

Как загрузить файл сfeature_table.txt корректно

- Откройте нужную чистую страницу в Google sheet
- В меню File => Import => Upload => Browse => выберите файл => Import location: Replace current sheet => Import data

Объяснение почему у некоторых студентов не получилось при следовании старой инструкции см. на следующих страницах

При копировании текста файла на страницу Google sheet ВАЖНО какой программой открыт файл ..._feature_table.txt

- При щёлкании на имя файла вызывается программа, которая ассоциирована с расширением .txt
- Если это Libre office (writer), Word (про Mac не знаю) или другой продвинутый редактор текстов, то ничего не получится; почему см. с 4го слайд
- Программа **Libre office calc** предназначена для таблиц (как Excel в windows). **Её можно использовать:** открыть её и в ней открыть файл _feature_table.txt и скопировать его в Google sheet
- **Можно использовать Блокнот (win), Mousepad (linux).**
- Как открыть правильной программой см. далее

Take home message

Всегда следите за тем, какой программой открываете файл.
Не соглашайтесь слепо с тем, что за вас решает компьютер

Выбор просмотрщика текста для копирования таблицы в Google sheet

- Курсор на файл .txt
- Щёлкнуть ПРАВОЙ кнопкой => Open with (Открыть с помощью) => Выбрать другое приложение
- То, которое сейчас стоит по умолчанию, стоит сверху
- Если не годится умолчательное выберите подходящее из списка или среди Других приложений

В чем разница между редакторами текстов?

- Каждый символ кодируется байтом (8 бит: от 0, до 255)
- В примитивном редакторе или просмотрщике символы так и показываются по строкам, один за другим. Для нашей цели это и нужно
- В продвинутом текстовом редакторе кодируется много чего ещё, размеры отступов и другое форматирование, которое не понятно
ЭТаблицам

Вид ..._feature_table.txt в примитивном редакторе. На символ Табулятор он реагирует сдвигом к следующей колонке, на символ Конца строки – переносом на следующую строку. Остальные символы подряд

#	feature	class	assembly	assembly_unit	seq_type	chromosome	genomic_accession	start	end	strand	product
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	201	2087	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	201	2087	+		WP_012619398.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	2282	3106	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	2282	3106	+		WP_009873964.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3143	3382	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3143	3382	+		WP_009873965.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3501	3986	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3501	3986	+		WP_009873966.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3986	4519	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	3986	4519	+		WP_014498806.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	4534	6072	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	4534	6072	+		WP_009873968.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	6105	6977	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	6105	6977	+		WP_009873969.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	7001	8398	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	7001	8398	+		WP_009873970.1	
gene	protein_coding	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	8419	8841	+			
CDS	with_protein	GCF_000183285.1	Primary Assembly	chromosome	NC_017253.1	8419	8841	+		WP_029585514.1	

Вид ..._feature_table.txt в Libre Office (writer). Как форматировать текст он решает сам

```
# feature class assembly assembly_unit seq_type chromosome
genomic_accession start end strand product accession non-
redundant_refseq related_accession name symbol GeneID locus_tag
feature_interval_length product_length attributes
gene protein_coding GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 201 2087 + mnmG
CWS_RS00005 1887 old_locus_tag=CWS_00005
CDS with_protein GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 201 2087 + WP_012619398.1 WP_012619398.1
tRNA uridine-5-carboxymethylaminomethyl(34) synthesis enzyme MnmG
mnmG CWS_RS00005 1887 628
gene protein_coding GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 2282 3106 + atpB
CWS_RS00010 825 old_locus_tag=CWS_00010
CDS with_protein GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 2282 3106 + WP_009873964.1 WP_009873964.1
F0F1 ATP synthase subunit A atpB CWS_RS00010 825 274
gene protein_coding GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 3143 3382 + atpE
CWS_RS00015 240 old_locus_tag=CWS_00015
CDS with_protein GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 3143 3382 + WP_009873965.1 WP_009873965.1
F0F1 ATP synthase subunit C atpE CWS_RS00015 240 79
gene protein_coding GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 3501 3986 +
CWS_RS00020 486 old_locus_tag=CWS_00020
CDS with_protein GCF_000183285.1 Primary Assembly chromosome
NC_017253.1 3501 3986 + WP_009873966.1 WP_009873966.1
F0F1 ATP synthase subunit B CWS_RS00020 486 161
```

