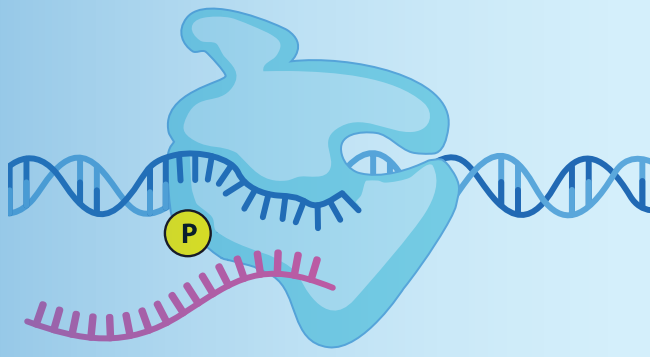
Methods in Molecular Biology 2880

Springer Protocols



Nalini Raghavachari
Natalia Garcia-Reyero *Editors*

Gene Expression Analysis

Methods and Protocols

Second Edition

MOREMEDIA

Humana Press

Chapter 9

Integrating Tissue Microarray to GeoMx® Digital Spatial Profiler : Spatial Transcriptomics Assay with Bioinformatics Analysis

Deshica Dechamma, Manju Moorthy, Vijayalakshmi Bhat, and Gopalakrishna Ramaswamy

Abstract

Genome-wide or high-plex gene expression is important to understand the organism, tissue, and cellular mechanism. From microarrays to next-generation sequencing (RNA-Seq) at the bulk level and then to the single cell level, gene expression studies have undergone a long transition. The current bulk gene expression and pathway-centric approach toward disease and therapeutics is moving toward spatial transcriptomics, which makes it possible to profile gene expression from the cellular microenvironment without any loss of spatial information. Spatial transcriptomics allows us to understand cellular interactions, cell type abundance, and profile expression differences between the region of interest and its microenvironment. The technology is revolutionizing oncology, developmental biology, neuroscience, preclinical studies, and many therapeutic approaches, especially immunotherapy. Taking into consideration the diverse spatial transcriptomics technologies available, the current chapter aims to delineate the NGS assay protocol for Digital Spatial Profiler (DSP) and follow bioinformatics analysis. While the workflow itself has been detailed elsewhere, in this chapter we are focusing on the integration of tissue microarrays, bioinformatics pipelines, and statistical approaches specific to GeoMx RNA assays as well as common errors that can occur while running a DSP RNA assay. The descriptions of these methods refer to the current version of the GeoMX DSP guidebook and GeoMx Data Analysis Manual, which can be downloaded from the documents section of NanoString website. These documents and user guides are continuously improved and updated; hence, it is important to regularly check the company's website for the most recent version.

Key words Spatial biology, Transcriptomics, Tissue microarray (TMA), Formalin-fixed, paraffin-embedded (FFPE) specimens, Fresh frozen (FF), Fixed frozen (FxF) , Digital Spatial Profiler (DSP), Whole Transcriptome Atlas (WTA), Cancer Transcriptome Atlas (CTA), next-generation sequencing (NGS), Abs (antibodies), ROI (region of interest), Bioinformatics analysis for spatial transcriptomics

Authors Deshica Dechamma and Manju Moorthy have equally contributed to this chapter.

Nalini Raghavachari and Natalia Garcia-Reyero (eds.), *Gene Expression Analysis: Methods and Protocols*, Methods in Molecular Biology, vol. 2880, https://doi.org/10.1007/978-1-0716-4276-4_9,

1 Introduction

The GeoMx DSP is one of the commercially available instruments from NanoString for performing spatial transcriptomics and proteomics. Despite being economical than other spatial experiments, it can be cost prohibitive and time consuming to assay multiple samples at a time. Tissue microarrays are composite paraffin blocks constructed by extracting cylindrical tissue core "biopsies" from different paraffin donor blocks and re-embedding these into a single recipient (microarray) block at defined array coordinates [1]. It allows researchers to study hundreds of tissues simultaneously under identical experimental conditions. Combining TMA technology with the GeoMx DSP experiment protocol will allow us to obtain high-throughput spatial data from multiple samples at a time.

The quantification of the RNA expression levels is possible within the Digital Spatial Profiler (DSP) technology by counting the barcodes digitally through two ways, either using nCounter® Analysis System or sequencing on a next-generation sequencing platform. While the former have been explained elsewhere [2, 3], the latter requires in-depth explanation, especially in terms of bioinformatics analysis. There has been a review of analysis methods used in spatially resolved transcriptomics, especially for barcode-based technology [4] but not specific to GeoMx DSP. It is important to note that though GeoMx Digital Spatial Profiler system provides graphical user interface-based data analysis pipelines, it is not that everyone would have access to the machine, making it necessary to understand the analysis workflow using standalone software outside the machine. Here we cover the standard bioinformatics workflow for GeoMx DSP data analysis using standalone software and R packages, ensuring to point out trouble-shooting methods in each step of the data analysis.

