Supplementary Information for:

Assessing micrometastases as a target for nanoparticles using 3D microscopy and machine learning

Benjamin R. Kingston[1,2], Abdullah Muhammad Syed[1,2], Jessica Ngai[1,2,3], Shrey Sindhwani[1,2], and Warren C.W. Chan[1,2,3,4,5*]

[1] Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada
[2] Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada
[3] Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario M5S 3E5, Canada
[4] Department of Material Science and Engineering, University of Toronto, Toronto, Ontario M5S 1A1, Canada
[5] Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6

*contact: warren.chan@utoronto.ca

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supplemental Videos S1 to S5
Supplemental Dataset Excel File 1

www.pnas.org/cgi/doi/10.1073/pnas.1907646116

MATERIALS AND REAGENTS

4T1 cells were generously donated from Dr. Reginald Gorczynski at the University of Toronto. Fetal bovine serum (FBS), Roswell Park Memorial Institute (RPMI 1640) media and trypsin (0.25% w/v) were purchased from Gibco, ThermoFischer Scientific (Burlington, Canada). Penicillin-streptomycin solution was from FroggaBio Inc. (Toronto, Canada). 6-8 week old female Balb/c mice were purchased from Charles River (Canada). Phosphate buffered saline (PBS), and Triton X-100 were purchased from Bio Basic Canada Inc. (Markham, Canada). Gold (III) chloride trihydrate ($HAuCl_4$ $3H_2O$ – 99.9% trace metal basis), sodium citrate tribasic dihydrate (99.0% ACS grade), hydroquinone (99.5% ReagentPlus), sodium nitrite, sodium azide, 2'2-thiodiethanol (TDE), sodium dodecyl sulfate (SDS), bovine serum albumin (BSA), normal goat serum, and Tween-20 were purchased from Sigma Aldrich (Oakville, Canada). Methoxy poly(ethylene glycol) thiol 5 kDa (5KDamPEG-SH) was purchased from Laysan Bio Inc. (Arab, USA). Cy3-NHS was purchased from Click Chemistry tools (Scottsdale, USA). *Griffonia simplicifolia* Lectin I (GSL-1) was purchased from Vector Laboratories (Burlingame, USA). Amicon 30kDa filters are from Millipore. Heparin, acrylamide, boric acid, 4-(2-hydrooxyethyl)-1-piperazineethanesulfonic acid (HEPES) and formaldehyde are from Bioshop Canada Inc (Burlington, Canada). VA-044 initiator is from Wako Chemicals (Richmond, USA). DAPI is from Life Technologies (Burlington, Canada). Nap-5 columns are from GE Healthcare (USA). Ki67 antibody (clone 11F6) and secondary goat anti-rat IgG Alexa-Flour 647 antibody were purchased from BioLegend (San Diego, USA).

METHODS

Spontaneous metastatic breast cancer mouse model

Tumours were induced by injecting 1 million 4T1 (mammary carcinoma cells – donated generously from Dr. Reginald Gorczynski at the University of Toronto) in 100 µL of serum free RPMI media into the right inguinal mammary fat pad of 8-week-old female Balb/c mice (Charles River). Tumours were grown for 3 weeks to allow time for spontaneous micrometastases to form in the liver and lungs. 2 x 10^{12} PEGylated 50nm gold nanoparticles were injected via the tail vein in 150 µL of sterile PBS and allowed to circulate for 1, 4 or 24 hours. Prior to sacrifice and transcardial perfusion fixation (as described below) 150 µg of Cy3-GSL-1 was injected via tail vein to label the blood vessels of the animal. All animal work was done in accordance with and approved by the University of Toronto Animal Care Committee.

Nanoparticle synthesis and characterization

Gold nanoparticles were synthesized as described in previous methods (1). 50 nm gold nanoparticles were made using 15 nm gold nanoparticle seeds. To synthesize the 15 nm seeds 100 mL of 0.25M chloroauric acid (Sigma-Aldrich) was brought to a boil with stirring and then 1 mL of 3.3 % w/v sodium citrate dibasic trihydrate (Sigma-Aldrich) was added and allowed to stir for 8 min before cooling on ice. Size and concentration are confirmed using dynamic light scattering (DLS) and UV-Visible absorbance spectroscopy, respectively. To synthesize 50 nm gold particles 967 µL of 25 mM chloroauric acid, 93.75 mL of deionized water, 967 µL of 15 mM sodium citrate tribasc dihydrate and 3.35 mL of

