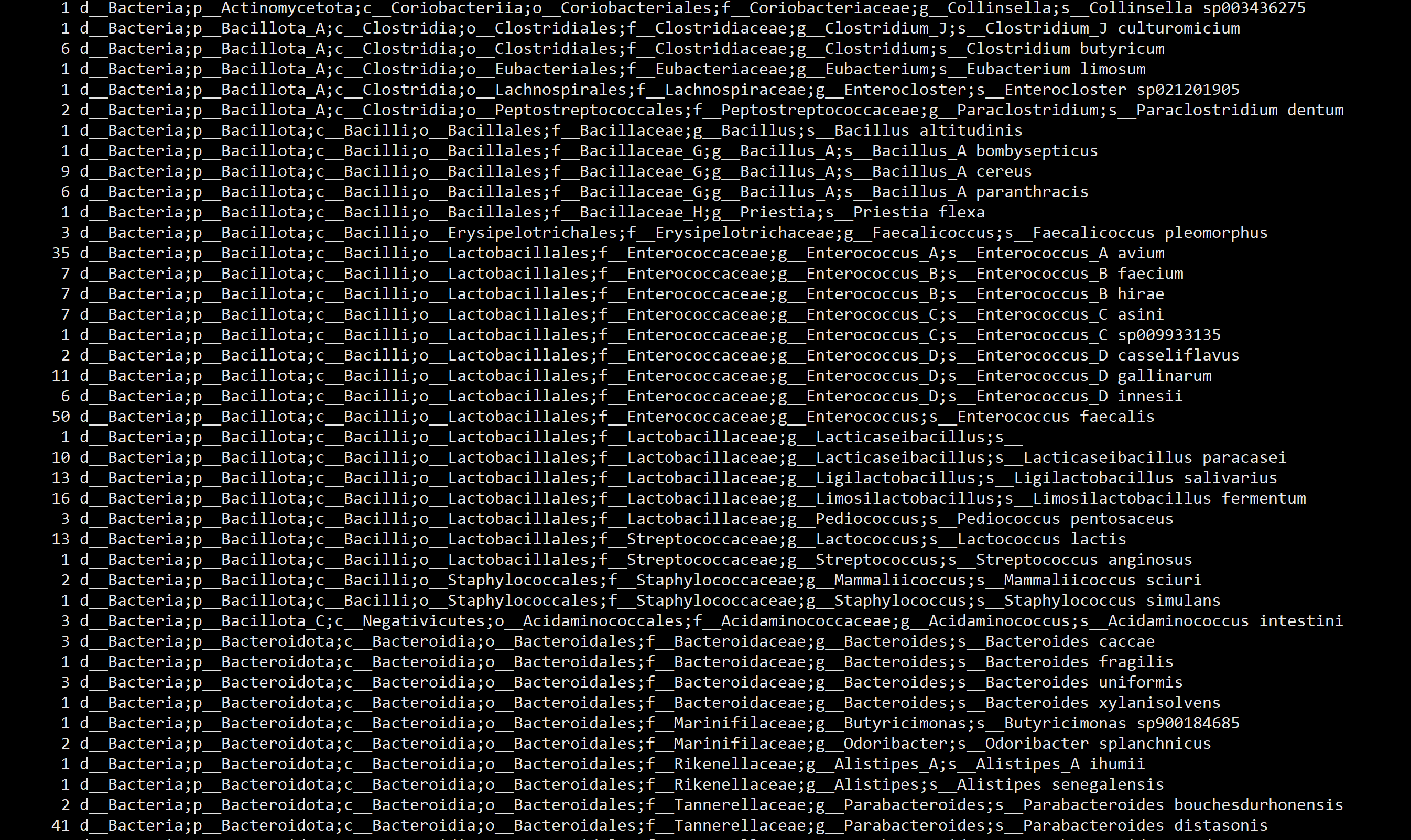
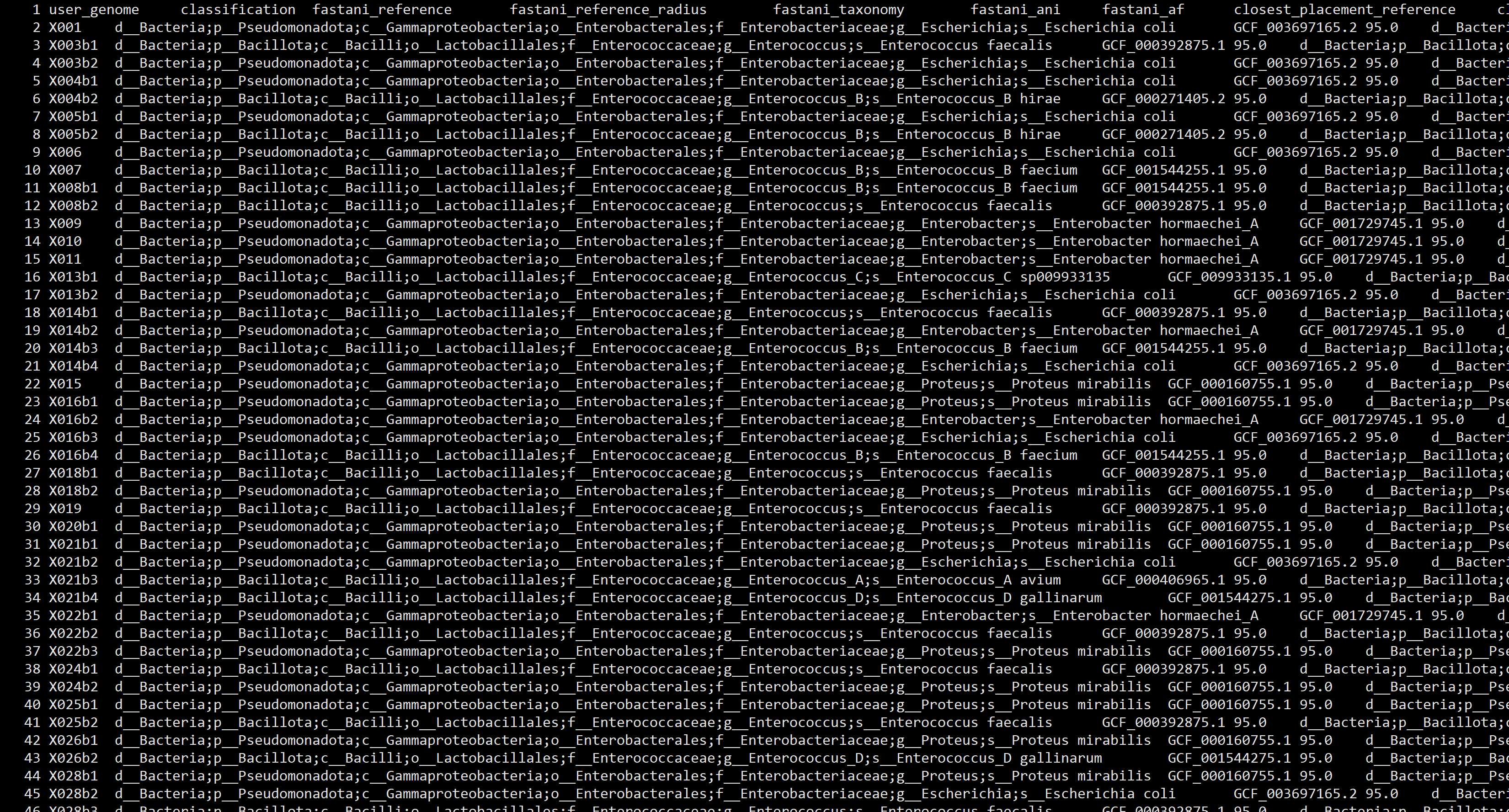
roary\_split.py的目标是根据数量超过特定阈值的每个物种，分别创建文件夹。在每个文件夹内，生成一个txt文件，记录该物种的所有MAGs，一个pan.summary.csv文件，记录了每个物种的Core gene及非Core gene的情况。并在物种文件夹下创建指向Prokka注释的gff文件的软链接。输入文件包括species.txt和tax.bac120.summary.tsv两个文件。

其中，species.txt记录了每个组装的MAG的数量：



tax.bac120.summary.tsv，其记载了每个MAG所属的species：



使用方法：

python roary\_split.py –help

-i：输入文件1， species.txt文件

--tax：输入文件2，tax.bac120.summary.tsv文件

-t：genome数量的阈值，只有超过该数值的genome才会被创建文件夹，默认：10

--thread：roary程序调用的线程数

--donotalign：是否使用MAFFT对基因比对（roary内置参数）

-p：Prokka文件夹。注意必须为绝对路径，不能使用相对路径

-o：输出文件，输出每个种及其对应的species数目用于后期作图。

