nCounter® TCR Diversity Report in ROSALIND®

Overview

The NanoString TCR Diversity Analysis evaluates the expression of the variable regions (alpha, beta, gamma, and delta) within T cell receptors. The variable regions are assessed for overall expression (above or below background) and normalized to a panel standard which allows for more precise quantification of these variable regions. We also calculate the T cell Receptor (TCR) score, an estimate of TCR Diversity.

The TCR score calculates the diversity of TCR beta variable regions within a sample. The score is based on the Shannon Diversity index calculation, which is a mathematical measure of species diversity in a community. We use this ecological calculation to account for the abundance and evenness of the variable regions present within a given sample versus the population of T cell receptors within a given dataset. The score given is relative within a dataset, in which a higher TCR score means there is a more diverse population of variable regions, or a less clonal population. A lower score means there is less diversity, or a more clonal population. True clonality can only be determined using the full sequence of the TCR, but we can estimate clonality by measuring a low diversity of TCR beta variable regions. The TCR Diversity Report includes a p-value calculation associated with the TCR diversity score; as such the sample size is recommended to be five or more per each group entered in the analysis.

Normalization Options

TCR Diversity analysis with nCounter has a unique need for normalization compared to standard gene expression applications. Counts on the nCounter platform represent unique molecules observed from each sample, and in this way they are quantitative. Each probe in the assay has its own efficiency rate, which means that the exact number of counts observed for each target in the assay cannot be directly translated into a certain number of targets present in the original sample. Given this, it is not accurate to directly compare counts across targets in an assay, only to compare counts across samples for a single target. In the instance of TCR diversity, comparing quantities of targets within a single sample is required. This is accomplished by performing panel standard normalization. Panel standards represent an approximately uniform quantity of synthetic targets for each probe pair in the panel (for more info, see https://www.nanostring.com/support-documents/panel-standard-and-calibration-sample-usage/). Normalization of all TCR variable region segments to their corresponding panel standard targets enables comparison of data in the form of heatmaps and calculation of a TCR Score.

The report is generated with three versions: Normalization-None, Normalization – constantNormalize, and Normalization – cd3Normalize. The "None" option indicates that no additional normalization was done beyond the normalization to the panel standard. The other two options enable additional normalization to T cell quantity in the form of CD3 or the TCR constant region. CD3 normalization will normalize the panel standard-normalized count for each target to the geometric mean of the three CD3 genes in the panel (CD3D, CD3E, and CD3G). Constant region normalization will normalize the counts for each of the TCR variable regions to the corresponding TCR constant region (TRAC for all alpha variable regions, TRBC1/2 for all beta variable regions, etc). Typically, results are similar regardless of normalization. However, if your samples are uneven with respect to T cell quantity, this normalization to CD3 or the TCR constant region is preferable.

Beginning a TCR Analysis on ROSALIND®

TCR analysis on ROSALIND requires 5 easy steps and takes less than 10 minutes from the start of the experiment set-up to the final deliverable of the complete TCR report.

1. Set Up Your Project Library
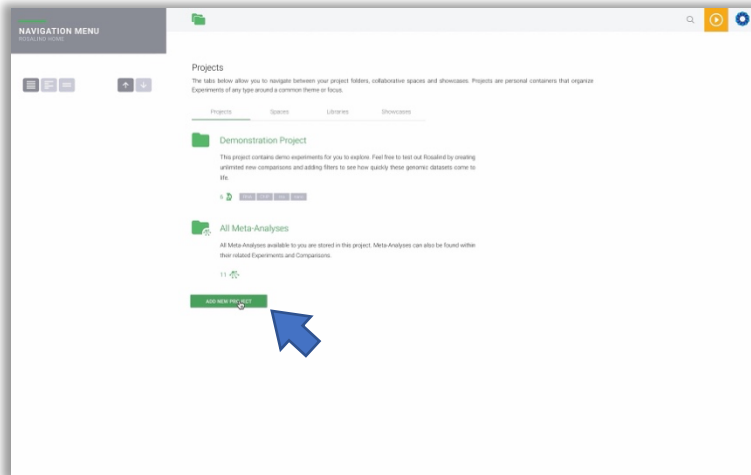
After logging in to your account (https://nanostring.rosalind.bio/login) select to ADD NEW PROJECT from the main ROSALIND home screen.

