Figure S2:

(A) The numbers of genes detected in bulk-OSCC-GB (n=40) and single-cell OSCC-GB. Number of genes detected are on the x-axis and sample ID on the y-axis. The scRNA-seq library was generated using Chromium Single Cell 3' Library Kit (10X Genomics). (The Single Cell 3' assay captures polyadenylated (polyA) transcripts, which include most eukaryotic messenger RNA (mRNA), and some long noncoding RNAs and antisense transcripts. Bulk RNA sequencing libraries were prepared from rRNA-depleted samples using TrueSeq RNA Sample Preparation Kit (Illumina) that captures the coding transcriptome. BioMart 1 was used to extract the list of protein-coding genes from Ensembl, and only these protein-coding genes were used for counting the total number of genes detected in single-cell data and compared with the total number of coding genes detected from bulk RNA-seq data.) (B) Scatter plots showing the correlation between the average gene expression of common genes in single-cell OSCC-GB (n=12) (x-axis) and bulk-OSCC-GB (n=40) (y-axis). FPKM and UMI values of bulk data and single-cell data respectively were converted to TPM values. Genes that overlapped between single-cell and bulk data (n= 14,716) were used to evaluate the correlation (using Pearson's correlation) at the average expression of these genes over single-cell data and bulk data.