Diabetic Complications Consortium

Application Title: In-situ characterization of cell injury in Diabetes using tissue cytometry and machine learning

Principal Investigators: Tarek M. Ashkar (El-Achkar) and Mohammad Al Hasan

1. Project Accomplishments:

We have achieved very significant progress during the funding period of the DIACOMP pilot award and developed methodologies to identify subvisual injury in diabetes based on nuclear staining as discussed below. We are pleased to report that we have already published or deposited in pre-print three manuscripts and currently preparing a large manuscript directly related to the specific aims proposed in this grant. Because of this pilot award, we were able to generate preliminary data and submit an R21(PAR-20-140) proposal for the July 2022 review cycle in PBKD. Work supported by this award was also presented as an oral presentation in Kidney Week 2021, and an abstract was submitted for Kidney Week 2022. The outputs supported by this award harmonize with ongoing efforts by our group such as our work in the Kidney Precision Medicine Project consortium.

2. Specific Aims:

Specific Aim 1. Define the landscape of diabetes-induced cell injury in situ using an imaging-based approach that combines tissue cytometry and machine learning.

Results: We have developed a framework to identify injured cells using machine learning classification by linking transcriptomics to image-based features. We first demonstrate this approach

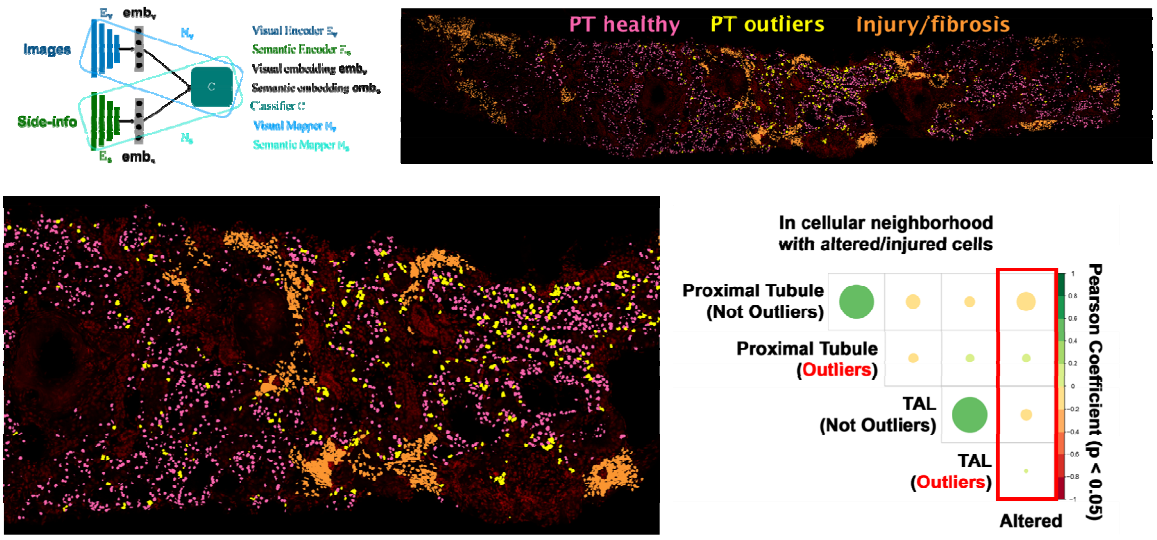

Figure 1: Detecting injured cells in image volumes using only DAPI staining based on machine learning (outlier detection algorithm) combined with transcriptomics as side information. Top left schematic shows the design of this approach using encoders for visual and semantic embedding and a classifier. The image on the top right is from a diabetic biopsy where the "outlier" injured proximal tubules classified based on their nuclear staining are mapped, and correlate with areas of injury and fibrosis. Image on bottom left is an enlarged area, and statistical correlations from 4 biopsies is shown on the bottom right.

using an outlier detection algorithm for visual embedding and classification (Figure 1), where transcriptomics data (healthy vs. altered cells) are used as side information for semantic embedding. The learning occurs by minimizing the error function between the visual and semantic embedding classes. Our findings show that we can detect injured PT and TAL cells in diabetes, and this was demonstrated by the statistically significant association of the injured ("outlier") TAL and PT with the altered areas (altered morphology and fibrosis) in the biopsy. This finding will be the basis of a manuscript in preparation, which will be drafted in collaboration with the Kidney Precision Medicine Project (KPMP) consortium because of the use of high-quality 3D imaging data of consortia tissue our group has been collecting. A concept proposal has been already submitted to the KPMP publication and presentation committee, and an abstract on these findings has been submitted to Kidney week 2022. We are currently extending this approach to be bias free, by using zero-shot learning rather than outlier detection and incorporating the transcriptomic side information as a continuous rather than in categories.

> **Specific Aim 2**. Establish that tubular injury in diabetes differentially affects thick ascending limb (TAL) cells.

