Supplementary Materials for

Multiplexed video hyperspectral microendoscopy reveals tumor–immune dynamics *in vivo*

Bryan Q. Spring[1,2,3]*, Mohammad Ahsan Saad[4]†, Rebecca C. Harman[1,2]†, Eric M. Kercher[1,2]†, Ryan T. Lang[1,2]†, Kai Zhang[1,2]†, Julia Tatz[1,2], Qianqian Fang[1,3,5], Jason Sutin[6], Zhiming Mai[4], Akilan Palanisami[4], and Tayyaba Hasan[4,7]*

Corresponding authors: b.spring@northeastern.edu or thasan@mgh.harvard.edu

The PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Measurement basis optimization and simulation-guided optical design

The pixel spectrum without measurement noise is described as a 1D vector (S) and expanded in an orthonormal basis $\Psi = [\psi_1 \psi_2 \cdots \psi_N]$ as follows:

$$S = \sum_{i=1}^{N} x_i \psi_i$$

where the coefficient sequence x is obtained from the spectrum inner product, $x_i = \langle S, \psi_i \rangle$. The first M terms define the "sensing" basis, and the remaining terms define the "non-sensing" basis:

$$S = \sum_{i=1}^{M} x_i \psi_i \quad + \sum_{j=M-N+1}^{N} x_j \psi_j$$

An optimal sensing basis minimizes the error between the spectrum and the sensing basis reconstruction:

$$\min \left| S - \sum_{i=1}^{M} x_i \psi_i \right|$$

with increasing M (corresponding to the number of detector channels) improving the spectral resolution and reconstruction. With a suitably sparse basis, the spectral reconstruction from the sensing basis can then be unmixed using conventional non-negative least squares fitting (NNLS). This framework becomes more complicated in the *in vivo* photon starved regime due to Poisson shot noise (ϵ_P) which has a non-linear dependence on the photon count ($\Delta t \langle S, \psi_i \rangle$). To mitigate the disruption of the shot noise on the spectral reconstruction and unmixing, an extra term is added to the minimization for simulations of the optical design:

$$\min \left| \Delta t \sum_{i=1}^{M} x_i \psi_i + \sum_{i=1}^{M} \epsilon_P(\Delta t \langle S, \psi_i \rangle) - \sum_{k=1}^{5} \beta_k \Phi_k \right|$$

where Δt is the exposure time for pixel spectral measurement, Φ is the fluorophore emission spectra and β is the unmixed fluorophore abundance. The above equation is written to make explicit the dependence of the shot noise on the choice of sensing basis. Unlike the noise-free situation, increasing M does not necessarily improve the minimization, as the non-linear dependence of the shot noise on signal can dominate the other terms. This minimization is non-trivial to solve analytically, motivating the use of Monte Carlo techniques for the inclusion of the noise term followed by NNLS for the spectral unmixing simulations. For each candidate measurement and basis frame rate, 10^4 Monte Carlo iterations were performed to estimate mean square error (fig. S1).

SMIRC design and implementation

SMIRC was developed to enable multiplexed *in vivo* spectral imaging under photon-limited conditions. The selection of fluorophore basis spectra, detector channel count, and acquisiton conditions was guided by simulation-based optimization described above. Because low photon counts preclude dense spectral sampling under conventional Shannon–Nyquist constraints, spectral acquisition was optimized using compressive sampling principles (*17*). Basis selection and spectral channel configuration were determined using Monte Carlo simulations that incorporated signal-dependent Poisson shot noise (see Supplementary Methods). Instrument design conditions were systematically examined using different combinations of fluorophores, excitation conditions and frame rates using MATLAB (The MathWorks, Natick, Massachusetts, USA) implemented on a Linux cluster (fig. S1). Zemax optical (Zemax LLC; Kirkland, Washington, USA) simulations performed in parallel optimized for chromatic aberration and provided optical parameters for the Monte Carlo calculations.

These calculations then informed the implemented optical design (figs. S2 and S3) where the output of a 636 nm diode laser (a special order OBIS 636 nm LX 100 mW laser system with a fiber pigtail; Coherent, Wilsonville, OR) was raster scanned over the active area of the distal end of a 0.35 NA coherent fiber-optic bundle (0.95-mm-diameter FIGH-30-850N Fujikura Image Fiber with ~30,000 cores and 3 μm core-to-core separation, which defines the number of resolvable pixels; Myriad Fiber Imaging Technology, Inc.; Dudley, Massachusetts, USA) using a 10× 0.45 NA plan-apochromat microscope objective (overfilled back aperture; Carl Zeiss Microscopy, LLC; Thornwood, New York, USA) with 0.7 mW of delivered power unless noted otherwise.

Raster scanning was implemented with a 36-facet polygonal mirror (DT-36250-020; Lincoln Laser; Phoenix, Arizona, USA) for the fast x-axis scan, and a galvanometric mirror (QS-7; Nutfield; Hudson, New Hampshire, USA) for the slow y-axis scan. The maximum field-of-view for a single image is ~0.79 mm, limited by the diameter of the fiber bundle active area. Programmable hardware for horizontal (x-axis) and vertical (y-axis) scanning mirror synchronization with the data stream acquisition enabled flexible frame dimensions and image reconstruction. The scanned laser beam is injected into the individual multimode cores composing the fiber bundle, which serve as light conduits to and from the tissue. This single laser line was used to excite the panel of targeted fluorophore conjugates.

The fiber cores, combined with a 100 μm pinhole at the entrance of the spectrograph in the descanned pathway, provide partial optical depth sectioning by rejecting out-of-focus fluorescence captured by neighboring fiber cores. However, due to the absence of a microlens objective, the excitation light diverges from the fiber core tip into tissue, resulting in "semi-confocal" performance. The lateral sampling (pixel size) was 1.1 μm, such that each fiber core was sampled with ∼7 pixels (*i.e.*, the lateral optical resolution is ∼3 μm). The axial resolution is ~160 μm in a phantom (fig. S3F), though limited fluorescent probe penetration (fig. S16) and scattering in tissue enhance effective axial resolution *in vivo* (*53*).

Fluorescence was detected using a custom 16-channel linear-array photomultiplier tube (PMT) system, of which 15 channels yielded usable data because one channel exhibited a persistent fault. The detector incorporated a PMT with extended sensitivity to wavelengths up to 920 nm (~10% quantum efficiency from 600–800 nm), modified with a 20 MHz amplifier (H01515M-20, Hamamatsu; Shimokanzo, Japan), and coupled to a software-controlled, tunable Czerny–Turner spectrograph (Acton SP-2156 Imaging Spectrograph; Princeton Instruments; Acton, Massachusetts, USA) with an f/4 optical configuration and a reflective diffraction grating (Richardson grating 53-*-396R, 500 grooves/mm, 770 nm blaze, >70% transmission from 600–

800 nm). This configuration enabled collection of a 642–807 nm spectral range in approximately 11 nm increments with single-photon sensitivity. Dichroic (LPD01-633RU-25) and notch (NF03-633E-25) filters were purchased from Semrock (Rochester, New York, USA). Four PCI Express (PCIe) 10-bit frame grabbers (Solios eA; Matrox Electronic Systems Ltd.; Dorval, Quebec, Canada), each equipped with four input channels, digitized the 15 PMT channels in parallel to enable real-time acquisition.

Laser scan engine

To relink the data streams with their spatial image coordinates, a timing laser system that detects the start of the polygonal line scan (Lincoln Laser BMC-7; 10,000 rpm, 36 facet mirror) with sub-µsec precision was used to synchronize the galvanometric mirror scan (Nutfield Technology QS-7). As the polygon rotated, the appearance of a new polygonal mirror facet was detected with a split photodiode (Advance Photonix Inc SD 113-2421-021) coupled to a wide-band current feedback operational amplifier in an inverting gain configuration (Analog Devices AD811) and inverting electronic Schmitt triggers (Texas Instrument CD74HC14) for back-end TTL pulse generation. The timing of these scan events is critical and must be implemented in hardware to avoid image jitter. This is typically done using hardwired digital electronics for vertical and horizontal synchronization pulse trigger generation (*54*). However, the resulting data acquisition rates were substantial and placed the instrumentation computer on the edge of stability. Optimizing data transfer rates while maintaining stable system operation required repeated timing-hardware redesign, which proved impractical using conventional hardwired approaches. Standard externally clocked solutions were more easily modified but had unacceptable jitter due to the high acquisition clock speeds. To accelerate the optimization, a new programmable hardware implementation was developed (fig. S2C). The galvanometric mirror and frame grabber trigger waveforms were loaded at runtime into the hardware FIFO buffer queue of a waveform generator (National Instruments PCI 6259) with simultaneous analog/digital output capabilities, and the subsequent waveform output clocked by the start of polygonal scan laser trigger. These synthesized waveforms were also used to trigger synchronized burst acquisition of the frame grabbers (Matrox Solios eA/XA). This implementation provided jitter-free synchronization of the raster scanning system with the image acquisition and enabled run-time raster scan parameter flexibility while obviating the problems associated with externally clocked trigger waveform generation. Furthermore, this design allows enhanced dynamic range by allowing a trade-off between speed and photon collection efficiency not available in conventional hardwired circuit topologies (*54*).

Spectral cube acquisition and spatial filtering

For each image frame, fluorescence signals in each PMT spectral channel were digitized at sampling rates of 16 or 32 MHz, corresponding to temporal oversampling of each spatial pixel by three or six sequential analog-to-digital samples during raster scanning. Given the image format (900 × 351 pixels) and acquisition rate (16.7 frames per second), each reconstructed pixel (1.1 × 1.1 µm) samples a region smaller than the fiber core–to–core spacing (3 µm), such that approximately seven pixels are acquired per fiber core. As a result, each fiber core is sampled by 3 or 6 temporal samples per pixel multiplied by ~7 spatial samples per core, yielding approximately 21 or 42 spectral samples per fiber core per frame. During reconstruction, temporally oversampled measurements were averaged to generate the final pixel intensity, effectively integrating the detected photon stream over the

pixel dwell time. No additional temporal averaging is performed beyond this reconstruction step. Prior to spectral unmixing, the reconstructed spectral image cube was processed using a spatial Gaussian convolution filter (6 × 6 pixels, σ = 4 pixels). This filter acts as a low-pass spatial frequency operation that suppresses high-frequency modulation arising from the discrete fiber cores, removing the fiber-core pattern. Importantly, this spatial filtering operates on the already reconstructed, oversampled image and preserves the fundamental optical resolution set by the fiber core spacing while integrating the multiple spatial samples acquired per fiber core.

Effective photon count estimation

Effective mean photon counts (N_{eff}) were estimated using a variance–mean signal analysis to characterize the relationship between detected signal statistics and the underlying photon-limited noise processes of the SMIRC microendoscope. This approach leverages the fact that, under shot-noise–limited conditions, the variance of the detected signal (σ_d^2) scales linearly with its mean (μ_d), allowing conversion from measured digital levels to effective photoelectron and photon counts via a proportionality constant. In brief, incident photons (N_γ) at the detector give rise to photoelectrons according to the detector quantum efficiency, η, ($N_{\mathrm{e}} \approx \eta N_\gamma$, where η varies with wavelength and spectral channel), followed by amplification and digitization through the PMT and analog-to-digital conversion (ADC) chain. The measured pixel value can therefore be expressed as the sum of an additive baseline offset and a gain-scaled contribution proportional to the number of detected photoelectrons. Under Poisson statistics, the variance of the photoelectron count equals its mean, which propagates through the detection chain to yield a linear relationship between the variance and mean of the measured signal (*55*):

$$\sigma_d^2 = \mathrm{R}_d^2 + \frac{G}{g_{ADC}}\left(1 + \left(\frac{\sigma_G}{\mathrm{G}}\right)^2\right)\mu_d,$$

where R_d^2 is the ADC read-out noise, g_{ADC} is the ADC photoelectron to digital level conversion factor, and G is the detector (PMT) effective gain with variance σ_G^2. To experimentally determine the effective gain factor relating digital levels to photoelectrons, uniform fluorophore-in-solution images were acquired over a range of excitation powers, and the variance and mean signal were computed from homogeneous regions of interest (Fig. 2A). The slope (m) of the resulting variance–mean relationship ($\sigma_d^2 = \mathrm{b} + m\mu_d$) provides the conversion factor between measured digital levels and effective photoelectrons, after accounting for any constant baseline offset (b) and assuming that fractional PMT gain fluctuations are small ($\sigma_G/\mathrm{G} \ll 1$):

$$N_{\mathrm{e}} \approx \frac{\mu_d}{(G/g_{ADC})} = {\mu_d}/{m},$$

This calibration enables estimation of the effective photoelectron count per pixel directly from either the mean signal or the measured variance, *e.g.*:

$$N_{\mathrm{eff}} \equiv N_{\mathrm{e}} \approx {\mu_d}/{m},$$

Effective photon counts were then obtained from mean signal and variance measurements, yielding an order-of-magnitude estimate of the effective photon budget per pixel relevant for assessing shot-noise–limited performance. Because the goal of this analysis is to quantify the

photon-limited operating regime and spectral unmixing feasibility rather than absolute radiometric flux, photon counts are reported as *effective* photons per pixel rather than *incident* photons.

Spectral unmixing and image processing

After reconstruction, hyperspectral image cubes were spectrally decomposed using a newly developed graphics processing unit (GPU)–accelerated implementation of fast non-negative least squares (FNNLS; figs. S4–S7). For each imaging experiment, a basis spectra library was assembled by measuring each fluorophore independently under matched imaging conditions. The library also included detector dark spectrum and excitation laser reflection as additional basis components (*56*). For *in vivo* studies, reference spectra were obtained from single-fluorophore conjugate injections to account for tissue-dependent optical distortions and background contributions (fig. S32).

At each pixel, the measured spectrum was modeled as a linear combination of the basis spectra, and fluorophore abundances were estimated by solving a non-negative linear least-squares problem (*57*). The non-negativity constraint was imposed to enforce physically meaningful solutions and suppress nonphysical negative cross-talk among spectrally overlapping components (*58*). Because conventional CPU implementations of non-negative least squares are prohibitively slow for hyperspectral video (*59–61*), the algorithm was parallelized in CUDA and executed on two TITAN V GPUs, yielding an approximately 18,000-fold increase in processing speed relative to conventional CPU-based implementations (fig. S7A). Average processing time was ~5 ms per hyperspectral image cube using two TITAN V GPUs (fig. S6C), exceeding the acquisition rate and enabling faster-than-acquisition spectral decomposition (performed offline in the present study).

Before unmixing, images were normalized to excitation power, and saturated pixels exceeding the 10-bit detector range were excluded. Unmixed component images were rendered as pseudo-colored basis maps and combined into composite images for visualization.

Wide-field hyperspectral image mosaicking

For wide-field imaging, sequential SMIRC frames were assembled into mosaics using a multi-channel feature-based registration framework derived from our previously described multichannel normalized cross-correlation (NCC) algorithm (*36*), augmented with feature matching, inlier pruning, and adaptive geometric transformation estimation, and integrated with graph-cut–based stitching applied to spectrally unmixed molecular abundance maps. Importantly, the algorithm does not assume purely translational motion but instead selects the appropriate transformation model (translation or affine) based on data-driven criteria.

Before mosaicking, a flat-field correction was applied by normalizing pixels of each unmixed channel by its temporal mean across the video, thereby reducing vignetting and illumination inhomogeneity.

The algorithm uses multiple frame-pair alignment stages to minimize registration error (fig. S25). An initial estimate of inter-frame motion is obtained using multichannel NCC, computed independently for each unmixed channel (*62*) and averaged to provide a robust global estimate of displacement, as previously described (*36*).

In parallel, local feature correspondences are extracted to enable more general (non-translational) motion estimation when required. For each channel, features in both frames are identified using the speeded-up robust features (SURF) descriptor (*63*), and candidate matches

are obtained using an approximate nearest-neighbors search with relaxed matching criteria (*64*). Matches from all channels are pooled into a single cohort, leveraging the spatial co-localization of spectral channels to increase robustness, and a single geometric transformation is computed and applied across all channels.

Following the approach of Yin *et al*. (*65*), matches are organized into a two-dimensional histogram according to their x and y displacement vectors. Peaks in this histogram represent clusters of matches with consistent motion estimates. The five largest peaks (feature groups) are evaluated using a cost function that quantifies alignment consistency:

$$\text{Alignment Cost} = \text{NCC Cost} + \text{Acceleration Cost} + \text{Size Cost}.$$

Here, NCC Cost is defined as the difference between the displacement of the feature group and the NCC-derived estimate. Acceleration Cost is defined as the difference between the direction of motion of the feature group and that of the previous frame alignment, where motion direction is computed as $\tan^{-1}(\Delta y/\Delta x)$; this term enforces temporal smoothness of probe motion and penalizes abrupt directional changes. Size Cost reflects the relative size of the feature group compared to the largest group in the set.

Minimization of this cost function selects a feature group that (i) agrees with the NCC-based estimate, (ii) maintains smooth temporal evolution of motion, and (iii) contains a large number of matched features. From the selected group, the geometric transformation is estimated using the MLESAC method (*66*), with the transformation model adaptively chosen (translation or affine) based on inlier support and residual error, rather than being restricted a priori.

Image stitching is performed using a graph-cut method (*67*, *68*), as previously implemented for mosaic construction (*69*), based on the work of Boykov and colleagues (*70*, *71*). Individual channel-wise mosaics are generated using this grayscale graph-cut implementation and subsequently combined to form the multiplexed mosaic.

To further improve robustness, a hierarchical alignment strategy is employed in which NCC provides a fast initial estimate, and feature-based refinement is invoked selectively. In practice, the majority of frame pairs can be aligned accurately using NCC alone; however, errors occur when motion deviates from pure translation. Such misalignments typically manifest as abrupt changes in the magnitude or direction of estimated motion, resulting in spikes in inferred acceleration (fig. S26).

Misaligned frame pairs are automatically identified as outliers in the estimated acceleration using MATLAB's *isoutlier* function. Because true probe motion during handheld imaging is typically smooth, with gradual directional changes over successive frames, these outliers provide a reliable indicator of alignment failure.

For these detected cases, the filtered SURF-based method is applied to refine alignment. This refinement step introduces additional degrees of freedom via affine transformations, enabling correction of rotational, shear, and scale components of motion, thereby improving overall mosaic accuracy while preserving stability in well-aligned regions (fig. S27).

Tissue phantom for mosaicking validation

A heterogeneous tissue phantom was prepared using a "virtual H&E" staining protocol adapted from Cahill *et al*. (*72*) . Thinly sliced chicken breast muscle was stained with acridine orange (200 μM) and sulforhodamine 101 (60 μM) in 50% ethanol for 10 min, rinsed in 50%

ethanol for 1 min, and mounted on a cavity slide (S175201A, Fisher Scientific) using O.C.T. compound (23-730-571, Fisher Scientific).

Human epithelial cancer cells (OVCAR5) were fixed in 4% paraformaldehyde, quenched with 0.1 M glycine, permeabilized with 0.1% (v/v) Triton X-100, and labeled with Alexa Fluor NHS 610 or 633 ester in 0.1 M sodium bicarbonate buffer. Labeled cells were deposited onto the stained tissue to generate a composite sample mimicking fluorescence microendoscopy of tumor cells over a structured stromal background. The combination of diffuse tissue texture and sparse, high-contrast cellular features provides a controlled testbed for evaluating mosaicking accuracy under conditions that challenge feature detection, matching, and alignment.

Imaging was performed using a Zeiss LSM 880 confocal microscope equipped with nonlinear excitation. The sample was excited at λ_{exc} = 890 nm, and emitted fluorescence (400–700 nm) was collected using a 32-channel spectral detector. Time-series imaging was acquired at maximum speed (0.471 s per spectral image cube) while the stage was translated in an approximate raster pattern, generating a sequence of overlapping fields of view for mosaicking.

A manually constructed ground-truth mosaic was used as a benchmark to evaluate mosaicking performance. Alignment error was quantified as the Euclidean distance between each estimated frame position and its corresponding ground-truth position (instantaneous error), and as the cumulative deviation of each frame from its true location over the sequence (cumulative error).

Power spectral analysis of spatial heterogeneity

Spatial heterogeneity of biomarker expression was quantified by two-dimensional Fourier power spectral analysis of wide-field unmixed image mosaics (Figs. 5A and 5B; figs. S28A–C, 29A, 29B, and S30). Square regions of mosaicked, unmixed images were multiplied by a Hamming window to reduce spectral leakage prior to Fourier transformation (*73*). The squared magnitude of the discrete Fourier transform was radially averaged to obtain isotropic power spectra as a function of spatial period. Power spectra were plotted in log–log space and fit with a power-law model, $P(f) \propto 1/f^{\Gamma}$, where Γ denotes the spectral slope. Higher Γ values indicate that a larger fraction of variance occurs at longer spatial scales, consistent with greater large-scale heterogeneity (Fig. 5C; figs. S28D, S29C, and S29D). As a control, pixel-scrambled versions of the unmixed mosaics were generated and analyzed using the same procedure. The two-dimensional discrete Fourier transform was computed, and the squared magnitude of the transform was calculated for each spatial frequency, $1/(2n)$, where n ranges from 1 to 2^L and L is the image side length in pixels. The resulting power spectrum was obtained by rotational averaging. These values were then plotted against their corresponding length scales, converted to μm using the pixel size of the microendoscope.

Image quality metrics

For the comparison of SMIRC and direct confocal imaging without the fiber bundle, image quality metrics were computed from the unmixed CD4 channel. Root-mean-square (RMS) contrast was calculated as the standard deviation of pixel intensities divided by the full 8-bit dynamic range of the unmixed CD4 channel (256). Signal-to-background ratio (SBR) was calculated as:

$$\mathrm{SBR} = 20\log_{10}\left(\frac{I_{\max} - \mu_{\mathrm{bg}}}{\sigma_{\mathrm{bg}}}\right),$$

where $I_{\max}$ is the maximum pixel intensity, and μ_{bg} and σ_{bg} are the mean and standard deviation of background pixels, respectively, identified using Otsu's three-class thresholding with third class sent to background. Statistics were computed across all frames for each imaging condition.

