Escalating High-dimensional Imaging using Combinatorial Channel Multiplexing and Deep Learning

Raz Ben-Uri*[1], Lior Ben Shabat*[1,2], Omer Bar-Tal[2], Yuval Bussi[1,2], Noa Maimon[1], Tal Keidar Haran[1,3], Idan Milo[1], Ofer Elhanani[1], Alexander Rochwarger[4], Christian M. Schürch[4,5], Shai Bagon[2], Leeat Keren[1]

Abstract

Understanding tissue structure and function requires tools that quantify the expression of multiple proteins at single-cell resolution while preserving spatial information. Current imaging technologies use a separate channel for each individual protein, inherently limiting their throughput and scalability. Here, we present CombPlex (COMBinatorial multiPLEXing), a combinatorial staining platform coupled with an algorithmic framework to exponentially increase the number of proteins that can be measured from C up to $2^C - 1$, and is applicable to any mass spectrometry-based or fluorescence-based microscopy platform. In CombPlex, every protein can be imaged in several channels, and every channel contains agglomerated images of several proteins. These combinatorically-compressed images are then decompressed to individual protein-images using deep learning and optimization. We perform feasibility experiments *in silico* and achieve accurate (F1=0.98, R=0.99) reconstruction for compressing the stains of twenty-two proteins to five imaging channels. We test our approach experimentally and obtain accurate (F1=0.97, R=0.93) images of seven proteins using three channels, both in fluorescence microscopy and in mass-based imaging. We demonstrate that combinatorial staining coupled with deep-learning decompression can serve to escalate the number of proteins measured using any imaging modality, without the need for specialized instrumentation. Coupling CombPlex with instruments for high-dimensional imaging could pave the way to image hundreds of proteins at single-cell resolution in intact tissue sections.

1. Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel
2. Department of Mathematics and Computer science, Weizmann Institute of Science, Rehovot Israel
3. Department of Pathology, Hadassah Medical Center, Jerusalem, Israel
4. Department of Pathology and Neuropathology, University Hospital and Comprehensive Cancer Center Tübingen, Tübingen, Germany
5. Cluster of Excellence iFIT (EXC 2180) “Image-Guided and Functionally Instructed Tumor Therapies”, University of Tübingen, Germany

* These authors contributed equally

Introduction

Immunohistochemistry, whereby antibodies are used to quantify the single-cell spatial distribution of proteins in tissue sections, is a pillar of pathology, informing patient diagnosis, clinical decision making as well as discovery studies [1]. Traditional chromogenic immunohistochemistry is limited to visualizing one or two proteins at a time, whereas standard immunofluorescence methods are limited to three to five proteins due to overlapping emission spectra [2]. However, to uncover complex cellular structures, interactions, and functions, there is a need to simultaneously visualize many proteins *in situ*. To address this gap and increase the number of proteins measured in a single tissue, various multiplexed imaging approaches have been recently developed, gaining interest from both the scientific and medical communities ([3,4,13–16,5–12] reviewed in [17,18]).

While there are many multiplexed imaging methods, differing in the antibody tagging and detection strategies used, they can be broadly grouped into two classes. Mass-based approaches use antibodies conjugated to heavy metals, which can be resolved using a mass spectrometer [16,19]. These methods require expensive instrumentation, and their multiplexing capability is limited to a few dozen proteins by the number of available metals for conjugation and detection. Cyclic fluorescence-based approaches increase the number of proteins measured by performing consecutive imaging cycles, each illuminating two or three proteins, followed by removal of the signal by chemical or enzymatic reactions [3,4,13,14,5–12]. These methods are also effectively limited since the number of imaging cycles increases linearly with the number of proteins visualized, which in turn can result in tissue deterioration and prohibitively long imaging times. To date, there remains a need to increase the multiplexing capacity in imaging for any instrument.

One common theme of existing multiplexed imaging approaches is the use of a separate channel (e.g., fluorescent color, mass tag etc.) for each individual protein, inherently limiting the scalability of these methods (**Fig. 1A**). Combinatorial approaches, whereby each target is encoded by several channels (e.g., a sequence of colors over multiple rounds), can dramatically increase throughput (**Fig. 1B**). Such methods effectively generate a target-specific barcode and have been successfully used to measure thousands of mRNAs [20,21]. One assumption that lies at the basis of these combinatorial approaches is the ability to optically resolve individual molecules ('spot counting') to identify the underlying barcodes. This assumption can be upheld for the mRNAs of many genes if using a high enough magnification (40-100X). If this assumption is invalidated and barcodes spatially overlap, it becomes challenging to differentiate between different genes [21]. Since proteins are on average 10,000-fold more abundant than mRNAs [22], the spot counting assumption is invalidated for most proteins under standard magnifications, and therefore to date combinatorial staining methods have not been applied to imaging proteins.

Mathematically, resolving the spatial distribution of P proteins from combinatorial staining using C channels $(P > C)$ while allowing spatial overlap, is underdetermined, meaning

that it has an infinite number of possible solutions (see details in the results). Work in signal processing has shown that solutions for underdetermined linear systems can be approximated if the underlying data is structured. For example, in compressed sensing [23], the assumption that the underlying data is sparse is often used to constrain the infinite plausible solution space. Previous work has leveraged the structure of gene expression modules to apply compressed sensing to increase the throughput of both single cell RNA sequencing [24] and single molecule mRNA imaging [25]. However, one limitation of compressed sensing is that it explicitly formulates the constraints on the solution space, and as such is highly sensitive to the predefined structure of the data. Thus, these methods require elaborate setups such as performing single cell RNA sequencing (scRNASeq) prior to imaging, which is expensive and often incompatible with availability of clinical material. While similar approaches have been suggested for proteins [26], in effect they have never materialized beyond simulations of cells in suspension, presumably due to the inherent limitations of the method.

In this work we describe CombPlex (COMBinatorial multiPLEXing), an experimental foundation for combinatorial staining of proteins in tissues coupled with an algorithmic framework for accurately decompressing the combined signal back to individual proteins. In CombPlex, each protein can be measured in several channels, and each acquired image presents a signal that is a combination of several proteins. CombPlex utilizes a deep neural network to learn meaningful priors on the underlying protein images, and then uses these as constraints in an optimization scheme to accurately reconstruct individual protein images from combinatorically-compressed measurements. We devise the experimental protocols to perform combinatorial staining and show that we can accurately multiplex multiple proteins on a single channel without loss of signal. We demonstrate accurate ($F1>0.97$, $R>0.93$) reconstruction when staining for seven proteins using three fluorescent channels, indicating that the approach is compatible with conventional fluorescent microscopes that are widely used in research and in the clinic. We further test it in cyclic fluorescence (CODEX) and mass-based imaging (MIBI) and show that the approach is applicable to a wide variety of platforms and specimens. Using simulations compressing 22 proteins to five imaging channels, we show that the neural network learns several degrees of structure in the data including the compression scheme, protein coexpression and spatial expression patterns. Altogether, we demonstrate how combinatorial staining coupled with deep-learning decompression can serve to escalate the number of proteins measured using any imaging modality, without the need for specialized instrumentation. Coupling CombPlex with instruments for high-dimensional imaging can pave the way to image hundreds of proteins at single-cell resolution in intact tissue sections.

Results

Combinatorial channel multiplexing escalates protein imaging

We formally define CombPlex as a method that allows to measure *P* proteins using *C* measurement channels, where *P* > *C*. A measurement channel can refer to, e.g., a particular color in fluorescent microscopy or a particular mass range in mass-based imaging. Each protein is measured using a unique combination of channels. This results in *compressed images*, in which each channel contains a combination of the signals of several proteins (**Fig. 1B,C**).

Formally, the multiplexed-image compression can be estimated as:

$$(1) \quad y = A * x$$

where y is a multiplexed-image, of dimensions height x width x channels, with C channels. A defines the compression matrix of size $C \times P$, x is the set of P single-protein images and $*$ is the convolution operator (Methods). This staining scheme exponentially increases the number of proteins that can be imaged using a fixed number of channels. By exploiting all possible combinations, we get $P = 2^C - 1$. So, for example, using three-color fluorescent microscopy one can measure seven proteins ($2^3 - 1$) as illustrated in figure 1B. A multiplexed imaging instrument with forty channels could potentially measure 10^{12} targets, orders of magnitude more than needed to measure the entire proteome. Although this is limited by experimental considerations, such as antibody availability and cost, it could drastically increase multiplexing capabilities.

Given a set of compressed images y and a pre-defined compression matrix A, we want to recover the original protein images, i.e., solve for x **(Fig. 1D)**. Altogether this defines a set of *C* linear equations with *P* variables for each pixel. Since $P > C$, this problem is underdetermined and has a degenerate solution space with an infinite number of solutions.

We reasoned that although the mathematical problem of compressing protein images is underdetermined, additional properties of the images may be leveraged to constrain the solution space and recover the true x. For example, images of proteins are often sparse, with only a minority of pixels containing a signal for any particular protein (**Fig. 1E**). Moreover, protein expression is continuous in space. If a pixel stains positive for a particular protein, it is likely that nearby pixels will also stain positive for that same protein (**Fig. 1F**). In addition, protein expression is structured, with particular proteins tending to cooccur together or be mutually exclusive. For example, pixels that stain positive for CD8, expressed on T cells, will also stain positively for CD45, expressed on all immune cells, but not vice-versa (**Fig. 1G**). Finally, different proteins vary in their staining patterns (e.g., membranous vs. nuclear) and their overall abundance (strong vs. weak) (**Fig. 1H**). The unique staining pattern of proteins enables a trained person to recognize a variety of widely used antibodies, even without labels. For example, a skilled biologist can visually separate the signals of Ki-67, a nuclear protein, from Keratins, which are expressed on the membrane, even when both are stained together in a single image (**Fig. 1I**). Therefore, it should

be equally possible to train an algorithm to do the same. We hypothesized that the various layers of structure in biological images could serve as priors to accurately reconstruct the images of single proteins from combinatorically-compressed images, even though equation (1) is underdetermined.