2 Materials

Consent for the use of tissue/biopsy samples is required from the patients and/or repository as per the ethical approval and guidelines. The preparation and use of the experiment materials should be performed at room temperature, unless indicated otherwise. All equipment and environment used in these processes needs to be DNase/RNase free.

Spatial transcriptomics assay performed through GeoMx involves multiple steps and appropriate preparation would be needed. Here in the workflow we provide the steps involved in a typical TMA-based experiment on GeoMx DSP platform (Fig. 1).

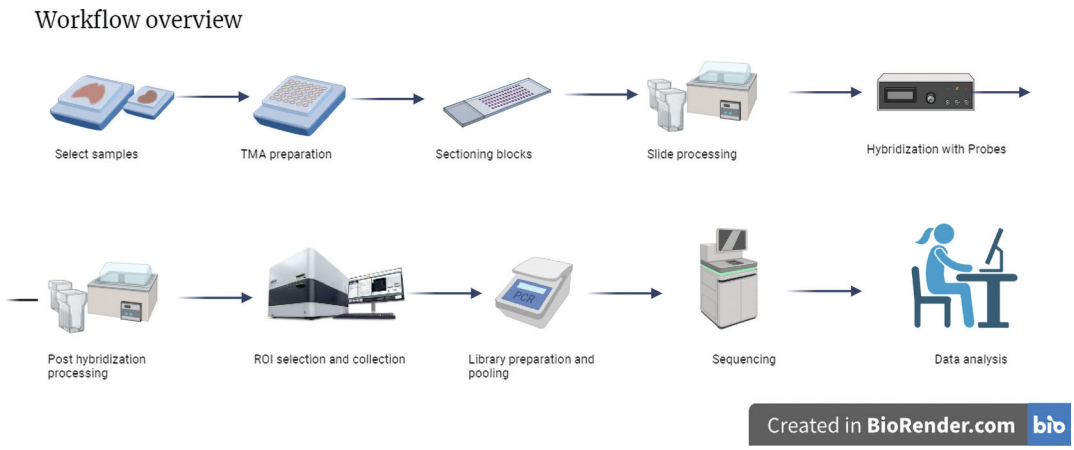


Fig. 1 Complete workflow of GeoMx® Digital Spatial Profiler (DSP) system

2.1 TMA Preparation

Equipments:

- Quick-Ray needle.
- Refrigerator.

Materials

- Donor blocks.
- Recipient block (Quick-Ray or similar one).
- Spatula.
- Gloves.
- Good grade paraffin wax.

2.2 Tissue Sectioning

Equipments

- Microtome.
- Floatation tank.
- Baking oven.

Materials

- Sectioning blades.
- DEPC-treated water.
- Ice block.
- Positively charged slides (Fisher Scientific, 12-550-15 or Leica Biosystems, S21.2113.A).

2.3 Slide Preparation

Equipments

- Hybridization oven including hybridization chamber*: HybEZ II Hybridization System or RapidFISH Slide Hybridizer (ACD-Bio, 321710/321720, Boekel Scientific, 240200 for 120 V).
- Water bath (programmable to at least 37 °C).

- 5-quart steamer—Russell Hobbs 19270-56 or other.
- Digital thermometer.
- Picofuge.
- Vortexer.
- Pipettes—10 μL, 20 μL, 200 μL, 1000 μL.

Materials and Reagents

- Coplin jars.
- Filter tips (DNase/RNase free).
- Microcentrifuge tubes (DNase/RNase free).
- Superfrost Plus microscope slides or Leica BOND Plus microscope slides (more adhesive; recommended for tissues prone to detaching from slides).
- Slide staining jars (Coplin jars) (qty 16; recommend at least two of plastic for use in the steamer) and slide holder inserts.
- Humidity chamber.
- Simport, M920-2 (select black lid) (or comparable).
- HybriSlip hybridization covers (22 mm × 40 mm × 0.25 mm) Grace Bio-Labs, 714022.
- RNase AWAY® or 10% Bleach (RNaseZap® is not a substitute) Thermo Fisher, 7003PK. RNAseZap ThermoFischer AM9780.
- Heat/cold protectant handling gloves.
- Forceps (for slide handling).
- Aluminum foil.
- Kimwipes™.
- USB drive v3.0, 64 GB or higher (able to be NTFS formatted).

