Supplementary File 1, related to Step 1 ("Identification of TCR sequences in single-cell TCR-seq data") in the "Before you begin" section.

Examples of the file formats used in this step.

The input file for Cell Ranger analysis should have the following format:

[Sample Name]_S1_L00[Lane Number]_[Read Type]_001.fastq.gz

Example of the csv file ("filtered_contig_annotations.csv") containing the information on TCR sequences. Columns are separated by commas. Reported below are the header line and the first three clonotypes of a representative file.

barcode,is_cell,contig_id,high_confidence,length,chain,v_gene,d_gene,j_gene,c_gene,full_length,productive,fwr1,fwr1_nt,cdr1,cdr1_nt,fwr2,fwr2_nt,cdr2,cdr2_nt,fwr3,fwr3_nt,cdr3,cdr3_nt,fwr4,fwr4_nt,reads,umis,raw_clonotype_id,raw_consensus_id,exact_subclonotype_id

AAACCTGAGCGTCTAT-1,true,AAACCTGAGCGTCTAT-1_contig_1,true,472,TRB,TRBV28,,TRBJ2-7,TRBC2,true,true,DVKVTQSSRYLVKRTGEKVFLECVQD,GATGTGAAAGTAACCCAGAGCTCGAGATATCTAGTCAAAAGGACGGGAGAGAAAGTTTTTCTGGAATGTGTCCAGGAT,MDHEN,ATGGACCATGAAAAT,MFWYRQDPGLGLRLIYF,ATGTTCTGGTATCGACAAGACCCAGGTCTGGGGCTACGGCTGATCTATTTC,SYDVKM,TCATATGATGTTAAAATG,KEKGDIPEGYSVSREKKERFSLILESASTNQTSMYL,AAAGAAAAAGGAGATATTCCTGAGGGGTACAGTGTCTCTAGAGAGAAGAAGGAGCGCTTCTCCCTGATTCTGGAGTCCGCCAGCACCAACCAGACATCTATGTACCTC,CASRALMGGAYEQYF,TGTGCCAGCAGAGCCCTGATGGGAGGGGCCTACGAGCAGTACTTC,GPGTRLTVT,GGGCCGGGCACCAGGCTCACGGTCACAG,935,24,clonotype47,clonotype47_consensus_1,1

AAACCTGAGCGTTCCG-1,true,AAACCTGAGCGTTCCG-1_contig_1,true,538,TRA,TRAV12-2,,TRAJ11,TRAC,true,true,QKEVEQNSGPLSVPEGAIASLNCTYS,CAGAAGGAGGTGGAGCAGAATTCTGGACCCCTCAGTGTTCCAGAGGGAGCCATTGCCTCTCTCAACTGCACTTACAGT,DRGSQS,GACCGAGGTTCCCAGTCC,FFWYRQYSGKSPELIMF,TTCTTCTGGTACAGACAATATTCTGGGAAAAGCCCTGAGTTGATAATGTTC,IYSNGD,ATATACTCCAATGGTGAC,KEDGRFTAQLNKASQYVSLLIRDSQPSDSATYL,AAAGAAGATGGAAGGTTTACAGCACAGCTCAATAAAGCCAGCCAGTATGTTTCTCTGCTCATCAGAGACTCCCAGCCCAGTGATTCAGCCACCTACCTC,CAVRGYSTLTF,TGTGCCGTGAGAGGATACAGCACCCTCACCTTT,GKGTMLLVSP,GGGAAGGGGACTATGCTTCTAGTCTCTCCAG,1000,16,clonotype7,clonotype7_consensus_2,1

AAACCTGAGCGTTCCG-1,true,AAACCTGAGCGTTCCG-1_contig_2,true,469,TRB,TRBV28,,TRBJ2-7,TRBC2,true,true,DVKVTQSSRYLVKRTGEKVFLECVQD,GATGTGAAAGTAACCCAGAGCTCGAGATATCTAGTCAAAAGGACGGGAGAGAAAGTTTTTCTGGAATGTGTCCAGGAT,MDHEN,ATGGACCATGAAAAT,MFWYRQDPGLGLRLIYF,ATGTTCTGGTATCGACAAGACCCAGGTCTGGGGCTACGGCTGATCTATTTC,SYDVKM,TCATATGATGTTAAAATG,KEKGDIPEGYSVSREKKERFSLILESASTNQTSMYL,AAAGAAAAAGGAGATATTCCTGAGGGGTACAGTGTCTCTAGAGAGAAGAAGGAGCGCTTCTCCCTGATTCTGGAGTCCGCCAGCACCAACCAGACATCTATGTACCTC,CASRGQTKTYEQYF,TGTGCCAGCAGGGGACAGACCAAGACCTACGAGCAGTACTTC,GPGTRLTVT,GGGCCGGGCACCAGGCTCACGGTCACAG,1516,16,clonotype7,clonotype7_consensus_1,1

Example of the text file ("barcodes.txt") containing the information on cell barcodes. This is a text file with only one column without header containing the selected cell barcodes. Reported below are the first 15 rows of a representative file.

AGGCCACAGTCCATAC-1
AGGCCACAGTCCATAC-1
AGGCCACAGTCCATAC-1
AGGCCGTCAGCTCGAC-1
AGGCCGTCAGCTCGAC-1
AGGCCGTCAGCTCGAC-1
CCCATACAGGAGCGTT-1
CCCATACAGGAGCGTT-1
CGAGCCACACCGTTGG-1
CGAGCCACACCGTTGG-1
CGAGCCACACCGTTGG-1
CGCTATCTCGCGGATC-1
CGCTATCTCGCGGATC-1
CGCTATCTCGCGGATC-1
CTCGTACAGAACAATC-1