The purpose of the ANNOVAR scrip is to annotate the traits and report functional importance in annovar format.

The main parts of the script are in these 3 steps.

# define a function (runsteps)

# we will use this function multiple times

def runsteps(targetfile):

filename = targetfile.split('.')[0] # split file by '.'

print(filename);

snpfile = filename + '.snplist.txt';

print(snpfile);

#argument1 = targetfile + ">" + snpfile;

#run linux cmd

open (snpfile, "w")

subprocess.call(["cut", "-f3", targetfile], stdout=snpfile);

sumstatsfile = filename + '.sumstats.add.avinput.txt';

print(sumstatsfile);

argument2 = './convert2annovar.pl -format rsid ' + snpfile + ' -dbsnpfile humandb/hg19\_snp138.txt ';

print(argument2);

args1 = shlex.split(argument2)

print(args1);

#run perl script

open (sumstatsfile, "w")

p = subprocess.Popen(args1, stdout=sumstatsfile)

argument3 = './table\_annovar.pl ' + sumstatsfile + ' humandb/ -buildver hg19 -out ' + filename + ' -remove -protocol refGene,cytoBand,exac03,avsnp147,dbnsfp30a -operation gx,r,f,f,f -nastring . -csvout -polish -xref example/gene\_xref.txt';

print(argument3);

args2 = shlex.split(argument3)

print(args2);

#run perl script

p = subprocess.Popen(args2, stdout=subprocess.PIPE)

By inputting the trait target files, this code processes the file with annovar related systems and allows the trait file to be converted into an annovar annotated format.