Longitudinal exploration of the epithelium in a 4NQO-induced esophageal cancer murine model with single cell RNA sequencing reveals distinct cellular populations that expand with cancer

Adam Karami, Mohammad Faujul Kabir, Alena Klochkova, Anbin Mu, Mary Grace Murray, Brittany Yancey, Kelly A. Whelan

Fels Institute for Cancer Research and Molecular Biology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA

Contact: adam.karami@temple.edu

ABSTRACT

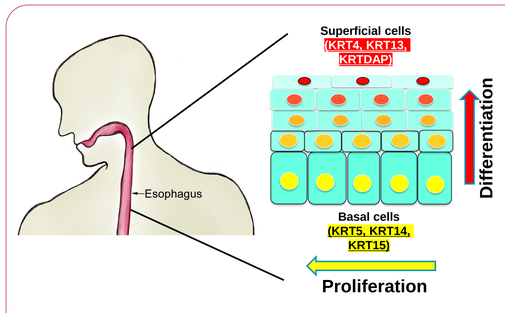
Introduction: At the tissue level, epithelial alterations in the pathology of esophageal cancerhas been characterized. However, the changes at the cellular level is - as of yet - poorly understood. To study the extent of epithelial cellular changes that happen in esophageal cancer, we administered 4NQO in mice as a carcinogen and performed single cell RNA sequencing on their epithelial tissue at different time points to observe the molecular alterations at the cellular level.

Methods: 4NQO was administered in the drinking water of the test group mice. Epithelial tissues were collected from mice before 4NQO administration (normal) (n = 12), as well as 4 months and 16 weeks (Precancer) (n = 6) and 4 months and 21 weeks (Cancer) (n = 4) after 4NQO administration. RNA was extracted, and libraries were prepared with InDrop. We used UMI-tools to identify likely cell barcodes and UMIs from the raw reads, and the UMI-extracted reads were aligned by STAR. The alignments were assigned to genes with FeatureCounts. UMI-tools was again used to retrieve the counts for each gene in each cell. Analysis of the count matrices was done with the R package Seurat. Cells mitochondrial expression percentage over 20% were filtered out and the datasets were batch corrected at the level of treatment weeks. We used UMAP to find principal components and reduce dimensionality. Monocle was used to calculate the reversed graph embedding for the pseudotime projection. Finally, pathway analyses using differentially expressed genes were done with Ingenuity Pathway Analysis (IPA) from QIAGEN.

Results: Unsupervised classification defined 8 clusters as epithelial cells. Visualization with UMAP clearly separates cells treated with 4NQO and cells without the treatment. Among the 8 epithelial clusters, 5 were enriched for basal cell markers, and 2 were enriched for superficial cell markers. The remaining suprabasal epithelial population co-expressed basal and superficial markers. Among the basal clusters, Basal 3 and 4 are continuously enriched as 4NQO treatment goes on. Basal 5 is most enriched at precancer, but is not found in untreated cells. Superficial 1 is depleted with 4NQO, while Superficial 2 is enriched. Pathway analyses with IPA revealed the clusters' differential regulation of pathways within its respective cell type, revealing further the clusters' heterogeneity. Pseudotime analysis of the clusters reveals Basal-2 as the bridge between wild type and cancer cells. Pathway analyses on the genes involved in the branchpoints reveal biological similarities and differences between wild type epithelial differentiation and its tumor counterpart, as well as how they compare to cancer proliferation.

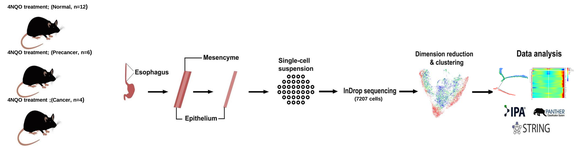
Conclusions: The analyses performed in this study reveal the different cell states of the epithelium under 4NQO, with clusters of basal (3 and 4) and superficial cells (2) that continuously expand throughout carcinogenesis. Pseudotime analyses show the genes and pathways involved in the trajectories of epithelial cells as cancer progresses.

