

# Industrialized, Artificial Intelligence-guided Laser Microdissection for Microscaled Proteomic Analysis of the Tumor Microenvironment

Dave Mitchell<sup>\*1,2</sup>, Allison L. Hunt<sup>\*1,4</sup>, Kelly A. Conrads<sup>1,2</sup>, Brian L. Hood<sup>1,2</sup>, Sasha C. Makohon-Moore<sup>1,2</sup>, Christine Rojas<sup>1</sup>, G. Larry Maxwell<sup>1,3,4</sup>, Nicholas W. Bateman<sup>1,2,3</sup>, Thomas P. Conrads<sup>1,3,4</sup>

<sup>1</sup>Women's Health Integrated Research Center, Gynecologic Cancer Center of Excellence, Department of Obstetrics and Gynecology, Uniformed Services University and Walter Reed National Military Medical Center <sup>2</sup>The Henry M. Jackson Foundation for the Advancement of Military Medicine <sup>3</sup>The John P. Murtha Cancer Center Research Program, Department of Surgery, Uniformed Services University <sup>4</sup>Women's Health Integrated Research Center, Inova Women's Service Line, Inova Health System

\*These authors contributed equally

## Corresponding Authors

Nicholas W. Bateman

batemann@whirc.org

Thomas P. Conrads

conrads@whirc.org

## Citation

Mitchell, D., Hunt, A.L., Conrads, K.A., Hood, B.L., Makohon-Moore, S.C., Rojas, C., Maxwell, G.L., Bateman, N.W., Conrads, T.P. Industrialized, Artificial Intelligence-guided Laser Microdissection for Microscaled Proteomic Analysis of the Tumor Microenvironment. *J. Vis. Exp.* (184), e64171, doi:10.3791/64171 (2022).

## Date Published

June 3, 2022

## DOI

10.3791/64171

## URL

jove.com/video/64171

## Introduction

## Abstract

The tumor microenvironment (TME) represents a complex ecosystem comprised of dozens of distinct cell types, including tumor, stroma, and immune cell populations. To characterize proteome-level variation and tumor heterogeneity at scale, high-throughput methods are needed to selectively isolate discrete cellular populations in solid tumor malignancies. This protocol describes a high-throughput workflow, enabled by artificial intelligence (AI), that segments images of hematoxylin and eosin (H&E)-stained, thin tissue sections into pathology-confirmed regions of interest for selective harvest of histology-resolved cell populations using laser microdissection (LMD). This strategy includes a novel algorithm enabling the transfer of regions denoting cell populations of interest, annotated using digital image software, directly to laser microscopes, thus enabling more facile collections. Successful implementation of this workflow was performed, demonstrating the utility of this harmonized method to selectively harvest tumor cell populations from the TME for quantitative, multiplexed proteomic analysis by high-resolution mass spectrometry. This strategy fully integrates with routine histopathology review, leveraging digital image analysis to support enrichment of cellular populations of interest and is fully generalizable, enabling harmonized harvests of cell populations from the TME for multiomic analyses.

The TME represents a complex ecosystem populated by a highly diverse array of cell types, such as tumor cells, stromal cells, immune cells, endothelial cells, other mesenchymal cell types, and adipocytes, along with a complex extracellular matrix<sup>1</sup>. This cellular ecosystem varies within and across different disease organ sites, resulting in complex tumor heterogeneity<sup>2,3</sup>. Recent studies have shown that heterogeneous tumors and tumors with low tumor cellularity (low purity) often correlate with poor disease prognosis<sup>2,3</sup>.

To understand the molecular interplay between tumor and non-tumor cell populations within the TME at scale, standardized and high-throughput strategies are necessary to selectively harvest distinct cellular populations of interest for downstream multiomic analysis. Quantitative proteomics represents a rapidly evolving and increasingly important technique to further the understanding of cancer biology. To date, the preponderance of studies employing proteomics have done so with proteins extracted from whole tumor tissue preparations (e.g., cryopulverized), leading to a paucity in the understanding of proteome-level heterogeneity in the TME<sup>4,5,6</sup>.

The development of sample collection strategies that seamlessly integrate with and harness information from clinical pathology workflows will enable a new generation of histology-resolved proteomics that are highly complementary to gold-standard, diagnostic pathology workflows. LMD enables direct and selective collection of cellular subpopulations or regions of interest (ROIs) through microscopic inspection of histologically stained tissue thin sections<sup>7</sup>. Recent major advances in digital pathology and AI-enabled analysis have demonstrated the ability to identify unique compositional features and ROIs within the TME in an

automated fashion, many of which correlate with molecular alterations and clinical disease features, such as resistance to therapy and disease prognosis<sup>8</sup>.

The workflow described in the protocol presented here leverages commercial software solutions to selectively annotate tumor ROIs within digital histopathology images, and utilizes software tools developed in-house to transfer these tumor ROIs to laser microscopes for automated collection of discrete cellular populations of interest that seamlessly integrates with downstream multiomic analysis workflows. This integrated strategy significantly diminishes LMD operator time and minimizes the duration for which tissues are required to be at ambient temperature. The integration of automated feature selection and LMD harvest with high-throughput quantitative proteomics is demonstrated through a differential analysis of the TME from two representative epithelial ovarian cancer histologic subtypes, high-grade serous ovarian cancer (HGSOC) and ovarian clear cell carcinoma (OCCC).

## Protocol

All study protocols were approved for use under a Western IRB-approved protocol "An Integrated Molecular Analysis of Endometrial and Ovarian Cancer to Identify and Validate Clinically Informative Biomarkers" deemed exempt under US Federal regulation 45 CFR 46.102(f). All experimental protocols involving human data in this study were in accordance with the Declaration of Helsinki. Informed consent was obtained from all subjects involved in the study.

