

Abstract

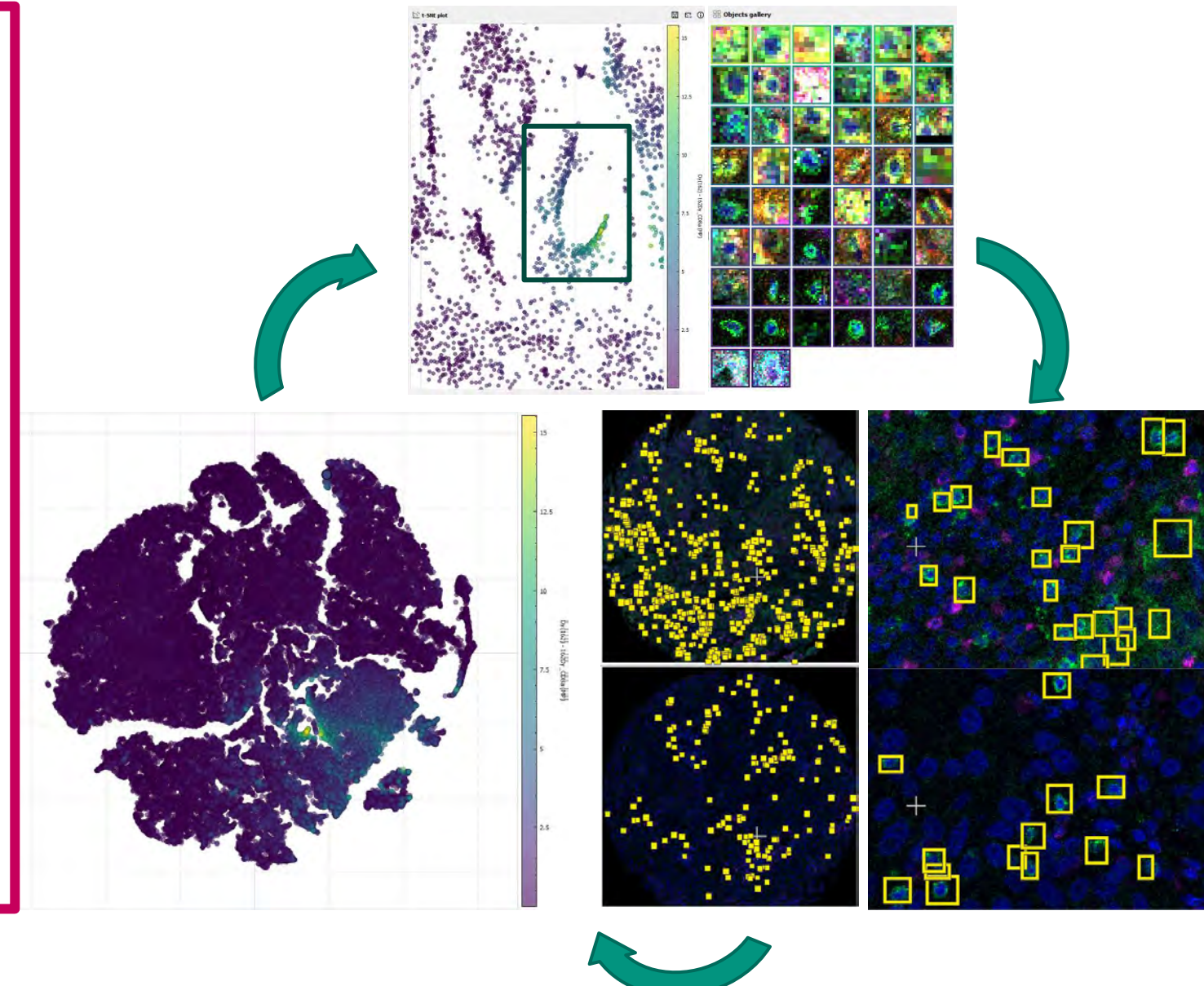
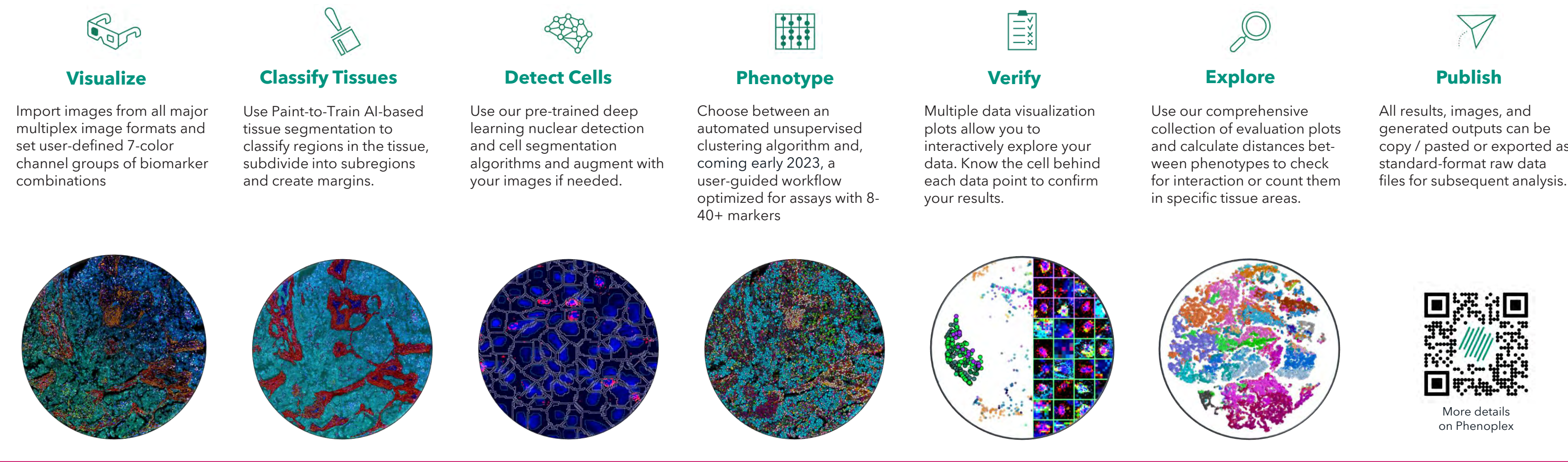
The growth in cancer immunotherapy agents requires an understanding of the immune contexture of the tumor microenvironment (TME). Understanding immune contexture requires multiplex staining, imaging, and analysis to obtain multi-marker phenotypes of specific cells and analyze their biodistribution in the TME. Imaging Mass Cytometry (IMC) is the method of choice for single-step staining and highplex imaging of FFPE tissues. FFPE tissue is autofluorescent, which limits the utility of immunofluorescence methods. Lung and colorectal tissue (and bone, skin, etc.) are highly autofluorescent, and therefore good targets for IMC. However, developments in analysis software for highplex imagery have not kept pace with imaging advances. We present a comprehensive workflow designed specifically for highplex image analysis, covering tissue segmentation, cell segmentation based on IMC DNA images, cellular phenotyping, and spatial analysis.

