Benchmark

BioTechniques

Comparison of optimized methodologies for isolating nuclei from esophageal tissue

Lucy M Kimbley[1], Rachel Parker[1], Jack Harrington[1], Robert C Walker[1], Ben Grace[1], Jonathan J West[1,2,3], Tim J Underwood[1,3] & Matthew JJ Rose-Zerilli*[,1,3]

*[1]Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; [2]Centre for Hybrid Biodevices, University of Southampton, Southampton, UK; [3]Institute for Life Sciences, University of Southampton, Southampton, UK; *Author for correspondence: mjrz@soton.ac.uk*

BioTechniques 72: 00–00 (March 2022) 10.2144/btn-2021-0036

First draft submitted: 16 April 2021; Accepted for publication: 17 January 2022; Published online: 14 February 2022

ABSTRACT

Single-nuclei RNA sequencing allows single cell-based analysis in frozen tissue, ameliorating cell recovery biases associated with enzymatic dissociation methods. The authors present two optimized methods for isolating and sequencing nuclei from esophageal tissue using a commercial EZ and citric acid (CA)-based method. Despite high endogenous RNase activity, these protocols produced libraries of expected fragment length (average length EZ: 745 bp; CA: 1232 bp) with comparable complexity (median Transcript/Gene number, EZ: 496/254; CA: 483/256). CA nuclei showed a higher proportion of ribosomal gene reads, potentially reflecting co-isolation of nuclei and adherent ribosomes. The authors identified 11 cell lineages in the combined datasets, with differences in cell type recovery between the two methods, providing utility dependent on experimental needs.

METHOD SUMMARY

The authors present a method for isolating intact nuclei from frozen esophageal tissue for high throughput single-nuclei RNA sequencing using microfluidic-based droplet partitioning. This protocol overcomes high levels of endogenous RNase activity in this barrier epithelium and would therefore be applicable to other tissue types affected by similar issues.

KEYWORDS:

esophagus ● nuclei ● RNA ● sequencing ● single cell ● tissue sampling ● transcriptome

Single-nuclei RNA sequencing (snRNA-seq) allows the identification of cell types in heterogeneous tissues that have been frozen before downstream processing and analysis. Commercial buffers exist to isolate nuclei, as well as traditional salt and detergent-based methods [1]. These buffers must be gentle enough to lyse cell membranes while preserving nuclear membrane integrity. Nuclei can then be purified from cellular debris and contaminating cell-free RNA through multiple centrifugation steps, fluorescence-activated cell sorting (FACS) or density gradients (e.g., sucrose or iodixanol) before proceeding with microfluidic encapsulation and RNA barcoding [1,2].

As a barrier epithelial tissue, the esophagus contains RNases, which may play a role in mediating antiviral defense [3]. From the authors' single-cell RNA sequencing datasets, they have identified the cellular source of the expression of these RNases (Figure 1A–C). The authors identified high levels of endogenous RNases in outer keratinocyte populations, which can contribute to the loss of RNA integrity. This is in keeping with the role for RNase 7 in antiviral activity in keratinized cells [4,5]. Upon disruption of cell membranes during freezing and subsequent nuclei extraction, there is significant and rapid degradation of nuclear RNA (Figure 1D), which is likely to be associated with the release of reactivated RNases [6]. The authors have optimized two nuclei isolation methods to protect nuclear RNA integrity in frozen esophageal tissue (Figure 2) and demonstrate that both methods can recover a high diversity of cell populations following snRNA-seq.

To assess these methods, the authors obtained sequencing data from nuclei from a single patient, using one sample of snap-frozen normal human esophagus obtained by surgical resection (see ethical disclosure statement). Using the Nuclei EZ prep nuclei isolation kit (Sigma), according to published methods used for isolating nuclei for snRNA-seq by Dronc-seq [7], the authors found that nuclei from frozen esophagus lacked sufficient quality to generate cDNA libraries (Figure 1D). To overcome the effects of endogenous enzymes, they initially added an RNase inhibitor to the nuclei isolation buffer (0.2% v/v Lucigen NxGen RNase inhibitor) and transferred the frozen tissue to an RNA rescue medium (RNALater-ICE) without defrosting and then maintained it in the solution at -20°C until nuclei extraction. This enabled stabilization of the frozen tissue, providing protection during immediate removal from cold storage, allowing the inhibitors in the lysis buffer to take effect. This method successfully yielded cDNA libraries, and with additional optimization (0.2% v/v Lucigen NxGen RNase inhibitor, 0.2% v/v SUPERase-in, 1x Roche Complete protease inhibitor cocktail, 4 mM DTT) (Figure 1D & Figure 2A) the final optimized EZ protocol resulted in cDNA libraries of sufficient quality.

