

Moreover, the lack of precision limits the ability to evaluate more complex properties such as immune cell distribution patterns. Therefore, there remains a need to develop ways to identify and characterize tumor-infiltrating immune cells in a biological sample.

SUMMARY

[0006] In one [aspectexample](#), this disclosure features methods of analyzing immune cell infiltration in a cancer stromal region of a biological sample (e.g., sample obtained from a subject), including: (a) identifying a cancerous region or an analyte associated with the cancerous region in the biological sample; (b) identifying a stromal region or an analyte associated with the stromal region in the biological sample; (c) identifying one or more immune cells or an analyte associated with an immune cell in one or more locations in the biological sample; and (d) using (i) the identified cancerous and stromal regions or associated analytes thereof in the biological sample and (ii) the identified one or more immune cells or associated analytes thereof to analyze immune cell infiltration in the cancer stromal region of the biological sample (e.g., sample obtained from the subject).

[0007] In some [embodimentexamples](#), the identifying the cancerous region, the identifying the stromal region, and/or the identifying immune cells includes: (a) generating a dataset from the biological sample, wherein the dataset includes one or more of: (i) analyte data for a plurality of analytes captured from a plurality of spatial locations in the biological sample; (ii) image data including images of the plurality of spatial locations of the biological sample; and (iii) registration data linking the analyte data to the image data; and (b) using the dataset to identify the cancerous region, the stromal region, and/or the immune cells in the biological sample.

[0008] In some [embodimentexamples](#), (b) includes providing the dataset to a trained machine learning module, wherein the trained machine learning module is trained at least in part from training data including reference analyte datasets from one or more reference samples, wherein the one or more reference samples include (1) one or more reference cancerous regions, (2) one or more reference stromal regions, and (3) one or more reference immune cells.

[0009] In some [embodimentexamples](#), the abundance of immune cells is determined via the trained machine learning module.

[0010] In some [embodimentexamples](#), the cancerous region includes one or more of a benign tumor, a pre-metastatic tumor, a malignant tumor, and one or more inflammatory cells.

[0011] In some [embodimentexamples](#), the stromal region includes one or more of connective tissue, blood vessels, and inflammatory cells.

[0012] In some embodimentexamples, the method further includes permeabilizing the biological sample.

[0013] In some embodimentexamples, the analyte associated with the cancerous region, an analyte associated with the stromal region, and/or an analyte associated with an immune cell is a nucleic acid. In some embodimentexamples, the nucleic acid is RNA. In some embodimentexamples, the RNA is an mRNA.

[0014] In some embodimentexamples, the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell is detected by the steps including: contacting the biological sample with a substrate including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a spatial barcode and a capture domain; hybridizing the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell to the capture probe; and determining (i) all or a part of a sequence corresponding to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof, and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the abundance and/or spatial location of the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof in the biological sample.

[0015] In some embodimentexamples, the determining step includes sequencing.

[0016] In some embodimentexamples, the analyte associated with the cancerous region, an analyte associated with the stromal region, and/or an analyte associated with an immune cell is a protein. In some embodimentexamples, the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell is detected by the steps including: attaching the biological sample with a plurality of analyte capture agents, wherein an analyte capture agent of the plurality of analyte capture agents includes: (i) an analyte binding moiety that binds specifically to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell; (ii) an analyte binding moiety barcode; and (iii) an analyte capture sequence, wherein the analyte capture sequence binds specifically to a capture domain; contacting the biological sample with a substrate, wherein the substrate includes a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes (i) the capture domain and (ii) a spatial barcode; hybridizing the analyte associated

with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell to the capture probe; and determining (i) all or a part of a sequence corresponding to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof, and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the abundance and/or spatial location of the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof in the biological sample.

[0017] In some [embodimentexamples](#), the determining step includes: sequencing (i) all or a part of a sequence corresponding to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof, and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the abundance and/or spatial location of the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof in the biological sample.

[0018] In some [embodimentexamples](#), the analyte binding moiety is an antibody or antigen-binding fragment thereof, a cell surface receptor binding molecule, a receptor ligand, a small molecule, a T-cell receptor engager, a B-cell receptor engager, a pro-body, an aptamer, a monobody, an affimer, or a darpin.

[0019] In some [embodimentexamples](#), the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell is detected using *in situ* sequencing.

