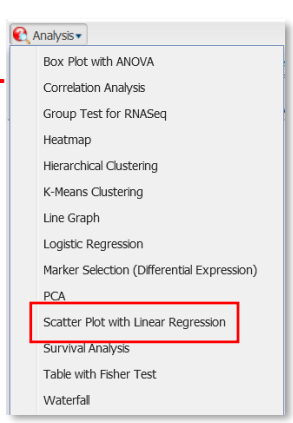
Using Gene Expression Data

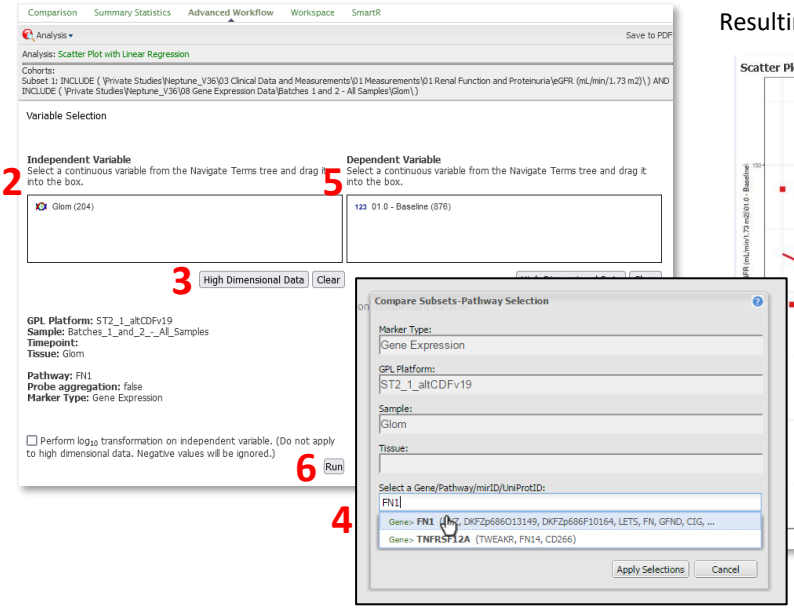


Explore genes of interest in gene expression data using the **Advanced Workflow** tab. Use a single gene or upload a gene list to compare multiple genes. To create a gene list, use the **Gene Signature/Lists** link at the top right-hand side of the application (**A**) and follow the instructions (only fields with an asterisk are required; use underscores instead of spaces in Gene List names). To use a gene list, type the name of a list into the "Select a Gene" box in the "High Dimensional Data" popup window and select it from the dropdown list (**B**).

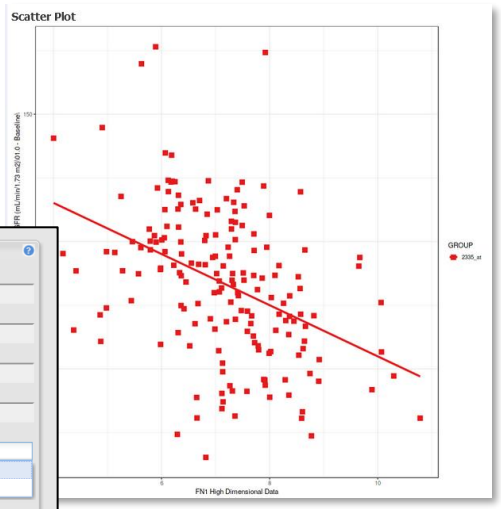
Example 2: Create a scatter plot of eGFR at visit "01.0 – Baseline" against Gene Expression levels on FN1.

- Subset to desired patients by dragging "Glom" and "eGFR (mL/min/1.73 m2)" into the first two Subset boxes on the **Comparison** tab.
- Click on the **Advanced Workflow** tab and select "Scatter Plot with Linear Regression" from the "Analysis" menu. (**1**)
- Drag "Glom" into the "Independent Variable" box. (**2**)
- Click on the "High Dimensional Data" button under the "Independent Variable" box. (**3**)
- Type "FN1" in the gene box and then select it from the dropdown list. (**4**) Apply selections.
- Drag the visit "01.0 – Baseline" from below the eGFR folder into the "Dependent Variable" box. (**5**)
- Click "Run". (**6**)



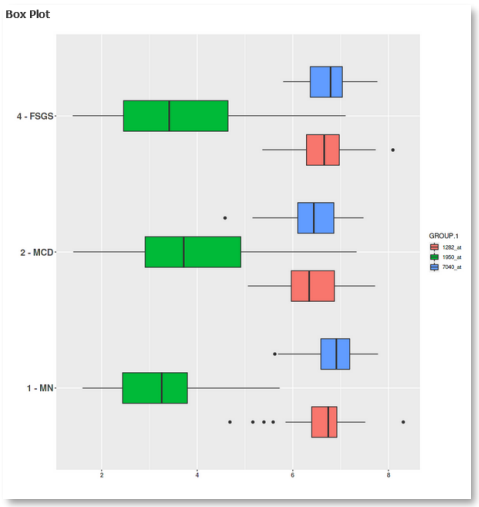


Resulting graph:



Example 3: Create a box plot to show expression of three genes in disease cohorts of MN, MCD, and FSGS.

1. Create a gene list with Col4A1, EGF, and TGFB1 using the **Gene Signatures/Lists** tab as explained above. Then subset to desired patients by dragging "MN", "MCD", and "FSGS" into the first Subset box on the Comparison tab.
2. Click on the Advanced Workflow tab and select "Box Plot with ANOVA".
3. Drag "Glom" into the Independent Variable box.
4. Click on the "High Dimensional Data" button.
5. Type the name of your gene list into the gene box and select it from the dropdown menu.
6. Drag "MN", "MCD", and "FSGS" into the Dependent Variable box.
7. Click "Run". Resulting graph is shown at right.



NEPTUNEtmSupport@umich.edu
Last updated: 11/5/2024