**CAUTION:** The following reagents used throughout the protocol are known or suspected carcinogens and/or contain hazardous materials: ethanol, DEPC water, Mayer's Hematoxylin solution, Eosin Y solution, methanol, acetonitrile,

and formic acid. Proper handling, as described in the respective safety data sheets (SDS), and use of appropriate personal protective equipment (PPE) is mandatory.

## 1. Generating the default shape list data (.sld) file containing calibrator fiducials

**NOTE:** The protocol steps described in this section are specific to use with an inverted laser microscope and the associated software (see the **Table of Materials**). Creation of a default .sld file is only necessary once per laser microscope. The resultant file can be used for cutting fiducials into all PEN slides used thereafter. Approximate time: 5 min (once only).

1. Open the LMD software and load the polyethylene naphthalate (PEN) membrane slide onto the LMD stage face down, with the label nearest to the user. Unselect the **Close line(s)** box on the right side of the program window.
2. Use the **PtoP** (point to point) function under high magnification (63x) to draw three "V" arrows to serve as calibration fiducials. Starting at one external point on the V, draw a line to the mid-point of the V and single click. Then, draw a second line from the central point of the V to the end of the second external V point, and double-click to create a single, unclosed V shape from the two lines.  
**NOTE:** These calibration fiducials should be placed in three corners of the slide: front right, rear right, rear left.
3. Select the **AF (autofocus) before cut** option. Cut the slide in position 1, move it into each of the remaining slide positions, and trace precisely over the calibration cuts.
4. Save the .sld file and select the **Save without calibration** option from the pop-up dialog box to avoid cutting the calibration fiducials into the membrane.

**NOTE:** A representative .sld file containing standard calibrator fiducials for four slide positions is provided in **Supplemental File 1**.

## 2. LMD slide(s) preparation

**NOTE:** The protocol steps described in this section are specific to use with an inverted laser microscope and the associated software (see the **Table of Materials**). Approximate time: 5 min.

1. Ensure that the slide is completely dry before cutting the reference calibration fiducials. Open the LMD software and open the default calibration .sld file under the **Import shapes** option.
2. Select the **AF (autofocus) before cut** option. Load the slide(s) with the tissue facing down and the label slide nearer to the operator into the slide holder on the LMD stage.
3. Using the laser microscope and default calibration .sld file, cut calibration fiducials into the PEN membrane.
  1. **OPTIONAL:** Cut the calibration fiducials into the PEN membrane either before or after the tissue section(s) is/are placed on the slide. If the calibration fiducials are cut prior to tissue placement, ensure that the tissue and/or fixative does not overlap with the calibrators when the tissue is placed on the slide in Step 2.5. If calibration fiducials are cut after tissue placement, stop after the completion of Step 2.4 and proceed to section 3.
4. Review all calibrators individually to ensure that each cut is complete and visible.  
**NOTE:** Use the **Move and Cut** feature to manually direct the laser over any calibration fiducials that did not fully cut through the PEN membrane.

- Place the frozen or formalin-fixed, paraffin-embedded (FFPE) tissue section onto the slide containing the calibration fiducials.

### 3. Tissue staining

**NOTE:** Approximate time: 30 min.

- Fix the frozen LMD tissue slides in 70% ethanol (EtOH) containing phosphatase inhibitor cocktail reagents for 5 min.
- Wash the slides in diethyl pyrocarbomate (DEPC) water containing phosphatase inhibitor cocktail reagents for 1 min.
- Wash the slides in DEPC water for 1 min.
- Incubate the slides in Mayer's Hematoxylin solution for 3 min.
- Rinse the slides in DEPC water for 3 min.
- Rinse the slides in a fresh exchange of DEPC water for 1 min.
- Incubate the slides in aqueous Eosin Y solution for 1 s.
- Rinse the slides 2 x 5 s in 95% EtOH.
- Rinse the slides 3 x 10 s in 100% EtOH.
- Wipe the excess EtOH from the back of the slides and allow the slides to air-dry.
- Store the slides at -80 °C if LMD is not to be performed immediately.

### 4. Slide imaging

**NOTE:** The protocol steps described in this section are specific for slides scanned (see the **Table of Materials**) and the resultant images saved as .svs files. Use any scanner and its associated software that generate image files in a format

that the image analysis software (see the **Table of Materials**) can open. File types using pyramidal tiffs that are supported include JPG, TIF, MRXS, QTIFF, component TIFF, SVS, AFI, SCN, LIF, DCM, OME.TIFF, ND2, VSI, NDPI, NDPIS, CZI, BIF, KFB, and ISYNTAX. Approximate time: 5 min.

- Turn on the scanner and open the slide scanner software. Load the slide with the tissue facing up onto the single slide stage in the scanner. Ensure the slide is completely dry and gently place a coverslip over the tissue. Do not use ethanol or immersion oil under the coverslip.
- Capture the micrograph image using settings calibrated to adjust for the PEN membrane instead of glass background and to ignore background membrane staining, per manufacturer's instructions.
- Adjust the imaging area by dragging and resizing the inner green perimeter to capture the entire PEN membrane area, as needed. Add four focus points onto the tissue by double-clicking onto the snapshot overview image, and three focus points on the membrane near the calibration fiducials (one focus point per each of the three calibration fiducials).
 

**NOTE:** The four focus points can be placed nearly anywhere on the tissue section, though placing onto tissue that is too darkly stained and appears black may cause the scan to fail.
- Under the **View** menu, select **Video Monitor**. Manually adjust the focus using the fine and/or macro focus slider, as needed, for each point around the LMD tissue. Capture the image scan under high magnification (20x). Confirm that all calibration fiducials are visible and clear in the saved image.

