

Hu680035kAtoU95 Documentation

Module name: Hu680035kAtoU95
Description: Converts a list of Affymetrix Hu6800/Hu35KsubA probes to the corresponding Affymetrix U95Av2 probes.
Author: GenePattern gp-help@broad.mit.edu
Date: 4/15/03
Release: 1.0

Summary: This task takes a list of unique feature identifiers from Affymetrix Hu6800 (aka HuFL) and Hu35KsubA microarrays and returns the corresponding feature identifiers on the Affymetrix U95Av2 array.

Gene lists in either .pol, .tag or .grp file formats are compatible with this module. Output files have the same format as the input file. A .pol (*ie* parametrized ordered list) file is a tab-delimited format in which the first column is a ranking, the second column is the unique feature identifier, the third column is a text field (usually a description of the feature), and the last column is a value upon which the rank position is based. The columns of the output .pol files show the rank of the corresponding query feature identifier, the U95Av2 feature identifier, the U95Av2 feature description, and the query feature identifier. The .grp file contains only one unique feature identifier per row. A .tag file is a tab-delimited format; the first column contains unique feature identifiers and the second column contains unspecified content, usually a feature description. The second column of an output .tag file has the query feature identifier concatenated onto the U95Av2 feature description.

Note that some query probe identifiers map to multiple U95Av2 probes, and some have no features associated with them.

The feature conversion provided by the task is based on a look-up table curated by Ken Ross.

The Global Cancer Map (GCM) was generated using the Hu6800 and Hu35KsubA platforms and this utility is especially useful for extrapolation of findings made in this dataset to datasets generated with the more recent Hu95Av2 array.

References:

- Ramaswamy *et al* (2003) A molecular signature of metastasis in primary solid tumors. *Nature Genetics* 33(1):49-54
- Ramaswamy *et al* (2001) Multiclass cancer diagnosis using tumor gene expression signatures. *PNAS* 98(26):15149-54

Usage/Example:

```
GCMtoU95 <-  
function(input.filename, output, server=defaultServer)  
{  
  return (runAnalysis("GCMtoU95",  
    "input.filename"=input.filename.pol/.tag/.grp, "output"=output,  
    server=server))  
}
```

GenePattern

Parameters:

Name	Description
output	name of the output file (a file extension will be appended)
input filename	the .pol, .tag or .grp file to be converted

Return Value: An R list with components:

1. output.pol/.tag/.grp: a file containing U95Av2 feature identifiers
2. STDOUT: a report of the number of hits
3. STDERR: the standard error report from the program

Platform dependencies:

Task type:	Preprocess&Utilities
CPU type:	any
OS:	any
Java JVM level:	any
Language:	Perl
Support files:	GCM_to_U95_lookup

Native command line: <perl> <libdir>GCM_to_U95.pl <input.filename> <output>
<libdir>