Lung and colorectal tissue sections with a 30-marker IMC panel of structural, tumor, stroma, immune cell, and immune activation markers were imaged (Hyperion[™], Standard BioTools). Highplex image analysis (Phenoplex[™], Visiopharm) was performed as a multi-step workflow in a single software package that includes: conversion of IMC images to pyramidal format; easy visualization methods for displaying different marker subsets; a point-to-train algorithm for tissue segmentation (into tumor, stroma, blood vessels, etc.); deep-learning-based nuclear segmentation pre-trained on IMC DNA channels; cellular phenotyping based on thresholds based on visual assessment of positivity; spatial biodistribution metrics for cell populations; and a flexible set of outputs for downstream analysis. Tissue segmentation was used to divide the tissue into tumor, stroma, and tumor margin regions, and these regions were used to compare the immune contexture through a series of t-SNE images partitioned by spatial region.

We demonstrate that a simple analysis workflow can be used for highplex images of different tissue types by users with no programming knowledge. Visualization templates for the marker subsets and the pre-trained IMC nuclear segmentation are reusable. A new tissue segmentation algorithm for each tissue type is required, as are new thresholds for biomarker positivity. Spatial biodistribution metrics, heatmaps and partitioned t-SNE plots were generated for each tissue type with a minimum of work.

Highplex IMC imaging of lung and colorectal tumor samples is a simple and effective means of obtaining high-parameter images without interfering autofluorescence. Having a comprehensive workflow for the analysis of this complex data makes obtaining useful results from highplex images more accessible to biologists and immunologists by circumventing the requirement for expert programming for each specific application.

