

**The Cancer Genome Atlas Pilot Project**  
**Broad Institute of MIT & Harvard**  
**Segmented Data Description**

**Platforms:** Affymetrix SNP 6.0 Array & Affymetrix HTA Array

**Segmented Data**

Copy number - Affymetrix SNP 6.0 Array

This is a tab-delimited file format that contains the output results of the GLAD module. The GLAD module, a SNP analysis module, runs the R package *Gain and Loss analysis of DNA* (GLAD) [Hupè et al., 2004] , which detects segments of the genome which have altered copy numbers. The GLAD file format is organized as follows:

1. The first line contains a list of labels identifying the columns.
  - o Line format: Sample (tab) Chromosome (tab) Start.bp (tab) End.bp (tab)  
Num.SNPs (tab) Seg.CN
2. The rest of the file contains one row of data for each altered chromosomal region.
  - o Line format: (sample) (tab) (chromosome) (tab) (startPosition) (tab)  
(endPosition) (tab) (numberOfSNPs) (tab) (regionCN)
  - o For example: MYNAH\_p\_Affy\_plate\_9\_Mapping250K\_Sty\_A02\_49084  
(tab) 17 (tab) 41419603 (tab) 36581538 (tab) 6427 (tab) 2.06

**Segmented Data**

Loss of Heterozygosity - Affymetrix SNP 6.0 Array (.seg-loh file)

This is a tab-delimited file format that contains the output results of a hidden Markov model to identify continuous segments of LOH. The file format is organized as follows:

1. The first line contains a list of labels identifying the columns.
  - o Line format: Sample (tab) Chromosome (tab) Start.bp (tab) End.bp (tab)  
Num.SNPs
2. The rest of the file contains one row of data for each region of LOH.
  - o Line format: (sample) (tab) (chromosome) (tab) (startPosition) (tab)  
(endPosition) (tab) (numberOfSNPs)