The Data Description

*golub.names.csv* is obtained by log\_10 transformation and standardization of arrays on acute leukemia gene expression data (Golub et al in [21]). *golub1.csv, golub2.csv, golub3.csv* and *leibiaoqian.csv* are obtained from *golub.names.csv. golubxin.txt* and *tezhengqunweizhi.txt* are obtained according to the gene groups after finishing grouping strategy.

The file *golub.names.csv* is the data of acute leukemia with 72 samples and 3571 genes.

The file *golub1.csv* is the data of BALL class with 38 and samples 3571 genes.

The file *golub2.csv* is the data of TALL class with 9 samples and 3571 genes.

The file *golub3.csv* is the data of AML class with 25 samples and 3571 genes.

The file *leibiaoqian.csv* is the labels of each sample.

The file *golubxin.txt* is the expanded data set with 72 samples and 10713 genes.

The file *tezhengqunweizhi.txt* is the position of groups which each gene belongs to.