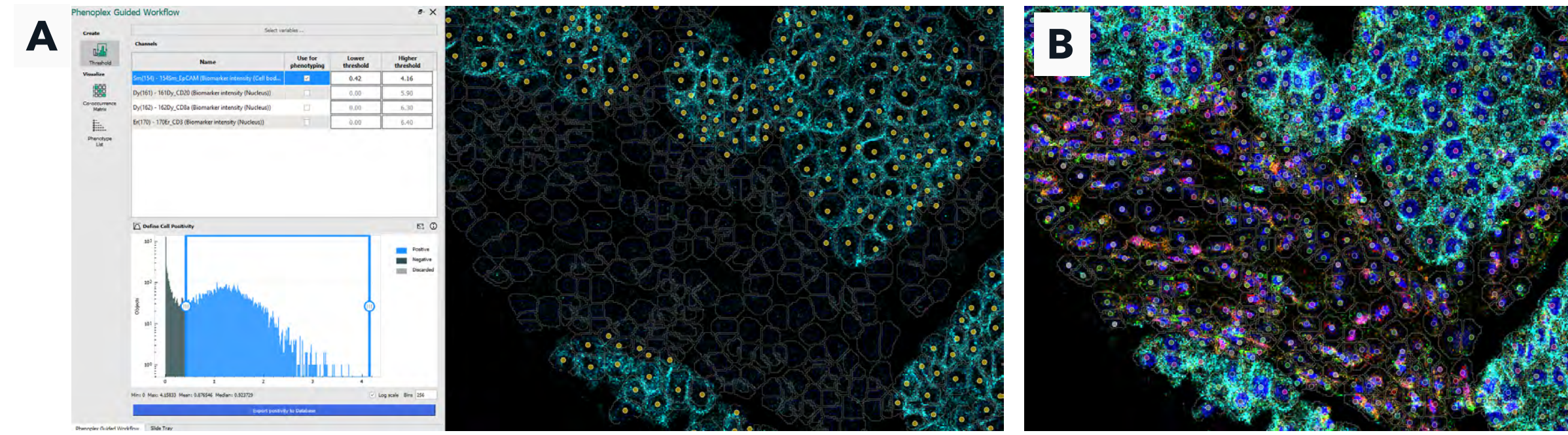
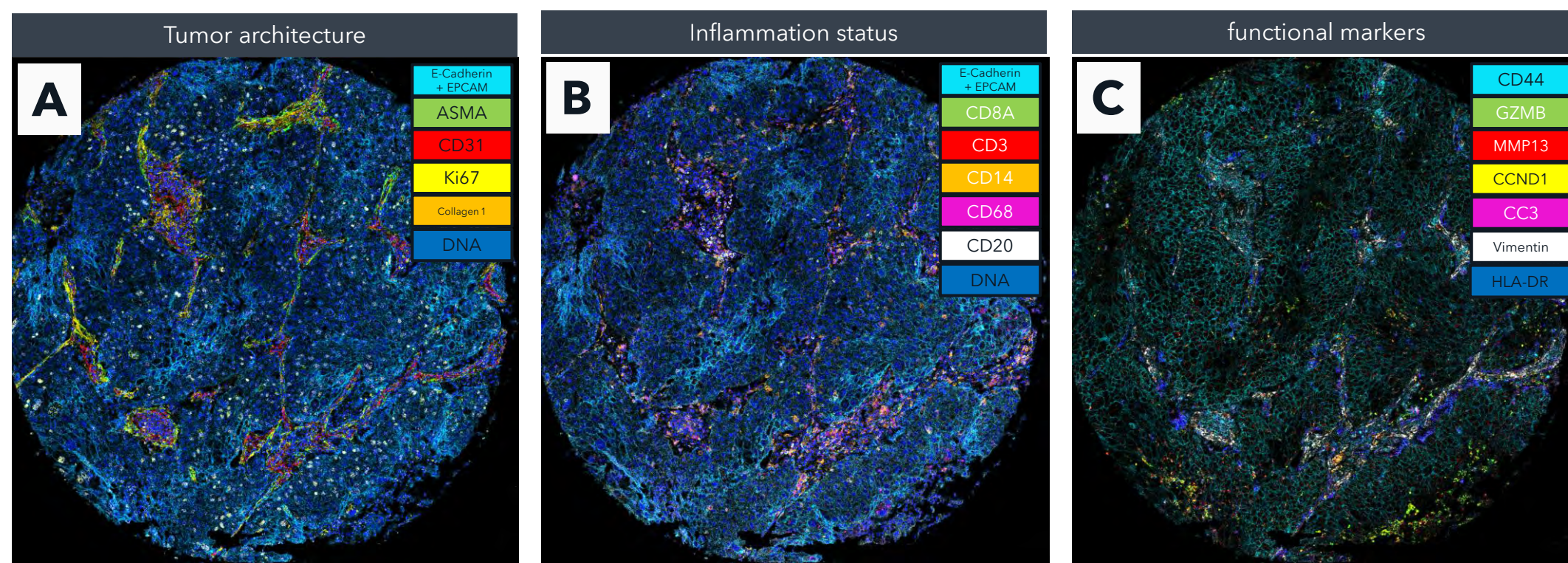
