Enhancing cellular specificity of RNA-ISH spot detection using dual-channel staining and AI-based analysis methods

Dan Winkowski,[1] Chun Hao Chen,[2] Fabian Schneider,[1] Joyce Yoo,[2] Regan Baird,[1] Aneesh Acharya,[2] James Mansfield[1]

1) Visiopharm A/S, Hørsholm, Denmark
2) Molecular Instruments, Los Angeles, CA, USA

Download a PDF of the poster

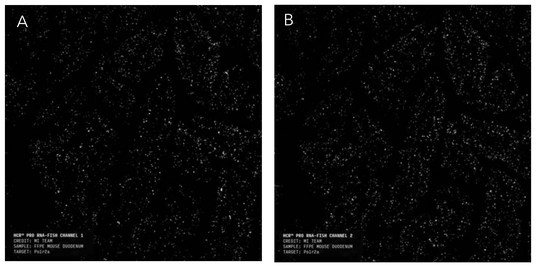
Introduction

RNA-ISH has rapidly achieved prominence within translational research, long holding promise as being critical for novel diagnostic and therapeutic applications where other biomarker types cannot be used. However, there remain key barriers to greater clinical adoption for RNA-ISH, including validation of observed "dots" and accurate identification of RNA within specific cells, particularly in cytoplasmic regions. These significant quantification challenges exist due to a lack of a robust assay validation methodology and difficulties in determining cell boundaries. Here we propose a novel collaboration to address both long-standing issues using a reagent-encoded approach (HCR™ Pro by MI) as well as downstream AI-enabled analysis software (Phenoplex by Visiopharm) to offer a proof-of-concept demonstration of end-to-end assay validation and quantification with high fidelity.

Approach

To validate this methodology, we divided the study into two discrete components:

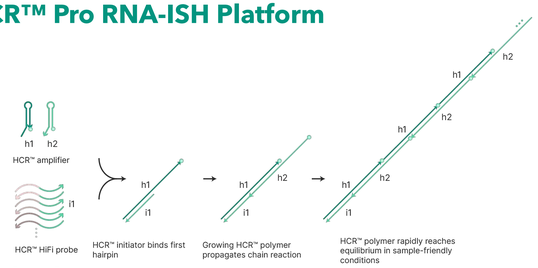
1. Performance characterization of the HCR™ Pro RNA-ISH assay through Dual Channel™ Validation (DCV)
2. Quantification of observed RNA "dots" with orthogonal methodologies (human-interpreted and AI-enabled).



Dual Channel™ Validation of the Polr2a HCR™ HiFi Probe in FFPE Mouse Duodenum Tissue.
(A) HCR™ Pro RNA-FISH targeting Polr2a in the 546 channel. (B) HCR™ Pro RNA-FISH targeting Polr2a in the 647 channel.

The analysis reveals a colocalization factor of 85% between the two channels, with an intensity correlation coefficient of 0.71. These findings validate the Polr2a HCR™ HiFi Probe's ability to specifically bind its target RNA with high sensitivity and specificity, effectively detecting genuine RNA transcripts for precise and reliable dot counting in quantitative analyses. While this does not validate that individual dots are individual transcripts, it does validate that dots can be used as proxies for target abundance.

HCR™ Pro RNA-ISH Platform



Methods

MI's HCR™ Pro RNA-ISH assays feature several key advancements, including: (1) truly protease-free sample pre-treatment for high-fidelity tissue preservation and plug-and-play compatibility with IHC/IF, and (2) incorporation of HCR™ HiFi Probes with DCV that characterizes the specificity and sensitivity of any given assay in any given tissue of interest. DCV is a generalizable assay validation tool that calls upon the property of linear amplification that has been validated in prior literature [1]. Using DCV, individual HCR™ HiFi Probes can be assigned a score that encodes its overall performance within an HCR™ Pro assay.

Additionally, compared to conventional RNA-ISH products that rely upon harsh protease digestion to allow for the diffusion of bulky detection reagents *in situ*, HCR™ Pro offers a distinct advantage through its use of small, background-suppressed DNA-encoded reagents that enable a truly protease-free sample preparation workflow, which safeguards tissue and epitope integrity. This allows for co-detection with virtually any protein marker on any given tissue section, maximizing the utility of valuable patient-derived samples and enabling the simultaneous use of IHC-based membrane stains for cell segmentation.

