A Novel AI-driven Spatial Genomics Platform Demonstrates Heterogeneity in NASH Liver

John Butler[1], Katie Konigsfeld[1], Krishna Murugesan[2], Casey Laris[2], Mary Nguyen[2], Marisa Sanchez[2], Bidhan Chaudhuri[1], Quin Wills[3], Jack O'Meara[3], Claire Weston[2]*

*Corresponding author (cweston@revealbio.com), [1]QuantumCyte, Sunnyvale, CA, [2]Reveal Biosciences, San Diego, CA, [3]Ochre Bio, Oxford, UK

REVEAL BIOSCIENCES
Data Powered Pathology

ABSTRACT

BACKGROUND: Non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) are common chronic liver disorders in the US. Spatial genomics and digital pathology continue to advance at a rapid pace, generating large amounts of data. There is a critical unmet need for advanced tools to expedite NASH research with respect to spatial data generation, visualization, and analysis, to facilitate an improved understanding of this disease and advance NASH therapeutics.

METHODS: In this study we present a novel method for the extraction and analysis of genomic information from specific regions of interest (ROI) within a series of formalin-fixed paraffin embedded (FFPE) NASH tissue sections. An artificial intelligence-based NASH algorithm first identified specific regions within the tissue, and segmentation results were confirmed by a Board Certified Pathologist. Automated lysate extraction was then performed in the selected ROIs using ink-jet technology and novel ink chemistry. The process was performed directly on the tissue slide and the crude lysates generated were transferred to a standard collection tube for subsequent sequencing. DNA was extracted from the selected areas before being purified and analyzed using a targeted multiplex amplicon NGS panel. The imageDx™: NASH AI-based digital assay was used to quantify micro- and macro-vesicular steatosis, lobular inflammation, hepatocellular ballooning, and fibrosis in each ROI. The CytoMask™ ink-jet based workflow was used to mask off selected ROIs and extract nucleic acid from those specific ROIs identified from the imageDx™: NASH output.

RESULTS: Crude lysates were generated and processed using SMartR RNAseq chemistry from Takara coupled to Illumina sequencing using MiSeq (150 X 2 paired end sequencing). 3' mRNA read counts were normalized to sequencing read depth (CPM), log2 scaled, and each gene then standardized to a mean of 0 and standard deviation of 1. Per sample, the change in gene expression was calculated as the difference between steatotic regions and healthy regions, or inflamed regions and healthy regions. Rank Product statistical significance was finally inferred as the geometric mean of the p-value bounds, according to the algorithm of Heskes et al (doi:10.1186/s12859-014-0367-1). imageDx™: NASH data is presented to correlate tissue morphology with gene expression in each ROI selected.

CONCLUSION: Novel spatial data visualization and analysis techniques allowed a comprehensive assessment of spatial genomic profiles within these NASH tissue samples and how they relate to NASH biology. This approach provides new tools to advance our understanding of spatial genomic profiles in liver disease.

METHODS

HISTOLOGY AND WHOLE SLIDE IMAGING: FFPE human NASH tissues were serial sectioned and H&E stained. Unstained sections were reserved for RNA isolation. Whole slide images (WSI) were generated using a Pannoramic SCAN (3D Histech) and uploaded to a cloud-based imaging platform, imageDx™ (Reveal Biosciences).

METHODS

AI-POWERED IMAGE ANALYSIS: Quantitative data were generated using imageDx™, an automated, integrated workflow that analyzes and manages WSI. The quality of each WSI was assessed with imageDx™ QC to detect focus and slide artifacts. Images that did not pass QC were re-scanned or excluded from the analysis.

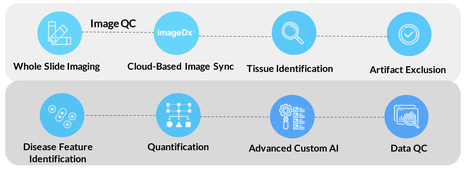

Figure 1. Digital Assay Workflow. The analysis process automated tissue identification, followed by segmentation of regions of interest.

REGION SELECTION: The tissues were analyzed for steatosis and inflammation using imageDx™ NASH, a collection of AI-based digital assays that provide quantitative histopathology from NASH human tissue. The algorithm identified tissue areas on the entire WSI, and positive lipid regions were visualized by an overlaid binary mask. Similarly, inflammatory cells were identified within lobular and portal regions, and immune cells were visualized by a binary mask.

Table 1. Quantitative data generated by imageDx™ NASH.

Tissue Feature	Data Output
Steatosis Macro-vesicular Micro-vesicular Mean Vesicle Size Mean Lipid per Hepatocyte Steatosis Score	Percentage (%), Area (mm^2) Percentage (%) Percentage (%) um^2 um^2 Algorithm-derived score
Ballooning Hepatocyte Density Ballooning Score	Ballooning cells per cm^2 Algorithm-derived score
Immune Cells Immune Cell Foci Inflammation Score	Count, Area (mm^2), Density (cells/mm^2) Foci count, Mean foci size Algorithm-derived score
Fibrosis Area Fibrosis Intensity Junction Branch Analysis Fibrosis Score	mm^2, Percentage (%) Intensity units Type 3 branch % Algorithm-derived score
Mallory Bodies	Presence or Absence
NAS Score	Steatosis score + Inflammation score + Ballooning score
Total Tissue Area Analyzed	mm^2

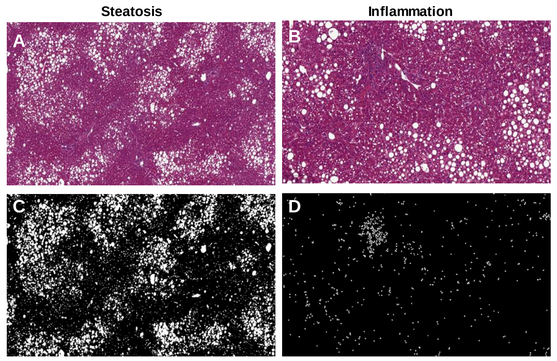

Figure 2. Steatosis and inflammation analysis by imageDx™: NASH. H&E stained WSI (A and B) of human NASH liver were analyzed, and regions of steatosis (C) and inflammation (D) were masked.