BACKGROUND



- Esophageal epithelium is comprised of a gradient of cells from proliferating basal cells to differentiated superficial cells.
- It is unknown how esophageal cancer affects these cellular gradients.
- 4NQO has been used as a reliable carcinogen to induce esophageal cancer, but its perturbation at the single cell level has yet to be studied.
- **Using a murine 4NQO-treated model, we interrogated the changes in the epithelium under esophageal cancer at the single cell level.**

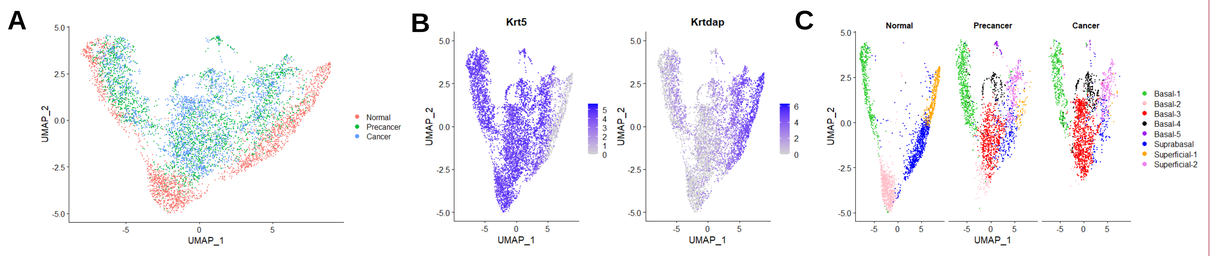
METHODS



A schematic of the experiment methods. 4NQO was administered in the drinking water of the experimental group. In total, 7207 cells passed quality control cutoffs and were used for downstream dimensionality reduction and further analyses.

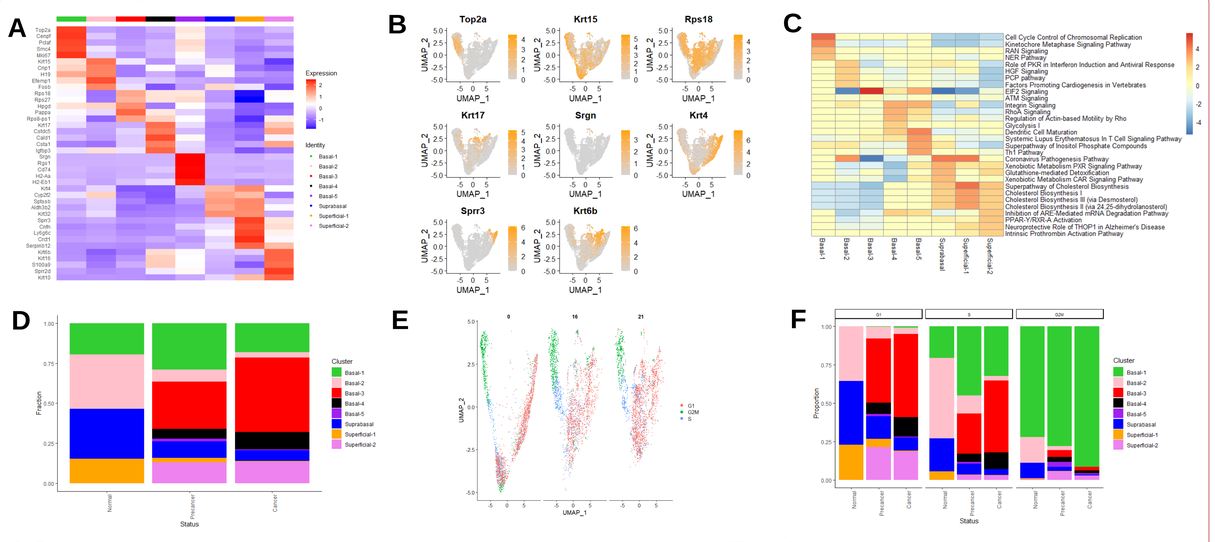
RESULTS

Figure 1. Identification of cell clusters of the esophageal epithelium across timepoints of 4NQO treatment duration (0, 16, and 21 weeks).



