Unlock the Power of Spatial Biology with phenoptrReports

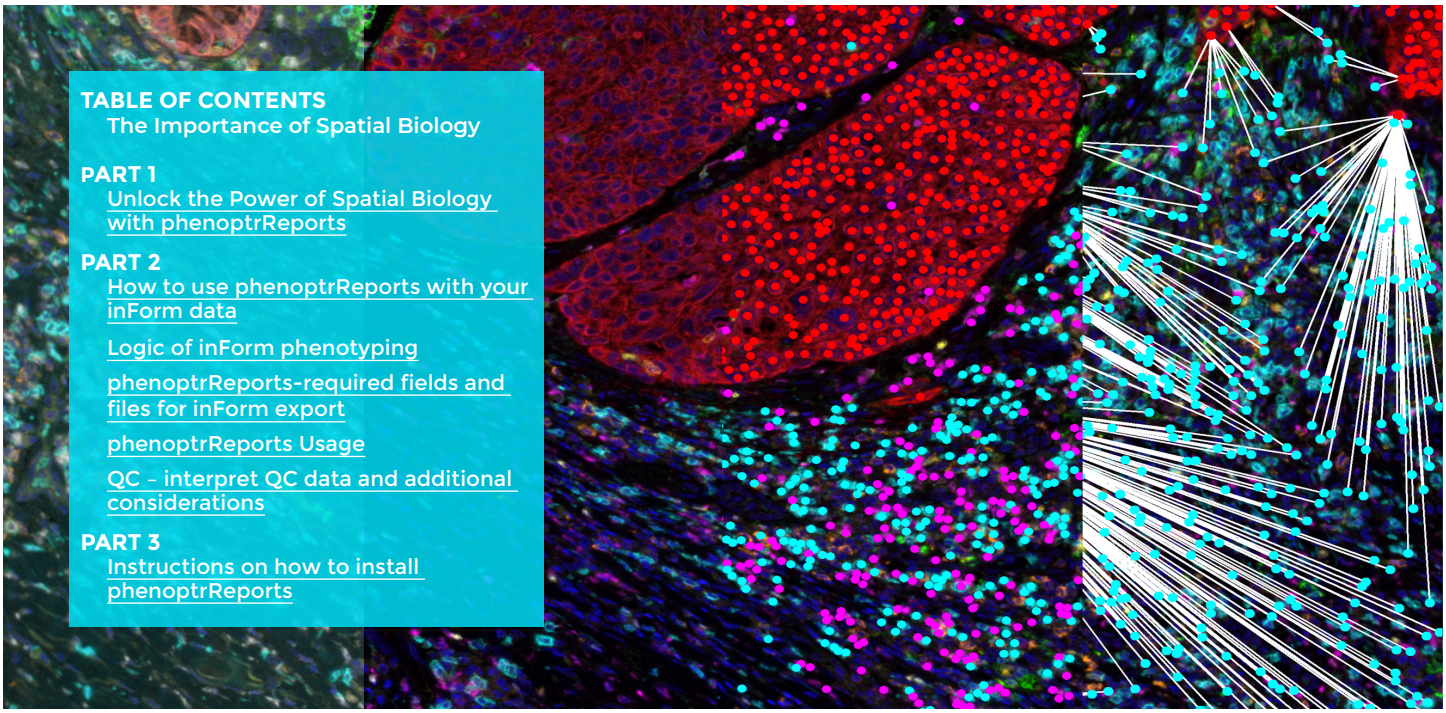


TABLE OF CONTENTS

Akoya's PhenoImager™ (formerly Phenoptics™) workflow offers an end-to-end solution for multiplex immunofluorescence, digital image analysis, and spatial biology.

THE IMPORTANCE OF SPATIAL BIOLOGY

Single cell analysis technologies have been transformative in understanding cellular heterogeneity within a larger population. However, accurately identifying cell lineage, functionality, and the important cellular molecular factors within the context of their resident tissues or organs remains difficult with current single cell technologies.

In 2020, Spatial Transcriptomics, and by extension Spatial "-omics" in general, was voted Method of the Year by the journal *Nature Methods*, underlining the growing importance of Spatial Biology. Akoya, the Spatial Biology Company, provides state-of-the-art spatial proteomics solutions for any stage to early discovery, to translational, and clinical research.

The heterogeneous cellular makeup of a tissue is a product of both the innate and adaptive immune systems and of the diverse communities that reside there. Spatial Biology can help define a tissue's heterogeneous nature by phenotyping single cells in context, across the whole tissue without alteration or disruption of the sample. In contrast to capturing subtle changes in isolated cell suspensions or whole tissue extracts, Spatial Biology captures information at the source, at the individual cell, within the intact spatial environment. With Spatial Biology, researchers capture a cell's identity and function, location, and relation to its environment.

In cancer biology, immunology, immuno-oncology, and beyond, spatial context matters. For instance, the lab of Bernie Fox at Providence Cancer Institute has used Spatial Biology to develop an immunoscore to more accurately predict a patient's response to immunotherapy. With this immunoscore, the spatial interaction of immunosuppressive and immunogenic cells at the tumor invasive margin can now be used for patient stratification, and in theory, a more accurate bench-to-bedside approach for assessing treatment effectiveness.

To enable true Spatial Biology, novel digital pathology and digital image analysis tools are needed. Together with Akoya's inForm machine learning software, phenoptrReports unlocks spatial analysis of multiplex immunofluorescence-stained tissues.

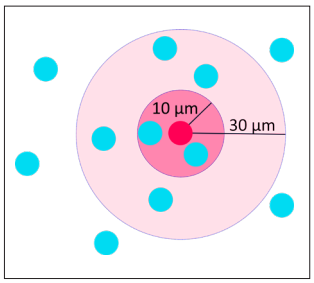


FIGURE 1: Spatial context matters. The number of regulatory T-cells within a specified radius of a given CD8+ cytotoxic T-cell informs whether that cell is active or suppressed. Adapted from Feng, Z. et al... Fox, B., Multiparametric immune profiling in HPV- oral squamous cell cancer. JCI Insight 2, (2017).

