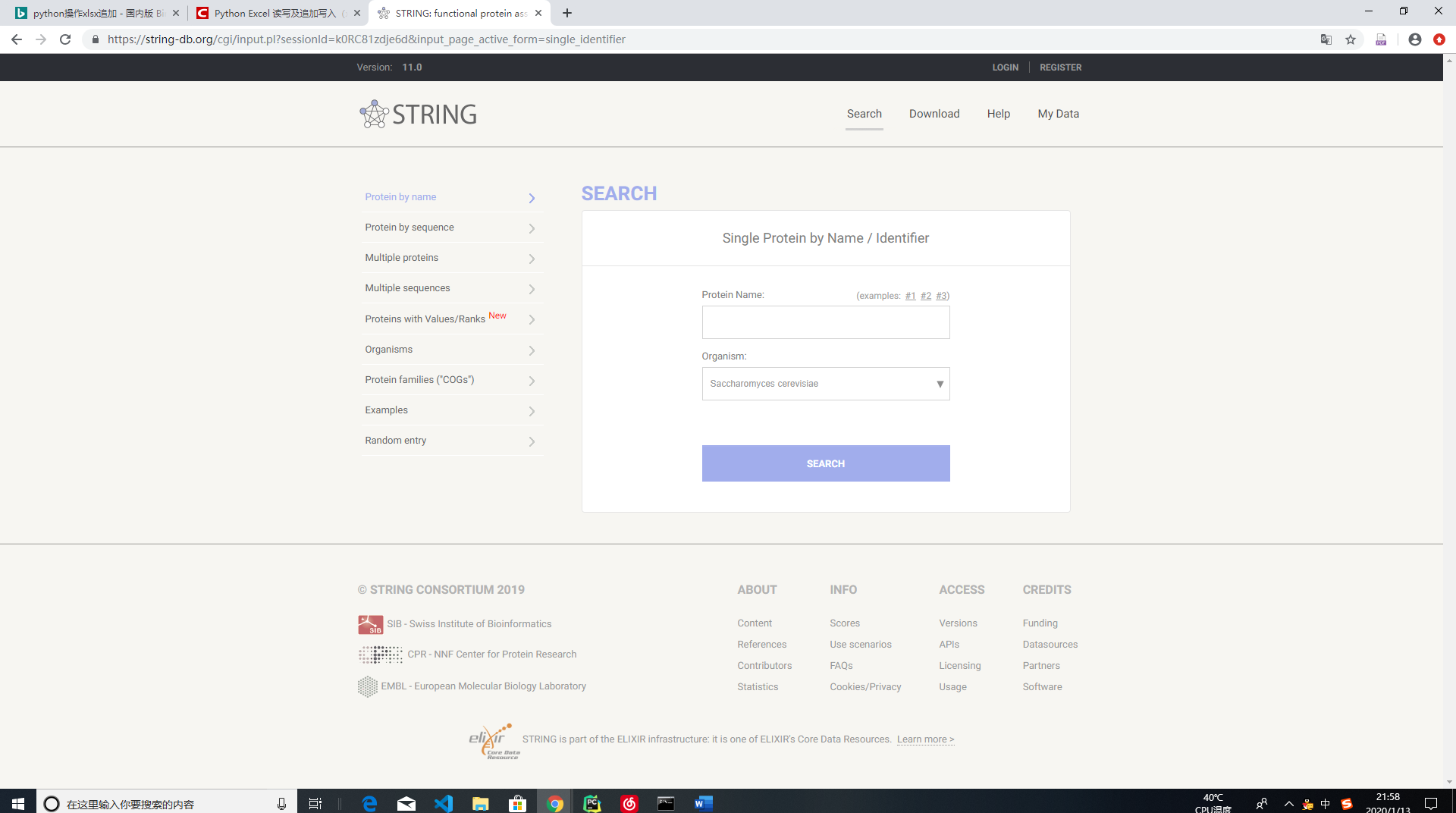
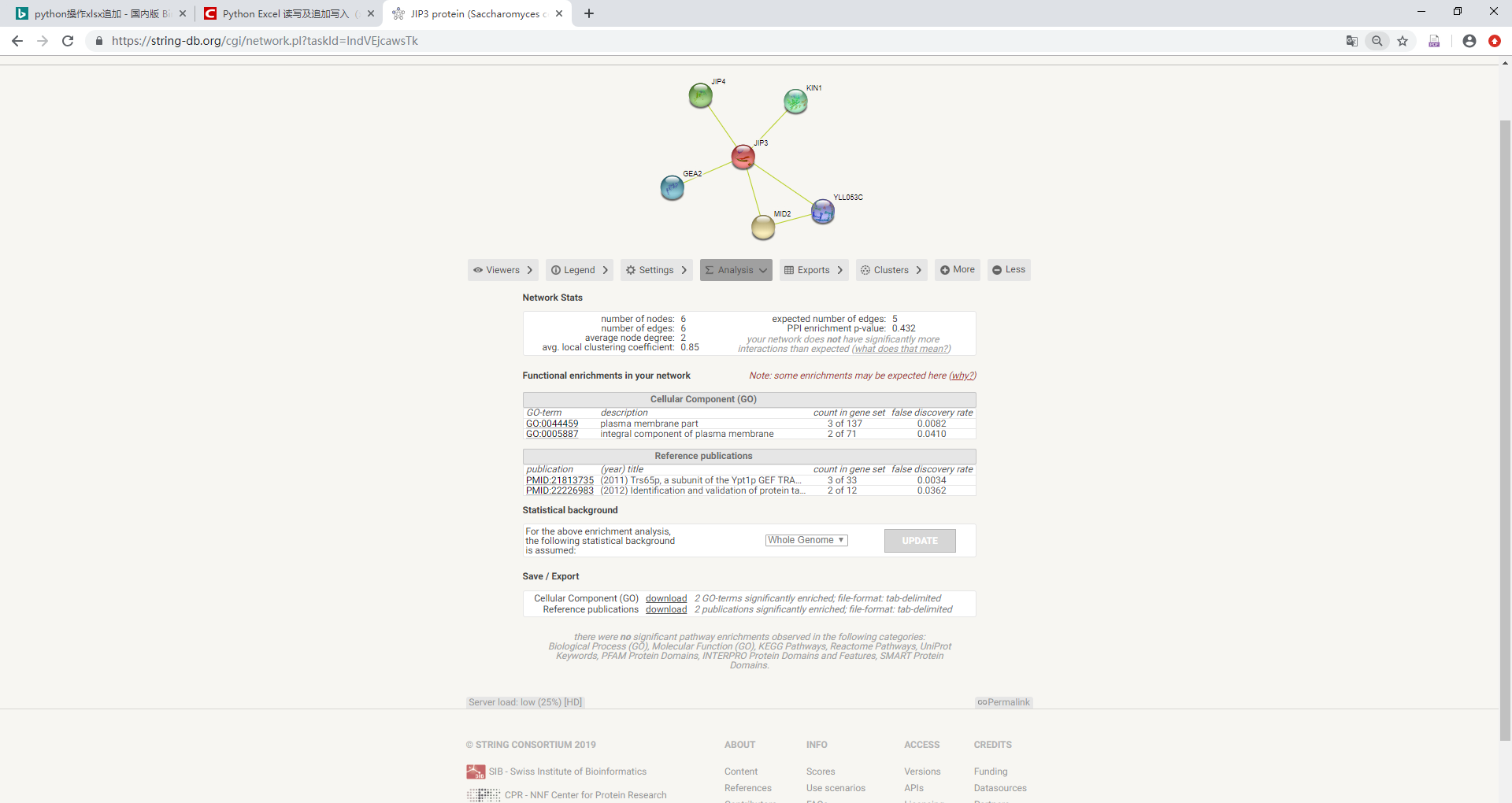
代码：**import** pandas **as** pd  
**import** os  
**import** re  
  