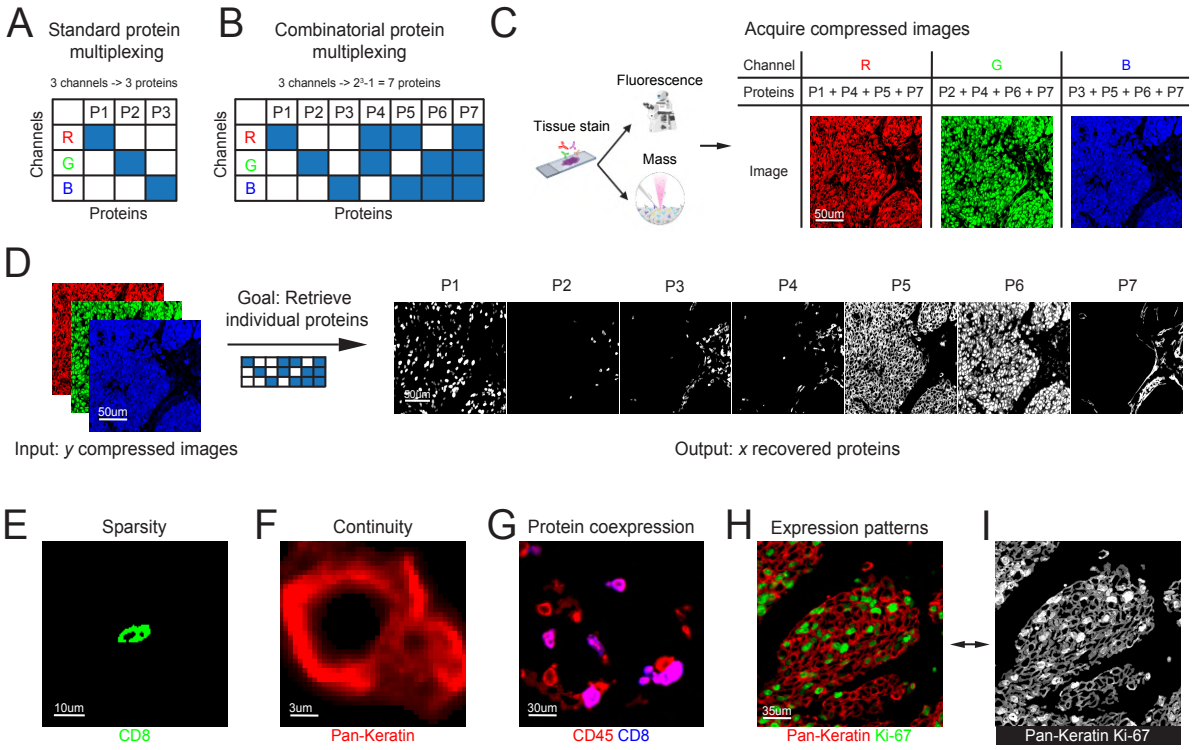

Figure 1: Combinatorial channel multiplexing escalates protein imaging

(A) In standard protein imaging, the image of each protein (columns) is acquired on a separate measurement channel (e.g. fluorescent color, rows). **(B)** In combinatorial protein multiplexing, the image of each protein (columns) is acquired on a unique combination of measurement channels (rows). This increases the number of proteins measured using a given number of channels exponentially. **(C)** Illustration of the acquisition of combinatorically-compressed images using fluorescent or mass-based approaches. Each channel in the image depicts the signal of multiple proteins. **(D)** Given compressed images acquired as detailed in (B-C) and the combinatorial staining matrix, the objective is to reconstruct the underlying single-protein measurements. **(E-I)** Examples of properties of protein-images that can serve as constraints during signal reconstruction. **(E)** CD8 exhibits sparse staining in tissues. **(F)** Keratins exhibit continuous signal. **(G)** CD8 (blue) is co-expressed with CD45 (red). **(H)** Pan-Keratin (red) exhibits membranous staining while Ki-67 exhibits nuclear staining (green). **(I)** Same as in (H). Although both proteins are shown in grayscale, they can be differentiated by their unique staining patterns.

***CombPlex,* a deep learning approach to decompress combinatorically-stained images, accurately reconstructs compressed images in simulations**

To devise a computational framework to decompress combinatorically-acquired images, we first performed *in silico* simulations. We downloaded a published dataset where 57 proteins were measured in colorectal cancer (CRC) specimens using co-detection by indexing (CODEX) cyclic fluorescence [27]. From this data, we chose 41 fields of view (FOVs) and 22 proteins that had specific staining, with good signal to noise ratio. We then simulated combinatorial measurements *in silico* by a linear combination of the protein-images according to the compression matrix illustrated in figure 2A. The matrix was designed to minimize the partition of proteins between channels and balance the number of proteins compressed in each channel (Methods). Overall, this resulted in 41 multiplexed images, each containing 5 channels, depicting the expression of 22 proteins (**Fig. 2B**, Methods).

We tested different approaches to decompress the simulated images and recover the original input images. We first tried to use compressed sensing, following literature that demonstrated success for similar problems in analysis of gene expression data (Methods [24,25]). However, using constraints for sparsity and signal continuity, compressed sensing failed to accurately recover the images of individual proteins (Median F1 scores of 0.28±0.26, **Fig. 2C,D**). We reasoned that more informative priors on the patterns and properties of protein expression could address the problem; but such priors are difficult to formulate mathematically. We therefore turned to deep learning approaches, in which a neural network learns the expected staining patterns of proteins, serving as a strong constraint for the solution space.

Next, we tried to pose the problem as a supervised learning task and use deep learning to solve it. We trained a convolutional neural network (CNN), which, given a compressed image with C channels (y), recovers the underlying single-protein images, $\hat{x}$ (Methods). We trained the network on the simulated compressed images described above. We used 33 random FOVs for training and evaluated the results on 8 images. This approach dramatically improved the overall results (F1 = 0.91±0.16, **Fig. 2E,F**), but performance varied significantly between proteins and FOVs. Investigation of the errors over training suggested that the network had a high incentive to maximize the reconstruction of highly abundant proteins. From a biologist's point of view, false predictions of signals are more crucial than the precise reconstruction of abundant signals. We therefore turned to a hybrid approach, using a CNN to identify the pixels that are positive for each protein, and then applying test-time optimization to reconstruct the actual values. The advantage of this approach is twofold: First, the network predictions are aligned with what is biologically most important, the locations of the proteins. Second, the additional optimization step to reconstruct the values gives our model the ability to adjust, at test time, to samples beyond the training set.

We designed CombPlex as a two-step process, a masking step, and an optimization step (**Fig. 2G**). In the masking step we use a CNN to recover the binary support of the individual protein images. Plainly, the goal of this stage is to determine *where* different proteins are expressed in the image, without recovering the actual values. The input to this step is a compressed image with C channels (y), and the output is P images with binary values $M_\theta(y)$. A value of 1 in a pixel in image p indicates that this pixel is expected to contain signal of protein p. To train the network we perform supervised learning using pairs of compressed images (y) and their corresponding individual-protein images (x). The goal is to minimize the binary cross-entropy between the binarized $B(x)$ and the reconstructed $M_\theta(y)$. In the second step, we perform optimization to solve equation (1) and recover the solution for x, using the binary images as constraints for the underdetermined equation system (Methods). In this step, we recover not only the areas where a signal is received, but its intensity value as well. The complete pipeline is illustrated in supplementary figure 1, and we provide details on each one of these steps in the Methods section.

We tested CombPlex on the simulated compressed images described above. We used 33 random images to train the CNN and evaluated the results on 8 images. We found that CombPlex achieved excellent results, with median F1 scores of 0.98±0.11, dramatically outperforming the reconstruction using compressed sensing or deep learning alone (**Fig. 2G,H**). CombPlex also achieved high correlation between the reconstructed values and the values measured in the ground truth images (Median Pearson R=0.99±0.06, **Fig. 2I**). The accuracy of reconstruction was moderately related to the abundance and intensity of the signal (R= 0.21, p= 0.004, **Supp. Fig. 2A,B**), suggesting improved results for highly-expressed proteins. Performance was not affected by the identity or subcellular localization of the protein, suggesting that the approach is compatible with a large variety of proteins. We reasoned that the two-step approach is advantageous in learning from prior data while still allowing flexibility to accommodate image-specific properties at test-time.

Next, we evaluated the sensitivity of CombPlex to the properties of the compression. In the simulations above 22 proteins were compressed to 5 channels, which is the minimal number of channels needed to encode 22 proteins. We tested whether relaxing the level of compression by increasing the number of channels affects CombPlex' performance. To this end, we simulated compression of the same 22 proteins to either 5, 7, 9 or 11 channels. We found that increasing the number of channels increased the accuracy of the reconstruction (**Fig. 2J**). For example, increasing the number of measured channels from 5 to 7 increased the F1 score for CD20 from 0.93 to 0.95 We conclude that there is a tradeoff between the degree of compression and accuracy of reconstruction, and that the compression can be tailored to achieve a desired level of reconstruction accuracy.

There are many options to compress 22 proteins to 5 channels. We therefore generated 10 permutations of the compression matrix A, for each one simulated compressed data and then evaluated CombPlex' performance in recovering the stains of the individual proteins. We found that CombPlex is relatively robust to the specific compression scheme used (**Fig. 2K**). Standard errors for the F1 scores of individual channels in different permutations of A ranged from 0.0004 to 0.008, similar to those obtained from retraining the network without changing A. We conclude that CombPlex is not very sensitive to the precise identity of A, allowing A to be largely determined by experimental considerations, such as availability of specific antibodies on specific channels etc.

Altogether, *in silico* simulations suggested that it is possible to compress the stains of many proteins to fewer measurement channels, and then recover the signal using deep learning and optimization.

CombPlex accurately reconstructs compressed cancer images measured by fluorescent microscopy

Following successful performance of CombPlex in simulations, we proceeded to test it experimentally. Most fluorescent imaging platforms easily allow to measure three channels, so we decided to test a compression scheme of seven proteins to three channels, which is the maximal level of compression. We chose seven target proteins with clinical relevance that are routinely used in immunostaining of tumors: Pan-Keratin, αSMA, Ki-67, p53, CD45, CD8 and HLA-DRDPDQ (HLA-II). These proteins constitute a variety of strong and weak, nuclear and membrane, and tumor and immune stains, and are a good representation of a plausible biologically-relevant immune-oncology panel. Importantly, this panel is expected to generate images with colocalization of proteins. For example, we expect to find pixels that are positive for CD45 and CD8 (cytotoxic T cells), Ki-67 and p53 (proliferating tumor cells), CD45 and HLA-II (myeloid cells) etc. Altogether, the selection of targets for the panel, and the compression scheme, utilizing all degrees of freedom in A, represent a realistic and difficult test case for the approach.