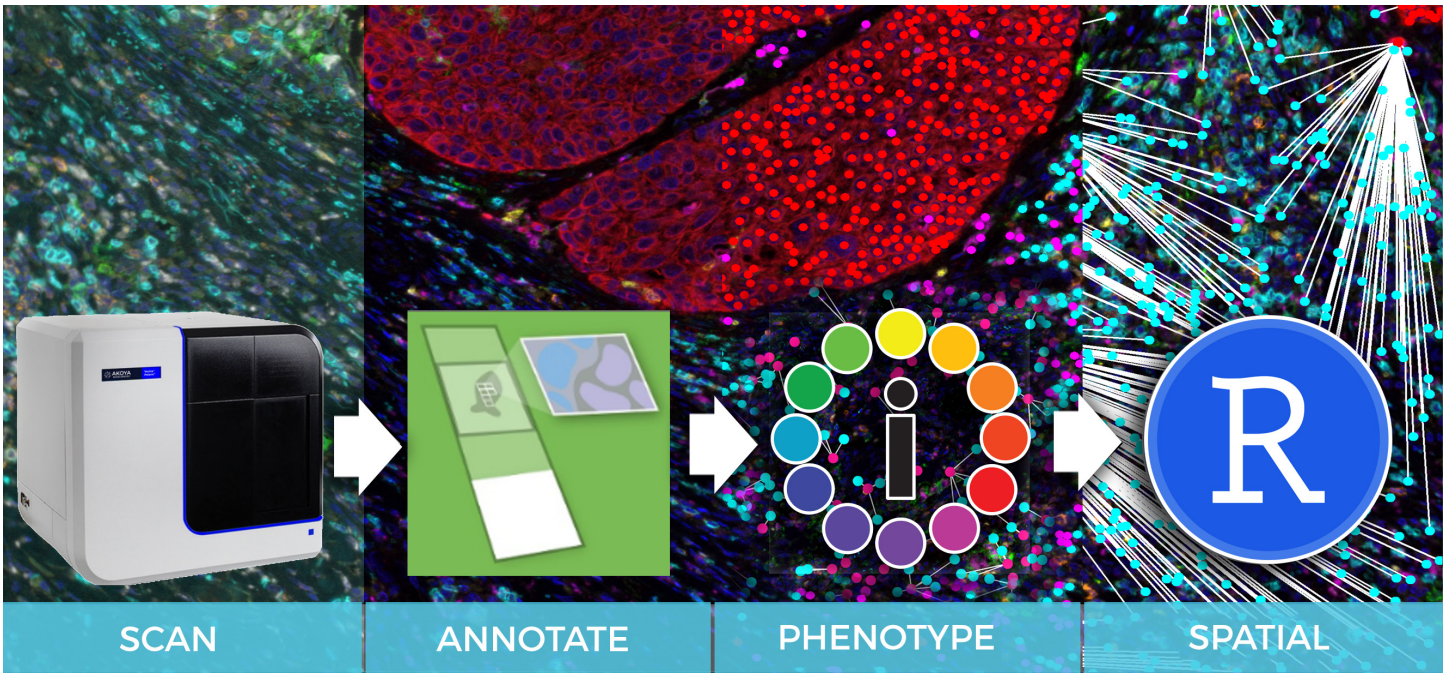


FIGURE 2: Outline of the Akoya PhenoImager™ digital pathology workflow. Images (multi- or single-plex) are acquired with an Akoya scanner, annotated in Phenochart software, tissue categories and cells are identified and phenotyped using inForm machine learning software, and spatial reports are created using phenoptrReports within the R-Studio environment.

PART 1

Unlock the Power of Spatial Biology with phenoptrReports

As part of the PhenoImager™ multiplex immunofluorescence and digital pathology workflow (FIGURE 2), Akoya offers phenoptrReports as an open-source tool. It ties directly into the phenotyping data produced by inForm machine learning software.

Built upon the popular open source platform R and RStudio, phenoptrReports allows researchers to expand their digital tissue analysis studies beyond phenotyping into the spatial biology realm. With phenoptrReports, you can:

Create reports

- Count cells across slide(s) and studies
- Count cells across tissue categories
- Report numbers for all cells
- Report numbers for specific phenotypes
- Uncover and verify rare phenotypes
- Create UpSet plots for all possible marker co-expression patterns in a study across slides and per slide (FIGURE 3)
- Find out which markers are most frequently co-expressed.

Calculate cell densities

- Cell densities for all cells or select phenotypes across the entire tissue or categories.

Perform spatial analysis

- Nearest Neighbor analysis: Learn, for example, where the nearest FoxP3+ regulatory T cell is in relation to each CD8+ cytotoxic T cell (FIGURE 4).
- Mutual nearest neighbor analysis: Learn, which two cells are mutually the closest cell to each other.
- Counts within: Calculate the number of select cells within a specified radius of another select phenotype, for one or multiple radii.
- Create Distance Heatmaps: View the Distance(s) of all available phenotypes to each other.
- Spatial Map Viewer: Visualize nearest neighbors and mutual nearest neighbors and export plots, and visualize cells with 2 select phenotypes that are touching (FIGURE 5).

Streamline quality control of multiplex immunofluorescence studies

- Assess robustness of libraries used for unmxing
- Report component levels for multiplex slides to validate and improve staining

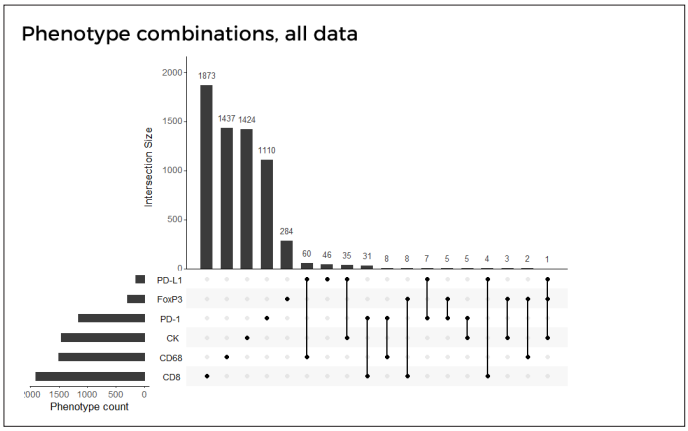


FIGURE 3: An ***UpSet plot*** highlights all biomarker co-expression patterns present in a given sample. The bars on the left identify cells expressing individual markers. The vertical bars identify the number of cells in each "intersection", or co-expression pattern. The dots and vertical lines resemble "intersections", i.e., co-expressed markers. For example, in column 6, two dots for PD-L1 and CD68 are connected by a line. This means, there are 60 cells in this image, that express both PD-L1 and CD68.

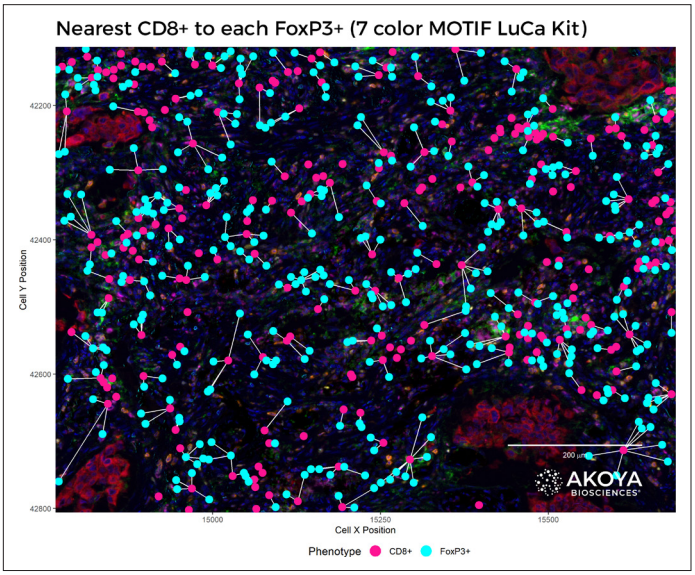


FIGURE 4: A nearest neighbor diagram shows cells in their spatial context. Here, the pink dots highlight CD8+ cytotoxic T-cells, and the teal dots show the position of FoxP3+ regulatory T-cells. The lines connect each FoxP3+ cell with the nearest CD8+ cell. Often, multiple FoxP3+ cells share one nearest CD8+ cell.