[0020] In some [embodimentexamples](#), the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell is detected using an antibody.

[0021] In some [embodimentexamples](#), the method further includes contacting the biological sample with one or more stains. In some [embodimentexamples](#), the one or more stains includes hematoxylin and eosin. In some [embodimentexamples](#), the one or more stains include one or more optical labels. In some [embodimentexamples](#), the one or more optical labels are selected from the group consisting of: fluorescent, radioactive, chemiluminescent, calorimetric, or colorimetric labels.

[0022] In some [embodimentexamples](#), the method further includes identifying one or more cancerous regions in the biological sample using the one or more stains of the biological sample. In some [embodimentexamples](#), the stain is specific to a cancer marker. In some instances, the cancer marker is pancytokeratin (Pan-CK or PAN-CK).

[0023] In some [embodimentexamples](#), the method further includes identifying one or more stromal regions within the one or more cancerous regions using the one or more stains of the biological sample. In some [embodimentexamples](#), the stain is specific to a stromal marker. In some instances, the cancer marker is CD45. In some [embodimentexamples](#), the image data is generated using a method including obtaining an image of the biological sample. In some [embodimentexamples](#), the method further includes registering the image data to a spatial location. In some [embodimentexamples](#), the method further includes identifying (1) the one or more cancerous regions and/or (2) the one or more stromal regions based on the image data. In some [embodimentexamples](#), the method further includes identifying the one or more immune cells based on the image data.

[0024] In some [embodimentexamples](#), the method further includes identifying the one or more cancerous regions via the trained machine learning module. In some [embodimentexamples](#), the method further includes identifying the one or more stromal regions via the trained machine learning module. In some [embodimentexamples](#), the method further includes identifying the one or more immune cells via the trained machine learning module.

[0025] In some [embodimentexamples](#), the analysis of immune cell infiltration in the cancer stromal region of the biological sample includes determining abundance of immune cells in the cancer stromal region in the biological sample.

[0026] In some [embodimentexamples](#), identifying the one or more cancer regions includes: (i) obtaining an image and registering the image data to the spatial location, (ii) using the spatial location of the determined sequences, or (iii) obtaining an image and registering the image data to the spatial location, and using the spatial location of the determined sequences; identifying the one or more stromal regions includes: (i) obtaining an image and registering the image data to the spatial location, (ii) using the spatial location of the determined sequences, or (iii) obtaining an image and registering the image data to the spatial location, and using the spatial location of the determined sequences; and identifying the one or more immune cells or associated analytes thereof in one or more locations in the biological sample includes: (i) obtaining an image and registering the image data to the spatial location, (ii) using the spatial location of the determined sequences, or (iii) obtaining an image and

registering the image data to the spatial location, and using the spatial location of the determined sequences.

[0027] In some [embodimentexamples](#), the abundance of immune cells in the cancer stromal region is determined as a percentage of cells in the cancer stroma area that are immune cells or a percentage of area of the cancer stroma that is occupied by immune cells.

[0028] In some [embodimentexamples](#), the abundance of immune cells in the cancer stromal region is determined using the spatial location of the determined sequence of the one or more cancerous regions, one or more stromal regions, and one or more immune cells.

[0029] In some [embodimentexamples](#), the using the spatial location of the determined sequences includes determining the sequence using *in situ* sequencing. In some [embodimentexamples](#), the abundance of immune cells in the cancer stromal region is determined using segmenting and (i) obtaining an image and registering the image data to the spatial location, (ii) using the spatial location of the determined sequences, or (iii) obtaining an image and registering the image data to the spatial location, and using the spatial location of the determined sequences.

[0030] In some [embodimentexamples](#), the determining includes: (a) identifying the amount of genes associated with immune infiltrating cells compared to known housekeepers normalized by number of cells per spatial location; (b) identifying the ratio of one or more tumor infiltrating lymphocytes (TILs) to one or more tumor infiltrating B cells (TIBs); and/or (c) calculating the abundance of tumor infiltrating immune cells in the biological sample based on the percentage of spatial locations including analytes associated with an immune infiltrating cells.

[0031] In some [embodimentexamples](#), the identification of the one or more immune cells includes segmenting immune cells from the image data.

[0032] In some [embodimentexamples](#), the further includes determining a cancer prognosis based on the immune infiltration.