2.4 nM 15 nm gold particle seeds were are added to an Erlenmeyer flask with stirring. 967 µL of 25 mM hydroquinone (Sigma-Aldrich) was then rapidly added and left to stir overnight at room temperature. The next day, Tween-20 (Sigma-Aldrich) is added to a final concentration of 0.05% v/v. 50 nm gold nanoparticles are washed 3 x with 0.05% v/v Tween-20 in 50 mL falcon tubes (1800 x g for 2 hours). Following the final wash 50 nm gold particles were stored at 4°C until PEGylation. A small fraction was used to characterize size by transmission election microscopy, and DLS. Concentration was measured using UV-visible absorbance spectroscopy.

PEGylated 50 nm gold nanoparticle preparation

PEGylation of gold nanoparticle was done as described previously (1). Briefly, a solution of 5000 Da methoxy-PEG-thoiol (Laysan Bio) at 10 mg/mL at a surface density of 5 PEG molecules per nm^2 surface area was mixed with the 50 nm gold nanoparticles. This reaction was incubated at 60°C for 2 hours. Excess PEG is then removed by washing the particles 3 x with sterile PBS by centrifugation (2300 x g for 45 min). The PEGylated 50 nm gold nanoparticles are then resuspended to a concentration of 1.33 x 10^{13} nanoparticles/mL in sterile PBS.

Cy3-GSL-1 Synthesis

Cy3-NHS (Click Chemistry Tools) was conjugated to *Griffonia simplicifolia* Lectin I (GSL-1) (Vector Labs) via standard NHS-amine chemistry as described previously (2). Briefly, 100 µL of 10 mg/mL GSL-1 was reacted with 100 ug of Cy3-NHS in 100mM pH 8.5 HEPES buffer overnight at room temperature. The Cy3-GSL-1 is then purified from unreacted Cy3 dye by NAP-5 column (GE Healthcare) followed by washing 3 x with sterile PBS in an Amicon 30kDa centrifugation filter (Millipore). Dye labelling is measured by UV-visible absorbance at 553 nm and 280 nm and determined to be 4-6 dye molecules per GSL-1 molecule. The final concentration is adjusted to 1 mg/mL and 150 µg is injected into mice 5 min prior to transcardial fixation perfusion.

Transcardial perfusion fixation

Transcardial perfusion was carried out as described previously (2). After mice are under deep sedation with 3% isofluorane in oxygen, the chest cavity is opened and heart exposed. 60 mL of 10 U/mL Heparin (Bioshop Canada Inc.), 0.5% w/v sodium nitrite (Sigma-Aldrich) in PBS is first perfused to remove blood. Next, 60mL of hydrogel monomer solution is perfused through the animal containing 4% formaldehyde (Bioshop Canada Inc.), 2% w/v acrylamide (Bioshop Canada Inc), and 0.25% w/v VA-044 axoinitiator (Wako Chemicals) in PBS. Tissues are then collected and incubated in the same hydrogel monomer solution for 7 days at 4°C.

Tissue crosslinking

Tissues were polymerized as described previously (2). Fresh hydrogel monomer solution was added to samples, and samples were degassed and filled with argon. Samples were then incubated at 37°C for 3 hours. Excess hydrogel monomer solution is discarded and samples are washed 2 x with borate solution (200mM sodium borate, pH 8.5 0.1% Triton-X 100 and 0.01% sodium azide) and stored in borate solution at 4°C.

Tissue clearing
Tissues were cleared as described previously (2). For 1 mm slices tissues were incubated in 4% w/v SDS (Sigma-Aldrich) in 200 mM sodium borate pH 8.5 for 10 days at 50°C. Larger slices or whole organs were incubated at 50°C for up to 3 weeks to ensure complete transparency.

Nuclear staining with DAPI
Nuclear staining was done as described previously (2). Briefly, DAPI (Life Technologies) at 10 μM in PBS, 0.1% Triton-X 100, 0.01% sodium azide solution was incubated with tissues at room temperature for 3 days. Then, excess DAPI is washed away by washing tissues with 10 mL of the PBS, Triton X-100, sodium azide solution at room temperature for 24 hours at room temperature, exchanging the buffer 3 times over the 24 hour period. When DAPI staining is done in conjunction with antibody staining it is added at the above concentration (10 μM) along with the secondary antibody.