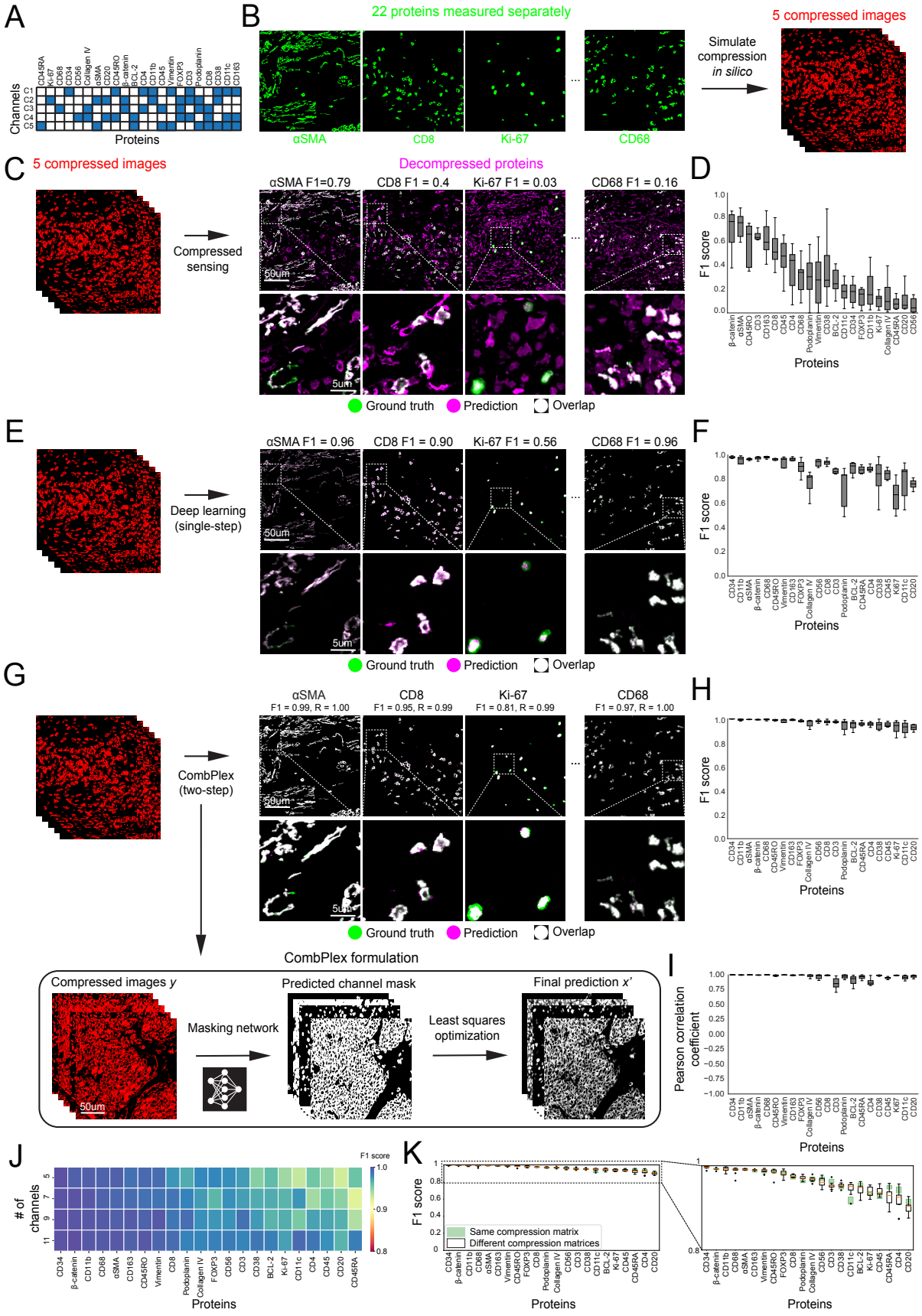
A
Channels
C1
C2
C3
C4
C5
Proteins
B
22 proteins measured separately
αSMA
CD8
Ki-67
CD68
Simulate compression in silico
5 compressed images
C
5 compressed images
Compressed sensing
Decompressed proteins
αSMA F1=0.79
CD8 F1 = 0.4
Ki-67 F1 = 0.03
CD68 F1 = 0.16
50um
5um
Ground truth
Prediction
Overlap
D
F1 score
Proteins
E
Deep learning (single-step)
αSMA F1 = 0.96
CD8 F1 = 0.90
Ki-67 F1 = 0.56
CD68 F1 = 0.96
F
G
CombPlex (two-step)
αSMA
F1 = 0.99, R = 1.00
CD8
F1 = 0.95, R = 0.99
Ki-67
F1 = 0.81, R = 0.99
CD68
F1 = 0.97, R = 1.00
H
CombPlex formulation
Compressed images y
Masking network
Predicted channel mask
Least squares optimization
Final prediction x'
I
Pearson correlation coefficient
J
# of channels
F1 score
K
Same compression matrix
Different compression matrices

Figure 2: *CombPlex,* a deep learning approach to decompress combinatorically-stained images, accurately reconstructs compressed images in simulations

(A) The compression matrix employed to condense the data from Schurch et al [1]. **(B)** *In silico* compression of 22 protein images into 5 combinatorically-compressed images. **(C)** A combinatorically-compressed image (red) is decompressed into individual protein images using compressed sensing. Accuracy is evaluated by calculating the F1 scores of the recovered images (magenta) compared to experimentally-measured single protein images (ground truth, green). White pixels were accurately recovered, whereas green and magenta indicate false negatives and false positives, respectively. **(D)** F1 scores (y-axis) for each one of the 22 proteins (x-axis) recovered by compressed sensing on eight test FOVs. **(E)** Same as (C), using a convolutional neural network (CNN) for reconstruction. The network uses MSE loss to recover the signal in a single step **(F)** F1 scores (y-axis) for each one of the 22 proteins (x-axis) recovered by a single-step CNN on the same eight FOVs shown in (C). **(G)** Same as (C), using the CombPlex algorithmic framework for reconstruction. CombPlex is designed as a two-step process, using a pre-trained convolutional neural network to output binary predictions, and then solving a least squares optimization problem using these binary masks as constraints. **(H)** F1 scores (y-axis) for each one of the 22 proteins (x-axis) recovered by CombPlex on the same eight FOVs shown in (C). **(I)** Pearson correlation coefficients (y-axis) between the ground truth images and the images recovered by CombPlex for the 22 proteins (x-axis) on the same eight FOVs shown in (C). **(J)** Simulations were performed to compress the 22 proteins (x-axis) into 5,7,9 or 11 channels (y-axis) and models were trained accordingly. Colors denote the median F1 scores across eight FOVs for the individual protein images recovered using CombPlex. Colors denote the median F1 scores across eight FOVs. **(K)** Ten models were trained using compression matrix A, while shuffling the columns to change the proteins' location in the matrix. Boxplots show the distributions of median F1 scores across eight FOVs (y-axis) for each of the 22 proteins (x-axis). Zoom (right) shows the range between 0.8-1. Green bars mark the 25th and 75th percentiles of the median F1 scores of 5 models trained with the same compression matrix.

We designed a matrix A to compress the staining of the seven target proteins (P = 7), into three images (C = 3). Experimentally, this is achieved by labeling several antibodies with the same fluorophore. We then used CO-Detection by IndEXing (CODEX) to measure on a single tissue section both the ground truth images of individual targets and the compressed images, where the same targets are measured in a combinatorial manner (**Fig. 3A**, Table S1, Methods). Altogether the experiment required ten channels: three for the compressed images and seven for the individual targets. We stained and imaged 132 breast cancer samples, from 66 different patients (**Supp. Fig. 3A**). We performed several analyses to gauge our protocol. First, we evaluated that the compressed images indeed contained signals overlapping with the expected proteins and did not contain signals of other proteins in the panel. For example, the ATTO-550 channel contained signal matching the single-channel images of Keratin, Ki-67, αSMA and HLA-II, but did not contain signal matching p53, CD8 or CD45 (**Fig. 3B, Supp. Fig. 3B**). Moreover, pixel values in the compressed images were well-correlated with a simulation of a compressed channel, generated by overlaying the corresponding images in which the proteins were measured individually (Methods**, Supp. Fig. 3C,D**). We noticed that pixels that contained signal of more than one protein were slightly lower than expected by the simulation (**Supp. Fig. 3D**). This suggests that there may be a physical limit to the number of antibodies that one can effectively multiplex

on a single channel. However, in the current compression scheme of multiplexing four proteins per channel, it was not significant. Altogether, we concluded that it is experimentally possible to multiplex the stains of several antibodies on a single measurement channel.

We evaluated CombPlex' performance in recovering the signal of the seven individual proteins from the compressed images. We trained the network on pairs of compressed images (y) and their corresponding individual-protein images (x). We used 100 paired images for training and evaluated performance on 32 held-out images. We found that CombPlex achieved very good reconstruction, with median F1 scores on the test data of 0.97±0.07 across channels, and Pearson correlation coefficients of 0.93±0.08 (**Fig. 3C-E**).

Next, we wondered whether we could circumvent the need to obtain training data in the form of compressed measurements acquired on the same tissue as the ground-truth single-protein measurements. Removing this requirement would be beneficial in making the approach more accessible, since pairs of compressed and single-protein images are not widely available and would have to be generated for each set of proteins and the specific compression matrix A. Simulating compressed images from widely-available sets of single-protein measurements would negate the need to acquire such data, and would also allow flexibility by enabling to use the same training data for different compression schemes.

We therefore devised a scheme to simulate compressed images from measurements of individual proteins (x) according to equation (1). For these simulations, we performed augmentations that are relevant to the domain measurement system. For example, we added noise to the compression matrix (A) to account for experimental differences in intensities between fluorescent channels. Additional augmentations are detailed in the methods section *Simulation of compressed images*. Overall, we found good agreement between real compressed images and our simulations (0.95<R<1.02, **Supp. Fig. 3E**). We then trained CombPlex again, this time on pairs of individual protein images (x) and simulated compressed images (y'). We then evaluated performance on the test set on real, experimentally-compressed images. We found that even though it was trained on simulations, CombPlex achieved very good reconstruction, with median F1 scores on the test data of 0.96±0.1 across channels, and Pearson correlation coefficients of 0.93±0.08, largely comparable to the results achieved when training on real compressed images (**Fig. 3F-G**). We conclude that it is possible to train CombPlex on simulations of compressed images generated from images in which the targets were measured individually, which is valuable for the utility and applicability of the approach.

Although CombPlex was overall highly accurate, we focused on the errors in prediction to better understand their properties and deduce the strengths and limitations of the approach. We separated the errors in prediction according to their type: False Negatives (FN) are pixels that contain staining for the protein in the ground truth images, but CombPlex fails to capture it, whereas False Positives (FP) are pixels in which CombPlex predicts signal that is absent in the

ground truth images **(Supp. Fig. 3F)**. Visual inspection of the images suggested that FNs and FPs tended to reside on the outskirts of real signal, correctly predicted by CombPlex (True Positives, TP). To test this more rigorously, for all channels across all test images we performed a 2-pixel dilation around TP pixels **(Fig. 3H)**. We found that the majority of errors, both FN and FP, were located within a 2-pixel radius from accurately-detected signal **(**Fraction = 0.66±0.18, **Fig. 3I)**. Moreover, FN errors were more prevalent in pixels with lower intensities as expected from the distribution of their corresponding stains (**Supp. Fig. 4,** Methods). Taken together, we deduce that CombPlex' errors mostly generate pixels with low intensity in the vicinity of a true signal.

We hypothesized that the errors made by CombPlex would not materially change the expression patterns of the proteins in the cells and would have minimal downstream effects, if any, on commonly-performed tasks such as segmentation and cell classification [27,28]. To test this explicitly, we segmented our images using Mesmer [29], and classified each protein in each cell as positive or negative. We then compared the cell classifications between the ground truth images and the images reconstructed by CombPlex (Methods). Overall, we found excellent agreement across all proteins (Balanced Accuracy = 0.98±0.009, **Fig. 3J-L, Supp. Fig. 5**). This suggests that even in cases in which the reconstruction is not perfect, it is robust enough to enable accurate routine analyses. We conclude that CombPlex achieved accurate recovery of individual protein signals of seven proteins measured using three fluorescent measurement channels.