Вид в ..._feature_table.txt в Libre Office (writer), т.е. в формате .odt. Документ перекодирован так, чтобы описать форматирование – отступы, длины строк, табуляторы заменены какими-то другими символами. Поэтому копировать из Libre office writer в Google sheet нельзя.

```
D:\debi_aba\education\y23\Blook3\sequences\GCF_000183285.1_ASM18328v1_feature_table.odt | ANSI | Ln
PK♦♦♦ sF]W^Ж2Q' ' mimetypeapplication/vnd.oasis.opendocument.textPK♦♦♦ sF]W → Configuration
a¶RS†¶ 'f@k LK,,t¶LeNKEЗ-yQ9БУнвЗ`SYfÿ<`p+¶ПШЮ€¶|e^aS°\fЯ|•hzoczUUEcЯ#”`ь2o0к»;уязю.Пя↓vİrvDl@OJгъĥзгт▼UGİPK•¶гчт¶€
дн†f”5V0¶Ж§)Mÿн←щqajđ6I=Люв ♥€}¶)}ow▼эИеркЯ+`к2єBІ¶¶ r0Rй¶İö}oПи±юRь!Q0
ід-•нГ▲<đ|Vнль{0B7<đГьr...f=ëB<B¶ -.-вН±лÿU1|Bwн†0г№6NDЖ¶¶дс ЯЯ←эб)эјуr8Fсє°Iћэх0нĚ¶;7b1·PR`и`s36$JрЖ0~юЧВМ▲
rVKoi“|yS<ндћ.../
Жİ(Хн▼00QSG;•0Лель▲гум+э•0-ЭЗЮА(ММы
M£,ÿкYыA4ь „,đİ.0' |»hYкщрэ*ЙZ†K•>Fbrб|€лФ!<ШS!?'9)«бпрj*ан`bьц0Obьf•İV[P`™7¶юÿë...цK=8<h”Dдбл▼5E?єЛЯЗSPSP@. f5iEб<Hн♦р↓!
УНГггг$0@ь9)3»0|вмwIo¶яюНЮ...|\RKÿудаZ%[->]’Н300%zx¶ЦІJ/+gİ0+İ@G&-zzY№←μ0>T>]эхCL&2,Zlÿ,0нOлxЦCb+|]»ZюшI’Y|ЗВІWУ”@WуQНн -İ
сс¶¶]№†6fГtttГль№f←!!Г.й!!ÿьЮв0†‘ћН0Г!!д>Жбо`60tv»ШПдкxм€O[ІЖ&_±•};пвным†<э$)g±зg¶¶$лдХБсf%€=zh•0fк1B•XNvГхо`к0х Oь0&↓ÿëuo>
сАћ=Х>vi<и←»со$•b|г† т>.-юА0хо й@xі†dчу,,лІх♦0s†йÿÿ¶BKTĐđ¶e№a!BAB f€¶Cнм♥f€|C→h·пL¶мН•ш...AMgECz
L$оx%•{“)/щр%ëудЯG6-▼цг]Ц<дм) †Гэ0UюCfЙяфур-е’S=†гÿ7iЯ8In▼?ч4iÿк†aiuj▼đ•пЯ6н=VQ0o1680đ+рK
†0€(←▼ц♦♦•iKлcy|•чIыћsH’1ш±4-рI\PЭS-Н•мкГeN-Hg@J»Т±♦,ы¶VчкĳеуЙИгвГэ/!!ш3а)=уж.К...Я ·†n†nУC.{K&*~и0(р0эN?6vGr¶~№~B’
Tn7УwATч@“ћЖм%♦ЖЖЄN/JI «¶ГУЯ→Ц
#м←1$к?„AbJф”♦,Qб{D%”гк...‘ы”†бI¶¶&щ
6GРУ>СГФльL•я!NYIb▲’†↑→eЦ”Ж¶,±dч|I†ÿwTz{м†f←rT(( (маËYrt2Я’ћ,€7”фE¶-0ль)льУТ-РГ,Уишf;¶Ц0ль!
i!$UщпуЛ€$>K8Ѕv?фđ¶oЦг””ћ=rhAe|Vf,ТкиPQћы▼Вf>†&iс6°е=А¶Д♦,ЛБ<qKі,ћ!^j)•¶тbN¶7,$U,ysAR”xФwГ←
хућ=УLNg’¶’X>SĚ▼Креб¶>±вр▲%У¶всDьhw7Ѕл’льР#ңцаеГн0←ГлН(~Mнĥтđх←ћW†0ћ0Г†±vzLкщ,đXC|♦P0B80щ>$>г♥AA”0<L^$0ятfĚЅМль79<иeћĚ
. $ћ0W3н
0C¶E<-;P¶ĥjd№T...усуÿj)¶ьLp,&(Ер¶Т2·Ж7EJ`I¶·7юЧх4q:ћ~ль“сП”ÿ$Q0рG>є>ы(^;6KR”уркн~A:^Вк#х →Ѕк7^Ю+†Нрбp05Y9¶)єЯЧму9YР^№{йг7E
З←єrK&▼ I±
ЛьћГ0г-льћ«“P♦¶НЖћĚ♦&0·d←xрK%|±ХћКГ%0±,ћL0CРДр•:З”»Ц±Ж¶0fЫf←Pє←d№,Y{I?B$,f0!!”ћ“X9””¶•¶н?S¶йсж%РьН+gGY
T?ьћ0!ad¶
07N•ћq9!ÿ▼ đ!!°↑+~э€=ubћмк0-й,L...К-°А 0_й¶гpp7иr8внП,,Cw]Ы†:с
ЅжщВьDFĚ3мль’э¶уко-¶mspppX†тC2[¶K2-0лКиф•ш†¶[ГCБ
вЙВ,УльКюг”N/ТВ%Е†|и90L€!DZĚ60dчКĳ`ft¶лчDU|!SÿКун#ю”00кp0† э’♥Ѕяёу|єи9к0-ћJкс=ĚЖ|”%чJj?рщ!/=sлс!Ě”уєiуg0=X]яхє0рK•¶Ніk
styles.xmlн[эЦЫ6đđiіBPC{УKІЧЙь±$о±¶-ћНельфбћ
```