FlowSorted.CordBlood.450k User's Guide
Methylation Dataset on Sorted Cord Blood Cells

Shan V. Andrews   Kelly M. Bakulski

Modified: January 30, 2016. Compiled: November 1, 2018

1   Introduction

The FlowSorted.CordBlood.450k package contains Illumina 450k measurements from 17 indivudals, all of whom contribute between 4-7 samples of distinct cell types. These cell types are B cells, CD4 T cells, CD8 T cells, granulocytes, monocytes, natural killer cells, and nucleated red blood cells. The package contains analogous objects for cord blood to those contained in the partner package FlowSorted.Blood.450k, and can be used in a similar manner to the those objects.

The primary use for these data is to estimate cell type proportions in 450k-based epigenome wide association studies in which DNA has been derived from cord blood. These data are used in `estimateCellCounts` function from the minfi package. In the function, the user can specify "Blood", or "CordBlood", in the `compositeCellType` argument, and the appropriate reference data will be loaded.

2   Data

These data are derived from 17 cord blood samples from a prospective study based in the Johns Hopkins Hospital. For more details on these samples, as well as the pipeline used to estimate cord blood cell type proportions, please see Bakulski et al. (2016). Raw 450k measurements on sorted cell populations are contained in an RGChannelSet.

3   Tables

In addition to the RGChannelSet, this package contains two additional tables. The first contains F statistics for each probe that survived several probe QC steps (detailed in Bakulski

et al. (2016)) demonstrating the extent to which methylation at that CpG site is associated with cell type. Users may seek to compare any sites of interest in an association study with this list to evaluate the potential for confounding by cell composition. The second table contains the 700 probes (100 for each cell type) that were selected to differentiate cell types. While these 700 probes were selected based on the reference RGChannelSet contained herein, in the implementation of the `estimateCellCounts` function probes are selected based on a combined RGChannelSet of user and reference data. Please see Bakulski et al. (2016) for details on the probe selection process.

4   References

Bakulski KM, Feinberg JI, Yang J, Brown S, Andrews SV, McKenney S, Witter F, Walston J, Feinberg AP, Fallin MD. *DNA methylation of cord blood cell types: Applications for mixed cell birth studies.* Epigenetics (2016), 11:354–362. DOI: http://dx.doi.org/10.1080/15592294.2016.1161875