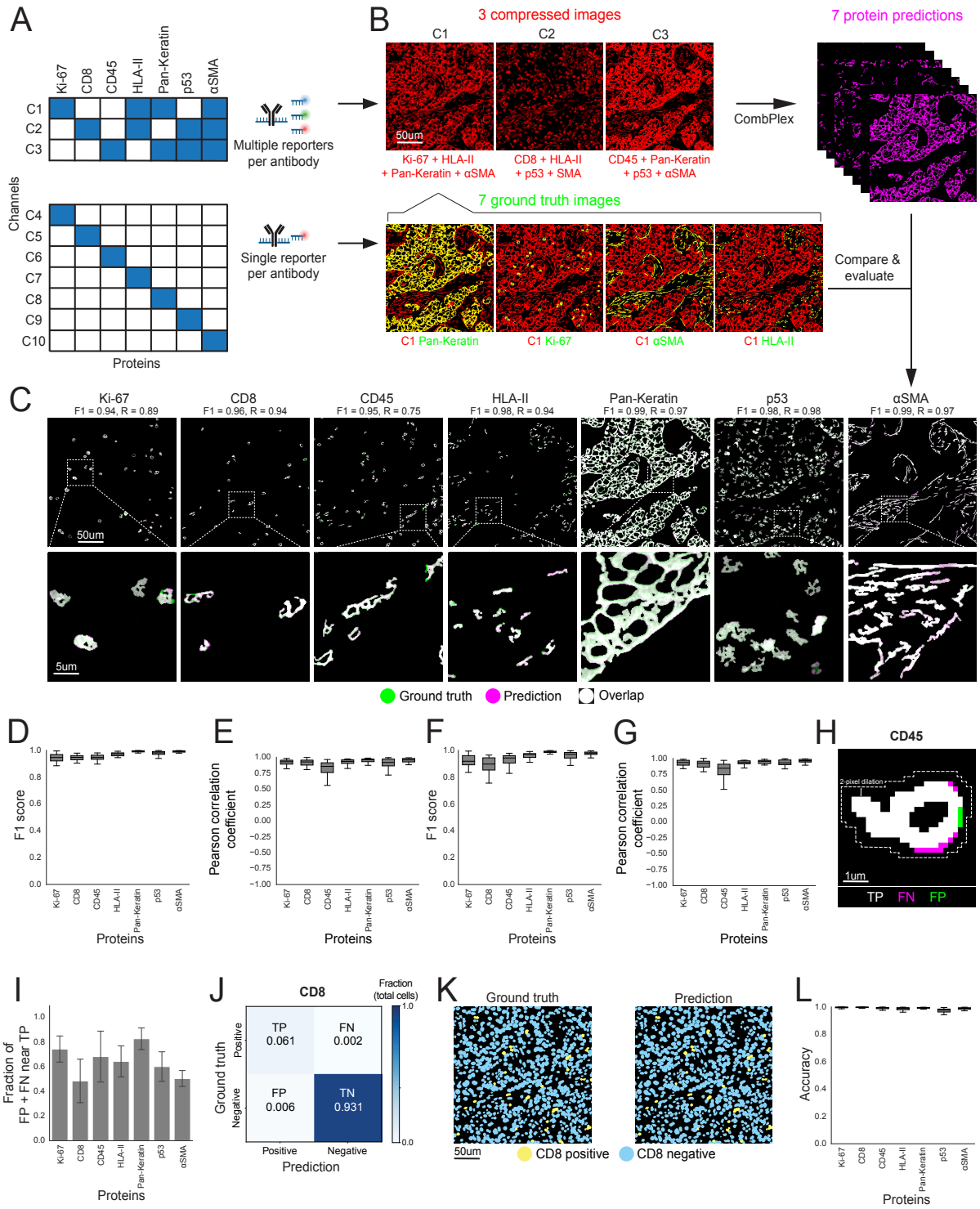

Figure 3: CombPlex accurately reconstructs compressed cancer images measured by fluorescent microscopy

(A) Experimental layout. CODEX cyclic imaging was used to image seven proteins on a single tissue section, both compressed to three measurement channels using multiple reporters per antibody (top) and individually, using a single reporter per antibody (bottom). **(B)** Top: For a single FOV, shown are the three

combinatorically-compressed images (red), decompressed by CombPlex to seven reconstructed single-protein images (magenta). Bottom: overlay of four experimentally-measured single-protein images (ground truth, green) with the compressed image that contains all of their signals (red). Yellow pixels indicate overlap. **(C)** CombPlex reconstruction (magenta) is overlaid with the ground truth single-protein images (green). White pixels indicate overlap. **(D)** F1 scores (y-axis) for each protein (x-axis) recovered by CombPlex on 32 test FOVs. The model was trained on experimentally-compressed images. **(E**) Same as (D), showing the Pearson correlation coefficients between the reconstructed and ground truth images. **(F)** Same as (D). The model was trained on simulated compressed images. **(G)** Same as (E). The model was trained on simulated compressed images. **(H-I)** For each protein in the 32 test FOVs, each pixel was compared between the ground truth images and the images reconstructed by CombPlex. Pixels were classified as TP (True Positives), FN (False Negatives), FP (False Positives) and TN (True Negatives). To evaluate the fraction of FP and FN pixels in the vicinity of TP signal, areas of TP signal were dilated by two pixels. The number of FP and FN pixels inside these 2-pixel dilations was divided by the total number of FP and FN pixels in the image to calculate their fraction (y-axis) for each protein (x-axis). **(J)** Cells were classified as positive or negative for CD8 based on either the ground truth or the reconstructed image. Shown is the confusion between both classification schemes across 32 test FOVs. **(K)** Cells classified as positive (yellow) and negative (cyan) for CD8 using ground truth (left) or reconstructed (right) images. **(L)** Segmented cells were classified as positive or negative for each protein based on either the ground truth or the images reconstructed by CombPlex. Shown are the accuracy scores (y-axis) for each protein (x-axis).

CombPlex accurately reconstructs compressed images measured by mass-spec imaging

Widely used multiplexed imaging approaches can be broadly grouped into two classes: fluorescence microscopy and mass-based imaging. We therefore tested CombPlex' compatibility with Multiplexed Ion Beam Imaging (MIBI), a mass-based imaging modality. We devised a method for attaching multiple metals to one antibody by loading polymers with pre-defined metal ratios and then conjugating them to the desired antibodies (methods, **Supp. Fig. 6A**). We measured on a single tissue section both the ground truth images of individual proteins and the compressed images, where the same targets were measured in a combinatorial manner (**Fig. 4A**). We stained and imaged 19 samples from both breast cancer and lung cancer. We evaluated that the compressed images indeed contained signals overlapping with the expected proteins, and did not contain signals of other proteins in the panel. For example, m/z 144 contained signal matching the single-channel images of αSMA, Keratin, Ki-67 and HLA-II, but did not contain signal matching CD8 or CD45 (**Fig. 4B, Supp. Fig. 6B**). Altogether, we concluded that it is experimentally possible to multiplex the stain of several antibodies on a single measurement channel also in mass-based imaging.

Encouraged by the results of using simulations to train the network in the fluorescence data, we further evaluated whether pre-existing data, not necessarily measured in the same experiment or in the same tissue of origin, could be used to train the network. We compiled a dataset from pre-existing images in the lab in which these proteins were measured individually. We included 143 images from breast cancer, lung cancer and melanoma. Since the melanoma

data did not contain staining for Keratin, we interchanged the Keratin images in this data for MelanA, reasoning that both stain tumor cells and the stains are visually similar (Methods, **Supp. Fig. 6C**). The images for training were acquired in different experiments, conducted at different times, by different students, using different antibody batches and different MIBI-TOF instruments.

We evaluated CombPlex' performance in recovering the signal of the individual proteins from the compressed images. We used the 143 simulated-compressed images for training and evaluated performance on the 19 experimentally-compressed images. We found that CombPlex achieved very good reconstruction, with median F1 scores on the test data of 0.92±0.15 across channels and median Pearson correlation coefficients of 0.91±0.08 (**Fig. 4C-E**). Next, we evaluated the data generated by CombPlex in cell segmentation and classification. We segmented our images using Mesmer [29], and classified each protein in each cell as positive or negative. We compared the cell classifications between the ground truth images and the images reconstructed by CombPlex (Methods), and found high agreement across all proteins (Balanced Accuracy = 0.97±0.01, **Fig. 4F-H, Supp. Fig. 7**). Similar to the results for the fluorescent data, this suggests that data generated by CombPlex yields cell-level classifications that are highly comparable to those obtained when proteins are measured individually.

Overall, our results indicate that CombPlex performs well also in mass-based imaging and that diverse pre-existing data can be used to train the algorithm. This greatly enhances the useability of the approach, negating the need to acquire data specifically for the purpose of training. Moreover, these results suggest that CombPlex is able to generalize between different tissues of origin and one model can be trained for a variety of pathologies.

Features contributing to CombPlex performance

In CombPlex, we use a neural network to constrain the solution space of an underdetermined equation system. We explored what does the network learn to aid it in its classification by performing various perturbations to the input images and training scheme. First, we evaluated whether the network learns the compression matrix A, since it was not directly supplied to the network during the training process. We used the published CRC CODEX dataset and the network trained on the A matrix in figure 2A. We then tested the network's performance on simulated compressed images that were generated by a different compression matrix A' (**Fig. 5A,** Methods). We found that the network's performance dropped by 98% compared to the original images (0.015±0.09 compared with 0.97±0.11, **Fig. 5B**). This indicates that the network indeed learns the compression scheme encoded by the matrix used for training.