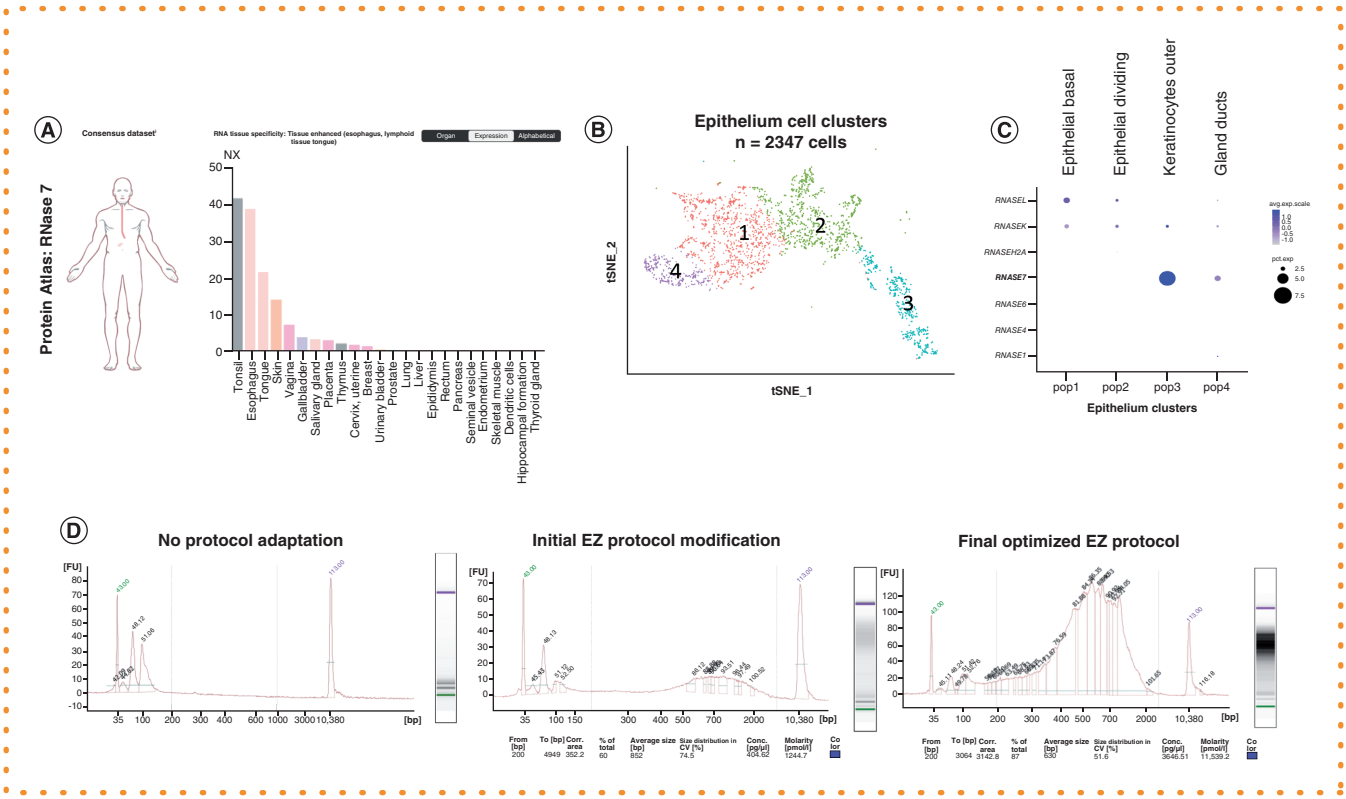


Figure 1. Identification of RNA inhibitors in esophageal tissue. (A) RNase 7 expression (NX: normalized expression) across tissue types (consensus dataset from the Protein Atlas; https://www.proteinatlas.org/). Image credit: Human Protein Atlas. **(B)** t-Distributed Stochastic Neighbor Embedding of epithelial cell populations from esophageal tissue analyzed by single-cell RNA sequencing. **(C)** RNase gene expression in epithelial cell populations from esophageal tissue analyzed by single-cell RNA sequencing. The size and color of the dot denote the percentage of cells and the average expression level, respectively. **(D)** Optimized nuclei isolation protocol for frozen esophageal tissue to mitigate reactivated endogenous RNase activity. cDNA library electrophoresis traces (high sensitivity DNA chip [Agilent Bioanalyzer] cDNA concentration quantified in fluorescence units) for pooled nuclei from a sample before, during and after successful protocol optimization.