Immunostaining cleared tissues
Ki67 staining was done by first blocking tissues with 5% w/v bovine serum albumin (BSA) in PBS, 0.1% Triton-X 100, 0.01% sodium azide solution overnight at room temperature with gentle agitation. Next, primary Ki67 antibody (clone 11F6, BioLegend) was added to tissues in 1 mL of 2% v/v goat serum, PBS, 0.1% Triton-X 100, 0.01% sodium azide at a 1:100 dilution and incubated at room temperature with gentle agitation for 3 days. Next, excess primary antibody is washed away by transferring the tissue into 10 mL of PBS, 0.1% Triton-X 100, 0.01% sodium azide solution for 24 hours at room temperature exchanging the washing buffer 3 times over the 24 hour period. Next, an anti-rat Alexa Fluor-647 secondary antibody (BioLegend) is added at a 1:200 dilution in 1 mL of 2% v/v goat serum, PBS, 0.1% Triton-X 100, 0.01% sodium azide and incubated with gentle agitation for 3 days at room temperature. Excess secondary antibody is washed away by transferring the tissue into 10 mL of PBS, 0.1% Triton-X 100, 0.01% sodium azide solution for 24 hours at room temperature exchanging the washing buffer 3 times over the 24 hour period. The tissue is then transferred to 67% 2'2-thiodiethanol (TDE) solution in 200 mM borate, 0.1% Triton-X, 0.01% sodium azide solution for refractive index matching for at least 24 hours at room temperature prior to lightsheet imaging.

3D Light-sheet imaging
Imaging was done with the Zeiss Lighsheet Z.1 microscope (CLARITY Plan-Neofluor Objective: 20X, NA = 1.0, refractive index = 1.45). Tissues are immersed in a 67% TDE solution as described above. Dark-field imaging of 50 nm gold nanoparticles was done as previously described using the 638 nm excitation laser at a power of 0.25% with the laser block and emission filters removed (2). Fluorescent imaging is done as using standard excitation lasers and emission filters for DAPI (excitation: 405 nm, emission filter: BP 420 nm – 470 nm), Cy3-GSL-1 (excitation: 561 nm emission filter: BP 575 nm – 615 nm) (Cy3 labelled *Griffonia simplicifolia* Lectin I) and Ki67 labelled with Alexa-Flour 647 secondary (excitation: 640 nm emission: LP 650 nm) using laser powers between 2% and 30%. For computational analysis a magnification of 0.7 using the 20X objective was used with a z-slice size of 1 μm keeping laser power for dark-field imaging at 0.25% constant for all images. A minimum of 4 images were taken for each animal liver, and 3 images for each

animal tumour for quantification. Rendering of 3D images and movies was done using Bitplane IMARIS 8.1.

Image segmentation and analysis

All code used in this paper is available on the Github repository located at the following link: (https://github.com/BenKingston/3D_met_NP_analysis), or from the developers of Ilastik (https://www.ilastik.org, or https://github.com/ilastik). Code was evaluated with MATLB 2017b and DIP image v2.7 (http://www.diplib.org or https://github.com/DIPlib/diplib).

Pre-processing

Image files from the Zeiss Lightsheet microscope were collected in the Zeiss .czi file format. In order to reduce file size and facilitate downstream processing the .czi files are converted in MATLAB to uint16 multipage tiff files for each individual channel (nuclei/DAPI, blood vessels/Cy3-GSL-1, micrometastases/Ki67, and nanoparticles/dark-field) from an x/y resolution of 1920 x 1920 to 627 x 627, and a final voxel size of 1 μm x 1 μm x 1 μm. Z-stack size varies between 800 μm to 1500 μm for each individual image captured. All following image processing steps are carried out only with the nuclei/DAPI, blood vessels/Cy3-GSL-1 and cancer cells/Ki67 channels to prepare them for segmentation. Each individual channel undergoes a local and global intensity correction to normalize the signal intensity throughout the image volume.

