



Published in final edited form as:

Cell. 2019 December 12; 179(7): 1661–1676.e19. doi:10.1016/j.cell.2019.11.013.

Deep learning reveals cancer metastasis and therapeutic antibody targeting in the entire body

Chenchen Pan^{1,2,18}, **Oliver Schoppe**^{3,4,18}, **Arnaldo Parra-Damas**^{2,18}, **Ruiyao Cai**^{1,2}, **Mihail Ivilinov Todorov**^{1,2,5}, **Gabor Gondi**⁶, **Bettina von Neubeck**⁶, **Nuray Bö ürcü-Seidel**⁷, **Sascha Seidel**⁸, **Katia Sleiman**^{4,9,10}, **Christian Veltkamp**^{4,9,10}, **Benjamin Förstera**^{1,2}, **Hongcheng Mai**^{1,2}, **Zhouyi Rong**^{1,2}, **Omelyan Trompak**⁷, **Alireza Ghasemigharagoz**², **Madita Alice Reimer**², **Angel M. Cuesta**⁸, **Javier Coronel**³, **Irmela Jeremias**^{11,12,13}, **Dieter Saur**^{4,9,10}, **Amparo Acker-Palmer**⁸, **Till Acker**⁷, **Boyan K. Garvalov**^{7,14}, **Bjoern Menze**^{3,4,17}, **Reinhard Zeidler**^{6,15}, **Ali Ertürk**^{1,2,16,19,*}

¹Institute for Tissue Engineering and Regenerative Medicine (iTERM), Helmholtz Zentrum München, 85764 Neuherberg, Germany

²Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig Maximilian University of Munich (LMU), 81377 Munich, Germany

³Department of Computer Science, Technical University of Munich, 81675 Munich, Germany

⁴Institute of Translational Cancer Research and Experimental Cancer Therapy, Klinikum rechts der Isar, Technische Universität München, 81675 Munich, Germany

⁵Graduate School of Systemic Neurosciences (GSN), 82152 Munich, Germany

⁶Helmholtz Zentrum München, Research Unit Gene Vectors, 81377 Munich, Germany

⁷Institute of Neuropathology, University of Giessen, 35390 Giessen, Germany

⁸Institute of Cell Biology and Neuroscience and Buchmann Institute for Molecular Life Sciences (BMLS), University of Frankfurt, 60323 Frankfurt, Germany

⁹Division of Translational Cancer Research, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

¹⁰German Cancer Consortium (DKTK), Partner Site Munich, Klinikum rechts der Isar, Technische Universität München, 81675 Munich, Germany

*Correspondence: erturk@helmholtz-muenchen.de.

AUTHOR CONTRIBUTIONS

C.P., A.P.D., and R.C. performed the tissue processing, clearing and imaging experiments. C.P. and M.I.T. stitched and assembled the full body scans. O.S. developed the deep learning architecture and performed the quantitative analyses. J.C. helped optimizing the network architecture. A.G. performed the image rendering and preliminary data analysis. O.S. developed annotation tools and the custom-made object detector, M.A.R. annotated the data. C.P. and O.S. reviewed the predictions identified by the algorithm. C.P., H.M., Z.R., and M.I.T. manually segmented the internal organs for registration. C.P., G.G., B.N., R.Z., N.B.S., O.T., S.S., T.A., A.M.C., A.A.P., K.S., C.V., D.S. and I.J. performed tumor transplantation experiments and bioluminescence imaging. B.F. helped to obtain the animal experiments license. B.K.G. helped with data interpretation. B.M. provided guidance in developing the deep learning architecture and helped with data interpretation. A.E., C.P., and O.S. wrote the manuscript. All the authors edited the manuscript. A.E. initiated and led all aspects of the project.

DECLARATION OF INTERESTS

A.E. has filed a patent related to some of the technologies presented in this work.