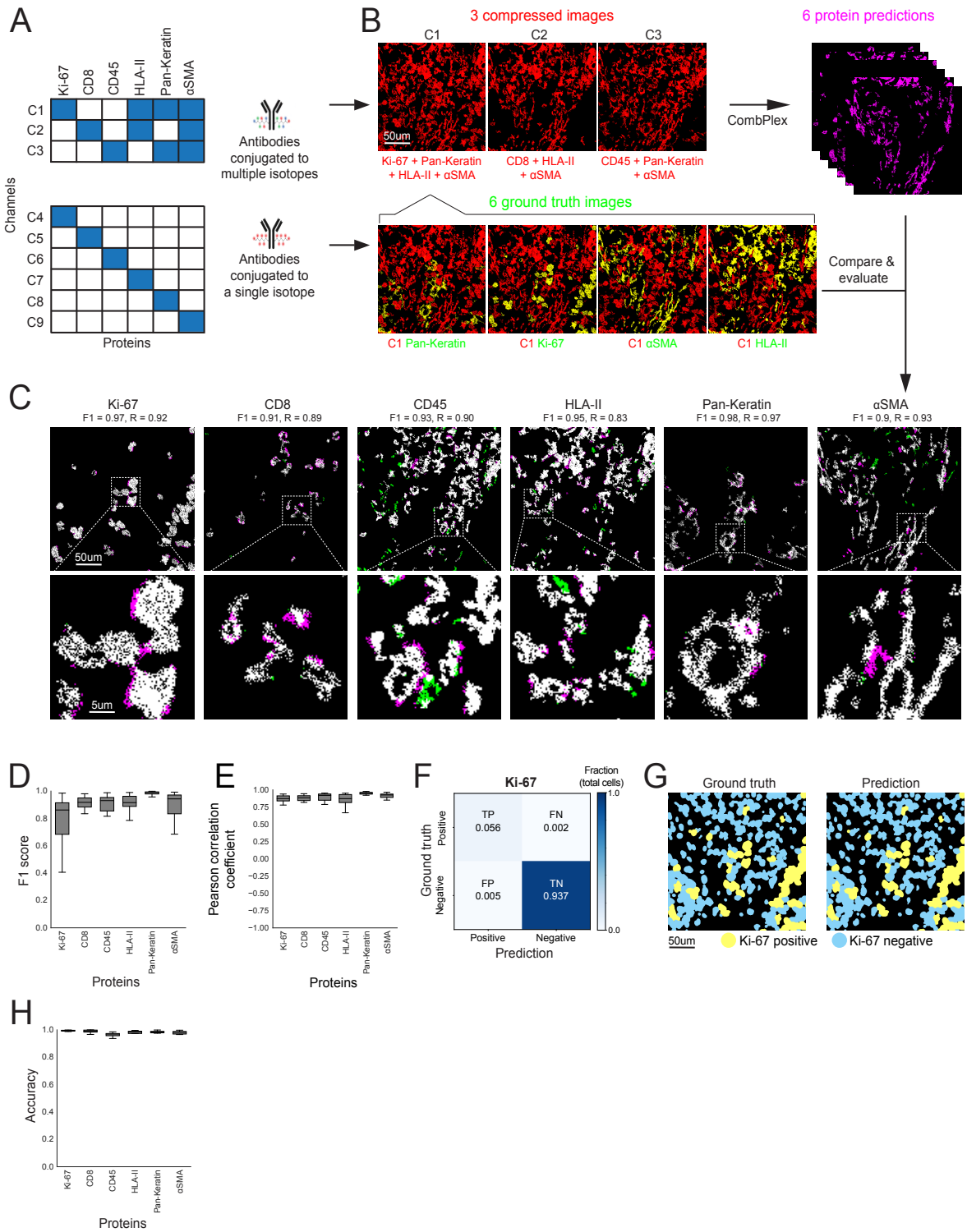

Figure 4: CombPlex accurately reconstructs compressed images measured by mass-spec imaging

(A) Experimental layout. MIBI-TOF was used to image six proteins on a single tissue section, both compressed to three measurement channels using multiple metal isotopes per antibody (top) and individually, using a single metal isotope per antibody (bottom). **(B)** Top: For a single FOV, shown are the three combinatorically-compressed images

(red), decompressed by CombPlex to six reconstructed single-protein images (magenta). Bottom: overlay of four experimentally-measured single-protein images (ground truth, green) with the compressed image that contains all of their signals (red). Yellow pixels indicate overlap. **(C)** CombPlex reconstruction (magenta) is overlaid with the ground truth single-protein images (green). White pixels indicate overlap. **(D)** F1 scores (y-axis) for each protein (x-axis) recovered by CombPlex on 19 test FOVs. The model was trained on compressed images generated by simulations on previously-available data. **(E)** Same as (D), showing the Pearson correlation coefficients between the reconstructed and ground truth images. **(F)** Cells were classified as positive or negative for Ki-67 based on either the ground truth or reconstructed single-protein image. Shown is the confusion between both classification schemes across 19 test FOVs. TP: True Positive, FN: False Negative, FP: False Positive and TN: True Negative. **(G)** Cells classified as positive (yellow) and negative (cyan) for Ki-67 using ground truth (left) or reconstructed (right) images. **(G)** Segmented cells were classified as positive or negative for each protein based on either the ground truth or the images reconstructed by CombPlex. Shown are the accuracy scores (y-axis) for each protein (x-axis).

Next, we explored which features of the data aided the network in converging to the correct solution. First, we evaluated the effect of spatial continuity. We took our test images from the CRC dataset and shuffled the pixels along the *x* and *y* dimensions. This generates images that overall have the same pixel values in each channel, but the spatial continuity of stains is disrupted (**Fig. 5C**). We found that when the size of the shuffled block is large enough to maintain single-cell morphology (>=16^2 pixels) the network manages to produce accurate predictions (Median F1=0.96±0.12). These drop dramatically when the size of the shuffled block decreases to one pixel (Median F1=0.24±0.22), indicating that the network learns spatial continuity (**Fig. 5D**). Next, we tested whether CombPlex learns priors on the staining properties of different proteins. To this end, we generated a toy example in which we generated a set of compressed images of two proteins, e.g., Ki-67 and αSMA (**Fig. 5E**). These images did not contain any additional information regarding coexpression of proteins, so to decompress them back to their single-protein images, the network has to learn their corresponding staining patterns. We trained CombPlex on 33 images and tested performance on 8 images. CombPlex demonstrated high recovery for both proteins (Median F1 = 0.99±0.005, **Fig. 5F,G**), indicating that it is able to learn to differentiate between the staining patterns of the different proteins.

Since CombPlex is based on supervised learning, we evaluated how robust the approach is to unexpected changes in the test data. First, we simulated an experimentally-plausible situation in which one antibody doesn't work. In current staining schemes, if one antibody doesn't work, only its cognate stain is affected. We were concerned that in combinatorial staining, if one antibody doesn't work then it may ruin the reconstruction of other antibodies. To this end, we used our CRC model, trained on a compression of 22 proteins to 5 channels. We then evaluated its performance on the simulated test data, each time setting the images of one protein to zero (**Fig. 5H**). Reassuringly, the F1 scores for the other proteins were not affected by this perturbation (Differences in F1 were 0.001±0.003 **Fig. 5I**). Encouraged by these results, we similarly simulated scenarios in which four or nine antibodies out of the 22 were missing, and found no reduction in the F1 scores for the remaining proteins (**Fig. 5H,I**). These results highlight

the robustness of the approach. Moreover, they suggest that a large model, trained on many proteins, could be used to decompress images that were combinatorically stained with only a subset of these proteins. This adds to the modularity of the approach, suggesting that a large pre-trained foundation model could accommodate the decompression of many different staining schemes visualizing different subsets of proteins.

Finally, we evaluated whether the network learns rules of coexpression between proteins, and whether it can accommodate biological deviations from these expected patterns. To this end, we searched in our data for biological differences in protein co-expression patterns. One such example was the expression of HLA-II on malignant cells. HLA Class II molecules are a class of major histocompatibility complex (MHC) proteins, normally found only on professional antigen-presenting cells [30]. However, in tumors, they are sometimes expressed by malignant epithelial cells [31]. In our data, we had images from patients in which the epithelial cells either expressed or did not express HLA-II (**Fig. 5J**). To test whether the network learns to either expect or not expect this co-expression we trained two models. A baseline model A was trained on images of tumors both positive and negative for HLA-II whereas model B was trained only on images in which the tumor cells were negative for HLA-II. The performance of both models was then evaluated on 11 test images in which the tumor cells were positive for HLA-II (**Fig. 5J**). This scheme simulates a situation in which the test set contains novel biology, not exhibited during training. We found that both models were equally successful at reconstructing HLA-II on antigen presenting cells, but there was a 2-9% reduction in performance in reconstructing HLA-II expression on tumor cells when such cells were not presented at all in the training data (**Fig. 5K**). We conclude that CombPlex can generalize to novel biology beyond the training set, but these capabilities are limited and, like in any supervised learning approach, ideally the training set should cover the full diversity of the domain.

Overall, we demonstrated that the network learns the compression matrix, spatial continuity of signals, the staining patterns of proteins and their co-expression, and leverages these properties in order to decompress combinatorically-multiplexed images. It can easily generalize to accommodate missing channels, but care should be taken with biological deviations that are not present in the training set.

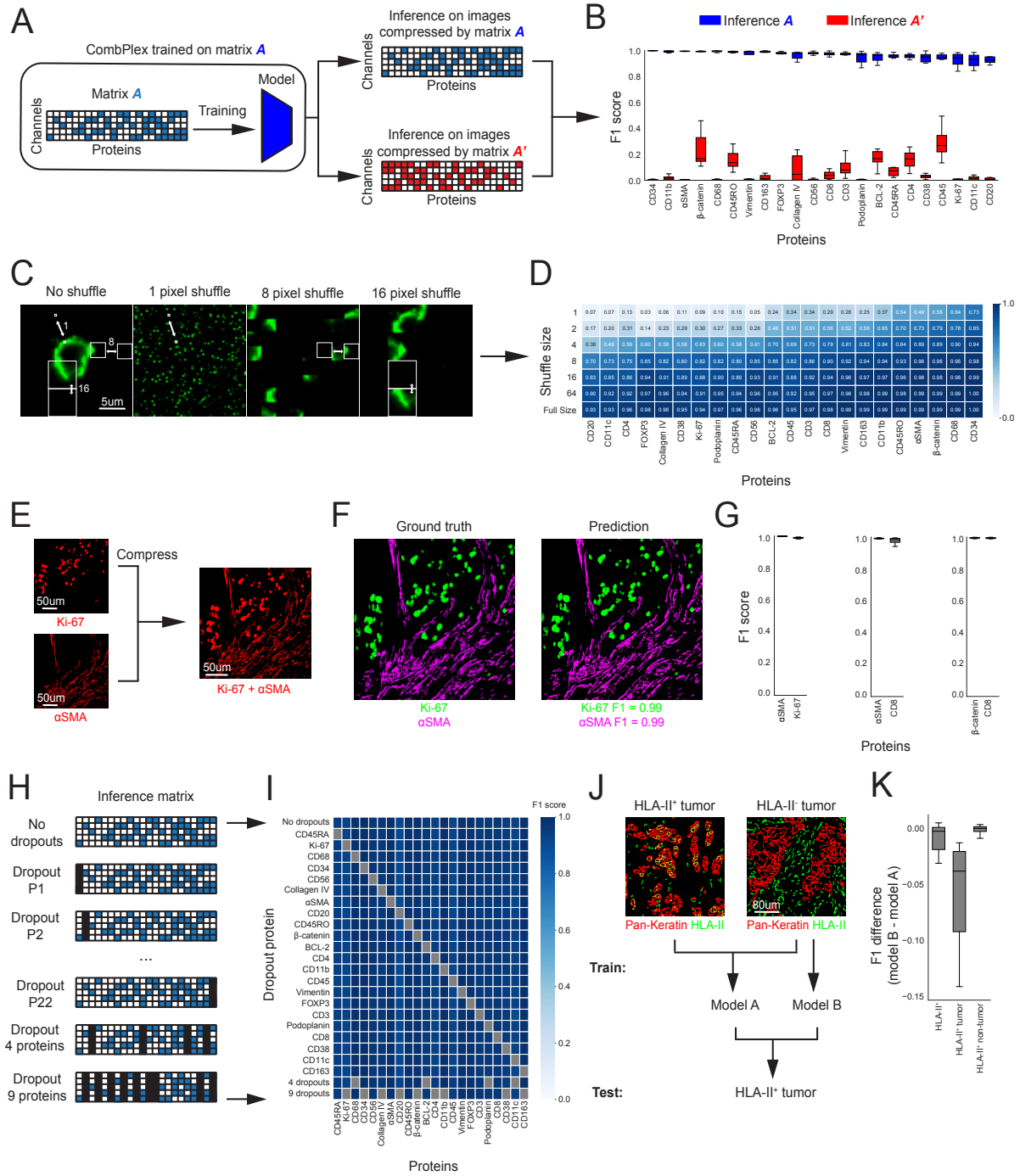

Figure 5: Features contributing to CombPlex performance, strengths and limitations