## 5. Automated feature selection using image analysis software

1. For whole tumor collections (approximate time: 5 min; case-dependent):

1. Open the image analysis software (see the **Table of Materials**). Select **Open Images** and from the pop-up window, select the .svs image file generated from scanning the slide on the AT2 scanner.

**NOTE:** A representative .svs image file is provided in **Supplemental File 2**.

2. Navigate to the **Annotations** tab. Select the **Pen** tool on the Annotation toolbar and draw a shape around the tissue.

3. Select the shape and right-click on the image. From the **Advanced** dropdown menu, select **Partitioning (Tiled)**. Set the **Tile Size** and **Space Between** to **500** and **40**, respectively, and select **OK** to generate the tiles. Select and delete the perimeter shape used to generate the tiles in step 5.1.2.

4. Select the **Layer Actions** dropdown menu | **Export** to save the tiled annotations as a .annotation file.

**NOTE:** A representative .annotation file for a whole tumor tissue collection is provided in **Supplemental File 3**.

5. Create a folder for the session or project, and save the .annotation file inside a subfolder labeled with the unique identifier for the slide.

6. Navigate to the **Annotations** tab. Select the **Layer Actions** dropdown | **Delete All Layers** to remove all annotations from the image. Select the **pen** tool and draw a short line from the inner tip of the arrowhead for each calibration fiducial. Draw the lines from the

marks in the following order: top left, top right, bottom right.

7. Select the **Layer Actions** dropdown menu | **Export** to save the line annotations as a .annotation file. Add **\_calib** to the filename and place the file in the subfolder that contains the coordinates for the tiled shapes.

**NOTE:** A representative \_calib.annotation file is provided in **Supplemental File 4**.

8. Copy the address for the main project or session folder. Open the XML import generating script, "Malleator" (available via <https://github.com/GYNCOE/Mitchell.et.al.2022>), using the IDLE integrated development environment and paste the project folder address between the quotes at the bottom of the script.

9. Select the **Run** dropdown menu | **Run Module** to execute the script.

**NOTE:** The .xml LMD import file will be generated inside the subfolder created for the image/slide. A representative .xml file is provided in **Supplemental File 5**.

2. For LMD-enriched collections only (approximate time: 15 min; case-dependent):

1. Open the image analysis software (see the **Table of Materials**). Select **Open Images** and from the pop-up window, select the .svs image file generated from scanning the slide.

2. Navigate to the **Annotations** tab. Select and use the **Rectangle** annotation tool to draw a box around the tissue.

3. Select the box annotation and right-click on the image. Select the **Advanced** dropdown menu |

- Partitioning (Tiled)** option. Set the **Tile Size** and **Space Between** to **500** and **40**, respectively, and select **OK** to generate the tiles. Select and delete the perimeter box annotation used to generate the tiles in step 5.2.2.
4. Select the **Layer Actions** dropdown menu | **Export** to save the tiled annotations as a .annotation file.
  5. Place a saved copy of the Python "Dapq" (available via <https://github.com/GYNCOE/Mitchell.et.al.2022>) algorithm, developed to merge the AI-classified annotation layers, into the same folder as the tiled annotations file.
  6. Copy the name of the tiled annotation file. Open the Python program using the IDLE integrated development environment and paste the name of the tiled annotation file between the quotations at the bottom of the program.
  7. Select the **Run** dropdown menu | **Run Module**. Wait for a new file to be generated that will have all the tiled annotations merged under a single layer.
  8. Open the image analysis software and navigate to the **Annotations** tab. Select the **Layer Actions** dropdown menu | **Delete All Layers** to remove all annotations from the image.
  9. Select the **Layer Actions** dropdown menu | **Import Local Annotation File**. In the pop-up window, select the merged .annotation file that was generated by the script. Ensure that all imported tiles are under the same annotation layer.
  10. Navigate to the **Classifier** tab and follow the manufacturer's instructions to generate shapes for the ROIs. Before running the classifier, select the desired annotation layer(s) (i.e., the tumor or stroma layer) by checking the ROI box or boxes on the **Annotations** tab, under the **Advanced Classifier Options**. Use the **Annotation Layer** option from the **Classifier Actions** menu to run the classifier.
  11. Once the Classifier analysis is complete, navigate to the **Annotations** tab and select the annotation layer generated from the analysis. Select the **Layer Actions** dropdown menu | **Delete All Layers But Current** to remove all other annotation layers from the image.
  12. Select the **Layer Actions** dropdown menu | **Export** to save the annotations as a .annotation file. Create a folder for the session or project and save the .annotation file inside a subfolder labeled with the unique identifier for the slide.  
**NOTE:** A representative .annotation file for a classified LMD enriched tissue collection is provided in **Supplemental File 6**.
  13. Navigate to the **Annotations** tab, select the **Layer Actions** dropdown | **Delete All Layers** to remove all annotations from the image. Select the **pen** tool and draw a short line from each calibration fiducial. Draw lines from the marks in the following order: top left, top right, bottom right.
  14. Select the **Layer Actions** dropdown menu | **Export** to save the line annotations as a .annotation file. Add **\_calib** to the filename and place the file in the subfolder that contains the coordinates for the tiled shapes.
  15. Copy the address for the main project or session folder. Open the XML import generating script, "Malleator" (available via <https://github.com/GYNCOE/Mitchell.et.al.2022>), using the IDLE

integrated development environment, then paste the project folder address between the quotes at the bottom of the script.

16. Select the **Run** dropdown menu | **Run Module** to execute the script.

**NOTE:** The .xml LMD import file will generate inside of the subfolder created for the image/slide.

## 6. Laser microdissection

**NOTE:** The protocol steps described in this section are specific to use with an inverted laser microscope and the associated software (see the **Table of Materials**). Approximate time: 2 h; case dependent.