Fluorophore–antibody and –ligand conjugates for SMIRC *in vivo*

Monoclonal antibodies (mAbs) were conjugated with N-hydroxysuccinimide ester chemistry to Alexa Fluor dyes (Thermo-Fisher) for 2 hours at room temperature before size-exclusion column purification. This reaction scheme led to the following fluorophore:antibody loading ratios: 5 for anti-mouse CD45 mAb (Biolegend, clone 30-F11; rat IgG2b κ); 4 for anti-human/mouse EGFR mAb (ImClone LLC, cetuximab; chimeric human/mouse IgG1); 8 for anti-human/mouse CD44 mAb (Bioxcell IM7; rat IgG2b κ); 4 for anti-CA125 mAb (MUC16; Abcam, X75; mouse IgG1); 7 for anti-human VEGF mAb (Roche, bevacizumab; humanized IgG1); 2 for anti-mouse CD8a mAb (BioLegend, clone 53-6.7, rat IgG2a, κ); 2 for IgG1 isotype control mAb (BioLegend RTK2071, rat IgG1, κ); 2 for IgG2a isotype control mAb (BioLegend RTK2758, rat IgG2a, κ); and, 2 for IgG2b isotype control mAb (BioLegend RTK4530, rat IgG2b, κ). The following fluorophore–mAb and –protein conjugates were purchased pre-conjugated to fluorophores: APC/Cyanine7–anti-mouse CD3 mAb conjugate (BioLegend, 17A2; rat IgG2b, κ); Alexa Fluor (AF) 647–anti-mouse CD4 mAb conjugate (BioLegend, GK1.5; rat IgG2b κ); AF647–peanut agglutinin conjugate (Thermo-Fisher L32460); and, AF633–transferrin conjugate (Thermo-Fisher T23362). The injected masses of the conjugates for *in vivo* imaging of the xenograft pancreatic tumor (AsPC1) and peritoneal metastasis (OVCAR5) mouse models were (Figs. 1b,c,d and 3a): 31 μg anti-CD45 mAb; 12.5 μg cetuximab; 5 μg anti-CD44 mAb; 25 μg anti-CA125 (MUC16) mAb; 13 μg bevacizumab; 1 μg peanut agglutinin; and, 20 μg transferrin. The injected masses of the conjugates for *in vivo* phenotyping of the xenograft peritoneal metastasis (OVCAR5) mouse model were (Supplementary Data Fig. 24a,b,c): 7.5 μg anti-CD45 mAb; 273 μg cetuximab; 10 μg anti-CD44 mAb; and, 20 μg transferrin. The injected masses of the conjugates for *in vivo* imaging of the syngeneic pancreatic tumor (KPC) mouse model were (Figs. 2a–e and 3b, as well as Supplementary Data Fig. 24d): 15 μg anti-CD3 mAb; 12.5 μg anti-CD45 mAb; 37.5 μg anti-CD4 mAb; and, 165 μg anti-CD8a mAb. The masses of isotype control antibodies used for the OVCAR5 and KPC models were: 35 μg IgG1 mAb; 48.6 μg IgG2a mAb; and, 73 μg IgG2b mAb. After synthesis, these conjugates were mixed into cocktails and injected *i.p.* in 1 ml of phosphate-buffered saline. The fluorophores used for each conjugate (listed by biomarker target) were as follows: AF633 for transferrin receptor; AF647 for CD4, T-antigen, VEGF, or CA125 (MUC16); AF660 for CD44; AF680 for CD45; AF700 for CD8 or EGFR; and, APC/Cyanine7 for CD3.

In vivo models and photodynamic priming

All animal procedures were approved by the Massachusetts General Hospital Institutional Animal Care and Use Committee. Peritoneal carcinomatosis was modeled by intraperitoneal injection of OVCAR5 cells (Fox Chase Cancer Center) into female Swiss nude mice (3.15×10^7 cells) (*74*, *75*). Orthotopic xenograft pancreatic tumors were established by injection of AsPC1 cells (ATCC CRL-1682; 10^6 cells in 50 μL BD Sciences Matrigel mixture) into the pancreas of male Swiss nude mice (*18*). For immunocompetent studies, FC1245 cells (KPC-derived; obtained from Dr. David Tuveson's laboratory, Cold Spring Harbor Laboratory; 5×10^4 cells in

50% Matrigel) were implanted orthotopically into C57BL/6J mice, yielding 50–75 mm^3 tumors within ~8 days (*32*, *76*, *77*).

Fluorophore–antibody and –ligand conjugates were administered intraperitoneally 1 h prior to imaging unless otherwise noted. For minimally invasive imaging, a 1-mm catheter was inserted into the peritoneal cavity and the microendoscope gently positioned against tissues of interest. Mice were anesthetized with isoflurane (3.5% in oxygen). Repeated imaging was performed over multiple days without observable adverse effects.

Photodynamic priming (PDP) was performed using verteporfin (0.25 mg/kg) (*33*, *78*). For peritoneal metastases, 690 nm laser irradiation (ML6500, Modulight) was delivered intraperitoneally via a diffusing fiber (90 mW/cm^2 for 55 s; 5 J/cm^2). For orthotopic KPC tumors, verteporfin was administered intravenously and tumors were irradiated at 100 mW/cm^2 for 750 s (75 J/cm^2) following a 60-min drug–light interval.

In vivo imaging and photodynamic priming experimental details

Minimally invasive imaging of the orthotopic pancreatic tumors was performed 43 days post-inoculation for the xenograft model (AsPC1) and 8 days post-inoculation for the syngeneic model (KPC). After insertion of the catheter, the microendoscope was manually guided toward the tumor, which was externally palpable by touch. Pancreatic tumors were identified intraoperatively by their strong fluorescence relative to surrounding tissues. Peritoneal carcinomatosis imaging was performed 9–14 days (Fig. 1D; figs. S14A S15, S27, S31A, and S31B), 16 days (figs. S28, S29A) and 20 days (Fig. 5A; fig. S29B and S30) post-inoculation. These time points were chosen to demonstrate imaging of early and late-stage dissemination. When examining treatment effects, PDP was performed on day 8 for the KPC model and on day 10 for the peritoneal carcinomatosis model. To provide an internal control, PDP was performed on the region of the peritoneal cavity near the right hind leg only. The left hind leg region was used as an untreated comparison. For mosaic imaging of peritoneal carcinomatosis and pancreatic tumors, the peritoneal cavity was surgically exposed and visible, mimicking open peritoneal surgery. The microendoscope was glided over a section of the peritoneal wall or pancreatic tumor for several seconds and SMIRC image cubes captured at 16.7 Hz.

Ex vivo histology and immunofluorescence

Serial cryosections of tissue biopsies from the xenograft (peritoneal wall or pancreas) and syngeneic (pancreas) models were stained *ex vivo* using standard H&E or immunofluorescence protocols. Briefly, primary antibodies for EGFR (Abcam, EP38Y), VEGF165 (R&D Systems, AF-293-NA), MUC16 (Cell Signaling Technologies, 29623), TfR (Abcam, ab84036), CD44 (BioXCell, IM7), CD45 (BioLegend, 30-F11; and, Abcam, ab10558), CD3 (Abcam, SP162), CD4 (ThermoFisher Scientific, PA5-87425; and, ThermoFisher Scientific, GK1.5), CD8a (BioLegend, 53-6.7), and CD11b (Abcam, EPR19387) were each used at ~10 μg/mL followed by overnight incubation at 4 °C. This was followed by incubation with their corresponding fluorophore conjugated secondary antibodies—goat anti-rabbit IgG H&L–AF488 conjugate (Abcam, ab150077), goat anti-rat IgG (H+L) cross-adsorbed secondary antibody–AF546 conjugate (ThermoFisher Scientific, A-11081), and donkey anti-goat IgG (H+L) cross-adsorbed secondary antibody–AF546 (ThermoFisher Scientific, A-11056). Rat IgG1 κ isotype (BioLegend, RTK2071), rat IgG2a κ isotype (BioLegend, RTK2758), and rat IgG2b κ isotype (BioLegend, RTK4530) were used as non-specific control *ex vivo* immunostains with goat anti-rat IgG (H+L) cross-adsorbed secondary antibody–AF546 conjugate (ThermoFisher Scientific,

A-11081). Rat IgG1-R-phycoerythrin (PE; BioLegend, RTK2071) was also used as a non-specific control *ex vivo* immunostain for select experiments. All mAbs were diluted in buffer (Dako Antibody Diluent with Background Reducing Components), and tissue sections were incubated with blocking buffer (Dako Protein Block) before staining. Tissue sections were then mounted using ProLong™ Glass Antifade Mountant with NucBlue™ Stain (ThermoFisher Scientific). H&E color imaging and immunofluorescence imaging of the tissue sections was performed using a commercial slide-scanning microscopy system (Hamamatsu NanoZoomer 2.0-RS) with a 20× 0.75 NA objective. The tumor sections were imaged in 40× mode (0.23 μm · $pixel^{-1}$) using a broadband lamp, a FITC filter cube (485 nm excitation, 525 nm emission) for Alexa 488, and a TRITC filter cube (560 nm excitation, 607 nm emission) for Alexa 546.

Cell count computation from SMIRC and *ex vivo* biomarker images

All *in vivo* and *ex vivo* SMIRC images of the KPC PDAC mouse model were spectrally unmixed and analyzed for cell counts. *In vivo* images were acquired with the microendoscope positioned at the tumor periphery. Following imaging, tumors were excised, sectioned perpendicular to the *in vivo* imaging plane, stained, and imaged *ex vivo* across the full tissue slice.

To facilitate comparison between *in vivo* and *ex vivo* measurements, tumor regions in *ex vivo* sections were identified using serial hematoxylin and eosin (H&E)–stained sections. A region adjacent to the tumor boundary (~35 μm width) was subdivided into tiled subregions (~0.055 mm^2 each), which were analyzed independently for cell counts. For comparison of SMIRC and direct confocal imaging without the fiber bundle, images were acquired from freshly excised tumors with either the microendoscope or a microscope objective oriented toward the same tissue surface.

Cell density (cells/mm^2) was quantified using CellProfiler (*79*). Identical computational steps were applied to both *in vivo* and *ex vivo* datasets. Image thresholding was performed using Otsu's three-class method (*80*), with the first and second classes assigned to foreground. A lower bound was applied to the threshold to suppress background signal and to avoid false cell identification in empty images. Threshold parameters were subsequently adjusted by manual inspection across representative images within each experimental group to ensure consistent cell detection.

Cell objects were identified following thresholding. Clumped cells were separated using local intensity maxima with a smoothing filter applied to reduce over-segmentation. Local maxima separated by less than ~3 μm were suppressed. Identified objects were constrained to diameters between ~5 μm and ~33 μm. In densely packed regions where individual cell boundaries are difficult to resolve, fluorescence intensity maxima corresponding to individual cells can be used to estimate cell counts, as implemented in CellProfiler (*81*) and previously validated for CD4+ and CD8+ T cell quantification in tissue sections (*82*).

For identification of CD3+/CD4+ and CD3+/CD8+ T cells *in vivo*, CD3+ cells were first segmented as described above. CD4+ or CD8+ channels were then converted into binary masks and used to filter CD3+ objects. A CD3+ cell was classified as CD3+/CD4+ or CD3+/CD8+ if >70% of its area overlapped with the corresponding CD4+ or CD8+ mask.

Co-expression patterns observed in this study and their relationship to cell population estimates are presented in figs. S19–S24. Marker selection and combinations used for immune cell phenotyping were based on established expression profiles reported in prior literature. CD3 is a highly specific T cell marker (*83*) and exhibits relatively sparse staining pattern compared to

the other markers. CD4 is expressed by T helper cells as well as monocytes, macrophages, and dendritic cells (*84*). CD8 is predominantly expressed by cytotoxic T cells but can also be found on natural killer cells and dendritic cells (*85*, *86*). CD11b is a canonical marker of myeloid lineage cells, including monocytes, macrophages, and certain dendritic cell populations, and is generally not expressed by T lymphocytes (*87*). CD45 is a pan-leukocyte marker expressed across myeloid and lymphoid lineages (*88*). Prior studies of KPC mouse models demonstrate variability in immune cell infiltration across KPC tumor cell lines (*89*) and a higher immune cell density at the tumor periphery (analyzed here) relative to the tumor core (*90*).

Equivalent diameter measurements derived from segmentation outputs were computed for all identified cell populations. Distributions of segmented cell diameters for each marker channel are shown in fig. S23. These measurements were compared to reported size ranges for immune cell types in prior studies. Monocytes and other myeloid cells typically exhibit diameters in the range of ~15–22 μm (*91*). Circulating T cells typically exhibit smaller diameters of ~6–8 μm, which increase to ~10–11 μm upon activation (*92*, *93*).

Nuclear-stain–assisted *ex vivo* cell count computation

We performed the following nuclear stain-based cell identification algorithm using the CellProfiler program. The composite image is split into two grayscale images, one with the nuclear stain intensities and one with the antibody stain intensities. All cells are first identified via the nuclear stain image by identifying nuclei. Nuclear stain image thresholding prior to cell identification is done using Otsu's three class method with first and second class sent to foreground and an adaptive window of ~23 μm. A manual minimum threshold is applied to avoid false cell identification in empty images for each experimental group. A manual threshold smoothing scale and threshold correction factor is applied to tune the threshold to correctly identify all cells for each experimental group. Cell identification is done after thresholding. Clumped nuclei are separated using local intensity maxima. A smoothing filter is applied for clump separation to avoid over-segmentation. Local maxima closer than ~1.2 μm are suppressed. Identified nuclei must have a diameter between ~1.5 μm and ~17 μm. To identify cells that are positive for the antibody stain, a biomarker mask is created from the antibody stain image. Thresholding to create the biomarker mask is done using Otsu's three class method with first and second class sent to foreground. A manual minimum threshold is applied to avoid false cell identification in empty images for each experimental group. A manual threshold smoothing scale and threshold correction factor is applied to tune the threshold to correctly identify all cells for each experimental group. The identified nuclei are then masked using the biomarker mask such that nuclei that do not overlap with the antibody mask are discarded. For T cell identification of CD3+CD4+ or CD3+CD8+ cells *ex vivo*, two biomarker masks are used to mask the identified nuclei: CD3+ CD4+ masks for CD3+CD4+ cells, and CD3+ and CD8+ masks for CD3+CD8+ cells.

Statistical analysis

Specific statistical tests and sample sizes are indicated in the figure captions. Analyses were performed in GraphPad Prism (GraphPad Software) or MATLAB (Mathworks). All reported *P* values are two-tailed.

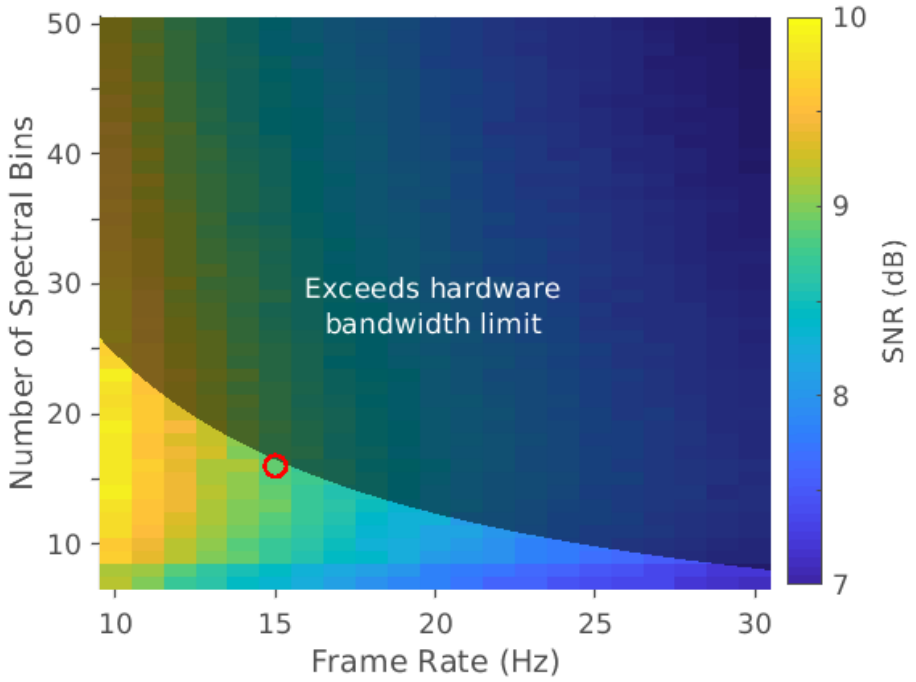

Fig. S1.
Spectral measurement basis determination. Spectral measurements were simulated to highlight the tradeoffs between imaging speed and spectral resolution. Monte Carlo calculations modeled spectral unmixing in the photon-starved regime influenced by illumination conditions and optical collection efficiency. These points were informed by Zemax modeling. Increasing frame rate or the spectral wavelength resolution reduced the number of photons per wavelength bin and increased the photon shot noise. Reducing spectral resolution blurred out spectral features and degraded the spectral unmixing. The combination of these effects led to the choice of measurement conditions with acceptable frame rate. The shaded area demarcates data acquisition rates larger than the maximum ability of the computer to transfer data to memory and constrained the target design regime (red circle). The simulated wavelength bins are spread out over 650 nm to 800 nm. Each point in the figure is calculated from 10,000 iterations. Alexa Fluor 633, 647, 660, 680 and 700 fluorophores were combined at concentrations designed to equalize their fluorescent contributions *in vivo* (~1 μM) for the calculation shown here.

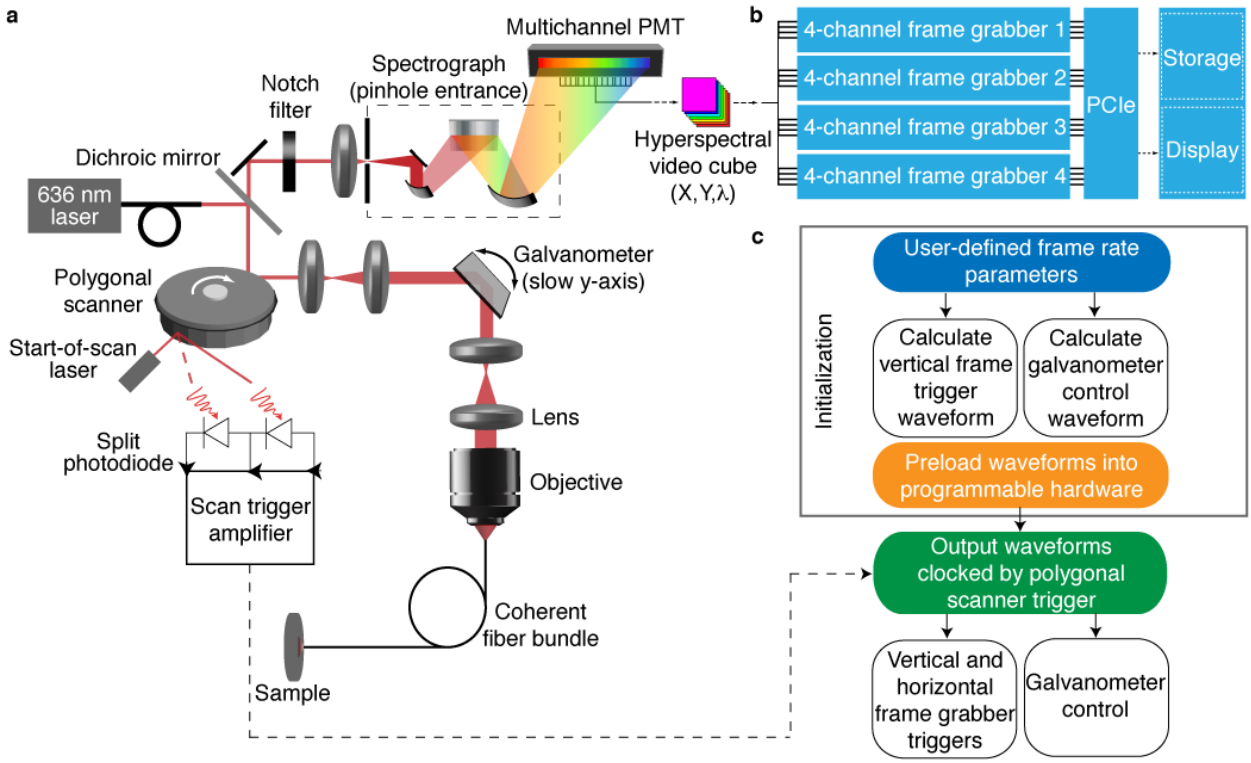

Fig. S2.

SMIRC layout. (**A**) The opto-mechanics consists of a fiber coupled diode laser beam that is expanded and scanned over the back aperture of an objective, which couples the light into a coherent fiber bundle. The collected fluorescence is de-scanned and focused through a pinhole to remove background light. A notch filter is used to attenuate the excitation wavelength, and the remaining light passes through a spectrograph. The spectrally dispersed light is then collected by a multichannel photomultiplier tube (PMT). To detect the start of the polygonal mirror x-scan, a second off-optical axis laser is detected by a split photodiode. The resulting start of scan (SOS) signal is run through a high slew rate hysteretic comparator to remove noise before being sent to the scan engine to time the galvanometric mirror and frame grabbers. (**B**) The fluorescence data stream is captured by a set of multichannel frame grabbers. The data stream is directly routed to memory and display components via the Peripheral Component Interconnect Express (PCIe) bus. (**C**) To synchronize the polygonal x-scan with the galvanometric mirror y-scan and the image formation, a scan engine receives the SOS timing pulse and generates an analog waveform which synchronizes the motion of the galvanometric y-scan. Simultaneously, a series of digital pulses is sent to the frame grabber initiating data acquisition and providing spatial information for the image cube reconstruction. Uniquely, the scan engine is initialized at run-time, allowing the scan parameters to be changed on the fly.

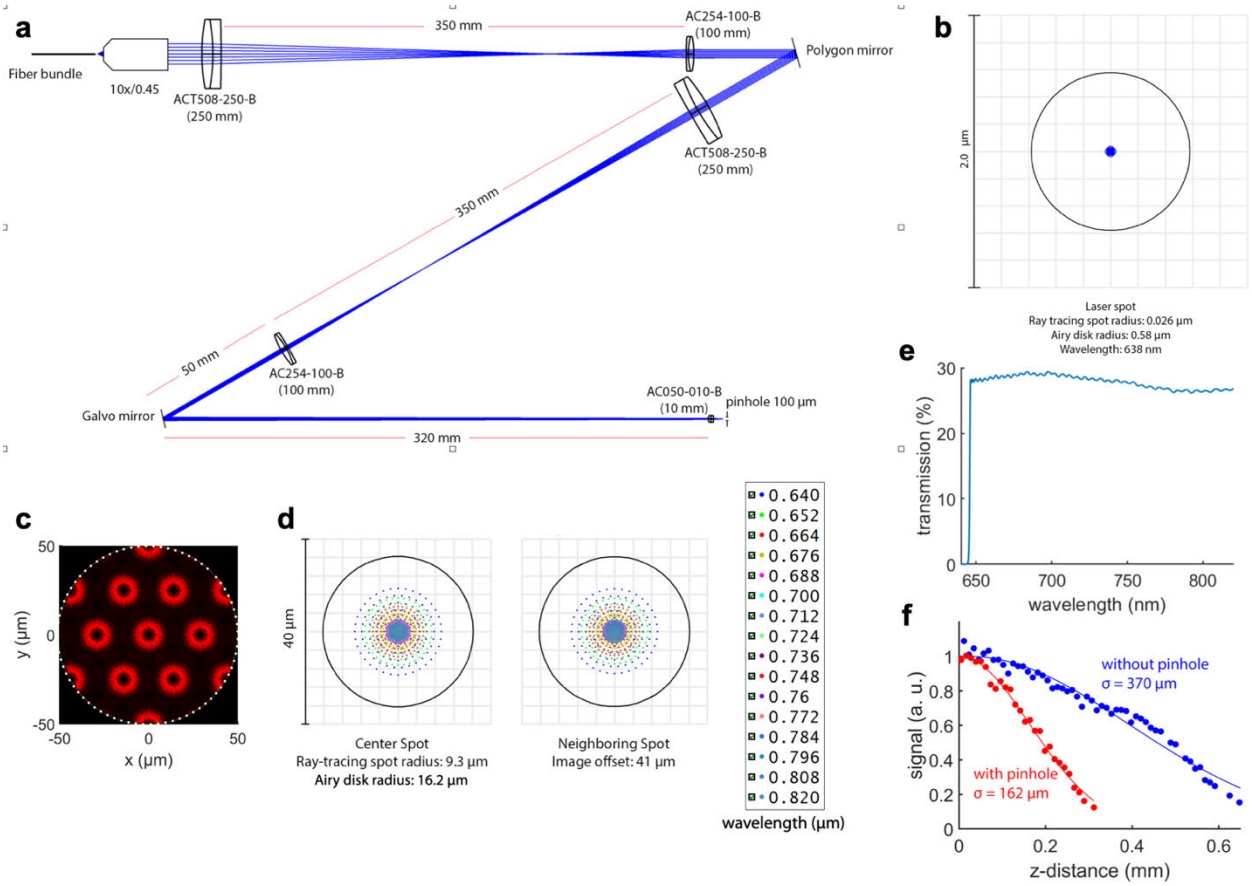

Fig. S3.