(A) For the data from Schurch et al [1], a model was trained on simulated compressed images using compression matrix A. Inference was performed on simulated compressed images using compression matrices A and A'. **(B)** A comparison of F1 scores (y-axis) for each protein (x-axis) in the experiment illustrated in (A) indicates that CombPlex learns the compression matrix. **(C-D)** CombPlex was trained on the data from Schurch et al [1]. Images of the test set were then shuffled with patches of varying sizes. (D) Shown are the median F1 scores for each protein (x-axis) on data shuffled with different patch sizes (y-axis). **(E)** Images of Ki-67 and SMA were used to simulate a naïve compression of two proteins in one image. **(F)** CombPlex was used to recover single-protein images from pairs of proteins compressed

to one channel as shown in (E). Shown are the ground truth images (left) and CombPlex' reconstruction (right). **(G)** Shown are the F1 scores (y-axis) for three 2-to-1 decompressions of pairs of proteins (x-axis) on a test set of 8 FOVs. **(H)** CombPlex was trained on the data from Schurch et al [1] and then evaluated on a test set of 8 FOVs. In each run, one or more proteins were set as empty to simulate a scenario of a malfunctioning antibody or utilization of a subpanel. **(I)** The median F1 scores of different proteins (x-axis) are presented across experiments in which different proteins have been dropped out (y-axis). **(J)** Top: Shown are example images of tumor and non-tumor appearances of HLA-II (green) intersected with Pan-Keratin (marking tumor, red). Two models were trained on the CODEX data presented in figure (3). Model A was trained on images of HLA-II$^+$ and HLA-II$^-$ tumors. Model B was trained only on images of HLA-II$^-$ tumors. Both models were evaluated on images of HLA-II$^+$ tumors **(K)** Shown are the differences in F1 scores (y-axis) between the models in (J). HLA-II reconstruction was evaluated in the entire image (left), tumor regions (middle) and non-tumor regions (right).

Discussion

Current approaches to image multiple proteins *in situ* are inherently limited in throughput by their design: using a separate measurement channel for each individual protein. Combinatorial staining holds great potential, because it exponentially increases the number of proteins that can be measured using any imaging platform from C to $2^C - 1$. For example, using three-color fluorescent microscopy one could measure seven proteins ($2^3 - 1$). Seven-color imagers, well-established in research and in the clinic, could serve to image dozens of proteins, and a multiplexed imaging instrument could accommodate hundreds of targets. However, to date, combinatorial staining was not applied to proteins, presumably due to the high abundance of proteins in cells. High protein density makes it impossible to resolve the signals of individual proteins under standard magnification conditions and thus mathematically infeasible to reconstruct the individual protein expression images.

Here, we presented CombPlex, an experimental approach to perform combinatorial staining of proteins, coupled with an algorithmic framework to accurately decompress the combined signal back to individual protein images. We showed that CombPlex allows us to decouple the staining of individual proteins stained together with over 90% accuracy, even in low, 20X magnification, both in fluorescent and in mass-based imaging. CombPlex improves upon previous attempts of combinatorial staining by using a design that combines supervised deep learning and test-time optimization. We showed that although the problem of reconstructing individual protein images from compressed measurements is mathematically underdetermined, CombPlex manages to achieve highly accurate reconstructions. Our analyses suggest that it does so by leveraging various features of the data, including spatial continuity, individual protein expression patterns and protein co-expression. While it is infeasible to encode such features explicitly, a neural network trained on image examples is able to learn them and exploit them as strong priors for probable solutions for solving the reconstruction task.

Although supervised deep learning leads to improved performance, it also serves as the Achilles' heel of the approach. CombPlex requires training data in the form of individual protein

images, coupled with their combinatorically-compressed counterparts. In the manuscript, we presented several results that could reduce the entry barrier posed by the need to acquire such training data. We devised a method to simulate combinatorically-compressed images *in silico* from individual measurements of the corresponding proteins. We demonstrated that a network trained on simulated images achieves high performance, almost the same as a network that was trained on experimental combinatorically-compressed images. Using simulations to train the network holds great advantages. It negates the need to acquire both single-protein images and combinatorically-compressed images on the same tissue. Here, we showed that we were able to train CombPlex on diverse, pre-existing data from various tissue sources and use this network to generate accurate reconstructions of experimentally compressed data. Moreover, we showed in simulations that CombPlex could accommodate 30% missing proteins without a reduction in performance for the remaining proteins. Putting these results together, we envision that a community resource could be generated in which hundreds of proteins would be measured individually on various tissues using currently-available multiplexed imaging approaches, such as cyclic fluorescence or mass-based imaging. This effort will only have to be done once, and could then serve as the basis to pre-train a large foundation model that accommodates the decompression of various staining schemes, visualizing different subsets of proteins. Subsequent studies could then use CombPlex to image multiple proteins on their respective samples, saving valuable time and money. Notably, such community efforts are well on their way, examples including the Human Protein Atlas, Single Cell Human Atlas, or the Human BioMolecular Atlas Program [32–34].

Protein imaging is a pillar of pathology. While recent technological developments have increased the number of proteins visualized *in situ*, these increases have been linear and thus far less than required for proteome-scale measurements. Looking forward, it is difficult to envision how proteome-wide measurements could be performed *in situ* without the use of combinatorics. As such, this study lays a foundation for a multitude of experimental, analytical and biological endeavors to undoubtedly follow. In the near future, there still exist open questions and challenges relating to combinatorial staining that need to be resolved, such as determining how many antibodies could be multiplexed without signal deterioration, devising methods to quickly titrate antibodies such that they could be imaged with similar exposure times etc. Farther-reaching projects could include applying CombPlex to measurements of abundant mRNAs that currently challenge the spot-counting requirements of barcoded *in situ* hybridization methods. We could also expect to see further increases in throughput using standardized combinatorically-multiplexed antibody panels, e.g., a standard immune profiling kit, which will be able to measure thirty proteins using five channels. These could be integrated with available instruments for multiplexed imaging to further increase their multiplexing capabilities. Finally, it will be enticing to observe further experimental and algorithmic developments that could reduce the amount of training data needed, simplify its acquisition or eliminate it altogether.

Methods

CombPlex overview

CombPlex consists of two components: 1) an experimental approach for combinatorial staining of proteins, and 2) an algorithmic framework for decompression of combinatorically-compressed images.

The experimental approach aims to generate compressed multiplexed images, where each channel contains the signal of multiple proteins. To generate these experimentally, multiple proteins are measured simultaneously on the same channel, according to a given compression matrix. Calibration of this staining regime should be optimized according to these principles:

- Protein signals in each channel must be visible under the same imaging conditions.
- A protein that is measured in more than one channel should display similar staining patterns across channels.

The algorithmic approach is divided into two steps: the decompression masking network, a deep learning-based step, and a least squares optimization phase. The former aims to estimate the regions where individual protein signal exists, while the latter completes the recovery by estimating pixel intensities in those regions.

The input for CombPlex is a multiplexed image denoted as $y \in R^{H \times W \times C}$, which is a compressed representation of the underlying signal $x \in R^{H \times W \times P}$ (where $P > C$). The output is a reconstructed signal $\hat{x}$ of the underlying signal x.

- H and W are the height and width dimensions, respectively.
- C signifies the number of channels in the measurement (e.g., colors captured in a microscope).
- P denotes the total number of distinct underlying proteins that were subjected to imaging during the experimental procedure.

An illustration of the approach is shown in Figure 2G and Supplementary figure 1.

<u>Algorithm</u>

Generating the compression matrix

The compression scheme is determined by a binary matrix A with C rows and P columns, representing the measurement channels and individual proteins, respectively. In this matrix, a value of 1 at position (i, j) indicates that protein j is stained in measurement channel i. To generate matrix A we employ three principles, ordered by their priority:

1. Each column vector must be unique such that different proteins are encoded differently.

2. The sum of each column is minimized to facilitate the process of partitioning the antibodies.
3. The sums of the rows are balanced, which aids in the computational task of distinguishing a higher number of protein signals within the same compressed image.

Simulations of combinatorically-compressed images

Given single-protein images x, we simulated compressed images y (**Supp. Fig. 1A**). Simulations for MIBI and CODEX were performed differently due to the distinct characteristics of the signal in each technology.

CODEX simulation process:

1. Matrix A was adjusted by replacing the non-zero values according to values calculated from experimentally compressed images. This step is needed to compensate for experimental differences between the channels. For example, AF750 will generally produce a weaker staining than AF647, even for the same concentration of antibodies. We used a small set of data pairs $\{x_i, y_i\}_{i=1}^{4}$ to minimize the following reconstruction loss and obtain $A_{optimal}$:
$$\omega^* = \underset{\omega}{argmin} \|(\omega \odot A) * x - y\|_2^2$$
$$A_{optimal} = \omega^* \odot A$$
Here, $\odot$ denotes pixel-wise multiplication and $*$ is the convolution operator. The optimization process utilizes gradient descent with a constant learning rate of 10^{-5}.
2. Gaussian noise was applied on Matrix A, followed by convolution with our single-protein images x:
$$y = A' * x\,, \qquad A' \sim N(A, 0.1 \cdot A)$$
This noise accounts for experimental differences between channels, resulting from either systematic differences in channel intensities or stochastic differences in antibody titers.
3. Pixel-wise Gaussian noise was applied on y to simulate experimental background:
$$y \sim N(y, 0.25y)$$
4. To account for signal saturation and interference, we factored the compressed images y with respect to the number of proteins expressed in each pixel. To calculate those factors, we used a small set of data pairs $\{x_i, y_i\}_{i=1}^{4}$, and counted how many proteins are expressed in each pixel of our FOV using x. Then, we divided the means of non-zero pixels in y and $A' * x$ for each protein count and channel and took the mean of those quotients over all the data pairs.

MIBI-TOF simulation process:

1. Average pooling with a kernel size of 5x5 was applied to smooth our single-protein signals.

2. Matrix A was adjusted as described in the CODEX simulation process steps 1-2.
3. Pixel-wise Poisson sampling was applied on y:

$$y \sim Poisson(y)$$

Training the decompression masking network

The goal of this step is to recover the binary support of the individual protein images (**Supp. Fig. 1B**). We perform pixel-wise classification for each pixel (i, j) in the image with respect to the presence or absence of each of the P proteins. Mathematically, we recover a function denoted as $M_\theta(y)$, which takes a compressed image $y \in R^{H \times W \times C}$ as input and outputs a binary image $B(x) \in R^{H \times W \times P}$ such that:

$$M_\theta(y) \approx B(x) \qquad \text{where} \qquad B_{ijp}(x) = \begin{cases} 1 & if\ x_{ijp} > 0 \\ 0 & Otherwise \end{cases}$$

The network is trained in a supervised manner using pairs $\{x_k, y_k\}_{k=1}^{N}$ of compressed images and their corresponding individual protein images. The compressed images can be either acquired experimentally or simulated as described in the *Simulations of combinatorically-compressed images* section.