Reagents

- DEPC-treated water Thermo Fisher, AM9922 (or comparable).
- 10X phosphate buffered saline pH 7.4 (PBS) Sigma-Aldrich, P5368-10PAK, P5368-5X10PAK (or comparable).
- 10% neutral buffered formalin (NBF). EMS Diasum, 15740-04 (or comparable).
- 100% deionized formamide Thermo Fisher, AM9342 or VWR, VWRV0606 (or comparable) 4 °C (bring it to RT before opening).
- 20X SSC (DNase/RNase free) Sigma-Aldrich, S6639.
- Proteinase K Thermo Fisher, AM2546, AM2548, or 25530049.
- 10X Tris–EDTA pH 9.0 (Antigen Retrieval Solution, 10X concentrate) Thermo Fisher (eBioscience™), 00-4956-58.
- Tris base Sigma-Aldrich, 10708976001 (or comparable).

- Glycine Sigma-Aldrich, G7126 (or comparable).
- CitriSolv or Xylene or D-Limonene ((R)-(+)-Limonene), Fisher Scientific, 04-355-121).
- Sigma-Aldrich, 183,164-100ML or 183164-500ML (or comparable).
- 100% ethanol (ACS grade or better).

Reagents-Supplied by NanoString: Contact your NanoString Sales Representative

- GeoMx RNA Slide Prep Kit.
- GeoMx Morphology Kit—Human or Mouse RNA compatible (available from various manufacturers) or custom.
- GeoMx Nuclear Stain Morphology Kit.
- For NGS readout: GeoMx Probe Mix for NGS readout (available from various manufacturers; green or white cap).

2.4 ROI Selection and Collection

- GeoMx DSP Instrument.
- Illumina NGS Sequencer.

2.5 Data Analysis

- Explained in detail section 4.0.

3 Methods

3.1 TMA Construction

The first thing to consider before preparation of TMA is how many samples pertaining to the study you want to accommodate in one slide; given that the scan area of GeoMx DSP is 35.3 mm long and 14.1 mm wide, the constructed TMA should fit in this area.

You could accommodate:

- $12 \times 6 = 72$ cores of 1 mm each.
- $9 \times 4 = 36$ cores of 2 mm each.
- $3 \times 6 = 18$ cores of 3 mm each.
- $2 \times 5 = 10$ cores of 5 mm each.

Selecting core size is dependent on study, morphology/heterogeneity of the tissue, data required, and various such factors.

Note: This is one workflow to construct a tissue microarray, other workflows, including an automated workflow are available, which can be used for further performing a DSP experiment.

The procedure must be carried out in an RNase-free environment.

Once the number of samples to accommodate in the TMA is finalized, the donor blocks are selected. H&E-stained slides, pathology reports, and other metadata, if any, for these blocks are necessary for the pathologist to select area(s), to be transferred into the recipient block.

- A pathologist/scientist should mark the areas on the slide with a felt-tipped waterproof pen, the marked areas are matched to corresponding donor blocks.
- After all donor blocks and areas are determined, a TMA template is created listing all the samples to be used in the TMA construction, in particular order, marking the coordinates for identification of orientation. It is a good practice to leave an empty core or place a very distinguishable tissue in the TMA to identify the orientation of the TMA.
- In a tray, arrange the donor blocks marked with area to be cored, according to the template created in the previous step.
- Take a recipient block based on the core size determined previously and place it on a benchtop.
- Confirm the donor block IDs against the TMA template.
- Ensure the recipient block is clean, check the inside wells for any residual paraffin. Examine the block against a lamp—check for any breaks.
- Set the Quick-Ray® needle to the determined core size, punch firmly and obtain the core of the previously marked area on the tissue block.
- Place it above the preidentified well on the recipient block and slowly release until the last end of the tissue reaches the tip of the needle, gently push the remaining tissue inside the well, with a blunt spatula. Use a sharp disposable scalpel to trim of any excess paraffin.
- Use RNaseZap and a lint-free wipe to clean the needle between each sample.
- Repeat the steps to fill the TMA block.
- Place the recipient block on the Glass slide with the cut section facing down, and incubate at 60 °C for 10 min. After 10 min remove the slide with the block without disturbing the setup and keep it at room temperature for 10 mins.
- Pour the wax (80–85 °C) on top to embed and allow it to completely cool (preferably in the refrigerator for 2 h).
- Trim the surface and proceed to take 3–5 μ sections for staining.
- Always stain H&E to check for representation of the tissue core in the TMA.
- Less than 100 cells of representative tissue may be inadequate representation.
- Store the TMA block in a cool and dry place. Avoid exposure to humidity and high temperature for adequate antigen preservation.

3.2 Tissue Sectioning

- Fill the floatation tank with DEPC-treated water and set the temperature to optimum (usually 38–42 °C).
- Place the TMA block on an ice block for 5–10 min, before placing it in the microtome.
- Insert a new blade to the microtome; trim and discard the first few sections from the block face.
- FFPE tissue sections should be cut 5 μm thick on a calibrated microtome.
- Tissue sections must be placed exactly in the center of the slide and be no larger than 35.3 mm long by 14.1 mm wide.
- The sections used should not have nicks or trapped air which might cause lifting of tissue.
- Air-dry mounted slides overnight or bake for 1 h at 60 °C; these slides can be stored at 4 °C/room temperature for up to 3 months without affecting the quality of the results that can be obtained from the tissue.

