子宫内膜癌补充生信数据

1. ZNF454需用以下四个区域进行ROC分析，做四条ROC曲线，四个区域（UCSC数据库中Human Feb. 2009 (GRCh37/hg19) Assembly版本）如下：chr5:178367763-178368288，chr5:178367621-178368725，chr5:178367763-178368725，chr5:178367621-178368288。
2. ZNF454的以上四个区域还需分别和CDO1的chr5:115151911-115152561区域（UCSC数据库中Human Feb. 2009 (GRCh37/hg19) Assembly版本）进行组合，做四条ROC曲线。

Response：

所给DMR的genome 坐标版本（GRCh37）与TCGA探针的坐标版本（GRCh38）不一致，故先将所给坐标map到GRCh38：

发现ucsc在线工具<https://genome.ucsc.edu/cgi-bin/hgLiftOver> 可以非常方便地完成这个任务。

在新坐标系下，给定的ZNF454的4个区域分布对应的坐标为：

chr5:178940762-178941287

chr5:178940620-178941724

chr5:178940762-178941724

chr5:178940620-178941287

所给CDO1对应的区域为：

chr5:115816214-115816864

查找TCGA各个探针对应的坐标，找到这5个区间对应的探针：

DMR1:

cg02165355|cg03234732|cg10575261|cg23037403|cg16536329|cg20778451|cg17840719

DMR2:

cg17741986|cg02165355|cg03234732|cg10575261|cg23037403|cg16536329|cg20778451|cg17840719|cg03355526|cg10902717

DMR3:

cg02165355|cg03234732|cg10575261|cg23037403|cg16536329|cg20778451|cg17840719|cg03355526|cg10902717

DMR4:

cg17741986|cg02165355|cg03234732|cg10575261|cg23037403|cg16536329|cg20778451|cg17840719

DMR5:

cg07405021|cg16265906|cg12880658|cg16707405|cg02792792|cg14470895|cg23180938|cg08516516|cg11036833

分别对ZNF454的4个DMRs及其与DMR5的组合的ROC曲线。

TCGA上子宫内膜癌对应的数据包含428例癌症样本和46例癌旁样本。