**import** csv  
**import** openpyxl  
**import** xlrd  
  
**from** xlutils.copy **import** copy  
  
**from** openpyxl **import** Workbook  
  
**from** xlutils.copy **import** copy  
**from** xlrd **import** open\_workbook  
  
**import** warnings  
  
  
warnings.filterwarnings(**"ignore"**)  
warnings.filterwarnings(**"ignore"**)  
  
  
**def** Get(s2,s):  
  
 *#把当前路径川递过来  
 #print(s)* MF = s+**"/enrichment.Function.tsv"** BP = s+**"/enrichment.Process.tsv"** CC = s+ **"/enrichment.Component.tsv"** *#print(F)  
 #print(MF,BP,CC)  
  
 #print(os.path.exists(MF))  
 #print(os.path.exists(CC))  
 #print(os.path.exists(BP))  
  
 #train=pd.read\_csv('enrichment.Function.tsv', sep='\t',header=0)* BP\_train = pd.read\_csv(BP, sep=**'\t'**,header=0)  
 CC\_train = pd.read\_csv(CC, sep=**'\t'**,header=0)  
 MF\_train= pd.read\_csv(MF, sep=**'\t'**,header=0)  
  
  
  
  
*##train["#term ID"]* ID\_BP = BP\_train[**"#term ID"**]  
 ID\_MF = MF\_train[**"#term ID"**]  
 ID\_CC = CC\_train[**"#term ID"**]  
 *# print(ID)  
  
 #print(ID\_MF) 读取出来  
 #print(ID\_BP)  
 #print(ID\_CC)  
  
 ##ID.to\_csv('data.csv')  
  
 #print(ID\_MF[0])  
 #print(ID\_MF[1])  
 #print(ID\_MF[2])  
 #print(ID\_MF[3])* mf=**""** cc=**""** bp=**""  
  
 for** num **in** range(0,len(ID\_CC)):  
 *#print(num)  
 #print(ID\_MF[num])* **if**(num+1!=len(ID\_CC)):  
 cc=cc+(ID\_CC[num])+**";"** *#print(mf)* **else**:  
 cc=cc+(ID\_CC[num])  
  
  
  
  
 **for** num **in** range(0,len(ID\_MF)):  
 *#print(num)  
 #print(ID\_MF[num])* **if**(num+1!=len(ID\_MF)):  
 mf=mf+(ID\_MF[num])+**";"** *#print(mf)* **else**:  
 mf=mf+(ID\_MF[num])  
  
  
  
  
 **for** num **in** range(0,len(ID\_BP)):  
 *#print(num)  
 #print(ID\_MF[num])* **if**(num+1!=len(ID\_BP)):  
 bp=bp+(ID\_BP[num])+**";"** *#print(mf)* **else**:  
 bp=bp+(ID\_BP[num])  
  
  
  
 print(mf+**"\n"**)  
  
 print(cc+**"\n"**)  
 print(bp)  
 *#yeast = pd.read\_csv("./processo.csv")  
  
 #print(yeast)  
 #yeast["Portain\_name"] = s2  
 #yeast["Gene othology(biological process)"]=bp  
 #yeast["Gene othology(molecular function)"]=mf  
 #yeast["Gene othology(cellular component)"]= cc  
  
 #print(yeast)  
 #YLR331C  
  
 #yeast\_1 = pd.read\_csv("./data.csv")  
 #yeast\_1.to\_csv("./data.csv",mode="a",index=False,columns=["Portain\_name","Gene othology(biological process)","Gene othology(molecular function)","Gene othology(cellular component)"])  
 #yeast = pd.read\_csv("./data.csv")  
 # print(yeast)  
 #print(s2)  
 #print(yeast["Portain\_name"])  
 #print(yeast["Gene othology(cellular component)"])  
 #print(yeast["Gene othology(molecular function)"])  
 # print("完成")  
  
 # with open("./data.csv","a+",newline='') as file:  
 # csv\_file = csv.writer(file)  
 # datas = [s2,bp,cc,mf]  
 # csv\_file.writerows(datas)* **"""  
 data = {"Portain\_name": s2,  
 "Gene othology(biological process)":bp,  
 "Gene othology(cellular component)":cc,  
 "Gene othology(molecular function)":mf  
 }  
 """** data = [s2,bp,cc,mf]  
 print(data)  
 a =[s2,bp,cc,mf]  
 *#for row in a:  
 #print(row)  
 #print(type(a))  
 #print(a)  
 #add(a)* add\_1(a)  
 *#print(type(a))  
 #for row in a:  
 #print(row)  
  
 #write\_excel\_xls\_append("./DATA.xls",data)  
 # read\_excel\_xls("./DATA.xls")  
  
  
  