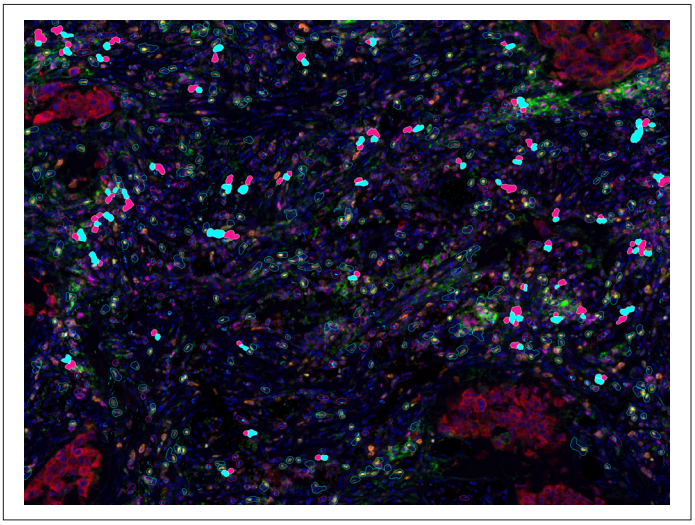


FIGURE 5: The spatial map viewer can visualize touching cells.

PART 2

How to use phenoptrReports with your inForm data

Logic of inForm phenotyping

inForm Machine Learning Software is a powerful tool for digital slide analysis. With inForm, researchers unlock advanced tissue analysis and phenotyping.

A typical inForm project starts with unmixing multispectral images using appropriate libraries, followed by tissue segmentation. Tissue segmentation divides the tissue into set categories, such as tumor, stroma, and others. This could be any area of the tissue that has common features, for example, germinal centers and t-cell zones in tonsil tissue, epidermis and dermis in skin cross-sections, or glomeruli in the kidney.

For phenotyping and cell counting, imaging data is converted into objects by cell segmentation. Cell segmentation identifies single cells based on their nuclear stain, cytoplasmic, and membrane markers. Once cells are converted into addressable objects, users can set up complex phenotyping schemas.

Lineage markers

When phenotyping cells using inForm, each cell can initially only belong to one "phenotype". Phenotypes are best selected based on expression of a single marker. It is essential to follow the phenoptrReports naming convention of adding a + sign behind each phenotype (e.g. CD3+), except for cells named "other". After inForm phenotyping, phenotypes can be combined in phenoptrReports to display and count more complex phenotypes, taking into account multiple marker positivity. So-called lineage markers are markers that define a specific cellular lineage – a cell should only be positive for one of the lineage markers used in a given multiplex panel at a time. Mutually exclusive lineage markers can be phenotyped together, while co-expressed markers need to be separated into distinct rounds of phenotyping. For example, in a panel that contains CD8, CD68, panCK, FoxP3, PD-1, and PD-L1, the markers CD8, CD68, and panCK would be mutually exclusive and can be phenotyped together, including a fourth phenotype named "other". "Other" cells are cells that do not fit in any of the phenotypes, for example stromal fibroblasts. phenoptrReports will automatically ignore "other" cells for spatial analysis, but include the data when calculating overall cell counts, densities, and percentages.

Functional markers

In contrast to lineage markers, a cell may potentially express multiple functional markers at the same time, or cells from multiple lineages can express the same functional marker. In our example, FoxP3, PD-1, and PD-L1 are functional markers that can sometimes be co-expressed. If functional markers are mutually exclusive, they can be phenotyped together, again adding an "other" phenotype. If functional markers are potentially co-expressed, the co-expressing markers need to be separated into a third, or even fourth, phenotyping schema.

Inform 2.6 vs older versions

For versions of inForm prior to 2.6, phenotyping for lineage and functional markers can be separated by creating copies of inForm projects, leaving unmixing, tissue, and cell segmentation unchanged, and only updating the phenotyping section of an inForm project. These projects can be saved under a new name, and phenotype-specific algorithms can be batch-analyzed sequentially.

With inForm 2.6, the ability to create multiple phenotyping schemas was added directly into inForm. Users can create one schema for lineage markers, another for mutually exclusive functional markers, and a third or fourth schema for overlapping markers. These schemas are saved directly to the inForm project or algorithm, allowing for faster and more efficient batch processing.

Regardless of the inForm version used, phenoptrReports functions identically.

phenoptrReports-required fields and files for inForm export

phenoptrReports receives data mostly from the cell segmentation files produced by inForm. When exporting inForm results, always export all data tables. In addition, for the spatial map viewer to work, phenoptrReports will require component and composite images for each image field.

By default, inForm exports all available data columns when saving data tables. phenoptrReports does not require all those columns. phenoptrReports expects the following data columns: Components, (except autofluorescence (AF)), Component Mean, Core ID when working with Tissue Microarrays, Tissue Categories (except "glass" or "background"), X and Y Position, Distance from Tissue Category Edge and Area (in square microns) as shown in Figure 6. All other fields can be unselected to save data storage. Recent updates to phenoptrReports include a *"Keep only phenoptrReports fields"* function in the *"consolidate and summarize"* view, which results in smaller consolidated data file sizes as well.

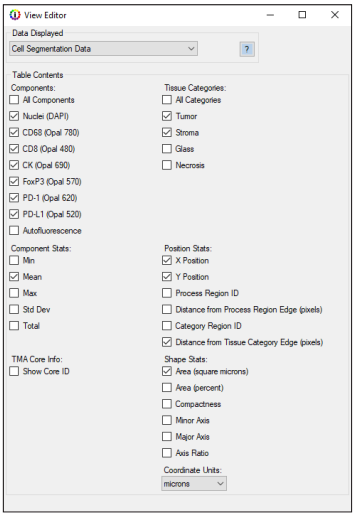
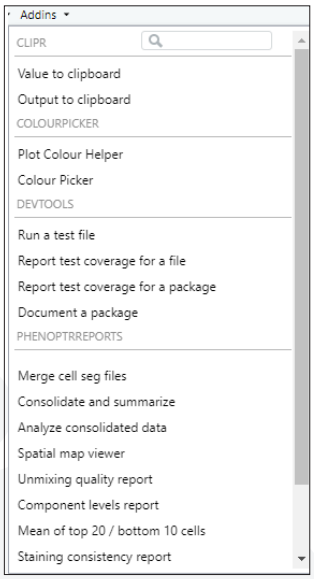


FIGURE 6: inForm Fields required for phenoptrReport analysis. The fields not selected can be omitted, resulting in a reduction of export file size.



phenoptrReports Usage

Please install phenoptrReports and prerequisites as listed in the appendix at the end of this document. Once installed properly, phenoptrReports will be available from the Addins menu in RStudio (FIGURE 7). The order in which the options are listed corresponds to the intended order of usage, however you can skip certain options if they are not needed for a project.