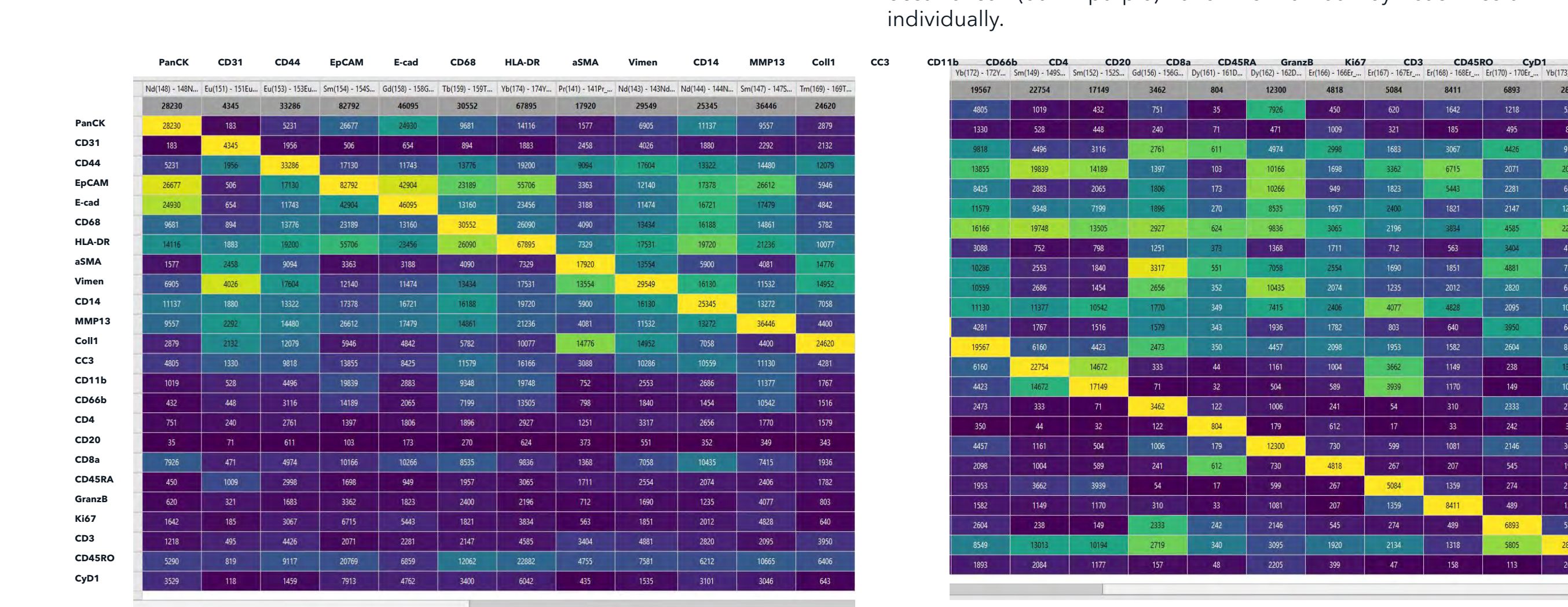
Phenoplex[™]: a new workflow for highplex image analysis



