The `titanQC` Package

Stijn Dickens, Hinrich Goehlmann, Willem Talloen, Suzy Van Sanden, Tobias Verbeke

April 13, 2012

1 General Introduction

There are a number of libraries and functions available for the QC of Affymetrix GeneChip arrays. With the development of the HT platform by Affymetrix it is now possible to profile 96 samples simultaneously using so-called peg arrays. These peg arrays are arranged in such a way that they have been made compatible with standard molecular biology 96-well plates. However, at this moment there are no packages that address potential spatial effects across the plates. The aim of the `titanQC` package is to provide users with functions that allow the assessment of potential gradients or artifacts visible only on the plate level rather than the individual sample level.

2 `getPosition`

This function retrieves the position of one single CEL file on the GeneTitan plate. For a set of CEL files on a GeneTitan plate, one can use the `getCelFilePosition` function. Using these functions the exact position of a sample on a GeneTitan plate is determined directly from the .CEL file without the need to consult other files.

3 `displayPlate`

This function generates a summary visualisation of a set of CEL files for a GeneTitan plate. It can also be used to, e.g., visualize the corresponding position in the incubator plate when a biological experiment has been performed in vitro. To check whether there might be confounding issues between problematic wells and the severity of a biological treatment, the user can assign colors to groups of samples (e.g., a color for the reference group, a different color for treatment 1 at timepoint 1, a different color for treatment 2 at timepoint 1, etc.) This color assignment is used throughout the whole package. In this view the group number plus the group color is presented in the context of their positions on the 96-well plate. This plot helps in visualisation of samples over a GeneTitan plate and helps in quickly retrieve and localise them. Ideally this is used only in combination with short, possibly numerical, .CEL file names to allow for optimal readability.

4 geneFiltering

This function provides as a preprocessing step an approach to gene filtering based on parameters such as control probesets, I/NI calls, fold changes and kurtosis cut-off. The function helps in eliminating probe errors or mismatches, ambiguities and irrelevant genes for cleaning up a dataset or ExpressionSet objects.

5 boxPlate

This function generates a boxplot of expression intensities for each well on the GeneTitan plate after preprocessing and gene filtering of the raw data. The plot helps to distinguish differences in expression intensities between all samples, and aids the descision to possibly eleminate strange observations.

6 statisticPlate

This function provides statistics to the ExpressionSet object that indicates certain types of deviations in the MA plots.

7 MAPlate aka the lips plot

This function plots the distribution of the intensity ratio versus the average intensity corrseponding to a .CEL file position on the GeneTitan plate. It gives a quick overview of the distribution of the data over the global plate and visualises peculiar behavour of samples which can be excluded for further analysis.

8 densityPlate

This function visualises the overall plate view of sample-specific density plots in a 96-well format. The graph shows how many genes have a certain signal intensity. The intensity values of the multiple oligos for each transcript are summarized to one single value per transcript. No normalization is done to visualize the distribution of raw intensities for every chip. This plot gives a global visualisation of signal intensities of all samples. Samples which show altered behaviour can be found quite easily.

9 plmPlate

This function visualises the overall plate view of PLM residuals and plots all affyPLM QC images on a grid respecting the plate layout positions. The plot helps in detecting sample specific artifacts, but also if artifacts show similar patterns throughout the plate (e.g. edge effects)