FIGURE 7: RStudio Addins menu. Once installed, phenoptrReports can be selected directly from the RStudio Addins menu.

Merge cell segmentation files

The *Merge cell seg files* Addin merges single image fields into one joint file (FIGURE 8). This is the same function as inForm Review/Merge, but without the ability to review/reject fields. If the data have already been processed with the inForm merge function, merging in phenoptrReports is not needed. To merge single image field data, select a folder that contains one or multiple 'AnnotationID'_cell_seg_data.txt files. (NOTE: when selecting the folder from within RStudio, it will appear empty). Once opened, phenoptrReports will automatically parse all 'AnnotationID'_cell_seg_data.txt files. A frequent error when selecting a folder with single field cell segmentation data files, is "*Please select a source directory which does not contain existing merge files.*" This means that phenoptrReports sees that there is a file with the word "merge" in the file name. Simply remove the merge_status.xml file from that folder and run the Addin again. Repeat merging for each phenotyping project in a study (NOTE: repeat merging is not required when using inForm 2.6 or higher).

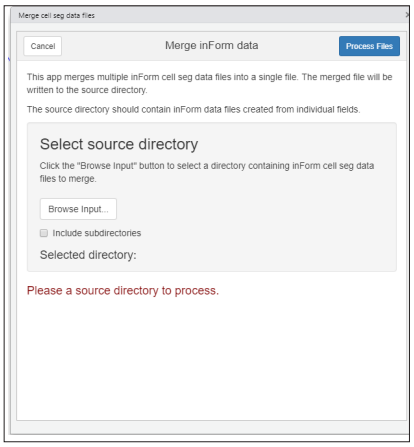


FIGURE 8: phenoptrReports merge screen

Consolidate and summarize

Before phenoptrReports can create reports with cell counts etc., merged data for lineage and functional phenotyping projects needs to be consolidated using the *consolidate and summarize* Addin (FIGURE 9). This Addin creates a single data table for the entire study and is essential for projects from any inForm software version, even if there is only one single merge_cell_seg_data.txt file. To save space, you may select *Keep only phenoptrReports fields*. This trims off unneeded data columns and results in substantially smaller file sizes.

When clicking on *Process Files*, the data in the merged files will be consolidated and three files will be created in the output directory:

- Consolidated_data.html: contains cell counts and UpSet plots for each image
- Consolidated_data.txt: contains raw data for each cell across each image
- Merge_cell_seg_data.html: contains the same information as Consolidated_data.html

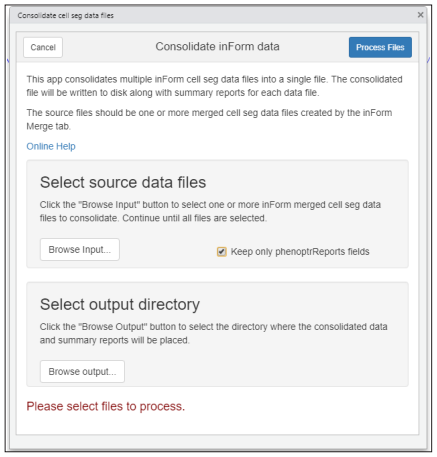


FIGURE 9: phenoptrReports consolidate data screen.

Common phenoptrReports messages:

Rows do not match – this happens when the number of cells in lineage and functional phenotyping projects does not match. Make sure to use the identical tissue- and cell segmentation algorithms for all phenotyping projects in a study. This error should not occur when using data generated with inForm 2.6 and higher.

Cannot allocate vector of size ###.# Mb – This error means that your computer does not have sufficient memory (RAM) installed to handle your project file sizes. The *Keep only phenoptrReports fields* selection can mitigate this error, as well as trying to split your study into smaller chunks. Alternatively, use a computer with more memory installed. We recommend at minimum 32 Gb of memory for larger analysis projects.

Other errors – phenoptrReports is open source and relies on many libraries maintained by independent programmers. Usually, re-installing phenoptr and phenoptrReports fixes broken dependencies.

Analyze consolidated data

When using the *Analyze consolidated data* function (FIGURE 10), select the consolidated data file from the previous step, as well as an output directory.

Optional: If cell densities are within the scope of the analysis, select a Merge_cell_seg_data_summary.txt file for your project. This file can correspond to any separate phenotyping project.

Optional: If you scored your images for a single biomarker, select the Merge_score_data.txt for your project. Please note that only H-Score type score data are supported.

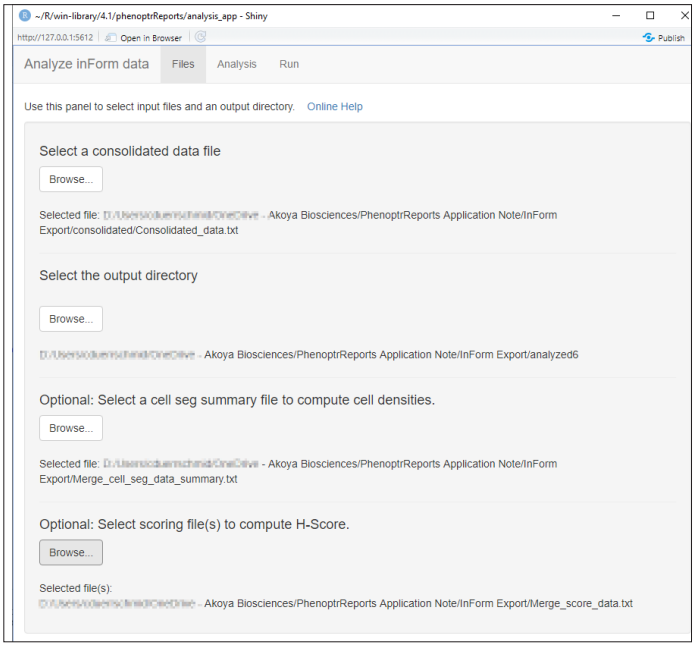


FIGURE 10: phenoptrReports file selection for analysis screen.