Semi-confocal, wide-pinhole axial resolution and achromatic performance. (**A**) The optical design using Zemax modeling (excluding the coherent fiber bundle). Lens product numbers (Thorlabs; Newton, New Jersey) are provided and the focal distance for each is listed in parentheses. Distances between lens elements are precise to 5 mm. Due to the lack of access to proprietary information, the microscope objective was simulated using an ideal paraxial lens (higher order aberrations were not modeled) with parameters matching the known focal distance, pupil size and position, and working distance (provided by a Zeiss engineer). (**B**) Simulated (Zemax) ray-trace spot diagram of the excitation laser at the plane of the coherent fiber bundle proximal end. The excitation laser spot is diffraction-limited (<1 µm Airy disk) and smaller than the diameter of a single fiber core (~3 µm), indicating efficient laser coupling. (**C**) Simulated fiber core image at the pinhole position (assuming each fiber core is a Gaussian source). A central fiber core at each laser scan position is injected with excitation light and emission from the central core and an additional layer of surrounding cores pass through the pinhole. (**D**) Simulated ray-trace diagram of wavelength-dependent spot sizes at the pinhole transmitted from the proximal fiber tip (from a single, central fiber core approximated as a point source), indicating diffraction-limited performance across the multispectral detection bandwidth. The 32-µm-diameter Airy disk at the pinhole plane indicates ~3 Airy unit operation (100 µm pinhole) compared to ideal confocal operation (~1 Airy unit, matched pinhole size to the zeroth order portion of the Airy disc in the image plane). (**E**) Simulated wavelength-dependent transmission efficiency through the confocal pinhole and downstream optics to the detector array plane across the multispectral detection range. The Zemax-modeled transmission accounts for losses from all

optical components, including the objective lens, relay optics, dichroic mirror, notch filter, diffraction grating, and the coherent fiber bundle. Component-specific, wavelength-dependent efficiencies were taken from manufacturer datasheets; product numbers are provided in panel (A) and in the *Materials and Methods* (*SMIRC design and implementation*). Because wavelength-dependent transmission data are not available for the coherent fiber bundle, it was conservatively modeled as a wavelength-independent 60% transmission for near-on-axis, low-NA photons; transmission decreases with incidence angle, such that multiply scattered light is expected to be transmitted less efficiently. This assumption is based on reported measurements of >60% transmission at 488, 543, and 633 nm(*35*). Transmission losses associated with the photomultiplier tube (PMT) quantum efficiency are not included in this model. (**F**) Measured axial resolution of the semi-confocal, wide-pinhole SMIRC system using 1 μm fluorescent microspheres in a weakly-scattering phantom (agarose gel), imaged through the distal end of the coherent fiber bundle. The sigma values of Gaussian fits are shown. The detector pinhole helps to eliminate out-of-focus light. Note that the intensity profile is attenuated more strongly with depth for *in vivo* imaging due to limited probe penetration and increased scattering within tissue.

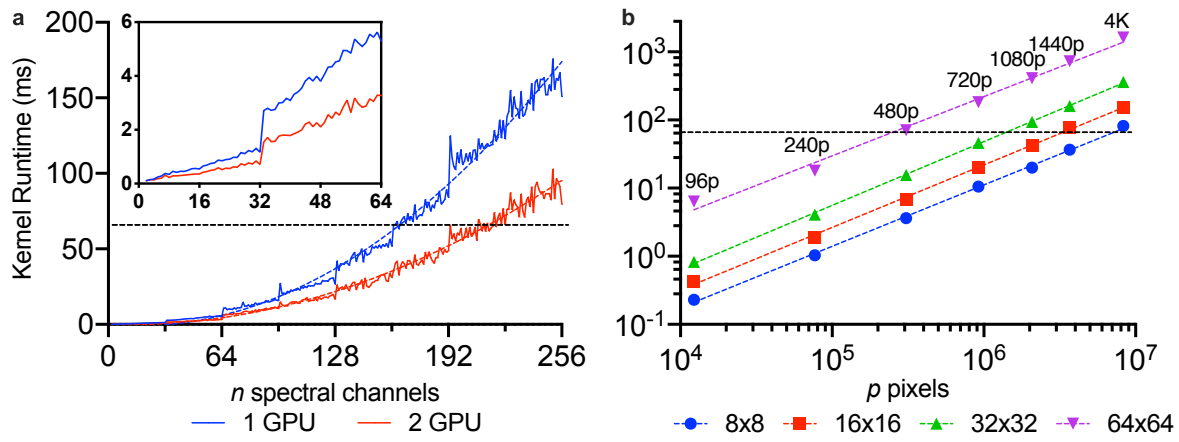

Fig. S4.
GPU-accelerated fast non-negative least squares (GPU-FNNLS) algorithm for high-speed spectral unmixing. Pseudo-random hyperspectral data cubes and endmember matrices were timed (10 trials per data cube) under GPU-FNNLS analysis (error bars too small for display). Black dashed lines indicate the 15-fps threshold, below which (shorter kernel runtimes) we consider the algorithm video-rate. (**A**) Runtime as a function of the number of spectral bands where the system is kept square ($m = n$, $p = 10^4$). Data are fit to quadratic trendlines (dashed, colored), which were found to be significantly better than a linear fit as determined by a sum-of-squares F test (*P<0.0001), demonstrating a quadratic dependence on the hyperspectral system size. Inset shows the first n = 2–64 trials. (**B**) Kernel runtimes (ms) computed with two GTX 1080Ti graphics cards for a range of standard image resolutions and system sizes common to hyperspectral cameras, with linear regression trendlines (colored, dashed lines). The slope of each fit (in ms/pixel) was converted to the algorithm's throughput (in GB/s): 3.26 ± 0.03 GB/s (8×8), 3.43 ± 0.10 GB/s (16×16), 2.95 ± 0.02 GB/s (32×32), and 1.31 ± 0.00 GB/s (64×64).

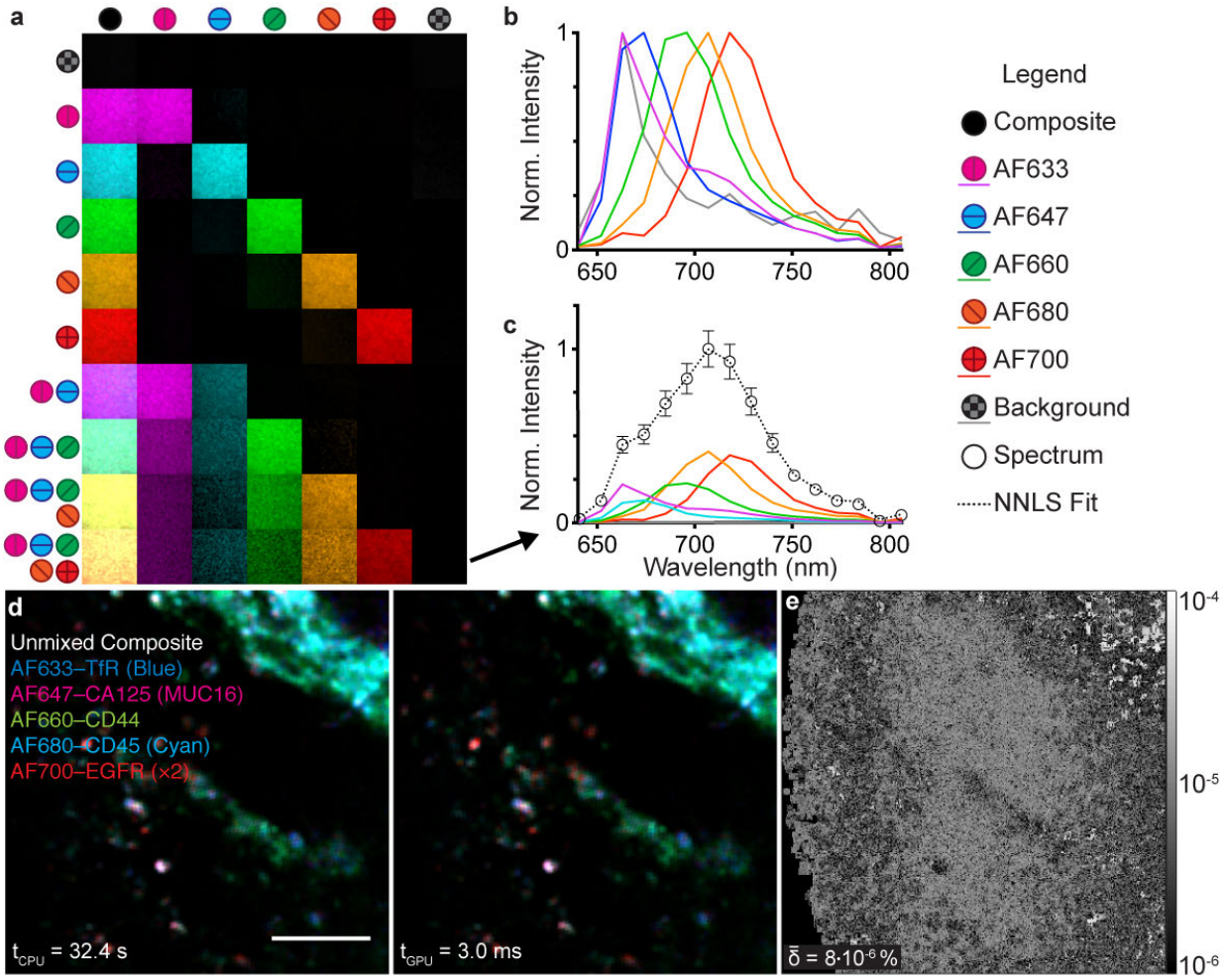

Fig. S5.

Video-rate spectral decomposition of 5 fluorophores imaged using SMIRC. (**A**) Measurements of single- and multi-fluorophore solutions *in vitro* resolve several spectrally distinct fluorescent dyes simultaneously after NNLS decomposition (16.7 fps acquisition time, 28-frame average, 200 × 200 μm per square). Each row contains a different combination of dyes indicated by symbols on the left. Subsequent columns to the right show the composite unmixing result and pseudo-colored basis maps corresponding to individual fluorophores labeled by symbols above. (**B**) Basis spectra library and (**C**) average pixel spectrum ± S.D. ($n = 1.5 \times 10^6$ pixels) for a 5-fluorophore mixture with the NNLS solution corresponding to the bottom row of **a** (indicated by arrow) demonstrate accurate decomposition of 6 basis spectra (5 dyes plus a mean background signature that includes reflected laser light). Legend applies to (A–C). (**D**) An exemplary in vivo hyperspectral image from a mouse model of metastatic epithelial ovarian cancer. Composite pseudo-colored results after CPU (left) or GPU (right) spectral decomposition are shown, with unmixing times indicated for CPU (FNNLS–single thread) and GPU (2 TITAN Vs) configurations. (**E**) Relative difference between the CPU and GPU unmixed results in (D); a mean of 8×10^{-6}% difference (averaged over the 5 tumor biomarker channels). (D and E) Scale bar, 100 μm. AF, Alexa Fluor; TfR, transferrin receptor; CA125 (MUC16), cancer antigen 125

(mucin 16); CD44, homing cell adhesion molecule (a receptor for hyaluronic acid); CD45, protein tyrosine phosphatase receptor type C; EGFR, epidermal growth factor receptor.

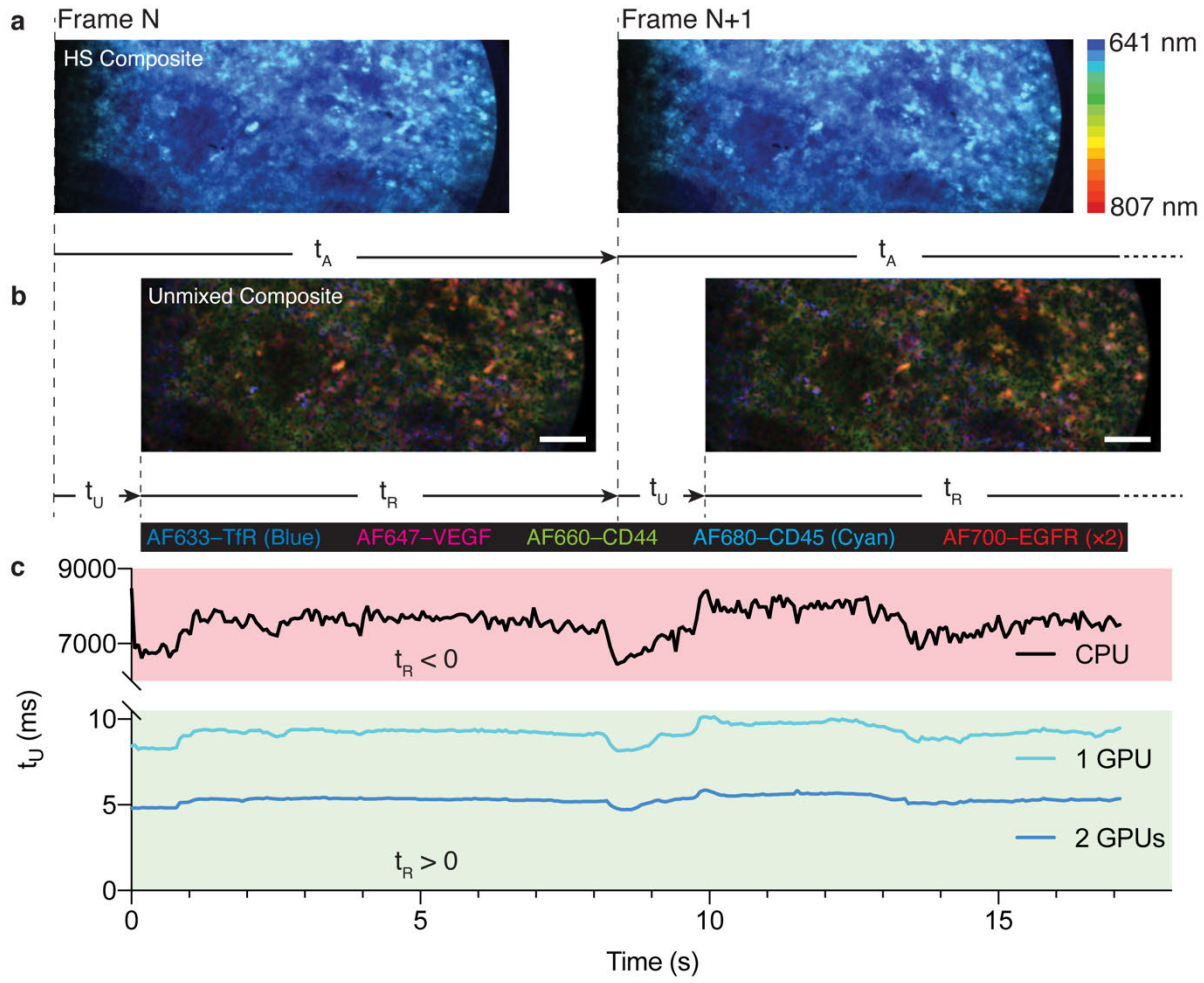

Fig. S6.

Sustained video-rate hyperspectral image cube decomposition. (**A**) Schematic of an *in vivo* SMIRC video cube stream of a metastatic mouse model of epithelial ovarian cancer acquired at acquisition time t_A per cube (frame). Composite raw image cubes, pseudo-colored by wavelength for 2D representation, are dominated by reflected laser light (blue pseudo-colored channels). (**B**) Analyzed images are available for display at unmixing time t_U after the raw image acquisition. This leaves a residual time $t_R = t_A - t_U$ during which the GPU is idle and may be used for subsequent image processing. If $t_R > 0$, real-time analysis is sustainable. Scale bar, 100 μm. (**C**) A SMIRC video cub stream (t_A = 60 ms per frame, 286 frames) was analyzed using a CPU (FNNLS-multi-thread) versus 1 or 2 GTX 1080Ti graphics cards. The unmixing time t_U per frame is plotted over the video acquisition time scale. GPU analysis is adequate to sustain real-time unmixing ($t_R > 0$), whereas CPU analysis is not. AF, Alexa Fluor; TfR, transferrin receptor; VEGF, vascular endothelial growth factor; CD44, homing cell adhesion molecule (a receptor for hyaluronic acid); CD45, protein tyrosine phosphatase receptor type C; EGFR, epidermal growth factor receptor.

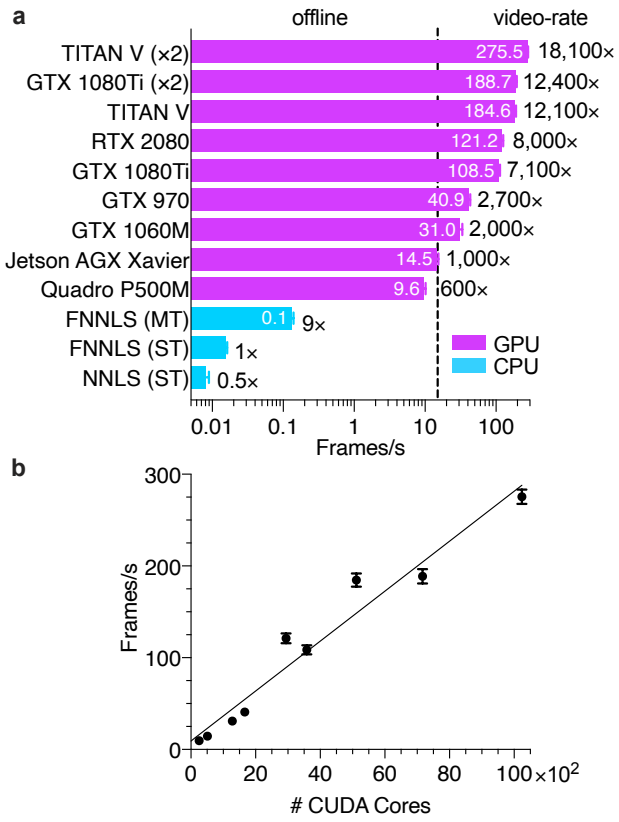

Fig. S7.

GPU-FNNLS benchmarks. (**A**) Hyperspectral video was analyzed using GPU-FNNLS across multiple CPU and GPU configurations. All but one GPU configuration reached video-rate speed, whereas CPU configurations were generally inadequate for online analysis. Results are the mean ± S.D. unmixing speed of the 286 frames in the video, plus fold increase in speed relative to the analogous CPU implementation. (**B**) Unmixing frame rates from (A) correlate with the number of CUDA cores available for parallel computation. Linear regression indicates 2.72 ± 0.01 fps are gained per hundred CUDA cores (intercept = 9.14 ± 0.59, $R^2 = 0.95$).

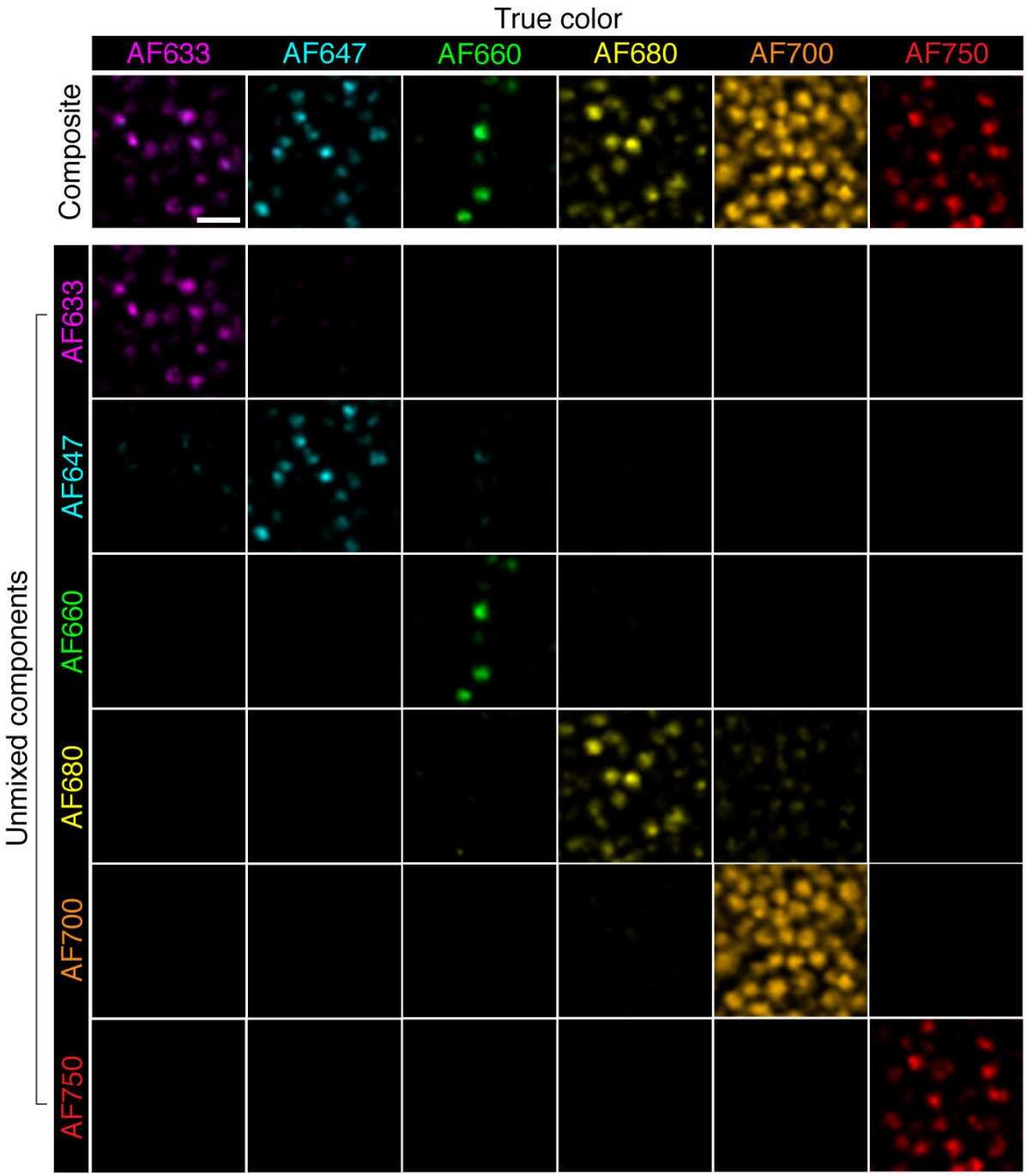

Fig. S8.
Single-cell resolution validation *in vitro*. Fixed OVCAR5 cells were directly conjugated to Alexa Fluor (AF) dyes with NHS-ester coupling and imaged with SMIRC. 6 different conjugated cell groups were created, listed in the top title row. The composite images are the summation of all photons collected by the system. In the matrix below, the composite images are unmixed, with each row indicating the unmixed contribution of the respective fluorophore (listed on the left title bar). Diagonal elements indicate correct unmixing analysis. Each image was taken with a 60 ms exposure. Scale bar, 20 μm.

