Poster # 2512

A multi-institution examination of concordance in spatial transcriptomics using the GeoMx® Cancer Transcriptome Atlas

Tyler Hether[1], Tim Howes[2], David Scoville[1], Charlie Glaser[1], Yanyun Li[2,3], Rami Vanguri[3], Neeman Mohibullah[3], Wan-Jung Chang[5], Todd Yoder[5], Minnal Gupta[5], Kathy Ton[1], Yan Liang[1], Ying Huang[4], Zach Herbert[4], Jason Reeves[1], Elizabeth Mittendorf[2,4], Simon Lacey[2,5], Travis Hollmann[2,3], Sarah Warren[1], and Theresa LaVallee[2]

1. NanoString Technologies, Inc., Seattle, WA 2. Parker Institute for Cancer Immunotherapy, San Francisco, CA 3. Memorial Sloan Kettering Cancer Center, New York, NY 4. Dana Farber Cancer Institute, Boston, MA 5. Center for Cellular Immunotherapies, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

STUDY OBJECTIVE & APPROACH

- Multiplexed spatial profiling can enable biological insights by characterizing gene expression within discrete physical locations of a tissue.
- To support collaborative studies that involve multiple institutions, we studied how consistent the results were for NanoString® GeoMx® spatial profiling with the Cancer Transcriptome Atlas, by analyzing a common set of samples across 4 different laboratories and in serial sections of 4 different samples: 2 cell pellet arrays (CPAs), healthy tonsil germinal centers, and healthy colon villi (see table below).
- Varying sizes for the selected Regions of Interest (ROIs) were also investigated to determine an appropriate minimum size for obtaining consistent results.
- The number of genes detected above Limit of Quantification (LOQ) and the correlations of their expression levels both between and within slides were measured to assess the quality and concordance of results.

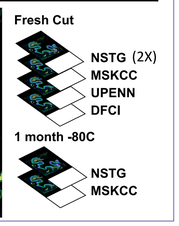
EXPERIMENTAL OVERVIEW

We examined intrasample and between sample heterogeneity using a mix of biological tissues and cell pellets. See the table below for more details.

Table 1. Set of tissue samples analyzed as replicates at each institution

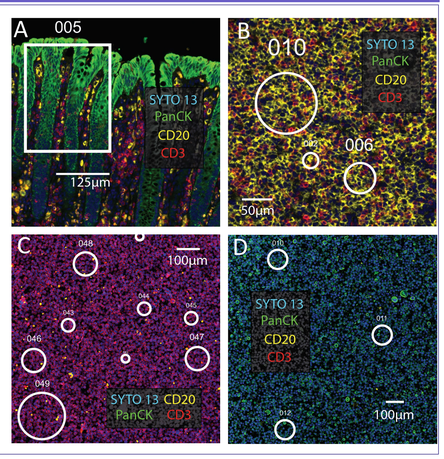
Tissue/Cell Pellet Array (CPA)	Experimental Rationale	# ROI/ROI Size/Shape
Colon	Assess intrasample heterogeneity	6 x 250 µm by 200 µm rectangles
Tonsil	Replicate ROIs to assess intrasample heterogeneity and 3 different sizes of ROI to assess variability across different amounts of biological tissue	4 x 25 µm diameter circles 4 x 50 µm diameter circles 4 x 100 µm diameter circles
8-core CPA See Table [CPA21] for details	Triplicate ROIs to assess intrasample heterogeneity and consistency across different cell lines	3 x 100 µm diameter circles on all cores
CCRF-CEM/HEK293 Dilution Series CPA 6 cores of varying percentage of HEK293	Assess intrasample heterogeneity and variability across different amounts of densely packed cell pellets	3 x 25 µm diameter circles on pure cell lines 3 x 50 µm diameter circles on all cores 3 x 100 µm diameter circles on pure cell lines 3 x 200 µm diameter circles on all cores 3 x 500 µm circles on pure CCRF-CEM

- In each FFPE block, serial sections were cut and sent to one of 4 laboratories.
- Two additional serial sections were cut and stored at -80°C and profiled by NanoString (NSTG) & MSKCC
- Throughout this study, slide names are labeled by the laboratory code, replicate number (1, 2), and preservation method (-80°C, fresh cut).



