Appendix 1 Supplementary method

Sample quantification and qualification

Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) was used to determine the concentration and evaluate the purity of RNA samples.

Agilent 2100 Bioanalyzer and 2100 RNA nano 6000 assay kit (Agilent Technologies) were used to evaluate the integrity of RNA samples.

Transcriptome sequencing library preparation

After the QC procedure, the RNA with poly-A in eukaryotic total RNA was enriched by TIANSeq mRNA Capture Kit (TIANGEN). Then, using the captured RNA as the starting sample, TIANSeq Fast RNA Library Kit (Illumina) was used to construct the transcriptome sequencing libraries. Briefly, the transcriptome sequencing library was constructed through RNA randomly fragmentation, cDNA strand 1 / strand 2 synthesis, end repair, A-tailing, ligation of sequencing adapters, size selection and library PCR enrichment.

Library quantification and qualification

Library concentration was first quantified using Qubit 2.0 fluorometer (Life Technologies), and then diluted to 1 ng/μL before checking insert size on an Agilent 2100 and quantifying to greater accuracy by quantitative PCR (Q-PCR) (library activity >2 nM).

Clustering and sequencing

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina sequencing platform and 150 bp paired-end reads were generated.

Data analysis

Quality control

Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter and trimming low quality base with Trimmomatic. At the same time, Q20, Q30 and GC content the clean data were calculated. All the downstream analyses were based on the clean data with high quality.

Reads mapping to the reference genome

Reference genome and gene model annotation files were downloaded from genome website directly. Index of the reference genome was built using Hisat2 v2.0.5 and paired-end clean reads were aligned to the reference genome using Hisat2 v2.0.5. We selected Hisat2 as the mapping tool for that Hisat2 can generate a database of splice junctions based on the gene model annotation file and thus a better mapping result than other non-splice mapping tools.

Quantification of gene expression level

Htseq-count was used to count the reads numbers mapped to each gene. And then FPKM of each gene was calculated based on the length of the gene and reads count mapped to this gene. FPKM, expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced, considers the effect of sequencing depth and gene length for the reads count at the same time, and is currently the most commonly used method for estimating gene expression levels.

Differential expression analysis

(For DESeq2 with biological replicates) Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq2 R package (1.16.1). DESeq2 provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value <0.05 found by DESeq2 were assigned a differentially expressed.

(For edgeR without biological replicates) Prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two conditions was performed using the edgeR R package (3.18.1). The P values were adjusted using the Benjamini & Hochberg method. Corrected P-value of 0.05 and absolute foldchange of 2 were set as the threshold for significantly differential expression.

© AME Publishing Company.
https://dx.doi.org/10.21037/tcr-2025-189

Table S1 MultiCox results

ID	HR	HR.95L	HR.95H	P
FCGR1A	1.006346726	1.002583954	1.01012362	0.000932406
ICOS	0.985345344	0.976779243	0.993986567	0.000920125
ICAM1	1.000338822	1.000052115	1.000625611	0.020542512
MMP9	0.999594336	0.999333778	0.999854962	0.002284908

© AME Publishing Company.
https://dx.doi.org/10.21037/tcr-2025-189