[0033] In some [embodimentexamples](#), the further includes scoring or determining the severity of the cancer in the subject based on the immune infiltration score.

[0034] In some [embodimentexamples](#), the determining includes identifying the ratio of one or more tumor infiltrating lymphocytes (TILs) to one or more tumor infiltrating B cells (TIBs) or one or more tumor infiltrating T cells to one or more tumor infiltrating B cells (TIBs).

[0035] In some [embodimentexamples](#), the further includes administering a therapeutic treatment (e.g., to a subject), wherein the therapeutic treatment includes surgery, chemotherapeutic agents, growth inhibitory agents, cytotoxic agents, agents used in radiation

therapy, anti-angiogenesis agents, cancer immunotherapeutic agents, apoptotic agents, antitubulin agents, or a combination thereof.

[0036] In some [embodimentexamples](#), the biological sample is obtained from a biopsy (e.g., from a subject). In some [embodimentexamples](#), the biological sample is obtained from a surgical excision (e.g., from a subject). In some [embodimentexamples](#), the biological sample is collected during an endoscopy or colonoscopy (e.g., from a subject). In some [embodimentexamples](#), the biological sample is a tissue section. In some [embodimentexamples](#), the biological sample is a tissue section on a slide. In some [embodimentexamples](#), the biological sample is a formalin-fixed, paraffin-embedded (FFPE) sample, a frozen sample, or a fresh sample. In some [embodimentexamples](#), the biological sample is an FFPE sample.

[0037] In some [embodimentexamples](#), the immune cells are selected from a B cell, a T cell, an NK cell, a monocyte, a macrophage, a neutrophil, a granulocyte, an innate lymphoid cell, or a dendritic cell, or a combination thereof.

[0038] In some [embodimentexamples](#), the analyte associated with the cancerous region is selected from an analyte from the AKT pathway, an analyte from the JAK-STAT pathway, and an analyte from the Notch pathway, or a combination thereof.

[0039] In some [embodimentexamples](#), the analyte associated with the cancerous region is selected from SCGB2A1, MKI67, BRCA1, BRCA2, PIKCD, CALML6, MYC, TP53, PALB2, RAD51, and MSH2, or a combination thereof. In some instances, the analyte associated with the cancerous region is selected from SCGB2A1, MKI67, BRCA1, BRCA2, PIK3CD, and CALML6, or a combination thereof. In some instances, the analyte associated with the cancerous region is selected from PRKCI, VTCN1, MECOM, TOP2A, SHDH, XPO1, TFRC, FUT8, SOX17, PBX1, EIF42, and WT1, or a combination thereof. In some instances, the analyte associated with the cancerous region is selected from VTCN1, MECOM, TOP2A, XPO1, FUT8, SOX17, PBX1, EIF42, and WT1, or a combination thereof. In some instances, the analyte associated with the cancerous region is TOP2A. In some instances, the analyte associated with the cancerous region is XPO1. Non-limiting examples of analytes disclosed in this paragraph can also include byproducts, precursors, and degradation products of such analytes thereof, and any combination of such analytes and byproducts, precursors, and degradation products thereof.

[0040] In some [embodimentexamples](#), the analyte associated with the stromal region is selected from VIM, EPCAM, FAP, and CDH1. In some [embodimentexamples](#), the analyte associated with the stromal region is selected from FAP, VCAN, ACTA2, and PDGFRB. In some [embodimentexamples](#), the analyte associated with an immune cell is selected from BLK, CD19,

FCRL2, MS4A1, KIAA0125, TNFRSF17, TCL1A, SPIB, PNOC, PTRPC, PRF1, GZMA, GZMB, NKG7, GZMH, KLRK1, KLRB1, KLRD1, CTSW, GNLY, CCL13, CD209, HSD11B1, LAG3, CD244, EOMES, PTGER4, CD68, CD84, CD163, MS4A4A, TPSB2, TPSAB1, CPA3, MS4A2, HDC, FPR1, SIGLEC5, CSF3R, FCAR, FCGR3B, CEACAM3, S100A12, KIR2DL3, KIR3DL1, KIR3DL2, IL21R, XCL1, XCL2, NCR1, CD6, CD3D, CD3E, SH2D1A, TRAT1, CD3G, TBX21, FOXP3, CD8A, CD8B, CD79A, CD79B, CD4, IGHA1, IGHG2, JCHAIN, IGKC, CD27, CD38, CD16, IL17RB, FANK1, CTLA4, MSR1, MRC1, NKG7, FCN1, TIGIT/LAG3. Non-limiting examples of analytes described in this paragraph can also include byproducts, precursors, and degradation products of such analytes thereof, and any combination of such analytes and byproducts, precursors, and degradation products thereof.

