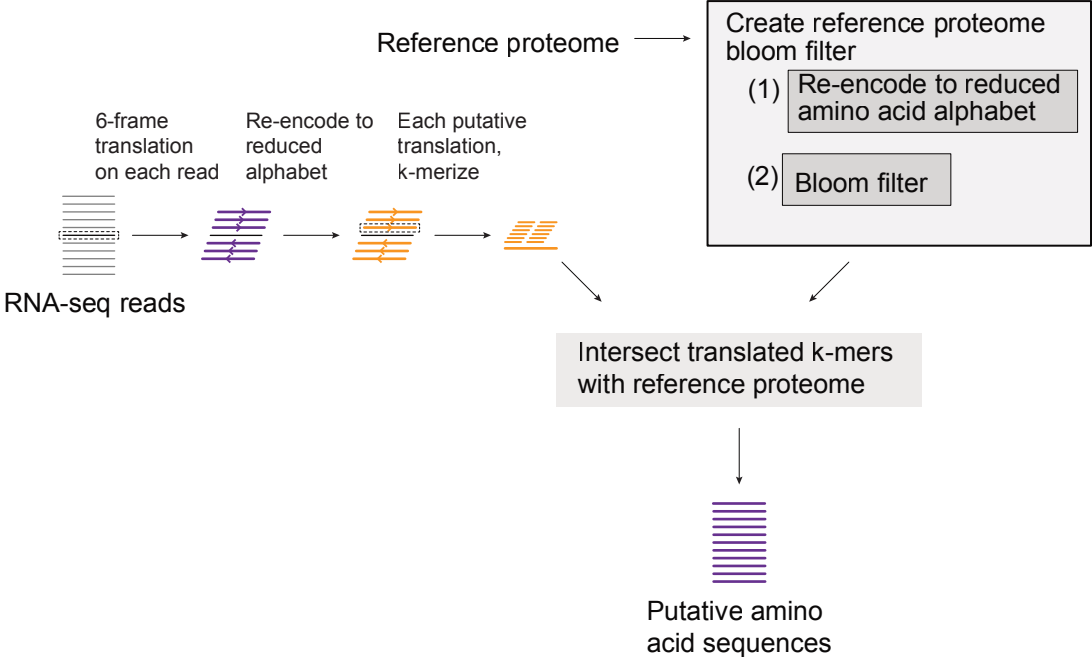


Supplemental Figure 3

- A**
1. Extract cell barcodes with sufficient nUMI from single-cell bam via `bam2fasta`
  2. Predict amino acid sequence from RNA-seq reads via `khtools extract_coding`



3. Randomly subsample amino acid k-mers via MinHash algorithm in `sourmash sketch`

- (b) `sourmash sketch` → MinHash protein k-mers
- (c) `sourmash compare` → Compute cell-cell similarity
- (d) `sourmash index` → Build sequence bloom tree (SBT)
- (e) `sourmash knn` → Build nearest neighbor graph
- (f) `sourmash umap` → Build UMAP off of knn