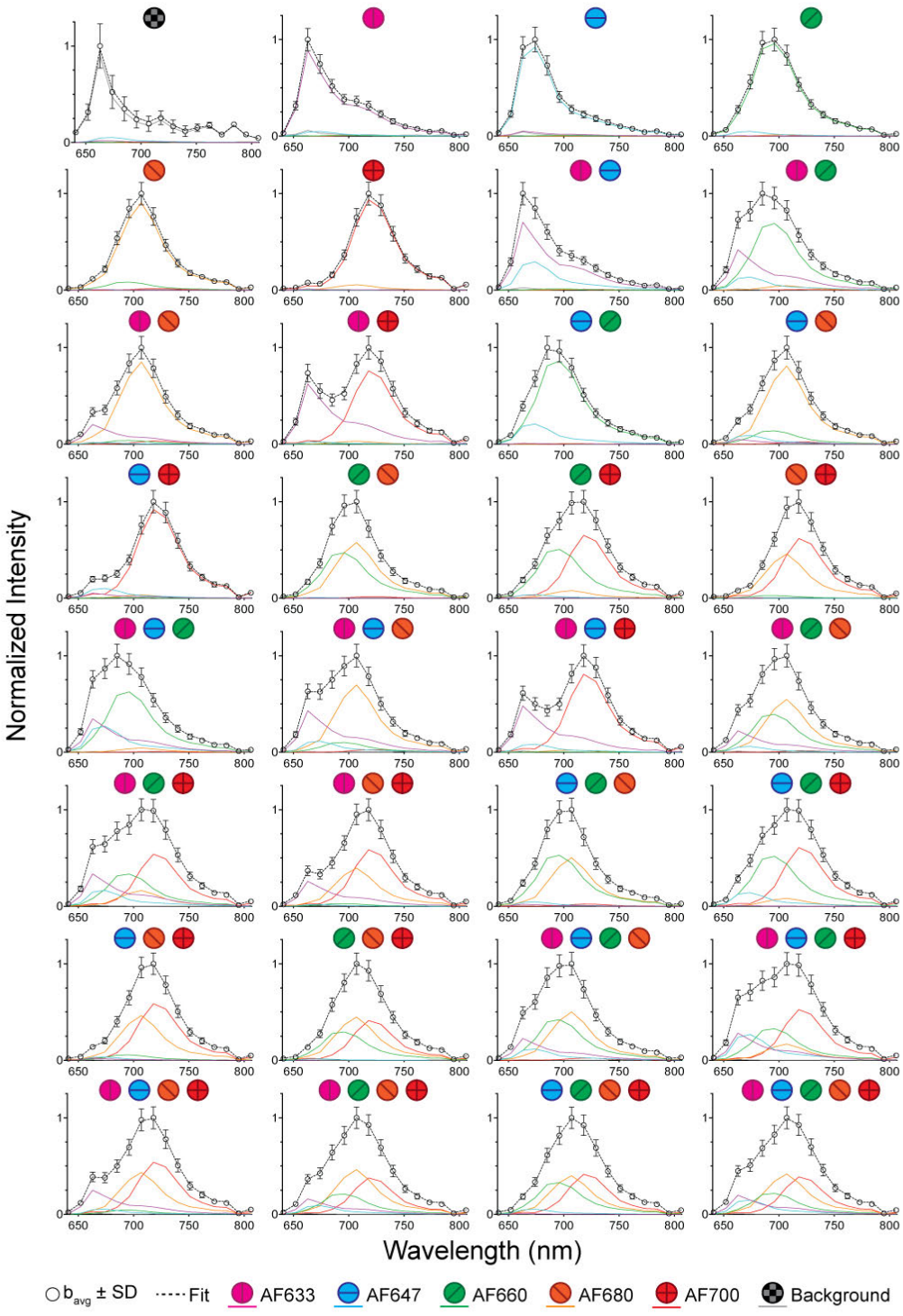

Fig. S9.
Spectral unmixing of 5 fluorophores in solution. The 32 possible permutations of 5 Alexa Fluor (AF) dyes were combined at concentrations of 0.1–10 μM and imaged with SMIRC. The raw spectra with the unmixed fluorophore spectra are shown, with each plot representing a unique fluorophore combination indicated by symbols at the top of each plot. Each fit was computed from a single 60 ms exposure. Error bars are standard deviation.

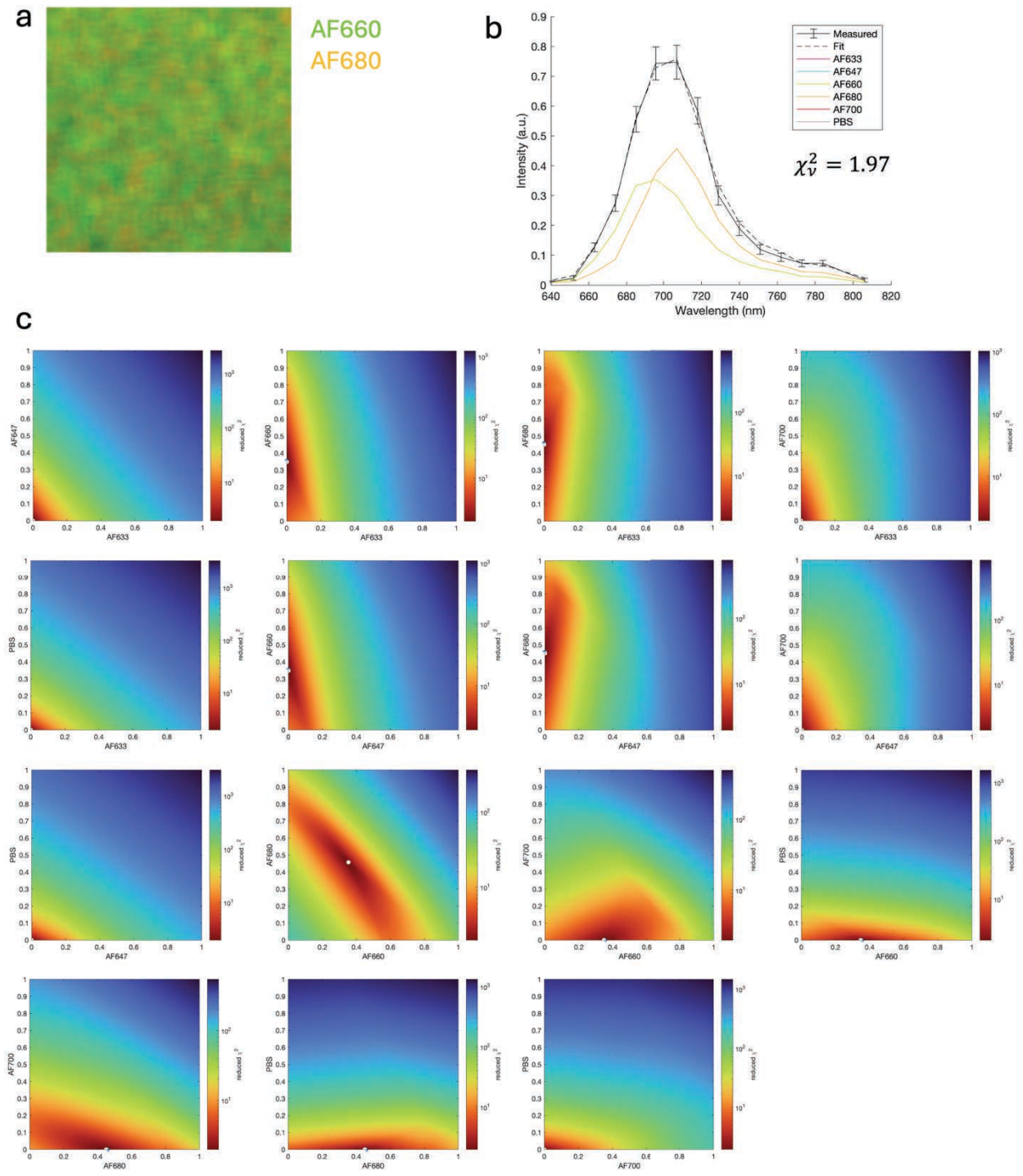

Fig. S10.
Multispectral unmixing goodness-of-fit for 2 spectrally overlapping fluorophores in solution using a basis set of 5 fluorophores. (**A**) Unmixed image of AF660 and AF680 dyes mixed in solution, imaged with the SMIRC microendoscope, and unmixed using AF633, AF647, AF660, AF680, AF700, and phosphate buffered saline (PBS) as basis spectra. Image dimensions, 110 μm × 110 μm (100 pixels × 100 pixels). (**B**) Measured emission spectra (black solid line), GPU-accelerated fast non-negative least squares (FNNLS) spectral decomposition fit to the measured

spectrum (dotted black line), and basis spectra contributions to the fit for AF633 (magenta), AF647 (cyan), AF660 (lime), AF680 (orange), AF700 (red), and PBS (grey) are shown for the center pixel (50, 50) of the image shown in (A). Reduced χ^2 (χ_ν^2) is 1.97 with 9 degrees of freedom (15 usable spectral channels – 6 basis spectra). AF660 and AF680 are accurately resolved, indicating that SMIRC can accurately identify cells labelled with 2 colocalized fluorescent species. Error bars represent standard deviations obtained via error propagation from the raw per-pixel uncertainty, modeled as $\sigma_{i,c} = \sqrt{m\, y_{i,c}}$ based on the empirically measured variance–mean relationship (slope m, Fig. 2A), where $y_{i,c}$ is the measured signal for pixel i in spectral channel c. This per-pixel uncertainty is propagated through microendoscope oversampling correction, spatial Gaussian filtering to remove the fiber-core pattern (low-pass spatial frequency filtering), and spectral image cube normalization. (**C**) χ_ν^2 heatmaps are shown for the fit in (B). In each map, the basis spectra contribution value for two dyes is fixed and the FNNLS fit is recalculated. χ_ν^2 is calculated for each new fit and plotted in the heatmap. The best-fit dye contribution values from the FNNLS fit are marked with a blue dot and the minimum χ_ν^2 value on the plot is indicated with a white dot.

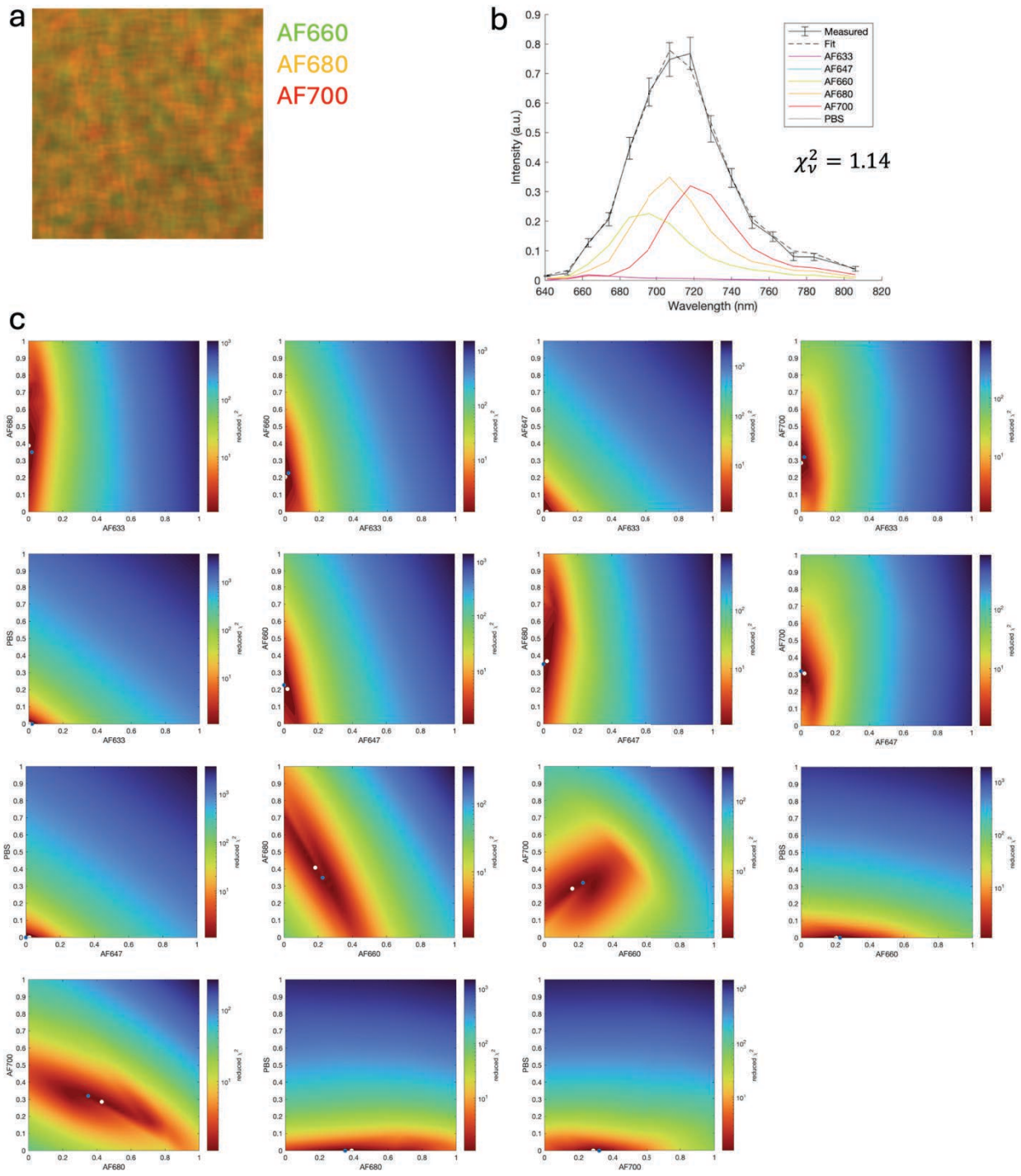

Fig. S11.
Multispectral unmixing goodness-of-fit for 3 spectrally overlapping fluorophores in solution using a basis set of 5 fluorophores. (**A**) Unmixed image of AF660, AF680, and AF700 dyes mixed in solution, imaged with the SMIRC microendoscope, and unmixed using AF633, AF647, AF660, AF680, AF700, and PBS as basis spectra. Image dimensions, 110 μm × 110 μm (100 pixels × 100 pixels). (**B**) Measured emission spectra (black solid line), GPU-accelerated fast non-negative least squares (FNNLS) spectral decomposition fit to the measured spectrum (dotted

black line), and basis spectra contributions to the fit for AF633 (magenta), AF647 (cyan), AF660 (lime), AF680 (orange), AF700 (red), and PBS (grey) are shown for the center pixel (50, 50) of the image shown in (A). Reduced χ^2 (χ_ν^2) is 1.14 with 9 degrees of freedom (15 usable spectral channels – 6 basis spectra). AF660, AF680, and AF700 are accurately resolved, indicating that SMIRC can accurately identify cells labelled with 3 colocalized fluorescent species. Error bars represent standard deviations obtained via error propagation from the raw per-pixel uncertainty, modeled as $\sigma_{i,c} = \sqrt{m\, y_{i,c}}$ based on the empirically measured variance–mean relationship (slope m, Fig. 2A), where $y_{i,c}$ is the measured signal for pixel i in spectral channel c. This per-pixel uncertainty is propagated through microendoscope oversampling correction, spatial Gaussian filtering to remove the fiber-core pattern (low-pass spatial frequency filtering), and spectral image cube normalization. (**C**) χ_ν^2 heatmaps are shown for the fit in (B). In each map, the basis spectra contribution value for two dyes is fixed and the FNNLS fit is recalculated. χ_ν^2 is calculated for each new fit and plotted in the heatmap. The best-fit dye contribution values from the FNNLS fit are marked with a blue dot and the minimum χ_ν^2 value on the plot is indicated with a white dot.

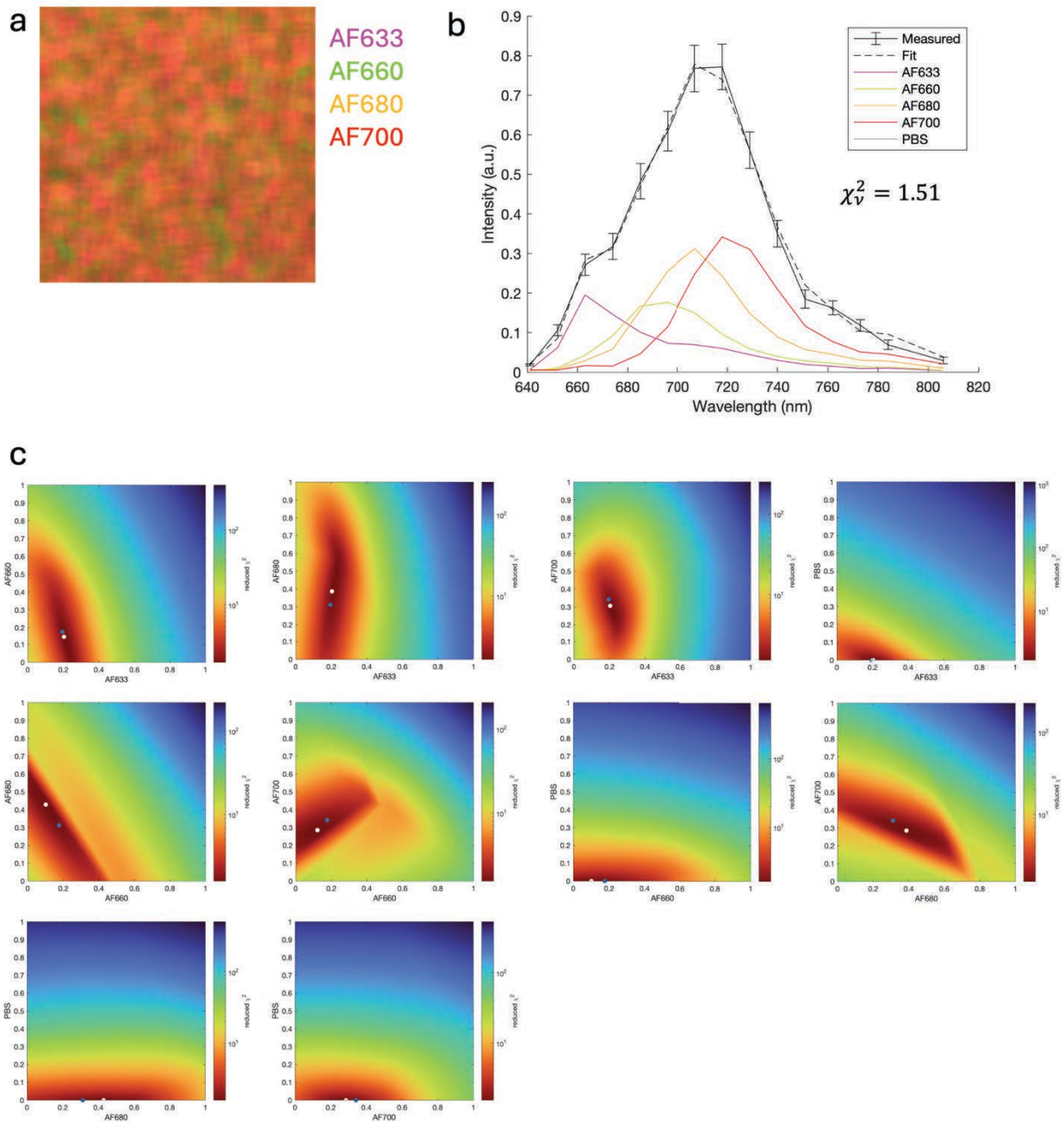

Fig. S12.
Multispectral unmixing goodness-of-fit for 4 spectrally overlapping fluorophores in solution using a basis set of 4 fluorophores. (**A**) Unmixed image of AF633, AF660, AF680, and AF700 dyes mixed in solution, imaged with the SMIRC microendoscope, and unmixed using AF633, AF660, AF680, AF700, and PBS as basis spectra. Image dimensions, 110 µm × 110 µm (100 pixels × 100 pixels). (**B**) Measured emission spectra (black solid line), GPU-accelerated fast non-negative least squares (FNNLS) spectral decomposition fit to the measured spectrum (dotted black line), and basis spectra contributions to the fit for AF633 (magenta), AF660 (lime), AF680 (orange), AF700 (red), and PBS (grey) are shown for the center pixel (50, 50) of the image shown in (A). Reduced χ^2 (χ_ν^2) is 1.51 with 10 degrees of freedom (15 usable spectral channels –

5 basis spectra). All 4 dyes are accurately resolved, indicating that SMIRC can accurately identify cells labelled with 4 colocalized dyes when using a 4-dye basis set. Error bars represent standard deviations obtained via error propagation from the raw per-pixel uncertainty, modeled as $\sigma_{i,c} = \sqrt{m\, y_{i,c}}$ based on the empirically measured variance–mean relationship (slope m, Fig. 2A), where $y_{i,c}$ is the measured signal for pixel i in spectral channel c. This per-pixel uncertainty is propagated through microendoscope oversampling correction, spatial Gaussian filtering to remove the fiber-core pattern (low-pass spatial frequency filtering), and spectral image cube normalization. (**C**) χ_ν^2 heatmaps are shown for the fit in (B). In each map, the basis spectra contribution value for two dyes is fixed and the FNNLS fit is recalculated. χ_ν^2 is calculated for each new fit and plotted in the heatmap. The best-fit dye contribution values from the FNNLS fit are marked with a blue dot and the minimum χ_ν^2 value on the plot is indicated with a white dot.

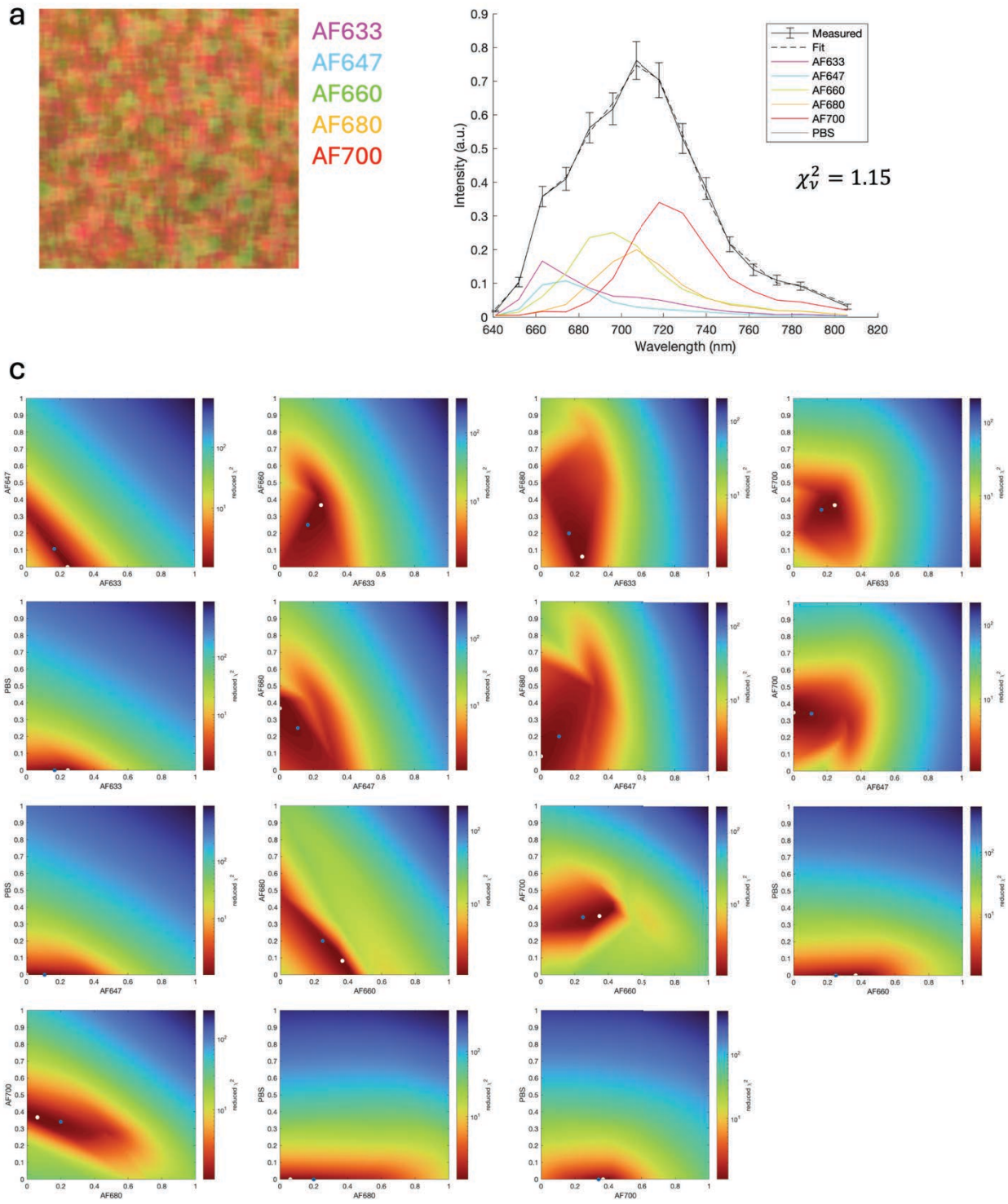

Fig. S13.
Multispectral unmixing goodness-of-fit for 5 spectrally overlapping fluorophores in solution using a basis set of 5 fluorophores. (**A**) Unmixed image of AF633, AF647, AF660, AF680, and AF700 dyes mixed in solution, imaged with the SMIRC microendoscope, and unmixed using AF633, AF647, AF660, AF680, AF700, and PBS as basis spectra. Image dimensions, 110 µm × 110 µm (100 pixels × 100 pixels). (**B**) Measured emission spectra (black solid line), GPU-

accelerated fast non-negative least squares (FNNLS) spectral decomposition fit to the measured spectrum (dotted black line), and basis spectra contributions to the fit for AF633 (magenta), AF647 (cyan), AF660 (lime), AF680 (orange), AF700 (red), and PBS (grey) are shown for the center pixel (50, 50) of the image shown in (A). Reduced χ^2 (χ_ν^2) is 1.15 with 9 degrees of freedom (15 usable spectral channels – 6 basis spectra). Error bars represent standard deviations obtained via error propagation from the raw per-pixel uncertainty, modeled as $\sigma_{i,c} = \sqrt{m\, y_{i,c}}$ based on the empirically measured variance–mean relationship (slope m, Fig. 2A), where $y_{i,c}$ is the measured signal for pixel i in spectral channel c. This per-pixel uncertainty is propagated through microendoscope oversampling correction, spatial Gaussian filtering to remove the fiber-core pattern (low-pass spatial frequency filtering), and spectral image cube normalization. (**C**) χ_ν^2 heatmaps are shown for the fit in (B). In each map, the basis spectra contribution value for two dyes is fixed and the FNNLS fit is recalculated. χ_ν^2 is calculated for each new fit and plotted in the heatmap. The best-fit dye contribution values from the FNNLS fit are marked with a blue dot and the minimum χ_ν^2 value on the plot is indicated with a white dot. All five AF dyes are identified by the FNNLS fit, but χ_ν^2 maps indicate χ_ν^2 minima that bring some dye contributions to 0, suggesting that SMIRC may not be able to accurately identify 5 colocalized dyes. If using 5 fluorescent dyes at once, dyes should be chosen such that a maximum of 3 spectrally overlapping dyes with a 5-dye basis set will stain the same cell to ensure unmixing accuracy.