Annotating phenotypes

In the next tab (FIGURE 11), Researchers can summarize data by Slide ID (whole slides) or tissue micro array annotation/core. Aside from Phenotypes that were defined using inForm (for example CD8+, CK+, or similar), phenoptrReports enables researchers to count and create reports for more complex phenotypes and multiple-positive cells. Phenotype definitions are composed of the names of phenotypes. Definitions must end in + or -, for example CD3+ or CD8-. The allowed operators are forward slash (/), comma (,), plus sign (+) and minus sign (-).

Some examples of valid phenotypes:

- Cells that are positive for two markers: ***CD8+/CD3+***
- Cells that are positive for either of two markers: ***CD8+,CD3+***
- Cells that are positive for one, but negative for another marker: ***CD8+/CD3-***

Phenotype definitions may also include valid expressions such as

~`Membrane PD-L1 (Opal 520) Mean`>5

In this case, this "phenotype" includes all cells that have PDL-1 expression of greater than 5 normalized counts in the membrane cell compartment.

Phenotypes can be combined with expressions, using the operators listed above, for example

CD68+/~`Membrane PD-L1 (Opal 520) Mean`>5

In this case, this cell type includes all CD68+ macrophages that also have PDL-1 expression of greater than 5 normalized counts in the membrane cell compartment.

After selecting the phenotypes for analysis, a custom script is displayed in the *Run* tab and will automatically run when starting the analysis by clicking on *Create Report*.

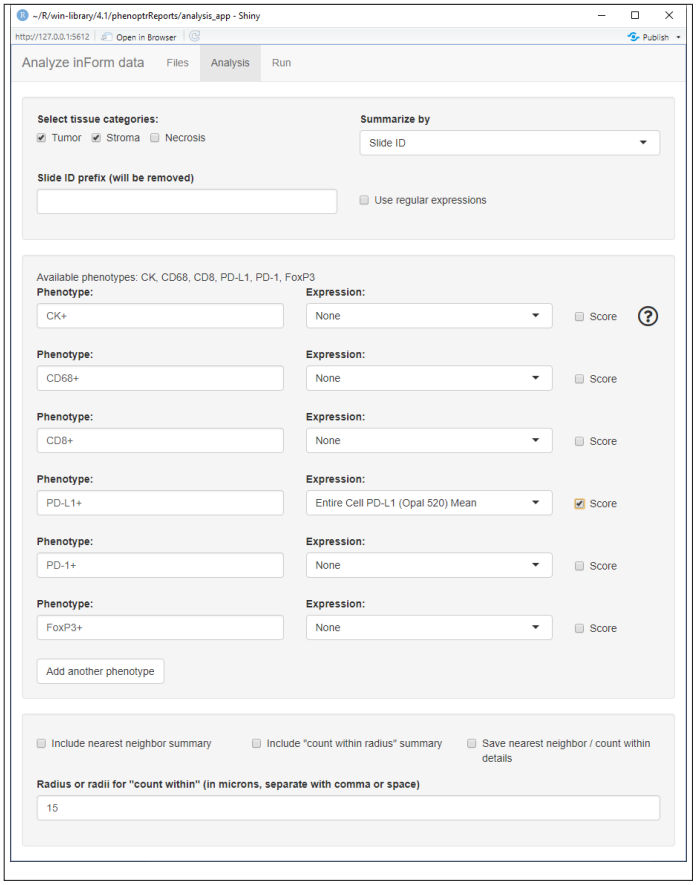


FIGURE 11: phenoptrReports main analysis screen.

The following files will be created in the output directory:

- **Charts.docx:** this file contains cell count charts and distance heatmaps
- **Results.xlsx:** this file contains tabular results for all cells and tissue categories (see interpretation below)
- **Sript.R:** this custom script can be re-used to perform identical analysis across multiple studies without needing to manually enter parameters
- **session_info.txt:** contains information on the R Studio session
- **count_within_All.txt**: this file will be created if *Include "count within radius" summary* and *Save nearest neighbor / count within details* was selected, additional files for each tissue category will be created, for example count_within_Stroma.txt
- **nearest_neighbors_All.txt:** this file will be created if *Save nearest neighbor / count within details* was selected, additional files for each tissue category will be created, for example nearest_neighbor_Stroma.txt

Interpret results.xlsx file:

- Slide Summary tab: list of slides and fields per slide
- Cell Counts and Cell Percents tabs: cell counts per slide per category in absolute numbers and percentage of total cells
- Cell Densities (optional): cell densities per slide/tissue microarray core, if a merge_cell_seg_data_summary.txt file was selected
- Mean Expression (optional): if mean expression of a given marker in a specified compartment was selected
- H-Score (optional): H-Score per slide or tissue microarray core, if a merge_score_data.txt file was selected
- Nearest Neighbors (optional): nearest neighbor data if *nearest neighbor analysis* was selected. Contains the following columns for each phenotype combination:

	NEAREST NEIGHBOR DISTANCES FOR PHENOTYPE PAIRS (MICRONS)							
Slide ID	**Tissue Category**	**From**	**To**	**Min**	**Mean**	**Median**	**Max**	**SD**
Slide ID	Tumor, Stroma, etc	Origin cell	Destination cell	Minimum distance	Mean distance	Median distance	Maximum distance	Standard deviation

Count Within (optional): count within data if *count within analysis* was selected. Contains the following columns for each phenotype combination:

	COUNT OF CELLS WITHIN THE SPECIFIED RADIUS							
Slide ID	**Tissue Category**	**From**	**To**	**Radius**	**From count**	**To count**	**From with**	**Within mean**
Slide ID	Tumor, Stroma, etc	Origin cell	Destination cell	Search radius	Count of cells in "from"	Count of cells in "to"	*	#

*Count of cells in "from" with at least one cell in "to" within the radius defined in "radius". #Mean number of "to" cells within "radius" of a single "from" cell.

QC – INTERPRET QC DATA AND ADDITIONAL CONSIDERATIONS

There are four image quality control tools built into phenoptrReports and can be found at the bottom of the Addins menu (Figure 7):

QC REPORT 1: Unmixing Quality

The unmixing quality report analyzes unmixed, **singleplex** images to help predict panel and library multiplex and unmixing quality. This report shows crosstalk between spectra and highlights potential problem areas during assay or library development. For each image, the brightest 3,200 pixels in the primary component are found (ignoring the 0.0100% very brightest). For these pixels, the mean signal level in each component is computed. The values are shown in the Signal Strength table (Table 4).

Crosstalk by image

The ***Crosstalk by image*** table (Table 1) evaluates how well your library images are predicted to unmix from one another. The input for this algorithm are library snapshots that have had spectra extracted and that spectral library is used to unmix the snapshots and export component tiffs with names including "Opal 570", "DAPI", etc. In the example in Table 1, there are not any expected unmixing issues using this spectral library, as the highest percentage level of non-signal to signal component, (1.6%) is still very low (colored red because it is the highest on the chart).

