TCGA.Merge - version 1

24 October 2011 - Robert Borkowski, robert.borkowski@utsouthwestern.edu

A module for collating TCGA Level 3 expression data, distributed in a one-file-per-sample format to a single, tab-delimited file with one probe per row, and one column per sample.

Be sure to check for proper line endings!

Parameters:

"input": An archive (tested with .zip) containing BOTH the "file_manifest.txt" generated by TCGA during the data download, and each of the sample files you wish to collate.

"output": The desired output file name. Will have ".txt" appended.

Known issues:

-Leaves behind the uncompressed .txt files.

-Will probably die if there are discrepancies between the file manifest and the other contents of the archive. Make sure they match if you are trying to make multiple files or trying to deal with a subset of your TCGA download.