Ilastik segmentation

Pre-processed images for nuclei, blood vessels and micrometastases channels were used for training the Ilastik (3) pixel classification supervised machine learning algorithm. This supervised machine learning classifier uses a set of 37 different filters that use the intensities, edges and textures of the image to identify the segmented feature of interest. A random forest classifier is used in the learning step to identify the labelled feature. More information about Ilastik can be found here: http://ilastik.org. For each animal and each channel, a single 3D image was randomly selected to train the pixel classification system. In the graphic user interface two labels were used to annotate: 1) The feature of interest (nuclei, blood vessels or micrometastases) and 2) background areas of the image. After 30 mins of this manual labelling the binary segmentation of 1) feature and 2) background voxels was batch applied to all respective images for that specific channel for that individual animal. This was repeated for each of the nuclei, blood vessel and micrometastasis channels for each animal. The result is a binary image stack for each of the nuclei, blood vessel and micrometastasis channels.

Post-processing

Following the Ilastik segmentation a number of artifacts appear that need to be removed with a post-processing script written in MATLAB. In all channels (nuclei, blood vessels, micrometastases and nanoparticle) imaging artifacts commonly appear on the surface of the tissue due to the presence of auto-fluorescent and scattering debris. To make sure these artifacts are removed the first step in post-processing is to define the boundaries of the tissue. This is done using the post-processed nuclei channel. First, a threshold of 40 is applied to the nuclei channel to create a rough binary image of the tissue outline. Next, an

erosion operation using an elliptical object of a diameter of 14 μm is used to remove the outermost layer at the tissue boundary. This eroded tissue boundary is then applied to the Ilastik segmented nuclei, blood vessel, micrometastasis channel, and grayscale nanoparticle intensity channel removing the outermost layer of the tissue boundary in these images. The next step in post-processing the nuclei channel is to perform a seeded watershed operation to separate merged nuclei. Seeds for the watershed algorithm are generated by first preforming an opening operation on the binary nuclei image, followed by a Euclidean distance transformation. The maxima of the distance transformed image are then dilated and used as seeds in the seeded watershed algorithm. This generates the segmented nuclei. To generate simulated cell boundaries the segmented nuclei are grown by 2 μm from the surface of the segmented nuclei, this generates the segmented cell boundaries. For blood vessels the Ilastik segmented blood vessels end up registering as larger than the actual boundaries of the vessels. To fix this the binary Ilastik segmented blood vessels are eroded by 2 μm. This generates the final segmented blood vessel channel. For the micrometastasis segmentation the Ilastik segmented channel first undergoes an opening operation to remove very small segmentation artifacts. Next, any micrometastasis under 30,000 μm^3 in size is removed. Micrometastases under this size consist of only a small number of cells (<10 cells) and could not be reliably segmented out from background noise. The resulting image is the segmented micrometastasis channel.

Segmentation validation

To validate the segmentation workflow the automatically segmented nuclei, blood vessels and micrometastases were compared to manually annotated images. For the nuclei and blood vessel channels three 2D slices from three different animals, and for the micrometastases five 2D slices from five different animals, were randomly selected for manual segmentation. The Dice similarity coefficient was then calculated comparing the automated and manual segmentations in MATLAB.