3.2.1 Morphology Marker Selection and Validation (Before Starting)

Understanding tissue morphology is the first step in a well-designed spatial experiment. In a GeoMx RNA Assay morphological markers are instrumental and are used in visualizing the tissue, selection, and collection of region of interests within a tissue.

The GeoMx instrument has four fluorescent channels: FITC, Cy3, Texas Red, and Cy5. These are optimal for using the fluorophores Alexa Fluor 488, 532, 594, and 647, respectively. Alternative fluorophores can be used; however, the compatibility of excitation and emission spectra should be confirmed before use in a GeoMx study.

Use of nuclei markers is essential for each study as it approximates the number of cells in each ROI. NanoString supplies SYTO13 as a nuclei marker which occupies the FITC channel, alternative nuclei markers like SYTO38 can be used upon validation to open up the FITC channel for other markers.

Directly conjugated primary antibodies offer the most straightforward options for fluorescent staining. When choosing an antibody for your study, it is recommended to begin with those already validated for IF/IHC assays by the antibody provider. Once selected, it is important that they are tested using the specific staining protocols used for GeoMx.

3.2.2 Prepare Reagents

- 95% EtOH: Prepare 500 mL of 95% ethanol by adding 25 mL of DEPC-treated water to 475 mL of 100% ethanol. NanoString recommends to prepare fresh each week.
- 1X PBS pH 7.4: Prepare 1 L of 1X PBS by combining 100 mL of 10X PBS and 900 mL of DEPC-treated water. Do not reuse.

- Proteinase K: Default concentration (1 μg/mL) is made by adding 10 μL of 20 mg/mL Proteinase K to 200 mL of 1X PBS made with DEPC-treated water. Alternative Proteinase K concentrations are decided empirically by tissue type and range between 0.1 mg/ml and 1 mg/mL.
- NBF stop buffer: Add 12.12 g Tris base and 7.5 g glycine to 1 L DEPC-treated water to yield 0.1 M tris, 0.1 M glycine. Store at RT up to 1 month.
- 2X SSC: Prepare 1 L of 2X SSC by combining 100 mL of 20X SSC and 900 mL of DEPC-treated water. Do not reuse. 2X SSC-T (optional) Prepare 250 mL of 2X SSC-T by combining 25 mL of 20X SSC, 2.5 mL of 10% Tween 20, and 222.5 mL of DEPC-treated water. Do not reuse.
- 4X SSC: Prepare 1 L of 4X SSC by combining 200 mL of 20X SSC and 800 mL of DEPC-treated water. Do not reuse.
- 1X Tris-EDTA pH (Antigen Retrieval Solution): Prepare 1 L of 1X Tris–EDTA pH 9.0 by combining 100 mL of 10X Tris–EDTA pH 9.0 and 900 mL of DEPC-treated water.

Day 1

3.3 Slide Preparation

3.3.1 Sectioning and Mounting

Tissue sections should be 5 μm thick and mounted on Superfrost Plus or BOND Plus slides. Tissue sections must be placed in the Scan Area in the center of the slide and be no larger than 35.3 mm long by 14.1 mm wide. Adhesive slide labels should not overlap Scan Area or slide gasket (Fig. 2).

3.3.2 Baking

Bake slides with mounted sections in a 60 °C drying oven for 30 min to 3 h prior to deparaffinization. Stand slides vertically during baking to allow excess paraffin to flow off. Longer baking times may be necessary for some tissues to sufficiently adhere to the slide (e.g., overnight at 37 °C followed by 2–4 h at 60 °C); this should be empirically tested.

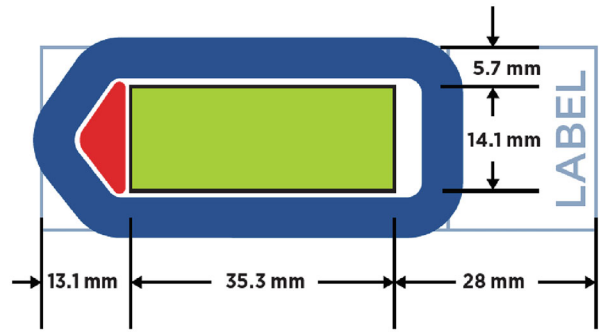


Fig. 2 Slide dimensions. Measure from the label edge of the slide as reference point. Place sections within the green highlighted area

- If your tissue type is fixed frozen, baking might not be necessary in all cases. You can still proceed to bake the slide if your tissue is not well adhered to the slide.
- If your tissue type is fresh frozen, keep your slides in 10% NBF solution overnight, wash with 1x PBS and then proceed to baking.