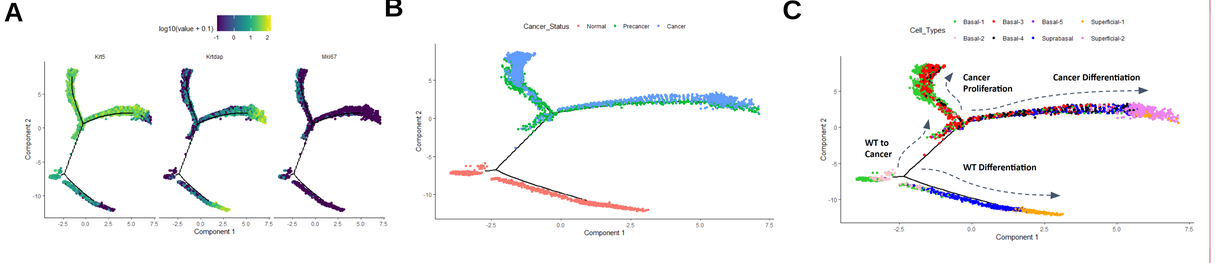
(A) UMAP display of the grouping of cells at different timepoints. Unsupervised clustering clearly separates between cells from mice that have been treated with 4NQO and cells from mice without the treatment. **(B)** Log1p expression of the basal marker Krt5 and superficial marker Krtdap across the epithelial dataset **(C)** UMAP display of the identified distinct cell clusters, separated by the different 4NQO timepoints.

Figure 2. Molecular characterization of the epithelial clusters identified in the mice with and without 4NQO



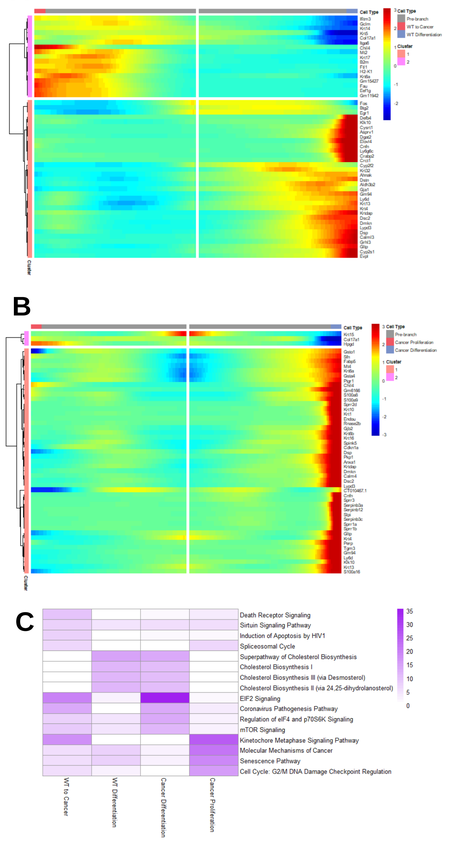
(A) Z-score scaled expression of the markers most upregulated in each of the cell clusters. **(B)** Log1p expression of the most upregulated distinguishing markers for each cell cluster. **(C)** Ingenuity Pathway Analysis (IPA) Z-scores of the canonical pathways most upregulated in each cell cluster. **(D)** Change in cluster compositions of cells in each 4NQO timepoint. **(E)** UMAP of epithelial cells in different 4NQO timepoints labeled by their cell cycle phases. **(F)** Change in cluster compositions of the different cell cycle phases in each 4NQO timepoint.

Figure 3. Pseudotime projection of epithelial cells with and without 4NQO



(A) Expression of the basal, superficial, and proliferation markers Krt5, Krtdap, and Mki67, respectively, in pseudotime **(B)** Pseudotime trajectory inference clearly separates cells before and after 4NQO treatment. **(C)** Cells in the pseudotime trajectories are labeled by the cell cluster that they each belong to, showing branch points with biological decision paths.

Figure 4. Genes and pathways associated with epithelial cellular trajectories under 4NQO



(A) Genes associated in the branchpoint between wild type differentiation and transition to cancer. **(B)** Genes associated in the branchpoint between cancer proliferation and cancer differentiation. **(C)** Top pathways involved in each trajectory in the epithelium under 4NQO.

CONCLUSIONS & FUTURE DIRECTIONS

- 4NQO-mediated carcinogenesis induces the development of basal and superficial cell clusters nonexistent in the wild type epithelium, and their markers and biological pathways have been identified
- Pseudotemporal analysis has identified differences between regulators of differentiation in wild type and cancer
- Future experiments will seek to validate the existence and localization of the identified 4NQO-induced cell clusters

ACKNOWLEDGEMENTS

We thank Jaroslav Jelinek and Hugh Huang from the Fox Chase Cancer Center for their help in scRNA sequencing