Source	DAPI	Opal 520	Opal 540	Opal 570	Opal 620	Opal 650	Opal 690
AF_1_component_data.tif	0.0%	0.4%	**0.8%**	0.1%	0.1%	0.0%	0.0%
Opal 520_2_component_data.tif	0.0%		0.2%	0.1%	0.0%	0.0%	0.0%
Opal 540_1_component_data.tif	0.0%	0.2%		0.4%	0.0%	0.0%	0.0%
Opal 570_1_component_data.tif	0.3%	0.0%	0.3%		**1.3%**	0.0%	0.0%
Opal 620_2_component_data.tif	0.0%	0.0%	0.0%	**1.2%**		**0.7%**	0.1%
Opal 650_1_component_data.tif	0.0%	0.0%	0.0%	0.0%	0.0%		**1.6%**
Opal 690_1_component_data.tif	**0.9%**	**1.5%**	0.0%	0.0%	0.0%	**0.8%**	

TABLE 1: Crosstalk by image table. This table shows crosstalk from each primary fluorophore to the other components in the same image. This is a measure of unmixing error.

Crosstalk by component

This section provides guidance on using unmixing error considerations when building your multiplex panel. Components used to assess expression levels have stricter requirements than those used for purely phenotyping (see guidance table below [TABLE 3]). Based on the guidance table below, it is probably best to pair Opal 690 with a phenotype marker rather than an expression-level/ functional marker. There are no considerations or concerns for the other Opals in this case. (NOTE: DAPI values do not raise any concern as there was not a DAPI single stain included in this library analysis so there is no relative positive control present.)

Source	DAPI	Opal 520	Opal 540	Opal 570	Opal 620	Opal 650	Opal 690
AF_1_component_data.tif	0.0%	0.2%	0.5%	0.1%	0.1%	0.0%	0.0%
Opal 520_2_component_data.tif	0.9%		0.2%	0.1%	0.0%	0.0%	0.0%
Opal 540_1_component_data.tif	5.1%	0.1%		0.6%	0.0%	0.0%	0.0%
Opal 570_1_component_data.tif	35.8%	0.0%	0.2%		1.3%	0.0%	0.0%
Opal 620_2_component_data.tif	1.7%	0.0%	0.0%	1.1%		0.5%	0.1%
Opal 650_1_component_data.tif	0.0%	0.0%	0.0%	0.0%	0.0%		2.2%
Opal 690_1_component_data.tif		0.7%	0.0%	0.0%	0.0%	0.6%	

TABLE 2: Crosstalk by component table. This table estimates crosstalk from other fluorophores to the primary fluorophore as a percentage of the primary fluorophore. This estimates the effect on unmixing error on other signals.

Guidance Table

Crosstalk values for a given component	Typical performance of component for different tasks	
	Expression level assessment	Cell lineage/phenotype classification
All values in a column are < 2%	Good	Good
All values in a column are < 5% Some values are > 2%	Marginal, depending on dynamic range of expression	Good
All values in a column are < 10% Some values are > 5%	Poor	Ok, may require more extensive classifier training
Any crosstalk values > 10%	Poor	Marginal, see Opal Assay Development Guide for advice

TABLE 3: Guidelines on interpreting results in table 2. Crosstalk by component. Components used to assess expression levels have stricter requirements than components used for phenotyping.

- *Pixel intensity in primary component* - This plot shows intensity distribution of primary fluorophore in each library image. It is helpful to ensure the intensities are balanced between the library signals.

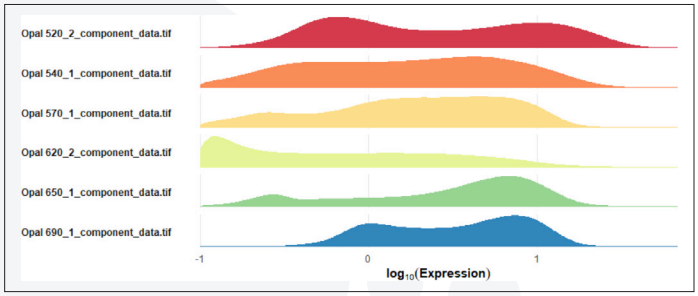


FIGURE 12: Plot showing pixel intensity distribution of the primary fluorophore in each library image.

- *Signal Strength* - Table 4 below shows a well-balanced library with all brightest pixels measured in the ~20-50 counts range.

Signal Strength Table

Source	DAPI	Opal 520	Opal 540	Opal 570	Opal 620	Opal 650	Opal 690
AF_1_component_data.tif	0.00	0.09	0.17	0.01	0.01	0.00	0.00
Opal 520_2_component_data.tif	0.00	47.27	0.07	0.03	0.00	0.00	0.00
Opal 540_1_component_data.tif	0.01	0.07	30.91	0.13	0.00	0.00	0.00
Opal 570_1_component_data.tif	0.07	0.01	0.06	22.25	0.28	0.00	0.00
Opal 620_2_component_data.tif	0.00	0.00	0.00	0.26	20.92	0.15	0.03
Opal 650_1_component_data.tif	0.00	0.00	0.00	0.00	0.00	31.07	0.51
Opal 690_1_component_data.tif	0.20	0.35	0.00	0.00	0.00	0.18	22.68

TABLE 4: Signal strength table. This table shows signal strength for each component, measured at the brightest pixels of the primary fluorophore. This is the raw data underlying the crosstalk tables.

QC REPORT 2: Component levels

Unlike the unmixing quality report, the component levels report analyzes unmixed, **multiplex** images to help evaluate staining and unmixing quality. This report shows the relative strength of each component and crosstalk between components and is a helpful tool to use during panel development and optimization. It aims to evaluate the following aspects of a multiplex panel by comparing distribution of intensity counts across ROIs:

- Qualification of signal balance across fluorophores
- Evaluation of heterogeneity of signals across an image
- Determination of spectral overlap/crosstalk
- Evaluation of signal-to-noise

If you are swapping a new marker into an existing, optimized panel, it can be helpful to compare component levels reports to decide when additional optimization is needed.

Qualifying well balanced signals

The *Pixel intensity by image section* measures the top 99.9th percentile pixel is usually reflective of 'positive' signal. If your multiplex panel is well-balanced, the 99.9th percentile values should be similar between channels. In the example below (FIGURE 13), you can see the 'positive' signals from CD68, CD8, CK, PD-1 and PD-L1 are all near each other and within the ~10-30 counts range (green square). The FoxP3 'positive' pixel is closer to 1 and may require additional optimization (yellow oval). The autofluorescence and DAPI 'positive' pixels are higher than the others, but do not call for concern.