[0041] In some [embodimentexamples](#), the one or more immune cells is selected from: (i) a CD3⁺ and CD4⁺T cell; (ii) a CD3⁺ and CD8⁺ T cell; (iii) a regulatory T cell including one or more of: CD4, Foxp3, IL17RB, CTLA4, FANK1, HAVCR1, CD25, CTLA-4, GITR, LAG-3, and CD127; (iv) a TH1 cell including one or more of: CD4, CD3D, S100A4, IL7R, and IFNG; (v) a TH2 cell including one or more of: CD4, IL7R, ICOS, CTLA4, TNFRSF4, and TNFRS18; (vi) a TH17 cell including one or more of: CD4, CD3D, IL17A, GZMA, and S100A4; (vii) a cytotoxic T cell including one or more of: CD8, CD3D, S100A4, IFNG, GZMB, GZMA, and IL2RB; (viii) a plasma cell including: one or more JCHAIN, MZB1, IGHA1, IGHG1, and IGKC; (ix) a monocyte including CD14⁺ CD16⁻; (x) a monocyte including CD14⁻ CD16⁺; and (xi) a natural killer cell including NKG7.

[0042] In some [embodimentexamples](#), the immune infiltrating cells is a tumor infiltrating B cell (TIB). In some [embodimentexamples](#), the TIB is selected from: (i) a plasma cell including one or more of: MZB1, IGLL5, IGHA1, IGHG1, JCHAIN, IGKC, IGHA2, IGLC2, IGLV3-1, and IGLV2-14; (ii) an Ig⁺ B cells including one or more of: IGHV3-74, SOCS3, JCHAIN, and SPARC; (iii) an activated B cell including: CD79B, HMGB2, HMGB1, HMGN1, and RGS13; (iv) a B cell including one or more of: MEF2B, RGS13, and MS4A1; and (v) a B cell including CD79A and CD79B. In some [embodimentexamples](#), the immune infiltrating cells is a plasma cell including one or more of: MZB1, IGLL5, IGHA1, IGHG1, JCHAIN, IGKC, IGHA2, IGLC2, IGLV3-1, and IGLV2-14.

[0043] In another [aspectexample](#), this disclosure features methods of determining immune cell infiltration in a biological sample including one or more cancerous regions and one or more stromal regions in a subject including: (a) generating a dataset from the biological sample obtained from the subject, wherein the dataset includes: (i) analyte data for a plurality

of analytes captured from a plurality of spatial locations of the biological sample, wherein an analyte in the plurality of analytes is an analyte associated with the cancerous region, an analyte associated with the stromal region, and/or an analyte associated with an immune cell; (b) providing the dataset to a trained machine learning module, wherein the trained machine learning module includes reference analyte datasets from one or more reference samples, wherein the one or more reference samples includes (i) a cancerous region from one or more cancerous regions, (2) a stromal region from one or more stromal regions, and (3) an immune cells from one or more immune cells; and (c) determining, via the trained machine learning module, the immune cell infiltration in the biological sample obtained from the subject.

[0044] In another [aspectexample](#), this disclosure features methods of determining immune cell infiltration in a biological sample including one or more cancerous regions and one or more stromal regions including: (a) generating a dataset from the biological sample obtained from a subject, wherein the dataset includes: (i) analyte data for a plurality of analytes captured from a plurality of spatial locations of the biological sample, wherein an analyte in the plurality of analytes is an analyte associated with the cancerous region, an analyte associated with the stromal region, and/or an analyte associated with an immune cell; (ii) image data including images of the plurality of spatial locations of the biological sample; and (iii) registration data linking the analyte data to the image data; (b) providing the dataset to a trained machine learning module, wherein the trained machine learning module includes reference analyte datasets from one or more reference samples, wherein the one or more reference samples includes (i) a cancerous region from one or more cancerous regions, (2) a stromal region from one or more stromal regions, and (3) an immune cells from one or more immune cells; and (c) determining, via the trained machine learning module, the immune cell infiltration in the biological sample.