 # write\_excel\_xlsx(path, sheet\_name, value):***def** write\_excel\_xls\_append(path, value):  
 index = len(value) *# 获取需要写入数据的行数* workbook = xlrd.open\_workbook(path) *# 打开工作簿* sheets = workbook.sheet\_names() *# 获取工作簿中的所有表格* worksheet = workbook.sheet\_by\_name(sheets[0]) *# 获取工作簿中所有表格中的的第一个表格* rows\_old = worksheet.nrows *# 获取表格中已存在的数据的行数* new\_workbook = copy(workbook) *# 将xlrd对象拷贝转化为xlwt对象* new\_worksheet = new\_workbook.get\_sheet(0) *# 获取转化后工作簿中的第一个表格* **for** i **in** range(0, index):  
 **for** j **in** range(0, len(value[i])):  
 new\_worksheet.write(i+rows\_old, j, value[i][j]) *# 追加写入数据，注意是从i+rows\_old行开始写入* new\_workbook.save(path) *# 保存工作簿* print(**"xls格式表格【追加】写入数据成功！"**)  
  
  
  
  
**def** write\_excel\_xls(path, sheet\_name, value):  
 index = len(value)  
 workbook = openpyxl.Workbook()  
 sheet = workbook.active  
 sheet.title = sheet\_name  
 **for** i **in** range(0, index):  
 **for** j **in** range(0, len(value[i])):  
 sheet.cell(row=i+1, column=j+1, value=str(value[i][j]))  
 workbook.save(path)  
 print(**"xlsx格式表格写入数据成功！"**)  
  
**def** add(a):  
 book = Workbook()  
 sheet = book.active *# get a first sheet* rows = ((1,2,3),  
 (2,3,4),  
 (4,5,6)  
  
 )  
 *#print(type(rows))  
 #print("---------")  
 #rows = a* **for** row **in** rows:  
 sheet.append(row)  
  
 book.save(**"./DATA.xlsx"**)  
  
**"""  
 book\_ro = open\_workbook("./DATA.xls")  
 book = copy(book\_ro) # creates a writeable copy  
 sheet1 = book.get\_sheet(0) # get a first sheet  
  
 colx = 1  
 for rowx in range(1):  
 # Write the data to rox, column  
 sheet1.write(rowx,colx, url)  
 sheet1.write(rowx,colx+1, count)  
  
 book.save("./DATA.xls")  
"""  
  
  
  
  
  
  
  
  
  
  
def** add\_1(a):  
 data = openpyxl.load\_workbook(**'DATA.xlsx'**)  
 *# 取第一张表* sheetnames = data.get\_sheet\_names()  
 table = data.get\_sheet\_by\_name(sheetnames[0])  
 table = data.active  
 print(table.title) *# 输出表名* nrows = table.max\_row *# 获得行数* print(**"已经有"**,nrows,**"行"**)  
 ncolumns = table.max\_column *# 获得最大列数* print(**"此时已经有"**,ncolumns,**"列"**)  
  
 *#print(table.cell(1,1)) #至少为一 （这是获取位置）  
 #print(table.cell(1,1).value) #获取这个位置的值  
  
 # 写入成功  
 #table.cell(1,2).value = "TEST"  
 #data.save("./DATA.xlsx")  
  
 #print(a[1])  
 #table.cell(1,2).value = a  
 #data.save("./DATA.xlsx")* **"""  
 for value in a:  
 print(value)  
 """** nrows = table.max\_row *# 获得行数* ncolumns = table.max\_column *# 获得行数* values = a  
  
 **"""  
 for value in values:  
 print(value)  
 print(len(values))  
 """** *#print(values)  
 #for i in range(0,len(values)):  
 #print(values[i])* **for** i **in** range(0,len(values)):  
 table.cell(nrows+1,i+1).value = values[i]  
 nrows= nrows + 1  
 data.save(**"./DATA.xlsx"**)  
  
 print(**"写入成功 "**)  
  
 *#行数加以  
 #nrows = nrows + 1  
  
 #data.save("./DATA.xlsx")* **"""  
 for value in values:  
 #table.cell(nrows+1,1).value = value  
 for j in range(1,ncolumns+1):  
 table.cell(nrows+1,j).value = value  
 #nrows = nrows + 1  
 nrows = nrows + 1  
 #nrows = nrows + 1  
 #print(value)  
  
 data.save('DATA.xlsx')  
  