RESULTS

AUTOMATED LYSATE EXTRACTION: Automated lysate extraction was performed in the ROIs using the CytoMask™ ink-jet based technology (QuantumCyte). The annotations from the imageDx™: NASH binary masks generated a corresponding print file that directed the application of a novel ink chemistry (CytoMask™ ink) using a Pixdro ink-jet printer on the tissue sections such that the areas outside the ROIs were covered with ink. The CytoMask™ was applied to the tissue sections post deparaffinization. A crude lysate was generated from the ROIs using a proteinase K based lysis buffer by adding 50 µl of lysis buffer directly on to the slide directly over the ROIs under a coverslip and incubating for 90' at 60°C. RNA was purified using the NucleoSpin totalRNA FFPE XS kit from Qiagen and processed for sequence analysis as described below.

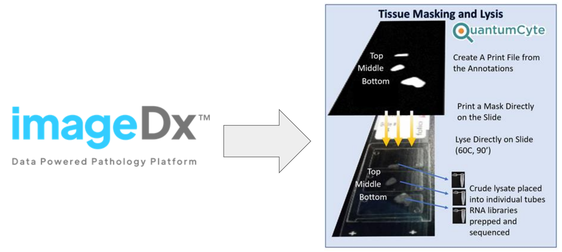

Figure 3: CytoMask Workflow. QuantumCyte's proprietary ink-jet based technology was used to mask and extract RNA from ROIs identified by the imageDx™: NASH output from serial sections.

RNA SEQUENCING AND GENOMIC ANALYSIS: Crude lysates were generated and processed using SMARTer® Stranded Total RNA-Seq from Takara coupled to Illumina sequencing using MiSeq (150 X 2 paired end sequencing). 3' mRNA read counts were normalized to sequencing read depth (CPM), log2 scaled, and each gene then standardized to a mean of 0 and standard deviation of 1. Per sample, the change in gene expression was calculated as the difference between steatotic regions and healthy regions, or inflamed regions and healthy regions. Rank Product statistical significance was finally inferred as the geometric mean of the p-value bounds, according to the algorithm of Heskes *et al* (doi:10.1186/s12859-014-0367-1).

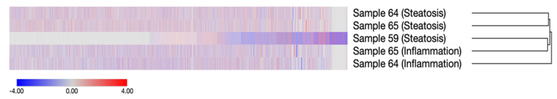

Figure 4. Differential Gene Expression in Steatotic and Inflamed Regions. Heat map analysis of the difference between regions with high vs. low levels of steatosis and high vs. low levels of inflammation within each sample. 17,965 genes are represented. Hierarchical clustering was performed using complete linkage and one minus Spearman rank correlation. Upregulated genes are shown in red, downregulated genes are shown in blue. The data indicates that there are variant genomic expression profiles between samples and indications within the same sample.

RESULTS

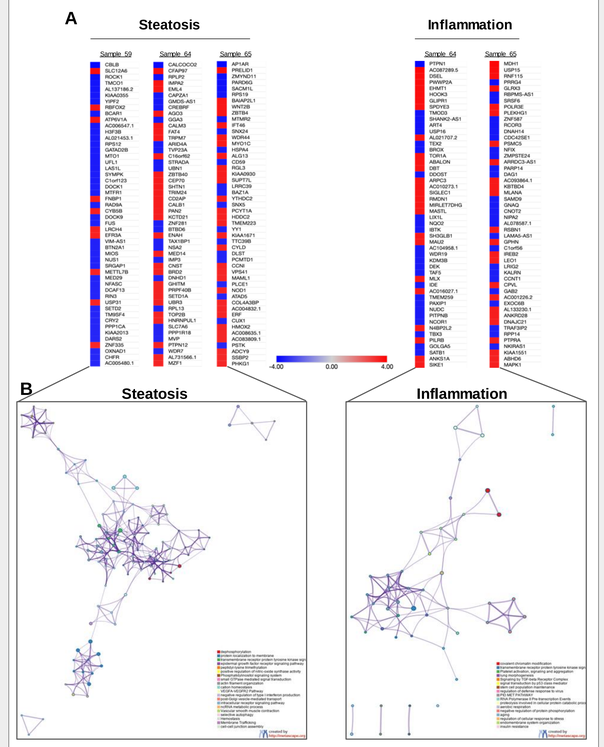

Figure 5. Pathway Enrichment Analysis. (A) Heat map analysis of the top 50 genes with the most differential expression in each sample for each indication. (B) Pathway enrichment analysis of the combined genes from (A) for steatosis and inflammation.

FUTURE DIRECTIONS

- Further insights into liver heterogeneity through the increased spatial resolution offered by this workflow may lead to more accurate NASH outcomes, with potential to accelerate personalized treatments and support tools into the clinical setting.
- The integrated digital histopathology and spatial genomic workflow is built to scale and adaptable for a broad range of disease indications.
- This work can accelerate the development of sophisticated technologies to integrate visualization capabilities into the workflow.

CONCLUSIONS

The presented study demonstrates successful integration of AI-driven pathology with spatial profiling for genomic analysis in human NASH samples. Variant profile differences were observed among the samples, suggesting successful detection of heterogeneity in NASH livers.

American Association for the Study of Liver Diseases - The Liver Meeting Digital Experience™ 2020 November 13-16, 2020 Publication Number 0506