

User Settings Parameters Explained

This document explains the main options in the `user_settings` section of your configuration file. These settings define your project name and the metadata variables that will be used in your analysis.

1. Project Name

- **project_name**

A short name for your project. This will be used as the main directory for storing results. A default is given based on the PI name, JGI proposal ID, and a keyword determined by a data analyst.

2. Variable List

- **variable_list**

A list of metadata variables (columns from `<dataset_name>.linked_metadata`) that you want to use in your analysis. These variables will be used for grouping samples, statistical tests, and plotting.

Warning: do not change or remove the "group" line, this is a placeholder that combines all supplied metadata variables to create unique groups of samples.

Note: The list of possible metadata variables will already be populated based on the information you provided to JGI during sample processing and data generation. To view how your metadata and samples are organized, run the workflow notebook up to the "*Link analysis datasets by finding corresponding samples and metadata*" section. You can then preview the metadata for your dataset by adding a new notebook cell and running the following command: `<dataset_name>.linked_metadata`)

Summary: Set `project_name` to determine your base output directory, and review the `user_settings->variable_list` of metadata variables you want to use during analysis - these will be pre-set, but can be subset as needed. These variables will be available for feature selection and downstream analysis.