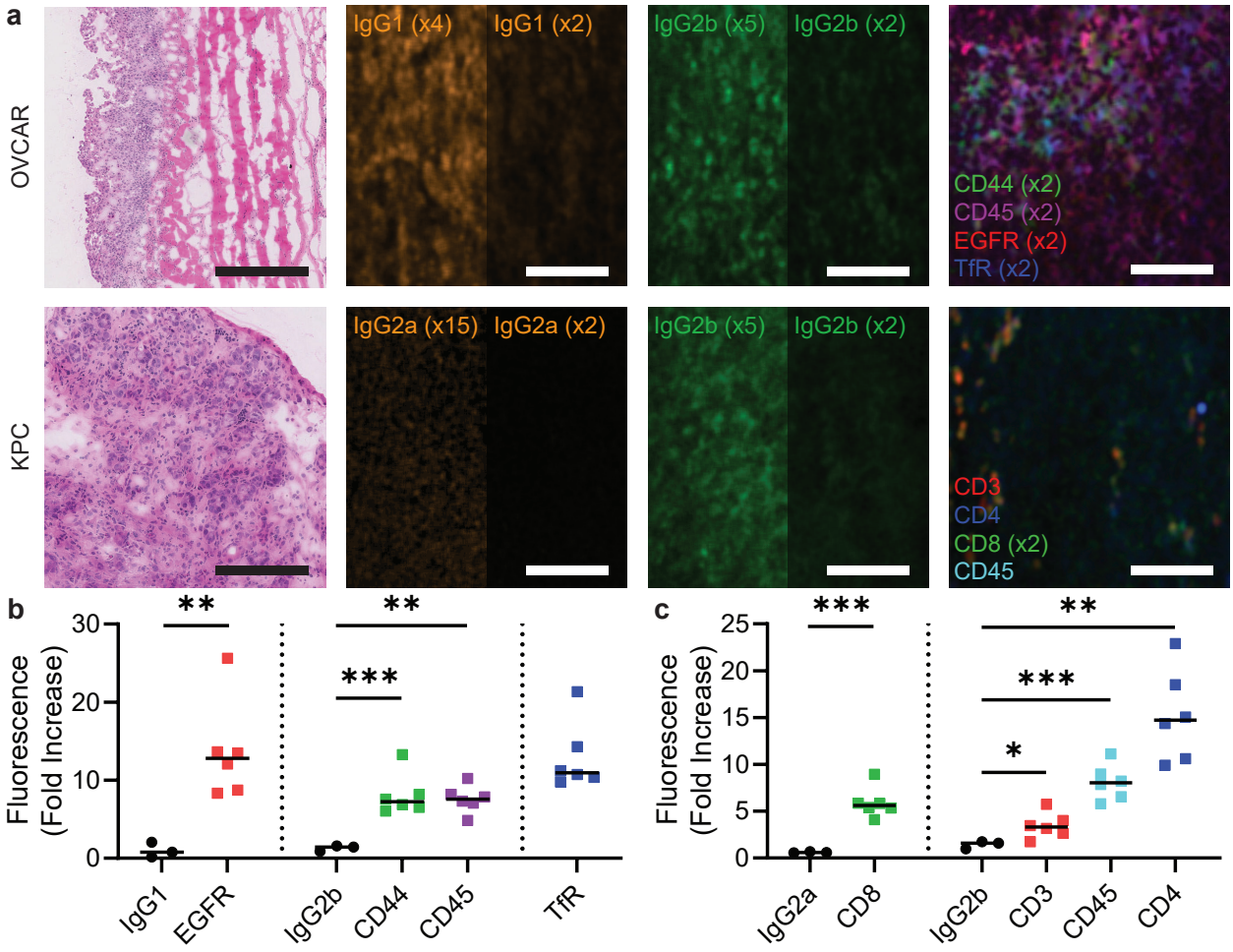

Fig. S14.

In vivo SMIRC imaging using non-specific antibody isotype–fluorophore conjugates to quantify non-specific background signal and demonstrate target-specific contrast. (**A**) *Top row, xenograft OVCAR5 mouse model:* hematoxylin and eosin (H&E)–stained peritoneal wall metastasis biopsy (*far left*); *in vivo* SMIRC image of a peritoneal metastasis in the same mouse (confirmed disease) using a non-specific control IgG1–AF700 antibody isotype–fluorophore conjugate (image with the maximum intensity shown, ~3× brighter than the mean non-specific control antibody isotype–fluorophore conjugate signal; image halves brightened 2× or 4×; acquired at laser power P = 70 mW), corresponding to the anti-EGFR, anti-CA125, and anti-VEGF antibody (IgG1)–fluorophore conjugates (*middle left*); *in vivo* SMIRC image of a peritoneal metastasis in the same mouse using non-specific control IgG2b–AF700 antibody isotype–fluorophore conjugate (image with the maximum intensity shown, ~2× brighter than the mean non-specific control antibody isotype–fluorophore conjugate signal; image halves brightened 2× or 5×; P = 70 mW) corresponding to anti-CD44 and anti-CD45 antibody (IgG2b)–fluorophore conjugates (*middle right)*; and, *in vivo* SMIRC image of a peritoneal metastasis (different mouse, H&E-confirmed disease; all biomarkers brightened 2×; P = 10 mW) using a panel of tumor biomarker-targeted antibody–fluorophore conjugates (*far right*). H&E scale bar, 400 μm; SMIRC, 100 μm. *Bottom row, syngeneic KPC mouse model:* H&E-stained primary pancreatic tumor biopsy (*far left*); *in vivo* SMIRC image of the pancreatic tumor in the same mouse using a non-specific control IgG2a–AF750 antibody isotype–fluorophore conjugate (image with the maximum intensity

shown, ~1.1× brighter than mean non-specific signal; image halves brightened 15× or 2×; $P = 50$ mW), corresponding to the anti-CD8a antibody (IgG2a)–fluorophore conjugate (*middle left*); *in vivo* SMIRC image of the pancreatic tumor in the same mouse using a non-specific control IgG2b–AF700 antibody isotype–fluorophore conjugate (image with the maximum intensity shown, ~2× brighter than the mean non-specific signal; image halves brightened 2× or 5×; $P = 50$ mW), corresponding to anti-CD3 (IgG2b), anti-CD4 (IgG2b) and anti-CD45 antibody (IgG2b)–fluorophore conjugates (*middle right)*; and, *in vivo* SMIRC image of a KPC pancreatic tumor (different mouse, H&E-confirmed disease; CD8 brightened 2×; $P = 40$ mW) using a panel of tumor biomarker-targeted antibody–fluorophore conjugates (*far right*). H&E scale bar, 200 μm; SMIRC, 100 μm. (**B**) Quantification of the cumulative fluorescence intensity of anti-EGFR, anti-CD44, and anti-CD45 antibody–fluorophore conjugates and anti-TfR ligand–fluorophore conjugates (normalized to excitation power and to their respective isotype controls) from *in vivo* SMIRC images of the peritoneal wall (A, top row) compared to non-specific control antibody isotype–fluorophore conjugates (IgG1 or IgG2b; $^{**}P<0.01$, $^{***}P<0.001$, one-way ANOVA with Dunnett's *post hoc* test). Notably, TfR-targeted signal is also significantly elevated relative to isotype controls (IgG1 or IgG2b, $P<0.01$). Data points represent individual mice ($n = 3$–6 mice per group; $n = 17$ images per mouse). (**C**) Quantification of cumulative fluorescence intensity of anti-CD3, anti-CD4, anti-CD8, and anti-CD45 antibody–fluorophore conjugates (normalized to excitation power and to isotype controls) from *in vivo* SMIRC images of syngeneic KPC pancreatic tumors compare to non-specific control antibody isotype–fluorophore conjugates (IgG2a or IgG2b; $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, one-way ANOVA with Dunnett's *post hoc* test). Data points represent individual mice ($n = 3$–6 mice per group; $n = 17$ images per mouse; solid lines indicate mean values). Together, these results demonstrate that non-specific accumulation of antibody isotype–fluorophore conjugates produces minimal background signal *in vivo*, whereas biomarker-targeted conjugates yield substantially higher fluorescence intensity. In pancreatic tumors and peritoneal metastases, targeted signals are typically 5–15× greater than isotype controls, confirming that SMIRC contrast arises predominantly from specific molecular binding rather than non-specific uptake.

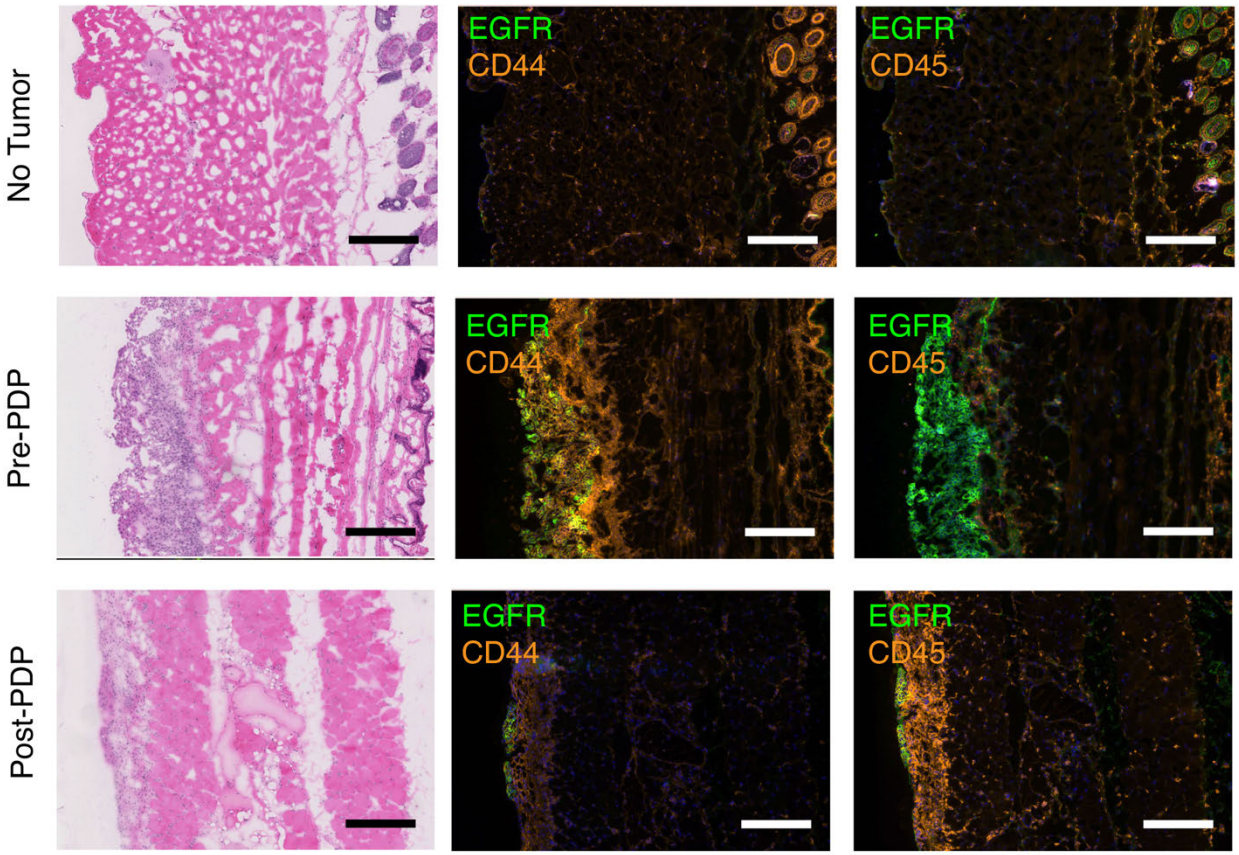

Fig. S15.
Validation of *in vivo* SMIRC imaging of the xenograft OVCAR5 peritoneal metastasis model *by ex vivo* histopathology and immunofluorescence staining. *Top row:* H&E biopsy tissue section of the peritoneal wall from a no-tumor control mouse; anti-EGFR and anti-CD44, and anti-EGFR and anti-CD45, immunostaining of serial, co-registered tissue sections adjacent to the H&E section. *Middle row:* H&E biopsy of the peritoneal wall from a xenograft OVCAR5 peritoneal metastasis mouse prior to PDT with the same *ex vivo* immunostains above applied to serial, co-registered tissue sections. *Bottom row:* H&E biopsy of the peritoneal wall from a xenograft OVCAR5 peritoneal metastasis mouse post-PDT with the same *ex vivo* immunostains above applied to serial, co-registered tissue sections. Note that the inner peritoneal wall faces left in each image and the right-hand side includes subcutaneous tissue with hair follicles, *etc*. Note that the immunofluorescence images also include a 4′,6-diamidino-2-phenylindole (DAPI) nuclear stain (blue). Scale bars, 250 μm.

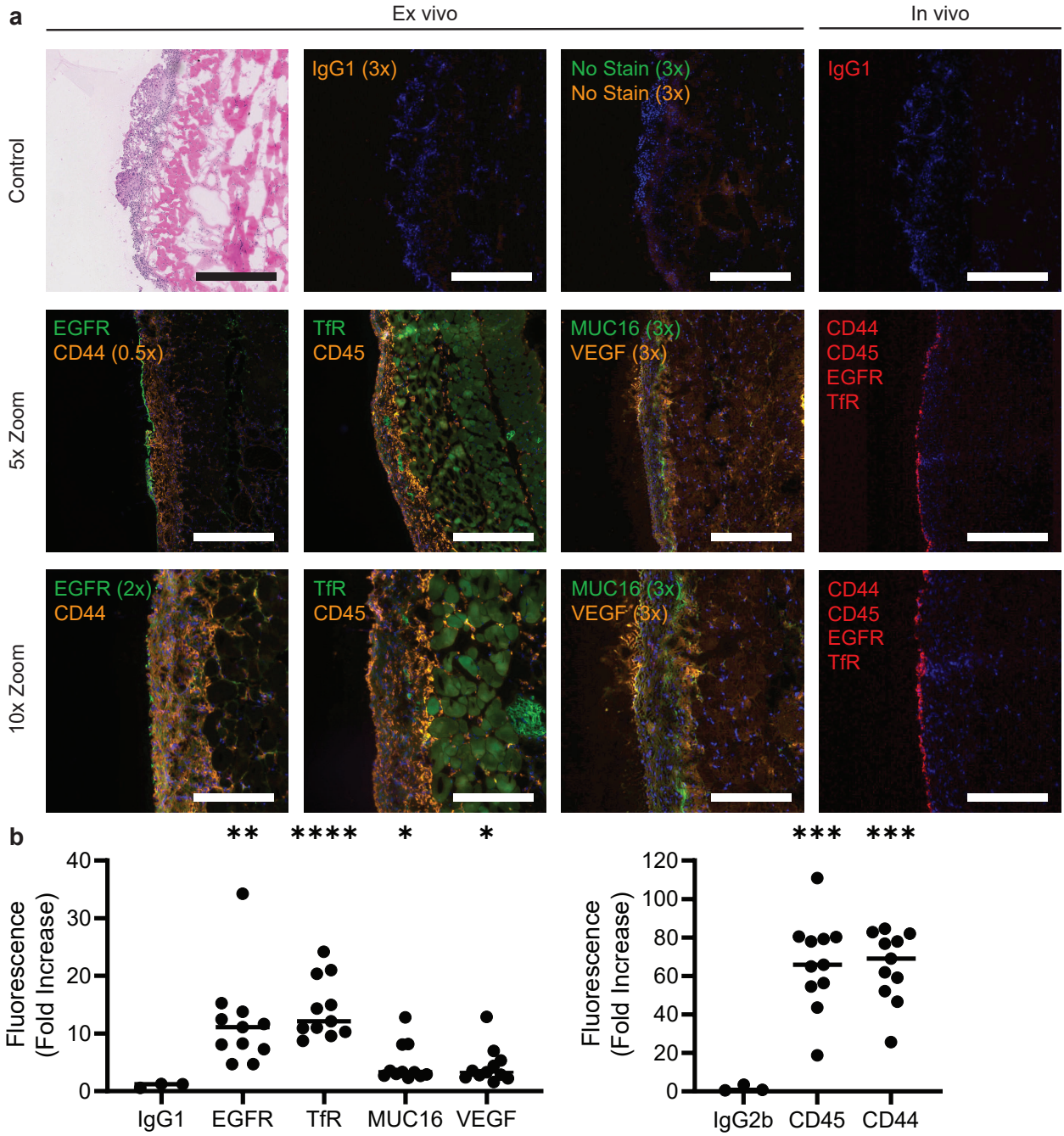

Fig. S16.
Validation of CD44, CD45, EGFR, TfR, MUC16, and VEGF tumor expression in the xenograft OVCAR5 peritoneal metastasis model *by ex vivo* immunofluorescence staining and *ex vivo* confirmation of the presence of the *in vivo* administered antibody- and ligand-fluorophore conjugates in the same tumors. (**A**) *Ex vivo, Control:* Serial sections from the peritoneal wall from a xenograft OVCAR5 peritoneal metastasis mouse, including a H&E biopsy tissue section, a control non-specific antibody isotype–fluorophore conjugate (IgG1-PE) immunostained section representative of *ex vivo* immunostain background, and unstained tissue sections (no stain) representative of autofluorescence background. *In vivo, Control: Ex vivo* image of the same serial section in a microscope channel that captures fluorescence of a control non-specific

antibody isotype–fluorophore conjugate (IgG1-AF700) administered *in vivo* representative of non-specific background. *Ex vivo, 5× Zoom and 10× Zoom:* Exemplary *ex vivo* immunostains of all six antibody- and ligand-fluorophore conjugates used in the xenograft OVCAR5 peritoneal mouse model confirm the presence of all six biomarkers in the tumor sites (brightened images are indicated). Note that TfR is expressed by adipocytes(*36*, *37*) and therefore appears within the peritoneal submesothelium (subserosa) below the tumor. *In vivo, 5× Zoom and 10× Zoom: Ex vivo* image of a serial section in a microscope bandpass channel (without spectral discrimination) that captures sum fluorescence of the antibody- and ligand-fluorophore conjugates administered *in vivo* (CD44, CD45, EGFR, and TfR) confirms that these probes stain the superficial layers of peritoneal tumors. Brightened channels are labeled as 0.5× to 3× with the label color corresponding to the biomarker. Scale bars for Control and 5× Zoom, 400 μm. Scale bars for 10× Zoom, 200 μm. Note the limited (~20 μm) penetration of the targeted fluorescent probes delivered *in vivo* (*i.p.* injection, 1 hour post-administration). (**B**) Quantification of the cumulative fluorescence intensity of each of the target tumor biomarkers within the *ex vivo* peritoneal wall tumor metastasis biopsies using anti-EGFR, anti-TfR, anti-MUC16, anti-VEGF, anti-CD44, and anti-CD45 antibodies versus their respective non-specific control antibody isotype (IgG1 or IgG2b) using appropriate secondary antibody–fluorophore conjugates (**P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001, one-way ANOVA with Dunnett's *post hoc* test). Data points represent one or more tissue sections analyzed per individual mouse (n = 3–5 mice per group; n = 1–3 peritoneal wall whole biopsy cross-section images per mouse, ~5–10 mm^2 in dimension per image; solid lines indicate mean values). A serial H&E section was used to locate the tumor boundary for each tissue section analyzed.

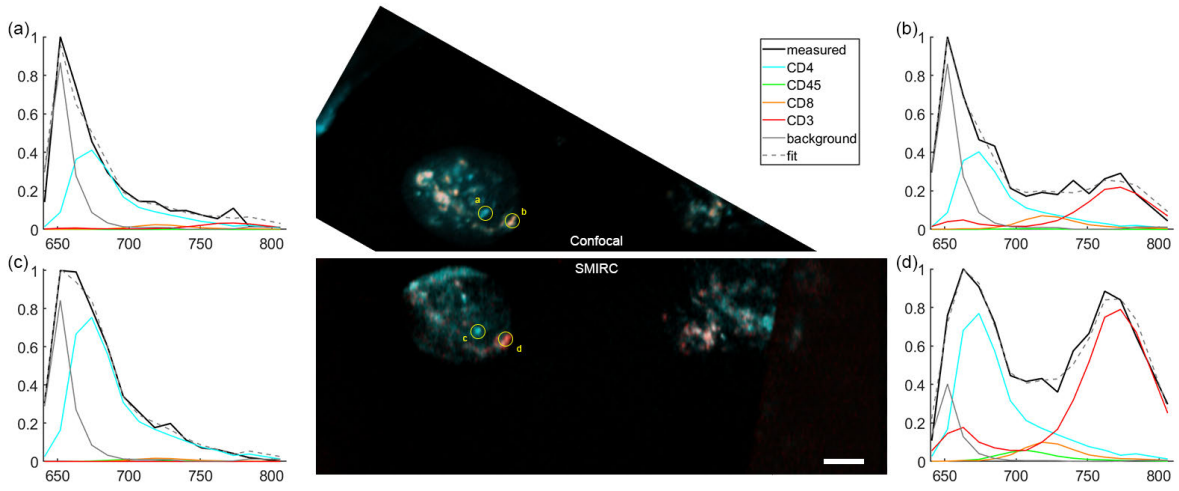

Fig. S17.
Comparison of SMIRC and direct confocal imaging without the fiber bundle: single-pixel spectra for CD3+/CD4+ versus CD3**–**/CD4+ cells within a co-registered field of view. Freshly excised pancreatic tumors from KPC syngeneic, orthotopic mouse models were imaged following intraperitoneal administration of antibody–fluorophore conjugates at 17 fps. Representative spectrally unmixed images from matched tissue regions acquired by direct confocal imaging (top) and SMIRC imaging through the fiber bundle (bottom) are shown. Scale bar, 50 μm. The confocal image was rotated using cross-correlation–based alignment to match the orientation of the corresponding fiber-based image; minor discrepancies arise from differences in optical point-spread function and imaging depth. (**A** and **B**) Processed single-pixel emission spectra (*Spectral cube acquisition and spatial filtering, Materials and Methods*) of a CD3–/CD4+ cell (A) and a CD3+/CD4+ cell (B) acquired using direct confocal imaging. (**C** and **D**) Corresponding processed single-pixel spectra from the same CD3–/CD4+ cell (C) and CD3+/CD4+ cell (D) acquired using SMIRC. For each spectrum, the measured emission (solid black line), GPU-accelerated fast non-negative least squares (FNNLS) fit (dotted black line), and individual basis spectra contributions for CD4 (cyan), CD45 (lime), CD8 (orange), CD3 (red), and PBS background (grey) are shown.