[0045] In some [embodimentexamples](#), the trained machine learning module is at least one of a supervised learning module, a semisupervised learning module, an unsupervised learning module, a regression analysis module, a reinforcement learning module, a self-learning module, a feature learning module, a sparse dictionary learning module, an anomaly detection module, a generative adversarial network, a convolutional neural network, or an association rules module.

[0046] In some [embodimentexamples](#), generating the dataset includes: contacting a biological sample (e.g., from the subject having cancer) with a substrate including a plurality of capture probes, wherein the biological sample includes (1) one or more cancerous regions, (2) one or more stromal regions, and (3) one or more tumor infiltrating immune cells, and wherein a

capture probe of the plurality of capture probes includes a spatial barcode and a capture domain; attaching an analyte from the biological sample to the capture probe; determining (i) all or a part of a sequence corresponding to the analyte, or a complement thereof, and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the spatial location and abundance of the analyte in the biological sample; and identifying a spatial location as being part of a cluster based on the determined sequences corresponding to the analytes at the spatial location and using the clusters to analyze immune cell infiltration in the cancer stroma of the subject having cancer.

[0047] In some [embodiment examples](#), a cluster one or more immune cells is identified using one of the methods selected from: nonlinear dimensionality reduction, t-distributed stochastic neighbor embedding (t-SNE), global t-distributed stochastic neighbor embedding (g-SNE), and uniform manifold approximation and projection (UMAP).

[0048] In some [embodiment examples](#), generating the dataset includes: attaching the biological sample with a plurality of analyte capture agents, wherein an analyte capture agent of the plurality of analyte capture agents includes: (i) an analyte binding moiety that binds specifically to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell; (ii) an analyte binding moiety barcode; and (iii) an analyte capture sequence, wherein the analyte capture sequence binds specifically to a capture domain; contacting the biological sample with a substrate, wherein the substrate includes a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes (i) the capture domain and (ii) a spatial barcode; hybridizing the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell to the capture probe; and determining (i) all or a part of a sequence corresponding to the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof, and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the abundance and/or spatial location of the analyte associated with the cancerous region, the analyte associated with the stromal region, and/or the analyte associated with an immune cell, or a complement thereof biological sample.

[0049] In some [embodiment examples](#), the analyte data is generated using *in situ* sequencing.

[0050] In another [aspect example](#), this disclosure features a kit including: (a) a histology stain; (b) a substrate including a plurality of capture probe, wherein an capture probe of the

plurality of capture probes includes a capture domain; and (c) instructions for performing any of the methods described herein.

[0051] In another aspectexample, this disclosure features a kit including: (a) an antibody that specifically binds to an antigen on an infiltrating immune cell; (b) a substrate including a plurality of capture probe, wherein an capture probe of the plurality of capture probes includes a capture domain; and (f) instructions for performing any of the methods described herein.

[0052] In another aspectexample, this disclosure features a kit including: (a) an antibody that specifically binds to an antigen on an infiltrating immune cell; (b) a second antibody that specifically binds to an antigen on a stromal cell; (c) a substrate including a plurality of capture probe, wherein an capture probe of the plurality of capture probes includes a capture domain; and (d) instructions for performing any of the methods described herein. In one aspect, the present disclosure refers to a kit comprising: (a) an antibody that specifically binds to an antigen in or on an infiltrating immune cell; (b) a second antibody that specifically binds to an antigen in or on a stromal cell; (c) a substrate comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises a capture domain and a spatial barcode; and (d) instructions for performing a method of analyzing tumor infiltrating immune cells in a cancer stromal region of a biological sample, the method comprising: (i) identifying a cancerous region or an analyte associated with the cancerous region in the biological sample; (ii) identifying a stromal region or an analyte associated with the stromal region in the biological sample; (iii) identifying one or more immune cells or an analyte associated with an immune cell in one or more locations in the biological sample; and (iv) using (A) the identified cancerous region or the analyte associated with the cancerous region in the biological sample, (B) the identified stromal region or the analyte associated with the stromal region in the biological sample, and (C) the identified one or more immune cells or the analyte associated with the immune cell to analyze tumor infiltrating immune cells in the cancer stromal region of the biological sample.

