How-to-Monocle in SeqGeq

Plugin Implemented by Miguel Velazquez-Palafox and Ian Taylor

Monocle is an algorithm developed and maintained by the ColeTrapnell lab.(1) The tool attempts to infer the progression of development in a population of cells (or "trajectories") by comparing single cell expression profiles. The platform also predicts states (i.e. subsets) within the progression mapping.

The current Monocle version developed to work with this plugin is version 2 from the Trapnell lab.

This document describes the intended use of the Monocle plugin for SeqGeq, tested in SeqGeq v1.4.

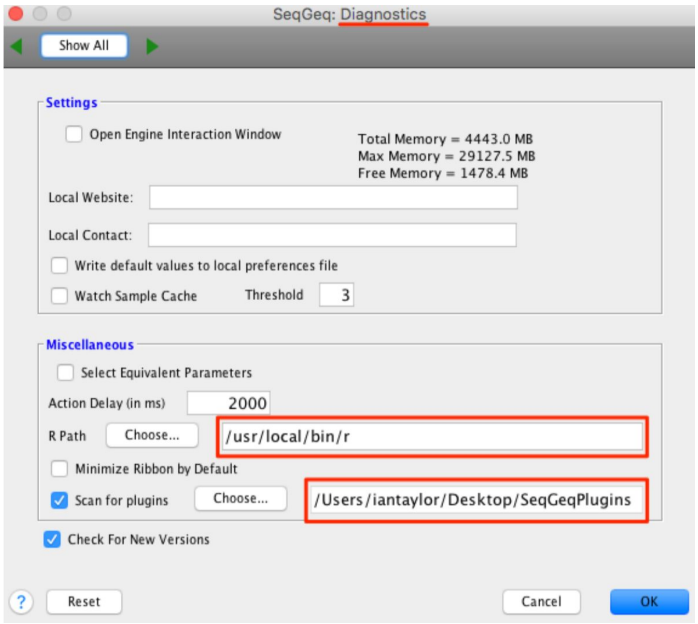
Installation

The Monocle packages need to be installed and will run in the R environment. This plugin was tested in R version 3.4.2. To install the required R packages, use the following commands in R:

```
source("https://bioconductor.org/biocLite.R")
biocLite("monocle")
```

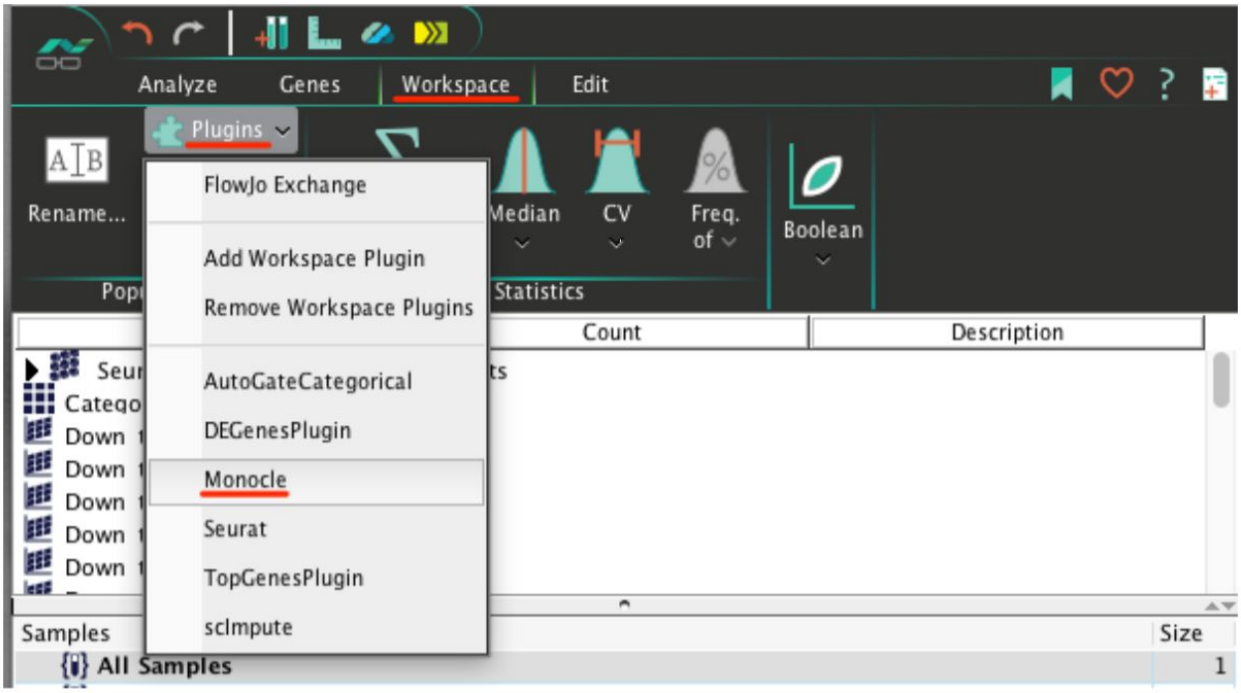
The Monocle.jar file will need to be placed in a user's SeqGeq Plugins folder.