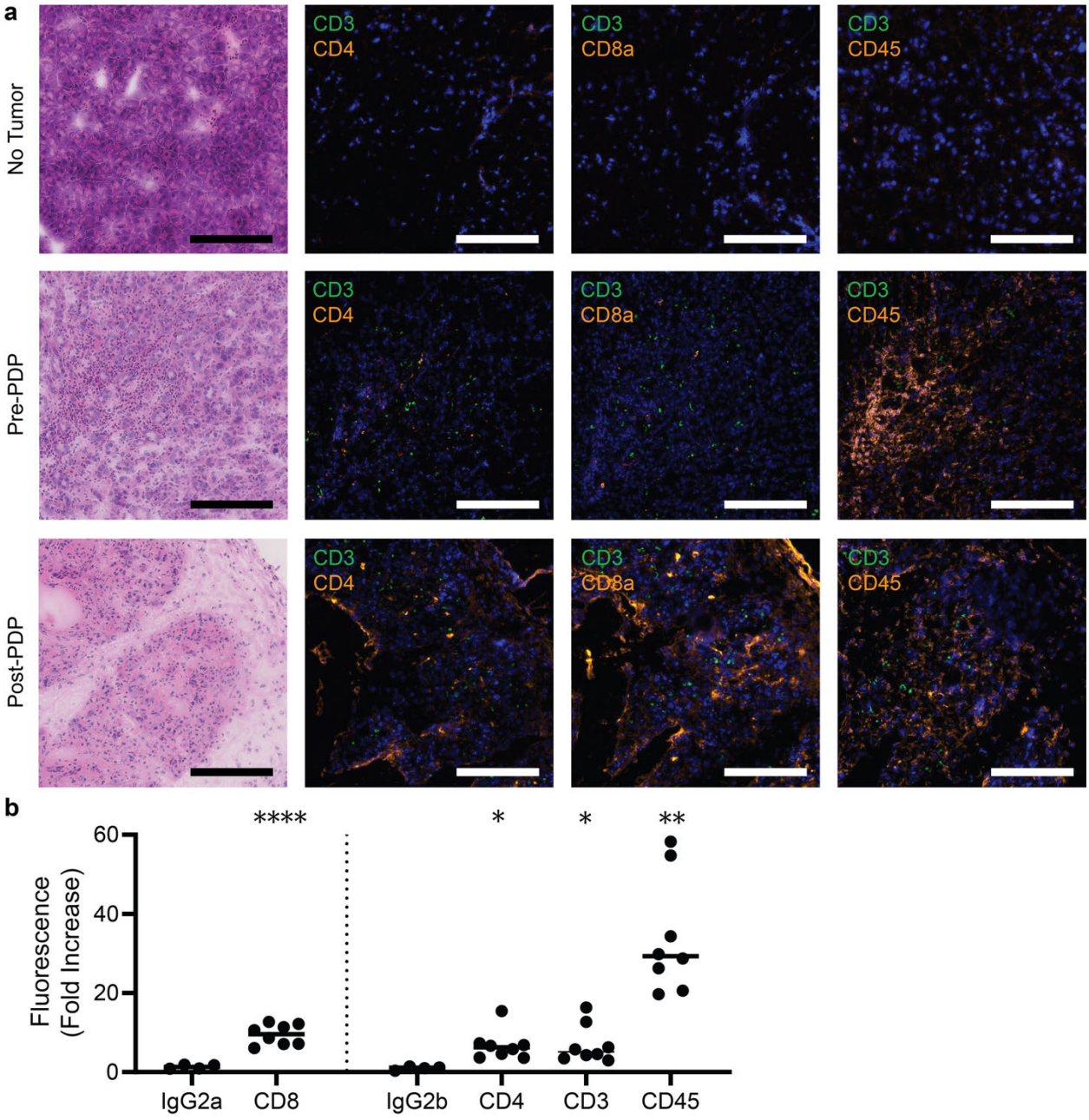

Fig. S18.
Validation of *in vivo* SMIRC imaging of the syngeneic KPC primary pancreatic ductal adenocarcinoma model by *ex vivo* histopathology and immunofluorescence staining. (**A**) *Top row:* H&E biopsy tissue section of the pancreas from a no-tumor control mouse; immunofluorescence staining for anti-CD3 and anti-CD4, and anti-CD3 and anti-CD8a, performed on serial, co-registered tissue sections adjacent to the H&E section. *Middle row:* H&E biopsy of the primary pancreatic tumor from a syngeneic KPC mouse prior to PDP with the same *ex vivo* immunostains above applied to serial, co-registered tissue sections. *Bottom row:* H&E biopsy of the primary pancreatic tumor from a syngeneic KPC mouse post-PDP with the same *ex vivo* immunostains above applied to serial, co-registered tissue sections. Immunofluorescence images include a DAPI nuclear stain (blue). Scale bars, 200 µm. (**B**) Quantification of the

cumulative fluorescence intensity of each of the target tumor biomarkers within the *ex vivo* KPC pancreatic tumor biopsies using anti-CD3, anti-CD4, anti-CD8, and anti-CD45 primary antibodies versus their respective non-specific control antibody isotype (IgG2a or IgG2b) using appropriate secondary antibody–fluorophore conjugates (*P<0.05, **P<0.01, ****P<0.0001, one-way ANOVA with Dunnett's *post hoc* test). Data points represent one or more tissue sections analyzed per individual mouse (n = 4–8 mice per group; 1–2 whole tumor cross-sections imaged per mouse, ~25 mm^2 in dimension per image; solid lines indicate mean values). A serial H&E section was used to locate the tumor boundary for each tissue section analyzed. These results demonstrate negligible numbers of T cells (CD3+, CD4+, or CD8+) in pancreatic tissue from no-tumor control mice, whereas abundant T cell infiltration is observed in KPC tumor biopsies, consistent with in vivo SMIRC measurements.

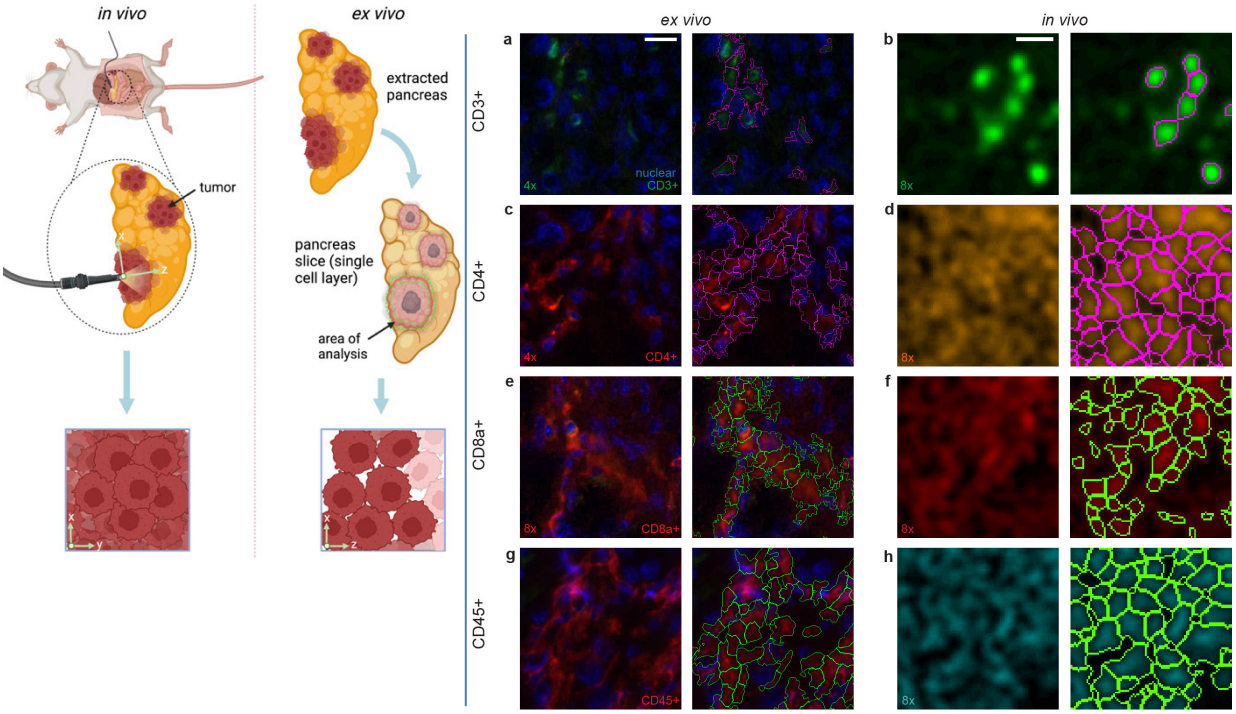

Fig. S19.
Representative images of cell identification for *in vivo* SMIRC images and for *ex vivo* immunofluorescence images of the syngeneic KPC primary pancreatic ductal adenocarcinoma model. *Left:* Schematic illustrating correspondence between *in vivo* SMIRC imaging of KPC pancreatic tumors and *ex vivo* tumor sections acquired in an orthogonal plane. The green dashed lines indicate the surface region of interest selected in ex vivo sections to best match the tumor surface regions imaged *in vivo*. *Right:* Representative images (brightened 1–4× as labeled) of cell identification *in vivo* and *ex vivo* for each biomarker. Images are taken near the tumor edge of untreated (pre-PDT) KPC pancreatic ductal adenocarcinoma primary tumors, either *in vivo* (x–z plane) or *ex vivo* (x–y plane). *In vivo* SMIRC images integrate signal over multiple cell layers (fig. S3F; axial resolution ~160 μm, further influenced by limited probe penetration and tissue scattering) whereas the *ex vivo* images of thin slices (5 μm) represent approximately a single-cell layer. For each example, the raw image (left) and the corresponding image with identified cells outlined (right) are shown: (**A**) CD3+ cells *ex vivo*; (**B**) CD3+ cells *in vivo*; (**C**) CD4+ cells *ex vivo*; (**D**) CD4+ cells *in vivo*; (**E**) CD8+ cells *ex vivo*; (**F**) CD8+ cells *in vivo*; (**G**) CD45+ cells *ex vivo*; and, (**H**) CD45+ cells *in vivo*. Scale bars 20 μm. To evaluate the sensitivity of SMIRC for cellular quantification, cell counts were computed for each biomarker (CD3, CD4, CD8, and CD45) in *in vivo* SMIRC images and compared to *ex vivo* immunofluorescence measurements in matched KPC tumor samples.

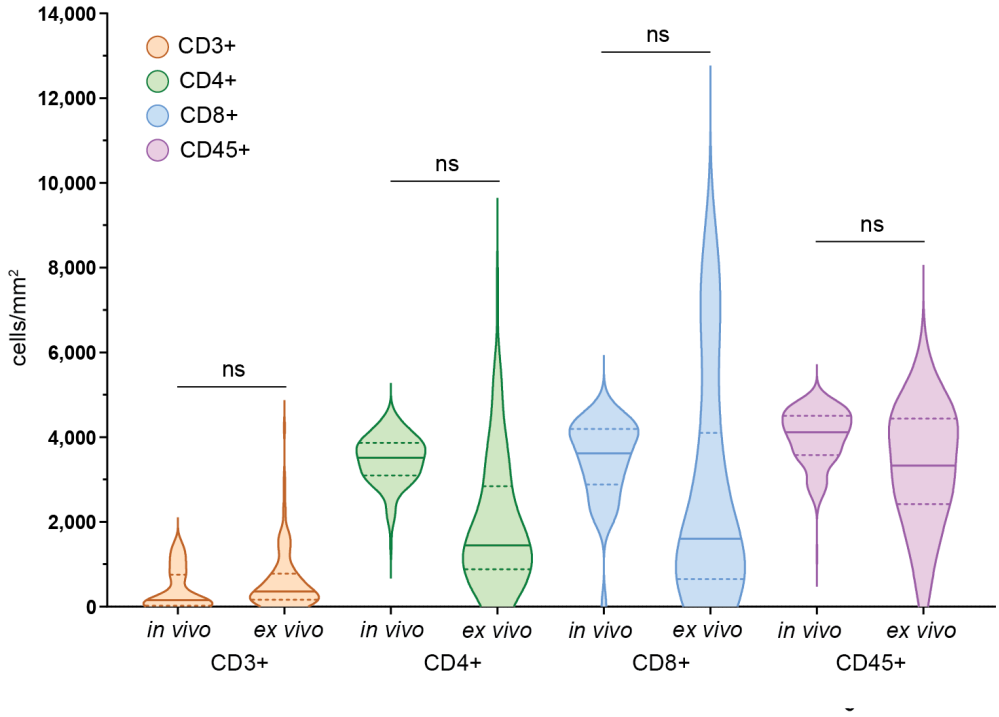

Fig. S20.
Validation of *in vivo* SMIRC imaging sensitivity by comparison of CD3, CD4, CD8, and CD45 cell counts measured *in vivo* versus *ex vivo* in the syngeneic KPC primary pancreatic ductal adenocarcinoma model. Violin plots of cell densities (cells/mm^2) are shown for each biomarker *in vivo* and *ex vivo* (n = 3 mice for each group; solid lines indicate medians; dotted lines indicate quartiles). Cell counts were analyzed near the tumor edge of untreated (pre-PDT) KPC pancreatic ductal adenocarcinoma primary tumors. All *in vivo* SMIRC images were included in the analysis. Tumors or cancerous cells were identified in *ex vivo* images and a slice around the tumor edge (~35 µm width) was broken up into tiled images for comparable analysis. Cell counting was done using a custom analysis pipeline in CellProfiler. Statistical analysis using nested one-way ANOVA showed no statistical differences (ns, P>0.05) between *in vivo* and *ex vivo* cell counts for any biomarker: CD3+ (P>0.999), CD4+ (P=0.5691), CD8+ (P=0.9298), CD45+ (P=0.7932). Additional comparisons revealed expected differences between biomarker populations (e.g., CD3+ *vs*. CD4+, CD8+, or CD45+), reflecting distinct immune cell subsets. Overall, *in vivo* SMIRC–derived cell counts agree closely with *ex vivo* measurements, with median values matching within approximately one quartile across all biomarkers, supporting the quantitative accuracy of *in vivo* SMIRC for cellular phenotyping.

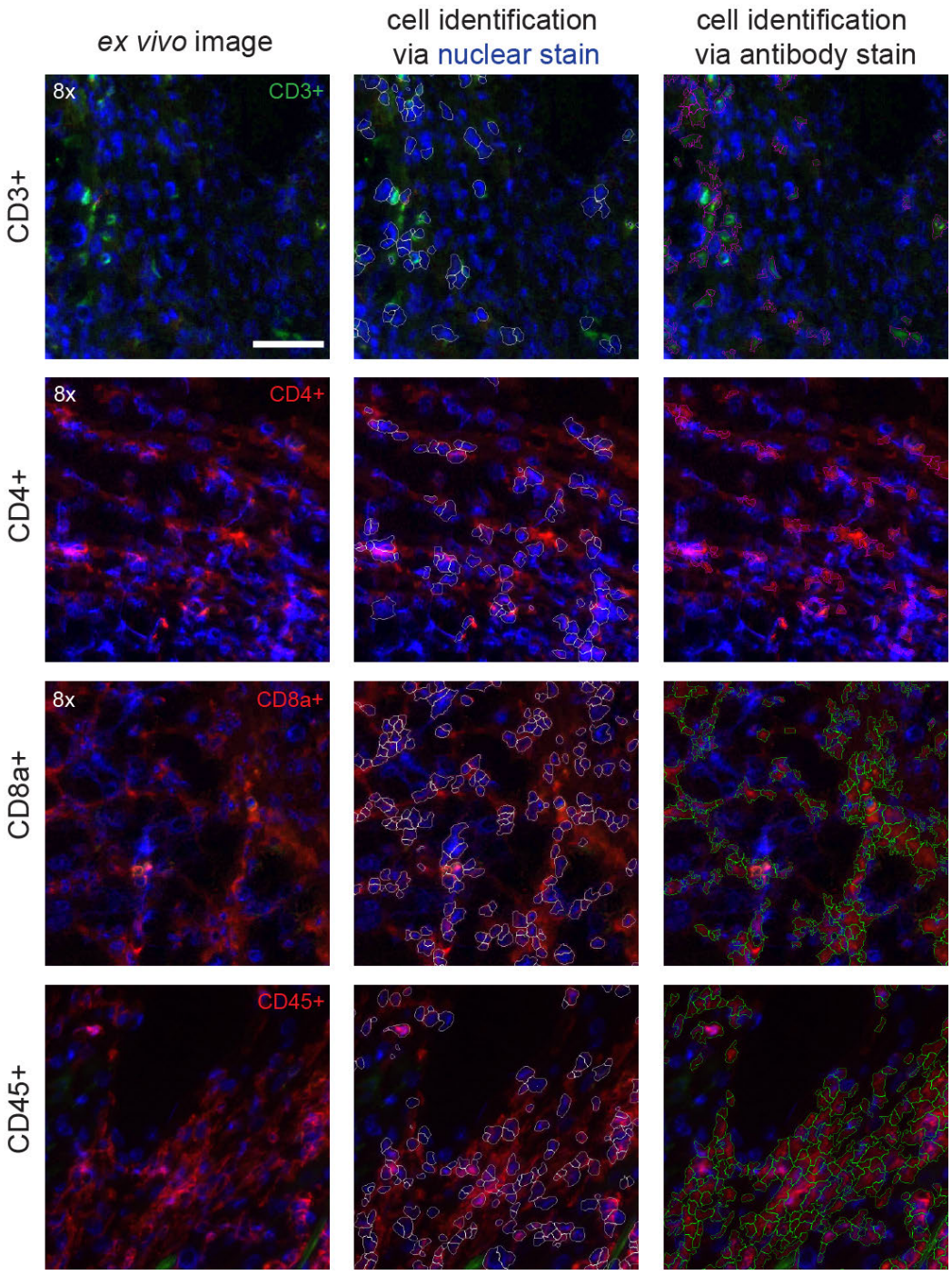

Fig. S21.

Comparison of automated CD3+, CD4+, CD8+, and CD45+ cell identification at the tumor periphery with and without nuclear staining to aid segmentation. *Ex vivo* immunofluorescence images of the syngeneic KPC primary pancreatic ductal adenocarcinoma model for each immune cell marker are shown (left column). Cells identified using a nuclear stain (DAPI) are outlined for each marker (middle column) and closely correspond to cells identified using antibody–fluorophore conjugates alone: anti-CD3 (brightened 4×), anti-CD4 (4×), anti-CD8a (4×), or anti-CD45) (right column). This comparison demonstrates that accurate cell identification can be

achieved without nuclear staining, consistent with prior studies showing that CellProfiler can reliably estimate cell counts from clustered immunofluorescence signals using intensity-based peak detection. This capability is critical for *in vivo* applications, where nuclear stains are generally not feasible due to the requirement for cell permeabilization. Scale bar, 50 μm.

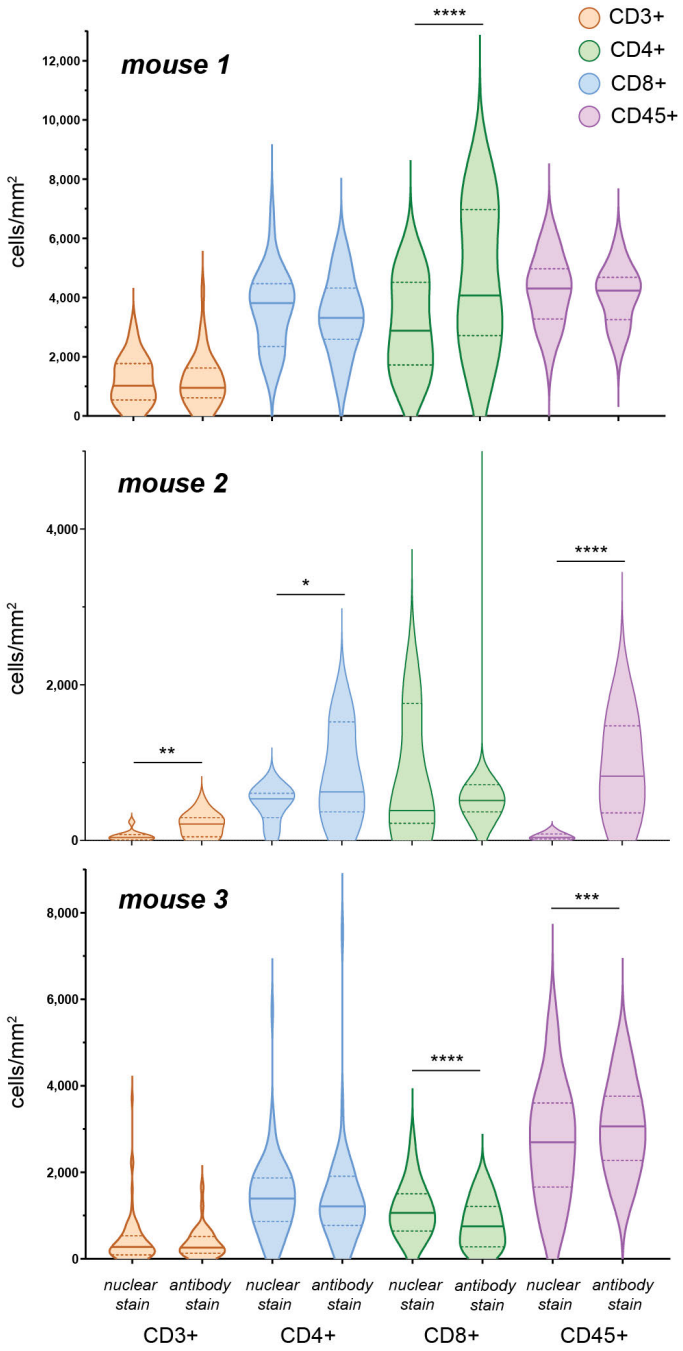

Fig. S22.
Mouse-by-mouse comparison of CD3+, CD4+, CD8+, and CD45+ cell counts from *ex vivo* immunofluorescence images of the syngeneic KPC pancreatic ductal adenocarcinoma model, with and without nuclear staining to aid segmentation. Violin plots of cell densities (cells/mm²) are shown for each biomarker on a per-mouse basis (solid lines indicate medians; dotted lines indicate quartiles). Cell counts were obtained near the tumor edge of untreated (pre-PDP) KPC pancreatic tumors. Tumors were identified in single-cell-layer *ex vivo* immunofluorescence images, and a region surrounding the tumor edge (~35 μm width) was subdivided into tiled regions for analysis. Cell counting was performed using two CellProfiler-based pipelines: (i) a nuclear-guided method, in which nuclei (DAPI) were identified and masked with corresponding

antibody-positive signal, and (ii) an antibody-only method, in which cells were identified directly from antibody–fluorophore signal using intensity-based segmentation. Comparison of tumor immune cell counts obtained with and without nuclear staining shows modest statistical differences without a consistent directional bias (*i.e.*, no systematic over- or under-counting by either method; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $**P < 0.0001$; paired Student's *t*-test). Overall, cell counts from both methods are in close agreement, and observed differences likely reflect variability in staining efficiency. Visual inspection of immunofluorescence images indicates incomplete or heterogeneous nuclear staining in some regions, which may contribute to discrepancies in nuclear-guided segmentation. These results support the reliability of antibody-based cell counting approaches and validate the accuracy of *in vivo* SMIRC-derived cell counts, for which nuclear staining is not feasible.

