**SUPPLEMENTARY TABLE**

Supplementary Table 6. Detailed information for datasets, platform and processing steps.

| Name     | Platform             | SampleNumber | Country      | DiseaseType        | Normalization |
|----------|----------------------|--------------|--------------|--------------------|---------------|
| TCGA     | IlluminaHiSeq        | 987          | Mostly US    | LUSC+LUAD          | Log2Counts    |
| GSE30219 | AffyU133P2           | 167          | France       | LUSC+LUAD+LULN     | RMA           |
| GSE41271 | IlluminaHumanWG-6v3  | 264          | US           | LUSC+LUAD          | Quantile      |
| GSE42127 | IlluminaHumanWG-6 v3.0 | 176        | US           | LUSC+LUAD          | Quantile      |
| GSE50081 | AffyU133P2           | 172          | Canada       | LUSC+LUAD          | RMA           |