CREATE TABLE test AS

SELECT

GSO.gene\_name,SA.id\_sample,

upper(REPLACE(sa.sample\_name, '-','')) as sample\_name,gsm.id\_mutation, SM.AA\_mut\_start, SM.AA\_MUT\_STOP,length(TR.TRANSCRIPT\_AA\_SEQ) as AA\_len,

SM.AA\_MUT\_SYNTAX,

GSO.swissprot\_accession,

SA.ID\_SOURCE\_SPEC,SA.ID\_source,

GSM.ID\_MUT\_SOMATIC\_STATUS,

gsm.ID\_MUT\_VERIF\_STATUS,

 SA.ID\_SOURCE\_TISSUE\_ORIGIN,

 TR.accession\_number, GS.WHOLE\_GENE\_SCREEN,

GS.WHOLE\_GENOME\_SCREEN FROM ANALYSED\_GENE\_SAMPLE T, GENE\_STUDY GS,

Gene\_sample\_mutation GSM,

sequence\_mutation SM ,

cosmic\_tumour\_sample\_overview SA,

gene\_study\_transcript GST,

transcript TR,

Gene\_SOM GSO

WHERE

GS.id\_gene\_study = T.id\_gene\_study AND

 GSM.id\_ags=T.ID\_AGS AND

SM.id\_mutation=GSM.id\_mutation AND

 SA.id\_sample=T.id\_sample AND

 GST.id\_gene\_study=GS.id\_gene\_study AND

 TR.id\_transcript = GST.ID\_TRANSCRIPT AND

GSO.id\_gene=TR.id\_gene AND

GSM.ID\_MUT\_SOMATIC\_STATUS in (1, 2) AND GSM.ID\_MUT\_VERIF\_STATUS=50 AND

SM.id\_mut\_type\_AA=22 AND

(GS.WHOLE\_GENE\_SCREEN='y' OR

GS.WHOLE\_GENOME\_SCREEN='y' OR

sa.sample\_name like 'TCGA%' OR

sa.sample\_name like 'ICGC%' OR

sa.sample\_name like 'CGP%' OR

sa.sample\_name like 'LUAD%' OR

sa.sample\_name like 'LC\_%')

ORDER BY

gs.id\_gene,SM.AA\_mut\_start,sa.sample\_name;