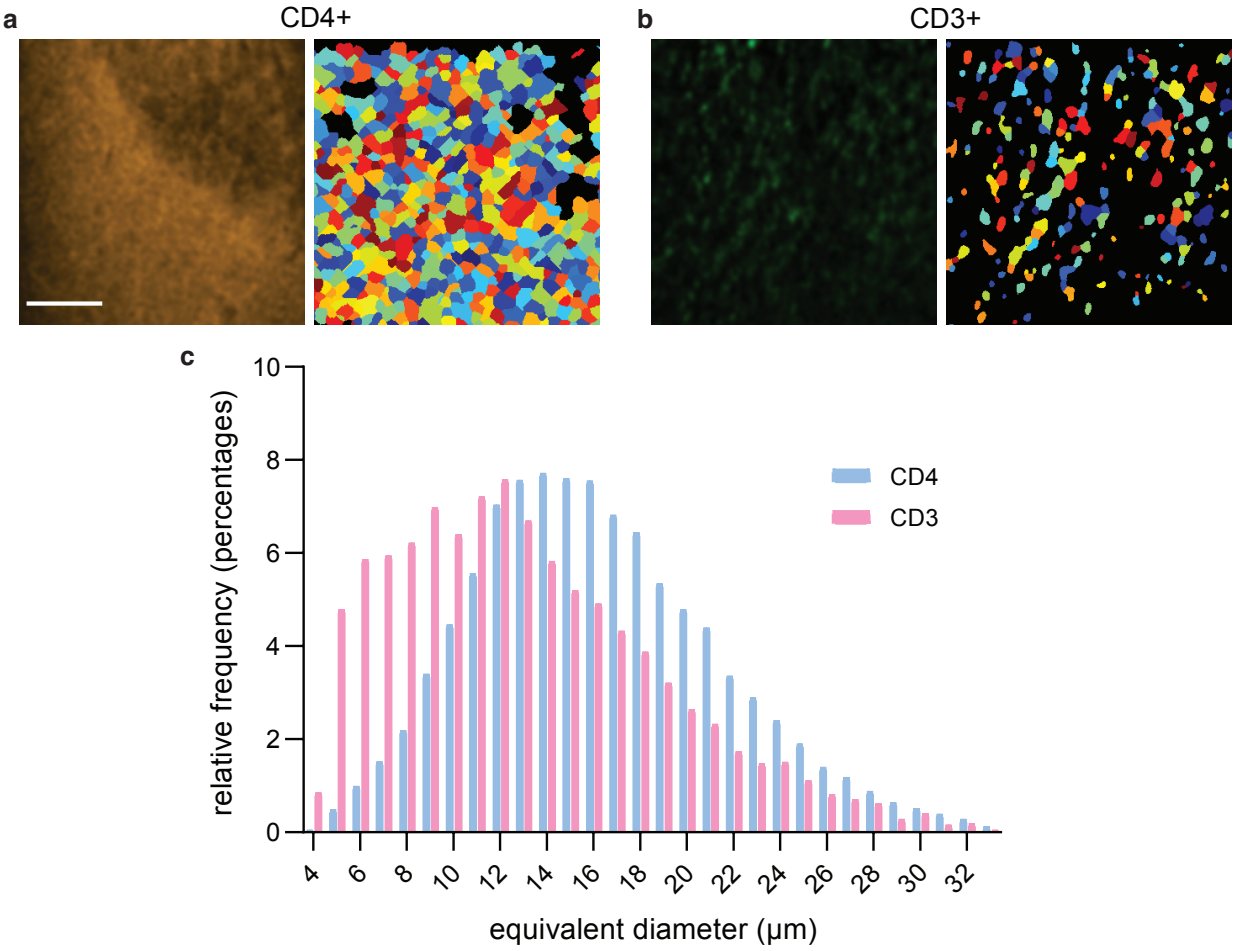

Fig. S23.
Quantification of cell equivalent diameter measurements for CD3+ and CD4+ cell counts from *in vivo* SMIRC images of the syngeneic KPC primary pancreatic ductal adenocarcinoma model. (**A**) *In vivo* SMIRC image of spectrally unmixed CD4 signal (left; brightened 2×) and corresponding identified cells (right; colors indicate detected cells). (**B**) *In vivo* SMIRC image of spectrally unmixed CD3 signal (left; brightened 2×) and corresponding identified cells (right; colors indicate detected cells). Scale bar, 100 µm. (**C**) Histogram of equivalent diameters for identified CD3+ and CD4+ cells from a representative mouse. CD3+ cells imaged by SMIRC exhibit a median equivalent diameter of 12.3 µm, whereas CD4+ cells exhibit a median equivalent diameter of 15.7 µm. The measured CD3+ diameters are consistent with activated T cells, which are larger than circulating T cells, while the larger CD4+ diameters are consistent with inclusion of CD4-expressing non-lymphocyte populations such as monocytes. Overall, the equivalent diameters of CD3+ and CD4+ cells identified *in vivo* by SMIRC are consistent with expected immune cell size ranges, supporting the biological validity of the segmentation and classification approach.

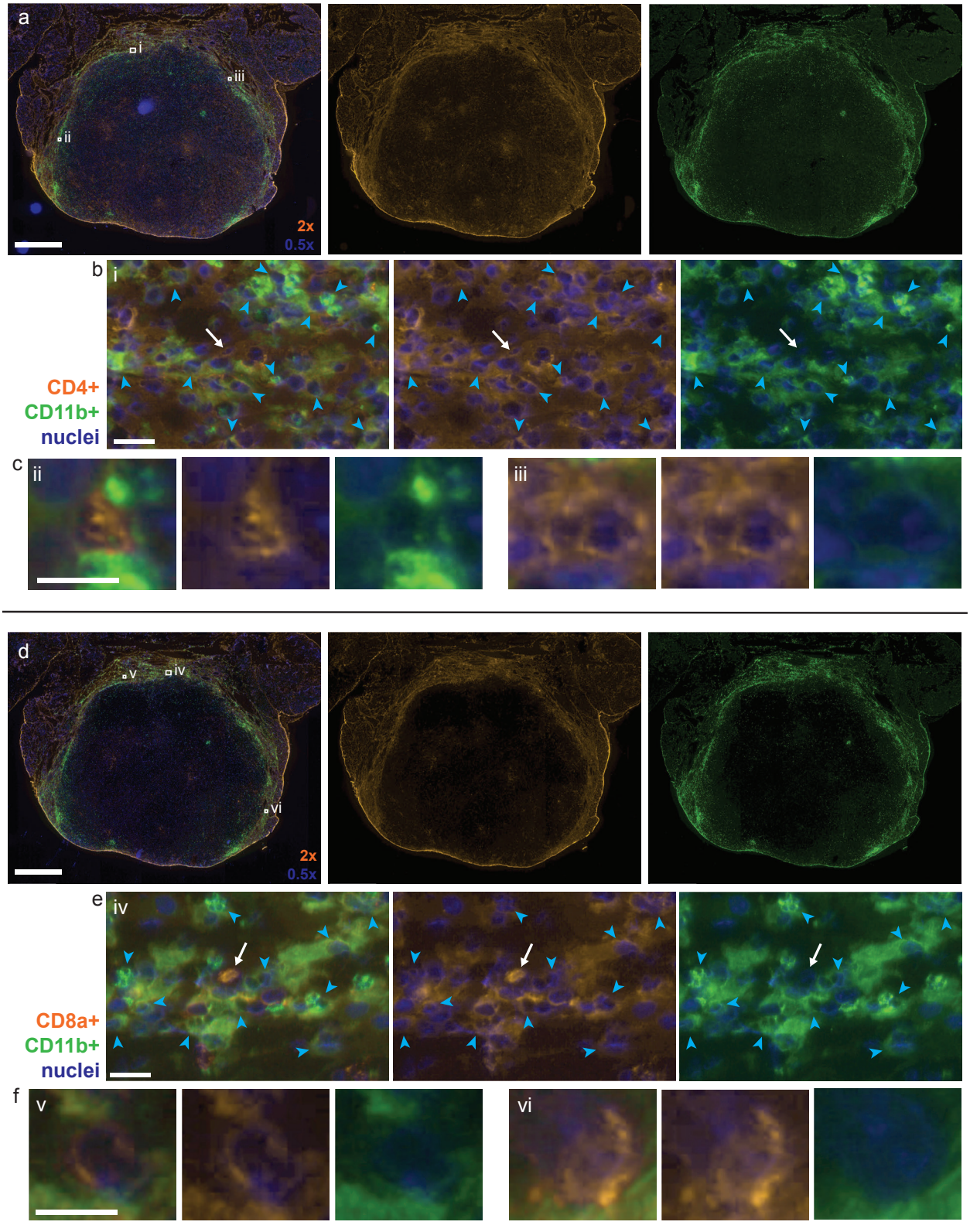

Fig. S24.
In the syngeneic FC1245 variant of the KPC primary pancreatic ductal adenocarcinoma model, CD4+ and CD8+ cell counts exceed CD3+ T cell counts due to an abundant population of

CD11b+ non-lymphocyte immune cells at the tumor periphery. *Ex vivo* immunofluorescence images of KPC tumor sections stained with either (**A–C**) anti-CD11b and anti-CD4 antibody–fluorophore conjugates or (**D–F**) anti-CD11b and anti-CD8a antibody–fluorophore conjugates. (A) Immunofluorescence staining for CD11b and CD4 demonstrates a high density of immune cells at the tumor periphery. Scale bar, 1 mm. (B) Zoomed-in region from (A) at the tumor edge. Most immune cells are CD11b+/CD4+ (consistent with monocyte/macrophage populations; examples indicated by blue arrowheads), whereas a smaller fraction are CD11b−/CD4+ T cells (white arrow). Scale bar, 25 μm. (C) Additional examples of CD11b−/CD4+ T cells. Scale bar, 25 μm. (D) Immunofluorescence staining for CD11b and CD8. Scale bar 1 mm. (E) Zoomed in region from (D) at the tumor edge. Most immune cells are CD11b+/CD8+ (consistent with non-lymphocyte populations such as myeloid or NK-lineage cells; blue arrowheads), whereas a smaller fraction are CD11b−/CD8+ T cells (white arrow). Scale bar, 25 μm. (F) Additional examples of CD11b−/CD8+ T cells. Scale bar, 25 μm. These data demonstrate that, in this KPC model variant, a substantial fraction of CD4+ and CD8+ cells are CD11b+ non-lymphocytes, explaining why CD4+ and CD8+ cell counts exceed CD3+ T cell counts. This contrasts with reports in other KPC models in which T cell markers more closely track CD3+ populations, and highlights model-dependent differences in tumor immune composition. Accordingly, these results emphasize the importance of multiplexed phenotyping and the use of CD3 to specifically identify T cells, as CD4 or CD8 expression alone is insufficient to distinguish lymphocyte populations.

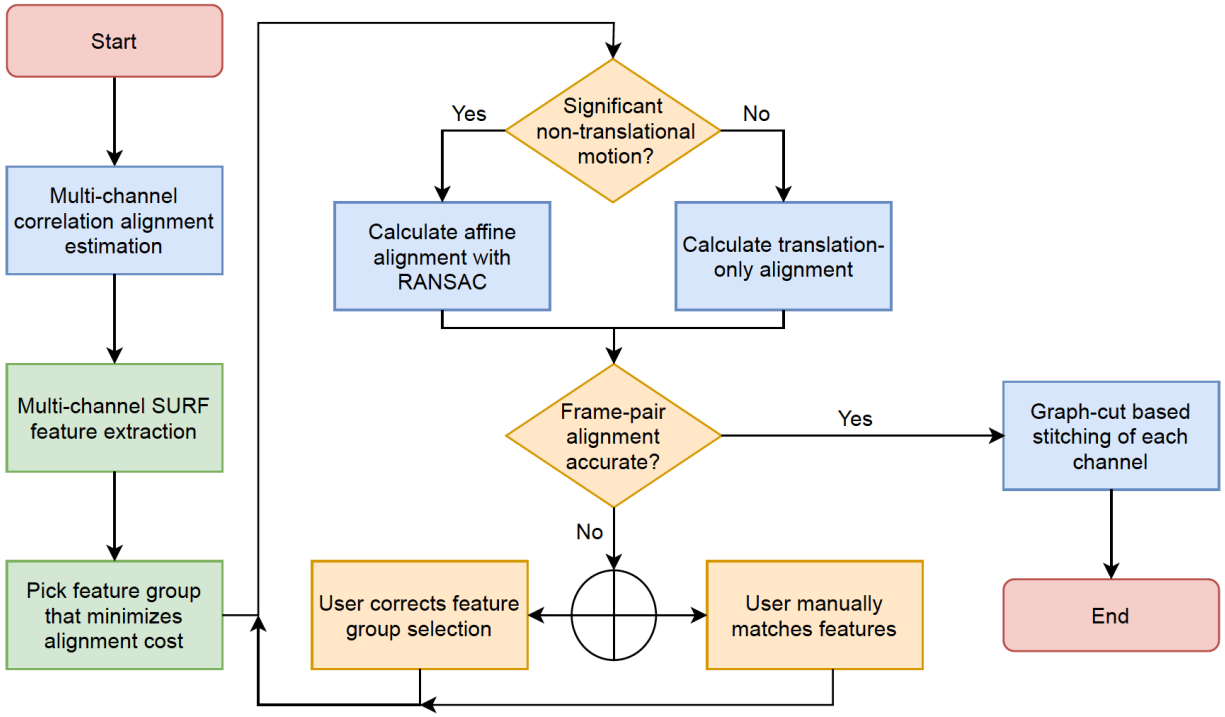

Fig. S25.
Overview of the mosaicking algorithm, which integrates previously published alignment methods (blue) with multichannel SURF-based feature analysis to improve registration accuracy. Alignment is performed using a cost-function–based framework that extends conventional approaches and enables consistent performance across different imaging conditions. User-guided decisions (yellow) are incorporated to identify cases of significant non-translational motion and to assess overall mosaic quality. When non-translational motion is detected, the algorithm adaptively introduces additional degrees of freedom (e.g., affine transformations) beyond simple translational correction to compensate for motion artifacts and improve alignment accuracy. The $\oplus$ symbol denotes a logical "or."

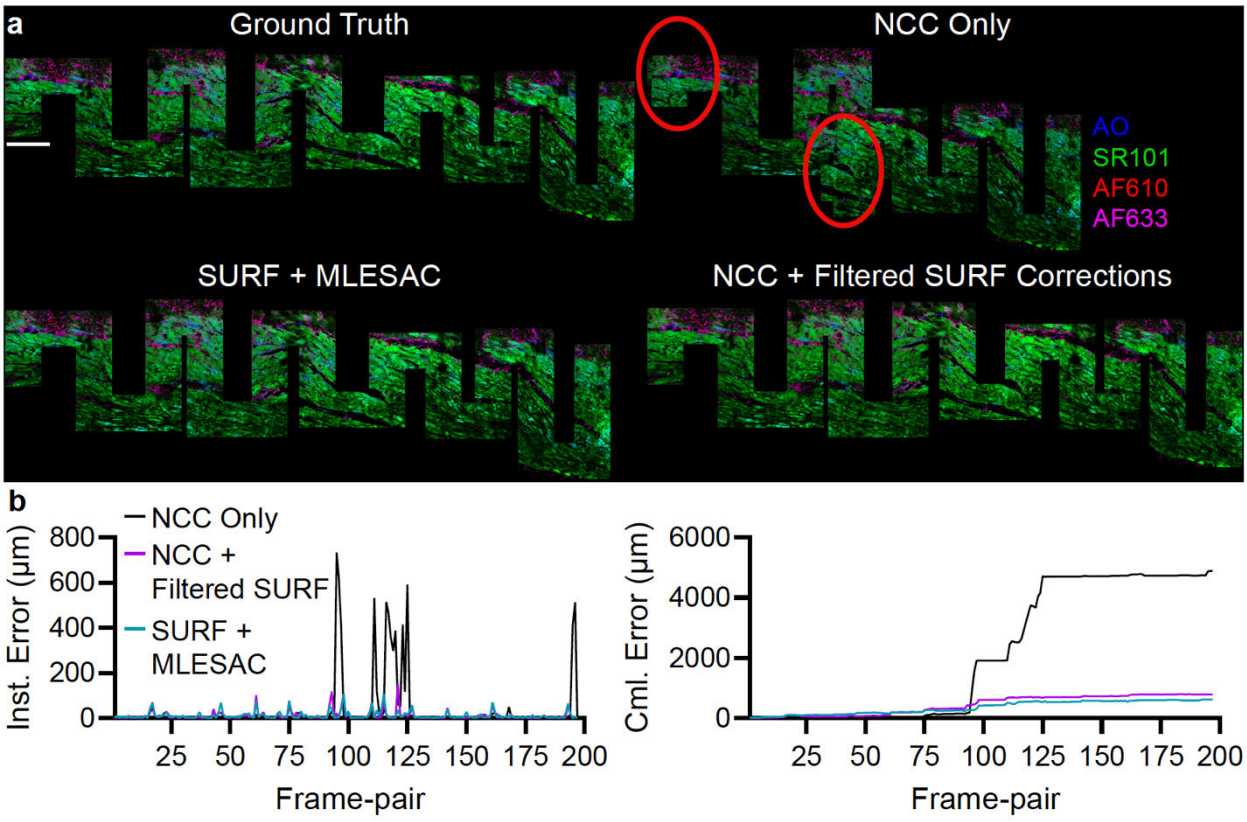

Fig. S26.
Illustration of improved alignment accuracy using the present mosaicking algorithm for a heterogeneous tissue phantom imaged by hyperspectral two-photon microscopy. (**A**) Comparison of mosaics generated from a time-series dataset acquired on a Zeiss LSM 880 confocal microscope (λ_{exc} = 890 nm; 32-channel detection, 400–700 nm) during manual raster scanning of a composite sample consisting of fluorescently labeled OVCAR5 cells deposited on "virtual H&E"–stained chicken breast tissue. Relative to a manually constructed ground-truth mosaic, the NCC-only configuration exhibits several critical alignment errors (red ovals), whereas application of filtered SURF-based corrections to the NCC initialization restores alignment accuracy. Standard SURF feature matching with MLESAC outlier rejection also performs well when restricted to translational transformations. (**B**) Instantaneous (inst., left) and cumulative (cml., right) alignment errors for each configuration relative to ground truth, calculated as the Euclidean distance between estimated and ground-truth frame positions. The heterogeneous background and sparse, high-contrast cellular features provide a challenging testbed for feature detection and alignment. Scale bar, 1 mm.

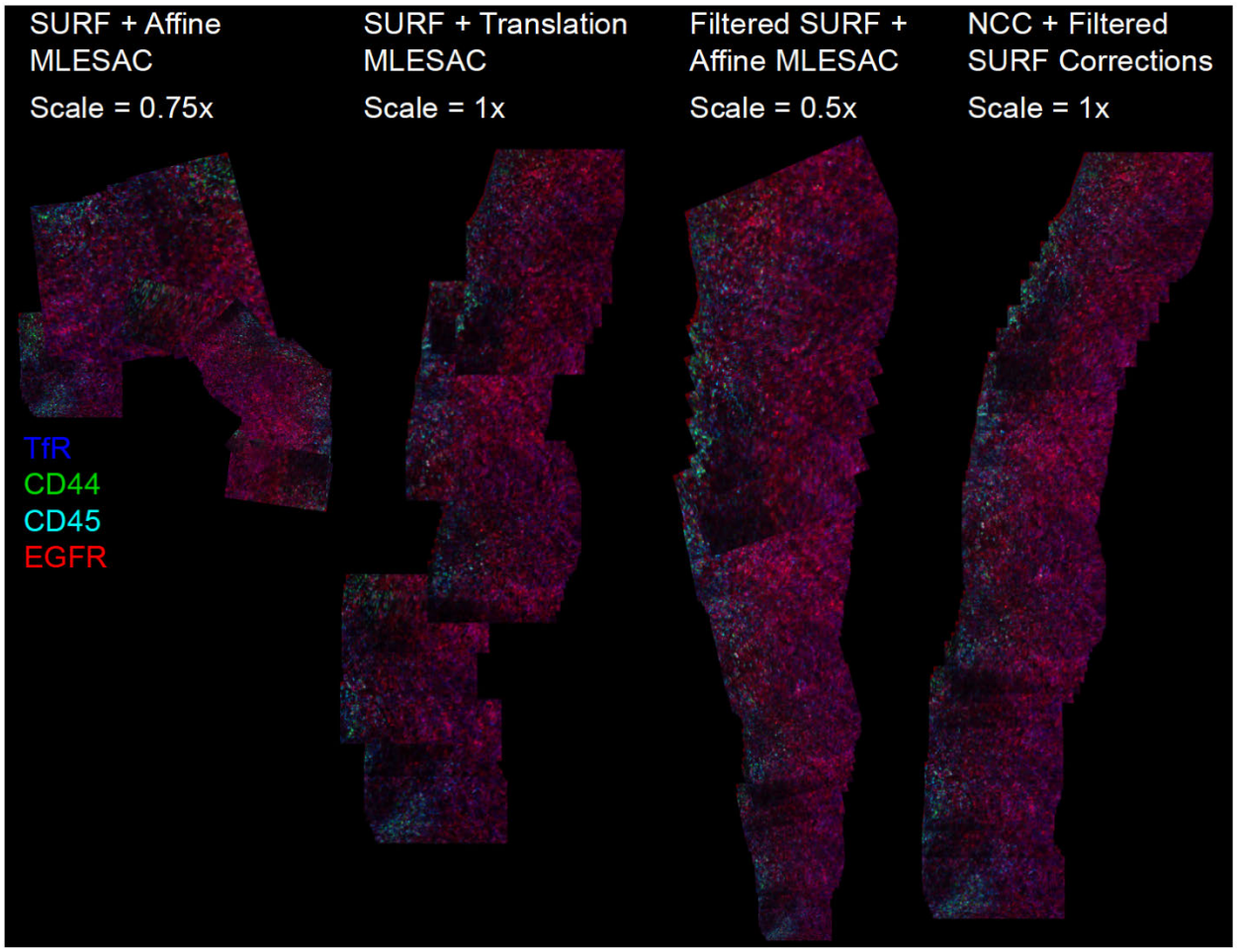

Fig. S27.
Alignment error across an example SMIRC imaging dataset from the OVCAR5 model of peritoneal carcinomatosis. Conventional SURF + MLESAC approaches fail to achieve robust alignment, even when homographies are restricted to translational motion. The filtered SURF + affine MLESAC scheme introduced here exhibits propagated scale error, leading to artificial enlargement of the reconstructed frame. In contrast, the NCC + filtered SURF correction approach yields the most accurate alignment, as additional degrees of freedom (*e.g.*, scale changes) are introduced only when required for correction, thereby preventing cumulative error propagation.

