# Feature Counter Plugin

This plugin uses HT-Seq count which is developed by Simon Anders at EMBL Heidelberg (Genome Biology Unit). HT-Seq is free software distributed under GNU GPL. The plugin calls htseq-count to get the number of reads overlapping features. This plugin can be use for several counting applications like RNA-Seq, miRNA analysis and other transcript counting applications.

This plugin is barcode aware and requires a bed file of known features. If you have a gtf file of exons or gff file of features, you need to convert it to a bed file prior to using this plugin. The bed file can be then uploaded to the torrent server using the regular upload process

## Upload feature file in BED format

This plugin will work with generic bed format.

http://genome.ucsc.edu/FAQ/FAQformat.html#format1

Column 4 is the feature name, column 5 is the score and column 6 is the strand.

If you downloaded the refGene.txt file from UCSC genome browser

http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/refGene.txt.gz

The format is below

738 NM_033196 chr19 - 20115226 20150277 20116813 20150154 4 20115226,20133812,20135058,20150151, 20118084,20133908,20135185,20150277, 0 ZNF682 cmpl cmpl 1,1,0,0,

892 NM_033194 chr17 + 40274755 40275371 40274868 40275348 1 40274755, 40275371, 0 HSPB9cmpl cmpl 0,

585 NR_026818 chr1 - 34610 36081 36081 36081 3 34610,35276,35720, 35174,35481,36081, 0 FAM138A unk unk -1,-1,-1,

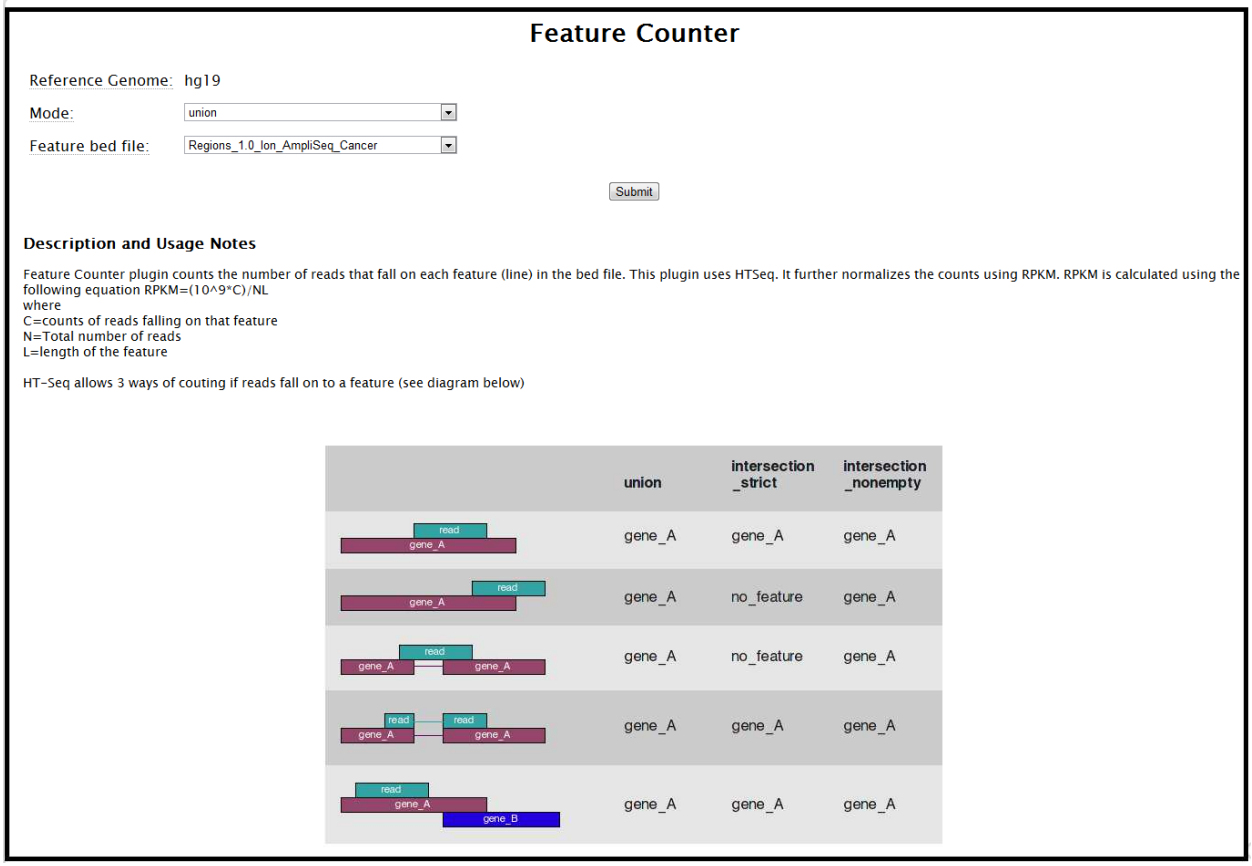
You can convert this to a bed format as below and upload it to the Torrent Server. You can set the score column to 0 or any other value between 0-1000. This value is ignored by the FeatureCounter plugin, but it is required in the bed file if you want the strand information to be used. If no strand information is given, featureCounter runs on stranded=no mode (does not use strand information).