REPRESENTATIVE IMAGES FROM EACH EXPERIMENT

- Representative ROI(s) from the four different experiments are shown. In all cases, ROIs are labeled by their 3-digit number
  - See Table 1 for ROI distribution
- A: Healthy colon villi
- B: Tonsil germinal center with three varying sizes
- C: Dilution Cell Pellet Array from that consists of 100% HEK-293. 3 of the 5 varying ROI sizes are shown
- D: Three 100µm diameter ROIs from cell line COLO201



MODELING SLIDE VARIATION

In multi-slide GeoMx analyses, a linear mixed model (LMM) is used for certain statistical analyses (e.g., Differential Expression) to account for the non-independence of samples (i.e., multiple ROIs are measured for a given slide). This is done by treating "slide" as a random effect in the LMM and contrasting different groups of interest as a fixed effect. For comparing expression levels prior to such statistical analyses, modeling slide variation directly can allow for more direct comparisons of individual ROIs and we show this technique in:

- Colon tissue, where slide effects were high relative to the "biological signal" (i.e., any differences in expression between samples) and
- In the 8-core CPA analysis, where slide effects were minimal compared to the biological signal (i.e., differences between cell lines).

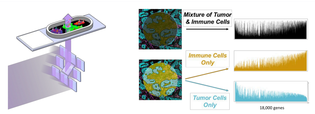
In the current study, care was taken to ensure that "slide" was not confounded with "cell line" or "target diameter". We note that experimental design should be taken into consideration when doing this slide-modeling approach so that the biological signal is not inadvertently removed from the data.

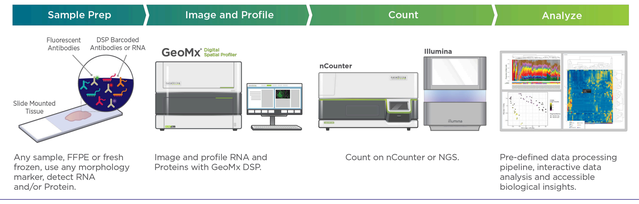
Slide-modeled algorithm:

- For a given gene, its log2 normalized expression was used as the response
- An intercept was modeled as the only fixed effect and slide was used as the random effect (with random intercept); this was done in the R package *lme4*.
- Both residuals and intercept's estimate were converted back to the linear scale
- Residuals were then multiplied by the intercept's estimate

GEOMX TECHNOLOGY FOR SPATIAL ANALYSIS IN TISSUES

NanoString's GeoMx® Digital Spatial Profiling (DSP) technology enables high throughput, spatially resolved analysis of gene or protein expression from tissues by profiling regions of interest (ROIs) selected based on fluorescently labeled visualization markers.





COLON: WITHIN AND BETWEEN SLIDE SIMILARITY

- All 42 ROIs passed QC
- 1,075 genes were consistently above background in >10% of ROIs
- PCA of normalized expression showed some slide-to-slide variation; this clustering is reduced by modeling slide effects (Fig 1).
- High concordance was observed between slides: Pearson's *r* = 0.906 - 0.995 (Fig 2).
  - In the least similar comparison, 64% of genes were below LOQ (Fig 3)
  - In the most similar comparison, 22% of genes were below LOQ (Fig 3)
- Within-slide correlation was similar to between-slide correlation (Fig 4):
  - Pearson's *r* within-slide = 0.925 - 0.995
- Modeling slide effects increased between-slide similarity without impacting within-slide variation (Fig 2; Fig 4)

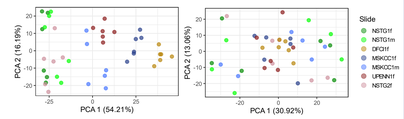


Fig 1: PCA on 1,075 genes using normalized (left) and slide-modeled (right) data.

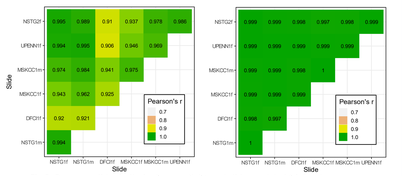


Fig 2: Between-slide similarity for raw (left) and slide-modeled (right) data.