Using a portion of the same sample stored in RNALater-ICE, the authors evaluated a citric acid (CA)-based isolation method, as described in Tosti *et al.* [8,9], where nuclei were isolated from pancreatic specimens, a tissue type with high RNase activity similar to esophageal tissue [10]. CA has historically been used for nuclei isolation and can improve the separation of nuclei from cytoplasmic contaminants [11] and yield greater quality RNA owing to its acidic pH and activity as an RNase chelator [12,13]. The tissue was removed from RNALater-ICE and homogenized in 25 mM citric acid and 0.25 M sucrose, following the protocol outlined in Figure 2B. This CA nuclei isolation was performed separately (+3 months) from the EZ isolation protocol. However, tissue stabilization using RNALater-ICE is a critical step in stabilizing nuclear RNA, and this step was not varied between methods. Again, the authors found that RNALater-ICE use yielded larger cDNA libraries with longer fragment lengths, suggesting a higher quality of recovered RNA (Figure 3A & B). Following sequencing, the resulting libraries from both isolation methods showed similar complexity (EZ median nTranscript/nGene: 496/254; CA: 483/256) (Figure 3C). The proportion of mitochondrial genes was higher than expected (EZ median: 25.1%; CA: 12.7%), given that these should be pure nuclei fractions. Previously published nuclei datasets have observed elevated mitochondrial gene expression, hypothesizing that mitochondria may associate with nuclear membranes [14]. The EZ method displayed approximately the anticipated proportion of intronic reads (median: 40.9%), given that nuclei contain pre-spliced transcripts. For the CA method, the proportion of intronic reads was lower than the EZ method (median: 28.6%), suggesting the presence of mature RNA. It has previously been found that CA-based methods may result in nuclei co-isolated with ribosomes or outer nuclear membrane fragments (Figure 3D) [15]. The CA nuclei showed a significantly higher percentage of reads originating from ribosomal genes (median: 7.6%) than the EZ method (median: 1.7%; $p < 0.001$ Mann–Whitney U test), potentially reflecting co-isolation with adherent ribosomes. This is further supported by stratification of the CA nuclei by percentage of ribosomal reads (Figure 3E & F), identifying two subpopulations of nuclei: those with a ribosomal gene proportion >8% showing a lower median intronic proportion (14%) and those with <8% ribosomal genes having a higher intronic proportion (37%). These two populations could reflect nuclei isolated with adherent ribosomes and pure nuclei, respectively.

For analysis, CA and EZ nuclei datasets from the same patient sample were combined (pre-combined analysis in Supplemental Figure 1). Following clustering with the Seurat package (v3.2.2) [16] using the first 15 principal components of variable gene expression, cell

