



## ReannotateGCT Documentation

**Description:** Reannotate a GCT file with a CHIP file

**Author:** Warren Kaplan (Garvan Institute of Medical Research, Sydney Australia), [w.kaplan@garvan.org.au](mailto:w.kaplan@garvan.org.au)  
Kevin Ying (Garvan Institute of Medical Research, Sydney Australia), [k.ying@garvan.org.au](mailto:k.ying@garvan.org.au)  
Mark Cowley (Garvan Institute of Medical Research, Sydney Australia), [m.cowley@garvan.org.au](mailto:m.cowley@garvan.org.au)

**Date:** 13<sup>th</sup> June, 2014

**Release:** 7.7

### Summary

Reannotate a GCT file, that is, replace the Description column of a GCT file with information from a CHIP file:

The Module matches the probe set ID (column 1) of the GCT file with the probe set ID (column 1) of the CHIP file. The Description column of the GCT is then filled with values from the CHIP file:

Gene Symbol :: Gene Title

Some Illumina CHIP files have a 4<sup>th</sup> column (Entrez Gene ID), which this module also works well with.

### References & Links

Official Broad documentation on GCT and CHIP file formats:

[http://www.broadinstitute.org/cancer/software/genepattern/tutorial/gp\\_fileformats.html#gct](http://www.broadinstitute.org/cancer/software/genepattern/tutorial/gp_fileformats.html#gct)

[http://www.broadinstitute.org/cancer/software/genepattern/tutorial/gp\\_fileformats.html#chip](http://www.broadinstitute.org/cancer/software/genepattern/tutorial/gp_fileformats.html#chip)

### Parameters (\* = required)

Name	Description
gct file*	The GCT file to be annotated.

chip file	A premade CHIP file to select from a drop down menu. This list of chip files may not be completely up to date, and we point out that the chipDatabase module generally contains a more up to date list of chip files.
custom chip file	A custom chip file a user can specify. If specified, this will be used instead of the “chip file” above.
output.file*	The name of the re-annotated GCT file Default: <gct.file_basename>.gct

## Input Files

### 1. gct file

The GCT file where the Description column needs to be annotated

### 2. chip file

Select from a drop down list premade CHIP files

### 3. custom chip file

A user specified CHIP file. The prob set IDs should match the ID's in the GCT

## Output Files

### 1. <gct.file\_basename>.gct

GCT file with the reannotated Description column

### 2. stdout.txt

A useful file containing information about the run. If some probe set ID's could not found in the CHIP file, a message will be written here.

## Example Data

<ftp://ftp.pwbc.garvan.unsw.edu.au:10021//pub/genepattern/modules/ReannotateGCT/chip/Illumina.HumanWG6.V3.chip>

<ftp://ftp.pwbc.garvan.unsw.edu.au:10021//pub/genepattern/modules/ReannotateGCT/noannot.gct>

## Citing this module

Kaplan, W., Ying, K., Cowley, M. *ReannotateGCT* – a GenePattern module for annotating a GCT file with a CHIP file (not published).

## Platform Dependencies

<b>Module type:</b>	Annotation
<b>CPU type:</b>	any
<b>OS:</b>	linux or mac Tested on CentOS 6.2 (former versions tested on Ubuntu 10.10) untested on Windows
<b>software</b>	nil
<b>Language:</b>	Python >= 2.6

## Installation

We have created this module such that it can download chip files from Garvan Institute, or from your own institute, via the PUB custom parameter.

### Simple installation

Install this module from a ZIP file, then create a custom variable:

```
PUB= ftp://ftp.pwbc.garvan.unsw.edu.au:10021/pub
```

### Local installation

If you want to have this module connect to chip files stored on your local systems, then:

1. download all chip files from:

<ftp://ftp.pwbc.garvan.unsw.edu.au/pub/genepattern/modules/ReannotateGCT/chip/>

2. store these on your local FTP server, in a folder called:  
genepattern/modules/ReannotateGCT/chip
3. Create a Genepattern custom variable 'PUB' pointing to the root where you downloaded the chip files:

```
PUB=ftp://ftp.your.institute.org/pub
```