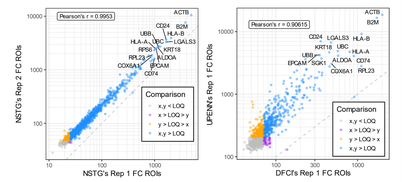


Fig 3: Details showing the highest and lowest similarity observed in the raw data. Each point is a gene and color denotes whether a gene was above LOQ in a given sample set. Left: average raw expression between slides NSTG1f and NSTG2f. Right: average raw expression between slides DFCI1f and UPENN1f.

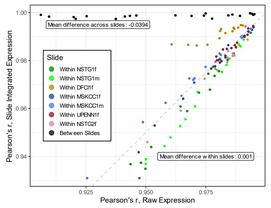


Fig 4: Pairwise comparisons within and between slides, before and after slide modeling.

CPA: HIGH CORRELATION BETWEEN SIMILAR DILUTIONS

- CPA dilution series analysis varied *both* the ROI target size and the proportions of two cell lines
- 325 ROIs passed QC
- 950 genes were consistently above LOQ in >10% of ROIs
- **ROI size analysis**:
  - Increasing ROI size increases the number of genes above LOQ (Fig 5)
  - Between sites and replicates there was a marginally significant difference in average Pearson correlation based on target diameter (P<0.049, Kruskal-Wallis). *Post hoc* Dunn Test showed 25µm was less concordant that 500µm. (Fig 6)
  - Within a given slide, concordance increased with ROI diameter (P < 1.3e-4, Kruskal-Wallis; Fig 7).
- **Dilutions analysis** (Fig 8):
  - In all slides, samples with similar composition (e.g., 0% vs 5% HEK293) had high correlation compared to samples with more divergent composition.
  - The largest differences were seen between pure CCRF-CEM and pure HEK293 cell lines; average Pearson's r = 0.82.

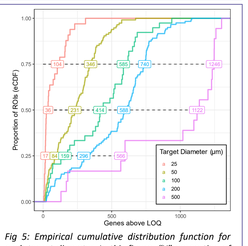


Fig 5: Empirical cumulative distribution function for each target diameter. In this figure, "Y" proportion of ROIs have at most "X" genes above LOQ. Number of genes above LOQ in 25%, 50%, and 75% of ROIs for each target diameter labeled. Data includes all CPA cores.

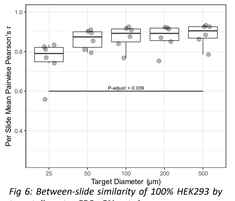


Fig 6: Between-slide similarity of 100% HEK293 by target diameter. FDR >5% not shown.

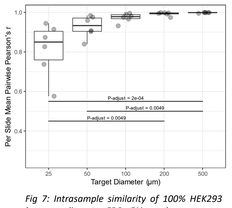


Fig 7: Intrasample similarity of 100% HEK293 by target diameter. FDR >5% not shown.

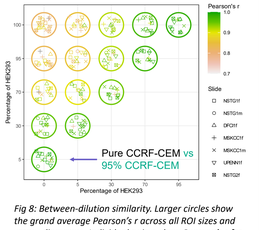


Fig 8: Between-dilution similarity. Larger circles show the grand average Pearson's r across all ROI sizes and target diameters. Individual points show Pearson's r for a given slide and target diameter.

TONSIL: HIGH CONCORDANCE ACROSS ROI DIAMETERS

- 83 ROIs passed QC for three ROI diameters (25, 50, and 100µm)
- 816 genes were above LOQ in >10% of ROIs
- The number of genes detected above LOQ varied with target diameter (Fig 9)
  - LMM Estimate: Genes Detected = 182 + 4.8*(diameter); P < 2e-16
- Gene expression between slides showed no difference across target diameter (Fig 10) with high concordance between slides (Fig 11)
  - Average Pearson's *r* = 0.971 – 0.988
- Within slides, concordance was overall high but 100µm ROIs were more consistent compared to 25µm ROIs (Fig 12).

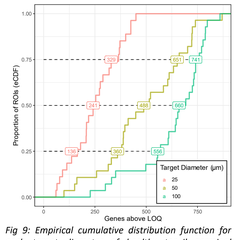


Fig 9: Empirical cumulative distribution function for each target diameter of healthy tonsil germinal centers. In this figure, "Y" proportion of ROIs have at most "X" genes above LOQ. Number of genes above LOQ in 25%, 50%, and 75% of ROIs for each target diameter labeled.

