SCLC mutation analysis steps -

Identifying significantly mutated genes in SCLC - Lists of silent and non-silent mutations were either obtained from published data directly or obtained with the following analysis pipeline. FASTQ files obtained from database center (EGAD00001000222) were processed with MuTect 1.1.4 to identify somatic mutations, the functional consequences of the mutations were annotated with Oncotator (v1.8.0.0) with default parameters. Finally, combined mutation data was processed with MutSigCV 1.0 with full coverage and covariates files as control. Significantly mutated genes are determined based on p-values calculated by MutSigCV.

Alignment - Raw sequences to hg19 human reference genome sequence (Bowtie2)

Post-alignment processing - Remove duplicates and process aligned files (Picard, Samtools)

Variant calling - Using MuTect on Tumor-Normal sample pairs

Variant annotation – Using Oncotator

List of silent and nonsilent mutations for MutsigCV

**MuTect –**

module add java/1.6.0\_31

export \_JAVA\_OPTIONS="-Xmx2g"

java -jar muTect-1.1.4.jar --analysis\_type MuTect --reference\_sequence ucsc.hg19.fasta --dbsnp dbsnp\_138.hg19.vcf --cosmic b37\_cosmic\_v54\_120711.vcf --intervals hg19intervals\_ucsc.bed --input\_file:normal reordered98739\_hg19.bam --input\_file:tumor reordered98735\_hg19.bam --out 98735\_hg19.call\_stats.txt --coverage\_file 98735\_hg19.coverage.wig.txt --vcf 98735\_hg19.vcf

**Oncotator –**

module load oncotator

oncotator --db-dir /ufrc/zhou/share/projects/Oncogenomics\_varsha/oncotator/oncotator\_v1\_ds\_Jan262014 -i transformed\_SCLC.muTect.vcf -o transformed\_SCLC\_onco.tsv hg19

**MutSigCV -**

<http://software.broadinstitute.org/cancer/software/genepattern/modules/docs/MutSigCV>

Format for MutsigCV -

gene patient effect categ

VCAM1 S00022 silent 3

PLOD1 S00022 nonsilent 4

NBPF7 S00022 nonsilent 3