The integration of Visiopharm's Phenoplex software, a powerful deep-learning solution powered by AI, along with co-detection of cellular membrane markers, further enhances the identification of cells and their boundaries. User-drawn annotations are used to train an algorithm for the precise recognition of cell boundaries based on membrane stains. This significantly improves cell boundary detection and therefore RNA spot localization, paving the way for more accurate spot counts and a deeper understanding of RNA expression within the complex tumor microenvironment.

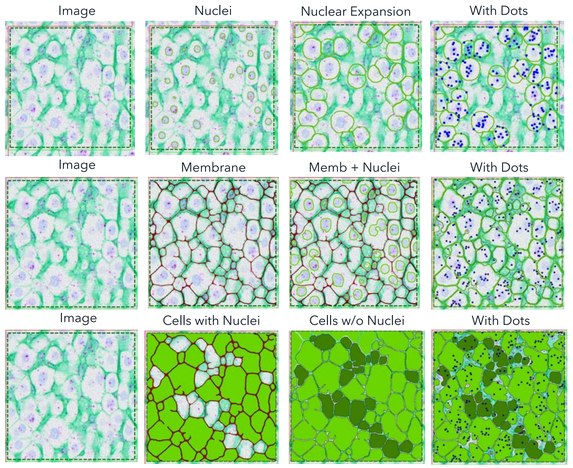
Here we present findings using orthogonal methods of dot counting and assignment: (1) human-interpreted scoring (N=4), (2) AI-enabled scoring based upon IHC membrane stain, and (3) software-enabled scoring based upon nuclear expansion from a standard hematoxylin stain.

Results: AI Counting

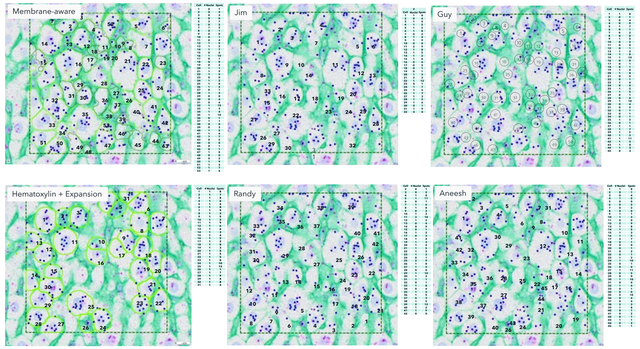
Hematoxylin-only. Top row. This method uses the hematoxylin counterstain as a basis to find nuclei, and then expands the nuclear region to create a cytoplasmic area around that, which is by far the most common cell segmentation method. The algorithm used is an AI / Deep Learning based method which was trained on hundreds of thousands of nuclei across a range of H&E and IHC images (and comes standard in Oncotopix Discovery). The left image shows a small area of the whole slide. The next image shows the nuclei that were found using the AI algorithm. The next images shows the expanded cytoplasmic area around each nucleus. The nucleus and cytoplasm together define what is used as a cell in this analysis. The final image shows the cell overlaid with the spots that were found using a separate AI-based algorithm.

Membrane-aware. Middle row. This algorithm used user-drawn annotations of cells to train a new AI algorithm that detects cell boundaries based on the membrane stain. This new algorithm was trained on one sample and then applied to all samples with excellent success. The nucleus was then found using the same AI algorithm as above, just to be able to determine the number of nuclei in each cell (0, 1, or 2). Spot counting per cell was done using the same spots as all other counting methods (human and machine).

Nuclear status of membrane-aware cells. Bottom row. Further visualization of which cells found using the membrane-aware algorithm contained nuclei (light green) and which did not (dark green).