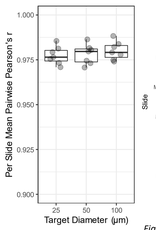


Fig 10: Between-slide average similarity for each target diameter. Kruskal-Wallis test: χ^2 = 0.83, df = 2, P < 0.66.

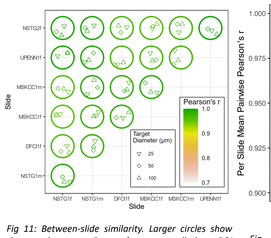


Fig 11: Between-slide similarity. Larger circles show the grand average Pearson's r across all three ROI diameters. Individual points show the between-slide concordance for a given ROI diameter.

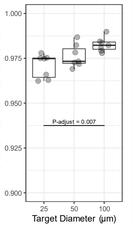


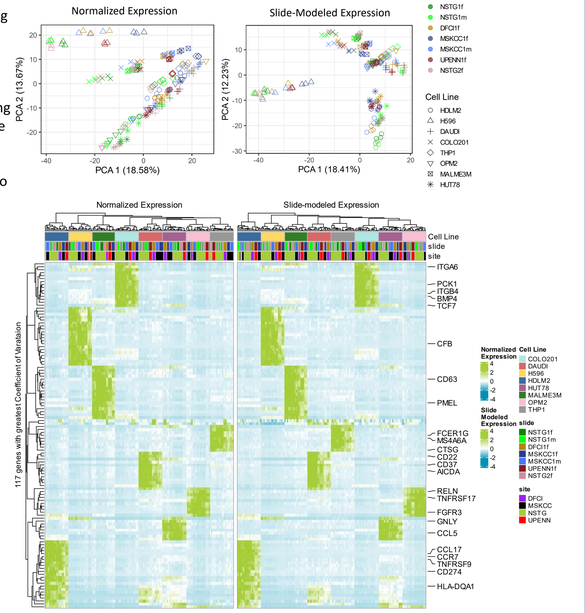
Fig 12: Within-slide average similarity for each target diameter. P-values are Benjamini & Hochberg adjusted based on Dunn Test. FDR >5% not shown.

8-CORE CPA: CELL LINE-SPECIFIC DIFFERENCES IN GENE EXPRESSION AND CONCORDANCE

- The 8-core CPA contained different cells lines of immunological significance (Table 2)
- 167 ROIs analyzed with 1,170 genes above LOQ
- Strong clustering observed based on cell lines (Fig 13). While no slide modeling was needed here, performing a batch correction did not erode the biological signal
- Comparing across slides (Fig 14):
  - Average Pearson's r was high with a range of 0.841 – 0.947
  - Similarity depended on the specific cell line with DAUDI and OPM2 showing lower overall slide-to-slide similarity compared to others. This suggests the presence of within-line heterogeneity that contributes to dissimilarity
- Across individual replicates we saw high concordance across all observations and line THP1 showing somewhat weaker within-slide correlation compared to line MALME3M (Fig 15).

Table 2: Cell lines profiled in this experiment. For each cell line and for each slide a total of three replicate ROIs of 100µm diameter were chosen.

Cell Line	Description
MALME3M	Fibroblast derived from malignant melanoma of the lung
H596	Human lung adenosquamous carcinoma
THP1	Monocytic leukemia cell line -- derived from blood of patient with acute monocytic leukemia
HDLM2	Hodgkin lymphoma
COLO201	Adenocarcinoma of the colon (Dukes Classification Grade D)
HUT78	T cell lymphoma -- Derived from peripheral blood of a patient with Sezary syndrome
OPM2	Multiple myeloma -- Peripheral blood from patient with multiple myeloma in leukemic phase
DAUDI	Burkitt's lymphoma



