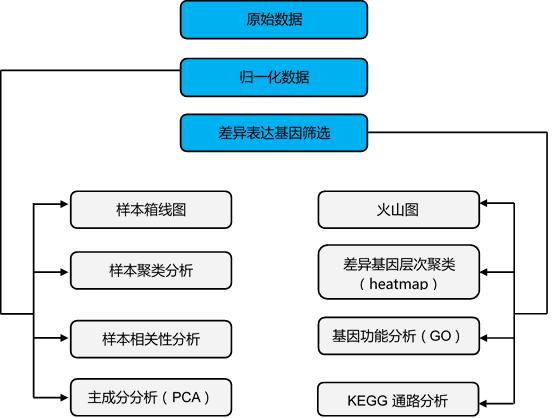
**数据分析**

1. **分析流程图**



1. **分析内容**

芯片扫描获得的原始数据由R软件中的limma包进行归一化处理后，利用箱线图观察各样本归一化数据的整体分布，并通过聚类图、样本相关性分析及PCA来评价样本的分组情况。采用Fold-Change（表达差异倍数）及统计学T检验（Student’s t-test）对差异表达基因进行筛选，利用散点图评估两组数据的整体差异及火山图展示显著性差异情况，并通过聚类热图直观地展示基因在不同样本中的表达情况。

对差异基因进行GO富集分析，达到对差异基因进行功能注释和分类的目的。同样对差异基因进行KEGG富集分析，有助于找到实验条件下显著差异变化的生物学调控通路。

对lncRNA靶基因进行预测（cis，trans）： 选取与lncRNA距离小于10kb的gene作为cis作用的靶基因；Trans预测采用物种对应的gene序列数据库，先通过blast选择出序列上具有互补性或相似性的序列，再利用RNAplex计算两序列之间的互补能量，选择e ≤−30的序列。

对circRNA吸附miRNA进行预测：circRNA分子富含microRNA结合位点，在细胞中起到miRNA海绵（ miRNA sponge）的作用，进而解除miRNA对其靶基因的抑制作用，升高靶基因的表达水平，这一作用机制被称为竞争性内源RNA（ceRNA）机制。通过对circRNA与miRNA可能结合的位点分析，预测circRNA吸附的miRNA分子，便于从ceRNA角度研究circRNA的分子功能。