Feature quantification from images

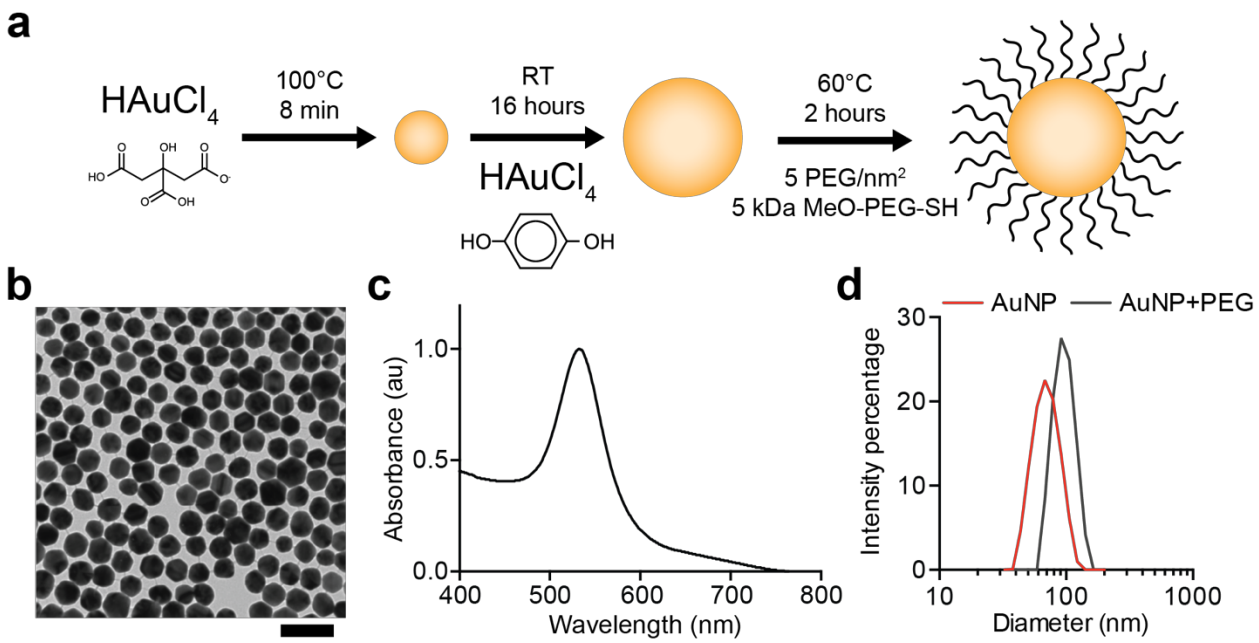
The segmented blood vessel, micrometastasis and nuclei channels along with the nanoparticle intensity channel are used to quantify the physiological and nanoparticle characteristics from the 3D images using a custom MATLAB script. To generate the individual micrometastasis data the segmented micrometastasis is labelled to assign each lesion a unique label. From this we extract the volume, surface area and in conjunction with the nanoparticle channel, the mean nanoparticle intensity of each lesion. To generate single cell information for nanoparticle intensity the dilated segmented nuclei are used as a mask applied to the nanoparticle intensity channel to extract the mean nanoparticle intensity per cell. To calculate the cell distance to the nearest vessel a Euclidean distance transformation is applied to the segmented blood vessel channel to create a greyscale image where the intensity is equal to the distance that voxel is from the nearest vessel. The dilated and segmented nuclei are then used as a mask and the average distance intensity per cell is calculated giving the cells distance from the nearest vessel. The nanoparticle intensity and distance from vessel for each cell is then calculated for each region defined by a micrometastasis. We can then export a list of cells for each micrometastasis with their corresponding mean nanoparticle intensities, cell distances and the number of cells within each lesion. From these parameters we can calculate the surface area to volume ratio,

sphericity, cell density, average cell distance to vessels, number of nanoparticle positive cells, and nanoparticle positive cell density. Nanoparticle positive cells were defined as cells having a mean intensity 3 times higher than the noise signal from cleared tissue that contains no nanoparticles (the background signal) which can be seen in SI Appendix, Fig. S3.