(Place two coplin jars- 1) 1x Antigen retrieval solution/Tris–EDTA; 2) DEPC-treated water in a steamer switched on to reach 100°C.)

3.3.3 Deparaffinize and Rehydrate FFPE Tissue Sections

- Immediately after baking, wash slides in xylene, three times for 5 min each, followed by three Ethanol washes for 5 min each, in the order: (1) 100% ethanol, (2) 100% ethanol, (3) 95% ethanol. Wash with 1x PBS for 5 min. (This is a stop point; slides can be stored in PBS for about 1 h.)
- Skip the Xylene washes and give three ethanol washes, in the order: (1) 50% ethanol, (2) 70% ethanol, (3) 95% ethanol, if your tissue type is fixed frozen or fresh frozen.

3.3.4 Antigen Retrieval

- Dip the slides in DEPC-treated water, proceed to transfer the slides to 1x Tris–EDTA—20 min (empirically tested temperatures for different tissue types).
- Proteinase K—15 min at 37 °C in a water bath.
- PBS wash—5 min.
- For fresh frozen or fixed frozen tissues—performing antigen retrieval at 85 °C might be ideal.

3.3.5 Tissue Fixation

- Immerse the slides in 10% NBF for 5 min followed by two 5 min washes in NBF stop buffer.
- Wash the slide with 1x PBS for 5 min.
- If your tissue type is fixed frozen, skip tissue fixation and directly proceed to apply probes.

3.3.6 Probing and Incubation

- Thaw probe stored in −20 to room temperature, flip mix and briefly centrifuge. Bring Buffer R to room temperature.
- Clean the hybridization chamber with RNase AWAY, followed by DEPC-treated water—take fresh lint-free wipes and layer them on the bottom of the chamber—wet them with 2x SSC or DEPC-treated water.
- Prepare probe mix, according to Table 1.
- Remove the slides one at a time, drain the excess fluid from the slide and place them In Hybridization chamber.

Table 1
Table for preparation of probe mix -where *n* represents the number of reactions

Buffer R	Panel	DEPC-treated water	Total volume
200 µL × *n*	25 µL × *n*	25 µL × *n*	250 µL × *n*

- Carefully apply 250 µL of probe mix onto the slide, and cover with HybriSlip.
- Incubate at 37°C overnight. (16–24 h).
- **Day 2**

3.3.7 Post Hybridization-Stringent washes

Place stringent wash solution jars in a water bath and bring it to 37°C.

- After 16–24 h incubation, remove the slides from the hyb chamber and carefully dip the slides in 2x SSC to remove the HybriSlip, the slides can be kept for a minute in 2x SSC or 2X SSC—Tween can be used if the HybriSlip does not come off the slide.
- Transfer the slides into stringent wash solution, make sure the solution is at 37 °C. Perform two washes with Stringent wash solution 25 min each.
- Wash the slides two times in 2X SSC for 2 min each.

3.3.8 Morphology Marker Addition

Thoroughly clean a humidity chamber with RNase AWAY and DEPC-treated water.

Drain excess liquid from slides by placing them vertically on lint-free wipes and the place it horizontally in a humidity chamber.

Carefully apply 250 µL of Buffer W onto the slide—make sure to cover the entire scan area where tissue is placed.

Close the humidity chamber and incubate for 30 min.

Remove nuclei marker SYTO13 from the freezer and thaw it to room temperature.

Prepare a morphology marker mix according to Table 2.

** Dilutions for additional markers should be empirically tested—the total volume of the marker mix should be made up to 220 µL with Buffer W.

Apply 200 µL of morphology marker mix to each slide and incubate for 1 h. Additional incubation times might be applicable depending on the antibodies being used and tissue type.

After 1 h, wash the slides two times in 2X SSC for 5 min each.

This is a stop point, Slides can be stored in 2x SSC at 4 °C for up to 21 days, before proceeding for collection.

Table 2
Table for preparation of morphology master mix- where *n* represents the number of reactions

Nuclear stain (SYTO 13)[a]	Morphology marker 1(PanCK)	Morphology marker 2(CD45)	Additional markers[b]	Buffer W	Total volume
22 μL × *n*	5.5 μL × *n*	5.5 μL × *n*	X μL × *n*	(up to 220 μL) × *n*	220 μL × *n*

[a]SYTO 13 occupies FITC channel, if FITC channel is required for other morphology marker used in the assay, SYTO 83 can be used as substitute to stain nuclei
[b]Any additional conjugated antibody, in interest of the experiment can be used. Volumes to be tested emperically

3.4 ROI Selection and Collection

- For RNA analysis, the area of an ROI should generally be at least 30,000 μm^2 (approximately equal to a square with side length 170 μm or a circle with diameter 200 μm). ROIs that will be segmented should be larger. As a general guideline, the minimum number of cells per ROI in an RNA assay (or, if segmenting, per segment) is 100 for NGS readout.
- Maximum ROI dimensions are 660 μm × 785 μm. A maximum of 380 ROIs can be placed on a single scan.
- In the scan workspace there is a sidebar toward the upper left corner of the screen where you can make adjustments to your scan.
- Go to render settings to adjust the intensity of each channel to visualize the tissue.