1. Load the marked membrane slide (containing calibration fiducials) with the tissue facing down and the label side nearer to the operator into the slide holder on the laser microscope stage.
2. Select **Import Shapes** from the **File** dropdown menu. Select the **.xml LMD import file** generated for the slide. Select **No** in the pop-up window to avoid loading reference points from file and **No** in the second pop-up window to avoid using any previously stored reference points for calibration.
3. Follow the prompts from the LMD application and align the calibration cross to each of the three calibration fiducials on the slide. Look for calibration fiducials that appear in the top left, top right, and bottom right of the slide image in the image analysis software that will correspond to reference points in the front right, rear right, and rear left corners, respectively, of the inverted LMD slide on the microscope stage. Switch between using the 5x objective lens to locate and the 63x objective to align each calibration fiducial. Select **No** on the pop-

up window to avoid saving the reference points to file and **OK** in the second pop-up window to confirm that the slide is inserted.

4. Move the 5x objective lens into position and select **Yes** in the pop-up window to use the actual magnification. Once the imported shapes appear, focus the camera onto the tissue.
5. Highlight and select all the shapes in the **Shapes List** window, drag them into place using one or two annotations within the field of view as references, and align the vertical z-axis for cutting with the laser.
6. Review the imported shapes and assign them to the appropriate tube position for collection. Press **Start Cut** to start the laser.

**NOTE:** Imported shapes in the .xml file will be automatically assigned to the "no cap" position in the **Shapes List** window. To harvest tissue, the imported shapes must be reassigned to a position containing a loaded tube.

## 7. Protein digestion by pressure-cycle technology (PCT)

**NOTE:** Approximate time: 4 h (3 h without vacuum centrifuge drying time).

1. Place 0.5 mL tubes containing the capped PCT MicroTubes containing LMD harvested tissue in 20  $\mu$ L of 100 mM TEAB/10% acetonitrile into a thermocycler and heat at 99  $^{\circ}$ C for 30 min, then cool to 50  $^{\circ}$ C for 10 min.
2. Spin the tubes for 30 s at 4,000  $\times$  g and then remove the MicroTubes from the 0.5 mL tubes. Using the MicroCap tool, remove and discard the MicroCaps from the PCT MicroTubes. Add trypsin (see the **Table of Materials**) at

a ratio of 1  $\mu\text{g}$  per 30  $\text{mm}^2$  tissue and insert a MicroPestle into the MicroTube using the MicroCap tool.

3. Transfer the MicroTubes into a barocycler cartridge and assemble the complete cartridge. Place the cartridge into the barocycler pressure chamber and secure the lid. Barocycle at 45,000 psi for 50 s and atmospheric pressure for 10 s at 50 °C for 60 cycles.
4. Once barocycling is complete, transfer the microtubes to a 0.5 mL microcentrifuge tube and centrifuge for 2 min at 4,000  $\times g$ .
5. Remove the MicroTube from the 0.5 mL microcentrifuge tube using the cap tool. Carefully remove the pestle using the cap tool and rinse the bottom half of the pestle with 20  $\mu\text{L}$  of liquid chromatography-mass spectrometry (LC-MS)-grade water and collect the wash into a clean 0.5 mL microcentrifuge tube.
6. Gently tap the MicroTube on the bench top to move the liquid to the bottom and transfer all of the solution from the MicroTube into the 0.5 mL microcentrifuge tube.
7. Add 20  $\mu\text{L}$  of LC-MS-grade water to the MicroTube and tap it gently on the bench top. Transfer the wash solution to the 0.5 mL tube and repeat this wash step once more.
8. Vacuum centrifuge to dry down the samples to  $\sim 2 \mu\text{L}$  and add 100  $\mu\text{L}$  of 100 mM TEAB, pH 8.0.
9. Determine the peptide concentration using a colorimetric assay (bicinchoninic acid (BCA) assay; see the **Table of Materials**) according to the manufacturer's protocol.

## 8. Tandem-mass tag (TMT) labeling and EasyPep cleanup

**NOTE:** Approximate time: 7 h 20 min (2 h 20 min without vacuum centrifuge drying time).

1. Bring the isobaric TMT labeling reagents to ambient temperature before opening. Add 500  $\mu\text{L}$  of 100% acetonitrile to each TMT vial (5 mg). Incubate for 10 min with occasional vortexing.
2. Dissolve 5  $\mu\text{g}$  of the peptide sample in 100  $\mu\text{L}$  of 100 mM TEAB, pH 8.0, and add 10  $\mu\text{L}$  of a given TMT reagent. Construct and include reference pools representing each individual sample in the experiment in each TMT multiplex set of samples to facilitate quantification of samples across multiple TMT multiplexes<sup>9</sup>. Incubate reactions for 1 h at ambient temperature with occasional shaking/tapping.
3. Quench the TMT labeling reaction by adding 10  $\mu\text{L}$  of 5% hydroxylamine and incubate for 30 min at ambient temperature with occasional tapping. After quenching, combine the TMT-labeled samples into one tube and dry to approximately 200  $\mu\text{L}$ .
4. Add 1,800  $\mu\text{L}$  of 0.1% formic acid. Check the pH with pH paper: if pH  $\sim 3$ , add 1 mL of 0.1% formic acid; if pH  $> 3$ , add 10-20  $\mu\text{L}$  of 5% formic acid until pH  $\sim 3$ . Add 0.1% formic acid to bring to a final volume of 3 mL.
5. Remove the tab at the bottom of the Peptide Clean-up Column, remove the cap, and place it in a 15 mL conical tube. Transfer the TMT-labeled sample to the column and proceed with cleanup as per the manufacturer's protocol.
6. Vacuum centrifuge to dry down the eluted peptides to  $\sim 20 \mu\text{L}$ , transfer to an LC vial using 25 mM ammonium bicarbonate for a final volume of 80  $\mu\text{L}$ , and proceed to offline fractionation.

