GenePattern

ExpressionFileCreator Documentation

Module name: ExpressionFileCreator
Description: Creates a gct or res file from a set of cel files
Author: Joshua Gould
Date: 10/05/2004
Release: 1.0

Summary: Creates a res or gct file from a set of cel files. The cel files are given in one zip file. The cel files contained in the zip file may also be zipped. Algorithms for converting from probe level data to expression level data are MAS5 and RMA. If a class.filename is supplied, the samples are reordered so that all samples that belong to the same class occur together and a cls file is created. The sample names are taken from the base name of the cel files. For example, the zip file my.cel.files.zip contains the files sample1.cel, sample2.cel, and sample3.cel. The class.filename could contain the following text:

sample1   Tumor
sample2   Normal
sample3   Tumor

References:
1. Affymetrix. Affymetrix Microarray Suite User Guide. Affymetrix, Santa Clara, CA, version 5 edition, 2001
2. Rafael A. Irizarry, Bridget Hobbs, Francois Collin, Yasmin D. Beazer-Barclay, Kristen J. Antonellis, Uwe Scherf, and Terence P. Speed. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics, 2003b. To appear.

Usage/Example: efc.out <- ExpressionFileCreator (input.filename='my.cel.files.zip')

Parameters:

Name	Description
input filename	A zip file containing CEL files or zipped CEL files
output file	The base name of the output file
method	The method to use
quantile normalization	(RMA only) Whether to normalize data using quantile normalization
background correct	(RMA only) Whether to background correct using RMA background correction
scale value	(MAS5 only) Value at which all arrays will be scaled to
compute present absent calls	(MAS5 only) Whether to compute Present/Absent calls
normalization	The normalization method to apply after expression values are computed
sample index:	(mean or median scaling only) The sample index (starting at 1) to be used as a reference for normalization
class filename	tab-delimited text file containing one sample name and class per line

Return Value: An R list with components:

1. gct or res file
2. cls file if class.filename is supplied

Platform dependencies:

    Task type: Preprocess&Utility

CPU type: any
OS: any
Language: R
Support files: expr.R

Native command line: <R> <libdir>expr.R create.expression.file -i<input.filename> -o<output.file> -m<method> -q<quantile.normalization> -b<background.correct> -s<scale.value> - c<compute.present.absent.calls> -n<normalization> -x<sample.index> -f<class.filename> -l<libdir>