  Note: If the binding sensitivity of your morphological marker is low, a grid-like background may be observed on decreasing the channel threshold. Set the threshold of the channel to maximum for a better view of the scan.

  The binding and intensity of morphological markers will not affect the RNA assay.
- GeoMx allows you to select standard geometric, polygonal and grid ROIs. Polygonal ROIs are close to placing a freehand ROI. Uploading an H&E image through overlay tool will help you to select ROIs based on the H&E image on the same slide.
- Post ROI/AOI selection libraries are prepared as per the instruction manual, pooled and sequenced following next-generation sequencing approach to get sufficient data. Post that bioinformatic analysis would be performed, which is explained in-depth in the following section.

Bioinformatics

4 Bioinformatics Software and Tools for Analyzing GeoMx DSP NGS Readouts

Materials

1. Windows/Mac/Linux -based machine with at least 4–16 GB of RAM and 200 GB of disk space (*see* **Note 1**).
2. FastQC, preferably belonging to latest version.
3. GeoMx NGS Pipeline software (*see* **Note 2**).
4. R programming language preferably belonging to latest version or atleast R version (> = 3.6) (https://cran.rstudio.com/) and RStudio (https://posit.co/download/rstudio-desktop/).
5. NanoString specific R packages GeomxTools, SpatialDecon and SpatialOmicsOverlay (latest version) all accessible through Bioconductor repository (*see* **Note 3**).
6. GeoMx DSP probe kit configuration (.pkc) file for the particular GeoMx panel used in the assay, which is a probe assay metadata describing the gene targets present in the data.
7. Annotation file in Microsoft excel format containing metadata for each ROI selected within the experiment. Other than the mandatory ROI ID in the first column, the annotation file should contain other useful information like tissue morphology, segment area, nuclei count per segment, characteristics (ex: test vs. control).

5 Step-by-Step guidelines for Analyzing GeoMx DSP Data from NGS Readouts

Methods

1. For each ROIs or segments present in a GeoMx experiment, individual NGS raw data in the format of .FASTQ would be provided. If segmentation have been performed for a ROI, the particular ROI would have two separate FASTQ files representing each segment or area of illumination (AOI). It is recommended that the raw data in FASTQ state should undergo a quality check using FastQC which can give an idea about the Phred score distribution, PCR duplicates, adapter contamination, and other sequencing-specific quality-related statistics.
2. GeoMx NGS Pipeline software after successful installation, following the instructions specific to the OS being used can be made used to convert .FASTQ files into GeoMx DSP specific data format (.DCC) or Digital Count Conversion format. We have to provide the path of the folder having .FASTQ files in .gz format per segment and path to the configuration file from the GeoMx DSP in .ini format, again downloadable from the GeoMx DSP system. The software would generate a zip file

containing all the DCC files per segment along with other statistics like number of reads present, trimmed, mapped and those which underwent deduplication, per segment.

3. Three main input files are required for analyzing GeoMx DSP NGS readout data using GeoMxTools R package [5]; the GeoMx raw count files (.DCC), ROI annotation containing metadata of the segments (*see* **Note 4**) and the PKC file containing probe annotation.

   (a) After loading the required R packages and above mentioned 3 input files, the first step as in any next generation sequencing assay is the quality check of both segments (ROIs) and gene features. Segments get filtered out on the basis of major parameters like Number of raw, trimmed, aligned, stitched, and deduplicated sequenced reads, percentage of sequencing saturation, geometric mean of negative control probes per segment, count of reads quantified from the no template control (NTC) used in the assay, number of nuclei and area of the ROI selected.

   In addition to these, parameters like gene detection rate also comes into picture where it has to be made sure that the segments are able to express genes above a preferable threshold (usually 5–10%) of the genes in the panel (*see* **Note 5**). The target gene QC happens at the probe level, where in the Whole Transcriptome Atlas (WTA) panel individual probes represent a single gene while in the Cancer Transcriptome Atlas (CTA) panel multiple probes represent a gene. Probes get filtered out in terms of their counts and if they act as an outlier, they are detected through Grubb's test.

   A1. One of the most important QC parameter calculated for both segment and genes is the Limit Of Quantification (LOQ), which is defined as the minimum quantity of gene expression quantifiable from a segment. Both the segments as well as genes, on the basis of the LOQ gets filtered out where the segments, are not able to express at least 5–10% of the genes in the panel above LOQ or for a gene if it is not found expressed above LOQ in at least 5–10% of the segments present in the study.