 """  
  
if** \_\_name\_\_ ==**"\_\_main\_\_"**:  
 **while True**:  
 filename = input(**"请输入你要操作的蛋白质名称"**)  
 s2=filename  
 s=**"./file/%s"** % filename;  
 **if**(os.path.exists(s)==**True**): *#如果存在该蛋白质的信息* Get(s2,s)  
*#print(s)* **else**:  
 print(**"不存在该蛋白质的信息，请补充"**)  
  
  
 *#Get(filename)*

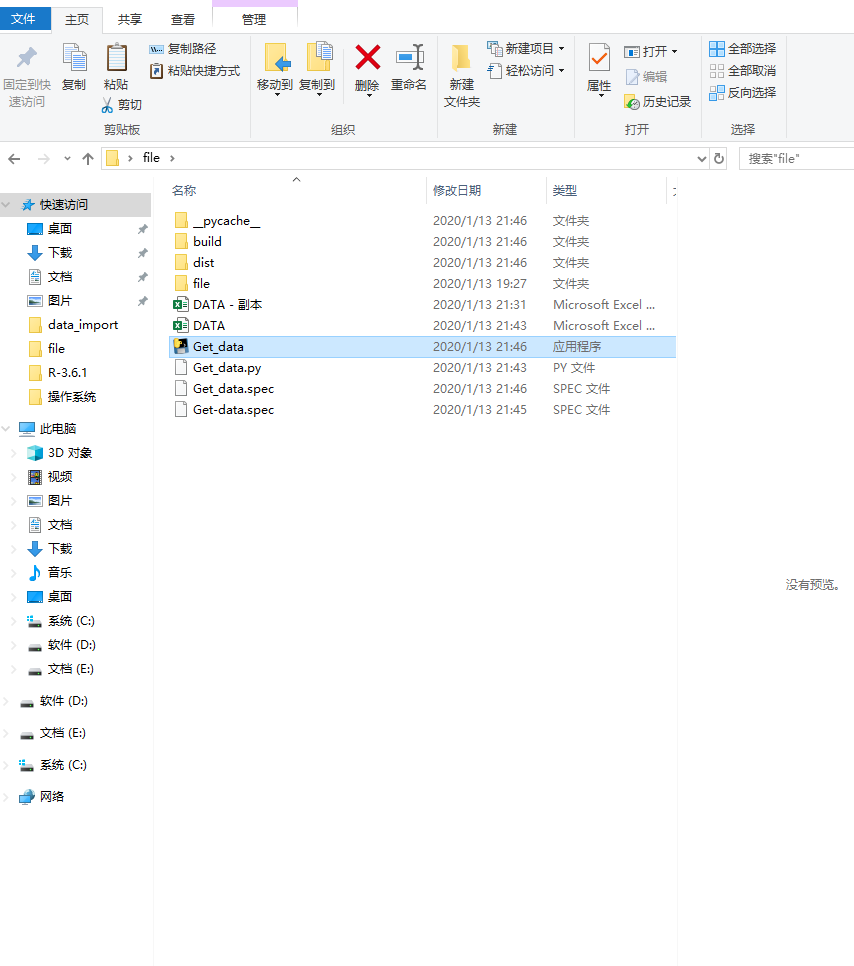
使用流程：进去网站<https://string-db.org/>搜索关于酿酒酵母蛋白质