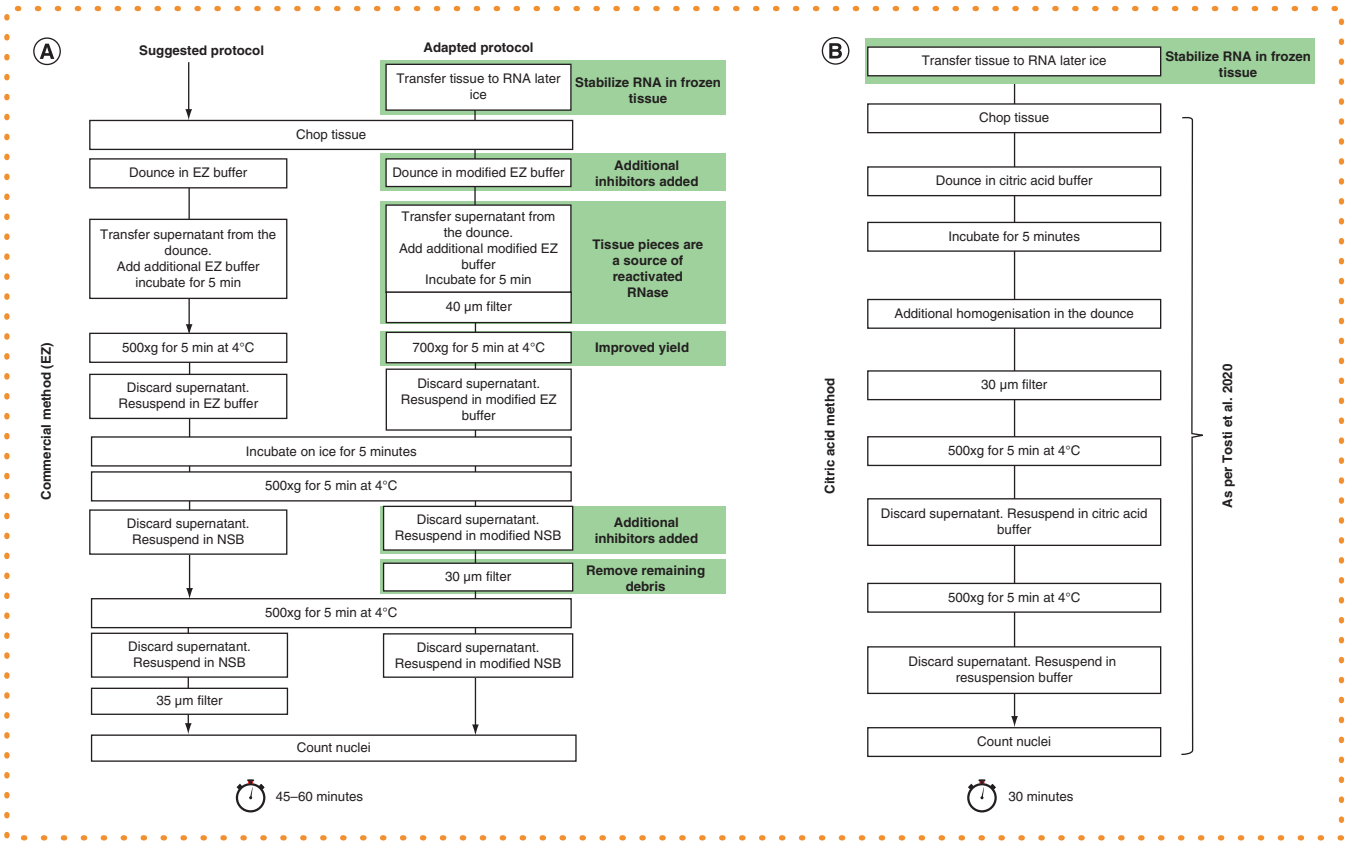


Figure 2. Optimization of nuclei isolation protocols. (A) Flowchart of adaptations to commercial protocol using modified EZ lysis buffer (EZ nuclei lysis buffer, 0.2% v/v Lucigen NxGen RNase inhibitor, 0.2% v/v SUPERase-in, 1x Complete protease inhibitor cocktail, 4 mM DTT) and modified nuclei suspension buffer (1xPBS, 0.01% BSA, 0.2% v/v RNase inhibitor, 1 mM DTT). **(B)** Citric acid protocol workflow. Citric acid buffer (0.25M sucrose, 25 mM citric acid), resuspension buffer (25 mM KCl, 3 mM $MgCl_2$, 50 mM Tris-buffer, 0.4 U/µl NxGen RNase inhibitor, 1 mM DTT, 0.4 U/µl SUPERasin).

types were identified using lineage marker genes from previous large-scale scRNA-seq esophageal datasets (Figure 4) [17]. Samples were sequenced separately, and batch effects were controlled for by regression of nCount and mitochondrial percentage. The authors were able to identify 11 cell clusters and one cluster of unassignable low-informative nuclei captures (Figure 4). Two clusters were uniquely observed in the CA dataset and marked by high ribosomal gene expression (stratified keratinocyte ribo-high and basal epithelial ribo-high clusters). The stratified keratinocyte and ribo-high populations clustered closely, with a highly correlated average gene expression profile between the populations (Pearson's correlation: 0.89). While there was co-clustering of the EZ- and CA-derived stratified keratinocytes, the CA-derived basal epithelial cells did not cluster with the primary basal epithelial cluster because of high expression of ribosomal genes (Figure 4C).

