

Table S1. Overview of datasets retrieved and preprocessed from TCGA and GEO

Dataset	Platform	Disease	Control
MMRF- COMMPASS	illumina	784	-
GSE39754	GPL5175	6	170
GSE4581	GPL570	414	-
GSE57317	GPL570	55	-
GSE136337	GPL27143	426	-