¹¹Research Unit Apoptosis in Hematopoietic Stem Cells, Helmholtz Zentrum München, German Center for Environmental Health (HMGU), 81377 Munich, Germany

¹²Department of Pediatrics, Dr. von Hauner Childrens Hospital, Ludwig Maximilians University, 81377 Munich, Germany

¹³German Consortium for Translational Cancer Research (DKTK), Partnering Site Munich, 80336 Munich, Germany

¹⁴Department of Microvascular Biology and Pathobiology, European Center for Angioscience (ECAS), Medical Faculty Mannheim, University of Heidelberg, 68167 Mannheim, Germany

¹⁵Department for Otorhinolaryngology, Klinikum der Universität München, Ludwig Maximilian University of Munich (LMU), 81377 Munich, Germany

¹⁶Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany

¹⁷Munich School of Bioengineering, Technical University of Munich, 85748 Munich, Germany

¹⁸These authors contributed equally

¹⁹Lead contact

SUMMARY

Reliable detection of disseminated tumor cells and of the biodistribution of tumor-targeting therapeutic antibodies within the entire body has long been needed to better understand and treat cancer metastasis. Here, we developed an integrated pipeline for automated quantification of cancer metastases and therapeutic antibody targeting, named DeepMACT. First, we enhanced the fluorescent signal of cancer cells more than 100-fold by applying the vDISCO method to image metastasis in transparent mice. Second, we developed deep learning algorithms for automated quantification of metastases with an accuracy matching human expert manual annotation. Deep learning-based quantification in 5 different metastatic cancer models including breast, lung, and pancreatic cancer with distinct organotropisms allowed us to systematically analyze features such as size, shape, spatial distribution, and the degree to which metastases are targeted by a therapeutic monoclonal antibody in entire mice. DeepMACT can thus considerably improve the discovery of effective antibody-based therapeutics at the pre-clinical stage.

INTRODUCTION

The metastatic process is complex and affects diverse organs (Hanahan and Weinberg, 2011; Lambert et al., 2017; Massague and Obenauf, 2016). As most cancer patients die of metastases at distant sites developing from disseminated tumor cells with primary or acquired resistance to therapy, a comprehensive and unbiased detection of disseminated tumor cells and tumor targeting drugs within the entire body is crucial (de Jong et al., 2014). Such technology would help to explore mechanisms affecting tumor metastasis and drug targeting in preclinical mouse models much more reliably, hence substantially contributing to the development of improved therapeutics. So far, such efforts have been hampered by the lack of 1) imaging technologies to reliably detect all individual metastases and disseminating tumor cells in mouse bodies, and 2) algorithms to quickly and accurately quantify large-

scale imaging data. Here, we developed an analysis pipeline that allows us to efficiently overcome these limitations.

First, we built upon recently developed tissue clearing methods for entire fixed mice (Cai et al., 2018; Pan et al., 2016; Tainaka et al., 2014; Yang et al., 2014) to address the imaging problem. Typically, fluorescent labeling of cancer cells *in vitro* or *in vivo* is achieved by endogenous expression of fluorescent proteins such as GFP, YFP, and mCherry, which emit light in the visible spectrum. However, many tissues in the mouse body show high autofluorescence in this range (Tuchin, 2016; Zipfel et al., 2003), which hinders reliable detection of single cancer cells or small cell clusters in mouse bodies based on their endogenous fluorescent signal. To circumvent this problem, we chose to implement the vDISCO technology (Cai et al., 2018), which enhances the signal of fluorescent proteins of cancer cells more than 100-fold in cleared tissues, enabling reliable imaging not only of large metastases but also micrometastases throughout the entire body.

Second, systematic analysis of metastasis in adult mouse bodies requires quantitative information such as location, size, and shape of all individual metastases. Manual detection and segmentation of numerous metastases in highly resolved full body scans is an extremely laborious task that may take several months per mouse for an expert annotator. In addition, automation by filter-based 3D object detectors is not reliable, as different body tissues have different levels of contrast (Pan et al., 2016), causing a high rate of false positive and false negative metastasis detections. Recent studies have demonstrated the high efficacy of deep learning-based analysis of biomedical images, compared to filter-based or manual segmentation methods (Camacho et al., 2018; Christiansen et al., 2018; Esteva et al., 2017; Kermany et al., 2018; Sullivan et al., 2018; Topol, 2019; Wang et al., 2019). To enable automated, robust, and fast mapping of all metastases in transparent mice, we developed an efficient deep learning approach based on convolutional neural networks (CNNs) and optimized it for vDISCO imaging data and metastasis distribution patterns.