   (b) After selecting the final set of segments and genes to be used for further analysis, normalization either based on Quartile 3 (Q3) method or on negative control probe counts (background) is applied. The normalized data can be visualized using multiple dimensional reduction plots like PCA, UMAP, and tSNE, to understand the grouping of the segments.

(c) Multiple differential expression test methods can be used for analyzing GeoMx DSP data while considering the fact that from a single tissue multiple ROIs (subsampling) can be selected which is not equal to biological replicates in a statistical test. The linear mixed effect model applied through the lme4 package within the GeoMxTools is the most common approach (*see* **Note 6**). The results of differential expression can be visualized using the usual representations like heatmaps, volcano plots, or MA plots.

(d) The downstream analysis of the differentially expressed genes is as any other bulk gene expression-based assay, which can include Gene Ontology and Pathway Over Representation Analysis (ORA) or Gene Set Enrichment Analysis (GSEA).

4. One of the most important applications of single-cell RNA-Seq assays is in identifying marker genes to represent different cell types for different conditions and sample types. In Bulk RNA-Seq experiments, we use the single-cell gene expression signature matrix to be merged with the bulk gene expression matrix, in order to deconvolute the absolute abundance of cell types within each sample. A similar approach but within each of the spatial segments profiled within GeoMx DSP, is possible with the use of SpatialDecon R package [6] (*see* **Note 7**).
5. Integrating back gene expression or any quantitative scores like pathway scores back to the spatial location is an outcome desirable after conducting a spatial transcriptomics experiment. Plotting overlays with data like gene expression over the GeoMx Scan where the ROIs have been selected (Fig. 3), is made possible through SpatialOmicsOverlay R package [7]. Inputs required to perform this include the GeoMx scan image in OME-TIFF format and Lab Worksheet exported from the GeoMx DSP instrument (*see* **Note 8**).

6 Notes

1. While selecting the core size for preparation of TMA, carefully select areas to be assayed and check if the core size covers the heterogeneity present in the tissue, selection of more cores from the same tissue might be necessary in case of highly heterogeneous tissue.
2. If using GFP as a marker in the study, using an anti GFP morphology marker will be ideal for the study, as the GFP signal gets quenched during slide preparation.
3. To avoid cross contamination, frequently use RNase AWAY to clean benchtop and change gloves frequently.

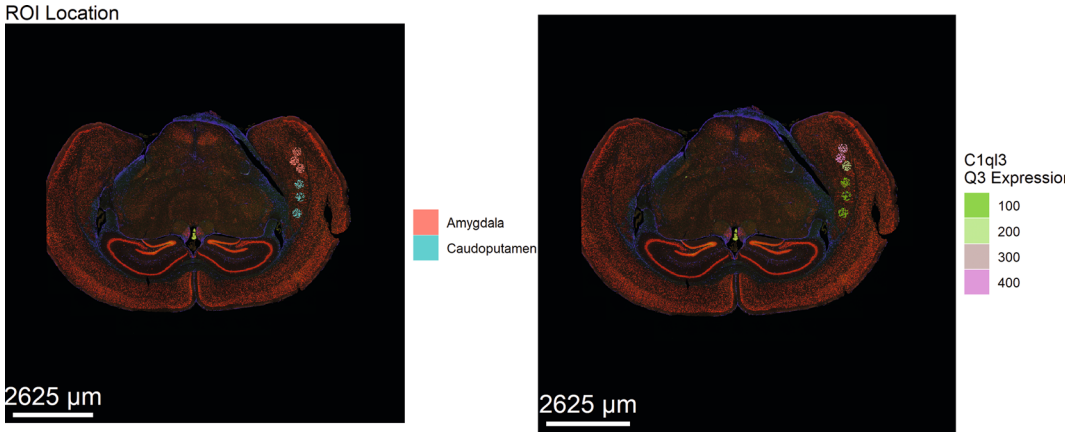


Fig. 3 Overlaying the expression of a differentially expressed gene (C1ql3), back to the ROI collected within the GeoMx® DSP scan. Expression of gene compared between two anatomical locations, Amygdala and Caudoputamen shows that C1ql3 is enriched within Amygdala compared to Caudoputamen. (The image has been generated using plotSpatialOverlay() function from SpatialOmicsOverlay R package. The data set in this plot was downloaded from Nanostring's Spatial Organ Atlas website (https:/nanostring.com/products/geomx-digital-spatial-profiler/spatial-organ-atlas/mouse-brain/), taken from Mouse brain samples)

