



## SegmentationCount Documentation

**Description:** Counts the number of segments in a SEG file.

**Author:** Gordon Saksena, [gp-help@broadinstitute.org](mailto:gp-help@broadinstitute.org)

### Summary

Counts the number of segments in the segmentation file, compares this count to a quality threshold cutoff of 2000, and outputs this information in a file. The threshold value of 2000 was selected based on looking at the distribution of data from 10,000+ samples, and selecting a bound between the main distribution and its outliers. The segmentation count is a way of measuring hyper-segmentation. Typical samples should contain hundreds of segments; if the seg file claims that it contains 10's of thousands, this is unlikely to reflect the underlying biology.

### Parameters

Name	Description
input file	A segmentation file in SEG format.
output file	The name of the output file. Default: <input.file_basename>.segcount.txt

### Output Files

1. <output file>(.segcount.txt)

A file that contains the segmentation counts and threshold cut off value for each sample.

### Platform Dependencies

**Module type:** SNP Analysis

**CPU type:** any

**OS:** any

**Language:** Python 2.5

### GenePattern Module Version Notes

Date	Version	Description
11/07/12	1	Initial version.

# GenePattern