## 9. TMT multiplex sample fractionation and pooling

**NOTE:** Approximate time: 3 h 30 min (1 h 30 min without vacuum centrifuge drying time).

1. Fractionate the TMT-labeled peptide multiplexes by basic reversed-phase chromatography into 96 fractions by developing an increasing linear gradient (0.69%  $\text{min}^{-1}$ ) of mobile phase B (acetonitrile) into mobile phase A (10 mM  $\text{NH}_4\text{HCO}_3$ , pH 8.0).
2. Generate 36 concatenated fractions by pooling sample wells. Vacuum centrifuge to dry fractions to  $\sim 2 \mu\text{L}$  and resuspend in 25 mM ammonium bicarbonate (final concentration 1.5  $\mu\text{g}/10 \mu\text{L}$ ), centrifuge at  $15,000 \times g$  for 10-15 min, and transfer to LC vials for MS analysis.

## 10. Liquid chromatography tandem mass spectrometry (LC-MS/MS)

**NOTE:** Approximate time: Instrument method and experimental design dependent.

1. Calibrate the mass spectrometer according to the manufacturer's instructions/protocols.
2. Prepare fresh mobile phases and standards and perform appropriate LC pre-run preparations (including but not limited to purge solvents, flush air, and leak test scripts for the referenced instrument [see the **Table of Materials**]). Equilibrate pre- and analytical columns and sample loop prior to starting analyses.
3. Prior to and in between serial TMT multiplex analyses, validate that the LC-MS system meets previously benchmarked performance metrics using quality assurance/quality control (QA/QC) TMT-labeled

peptide digests and (e.g., MSPE (see the **Table of Materials**), HeLa).

4. Load autosampler vials in the appropriate positions in the LC autosampler. Analyze individual fractions with an appropriate gradient/MS method. Intersperse a "wash" run with peptide standard (e.g., peptide retention time calibration [PRTC]) approximately once daily to evaluate chromatographic and mass spectral performance. Following the analysis of each TMT multiplex sample fraction series, run QA/QC TMT benchmark standards to evaluate system performance.
5. Run mass spectrometer evaluation routines after QA/QC TMT benchmark standards to assess post-sample performance and then calibrate system as in step 10.1 for next sample set.

## 11. Bioinformatic data analysis

**NOTE:** Approximate time: Experimental design dependent.

1. Transfer all sample data (e.g., .raw files) to an appropriate network storage/computer drive.
2. Search all fractions together using the desired data analysis application (e.g., Proteome Discover, Mascot) using appropriate parameters<sup>9</sup> against a species-specific protein reference database to generate peptide spectral matches (PSMs) and extract TMT reporter ion signal intensities. Filter PSMs based on appropriate quality control metrics, and aggregate normalized, median  $\log_2$ -transformed TMT reporter ion ratio abundances into global protein level abundances, as previously described<sup>3,9</sup>.
3. Compare protein alterations in conditions of interest using desired differential analysis software.

## Representative Results

Fresh-frozen tissue thin sections from two HGSOc and two OCCc patients were analyzed using this integrated AI-driven tissue ROI identification, segmentation, LMD, and quantitative proteomic analysis workflow (**Figure 1**). Representative H&E-stained tissue sections for each tumor were reviewed by a board-certified pathologist; tumor cellularity ranged from 70% to 99%. Tissues were thin-sectioned onto PEN membrane slides (**Supplemental File 2**) and pre-cut with calibrator fiducials (**Supplemental File 1**), enabling integration of positional orientation data from annotations generated in the image analysis software (see the **Table of Materials**) with Cartesian coordinate orientation in the LMD software. Following H&E staining, high-resolution images (20x) of the PEN slides containing the tissue plus calibrators were captured.

Tumor and stromal cell populations in the micrographs were segmented using image analysis software (see the **Table of Materials**) for selective harvest by LMD, along with harvests representing the entire tissue thin section (e.g., whole tumor tissue) (**Figure 1**). Non-discriminate annotations for whole tumor tissue collections were generated by partitioning the entire tissue section with tiles of  $500\ \mu\text{m}^2$ , leaving a  $40\ \mu\text{m}$  gap between tiles to maintain PEN membrane integrity and prevent the membrane from curling during LMD. On slides for histology-resolved LMD enrichment, the AI classifier in the image analysis software (see the **Table of Materials**) was trained to discriminate between tumor and stromal cells, along with the blank glass slide background. Representative tumor, stroma, and blank glass regions were manually highlighted, and the classifier tool was used to segment these ROIs throughout the entire tissue section. The segmented layers representing whole tissue, tumor epithelium, and