More keratinocytes were recovered by the CA method, particularly stratified keratinocytes (~14-fold increase in the percentage recovered) (Figure 4B), distinguished by their expression of *KRT4* (Figure 4C & D). Interestingly, the single-cell dataset shows that *RNase7* expression was highest in outer keratinocytes, suggesting that the CA method may result in improved keratinocyte membrane lysis and/or RNA integrity from nuclei compared with the EZ method. In contrast, the EZ method yielded a greater recovery of fibroblasts (sixfold) and immune cells (2.5-fold for combined immune populations), with populations expressing marker genes of natural killer (*PTPRC, KLRD1*) and macrophage (*CD163*) lineages (Figure 4C & D). Both methods identified two clusters of vascular smooth muscle cells (VSMCs) marked by both *ACTA2* and *MYH11* and discriminated by *A2M* expression, potentially suggesting an activated subset involved in an inflammatory VSMC response (Figure 4C & D) [18]. Nuclei doublet estimates by DoubletFinder [19] were generally less than 4.8% across most cell populations, bar the outer keratinocyte cluster, where doublet estimates were 77.8% (Figure 4E). This finding was corroborated by a lower genes/transcript ratio, clearly identifying them as nuclei doublet captures, and the absence of mixed lineage marker expression suggests these are homotypic doublets. Alternatively, this could be due to keratinocytes being more resistant to lysis, or greater levels of adherent cytoplasm. Both methods provided representation of stromal, immune and vascular cell types known to make up the architecture of the normal esophagus.

In the datasets, the authors have corroborated identification of the cell types they recovered using markers established in large-scale scRNA-seq studies of the esophagus. In their hands, the EZ method recovered a greater number of immune and fibroblast cells, while