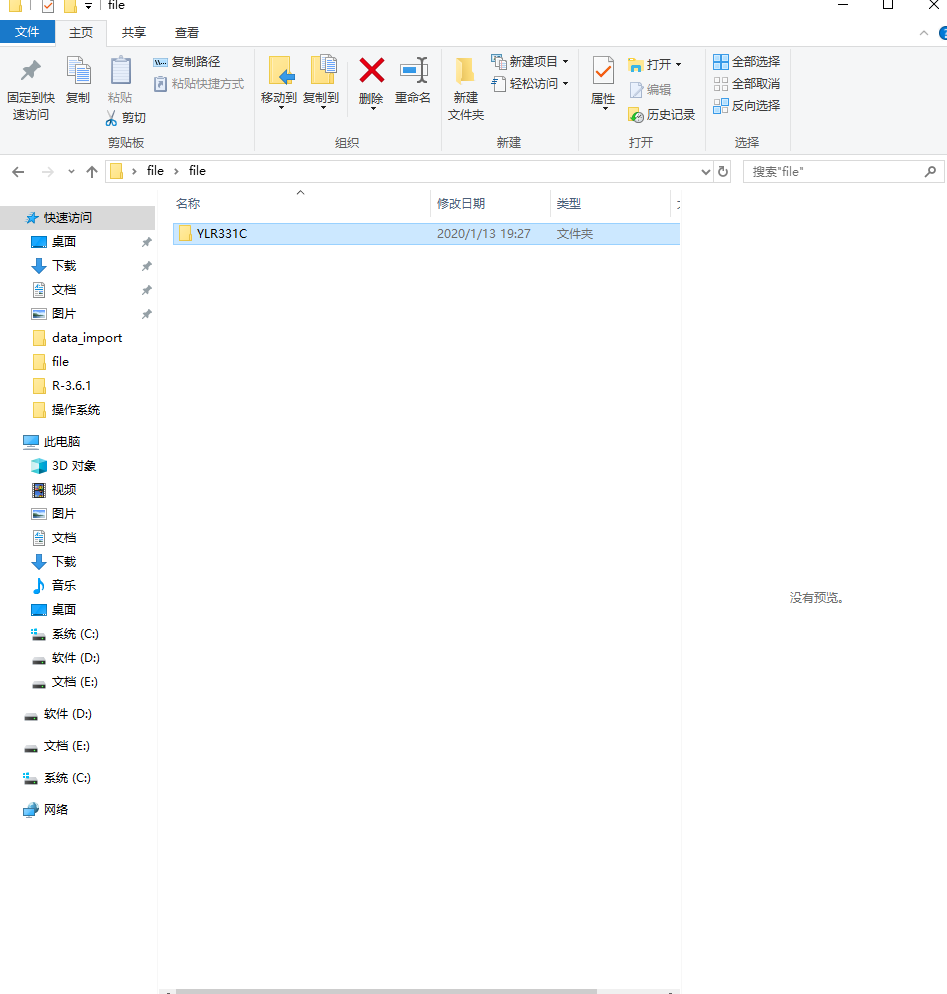


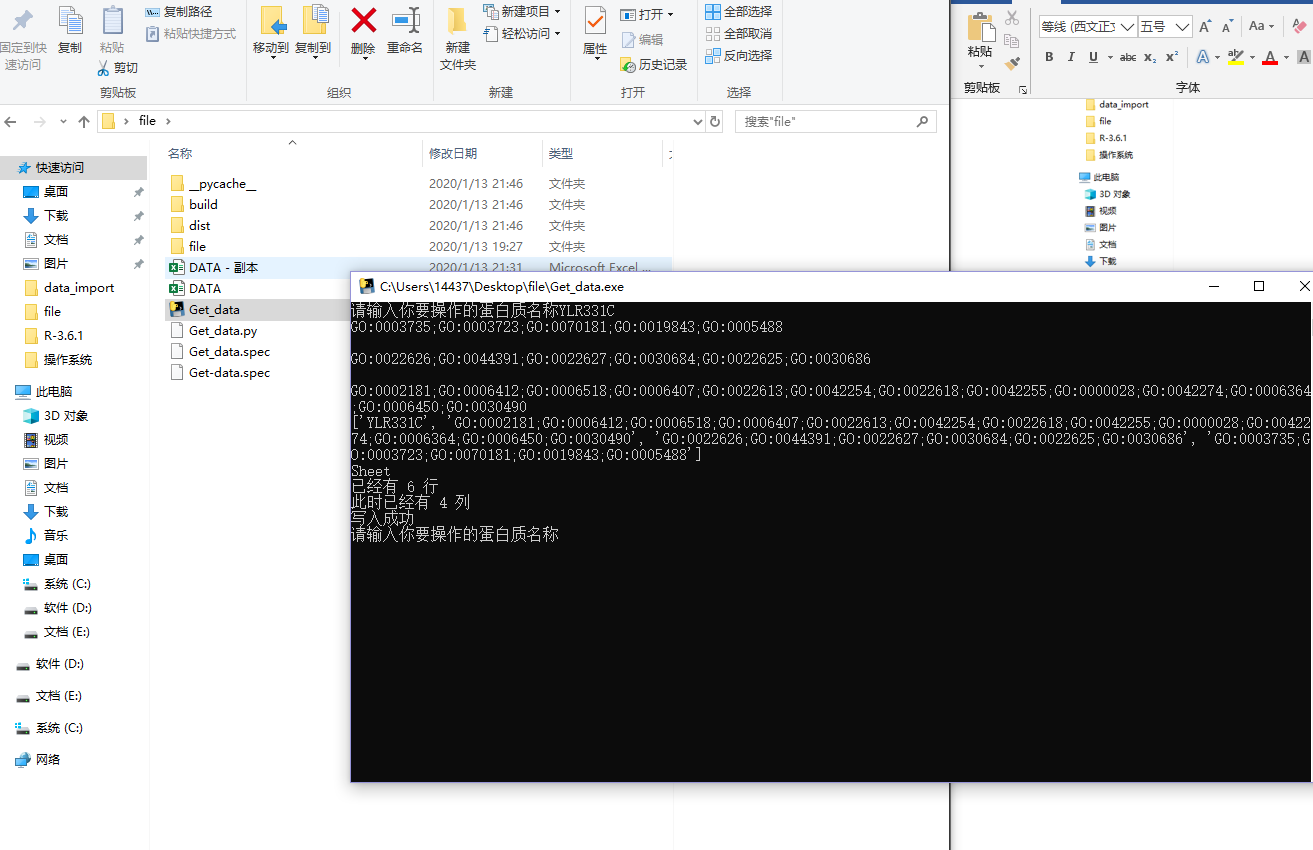
下载蛋白质本体的三种GO本体数据 （包括细胞成分 生化过程 功能作用）