[0053] In another aspectexample, this disclosure features computer implemented methods, where the methods include: (a) generating a dataset of a plurality of biological samples, wherein the dataset includes, for each biological sample of the plurality of biological samples: (i) analyte data for a plurality of analytes captured at a plurality of spatial locations of a reference biological sample; (ii) image data of the reference biological sample;

and (iii) registration data of the imaged data linking to the analyte data according to the spatial locations of the reference biological sample; wherein the reference biological sample includes (1) one or more cancerous regions in the reference biological sample, (2) one or more stromal regions within the one or more cancerous regions, and (3) a plurality of tumor infiltrating lymphocytes (TILs); (b) training a machine learning module with the dataset, thereby generating a trained machine learning module; and (c) determining immune cell infiltration in a biological sample via the trained machine learning module. In one aspect, the present disclosure refers to a computer implemented method comprising: (a) generating a dataset from a plurality of biological samples, wherein the dataset comprises, for a biological sample of the plurality of biological samples: (i) analyte data for a plurality of analytes captured at a plurality of spatial locations of a reference biological sample; (ii) image data of the reference biological sample; and (iii) registration data of the imaged data linking to the analyte data according to one or more spatial location in the reference biological sample; wherein the reference biological sample comprises: (1) a cancerous region in the reference biological sample, (2) a stromal region within the cancerous region, and (3) one or more immune cells; (b) training a machine learning module with the dataset, thereby generating a trained machine learning module; and (c) determining tumor infiltrating immune cells in the biological sample using the trained machine learning module.

[0054] In another aspectexample, this disclosure features systems, where the systems include: (a) a storage element operable to store a dataset of a plurality of biological samples, wherein the dataset includes, for each biological sample: analyte data for a plurality of analytes captured at a plurality of spatial locations of a reference biological sample; image data of the biological sample; and registration data of the imaged data linking to the analyte data according to the spatial locations of the reference biological sample; wherein the biological sample includes (1) one or more cancerous regions in the reference biological sample, (2) one or more stromal regions within the one or more cancerous regions, and (3) the a plurality of tumor infiltrating lymphocytes (TILs); and (b) a processor operable to process the dataset

through a machine learning module to train the machine learning module, to determine immune cell infiltration in a biological sample. In one aspect, the present disclosure refers to a system comprising: (a) a storage element operable to store a dataset of a plurality of biological samples, wherein the dataset comprises: analyte data for a plurality of analytes captured at a plurality of spatial locations of a reference biological sample; image data of the biological sample; and registration data of the imaged data linking to the analyte data according to the spatial locations of the reference biological sample; wherein the biological sample comprises (1) a cancerous region in the reference biological sample, (2) a stromal region within the cancerous region, and (3) the a plurality of tumor infiltrating immune cells; and (b) a processor operable to process the dataset through a machine learning module to train the machine learning module, to determine immune cell infiltration in a biological sample.

[0055] All publications, patents, patent applications, and information available on the internet and mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, patent application, or item of information was specifically and individually indicated to be incorporated by reference. To the extent publications, patents, patent applications, and items of information incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

[0056] Where values are described in terms of ranges, it should be understood that the description includes the disclosure of all possible sub-ranges within such ranges, as well as specific numerical values that fall within such ranges irrespective of whether a specific numerical value or specific sub-range is expressly stated.

[0057] The term “each,” when used in reference to a collection of items, is intended to identify an individual item in the collection but does not necessarily refer to every item in the collection, unless expressly stated otherwise, or unless the context of the usage clearly indicates otherwise.

[0058] The singular form “a”, “an”, and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes one or more cells, including mixtures thereof. “A and/or B” is used herein to include all of the following alternatives: “A”, “B”, “A or B”, and “A and B”.

[0059] Various embodiment examples of the features of this disclosure are described herein. However, it should be understood that such embodiment examples are provided merely by way of example, and numerous variations, changes, and substitutions can occur to those skilled in the art without departing from the scope of this disclosure. It should also be understood that various alternatives to the specific embodiment examples described herein are also within the scope of this disclosure.

DESCRIPTION OF DRAWINGS

[0060] The following drawings illustrate certain embodiments of the features and advantages of this disclosure. These embodiments are not intended to limit the scope of the appended claims in any manner. Like reference symbols in the drawings indicate like elements.

[0061] **FIG. 1** is a schematic diagram showing an example of a barcoded capture probe.