Make sure to select "Partner Recommended Analysis" or the "Design a New Experiment" option and choose the NanoString nCounter Data Analysis / NanoString TCR Report options during the experiment set up.
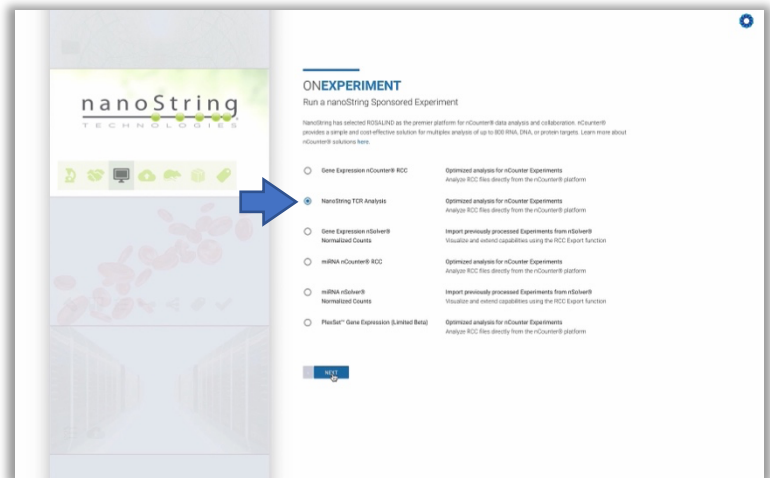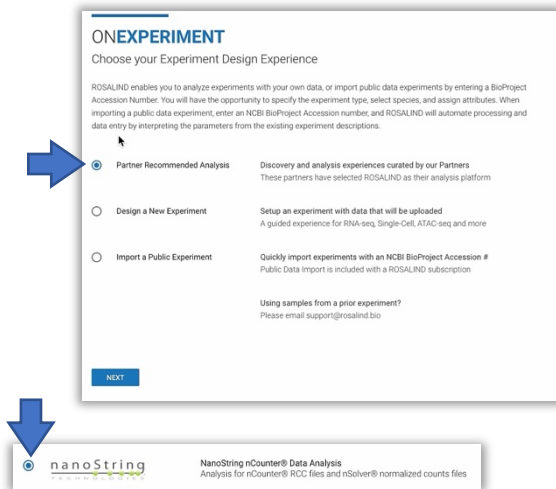
2. Upload Your Files
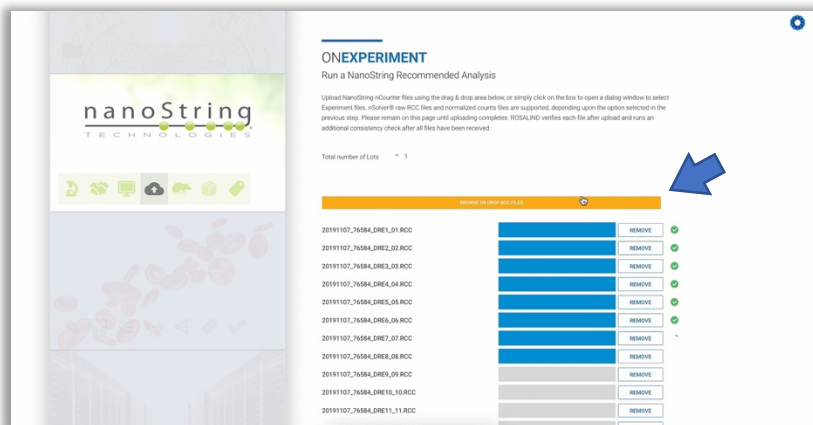
Drag and drop your RCC files (at least 12 files) into the upload area provided, and ROSALIND will validate the data. The validation ensures the panel run was either TCR Diversity or CAR-T Characterization, and all samples were run with the same panel.

3. Generate Your Attributes File

The attributes file is a convenient way to inform the analysis of all possible annotations for your samples. A template can be found here, showing the attribute details for samples, including a time course and the panel standard: https://rosalind.bio/hubfs/TCR_example_attributes_file.csv

Samples should be annotated by the convention shown in the attribute file provided in the link above. The reference sample(s) should be denoted "1", and the other samples "0" in the "PanelStandard" column.

ROSALIND provides a way to create a CSV of defined attributes by clicking **GENERATE ATTRIBUTE FILE** . A sample attribute file will be designed and downloaded for your experiment. When your attributes file is ready, click **UPLOAD ATTRIBUTE FILE** to upload your final version of this file.