Resolving these two bottlenecks allowed us to build an integrated, highly automated pipeline for analysis of metastasis and tumor-targeting therapeutics, which we named DeepMACT (Deep learning-enabled Metastasis Analysis in Cleared Tissue). Using DeepMACT, we detected cancer metastases and even individual disseminated tumor cells in mouse bodies, including many metastases previously overlooked by human annotators. Furthermore, this enabled analyzing the targeting efficiency of a therapeutic antibody against carbonic anhydrase XII on the level of individual metastases. As a scalable, easily accessible, fast, and cost-efficient method, DeepMACT enables a wide range of studies on cancer metastasis and therapeutic strategies. To facilitate adoption of DeepMACT, a step-by-step handbook (Methods S1), the protocols for clearing and imaging, the deep learning algorithm, the training data, and the trained model are available online to address diverse questions in cancer research.

RESULTS

Focusing on a clinically relevant tumor model, we transplanted human MDA-MB-231 mammary carcinoma cells, expressing mCherry and firefly luciferase, into the mammary fat

pad of NOD *scid* gamma (NSG) mice and allowed the tumors to grow and metastasize for 6-10 weeks (Figure 1A) (Iorns et al., 2012; von Neubeck et al., 2018). Furthermore, we injected the fluorescently-tagged 6A10 therapeutic antibody that has been shown to reduce tumor burden in this model (Gondi et al., 2013; von Neubeck et al., 2018). To comprehensively assess cancer cell dissemination and therapeutic antibody targeting in mouse bodies at the level of individual micrometastases, we developed DeepMACT. In short, we transcardially perfused the animals using standard PFA fixation and applied the vDISCO method to enhance the fluorescent signal of tumor cells. After light-sheet microscopy, the 3D image stacks of entire transparent **mouse** bodies were analyzed using deep learning algorithms. The DeepMACT pipeline consists of 1) vDISCO panoptic imaging of cancer metastases in transparent mice and 2) deep learning-based analysis of cancer metastasis and antibody drug targeting (Figure 1B).

DeepMACT step 1: vDISCO imaging of cancer metastases in optically cleared mice

We previously developed the vDISCO technology to image single fluorescent cells in mouse bodies through intact bones and skin (Cai et al., 2018). The vDISCO method utilizes bright fluorescent dyes conjugated with nanobodies to enhance the fluorescent signal of the cells that is weakened during the fixation and clearing process. Here, we first applied vDISCO to increase the fluorescence signal of mCherry-expressing cancer cells. By enhancing the tumor cell fluorescence with anti-mCherry or anti-GFP nanobodies conjugated to Atto-594 or Atto-647N dyes, we found that nanobodies can increase the signal strength of cancer cells over 100 times compared to imaging the endogenous mCherry signal after clearing (Figure S1). Owing to this significant enhancement in signal contrast, we could readily detect micrometastases buried in centimeters-thick mouse bodies (Figure S1F-L) e.g., in deep brain and spinal cord regions through the intact skull and vertebrae (Figure S1F,I, yellow arrowheads). To confirm the specificity of vDISCO enhancement of the signal from mCherry expressing cancer cells, we performed the following experiments: 1) we stained control mice without a tumor transplant, thereby lacking mCherry expression, and found no labeling in any of the analyzed organs (Figure S2A); 2) we analyzed the primary tumors and lung metastases from the mouse bodies by staining them using a specific anti-luciferase antibody, which confirmed that endogenous mCherry fluorescence co-localized with both the signals from nanobodies and from the anti-luciferase antibody (Figure S2B,C).

