**Sample names**

Fertile human sperm: AAhu-FFfer-TTsp-SS#, fer

Infertile human sperm: AAhu-FFinfer-TTsp-SS#, infer

SS#: replicate number

**Analysis workflow:**

The raw data was processed by Trimmomatic (v0.33) to keep the good quality reads. The clean reads were then aligned to UCSC hg19 genome by HiSAT2(v2.1.0). Transcripts were assembled by StringTie (v1.3.4d). The differential expression analysis was performed with R Ballgown (R package v2.10.0).

**Figure description**

**Fig.1 Transcript length distribution**

The X-axis is the transcript length for each gene (Unit: bp). The Y-axis is the gene density for each gene at the corresponding transcript length. Density plot shows the distribution of transcript length across biological replicates.

**Fig.2 FPKM distribution**

Box plot shows the gene FPKM distribution for across biological replicates as a quality control measurement for RNA-Seq data.

**Fig.3 Fold change distribution**

Probability density histogram showing the distribution of log2 fold change (infer/fer) for each gene. The dash line indicates fold change = 2 (log2 fold change = 1) as a cut-off.

**Fig.4 FPKM Scatterplot**

Scatterplot matrix showing the pairwise scatter plots of the log2 FPKM value

Between infer (Y-axis) and fer (X-axis).

**Fig.5 Transcript P value distribution**

Probability density histogram showing the distribution of P value for each transcript between fer and infer comparison. Dash line indicates P value = 0.05 as a cut-off.

**Fig.6 Gene P value distribution**

Probability density histogram showing the distribution of P value for each gene between fer and infer comparison. Dash line indicates P value = 0.05 as a cut-off.