All sample attributes should be character-based (i.e. non-numerical values) unless you are requesting a time-course assay. Use of BOTH "TimeSeries" and "TimeSeriesLabel" attributes is required for reports with time series defined. "TimeSeries" is a numerical value that orders the data, and "TimeSeriesLabel" is a string that will label the graphs (e.g. 12h, 1d, 24h).

4. Review and Launch Your Experiment


Review your sample annotation table to ensure accuracy and either modify in-line on the table or choose **UPLOAD NEW FILES** to make changes before saving and launching your experiment. Once you have confirmed your data accuracy, select **SAVE ATTRIBUTES**

You will receive an email in 5-10 minutes stating that your TCR report is ready to be unlocked.

5. "Unlock" and Download Your TCR Report

While your report is locked, you will notice an "Analysis Locked" navigational item where you can use the Analysis Units provided by NanoString to access your final TCR report download.
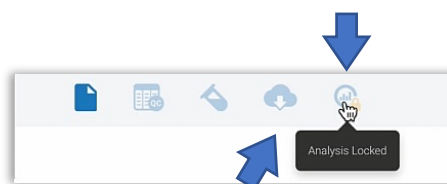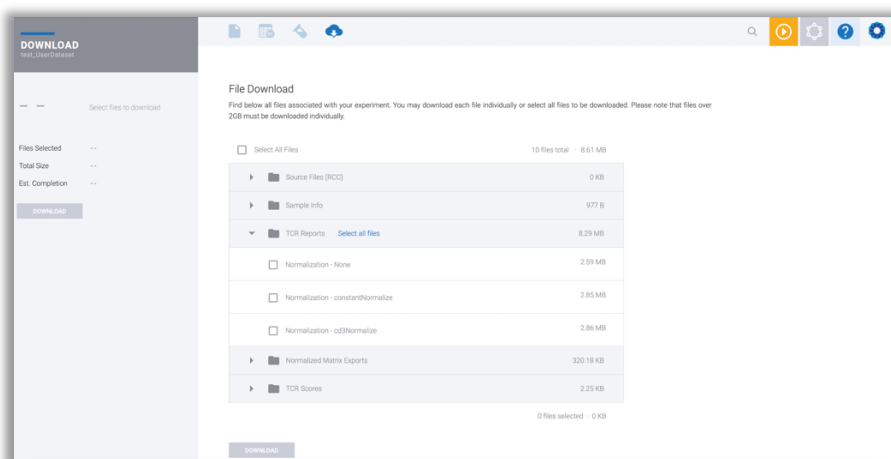
Once the report is complete and unlocked, a download icon will be present at the top of the screen. Click on that icon, and then select the option for "TCR Report". Three TCR reports will be available based on the 3 options for normalization, including the TCR exports of relevant data in a flat csv format. Within each html report, images are available for individual download using an ⬇ icon at the side of each figure. Other metrics and values are also available for download on the same page as the

TCR Score

TCR Diversity indexes can be calculated in a few manners, but for this analysis a version of the Shannon Diversity index has been chosen. Shannon Entropy is calculated as in the equation here:

$$H = -\Sigma p_i * ln(p_i)$$

The steps required to calculate the TCR Score are as follows:

1. Start with the normalized counts (panel standard only, CD3, or Constant) for each beta variable region. As these values have been log_2 transformed, the calculations below will reverse that transformation in the calculation.
2. Compute the p_i (proportion of each variable segment): (2^normalized count)/(number of beta variable regions). The number of beta variable regions is 46.
3. Calculate the $p_i*ln(p_i)$ for each beta variable segment.
4. Sum the $p_i*ln(p_i)$ for each sample to compute the TCR score.

Plots in the TCR Score tab

Boxplots

In the TCR Score tab, comparisons of the TCR scores for all possible combination of the parameters in the attribute file is included. For grouping variables with 2 levels, the p-value from Mann-Whitney-Wilcoxon analysis is displayed in the boxplot. For grouping variable with more than 2 levels, the p-value from Kruskal-Wallis analysis is displayed. Note that p-values for comparisons with fewer than 5 samples per group can be inaccurate.

The boxplots are colored differently depending on the value of the grouping variable. The grouping variables are displayed in the legend at the right side of the boxplot.

Time Series Plots

If time series attributes are specified in the attribute file, a time series plot of time point vs TCR score is displayed.