The user's SeqGeq Diagnostics preferences will need to include file paths to both the SeqGeq Plugins folder and the local R environment:



Running the Plugin

In order to run the Monocle plugin you'll first need to select the population of interest within the SeqGeq workspace. Then visit the Workspace tab in SeqGeq's workspace, and select Monocle from the Plugins dropdown menu there:



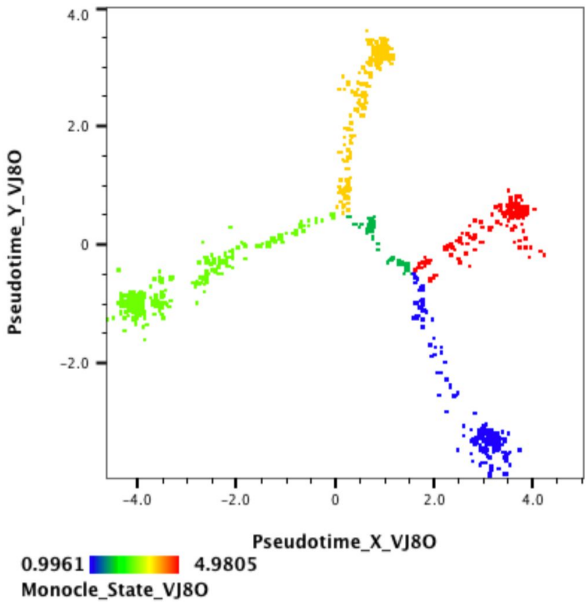
This will bring up a plugin dialog in which gene filters can be set, including: Minimum mean gene expression across the population, minimum threshold of cell number and percentage expression. You'll also need to select the genes with which you want the algorithm to compare between cell states. Mousing over options in that dialog will bring up tool tips with a brief description of what the function does.

The Monocle algorithm is computationally intensive, and may take some time to calculate, dependent on the number of cells and parameters selected.

Outputs

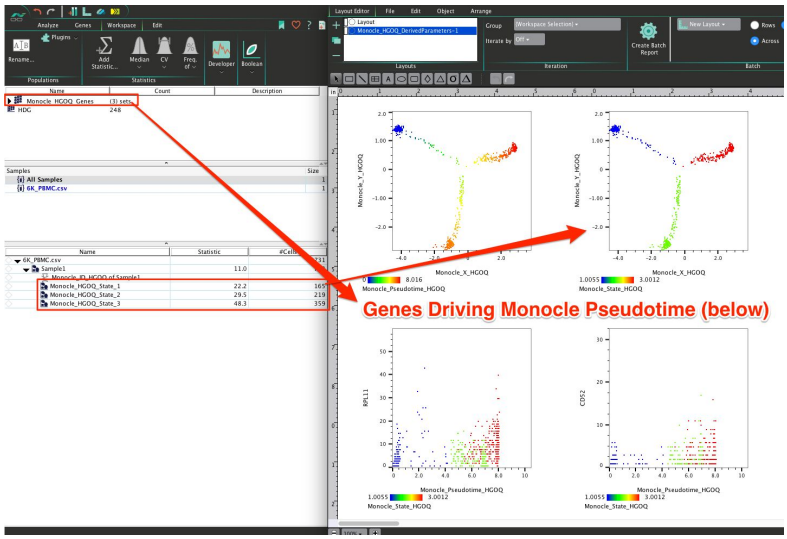
Monocle generates a set of synthetic parameters: Pseudotimes x and y illustrating a tree of differentiation, a predicted state, which typically maps onto the pseudotime parameters well, Size Factor which normalizes an expression matrix across cells, Genes Expressed which illustrates the

gene filters set initially, and differentially expressed genes per state (cluster) driving trajectories of differentiation there:



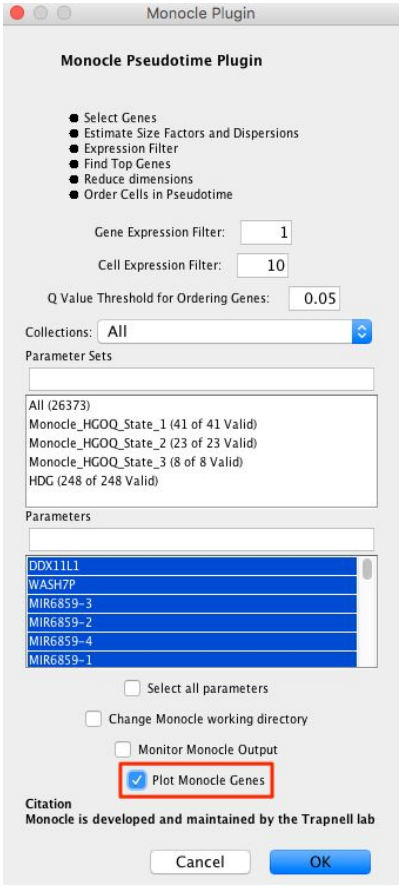
Monocle keeps track of its own run number using a randomized four character code. This allows you to re-run monocle on the same population multiple times, or within the same sample, while retaining previous runs.

As of Monocle v2.0 the plugin also automatically produces graphs in SeqGeq's Layout Editor illustrating the outputs of the plugin, including branching predicted by differentially expressed genes in each of the clusters or "states", and across the predicted pseudotime:



Note: To see each of the genes predicted to drive the development of states illustrated across

pseudotime, check the box “Plot Monocle Genes” within the plugin dialog:



Note: This new release of Monocle automatically gates state clusters detected by the algorithm, and genesets predicted to be important in driving the trajectories of those states.

References

1. Trapnell, Cacchiarelli et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature Biotechnology. 2014

2. cole-trapnell-lab.github.io/monocle-release/docs/