Step 5: Verify. An important task after cellular phenotyping is to be able to look at the resulting combinations of biomarkers and verify that the phenotyping results match the actual cellular expression patterns. For this verification step, we have a number of interactive plotting tools for viewing cells in the image, viewing cells in tables/graphs, and viewing a gallery of specific cells. These data all interact; cells selected in one are highlighted in the others. A biomarker co-expression table is a good way to investigate co-expressing cells.

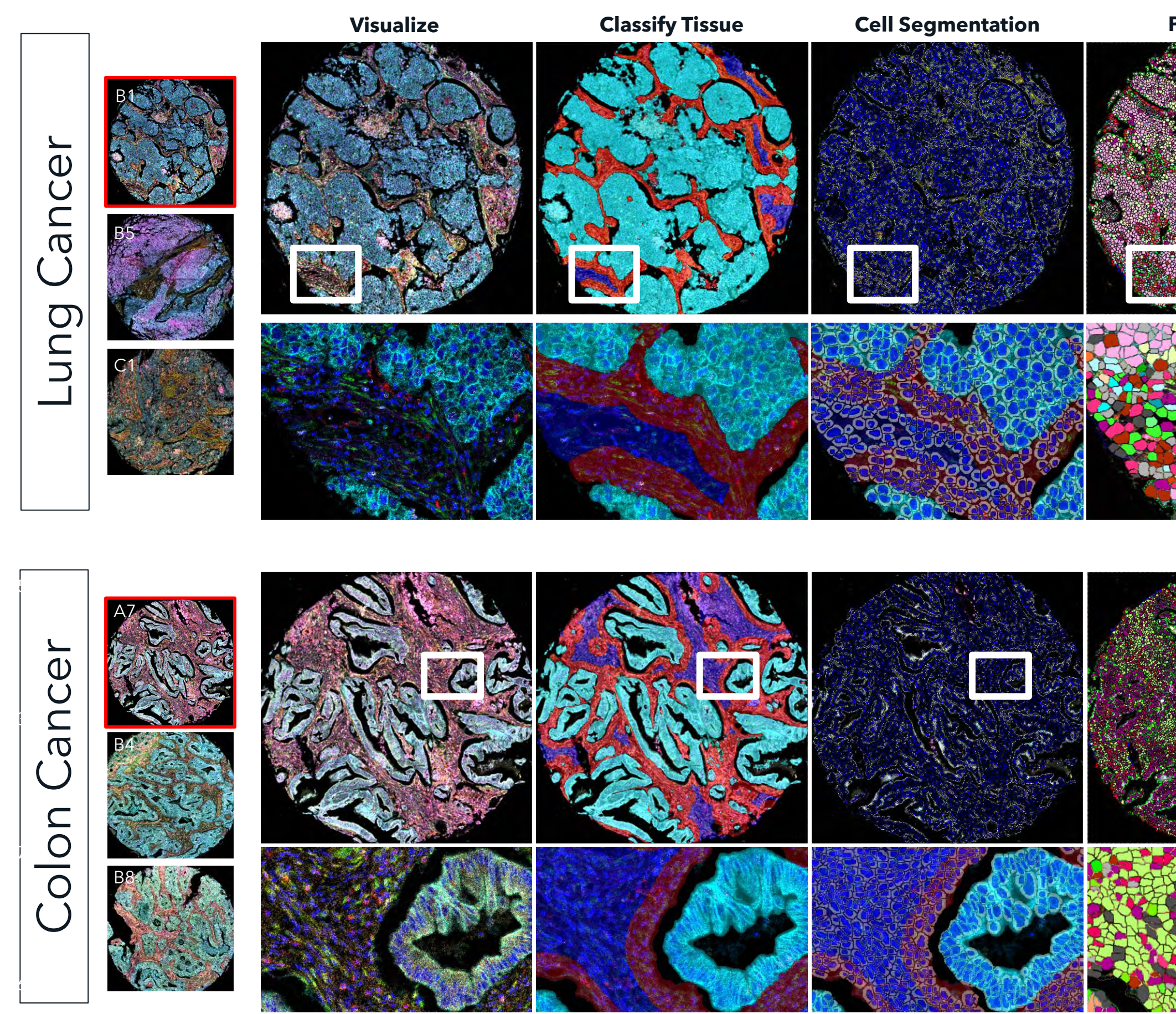
Figure 4 (left) shows a typical QC workflow using the Data Exploration tool. Users can visualize their data as t-SNE plots and color them according to the intensity of biomarkers, here CD8a, to understand clusters. Users are then able to zoom into the plots and select cell objects of interest in the plotting area. The selected objects are visualized in a cell gallery and are highlighted (yellow boxes) within the tissue samples to investigate their spatial localization within the tissue further.

