bcellViper, a data package required for the examples and vignette of the viper package

Mariano J. Alvarez, Federico Giorgi, Andrea Califano

Department of Systems Biology, Columbia University, New York, USA

October 30, 2025

1 Overview of bcellViper data package

The *bcellViper* data package provides some example datasets and a small B-cell context-specific transcriptional regulatory network (interactome).

Human B-cell expression dataset   The human B-cell dataset (Gene Expression Omnibus series GSE2350) [1] consists of 211 normal and tumor human B-cell phenotypes whose expression was profiled on Affymatrix HG-U95Av2 arrays, and it is contained in an ExpressionSet object with 6,249 features x 211 samples. The features (probe-clusters) were generated by the cleaner algorithm [2]. We can access this dataset with the following code:

```
> library(bcellViper)
> data(bcellViper)
> print(dset)

ExpressionSet (storageMode: lockedEnvironment)
assayData: 6249 features, 211 samples
  element names: exprs
protocolData: none
phenoData
  sampleNames: GSM44075 GSM44078 ... GSM44302 (211 total)
  varLabels: sampleID description detailed_description
  varMetadata: labelDescription
featureData: none
experimentData: use 'experimentData(object)'
Annotation:
```

B-cell context-specific transcriptional network   The B-cell interactome represents 172,240 inferred regulatory interactions between 621 transcription factors and 6,249 target genes. It is contained in a *regulon* class S3 object, and methods to access it are included in the *viper* package.

```
> targets <- unlist(lapply(regulon, function(x) names(x$tfmode)), use.names = FALSE)
> cat("Regulators: ", length(regulon), "\nTargets: ", length(unique(targets)),
+       "\nInteractions: ", length(targets), "\n", sep="")

Regulators: 621
Targets: 6249
Interactions: 172240
```

B-cell ARACNe results   A subset of the results of running ARACNe [3] on the B-cell dataset are included in the bcellViper package in the adjacency matrix format generated by ARACNe. The following code shows how this matrix can be parsed into a *regulon* S3 class object by the `aracne2regulon` function from the *viper* package.

References

[1] Basso,K. et al. (2005) Reverse engineering of regulatory networks in human B cells. Nat. Genet., 37, 382-90.

[2] Alvarez,M.J. et al. (2009) Correlating measurements across samples improves accuracy of large-scale expression profile experiments. Genome Biol., 10, R143.

[3] Margolin,A.A. et al. (2006) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics, 7 Suppl 1, S7.