The objective is to minimize the binary cross-entropy (BCE) loss between the real $B(x)$ and the model output $M_\theta(y)$ in a set of N images. Formally, finding a set of parameters θ to minimize:

$$L(\theta) = -\frac{1}{NHWP} \sum_{k,i,j,p} B\left(x_{kijp}\right) \log\left(M_\theta\left(y_{kijp}\right)\right) + \left(1 - B\left(x_{kijp}\right)\right) \log\left(1 - M_\theta\left(y_{kijp}\right)\right)$$

The network follows the UNet 3+ architecture, an extension of the original UNet architecture that introduces full scale skip connections and convolutional blocks [35]. The symmetric encoder-decoder structure cosists of five scales in terms of depth, with convolutional layers having a kernel size of 3x3. Batch normalization layers and ReLU activation functions are applied after each convolutional layer. We used Adam optimizer, constant learning rate value of 0.001 and batches of 4 FOVs per iteration. Data augmentations were used to enrich the datasets: random crop of size 512x512, random flips and rotations by 90 degrees.

Inference using trained model and least squares optimization

During inference, we use a trained model M_θ and a set of compressed images $\{y_i\}_{k=1}^{N}$ from which we want to obtain $\hat{x}$ (**Supp. Fig. 1C**).

First, we apply the trained model on y to obtain $M_\theta(y)$. Then, we recover x given the non-zero signal regions in $M_\theta(y)$. Mathematically, we solve a constrained least-square problem as follows:

$$\omega^* = \underset{\omega}{argmin} \|A * (M_\theta(y) \odot \omega) - y\|_2^2$$

For the CODEX experiment, we compensated the reconstructed compressed images with respect to the factors used in step 4 of the CODEX simulation process:

$$\omega^* = \underset{\omega}{argmin} \|A * (M_\theta(y) \odot \omega) \odot S - y\|_2^2$$

To optimize the objectives above, we perform gradient descent with the Adam optimizer and a constant learning rate of 0.001. The output is $\hat{x} = M_\theta(y) \odot \omega^*$, derived only from A, M_θ and y.

CODEX *in silico* experiment

Data

Analysis was performed on the data from Schurch et al., in which 57 proteins were measured in colorectal cancer (CRC) specimens using co-detection by indexing (CODEX) cyclic fluorescence [27]. From the publicly available dataset, 22 proteins which had specific staining and a high signal-to-noise ratio, were chosen for compression (**Supp. Table 1A**). Multiplexed image sets from 41 FOVs were used in the following steps. Images were denoised using the Ilastik software, by training a pixel-classifier for each protein [36]. Briefly, all FOVs for a given protein were loaded into Ilastik and used for training the pixel classifier by manually marking real signal and noise. After achieving satisfactory classifications, the trained model was used to produce binary segmentation masks for each FOV. Raw images were intersected with the segmentation masks to produce clean images, containing the original pixel values where the segmentation mask pixels equaled 1. Images were aggregate filtered by applying average smoothing with a sliding window and removing signal clusters with a radius smaller than 100 pixels (**Supp. Table 1A**).

Simulation and decompression

The compressed images y were simulated using the single-protein images x according to: $y = A * x$. Matrix A was generated following the *"Generating the compression matrix"* methods section. To train the masking network, the dataset was split into a training set of 33 FOVs and a test set of 8 FOVs. The model was trained for 115K epochs using a GPU of type Nvidia Quadro RTX8000.

CODEX compression experiment

Antibody conjugation

A summary of antibodies, barcode oligos, reporter fluorophores, and concentrations can be found in Supplementary Tables 1B,C. Oligo-conjugated primary antibodies were prepared 50µg at a time using the Akoya Antibody Conjugation Kit (Akoya Biosciences, Delaware, USA) according to the manufacturer's recommended protocol. Biotin-conjugated antibodies were prepared 100µg at a time using the ChromaLINK One-Shot Antibody Biotinylation Kit (Vector labs, Newark, USA).

CODEX staining

Breast cancer TMA glass slides were acquired from TissueArray (Maryland, USA). Slide tissue sections were baked at 70°C for 20 min. Tissue sections were deparaffinized with 3 washes of fresh xylene. Tissue sections were then rehydrated with successive washes of ethanol 100% (2x), 95% (2x), 80% (1x), 70% (1x), and ultra-pure DNase/Rnase-Free water (Bio-Lab, Jerusalem, Israel), (3x). Washes were performed using a Leica ST4020 Linear Stainer (Leica Bio-systems, Wetzlar, Germany) programmed to 3 dips per wash for 30 seconds each. The sections were then immersed in epitope retrieval buffer (Target Retrieval Solution, pH 9, DAKO Agilent, Santa Clara, CA) and incubated at 97°C for 40 min and cooled down to 65°C using Lab vision PT module (Thermofisher Scientific, Waltham, MA). Slides were stained using Akoya Staining Kit for Phenocycler (Akoya Biosciences, Delaware, USA) according to the manufacturer's recommended protocol. Details on the staining conditions for the primary antibodies are provided in Supplementary Table 1C.

CODEX imaging

CODEX imaging was performed using a Phenocycler Fusion system (Akoya Biosciences, Delaware, USA). Phenocycler reporter plate and Flow Cell assembly were performed according to the manufacturer's protocol. Compressed images were created by hybridizing multiple reporters loaded with the same fluorophore in the same imaging cycle. Details on the layout of the reporter plate and imaging conditions are provided in Supplementary Table 1C. For each antibody, the working titer was determined as the titer resulting in the detection of all antibodies in each compressed channel. Titers were assessed by overlaying compressed and ground truth images and evaluating successful compression both qualitatively and quantitatively by measuring signal correlation (**Supp. Fig. 3**).

Multiplexed image sets were extracted by manually marking individual cores using QuPaths' TMA De-arrayer with a diameter of 1.2mm and a custom groovy script, which exports them as tiff files [37]. Single-protein images were denoised and aggregate-filtered as described above (**Supp. Table S1B**). Compressed segmentation masks were produced by a binary summation of all single-protein images in each compressed channel, according to the compression matrix. Raw experimentally compressed images were intersected with the compressed segmentation mask to produce clean compressed images.

Dataset and model settings

The dataset included 132 1mm^2 FOVs and was split into a training set of 100 FOVs and a test set of 32 FOVs. We trained two models: the first was trained on experimentally compressed images, and the second was trained on simulated compressed images as described in the *Simulations of combinatorically-compressed images* section. Both models were evaluated on the same experimentally compressed images. The model was trained for 110K epochs using a GPU of type Nvidia Quadro RTX8000.

Cell classification

Cell segmentation masks were generated using Mesmer and the nuclear channel (DAPI) for each FOV [29]. Intensity values were transformed using arcsinh transformation with a factor of 0.1, and normalized to 0-1 using min-max normalization between the 1st and 99.5th percentiles. Thresholds for gating positive/negative were chosen for each protein in the ground truth images using visual inspection. Optimal thresholds for the reconstructed images were chosen using a ROC curve analysis, which maximized accuracy between the classification in the ground truth and reconstructed images.

MIBI-TOF experiment

TMA preparation for MIBI-TOF

Samples of ten breast carcinoma and seven lung carcinoma FFPE biopsies from ten patients were obtained from Hadassah Hospital Pathology department. IHC staining for tumor cells (Pan-Keratin) and immune cells (CD45) was performed on all samples as described below. IHC results were reviewed for expression of both proteins and 2mm cores were compiled into a TMA.

Antibody conjugation

A summary of antibodies, reporter isotopes, and concentrations can be found in Supplementary Table S3D. Metal conjugated primary antibodies were prepared 100ug at a time using the Maxpar X8 Antibody Labeling Kit (Fluidigm, California, USA) according to the manufacturer's recommended protocol. Multi-metal conjugates were prepared using the same kit and protocol, by mixing multiple isotopes to a final volume of 5µl and conjugating them to a single polymer, as detailed in supplementary table S1D and supplementary figure 6A. For each antibody, the ratio between isotopes was calculated to compensate for inherent differences in their respective ionization efficiencies (I/E) [15] as listed in table S1E. Following metal labeling, antibodies were diluted in Antibody Stabilizer PBS solution (Boca Scientific, Washington, USA) to 0.5 mg/ml and stored long-term at 4°C.

MIBI-TOF staining

Staining was performed as detailed in Elhanani et al [38]. Tissue sections (5µm thick) were cut from FFPE tissue blocks of the TMA using a microtome, and mounted on gold-coated slides (Ionpath, Menlo Park, CA) for MIBI-TOF analysis. Slide tissue sections were baked at 70°C for 20 min. Tissue sections were deparaffinized with 3 washes of fresh xylene. Tissue sections were then rehydrated with successive washes of ethanol 100% (2x), 95% (2x), 80% (1x), 70% (1x), and ultra-pure DNase/RNase-Free water (Bio-Lab, Jerusalem, Israel), (3x). Washes were performed using a Leica ST4020 Linear Stainer (Leica Bio-systems, Wetzlar, Germany) programmed to 3 dips per wash for 30 s each. The sections were then immersed in epitope retrieval buffer (Target Retrieval Solution, pH 9, DAKO Agilent, Santa Clara, CA) and incubated at 97°C for 40 min and cooled down to 65°C using Lab vision PT module (Thermofisher Scientific, Waltham, MA). Slides were washed for 5 min with a wash buffer made with ultra-pure DNase/RNase-Free water (Bio-Lab, Jerusalem, Israel) and 5% 20X TBS-T pH7.6 (Ionpath, Melo Park, CA). Sections were then blocked for 1h with 3% (v/v) donkey serum (Sigma-Aldrich, St Louis, MO) diluted in TBS-T wash buffer. Metal-conjugated antibody mix was prepared in 3% (v/v) donkey serum TBS-T wash buffer and filtered using centrifugal filter, 0.1 µm PVDF membrane (Ultrafree-Mc, Merck Millipore, Tullagreen Carrigtowhill, Ireland). Two panels of antibody mixes were prepared. The first panel contained most of the metal-conjugated antibodies, and was incubated overnight at 4°C in a humid chamber. The second mix contained antibodies for dsDNA and αSMA. Following overnight incubation, slides were washed twice for 5 min in wash buffer and incubated with the second antibody mix for 1h at 4°C. Slides were washed twice for 5 min in wash buffer and fixed for 5 min in diluted glutaraldehyde solution 2% (Electron Microscopy Sciences, Hatfield, PA) in PBS-low barium (Ionpath, Menlo Park, CA). Tissue sections were dehydrated with successive washes of Tris 0.1 M (pH 8.5), (3x), Ultra-pure DNase/RNase-Free water (2x), and ethanol 70% (1x), 80% (1x), 95% (2x), 100% (2x). Slides were immediately dried in a vacuum chamber for at least 1 h prior to imaging.