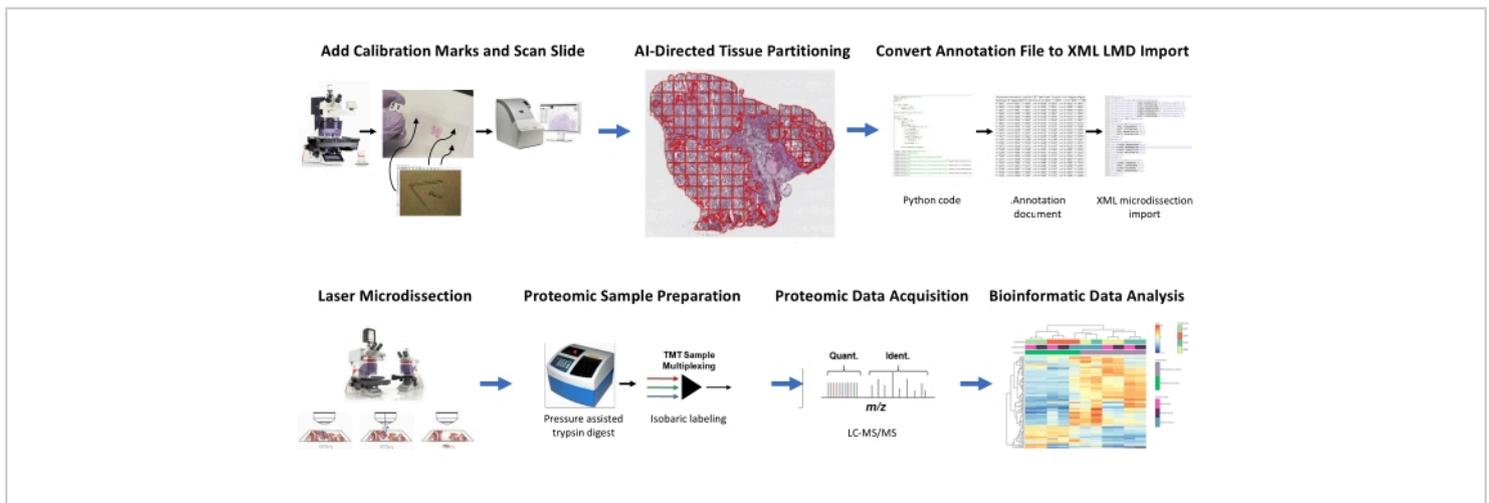
stroma were saved separately as individual .annotation files (**Supplemental File 3** and **Supplemental File 6**). In a separate copy of the image file (without the partitioned ROI annotations), a short line from the centermost tip of each of the three fiducial calibrators was annotated and saved as a .annotation file using the same file name as each of the LMD annotation layer files but appended with the suffix "\_calib" (**Supplemental File 4**). These lines were used to co-register the position of the PEN membrane calibrators with the annotation shape list data drawn in the image analysis software.

The present study provides two algorithms, "Malleator" and "Dapø" in Python to support this AI-driven LMD workflow, which are available at <https://github.com/GYNCOE/Mitchell.et.al.2022>. The Malleator algorithm extracts the specific Cartesian coordinates for all individual annotations (tissue ROI and calibrators) from the paired .annotation files and merges these into a single Extensible Markup Language (XML) import file (**Supplemental File 5**). Specifically, the Malleator algorithm uses the directory name from a parent folder as input to search all subdirectory folders and generates .xml files for any subfolders that do not already have a .xml merged file. The Malleator algorithm merges all annotation layers in the image analysis software (see the **Table of Materials**) into a single layer and converts the AI-generated shape list data, which is saved as proprietary .annotation file type, into .xml format compatible with the LMD software. After merging the annotation and calibrator files, the algorithm-generated .xml file is saved and imported into the LMD software. Slight adjustments are necessary to manually adjust the alignment of annotations, which also serves to register the vertical (z-plane) position of the slide stage on the laser microscope. The Dapø algorithm is used specifically for LMD-enriched collections. Partitioned

tiles are automatically assigned to individual annotation layers by the image analysis software. The Dapö algorithm merges all partitioned tiles into a single annotation layer prior to use of the Classifier tool, thereby reducing the Classifier analysis run time for LMD enriched collections.

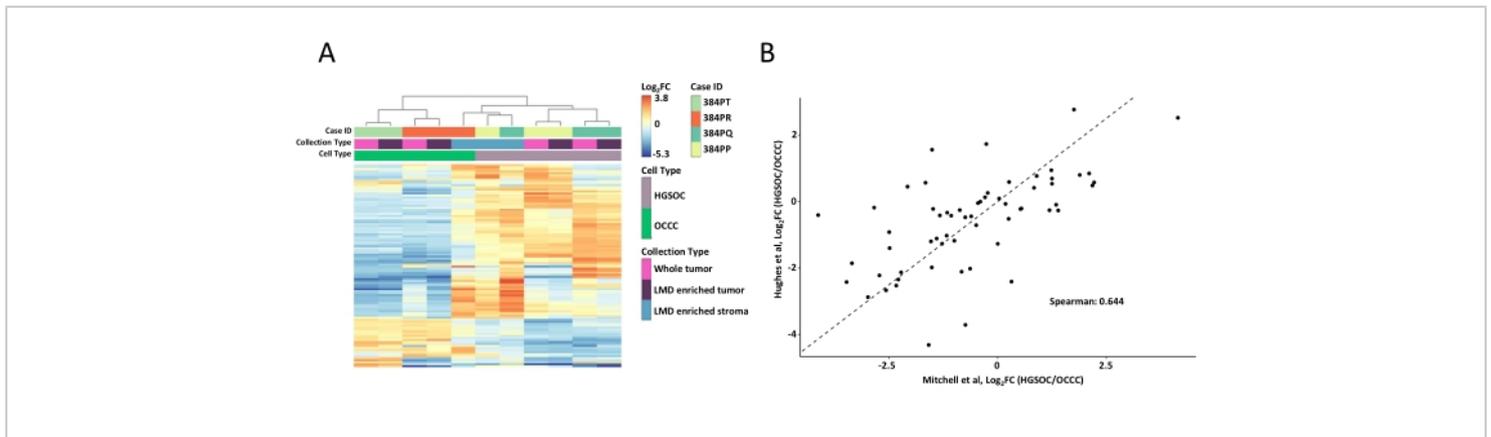
The whole tumor and LMD-enriched tissue samples were digested, labeled with TMT reagents, multiplexed, fractionated offline, and analyzed *via* quantitative MS-based proteomics as previously described<sup>9</sup>. The mean peptide yield (43-60 µg) and recovery (0.46-0.59 µg/mm<sup>2</sup>) for samples harvested using this AI-driven workflow were comparable with previous reports<sup>9,10</sup>. A total of 5,971 proteins were co-quantified across all samples (**Supplemental Table S1**). Unsupervised hierarchical clustering using the 100 most variable proteins resulted in segregation of the HGSOC