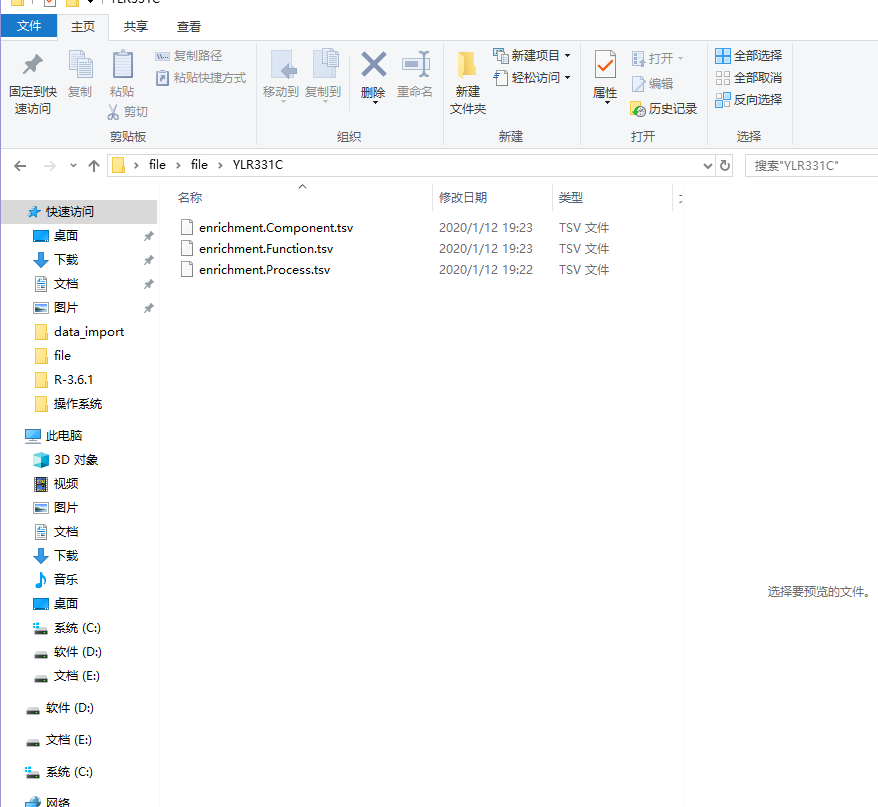


保存到相应file 文件夹 保存类型为文件夹名字 为蛋白质名称 里边包含三种 go本体









点击Get\_data 运行 输入蛋白质名称 然后运行就保存到excel中然后 数据用R语言进行比较