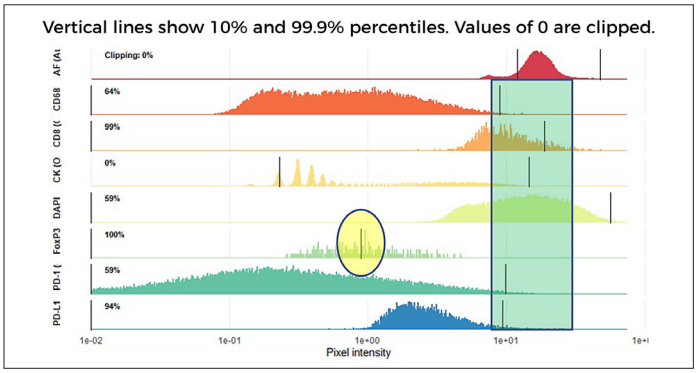


FIGURE 13: phenoptrReports signal distribution.

This can be a useful tool when used during panel development to evaluate how well-balanced your signals are across pixels in the image. This function is complementary to the "counts" tool in inForm and Phenochart softwares, which measures per pixel intensities or standardized counts.

Evaluating heterogeneity of intensity across an image set

The *Pixel intensity by component* section allows you to quickly assess if there are any outliers or large-scale differences in staining intensity across the image set. In the examples below (FIGURE 14,FIGURE 15) you can see the consistency of DAPI staining across images whereas you have more variability across samples when focusing on PD-L1.

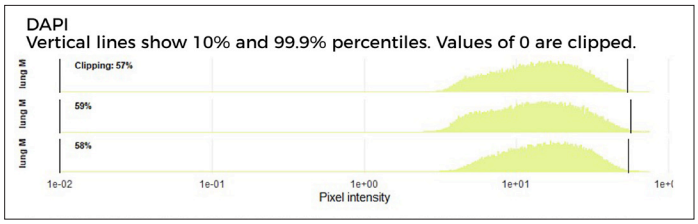


FIGURE 14: DAPI signal distribution

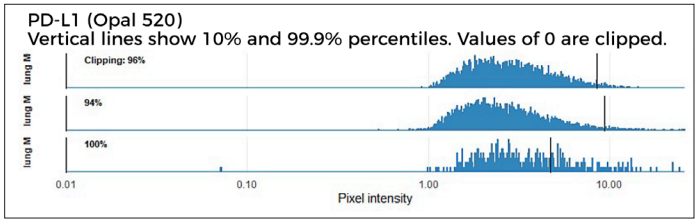


FIGURE 15: Opal520 signal distribution.

Determining spectral overlap/crosstalk

FIGURE 16 is an example of how the Pairs Plots Component is used to evaluate crosstalk between channels per ROI.

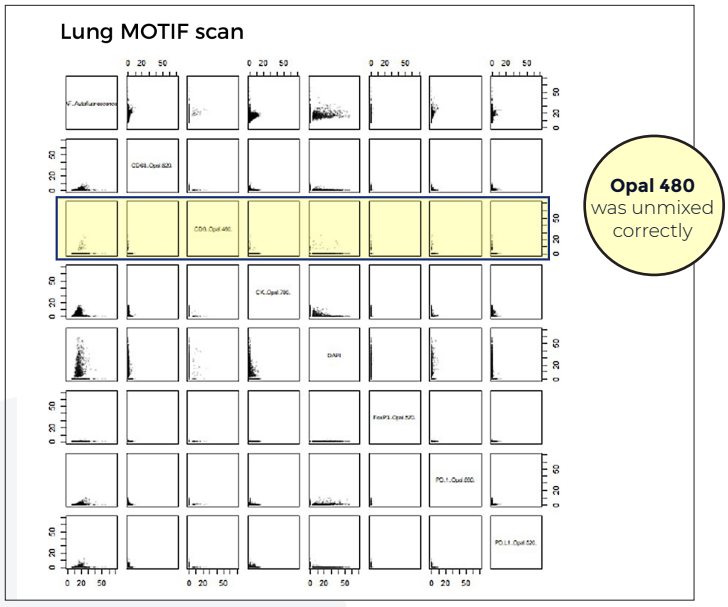


FIGURE 16: Pairs Plots Component distribution — This figure can give a quick indication on whether one Opal is bleeding into another channel after unmixing. For example, if **Opal 480** was unmixed correctly, dots along the horizontal or vertical axis (i.e. little to no correlation) would indicate clean unmixing, while dots along the diagonal (i.e. tight correlation) indicate crosstalk between channels. Here Opal 480 is cleanly unmixed from all other Opals and showing only slight overlap with DAPI and AF.

Signal-to-noise ratio

In order to generate the signal-to-noise output, you must enter multiple values into the report options box (i.e. 0.1, 0.999) to define which pixels are 'background' noise and which are 'positive' signal. If left at the default of a single value of 0.999 (FIGURE 17, without adding the noise value of 0.1), then the signal-to-noise ratio (TABLE 5) will not be generated.

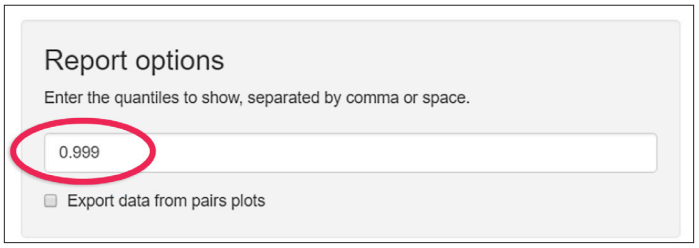


FIGURE 17: Report options.

Source	AF (Autofluorescence)	CD68 (Opal 620)	CD8 (Opal 480)	CK (Opal 780)	DAPI	FoxP3 (Opal 570)	PD-1 (Opal 690)	PD-L1 (Opal 520)
lung MOTiF_Scan1_[7456,42065]	5.5	Inf	Inf	105.9	Inf	Inf	Inf	Inf
lung MOTiF_Scan1_[9417,40618]	3.9	Inf	Inf	63.0	Inf	Inf	Inf	Inf
lung								

TABLE 5: Ratio of the 100%ile signal level to the 10%ile signal level.

TABLE 5 above shows excellent signal-to-noise for all components and accurate unmixing. This is indicative of reliable and robust data output from the automated algorithms. For Opal fluorophores, a general guideline is that ratios below 30 are considered poor signal-to-noise and can likely be improved through re-optimizing a staining or unmixing protocol.

QC REPORT 3: Mean of top 20 / bottom 10 cells

The "Mean of top 20/ bottom 10" analysis is performed on multiplexed fluorescent IHC images. This addin can be used in a few ways including evaluating the relative quality of your panel and likelihood of crosstalk between adjacent fluorophores. Input for this analysis include the merged cell seg data file and a user generated "Configuration.txt" file that includes information about all markers and Opal pairings (FIGURE 18).

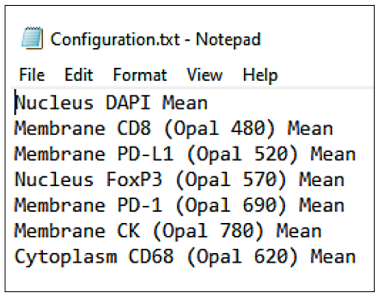


FIGURE 18: Example of a configuration file for the phenoptrReports 20/10 tool.