and OCCC histotypes from the LMD-enriched and whole tumor samples (**Figure 2A**), similar to that previously described<sup>11</sup>. By contrast, the LMD-enriched stroma samples from both HGSOC and OCCC clustered together and independently from the LMD-enriched tumor and whole tumor samples. Among the 5,971 quantified proteins, 215 were significantly altered (LIMMA adj.  $p < 0.05$ ) between whole tumor collections from HGSOC and OCCC specimens (**Supplemental Table S2**). These altered proteins were compared with those identified to differentiate HGSOC and OCCC tumor tissue by Hughes et al.<sup>11</sup>. Of the 76 signature proteins quantified by Hughes et al., 57 were co-quantified in this dataset and were highly correlated (Spearman Rho = 0.644,  $p < 0.001$ ) (**Figure 2B**).



**Figure 1: Summary of the integrated workflow for automated tissue region of interest selection for laser microdissection for downstream quantitative proteomics.** Calibration fiducials are cut onto PEN membrane slides to co-register positional orientation data from AI-derived segments of tissue ROI in the image analysis software, HALO, with horizontal positioning on the LMD microscope. The Malleator algorithm is used to merge the annotated segmentation data across all annotation layers for a slide with the \_calib reference file, and to convert it to a .xml file compatible with the LMD software. LMD-harvested tissue for proteomic analysis is digested and analyzed by high-throughput quantitative proteomics as previously described<sup>9</sup>. Abbreviations: LMD = laser microdissection; ROI = region of interest; TMT = tandem mass tag;

Quant. = quantification; Ident. = identification; LC-MS/MS = liquid chromatography-tandem mass spectrometry. [Please click here to view a larger version of this figure.](#)



**Figure 2: Analysis of the proteins in LMD-enriched and whole tumor samples.** (A) Unsupervised hierarchical cluster analysis of the 100 most variably abundant proteins in HGSOV and OCCV LMD enriched and whole tumor samples. (B) Correlation of log<sub>2</sub> fold-change protein abundances between HGSOV and OCCV whole tumor harvests in the present study (Mitchell et al., x-axis) and a similar study by Hughes et al. (y-axis)<sup>11</sup>. Abbreviations: LMD = laser microdissection; HGSOV = high-grade serous ovarian cancer; OCCV = ovarian clear cell carcinoma; log<sub>2</sub>FC = log<sub>2</sub>-transformed proteomic abundance. [Please click here to view a larger version of this figure.](#)

**Supplemental Table S1: Abundances of 5,971 proteins co-quantified across all LMD enriched and whole tumor samples from HGSOV and OCCV tissue specimens.** Abbreviations: LMD = laser microdissection; HGSOV = high-grade serous ovarian cancer; OCCV = ovarian clear cell carcinoma. [Please click here to download this Table.](#)

**Supplemental Table S2: Differentially expressed proteins (215) in whole tumor collections from HGSOV vs OCCV (LIMMA adj. p < 0.05).** Abbreviations: HGSOV = high-grade serous ovarian cancer; OCCV = ovarian clear cell carcinoma. [Please click here to download this Table.](#)

**Supplemental File 1: Representative shape list data (.sld) file containing standard calibrator fiducials for four slide**

**positions.** The file can be imported into the LMD software. [Please click here to download this File.](#)

**Supplemental File 2: Representative .svs image file of a H&E-stained high-resolution (20x) tissue section.** The file can be opened and viewed using image analysis software or LMD software. Abbreviation: H&E = hematoxylin and eosin; LMD = laser microdissection. [Please click here to download this File.](#)

**Supplemental File 3: Representative .annotation file of partitioned whole tumor segments.** The file can be imported into image analysis software. [Please click here to download this File.](#)

**Supplemental File 4: Representative \_calib.annotation file of calibrator fiducial segments.** Coordinate information

represents oriental positioning of the short calibrator lines drawn from each arrowhead fiducial. The file can be imported into image analysis software. [Please click here to download this File.](#)

**Supplemental File 5: Representative extensible markup language (.xml) file generated by the Malleator algorithm.**

The file can be imported into the laser microdissection software. [Please click here to download this File.](#)

**Supplemental File 6: Representative .annotation file of partitioned AI-classified segments for LMD-enriched collections.** The file can be imported into image analysis software. Abbreviations: AI = artificial intelligence; LMD = laser microdissection. [Please click here to download this File.](#)

## Discussion

While there have been multiple study precedents aimed at developing and/or improving workflows for enrichment of target cellular subpopulations from FFPE and/or fresh-frozen tissues and methodologies for maintaining sample quality during processing<sup>9,12,13,14,15</sup>, a substantial need exists to develop automated strategies for preparing clinical tissue specimens for molecular analyses to decrease variability and increase reproducibility. This workflow describes a standardized, semiautomated protocol that integrates existing image analysis software tools (see the **Table of Materials**) for histology-resolved harvest of discrete cell populations by LMD from clinical tissue specimens.

Spatially-resolved LMD enrichment of ROIs capturing discrete cell populations represents a next-generation tissue processing step prior to multiomic analyses to improve molecular characterization and identification and facilitate cell-selective biomarker discovery. This protocol improves upon existing methodologies by reducing the often-long

exposure of tissue sections to the ambient environment that is associated with manual segmentation of ROI by a histologist (which can take >1-2 h prior to the LMD collection). This workflow instead allows ROI to be preidentified by AI-guided classification and segmentation. Limiting tissue dwell time will decrease spurious variations in the assessments of highly labile molecular targets, such as phosphopeptides and mRNA, or for antibody-based analytical techniques that rely on a target protein being in its native conformation for detection.