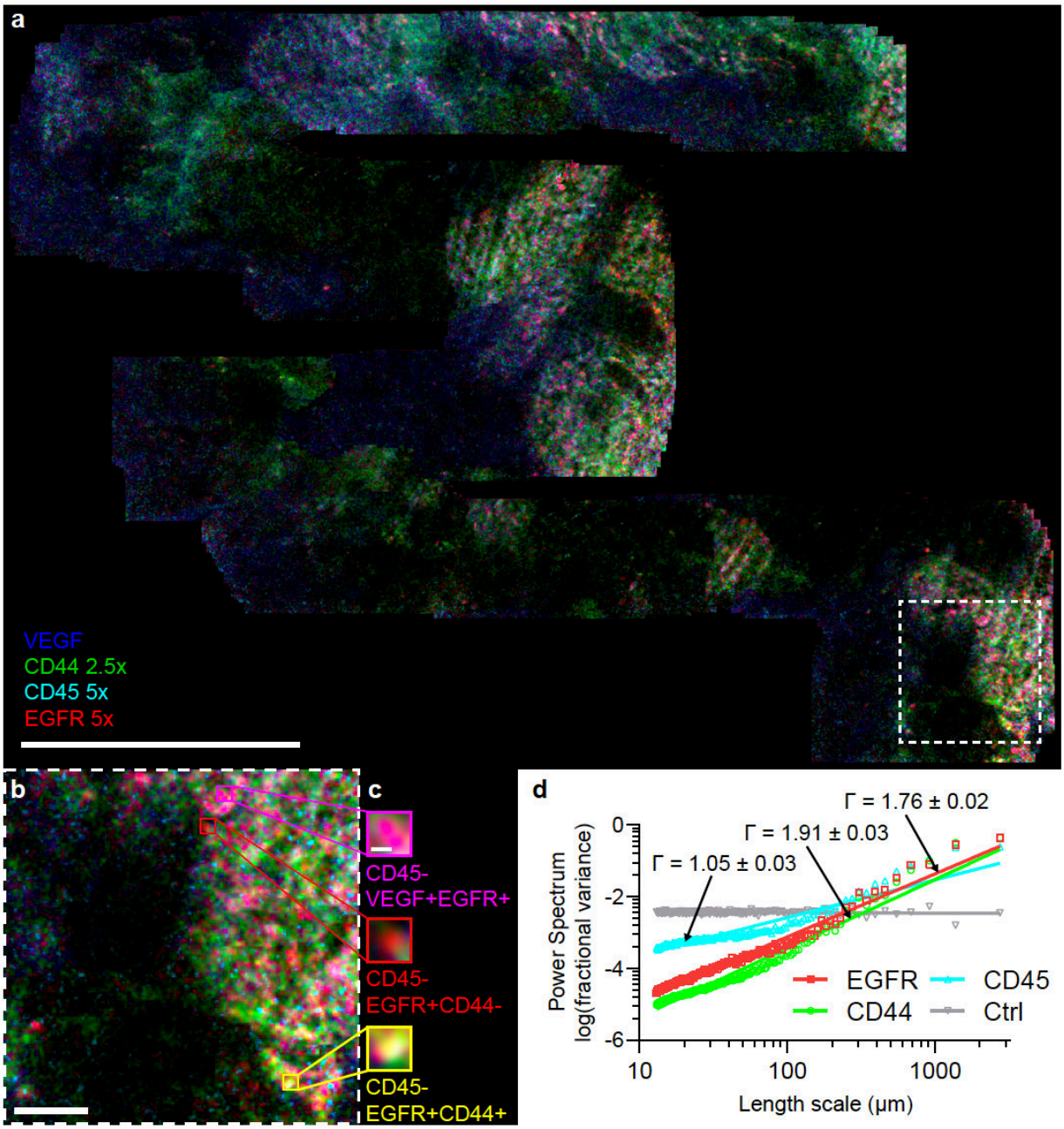

Fig. S28.
Micro-image mosaicking of SMIRC data in the OVCAR5 mouse model of peritoneal carcinomatosis reveals multiple CD45− cancer cell phenotypes and pronounced spatial biomarker heterogeneity during metastatic growth at day 16. (**A**) Four-biomarker mosaic of peritoneal micrometastases imaged in vivo by microendoscopy 16 days post–tumor inoculation. Scale bar, 1 mm. (**B**) Representative region of interest from (A) showing the boundary of a metastasis. Scale bar, 100 μm. (**C**) Higher-magnification view of (B) highlighting cellular features, where biomarker colocalization produces red, magenta, and yellow signals. Scale bar, 10 μm. (**D**) Power spectral analysis of the EGFR, CD44, and CD45 channels in (A), enabling quantification of spatial heterogeneity via the log–log slope (Γ) of a power-law model (solid lines).

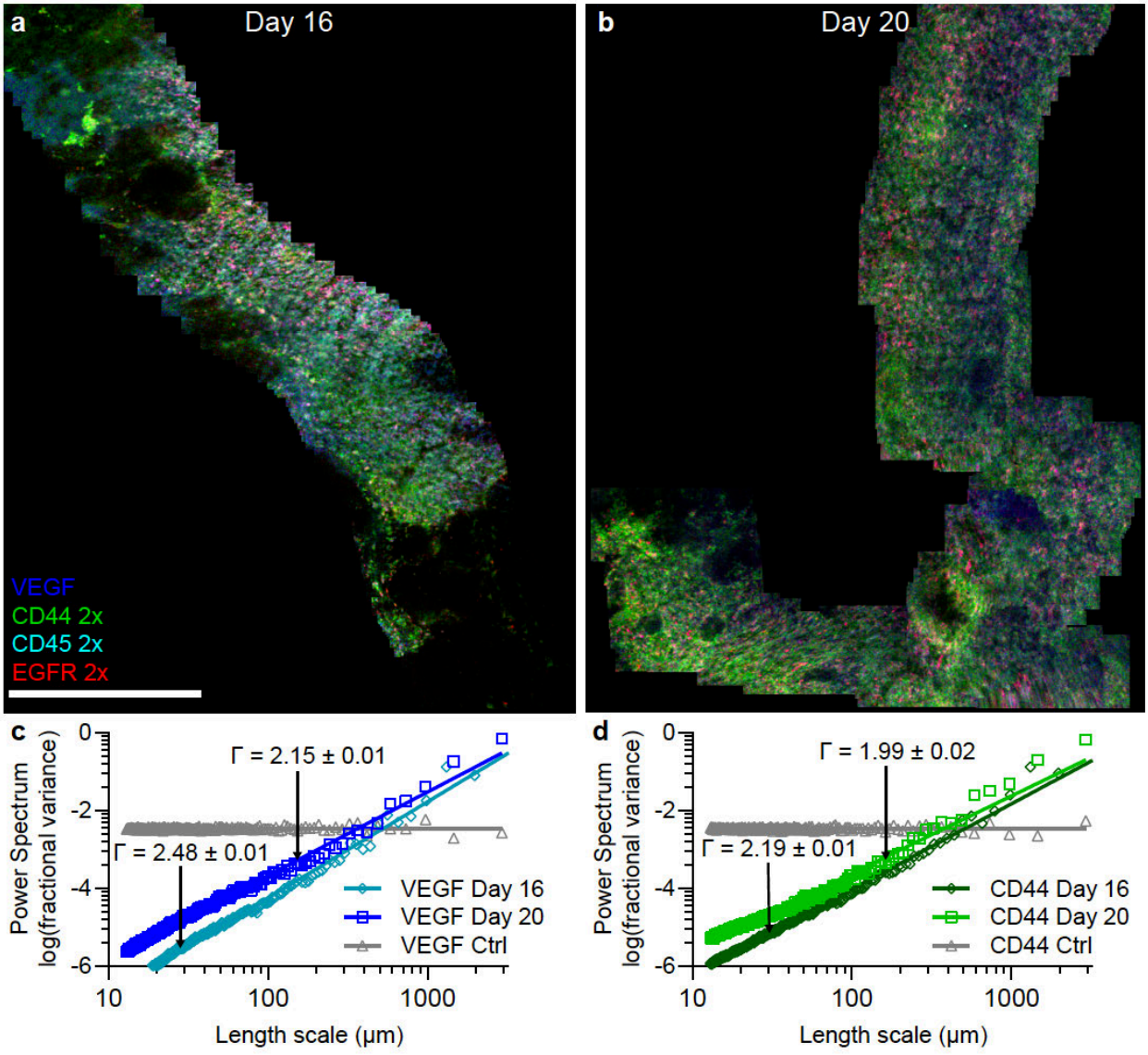

Fig. S29.
Longitudinal micro-image mosaicking of SMIRC data in the OVCAR5 model reveals progression from spatially heterogeneous micrometastases to more confluent late-stage disease. (**A**) Mosaic from microendoscopic imaging traversing the peritoneal wall at mid-stage disease, 16 days post–tumor inoculation. (**B**) Corresponding mosaic acquired at 20 days post–tumor inoculation, demonstrating more widespread, coalescent late-stage disease. Scale bar, 1 mm. (**C–D**) Power spectral analysis of the VEGF and CD44 channels in (A–B) shows that $\Gamma_{16\ \mathrm{day}} > \Gamma_{20\ \mathrm{day}}$ (P<0.0001, F-test on Γ), indicating greater spatial heterogeneity at earlier time points. This is consistent with smaller metastases exhibiting more frequent and pronounced spatial “hot” and “cold” regions of tumor growth.

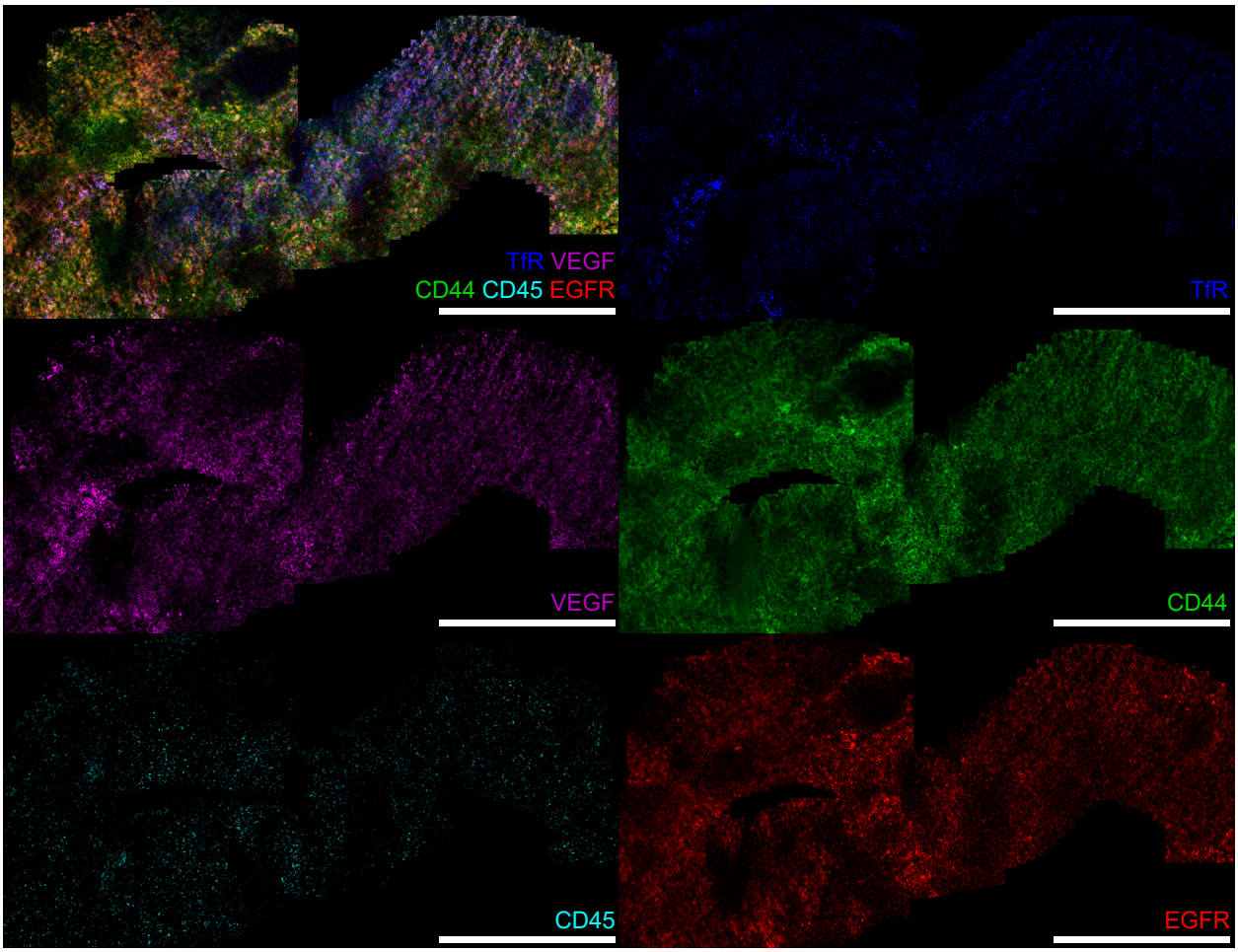

Fig. S30.
Micro-spectroscopic mosaic of individual biomarker channels in the xenograft peritoneal metastasis mouse model. Individual channels corresponding to Fig. 5A are shown separately for clarity. Scale bars, 1 mm.

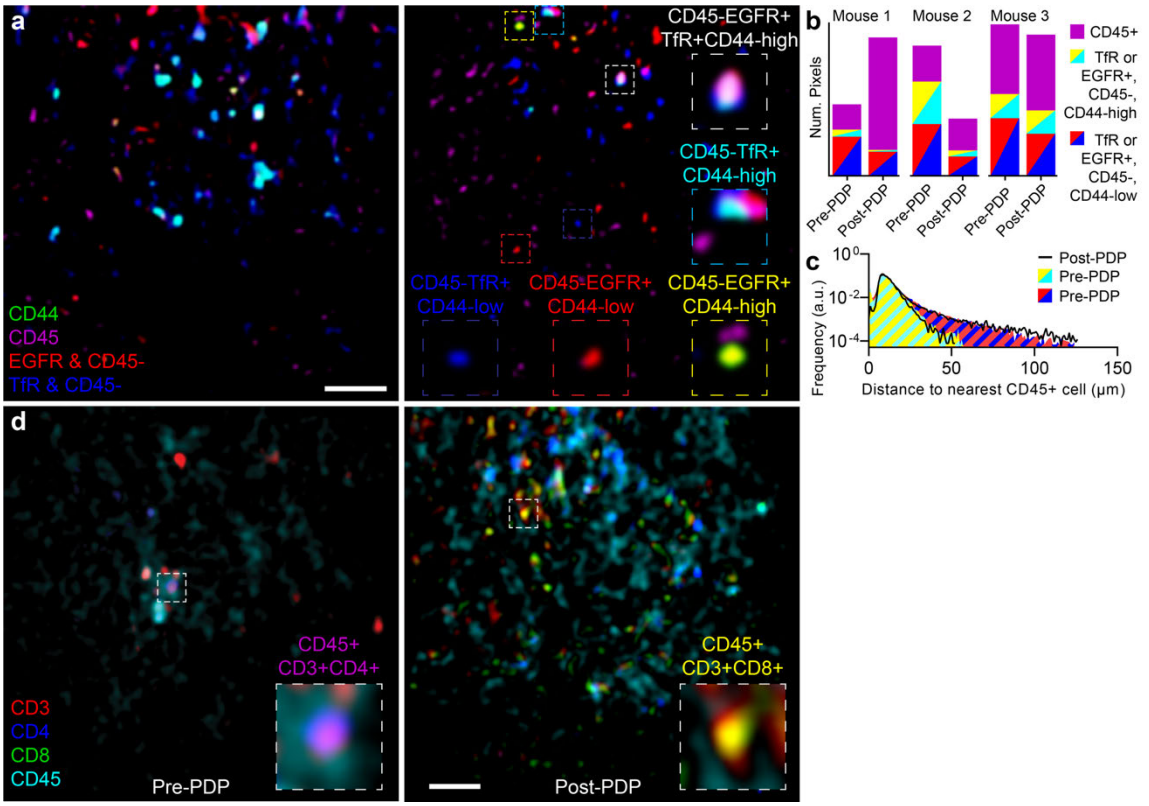

Fig. S31.
Tumor cell phenotype mapping and cancer–immune spatial analysis from in vivo SMIRC imaging. (**A**) Representative unmixed composite SMIRC images of untreated tumor sites on the peritoneal wall in a xenograft peritoneal metastasis (OVCAR5) model. Tumor biomarkers are pseudocolored following spectral unmixing. Scale bar, 50 μm. Dashed box insets (22 × 22 μm) highlight apparent cancer cell subtypes defined by biomarker co-expression. For visualization, unmixed basis images (EGFR, TfR, CD44, CD45) were thresholded (2× Otsu) to segment sparse cellular objects, and pixels overlapping CD45-positive signal were excluded. Images were further restricted to 45–55% of maximum intensity to emphasize bright cellular features. Insets were interpolated (3 × 3 grid, bicubic) for visualization. (**B**) Integrated peritoneal metastasis cell phenotype signal before and after photodynamic priming (PDP) for individual mice (n = 3 mice; n = 17 images per mouse per time point). Cell detection was automated using 8-connected component analysis on Canny edge maps of subtype images. (**C**) Cancer–immune spatial distance distributions show closer association of immune cells with CD44-high ($CD45^{-}$; $EGFR^{+}$ or TfR^{+}) cancer cells compared to CD44-low counterparts, independent of PDP ($P < 0.0001$, Mann–Whitney U test). Distances were computed as the Euclidean distance from each cancer cell object to the nearest immune cell. (**D**) Representative unmixed composite SMIRC images of an immunocompetent KPC pancreatic ductal adenocarcinoma tumor before and 4 days after PDP. Pre-PDP images show low baseline tumor-infiltrating T cell counts. Insets (26 × 26 μm) highlight a $CD45^{+}/CD3^{+}/CD4^{+}$ T cell (magenta). Post-PDP images demonstrate increased $CD45^{+}/CD3^{+}/CD8^{+}$ T cell infiltration, with inset showing a $CD45^{+}/CD3^{+}/CD8^{+}$ T cell (yellow). Scale bar, 50 μm. For visualization, CD4 and CD8 channels were thresholded at 2× Otsu, while CD3 and CD45 thresholds were set to 5% and 2.5% of maximum intensity, respectively. Images were further restricted to 33–72% of maximum intensity. Insets were interpolated (5 × 5 grid, bicubic).

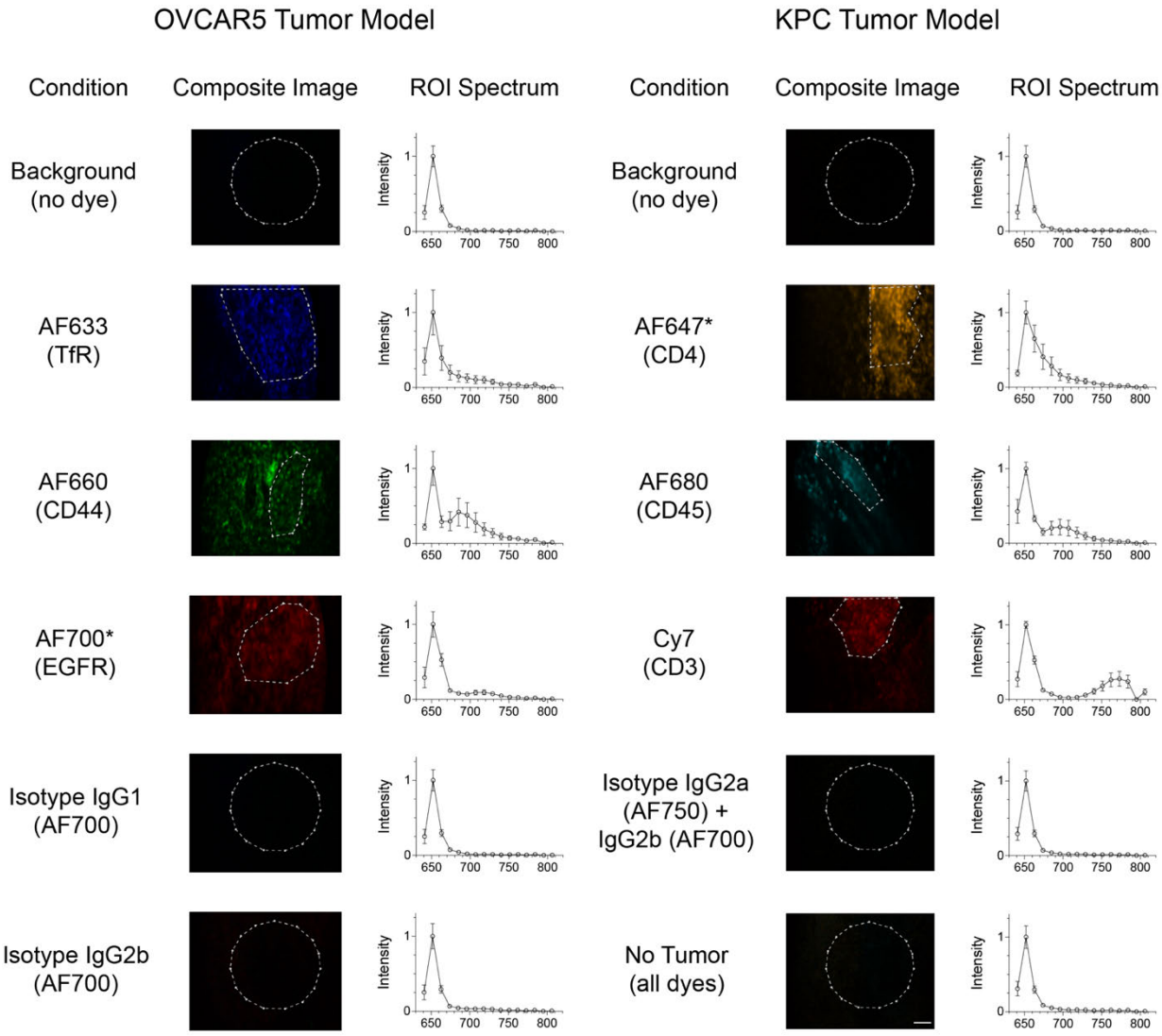

Fig. S32.

In vivo measurement and processing of fluorophore basis spectra. OVCAR5 and KPC tumor–bearing mice were first imaged without exogenous contrast agents to measure background signal, including laser reflection and tissue autofluorescence (notably in channels 1–2). Mice were then injected with individual antibody–fluorophore conjugates and imaged using SMIRC. Bright raw fluorescence signals were used to determine the spectral shape of each fluorophore. Background contributions were removed by unmixing the measured background signal, and regions of interest (ROIs) were manually selected around isolated fluorescent features. Mean spectra from each ROI are shown with standard deviation to indicate variability. Representative extracted basis spectra are shown in Fig. 1(B). For spectral unmixing, basis spectra were modified by setting channels 1–2 to zero, allowing the algorithm to account for spatial and temporal variations in background signal arising from fluctuations in laser power and tissue properties. This approach avoids the need for frame-by-frame background subtraction and improves robustness to inter- and intra-image variability. Distinct spectral features across channels 3–16 were sufficient for reliable fluorophore identification. Control experiments, including isotype-control conjugates in tumor-bearing mice and full probe cocktails in healthy mice, demonstrated negligible non-specific fluorescence.

Movie S1.

Online SMIRC video and micro-spectroscopic image mosaicking. Five tumor biomarker live animal video microendoscopy of a peritoneal carcinomatosis mouse model was performed at a SMIRC image cube acquisition rate of 17 Hz and is shown here at the same display rate. The right panel displays the SMIRC field-of-view (scale bar, 100 μm). The left panel displays the widefield mosaic formation as images are acquired (scale bar, 500 μm). The biomarker targets imaged are: TfR (dark blue), VEGF (magenta), CD44 (light green), CD45 (cyan) and EGFR (red).