| Chr | start | end | feature_name | score | strand |
|---|---|---|---|---|---|
| chr19 | 20115226 | 20150277 | NM_033196 | 0 | - |
| chr17 | 40274755 | 40275371 | NM_033194 | 0 | + |
| chr1 | 34610 | 36081 | NR_026818 | 0 | - |

## Running FeatureCounter plugin from the plugin framework

1. To run the FeatureCounter, download the plugin from the pluginstore, un-tar it and install it in your local plugins directory. After installing it, make sure that the plugin is enabled in the config tab of your torrent server. Also please make sure, you have python 2.6 or higher installed on the torrent server.

2. To run the plugin. Make sure the alignment is complete and the relevant bed file has been uploaded to the torrent server and then go to "Select Plugins To Run" and select FeatureCounter. When that is selected, a page to get the user parameters opens up.

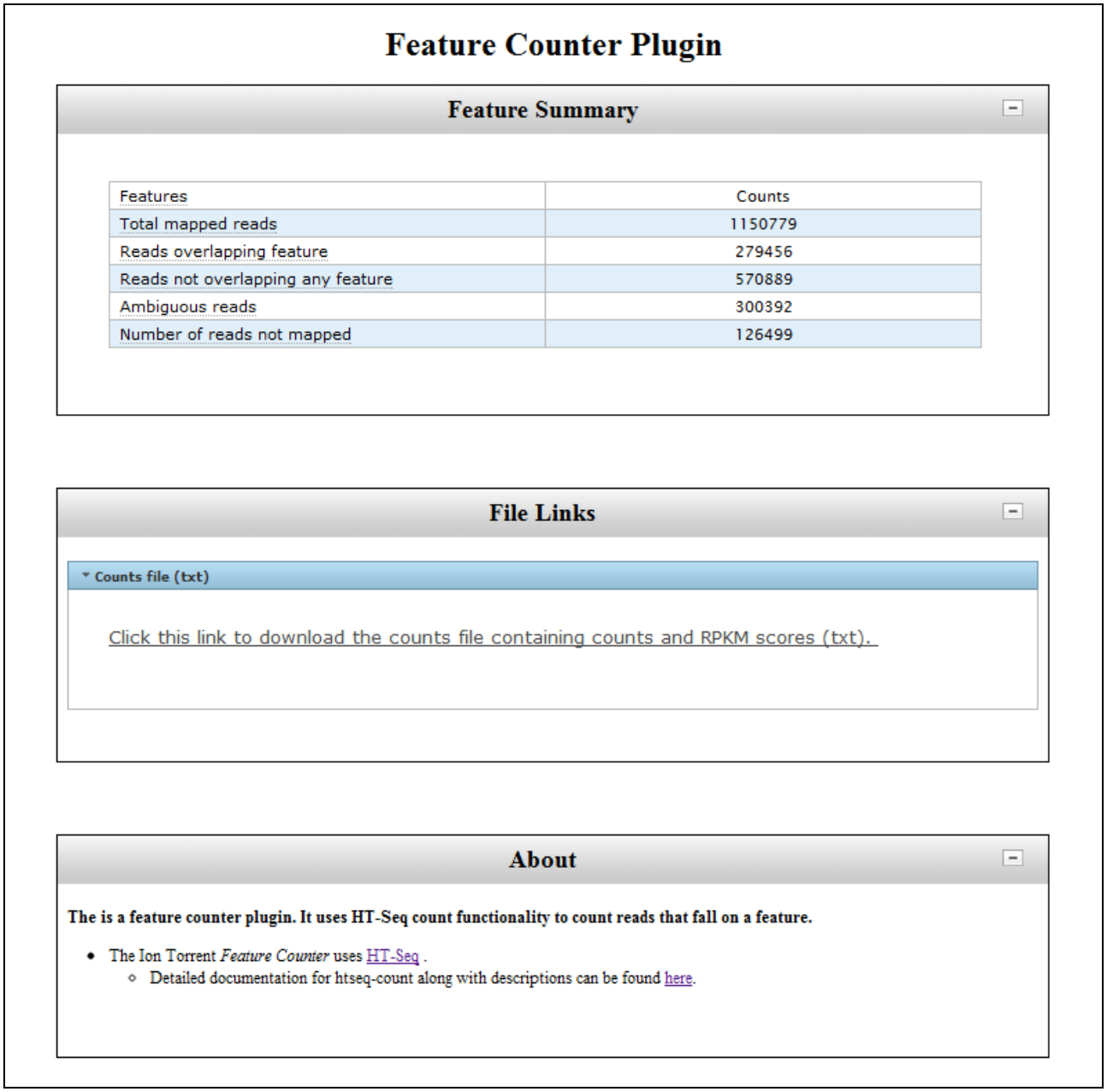



3. Select the mode. There are 3 modes available in htseq-count- union, intersection-strict, intersection-nonempty. You can get an explanation of the 3 modes using the instance.html or from the readme http://www-huber.embl.de/users/anders/HTSeq/doc/count.html

4. The feature bed file selectbox will get populated with the bed file entries for the particular reference genome. Make sure you have the relevant bed file uploaded before you start the plugin. Clicking on submit will kick off the feature counter.

## Interpreting results

1. When the plugin status says completed, html links to the plugin output will appear on your report page. This plugin is barcode aware. In case of a barcode experiment separate links will appear for each barcode. Click on the link to view results.



**Feature Counter Plugin**

**Feature Summary**

| Features | Counts |
|---|---|
| Total mapped reads | 1150779 |
| Reads overlapping feature | 279456 |
| Reads not overlapping any feature | 570889 |
| Ambiguous reads | 300392 |
| Number of reads not mapped | 126499 |

**File Links**

Counts file (txt)

Click this link to download the counts file containing counts and RPKM scores (txt).

**About**

**The is a feature counter plugin. It uses HT-Seq count functionality to count reads that fall on a feature.**

- The Ion Torrent *Feature Counter* uses HT-Seq .
  - Detailed documentation for htseq-count along with descriptions can be found here.

2. First table gives a summary of total mapped reads, number of reads overlapping feature, ambiguous reads etc. Reads that fall on more than 1 feature are classified as ambiguous reads.

3. File links provides a way to download the counts file in a .txt format. The output file is in the following format

   chr start end feature_id counts RPKM

   RPKM= 10^9*C/NL

   where
   C= count of the number of reads falling on that feature
   N= total mapped reads
   L= length of the feature