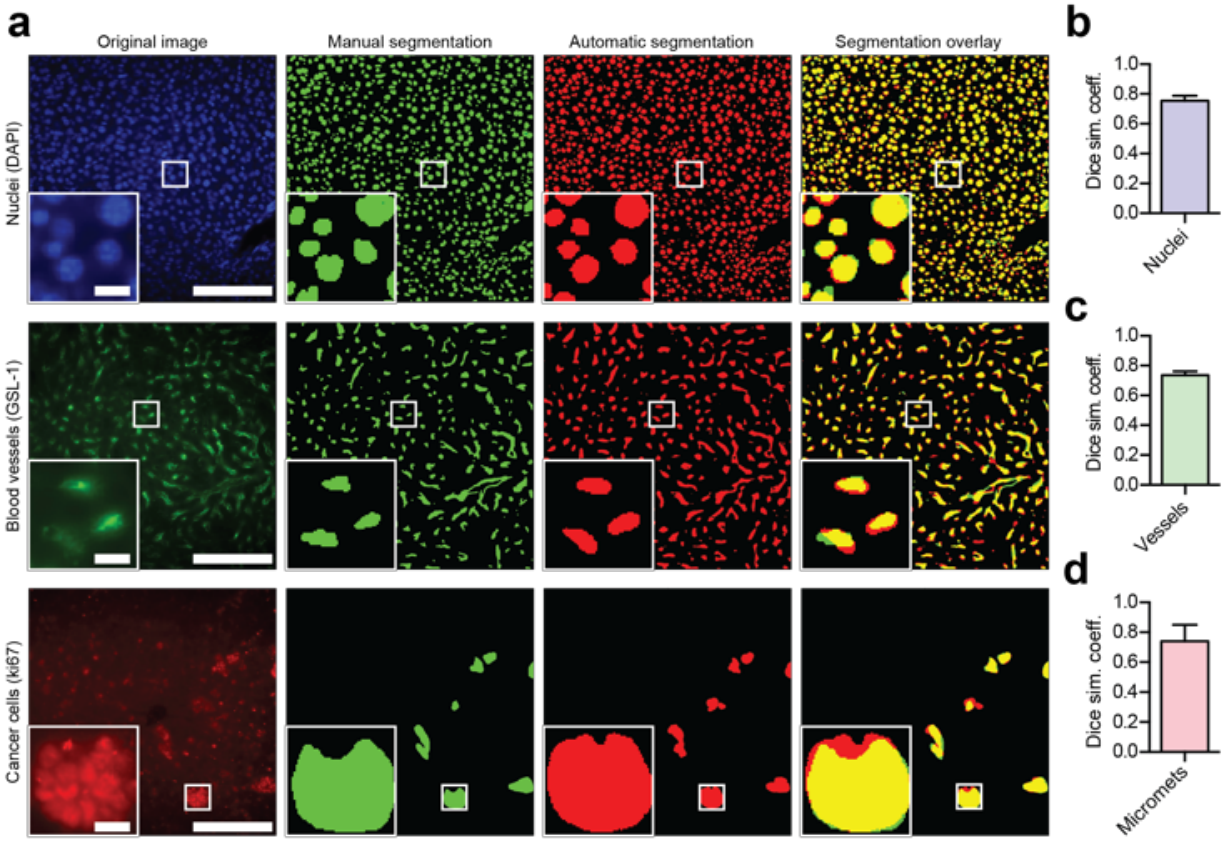
Support Vector Machine modelling

Prior to beginning any model generation the dataset of 1,301 micrometastases was split randomly into a training and cross-validation dataset (80%, 1041 micrometastases) to evaluate potential models, and a test dataset (20%, 260 micrometastases) to evaluate the final predictive models. The entire dataset is attached as Supplemental Excel File 1. All model generation and evaluation was done in MATLAB using the Statistics and Machine Learning toolbox. Initially, linear, quadratic, cubic, and fine, medium and coarse gaussian Support Vector Machine (SVM) models were trained using the training and cross-validation dataset to determine if these models could predict the mean nanoparticle intensity, density of nanoparticle positive cells, or number of nanoparticle positive cells from the inputs of volume, surface area, surface area to volume ratio, sphericity, total cell number, cellular density, mean distance to nearest blood vessel and hours post nanoparticle injection. Based on the performance of the models (SI Appendix, Fig. S4) the quadratic SVM model generally had the best performance. Separate quadratic SVM models that were trained to predict mean nanoparticle intensity, density of nanoparticle cells and number of nanoparticle cells per micrometastasis were then exported. The test dataset was then used to evaluate the prediction of these models with an unseen dataset.