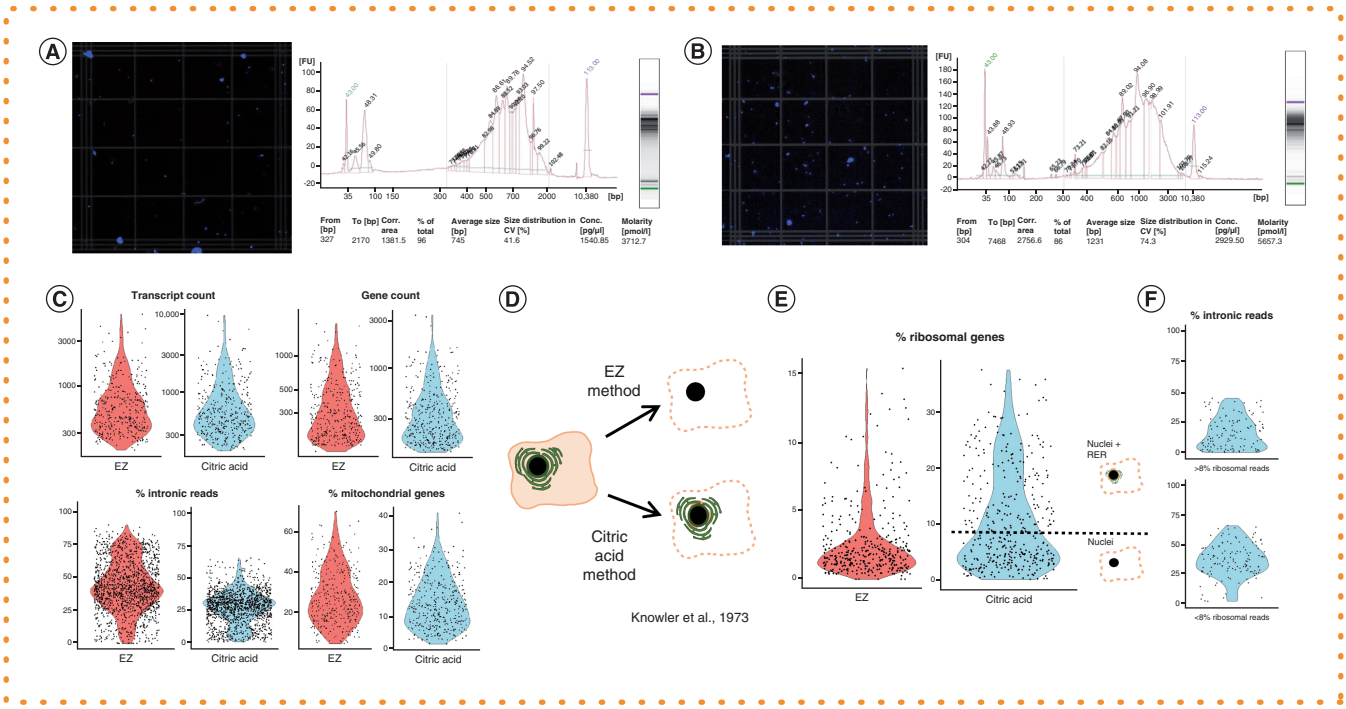
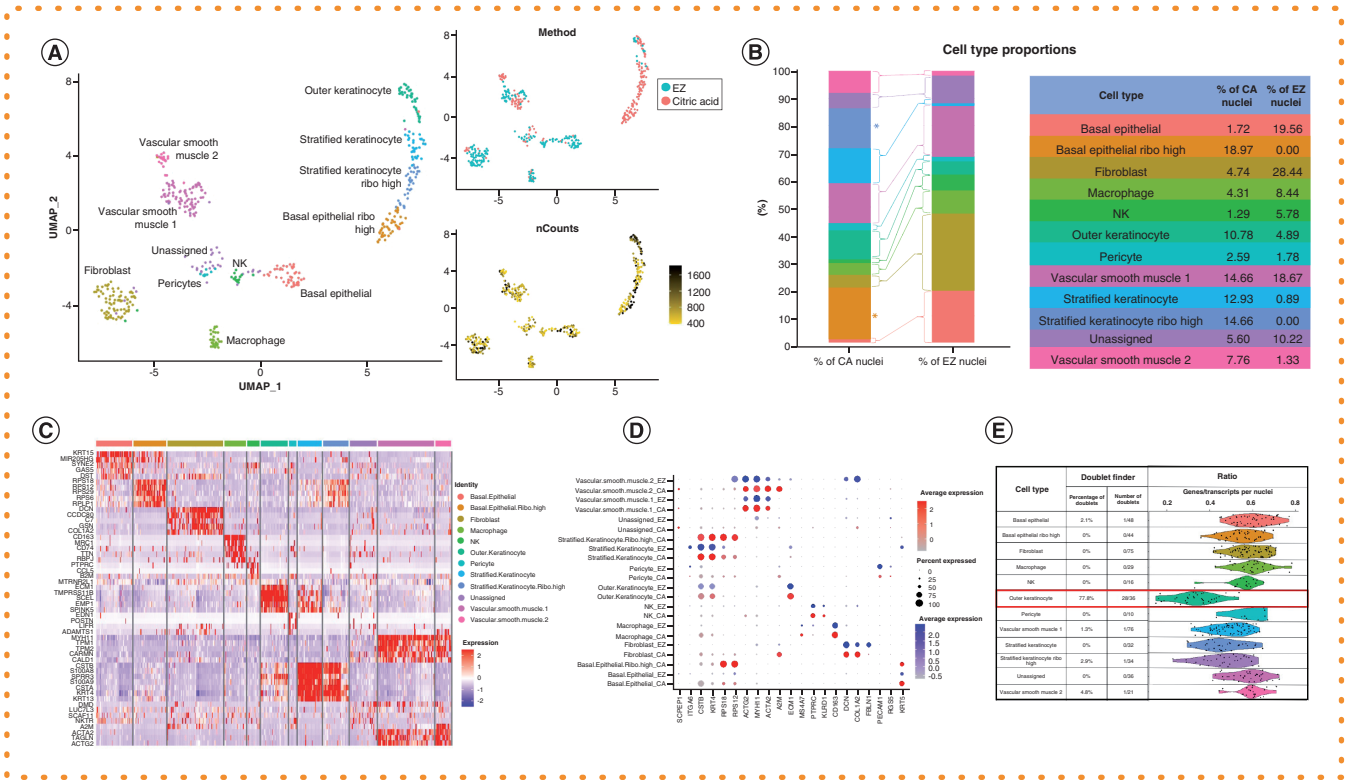


Figure 3. Comparison of EZ and citric acid (CA)-isolated nuclei preparations and library quality. (A & B) Representative images (4x magnification) of DAPI-stained nuclei and cDNA electrophoresis fragment traces for recovered nuclei pools for the **(A)** modified EZ method and **(B)** CA method, respectively. In both nuclei images, a 0.2 μl volume region is shown in both, but with different nuclei loading concentrations. **(C)** Percentage reads originating from ribosomal genes for single nuclei isolated by the EZ (median: 1.7%) or CA method (median: 7.6%). Dotted line at 8% representing a potential divide between nuclei isolated with adherent ribosomes and nuclei isolated alone. **(D)** Intronic read percentage for populations with > or <8% ribosomal reads. **(E)** Diagram demonstrating potential retention of endoplasmic reticulum and ribosomes when nuclei are isolated by the CA method compared with the EZ method. **(F)** Transcript count and feature count for single nuclei (points) isolated by the EZ (red violin) or CA method (blue violin). Log scale. EZ (median transcript count: 496; median gene count: 254). CA (median transcript count: 483; median gene count: 256). Percentage of intronic reads (median EZ: 40.9%; median CA: 28.6%) and mitochondrial genes (median EZ: 25.1%; median CA: 12.7%) for single nuclei isolated by the EZ or CA method.

