

Library preparation and sequencing

Library construction was performed using the Illumina TruSeq Stranded mRNA library preparation kit. Briefly, Total RNA samples were assessed for concentration using the Nanodrop 8000 (Thermo Scientific) and quality using the 2100 Bioanalyzer (Agilent). One microgram of total RNA per sample was used for poly-A mRNA selection using streptavidin-coated magnetic beads. cDNA was synthesized from enriched and fragmented RNA using reverse transcriptase (Super-Script II, Invitrogen) and random primers. The cDNA was further converted into double-stranded DNA, and the resulting dsDNA was enriched with PCR for library preparation. The PCR-amplified library was purified using Agencourt AMPure XP beads (Beckman Coulter). The concentration of the amplified library was measured with a NanoDrop spectrophotometer and an aliquot of the library is resolved on an Agilent 2100 Bioanalyzer. Sample libraries are multiplexed and sequenced on a NextSeq 500 platform (Illumina) using 75bp single-end sequencing. On average, about 20 million reads were generated from each sample.

Data analysis

Raw reads obtained from RNA-Seq were aligned to the transcriptome using STAR (version 2.5.0) (Dobin A et al., 2013) / RSEM (version 1.2.25) (Li B and Dewey CN, 2011) with default parameters, using a custom human GRCh38 (or mouse CRCm38) transcriptome reference downloaded from <http://www.gencodegenes.org>, containing all protein coding and long non-coding RNA genes based on human GENCODE version 23 (or Mouse GENCODE M8) annotation. Expression counts for each gene (TPM: transcripts per million) in all samples were normalized by the sequencing depth.

References

- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan 1;29(1):15-21.
- Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011 Aug 4;12:323.