Results: Manual Counting

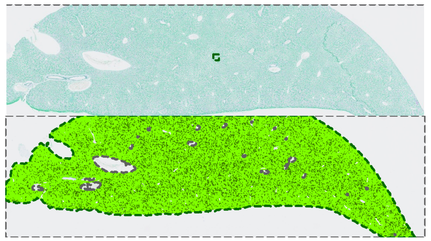


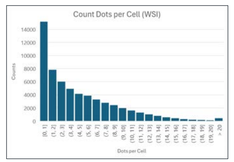
Person	# Cells	Average # nuclei	STD DEV	Average # spots	STD DEV
Jim	32	0.75	0.57	4.69	3.04
Guy	50	0.46	0.50	3.5	3.35
Randy	42	0.71	0.60	4.24	3.28
Aneesh	45	0.53	0.66	3.31	3.29
Membrane-aware	53	0.51	0.54	3.45	3.27
Hematoxylin + Expansion	31	1	0	5.87	2.81

Human scoring versus AI scoring. In order to test whether using the membrane stain as an outline of the cell for machine learning improves the counting of the number of RNA spots per cell, we compared four human spot counters to each of the AI-based algorithms. The humans were asked to find every cell that was not outside the boundary (the dotted square in the images), to place a number on that cell, and then to count the number of nuclei in that cell and the number of spots. To keep the exercise to only be about what is and what is not a cell, each person used the same spots as the basis for spot counting.

The results for Jim, Guy, Randy, and Aneesh can be seen in the cell number overlays on each image and in the tables beside each image. The summary statistics can be seen in the table below.

Although it is difficult to make a rigorous comparison between human reads, the humans were generally similar to each other, depending on how strict they were with edge cell counting. They each found cells that had 0, 1, and 2 nuclei, and for those cells that are obviously the same cell (despite the differing numbering schemes), counted a similar number of spots. What was different was the basic "hematoxylin only" counting method. That clearly misses many cells, splits multinucleated cells inappropriately, and clearly counts spots from outside a given cell's boundaries (see lower left image). So, the hematoxylin-only method undercounts cells and overcounts the number of spots per cell.





Whole-slide scoring results. The membrane-aware algorithm was applied to whole-slide images. Upper image shows the entire sample. First, a tissue segmentation was performed on the whole-slide image. Pancreatic tissue is shown in green and everything else (blank slide, vessels, junk, etc) is shown in white (middle image).

The number of dots per cell are show in the histogram (left). All cells are counted, regardless of number of nuclei (0, 1, 2). The largest population are cells with just one spot.

Results: Data Exploration

Visualizing cells with different numbers of spots. A tSNE plot was created using all cells from the whole-slide image analysis. That tSNE plot was then split into those cells who contained a nucleus (or nuclei, suplot at right) and those cells which did not contain any nucleus (left subplot). Cells with different number of spots appear in the same locations in the tSNE – and one can see that by overlapping the two subplots, cells with the same number of spots appear in the same place in the joint tSNE.

By selecting those cells which have no nucleus and which have 6 spots per cell, one can see in the thumbnail image gallery (right) how each of those cells looks. These are all cells that would have been missed by a "nucleus only" cell segmentation method. This use of tSNE plots as a means by which one can select cells to view in a thumbnail gallery is extremely useful, as is the 'splitting' of the tSNE plot into sub-tSNE plots.

Conclusions

Based upon these initial findings, we propose two primary conclusions:

1. A membrane stain-based approach is superior to a nuclear expansion-based approach
2. Use of a membrane stain in assisting a human quantifier is superior to relying solely on hematoxylin for cell identification

We suggest that RNA-ISH quantification can be generally improved by the simultaneous use of an IHC-based membrane stain whenever possible. MI offers a validated library of HCR™ Membrane Stains that can be used across many common tissue types and species.

These enabling features lay the groundwork for robustly characterized, clinical-grade RNA-ISH assay performance, as well as high-fidelity assignment of RNA dots to correct cell compartments using cell segmentation via a physical biomarker. This facilitates the translation of high-value RNA-ISH assays from research to clinical settings, ultimately enabling clinicians to make reliable, clinically meaningful decisions based on RNA quantification, regardless of analysis approach.

References

[1] Harry M. T. Choi, Maayan Schwarzkopf, Mark E. Fornace, Aneesh Acharya, Georgios Artavanis, Johannes Stegmaier, Alexandre Cunha, Niles A. Pierce; Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development 15 June 2018; 145 (12): dev165753. doi: https://doi.org/10.1242/dev.165753