This analysis computes the mean fluorescence signal intensity of selected markers in the 20 highest-expressing cells and the 10% lowest-expressing cells (referred to as "bottom 10%ile cells" below). Each pixel in an unmixed, multiplex MSI image contains calibrated and precise data that indicates relative abundance of the multiple proteins of interest. Thus, the fluorescence signal intensity on a given cell is indicative of the cell's relative protein expression. The resulting Excel workbook contains a separate tab showing these metrics.

For the example data used in this application note, TABLE 6 reports the mean signal intensity of the 20 highest-expressing cells for each marker. This estimates true positive signal levels.

	MEAN EXPRESSION OF TOP 20 CELLS						
Slide ID	Mean expression of top 20 cells	Membrane CD8 (Opal 480)	Membrane PD-L1 (Opal 520)	Nucleus FoxP3 (Opal 570)	Membrane PD-1 (Opal 690)	Membrane CK (Opal 780)	Cytoplasm CD68 (Opal 620)
lung MOTIF_Scan 1	50.6099	15.0263	5.2267	2.0574	7.8507	11.3441	6.7986

TABLE 6: Mean expression of top 20 cells.

For the example data used here, TABLE 7 reports the mean signal intensity of the bottom 10%ile lowest-expressing cells for each marker. This estimates background signal levels.

	MEAN EXPRESSION OF BOTTOM 10%ILE OF CELLS						
Slide ID	Nucleus DAPI	Membrane CD8 (Opal 480)	Membrane PD-L1 (Opal 520)	Nucleus FoxP3 (Opal 570)	Membrane PD-1 (Opal 690)	Membrane CK (Opal 780)	Cytoplasm CD68 (Opal 620)
lung MOTIF_Scan 1	9.0865	0.0000	0.0000	0.0000	0.0012	0.1294	0.0000

TABLE 7: Mean expression of bottom 10%ile cells.

For the example data used here, TABLE 8 reports the ratio of mean signal intensity in the 20 highest-expressing cells to the bottom 10%ile of cells for each marker. This estimates the signal-to-background noise.

	RATIO OF MEANS, TOP 20 CELLS / BOTTOM 10%ILE CELLS						
Slide ID	Nucleus DAPI	Membrane CD8 (Opal 480)	Membrane PD-L1 (Opal 520)	Nucleus FoxP3 (Opal 570)	Membrane PD-1 (Opal 690)	Membrane CK (Opal 780)	Cytoplasm CD68 (Opal 620)
lung MOTIF_Scan 1	5.5698	inf	inf	inf	6550.5881	87.6940	inf

TABLE 8: Ratio of 20/10 means.

For the example data used here, TABLE 9 reports the ratio of mean signal intensity in the 20 highest-expressing cells for two markers in adjacent fluorophore channels. This metric highlights ratios that exceed a specified range as this would indicate potential unmixing errors or crosstalk.

	RATIO OF MEANS, TOP 20 CELLS OF ADJACENT FLUOROPHORES					
Slide ID	Nucleus DAPI / Membrane CD8 (Opal 480)	Membrane CD8 (Opal 480) / Membrane PD-L1 (Opal 520)	Membrane PD-L1 (Opal 520) / Nucleus FoxP3 (Opal 570)	Nucleus FoxP3 (Opal 570) / Membrane PD-1 (Opal 690)	Membrane PD-1 (Opal 690) / Membrane CK (Opal 780)	Membrane CK (Opal 780) / Cytoplasm CD68 (Opal 620)
lung MOTIF_Scan 1	3.3681	2.8749	2.5405	0.2621	0.6920	1.6686

TABLE 9: Ratio of means, top 20 cells of adjacent fluorophores.

Generally, the best unmixing results are obtained when the average signal intensity levels of all fluorophores (excluding DAPI) are within the range of 10-30 normalized counts. Most importantly, the ratio between neighboring fluorophores should fall between 0.3-3 for multispectral IM3 images or 0.1-10 for MOTiF images. When one fluorophore is 3x/10x (IM3/MOTiF) brighter than its neighboring fluorophore, unmixing error and crosstalk are more likely to occur.

The signal-to-noise value helps to assess background levels in relation to real, positive signal. When higher background is present, resulting in a lower signal-to-noise ratio, cell scoring based on thresholding will be less reliable and will likely result in more false positives. Phenotyping may also be more challenging. The output tables will flag signal-to-noise values of 30 or less within the Opal 780 channel as this fluorophore tends to have lower intensities with higher background levels.

This report is useful when looking at batch controls across stainings to evaluate batch errors as well as to evaluate reproducibility across batches in staining intensities, background levels, signal to noise ratios and ratios of neighboring fluorophores to confirm values fall within the expected and recommended ranges.

QC REPORT 4: Staining consistency

The staining consistency report measures variation in the mean fluorescence signal intensity (indicative of expression) of a single marker across multiple images. It is used to assess consistency of staining within or across staining runs and can be a quality control measurement for autostainers. If using, for example, CD20 stained monoplexes on serial tonsil sections in all slots on the machine, this report can confirm consistency in staining across all slides. A drop in signal within 1 or more slides may point to an autostainer issue within a specific slide position or slide tray, suggesting that autostainer maintenance is needed. When analyzing slides from multiple runs of an assay, the report can indicate changes in assay performance over time.

The inputs required to run the *Staining Consistency report are*: 1) *merged cell seg data file* 2) designated marker 3) expression compartment of interest. To generate the *merged cell seg data file*, unmixed images must run through inForm cell segmentation algorithm for nuclear, cytoplasm and membrane segmentation, then data must be merged (see *Part 2. Merge cell segmentation files*), prior to running this report.

PART 3

Instructions on how to install phenoptrReports

1. Install RStudio.
   https://rstudio.com/products/rstudio/download/#download

2. Install R 4.x.x from the official R website.
   https://cran.r-project.org/bin/windows/base/

3. Install Rtools for R v4.
   https://cran.r-project.org/bin/windows/Rtools/rtools40-x86_64.exe

4. Install phenoptr and phenoptrReports by executing the following commands from the command line. Details can be found on github:
   https://akoyabio.github.io/phenoptr/
   and
   https://akoyabio.github.io/phenoptrReports/

   install.packages("remotes")

   remotes::install_github("akoyabio/phenoptr")

   remotes::install_github("akoyabio/phenoptrReports")

   remotes::install_github("akoyabio/rtree")

Note: In some installations of RStudio, the following error occurs when trying to install phenoptr and phenoptrReports from GitHub:

"Using github PAT from envvar GITHUB_PAT

Error: Failed to install 'phenoptr' from GitHub:

HTTP error 401.

Bad credentials

[...]"

This error can be temporarily corrected using the following commands:

Sys.unsetenv("GITHUB_PAT")

Sys.getenv("GITHUB_PAT")