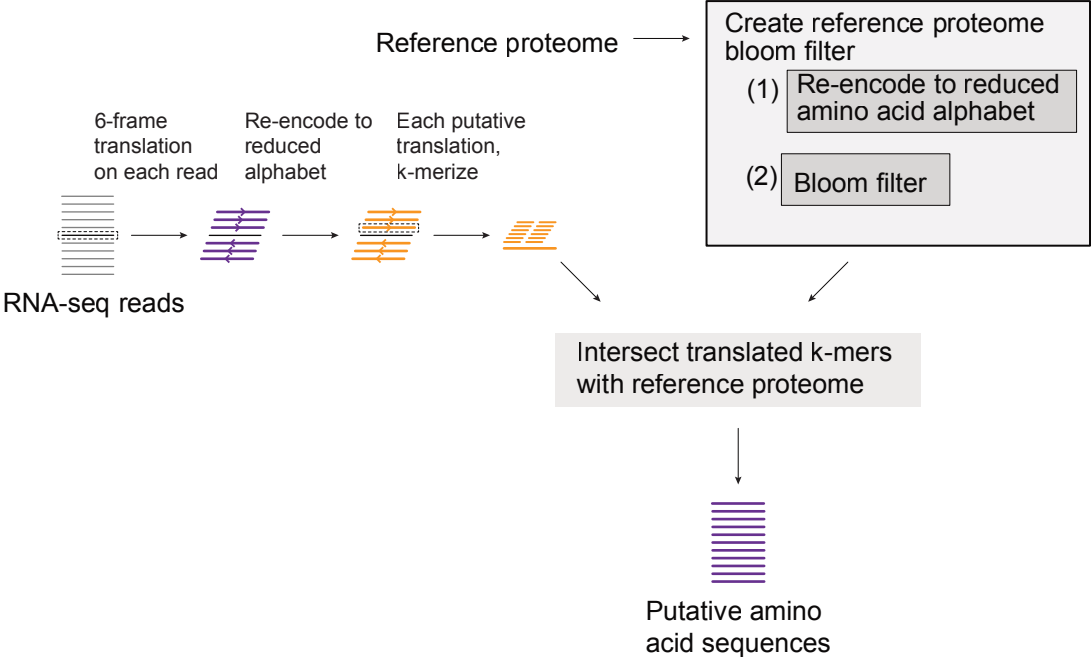


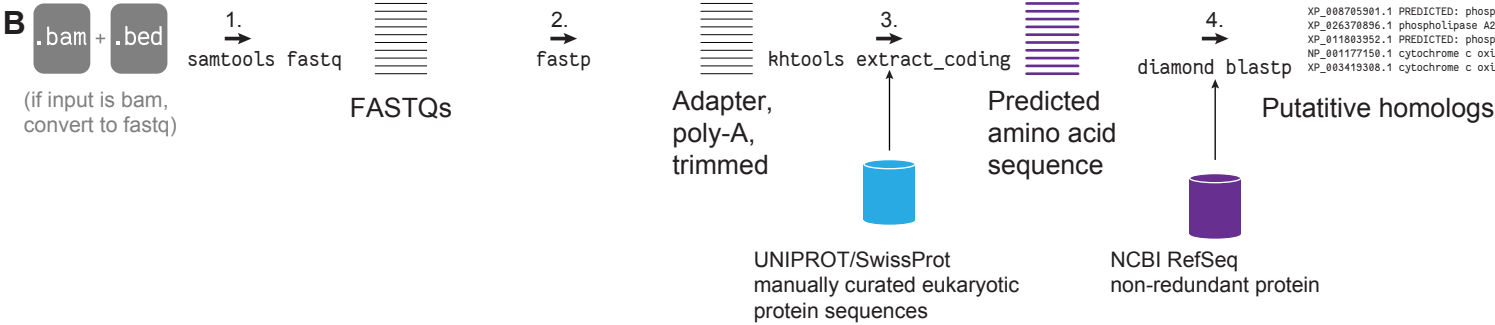
Supplemental Figure 3

- 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) | <code>sourmash sketch</code> | MinHash protein k-mers |
| (c) | <code>sourmash compare</code> | Compute cell-cell similarity |
| (d) | <code>sourmash index</code> | Build sequence bloom tree (SBT) |
| (e) | <code>sourmash knn</code> | Build nearest neighbor graph |
| (f) | <code>sourmash umap</code> | Build UMAP off of knn |



C