Quality metrics were computed using the function `calculateQCMetrics` of the R package *scater*.

The following metrics were considered:

(1) detected genes: the number of genes for which at least one read was mapped (which is indicative of library complexity)
(2) the total number of counts
(3) the percentage of counts mapping to the top 50 genes
(4) mitochondrial percentage: the percentage of reads mapped to mitochondrial genes
(5) mitochondrial ratio = mitochondrial percentage/detected genes

cells with fewer than 200 detected genes were removed
cells with high ratio of counts mapping to mitochondrial genes were removed [classified as high using k-means clustering algorithm (k = 2): `kmeans(mitochondrial.ratio, 2)`].
nonprotein-coding genes were removed
mitochondrially encoded genes were excluded
genes detected in fewer than 2 cells were excluded

Extensive manual inspection/curation was done during subclustering analyses for each cell type: multiple rounds of subclustering were used to identify potentially spurious clusters representing low-quality or doublet cells. Subclusters distinctly high number of total counts and mixed expression of markers from different cell types were tagged as potential doublets and not considered for downstream analyses.