MIBI-TOF imaging

Imaging was performed using a MIBIScope mass spectrometer (Ionpath, Menlo Park, CA) with a Xenon ion source. FOVs of 400µm^2 were acquired using a grid of 1024X1024 pixels, with 1msec dwell time per pixel.

For each antibody, the working titer was determined as the titer that resulted in the detection of all antibodies in each compressed channel. This was assessed by overlaying compressed and ground truth images and evaluating successful compression (**Supp Fig. 6B**).
Single-protein images were extracted, slide background-subtracted, denoised and aggregate filtered using MAUI [39]. Compressed segmentation masks were produced by a binary summation

of all single-protein images in each compressed channel, according to the compression matrix. Experimentally compressed images were intersected with the compressed segmentation mask to produce clean compressed images.

Dataset and model settings

The decompression masking network was trained and tested using MIBI-TOF images of compressed or single expression measurements of Pan-Keratin, αSMA, CD45, CD8, HLA-II, and Ki-67 from several datasets:

- *Triple-negative breast cancer dataset:* The dataset contains 44 800µm^2 images [40].
- *Non-small cell lung carcinoma dataset:* The dataset contains 84 800µm^2 images.
- *Melanoma lymph node dataset:* The dataset contains 16 800µm^2 images.
- *Breast and Lung Carcinoma TMA dataset:* The dataset contains 57 400µm^2 images.

The dataset included 200 FOVs of varying sizes and was split into a training set of 181 FOVs and a test set of 19 FOVs. Training was performed using simulated compressed images as described in the algorithm section. Inference was performed on experimental combinatorically-compressed images. The model was trained for 90K epochs and took 1-2 days using a GPU of type Nvidia Quadro RTX8000.

Cell classification

Cell segmentation masks were generated using Mesmer and the nuclear channel (dsDNA) for each FOV [29]. Intensity values were transformed using arcsinh transformation with a factor of 100, and normalized to 0-1 using min-max normalization between the 1st and 99.5th percentiles. Thresholds for gating positive/negative were chosen for each protein in the ground truth images using visual inspection. Optimal thresholds for the reconstructed images were chosen using a ROC curve analysis, which maximized accuracy between the classification in the ground truth and reconstructed images.

Analyses

F1 scores calculations

F1 scores were computed by comparing the GT single-protein images with their recovered image using CombPlex. This was done separately for each FOV and each protein using the following formula:

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

For MIBI data, prior to calculating the F1 score, we applied average pooling with a kernel size of 5x5 on the images to smooth the signal.

Reconstruction evaluation

The reconstruction scores were computed by comparing the GT single-protein images with their recovered image using CombPlex. This was done separately for each FOV and each protein, following these steps:

1. A 5x5 average pooling is applied to both the GT and recovered images.
2. The non-zero pixel values from the intersection of both images are flattened into two vectors.
3. Pearson correlation coefficient is calculated between these two vectors to determine the reconstruction score.

Compressed sensing

Gradient descent was performed to minimize the objective function, which consists of three components:

- MSE loss between the reconstructed compressed images, created by convolution of single protein prediction with A, and the simulated compressed images.
- Sparsity constraint on the single-protein reconstructed images, factored by $\alpha = 0.1$.
- Smoothing constraint on the single-protein reconstructed images, factored by $\beta = 0.001$.

Mathematically, we sought to find $\hat{x}$ that minimizes:

$$\mathcal{L}(\hat{x}) = \mathcal{L}_{MSE} + \alpha \cdot \mathcal{L}_{sparsity} + \beta \cdot \mathcal{L}_{smooth}$$

Where:

$$\mathcal{L}_{MSE} = \sum_{i,j,c} \parallel (A * \hat{x})_{ijc} - y_{ijc} \parallel_2^2$$

$$\mathcal{L}_{sparsity} = \sum_{i,j,p} \sqrt{\parallel \hat{x}_{ijp} \parallel_1}$$

$$\mathcal{L}_{smooth} = \sum_{i,j,p} \left|\hat{x}_{i+1,j,p} - \hat{x}_{i,j,p}\right| + \left|\hat{x}_{i,j+1,p} - \hat{x}_{\ i,j,p}\right|$$

Decompression using a one-step algorithm

A UNet 3+ network was trained as described for CombPlex, with the following changes: 1). Mean squared error (MSE) loss was used instead of binary cross-entropy (BCE). 2) The sigmoid activation function at the end was removed to allow for pixel-value reconstruction rather than binary classification. The objective that was minimized is expressed as follows:

$$L(\theta) = \frac{1}{NHWP} \sum_{k,i,j,p} \left(M_\theta\left(y_{kijp}\right) - x_{kijp} \right)^2$$

Where all the notations correspond to those described in the "Training the decompression masking network" methods section. The dataset, the process of simulating compressed images, the architecture of the model, and the hyperparameters employed were kept identical to those outlined in the CODEX in-silico experiment.

Figure creation

Illustrations were created using Biorender templates (https://biorender.com).

Data and code availability

The code for CombPlex can be found at: https://github.com/KerenLab/CombPlex. The data used in our experiments can be found at: https://www.ebi.ac.uk/biostudies/studies/S-BIAD873.

Acknowledgements

L.K. is supported by the Enoch, Azrieli, Sharon Levine and Rising Tide foundations, and grants funded by the Schwartz/Reisman Collaborative Science Program, European Research Council (948811), the Israel Science Foundation (2481/20, 3830/21) within the Israel Precision Medicine Partnership program. I.M is supported by a EU - Horizon 2020 - MSCA Individual Fellowship (890733). S.B. is a Robin Chemers Neustein AI Fellow and acknowledges funds from the Carolito Stiftung and the NVIDIA Applied Research Accelerator Program. C.M.S is supported by the German Research Foundation (INST 37/1228-1 FUGG, and Germany's Excellence Strategy - EXC 2180 – 390900677), Swiss Life Jubiläumsstiftung, Mach-Gaensslen Stiftung Schweiz, American Society of Hematology Research Restart Award, State of Baden-Württemberg within the Centers for Personalized Medicine Baden-Württemberg (ZPM). O.E. is supported by the Israel Cancer Research Foundation (ICRF).

Conflicts of Interest

C.M.S. is a scientific advisor to AstraZeneca plc, and is on the scientific advisory board of, has stock options in, and has received research funding from Enable Medicine, Inc., all outside the current work. R.B, L.B, O.B and L.K. are authors of patent P-621615-IL on combinatorial staining for multiplexed imaging.

Bibliography

1. Rodriguez-Canales, J., Eberle, F. C., Jaffe, E. S. & Emmert-Buck, M. R. Why is it crucial to reintegrate pathology into cancer research? *Bioessays* **33**, 490–498 (2011).
2. Odell, I. D. & Cook, D. Immunofluorescence techniques. *J. Invest. Dermatol.* **133**, (2013).
3. Tsujikawa, T. *et al.* Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. *Cell Rep.* **19**, 203–217 (2017).
4. Gerdes, M. J. *et al.* Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11982–7 (2013).
5. Radtke, A. J. *et al.* IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33455–33465 (2020).
6. Schubert, W. *et al.* Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat. Biotechnol. 2006 2410* **24**, 1270–1278 (2006).
7. Lin, J. R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife* **7**, (2018).
8. Kinkhabwala, A. *et al.* MACSima imaging cyclic staining (MICS) technology reveals combinatorial target pairs for CAR T cell treatment of solid tumors. *Sci. Reports 2022 121* **12**, 1–16 (2022).
9. Gut, G., Herrmann, M. D. & Pelkmans, L. Multiplexed protein maps link subcellular organization to cellular states. *Science* **361**, eaar7042 (2018).
10. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
11. Wang, Y. *et al.* Rapid Sequential in Situ Multiplexing with DNA Exchange Imaging in Neuronal Cells and Tissues. *Nano Lett.* **17**, 6131–6139 (2017).
12. Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* 1–11 (2019) doi:10.1038/s41587-019-0207-y.
13. Lin, R. *et al.* A hybridization-chain-reaction-based method for amplifying immunosignals. *Nat. Methods 2018 154* **15**, 275–278 (2018).
14. Stack, E. C., Wang, C., Roman, K. A. & Hoyt, C. C. Multiplexed immunohistochemistry, imaging, and quantitation: A review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* **70**, 46–58 (2014).
15. Keren, L. *et al.* MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci. Adv.* **5**, (2019).
16. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).

17. Elhanani, O., Ben-Uri, R. & Keren, L. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* (2023) doi:10.1016/J.CCELL.2023.01.010.

18. Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nat. Rev. Genet. 2022 2312* **23**, 741–759 (2022).

19. Keren, L. *et al.* MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci. Adv.* **5**, eaax5851 (2019).

20. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science (80-.).* **348**, aaa6090–aaa6090 (2015).

21. Eng, C. H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nat. 2019 5687751* **568**, 235–239 (2019).

22. Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers--the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **38**, D750-3 (2010).

23. Campagne, F., Dorff, K. C., Chambwe, N., Robinson, J. T. & Mesirov, J. P. Compression of Structured High-Throughput Sequencing Data. *PLoS One* **8**, 79871 (2013).

24. Cleary, B., Cong, L., Cheung, A., Lander, E. S. & Regev, A. Efficient Generation of Transcriptomic Profiles by Random Composite Measurements. *Cell* **171**, 1424-1436.e18 (2017).

25. Cleary, B. *et al.* Compressed sensing for highly efficient imaging transcriptomics. *Nat. Biotechnol. 2021 398* **39**, 936–942 (2021).

26. Sachs, K. *et al.* Compressed Sensing for Simultaneous Measurement of Multiple Different Biological Molecule Types in a Sample. (2013).

27. Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* **182**, 1341-1359.e19 (2020).

28. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, (2018).

29. Greenwald, N. F. *et al.* Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* (2021) doi:10.1038/S41587-021-01094-0.

30. Murphy, K., Weaver, C. & Janeway, C. *Janeway's immunobiology*. (Garland Science, 2016).

31. Johnson, D. B. *et al.* Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. *Nat. Commun.* **7**, 10582 (2016).

32. Consortium, H. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187 (2019).

33. Regev, A. *et al.* The Human Cell Atlas. *Elife* **6**, (2017).

34. Thul, P. J. & Lindskog, C. The human protein atlas: A spatial map of the human proteome.

Protein Sci. **27**, 233 (2018).

35. Huang, H. *et al.* UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* **2020**-**May**, 1055–1059 (2020).

36. Sommer, C., Straehle, C., Kothe, U. & Hamprecht, F. A. Ilastik: Interactive learning and segmentation toolkit. in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 230–233 (IEEE, 2011). doi:10.1109/ISBI.2011.5872394.

37. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Reports 2017 71* **7**, 1–7 (2017).

38. Elhanani, O., Keren, L. & Angelo, M. High-Dimensional Tissue Profiling by Multiplexed Ion Beam Imaging. *Methods Mol. Biol.* **2386**, 147–156 (2022).

39. Baranski, A. *et al.* MAUI (MBI Analysis User Interface)—An image processing pipeline for Multiplexed Mass Based Imaging. *PLoS Comput. Biol.* **17**, e1008887 (2021).

40. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387.e19 (2018).