Results: We have used our deep learning NephNet3D trained to classify TAL and proximal tubules in diabetes as healthy or altered (Figure 2). We found altered TAL more commonly than PT (although this needs to be validated in a larger number). We subsequently used an unsupervised approach to extract imaging features of altered TALs and define 2 subtypes that statistically associate with area of injury in the biopsy based on neighborhood analysis. In these biopsies, altered PTs did not have a positive association with areas of injury. These findings support our hypothesis, and were presented in Kidney week 2021 as an oral presentation, and will be part of a manuscript in preparation. We plan

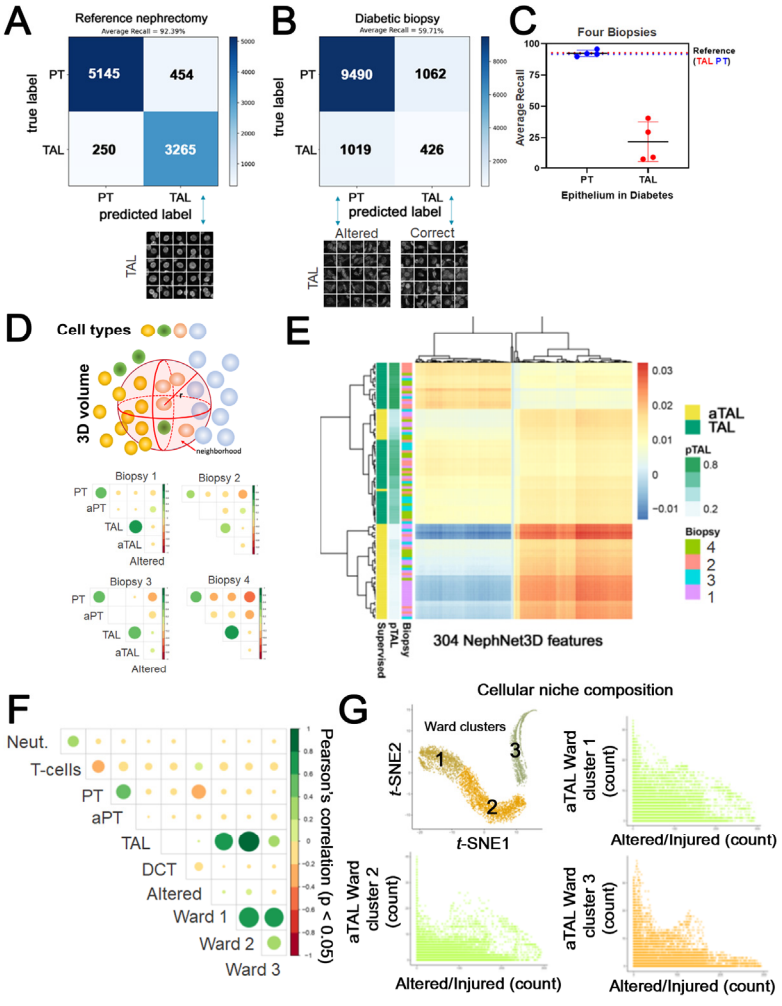

Figure 2. Supervised and unsupervised analysis of image features uncovers signature of diabetes in tubular epithelial nuclei. 3D nuclei of tubular epithelial cells were selected from reference human nephrectomies or renal biopsies of patients with diabetes with Volumetric Tissue Exploration and Analysis and classified with NephNet3D. A-C. Nuclei from the thick ascending limb of biopsies in four patients with diabetes were poorly classified versus reference tubules suggesting altered TAL (aTAL) and proximal tubule (aPT) epithelial cells based on poor average recall by NephNet3D (B, e.g. pTAL in E). D. The association of putative aPT and aTAL with areas of injury (Altered) were assessed in 3D niches. E. The 304 features calculated by NephNet3D, prior to supervised classification in the fully-connected layers, were used to cluster the nuclei. Hierarchical clustering (by rows) suggests multiple aTAL subtypes. F-G. Hierarchical clustering with Ward-linkage on putative aTAL from all four biopsies identifies three clusters (Ward 1-3), two of which positively correlate with areas of injury (Altered, green in G) and not neutrophils or T-cells in 3D niches (F).

to expand these studies to a larger sample size from the KPMP imaging dataset, and with additional extramural funding.

3. Publications:

1. El-Achkar TM, Winfree S, Talukder N, Barwinska D, Ferkowicz MJ, Al Hasan M. Tissue Cytometry With Machine Learning in Kidney: From Small Specimens to Big Data. Front Physiol. 2022;13:832457. doi: 10.3389/fphys.2022.832457. eCollection 2022. Review. PubMed PMID: 35309077; PubMed Central PMCID: PMC8931540.

2. Sabo AR, Winfree S, Bledsoe SB, Phillips CL, Lingeman JE, Eadon MT, Williams JC Jr, El-Achkar TM. Label-free imaging of non-deparaffinized sections of the human kidney to determine tissue quality and signatures of disease. Physiol Rep. 2022 Feb;10(3):e15167. doi: 10.14814/phy2.15167. PubMed PMID: 35133089; PubMed Central PMCID: PMC8822874.

3. Winfree S, McNutt AT, Khochare S, Borgard TJ, Barwinska D, Sabo AR, Ferkowicz MJ, Williams JC, Lingeman JE, Gulbronson CJ, et al. Integrated cytometry with machine learning applied to high-content imaging of human kidney tissue for in-situ cell classification and neighborhood analysis. *bioRxiv*. 2022:2021.2012.2027.474025. doi: 10.1101/2021.12.27.474025