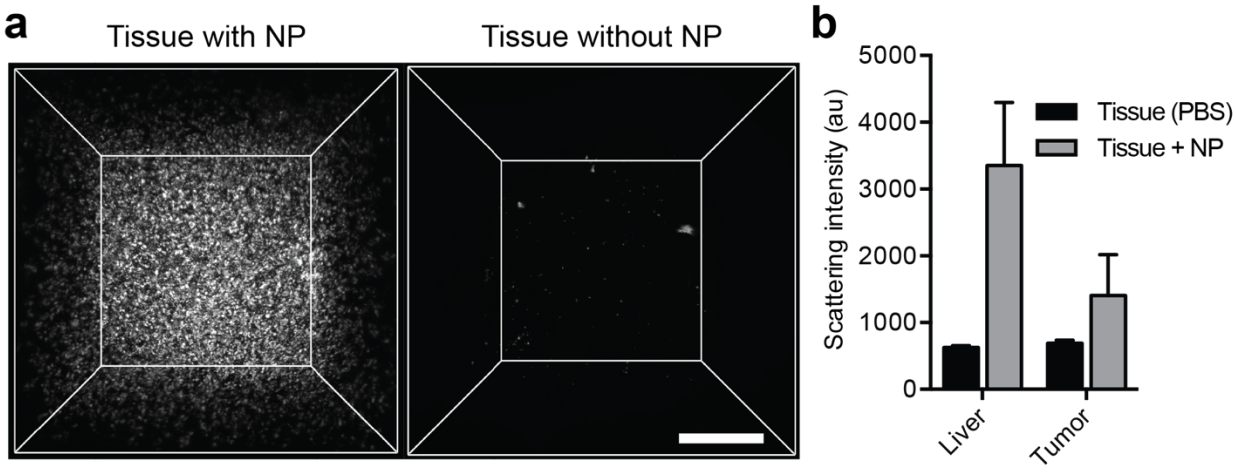
SUPPLEMENTAL FIGURES



Supplemental Figure S1. Gold nanoparticle synthesis and characterization. a) Schematic summarizing the key steps in the synthesis of 50nm gold nanoparticles coated in 5 kDa methoxy-PEG-thiol. b) Transmission electron micrograph of 50 nm gold nanoparticles. Scale bar 100 nm. c) UV-Visible absorbance spectroscopy of 50 nm gold nanoparticles. d) Dynamic light scattering (DLS) intensity distribution of gold nanoparticles before and after PEG conjugation.



Supplemental Figure S2. Nuclei, blood vessel and micrometastasis segmentation validation. a) Images of 2D sections of mouse livers containing micrometastases from full 3D images that were evaluated by both manual and automated segmentation. A 2D slice showing the original fluorescent image, manual segmentation, automated segmentation and an overlay of the manual and automated segmentation is shown for nuclei (top row), blood vessels (middle row) and micrometastases (bottom row). Scale bars are 200 µm for low magnification images and 20 µm for insets. Dice Similarity coefficient (Dice sim. coeff.) for the segmentations of the nuclei (0.75, n=3) (b), blood vessels (0.74, n=3) (c), and micrometastases (0.74, n=5) (d).



Supplemental Figure S3. Darkfield imaging of cleared tissues with and without gold nanoparticles. a) Cleared murine liver tissues imaged with darkfield via light-sheet microscopy with gold nanoparticles 1 hour post injection and after PBS injection. Scale bar 200 μm. b) The mean scattering intensity of cleared liver and primary tumour tissue 1 hour post injection of 50 nm gold nanoparticles or PBS injection. n=3. NP = nanoparticles.

a

Number of nanoparticle positive cells per micrometastasis

Model	RMSE	R^2	r	MSE	MAE
Quadratic SVM	32.0	0.71	0.84	1021	18
Linear SVM	33.8	0.68	0.82	1145	19
Cubic SVM	56.6	0.10	0.32	3203	20
Fine Gaussian SVM	56.5	0.10	0.32	3187	23
Medium Gaussian SVM	47.1	0.38	0.62	2216	19
Coarse Gaussian SVM	42.7	0.49	0.70	1823	20

b

Mean nanoparticle intensity per micrometastasis

Model	RMSE	R^2	r	MSE	MAE
Quadratic SVM	1603	0.31	0.56	2.6×10^6	1127
Linear SVM	1686	0.24	0.49	2.8×10^6	1220
Cubic SVM	1778	0.16	0.40	3.2×10^6	1168
Fine Gaussian SVM	1734	0.20	0.45	3.0×10^6	1261
Medium Gaussian SVM	1590	0.32	0.57	2.5×10^6	1140
Coarse Gaussian SVM	1670	0.26	0.51	2.8×10^6	1200

c

Density of nanoparticle cells per micrometastasis

Model	RMSE	R^2	r	MSE	MAE
Quadratic SVM	2.4×10^5	0.37	0.61	5.9×10^{10}	1.9×10^5
Linear SVM	2.5×10^5	0.32	0.57	6.3×10^{10}	2.0×10^5
Cubic SVM	2.5×10^5	0.31	0.56	6.5×10^{10}	1.9×10^5
Fine Gaussian SVM	2.7×10^5	0.20	0.45	7.5×10^{10}	2.1×10^5
Medium Gaussian SVM	2.4×10^5	0.37	0.61	5.9×10^{10}	1.9×10^5
Coarse Gaussian SVM	2.5×10^5	0.33	0.57	6.2×10^{10}	2.0×10^5