Figure 5 (below) shows the co-occurrence matrix following the thresholding step (Step 4). Each column represents one biomarker for which a threshold has been set. The total cell count for that biomarker is seen in the first grey row for each column. Biomarker co-occurrences are shown in rows. The number within each table field is the number of objects that are positive for these two biomarkers (co-occurrence) of the column and the row. Each column is color-coded from high co-occurrence (yellow) to low co-occurrence (dark purple) and normalized by each column individually.

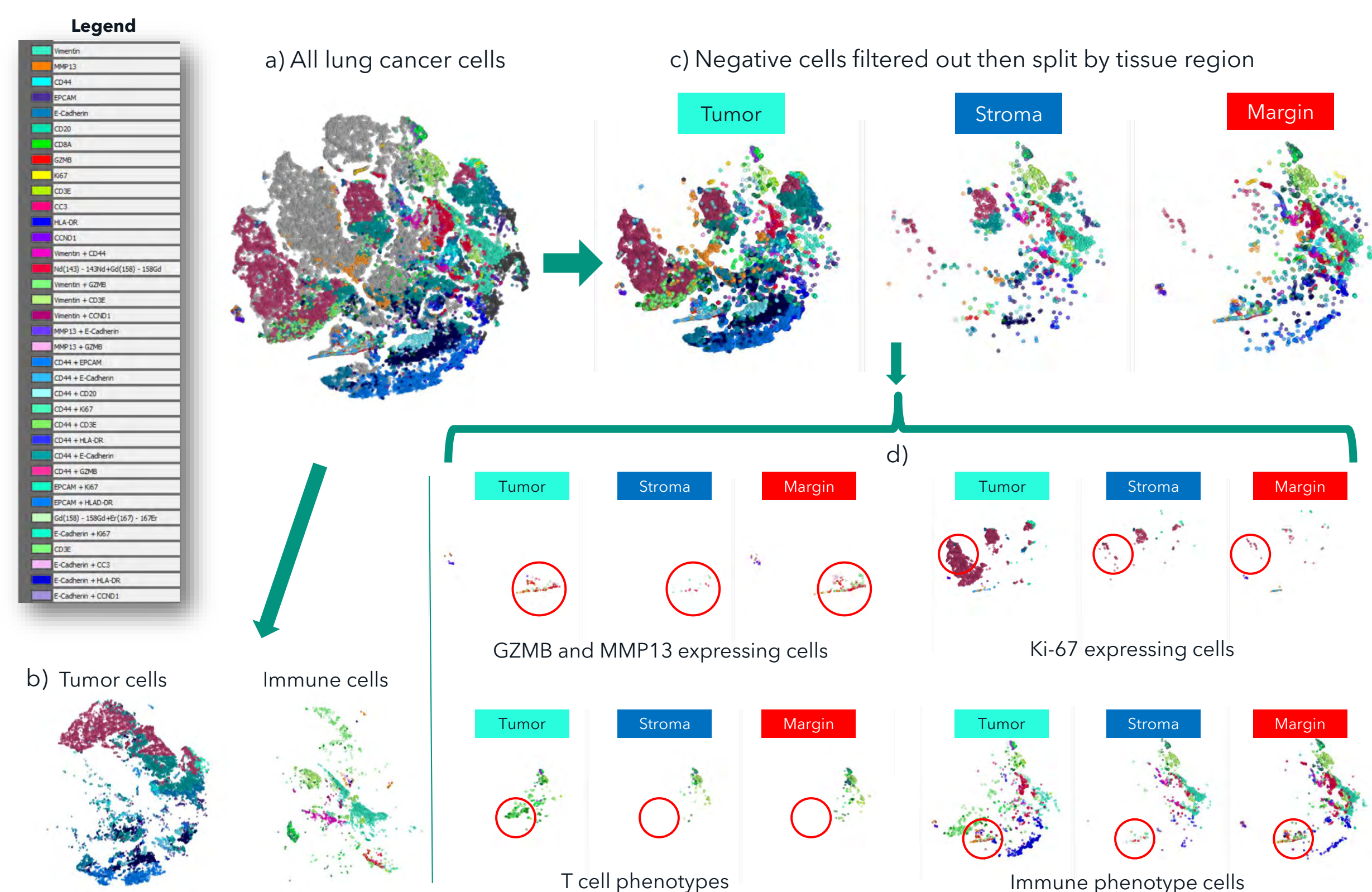


Step 2: Classify Tissue / Tissue Segmentation. One of the most critical steps. It is important to be able to count cells and events in different tissue compartments. Using a paint-to-train annotation method, a deep-learning algorithm was developed to differentiate tumor (cyan) from stroma (blue). A 50-µm margin zone (red) was denoted around the tumor region to give three morphologic regions. In addition, this step can also be used as general tissue detection step, so analysis time is not wasted on empty glass or be used to identify artifacts (folds, bubbles, dust, pen marks) to remove from analysis.

Step 3: Detect Cells / Cell Segmentation. Cell segmentation is based on pre-trained deep-learning algorithms. The software comes with three pre-trained versions: for hematoxylin (brightfield), DAPI (fluorescence) and, in this case, IMC (imaging mass cytometry). The pre-trained algorithms work very well for most cases (and that is what was used here), but for unusual cells or tissues, or when another marker is needed (CD68, the Standard BioTools IMCK cell segmentation kit, etc.), pre-trained algorithms can be augmented with new annotations from the user's imagery and the deep learning training continued until suitable results are obtained.



Step 6: Explore. Interactive plotting tools are useful for exploration, as are the many ways of developing t-SNE (t-distributed stochastic neighbor embedding) plots. The t-SNE graphs can be colored by any cellular property (phenotype, size, distance to boundaries, etc.). t-SNE plots can also be split into multiple versions by showing only subsets of the cells that make it up. The top left t-SNE (a) shows all lung cancer cells. The two t-SNE plots to the lower left show (b) only the tumor-related cells (CK+, e-cadherin+, etc.) and, at right, only the immune-related cells (CD3+, CD4+, CD20+, etc.). We can split the t-SNE according to which tissue compartment the cells were in (c: upper right). The bottom 12 t-SNE plots (d) show which cells have differential expression in the tumor, stromal and margin regions. Granzyme B and MMP13 are in the tumor/margin but not stroma. Ki-67 is primarily in the tumor. T cell and other immune cell phenotypes vary by tissue region as well.



Summary

- Non-programmers require a step-by-step analysis workflow to be able to extract relevant biological information from their spatial biology images.
- Highplex imaging using IMC is an excellent way to collect multimarker imaging data
- Phenoplex is a new workflow-based method for highplex image analysis