aracne.networks, a data package containing ARACNe-inferred gene networks from TCGA data

Federico M. Giorgi, Mariano J. Alvarez, Andrea Califano
Department of Systems Biology, Columbia University, New York, USA

June 8, 2016

1 Overview of aracne.networks data package

The *aracne.networks* data package provides context-specific transcriptional regulatory networks (also called interactomes or regulons) from TCGA datasets.

ARACNe networks This package contains 24 Mutual Information-based networks generated using ARACNe-AP [1] with default parameters (MI p-value=10^{-8}, 100 bootstraps, permutation seed=1). ARACNe-AP was run on RNA-Seq datasets normalized using Variance-Stabilizing Transformation [2]. The raw data was downloaded on April 15,2015 from the TCGA website [3]. We follow the TCGA naming convention (e.g. BRCA = Breast Carcinoma) to name the individual context-specific networks.

```
> library(aracne.networks)
> data(package="aracne.networks")$results[,"Item"]
```

```
 [1] "regulonblca" "regulonbrca" "reguloncesc" "reguloncoad" "regulonesca"
 [6] "regulongbm"  "regulonhnsc" "regulonkirc" "regulonkirp" "regulonlaml"
[11] "regulonlihc" "regulonluad" "regulonlusc" "regulonov"   "regulonpaad"
[16] "regulonpcpg" "regulonprad" "regulonread" "regulonsarc" "regulonstad"
[21] "regulontgct" "regulonthca" "regulonthym" "regulonucec"
```

Write a network to file The package contains a function to print individual networks into a file. Four columns will be printed: the Hub id, the Target id, the Mode of Action (MoA, a glorified fitted Spearman correlation that indicates the sign of the connection and ranges betrween -1 and +1), the Likelihood (essentially an edge weight that indicates how strong the mutual information for an edge is when compared to the maximum observed MI in the network, it ranges between 0 and 1). Details for the *regulon* object are present in the VIPER publication [4].

In the following example, we print the first 10 interactions from the Bladder Carcinoma (blca) network. The network genes are named as Entrez Gene ids.

```
> data(regulonblca)
> write.regulon(regulonblca,n=10)
```

```
Regulator        Target        MoA        likelihood
10002        2648        0.994689591270463        0.886774633189913
10002        677827        0.116175345640136        0.707841406455471
10002        80152        0.999770437015603        0.950286744281199
10002        284382        -0.0368424333564396        0.0419762049859333
10002        9866        0.972066598154448        0.442238853411591
```

10002	283422	-0.574084929385018	0.260828476620346
10002	221613	-0.0959242601820319	0.717904706549976
10002	348174	0.953943934091558	0.814491117578869
10002	373509	0.704691385719852	0.244337186726846
10002	8803	-0.959165656086931	0.831653033754096

References

[1] Giorgi,F.M. et al. (2016) ARACNe-AP: Gene Network Reverse Engineering through Adaptive Partitioning inference of Mutual Information. Bioinformatics doi: 10.1093/bioinformatics/btw216.

[2] Anders, S and Huber W. (2010) Differential expression analysis for sequence count data. Genome Biol 2010;11(10):R106

[3] Weinstein J.N. et al. (2013) The cancer genome atlas pan-cancer analysis project. Nature Genetics 45, 1113-1120 2013

[4] Alvarez M.J. et al. (2016) Inferring protein activity from gene regulatory networks. Nature Genetics, in press