the CA method also provided representation of keratinocyte populations, which will be useful for researchers studying these specific cell types. Further validation studies using additional patient samples to determine the extent of interpatient variability in cell type recoveries are required, but the results presented here allow direct comparison of cell type recoveries between methods using one patient sample.

The authors' observation of high mitochondrial genes expression has been observed previously in snRNA-seq datasets [14], supporting the assertion that mitochondria, and organelles including ribosomes, may become associated with nuclear membranes. However, cell viability and sample handling, preservation and differences in storage conditions or time may also be contributing factors, which the authors were unable to address in full. The localization of organelles may have additional benefits for the CA method. First, nuclei and attached ribosome sequencing has been shown to facilitate the study of rare cell types that are challenging to analyze by nuclei sequencing methods alone [20]. Second, the CA method may have utility for studies of RNA transcripts actively undergoing translation.

To facilitate analysis of primary tissues such as the esophagus, it is essential that dissociation methods are validated, so that cellular heterogeneity captured across studies can be corroborated. Validated methods that can be used with frozen tissue have advantages for the retrospective study of archived samples and for facilitating work in laboratories without access to fresh samples. The authors' protocols have immediate utility in studies of esophageal cancer, a disease with significant unmet clinical need that was previously challenging to study at the single-cell level. The authors have provided detailed laboratory protocols here: www.protocols.io (dx.doi.org/10.17504/protocols.io.btm6nk9e; dx.doi.org/10.17504/protocols.io.t9wer7e).



Cell type	% of CA nuclei	% of EZ nuclei
Basal epithelial	1.72	19.56
Basal epithelial ribo high	18.97	0.00
Fibroblast	4.74	28.44
Macrophage	4.31	8.44
NK	1.29	5.78
Outer keratinocyte	10.78	4.89
Pericyte	2.59	1.78
Vascular smooth muscle 1	14.66	18.67
Stratified keratinocyte	12.93	0.89
Stratified keratinocyte ribo high	14.66	0.00
Unassigned	5.60	10.22
Vascular smooth muscle 2	7.76	1.33

Figure 4. Cell type identification of nuclei. (A) Uniform Manifold Approximation and Projection (UMAP) clustering of nuclei isolated by both citric acid (CA) and EZ methods with cell types identified. UMAPs showing method of origin and transcript count for each nucleus. **(B)** Proportions of each cell type isolated by each method. *A unique population not represented in both methods. **(C)** Heatmap of top five genes expressed by each cell type. Cell type indicated by bar. **(D)** Dot plot showing expression of cell type marker genes in nuclei clusters, where color shows average expression and size of dot indicates percentage of cells expressing the marker gene. **(E)** Ratio of genes to transcript counts by cell type. Proportions of doublets in each cell type cluster are listed as determined by DoubletFinder assuming a doublet rate of 7%.

Author contributions

Conceptualization: MJJ Rose-Zerilli, TJ Underwood; methodology: R Parker, J Harrington, RC Walker, MJJ Rose-Zerilli, LM Kimbley; formal analysis: MJJ Rose-Zerilli, LM Kimbley; investigation: R Parker, J Harrington, RC Walker, B Grace, LM Kimbley; resources: JJ West; data curation: LM Kimbley; writing: original draft, LM Kimbley; writing: review and editing, MJJ Rose-Zerilli, TJ Underwood; visualization: LM Kimbley; supervision, MJJ Rose-Zerilli; project administration: MJJ Rose-Zerilli, TJ Underwood; funding acquisition: MJJ Rose-Zerilli, TJ Underwood.

Acknowledgments

The authors thank University Hospital Southampton patients for their participation in this study and the Southampton CR UK Tissue Bank. The authors would also like to thank the Wessex Investigation Sciences Hub Laboratory for access to their facilities and their technical support.