Supplemental Figure S4. Comparison of performance of different support vector machine (SVM) models in nanoparticle delivery prediction. Performance as measured by root mean squared error (RMSE), R squared (R^2), Pearson correlation (r), mean squared error (MSE), and mean absolute error (MAE) for the different SVM models tested on the training data-set with 5-fold cross validation for predicting the number of nanoparticle cells (a), mean nanoparticle intensity (b), and density of nanoparticle cells (c) per micrometastasis. Number of NP+ cells are in units of cells. Mean nanoparticle intensity is in arbitrary units. Density of NP+ cells in in number of cells per mm^3.

a

SVM model performance with test dataset

Model	RMSE	R^2	r	MSE	MAE
Number NP+ cells	27	0.88	0.94	716	16
Mean NP Intensity	1559	0.36	0.60	2.4×10^6	1114
Density NP+ cells	2.4×10^5	0.43	0.66	5.8×10^{10}	1.9×10^5

Supplemental Figure S5. Performance of final quadratic support vector machine (SVM) models in nanoparticle delivery prediction with test dataset. a) Performance of the three SVM models using the test dataset as measured by root mean squared error (RMSE), R squared (R^2), Pearson correlation (r), mean squared error (MSE), and mean

absolute error (MAE). NP = nanoparticle. NP+ = nanoparticle positive cells. Number of NP+ cells are in units of cells. Mean nanoparticle intensity is in arbitrary units. Density of NP+ cells in in number of cells per mm^3.

SUPPLEMENTAL VIDEOS

Supplemental Video S1. 3D video of Figure 1c showing an individual micrometastasis in the lung. Nuclei are in blue, blood vessels are in green, Ki67 is in red and nanoparticles are in white.

Supplemental Video S2. 3D video of Figure 1e showing individual micrometastases within the liver. Nuclei are in blue, blood vessels are in green, Ki67 is in red and nanoparticles are in white.

Supplemental Video S3. 3D video of Figure 6a showing the distribution of cells, blood vessels and nanoparticles within the primary tumour. Pseudo-colored images of the segmented cells are shown to depict the nanoparticle intensity of individual cells (dark red is low nanoparticle intensity, white is high nanoparticle intensity), and cell distance to the nearest vessel (blue is close to a vessel, pink is far away from a vessel). Nuclei are in blue, blood vessels are in green, and nanoparticles are in white.

Supplemental Video S4. 3D video of Figure 6b showing the distribution of cells, blood vessels, micrometastases and nanoparticles within the liver. Pseudo-colored images of the segmented micrometastasis cells are shown to depict the nanoparticle intensity of individual cells (dark red is low nanoparticle intensity, white is high nanoparticle intensity), and cell distance to the nearest vessel (blue is close to a vessel, pink is far away from a vessel). Nuclei are in blue, blood vessels are in green, Ki67 is in red and nanoparticles are in white.

Supplemental Video S5. 3D video of Figure 7a and 7b showing the distribution of different morphologies of micrometastases within the liver. Nuclei are in blue, blood vessels are in green, Ki67 is in red and nanoparticles are in white. Pseudo-colored images of the segmented micrometastasis cells are shown to depict the nanoparticle intensity of individual cells (dark red is low nanoparticle intensity, white is high nanoparticle intensity), and cell distance to the nearest vessel (blue is close to a vessel, pink is far away from a vessel).

SUPPLEMENTAL DATASET

Supplemental Excel File 1. Physiological and nanoparticle characteristics of all micrometastases. The complete dataset as well as the randomly separated: 1) training and cross-validation dataset, and 2) test dataset.

REFERENCES

1. Perrault SD, Chan WCW (2009) Synthesis and Surface Modification of Highly

Monodispersed , Spherical Gold Nanoparticles of 50 - 200 nm. *J Am Chem Soc* 131(47):17042–17043.
2. Syed AM, et al. (2017) Three-Dimensional Imaging of Transparent Tissues via Metal Nanoparticle Labeling. *J Am Chem Soc* 139(29):9961–9971.
3. Sommer C, Straehle C, Ullrich K, Hamprecht FA (2011) ilastik : Interactive Learning and Segmentation Toolkit. *Eighth IEEE Int Symp Biomed Imaging*:230–233.