4. FASTQ file size would be proportional to the total number of reads which in turn would depend on multiple factors like ROI area, number of nuclei within the ROI and type of genes being looked into (rare expressors or high expressors). The disk space within the machine used for data analysis hence depends upon the total number of ROIs chosen in a GeoMx DSP NGS experiment along with properties related to the ROIs.
5. GeoMx NGS Pipeline software is NanoString's standalone software which is also accessible through Illumina's DRAGEN™ platform accessed via BaseSpace™ Sequence Hub (cloud) or NextSeq 1000/2000 (local). Users analyzing their data in local system can get access to the executable file for installing the software by contacting the NanoString support team directly.
6. Other than GeoMxTools there are more packages in R which can be used for analyzing GeoMx DSP NGS readout data, like standR [8], which is a statistical package that uses Limma or EdgeR methods for statistical analysis while GeoMxTools uses a linear mixed-effect model (LMM) to handle multiple ROIs selected from the same tissue core.
7. It is mandatory that the segment annotation file contains the first column with a column title which matches the input arguments within GeoMxTools R package while importing the annotation file.

8. The thresholds for each QC parameter have been defined elsewhere [5] but is a decision taken majorly on the basis of the type of the study and prior knowledge on the tissues used for the GeoMx spatial transcriptomics assay. For example, if in case most of the segments are found with gene detection rate in the range less than 10% and at the same time we know the segments contain less number of nuclei, they correlate with one another and do not require any filtration.
9. The formula within the LMM model to perform differential expression tests has to be devised with care, since we need to account for multiple variables in an experiment like comparing segments from multiple tissues which are mutually exclusive and comparing segments from the same tissue. The addition of random slope and random intercept depends on these conditions.
10. SpatialDecon R package by default provides access to profile the abundance of 14–18 Immune cell types with the use of default single cell expression matrix, "safeTME." But it is best advised to use the background single cell expression matrix according to the tissue type, disease condition, organism and the type of cells interested in deconvoluting. NanoString provides a wider array of options to choose for the single cell expression matrix to be used for deconvoluting cell types in their ROIs, from the CellProfileLibrary Github repository.
11. SpatialOmicsOverlay R package, requires the input Lab Worksheet file where it is mandatory that the file name should contain the keyword "LabWorksheet." If the GeoMx experiment was carried out in a Tissue Microarray (TMA) model, where a single slide contains multiple tissues from which multiple ROIs are selected, then it is possible to crop the image visualizing the selected tissue core within the GeoMx scan.

References

1. Jawhar NM (2009) Tissue microarray: a rapidly evolving diagnostic and research tool. Ann Saudi Med 29(2)spiepr A3B2 twbch ":":spiepr A3B2 twbch123–127. https://doi.org/10.4103/0256-4947.51806. PMID: 19318744; PMCID: PMC2813639
2. Zollinger DR, Lingle SE, Sorg K, Beechem JM, Merritt CR (2020) GeoMx™ RNA assay: high multiplex, digital, spatial analysis of RNA in FFPE tissue. In: Nielsen BS, Jones J (eds) Situ hybridization protocols. Methods in molecular biology, vol 2148. Humana, New York. https://doi.org/10.1007/978-1-0716-0623-0_21
3. Wang N, Li X, Ding Z (2023) High-Plex spatial profiling of RNA and protein using digital spatial profiler. Methods Mol Biol spiepr A3B2 twbch ":"2660:69spiepr A3B2 twbch–83. https://doi.org/10.1007/978-1-0716-3163-8_6
4. Ospina O, Soupir A, Fridley BL (2023) A primer on preprocessing, visualization, clustering, and phenotyping of barcode-based spatial Transcriptomics data. Methods Mol Biol spiepr A3B2 twbch ":"2629:115spiepr A3B2 twbch–140. https://doi.org/10.1007/978-1-0716-2986-4_7
5. Griswold M, Ortogero N, Yang Z, Vitancol R, Henderson D (2024) GeomxTools: NanoString GeoMx tools. R package version 3.8.0. https://bioconductor.org/packages/devel/workflows/vignettes/GeoMxWorkflows/inst/doc/

GeomxTools_RNA-NGS_Analysis.html#Analyzing_GeoMx-NGS_RNA_Expression_Data_with_GeomxTools

6. Danaher P, Kim Y, Nelson B et al (2022) Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. Nat Commun spiepr A3B2 twbch ":"13:385spiepr A3B2 twbch. https://doi.org/10.1038/s41467-022-28020-5
7. Griswold M (2024) SpatialOmicsOverlay: spatial overlay for Omic data from Nanostring GeoMx data. R package version 1.4.0
8. Liu N, Bhuva DD, Mohamed A, Bokelund M, Kulasinghe A, Tan CW, Davis MJ (2024) standR: spatial transcriptomic analysis for GeoMx DSP data. Nucleic Acids Res 52(1):e2. https://doi.org/10.1093/nar/gkad1026. PMID: 37953397; PMCID: PMC10783521