Cutting neat calibrator fiducials onto the PEN membrane slide that are clearly visible in the scanned slide image is one of the key components enabling integration of the image analysis software (see the **Table of Materials**) with the LMD workflow. Ensuring that the calibrators have a precise ("clean") point at the bottom of the "V" shape allows for selection of a precise point in the image analysis software for the calibrator lines to be drawn from, as described in steps 5.1.6 and 5.2.13. Alignment of these points during import into the LMD software is critical for properly overlaying the annotations (facilitated through the generation of a compatible .xml file using the "Malleator" and/or "Dapø" algorithms) onto the relevant tissue ROI on the physical LMD slide. It is necessary to highlight all shapes and collectively "drag and drop" into place even when the alignment is precise upon import into the LMD software to register the vertical (z-plane) position of the slide stage on the laser microscope. Minor adjustments to the positioning of the annotations over the tissue ROI can also be made during this step, if needed.

A limitation of the current version of the Malleator algorithm is that it is not compatible with the predefined annotation shape tools provided by the image analysis software (see the **Table of Materials**), although future updates/versions

of the algorithm will aim to improve this compatibility. The .annotation file for shapes drawn using these tools contains only two sets of paired x and y coordinates for each annotation, without the complete spatial orientation around those points. Current use of these tools results in the annotations being converted to straight lines defined by only two points during the import process. Manual definition of tissue ROI segments is required for successful conversion to XML format and LMD import. This can either be performed by manually defining each ROI with individual free-hand polygonal annotations specific to the target area or by applying an approximated circular or rectangular annotation across all tissue ROI segments, if desired, and will be compatible with this workflow.

While the workflow presented here was demonstrated for proteomic analysis of fresh-frozen human cancer tissue specimens, this AI-driven LMD workflow can be equivalently used with FFPE tissues, non-cancerous tissue types, and those from non-human sources. It can also support other downstream molecular profiling workflows, including transcriptomic, genomic, or phosphoproteomic analyses. This workflow can also leverage other uses of the image analysis software (see the **Table of Materials**), including with capabilities associated with cell counting or other analytical modules, including the "Multiplex IHC" module or the "Tissue Microarray (TMA) Add-on". Future applications of this workflow may also benefit from predefining the number of cells per ROI segment, thereby ensuring equivalent cellular inputs across multiple collections, or by using alternative methods to define cellular ROIs of interest, such as by immunohistochemistry or cell sociology.

## Disclosures

T.P.C. is a ThermoFisher Scientific, Inc SAB member and receives research funding from AbbVie.

## Acknowledgments

Funding for this project was provided in part by the Defense Health Program (HU0001-16-2-0006 and HU0001-16-2-00014) to the Uniformed Services University for the Gynecologic Cancer Center of Excellence. The sponsors had no role in the design, execution, interpretation, or writing of the study. **Disclaimer:** The views expressed herein are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or U.S. Government.

## References

1. Motohara, T. et al. An evolving story of the metastatic voyage of ovarian cancer cells: cellular and molecular orchestration of the adipose-rich metastatic microenvironment. *Oncogene*. **38** (16), 2885-2898 (2019).
2. Aran, D., Sirota, M., Butte, A. J. Systematic pan-cancer analysis of tumour purity. *Nature Communications*. **6**, 8971 (2015).
3. Hunt, A. L. et al. Extensive three-dimensional intratumor proteomic heterogeneity revealed by multiregion sampling in high-grade serous ovarian tumor specimens. *iScience*. **24** (7), 102757 (2021).
4. Dou, Y. et al. Proteogenomic characterization of endometrial carcinoma. *Cell*. **180** (4), 729-748 (2020).

5. Zhang, H. et al. Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell*. **166** (3), 755-765 (2016).
6. Gillette, M. A. et al. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell*. **182** (1), 200-225 (2020).
7. Silvestri, A. et al. Protein pathway biomarker analysis of human cancer reveals requirement for upfront cellular-enrichment processing. *Laboratory Investigation*. **90** (5), 787-796 (2010).
8. Echle, A. et al. Deep learning in cancer pathology: a new generation of clinical biomarkers. *British Journal of Cancer*. **124** (4), 686-696 (2021).
9. Lee, S. et al. Molecular analysis of clinically defined subsets of high-grade serous ovarian cancer. *Cell Reports*. **31** (2), 107502 (2020).
10. Xuan, Y. et al. Standardization and harmonization of distributed multi-center proteotype analysis supporting precision medicine studies. *Nature Communications*. **11** (1), 5248-5248 (2020).
11. Hughes, C. S. et al. Quantitative profiling of single formalin fixed tumour sections: proteomics for translational research. *Scientific Reports*. **6** (1), 34949 (2016).
12. Espina, V. et al. A portrait of tissue phosphoprotein stability in the clinical tissue procurement process. *Molecular & Cellular Proteomics*. **7** (10), 1998-2018 (2008).
13. Espina, V., Heiby, M., Pierobon, M., Liotta, L. A. Laser capture microdissection technology. *Expert Review of Molecular Diagnostics*. **7** (5), 647-657 (2007).
14. Havnar, C. A. et al. Automated dissection protocol for tumor enrichment in low tumor content tissues. *Journal of Visualized Experiments*. (169) (2021).
15. Mueller, C. et al. One-step preservation of phosphoproteins and tissue morphology at room temperature for diagnostic and research specimens. *PLoS One*. **6** (8) (2011).