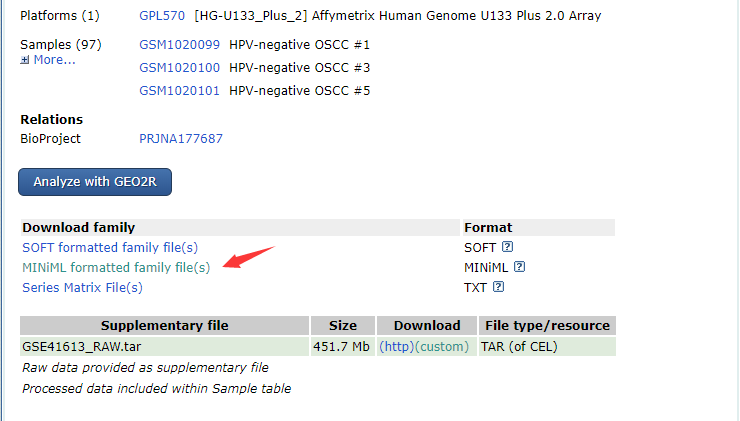
**①如果有自己比较熟练的下载和处理方式，可忽略本教程；**

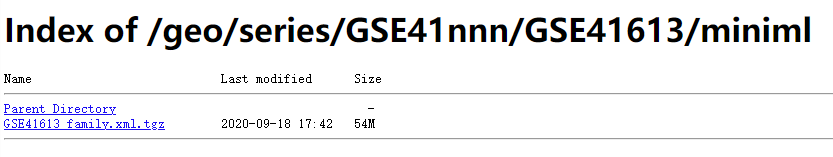
②本教程采用SangerBox对GEO基因芯片进行预处理，该工具的下载地址为：

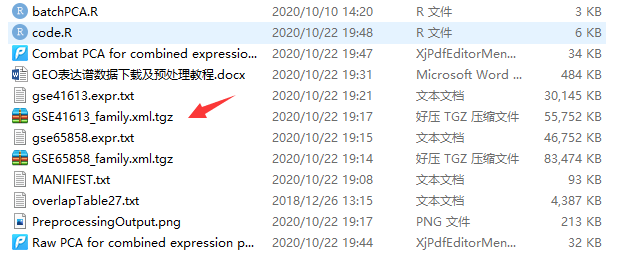
<http://soft.sangerbox.com/>

③下载以原文GSE41613为例，下载网页上的[MINiML formatted family file(s)](https://ftp.ncbi.nlm.nih.gov/geo/series/GSE41nnn/GSE41613/miniml/)，文件大小为54M





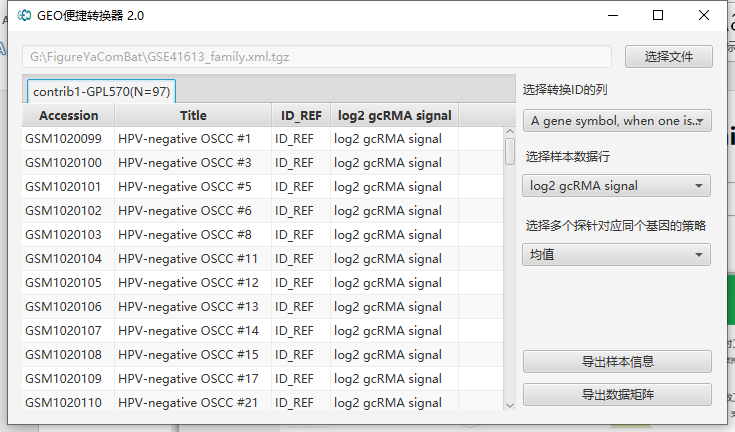




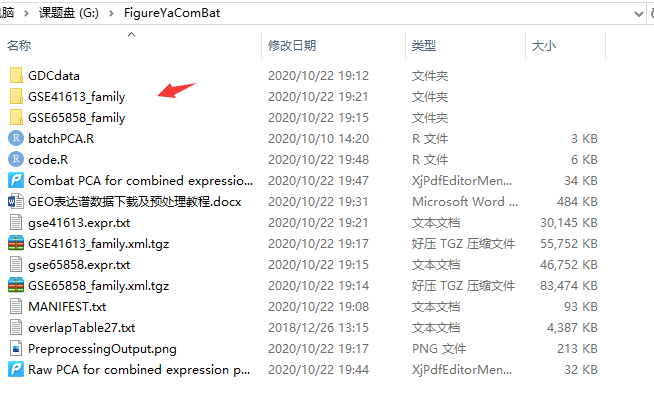
④打开SangerBox对应模块：

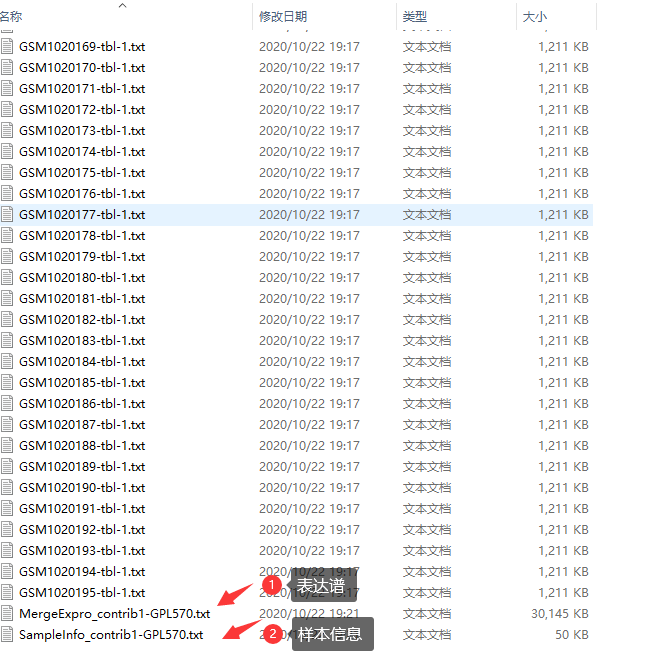


⑤选择文件路径，转换ID的列尽量选择包含“gene symbol”，数据行选择包含signal, RMA之类定量字样的，多探针策略一般为均值或中位数：



⑥导出样本信息和数据矩阵，会在该路径下生成一个文件夹，其中包含表达谱和样本信息：





**⑦这里我将表达谱重新命名为gse41613.expr.txt，移动到上一层目录中，并在R代码中做进一步处理**

