An Introduction to SingleCellAssay

Andrew McDavid

September 10, 2012

1 Philosophy

`SingleCellAssay` is an R/Bioconductor package for Fluidigm and friends. We are trying to represent assays that have multiple features (genes, markers) per well (cell, etc). The assays is mostly *complete* in the sense that most features are measured in each well. We'll complete them if they are not.

Internally, we store everything as one giant `data.frame` with names of special columns kept in a `mapping` that contains column names and keywords. It is in long-melted format, in feature-major order, so not especially space-efficient, but rather is intended to be very flexible.

Each well, feature and **TODO:** unit (phenotype) has covariates measured. These are kept in AnnotatedDataframes, which are generated from the basal data.frame, if so provided. **TODO:** If not provided, then they can be added after object creation.

2 Reading Data

Data imported in a Fluidigm instrument-specific format (whose details are undocumented, and probably subject-to-change) or in "long" (melted) format, in which each row is a measurement, so if there are N wells and M cells, then the `data.frame` should contain $N \times M$ rows.

```
library(lattice)
xyplot(rnorm(10) ~ rnorm(10))
```