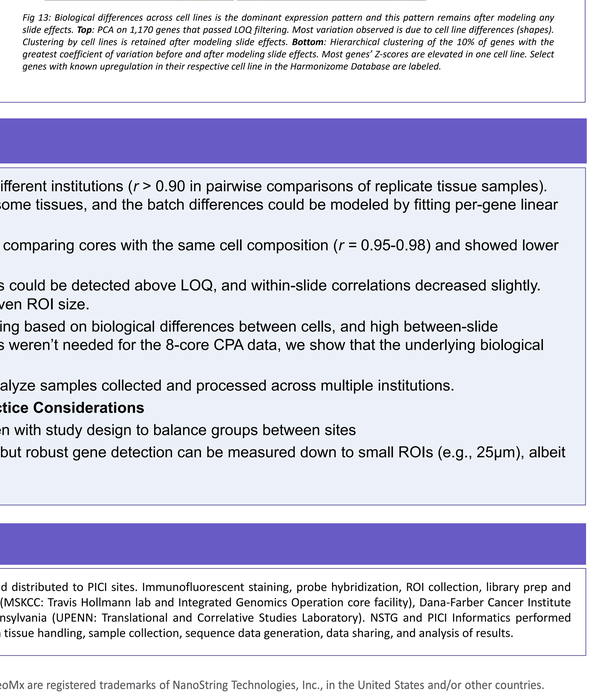


*Fig 13: Biological differences across cell lines is the dominant expression pattern and this pattern remains after modeling any slide effects. **Top**: PCA on 1,170 genes that passed LOQ filtering. Most variation observed is due to cell line differences (shapes). Clustering by cell lines is retained after modeling slide effects. **Bottom**: Hierarchical clustering of the 10% of genes with the greatest coefficient of variation before and after modeling slide effects. Most genes' Z-scores are elevated in one cell line. Select genes with known upregulation in their respective cell line in the Harmonizome Database are labeled.*

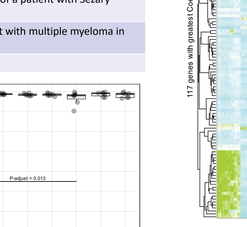


Fig 14: Between-slide average Pearson Correlation for each cell line investigated. P-values are Benjamini & Hochberg adjusted based on Dunn Test. FDR >5% not shown.

Fig 15: Within-slide average Pearson Correlation for each cell line investigated. P-values are Benjamini & Hochberg adjusted based on Dunn Test. FDR >5% not shown.

SUMMARY & BEST PRACTICES

- Gene expression levels showed high correlation across slides run at different institutions (*r* > 0.90 in pairwise comparisons of replicate tissue samples). Principal component plots showed some evidence of batch effects in some tissues, and the batch differences could be modeled by fitting per-gene linear models with varying intercepts across slides.
- Cell pellet dilution series showed appropriately high correlations when comparing cores with the same cell composition (*r* = 0.95-0.98) and showed lower correlations when comparing dissimilar cell compositions.
- When smaller ROI sizes were selected (as low as 25 µm), fewer genes could be detected above LOQ, and within-slide correlations decreased slightly. Nevertheless, high correlations were observed between slides for a given ROI size.
- Using an array of cancer-related cell lines, we observed strong clustering based on biological differences between cells, and high between-slide correlation within each cell line (*r* > 0.84). While modeling batch effects weren't needed for the 8-core CPA data, we show that the underlying biological signal is retained after such modeling.
- These findings support the use of GeoMx spatial transcriptomics to analyze samples collected and processed across multiple institutions.

Best Practice Considerations

- Data generated at different institutions can be compared if care is taken with study design to balance groups between sites
- ROIs with larger areas tend to have higher number of genes detected but robust gene detection can be measured down to small ROIs (e.g., 25µm), albeit with greater variability

ACKNOWLEDGMENTS

Tissue and cell pellet array slides were prepared by NanoString Translational Group (NSTG) and distributed to PICI sites. Immunofluorescent staining, probe hybridization, ROI collection, library prep and sequencing were performed independently at NSTG, Memorial Sloan Kettering Cancer Center (MSKCC: Travis Hollmann lab and Integrated Genomics Operation core facility), Dana-Farber Cancer Institute (DFCI: Elizabeth Mittendorf lab and Molecular Biology Core Facilities), and University of Pennsylvania (UPENN: Translational and Correlative Studies Laboratory). NSTG and PICI Informatics performed downstream data processing and concordance analysis. We thank all of the scientists involved in tissue handling, sample collection, sequence data generation, data sharing, and analysis of results.

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.