Financial & competing interests disclosure

TJ Underwood is supported by a Royal College of Surgeons of England and Cancer Research UK Advanced Clinician Scientist Fellowship (A23924). RC Walker is supported by a Cancer Research UK Clinical Research Training Fellowship (A25162). Development of Drop-seq technology in Southampton was funded by MRC grant (MC_PC_15078). This work was supported by a grant from the Chan Zuckerberg Initiative (Human Cell Atlas Pilot project: START: Standardization of Single-Cell and Single-Nucleus RNA-Seq Protocols for Tumors) to JJ West, TJ Underwood and MJJ Rose-Zerilli, and a grant from the Southampton CR UK Development Fund to JJ West, TJ Underwood and MJJ Rose-Zerilli. This investigation was supported by a collaboration with The General Hospital Corporation (Massachusetts General Hospital) under an award from The Chan Zuckerberg Initiative, an advised fund of Silicon Valley Community Foundation. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Human tissue samples were obtained from patients through University Hospitals Southampton NHS trust after informed consent. This study was approved by The Proportionate Review Sub-Committee of the North East – Newcastle & North Tyneside 1 Research Ethics Committee (REC no: 18/NE/0234).

Open access

References

1. Slyper M, Porter CBM, Ashenberg O *et al.* A single-cell and single-nucleus RNA-seq toolbox for fresh and frozen human tumors. *Nat. Med.* 26(5), 792–802 (2020).
2. Blobel G, Potter VR. Nuclei from rat liver: isolation method that combines purity with high yield. *Science* 154(3757), 1662–1665 (1966).
3. Ilinskaya ON, Mahmud RS. Ribonucleases as antiviral agents. *Mol. Biol.* 48(5), 615–623 (2014).
4. Harder J, Schröder J-M. RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin. *J. Biol. Chem.* 277(48), 46779–46784 (2002).
5. Rademacher F, Dreyer S, Kopfnagel V, Gläser R, Werfel T, Harder J. The antimicrobial and immunomodulatory function of RNase 7 in skin. *Front. Immunol.* 10, 2553 (2019).
6. Butler AE, Matveyenko AV, Kirakossian D, Park J, Gurlo T, Butler PC. Recovery of high-quality RNA from laser capture microdissected human and rodent pancreas. *J. Histotechnol.* 39(2), 59–65 (2016).
7. Habib N, Avraham-Davidi I, Basu A *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* 14(10), 955–958 (2017).
8. Tosti L, Conrad C. Nuclei isolation from snap frozen human pancreatic tissue using a citric acid buffer. *io.protocols* (2020).
9. Tosti L, Hang Y, Debnath O *et al.* Single-nucleus and *in situ* RNA-sequencing reveal cell topographies in the human pancreas. *Gastroenterology* 160(4), 1330–1344.e1311 (2021).
10. Farrell RE (Ed.). Chapter 7: Resilient ribonucleases. In: *RNA Methodologies (4th Edition).* Academic Press, CA, USA, 155–172 (2010).
11. Higashi K, Narayanan KS, Adams HR, Busch H. Utilization of the citric acid procedure and zonal ultracentrifugation for mass isolation of nuclear RNA from Walker 256 carcinosarcoma. *Cancer Res.* 26(7 Part 1), 1582 (1966).
12. Adamala K, Szostak JW. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* 342(6162), 1098–1100 (2013).
13. Nowotny M, Yang W. Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release. *EMBO J.* 25(9), 1924–1933 (2006).
14. Lake BB, Chen S, Hoshi M *et al.* A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* 10(1), 2832 (2019).
15. Knowler JT, Moses HL, Spelsberg TC. Comparison and characterization of nuclear isolation procedures as applied to chick oviduct. *J. Cell Biol.* 59(3), 685–695 (1973).
16. Stuart T, Butler A, Hoffman P *et al.* Comprehensive integration of single-cell data. *Cell* 177(7), 1888–1902.e1821 (2019).
17. Madissoon E, Wilbrey-Clark A, Miragaia RJ *et al.* scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* 21(1), 1 (2019).
18. Rodríguez-Calvo R, Ferrán B, Alonso J *et al.* NR4A receptors up-regulate the antiproteinase alpha-2 macroglobulin (A2M) and modulate MMP-2 and MMP-9 in vascular smooth muscle cells. *Thromb. Haemost.* 113(6), 1323–1334 (2015).
19. Mcginnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst.* 8(4), 329–337.e324 (2019).
20. Drokhlyansky E, Smillie CS, Van Wittenberghe N *et al.* The human and mouse enteric nervous system at single-cell resolution. *Cell* 182(6), 1606–1622.e1623 (2020).