FlowSorted.Blood.450k User's Guide
A Public Illumina 450k Dataset

Andrew E. Jaffe

Modified: October 19, 2013. Compiled: November 3, 2022

Introduction

The FlowSorted.Blood.450k package contains publicly available Illumina HumanMethylation450 ("450k") DNA methylation microarray data from a recent publication by Reinius et al. 2012 [1], consisting of 60 samples, formatted as an `RGChannelSet` object for easy integration and normalization using existing Bioconductor packages (the code for creating the R object from the raw `.idat` files is provided in the package as well at `inst/scripts/getKereData.R`). For example, this dataset may be useful as example data for other packages exploring, normalizing, or analyzing DNA methylation data.

Data

Researchers may find this package useful as these samples represent different cellular populations of whole blood generated on the same 6 male individuals using flow sorting, a experimental procedure that can separate heterogeneous biological samples like blood into pure cellular populations, like CD4+ and CD8+ T-cells. This data can be directly integrated with the minfi Bioconductor package to estimate cellular composition in users' whole blood Illumina 450k samples using a modified version of the algorithm described in Houseman et al. 2012 [2].

Tables

This package also contains several useful tables, including the degree of association between cellular composition and each CpG on the Illumina 450k: CpGs identified in whole blood for phenotypes or outcomes should be treated with caution, especially if the outcome of interest correlates with cellular composition. Lastly, there is a table containing the average DNAm by cell type for the probes determined to best estimate cellular composition, which can be passed to the cellular estimation function. These probes appear to be least susceptible to association with other phenotypes, given the very high degree of association with cellular composition, and can therefore be suitable "control probes" for removing unwanted variation [3].

References

[1] Lovisa E Reinius, Nathalie Acevedo, Maaike Joerink, Göran Pershagen, Sven-Erik Dahlén, Dario Greco, Cilla Söderhäll, Annika Scheynius, and Juha Kere. Differential DNA Methylation in Purified Human Blood Cells: Implications for Cell Lineage and Studies on Disease Susceptibility. *PloS One*, 7(7):e41361, 2012. doi:10.1371/journal.pone.0041361, PMID:22848472.

[2] E Andres Houseman, William P Accomando, Devin C Koestler, Brock C Christensen, Carmen J Marsit, Heather H Nelson, John K Wiencke, and Karl T Kelsey. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, 13(1):86, 2012. doi:10.1186/1471-2105-13-86, PMID:22568884.

[3] J A Gagnon-Bartsch and T P Speed. Using control genes to correct for unwanted variation in microarray data. *Biostatistics*, 13(3):539–552, 2012. doi:10.1093/biostatistics/kxr034, PMID:22101192.