Since the detection of smaller-sized tumor cell clusters, which may represent dormant cancer cells or incipient metastatic nodules, is critical, we next tested if vDISCO allows imaging cancer micrometastases in mouse bodies. In order to compare our approach to conventional methods, we also acquired bioluminescence images of mice before applying DeepMACT. In line with previous findings (Iorns et al., 2012), we detected the earliest large metastasis of transplanted MDA-MB-231 cells at the axillary lymph node of mice by bioluminescence (Figure 2A, Figure S3). However, bioluminescence imaging did not reveal any detailed information such as size or shape and failed to show the presence of micrometastases.

After bioluminescence assessment, we applied vDISCO using anti-mCherry signal enhancing nanobodies conjugated to Atto-647N and imaged the mouse bodies first using epifluorescence in 2D (Figure 2B-G), then using light-sheet microscopy in 3D (Figure 2H-

L). In epifluorescence, we could readily see both the primary tumor (Figure 2F) and the major metastases at the axillary lymph node (Figure 2D), which were also detected by bioluminescence imaging (Figure 2A), albeit as a bulk signal, lacking information on real size and shape. Importantly, our approach allowed the visualization of several micrometastases in the lungs with conventional epifluorescence imaging, which were not visible in bioluminescence (compare the magenta marked regions in Figure 2A with Figure 2B, and red arrowheads in Figure 2E; more examples shown in Figure S3). Thus, vDISCO followed by epifluorescence imaging, which can be completed within minutes, already provided greater details and sensitivity compared to bioluminescence imaging. Next, we imaged entire fixed transparent mice using a light-sheet microscope (Cai et al., 2018) in 3D to detect individual micrometastases throughout the body (Figure 2H). In the chest area, we could see various metastases not only in the lungs (yellow segmented region in Figure 2I) and lymph nodes, but also at the base of the neck and surrounding tissues (Figure 2J–L and Video S1). Importantly, light-sheet microscopy scanning allowed us to image even single disseminated tumor cells in the mouse body. Examples of single disseminated tumor cells resolved in full body scans are shown in Figure 2M (see also Video S2), which were further verified by high magnification light-sheet microscopy imaging showing the colocalization of each single tumor cells with a single nucleus stained by PI (Figure 2N). Thus, our approach allows for the first time to detect micrometastases in full body scans of mice in 3D down to the size of individual cells.

DeepMACT step 2: Deep learning for detection and quantification of metastases

We developed an optimized deep learning-based approach to detect and segment all cancer metastases in full-body scans of mice. This framework solves the 3D task of detecting and segmenting metastases in volumetric scans with CNNs that process 2D projections of small sub-volumes (Figure 3A). In brief, we first derived three 2D maximum intensity projections (aligned with the x-, y-, and z-axes) for each sub-volume in order to increase the signal-to-noise ratios (SNR). We fed the resulting projections to the CNN and obtained 2D probability maps, in which each pixel value represents the estimated probability that this pixel identifies a metastasis under the given projection. We then reconstructed a 3D segmentation from the three projections observing increased reliability in detecting true positive metastases while safely ignoring non-metastatic tissue that would produce false positives in the individual projections. For example, in Figure 3B, the green arrows show successful detection of a real metastasis and the red arrows show successful ignoring of a structure that could be mistaken for a metastasis from a single 2D projection. This approach was highly effective in detecting and segmenting metastases in the imaging data, yielding a binary mask for all metastases in the body.

The core of our architecture makes use of CNNs (Figure 3C), structurally similar to the established U-net (Ronneberger et al., 2015), which learn to distinguish metastases from the background signal. This is achieved by using a deep stack of encoding units, which detect characteristic cancer features, and a corresponding stack of decoding units, which segment each metastasis at pixel-level. Each encoding unit performs two convolutions, extracting information about the environment for each pixel and representing that information in a third dimension - the feature channels. Before being passed on to the next encoding unit, the

image is spatially down-sampled. Together, this means that the neural network is steadily increasing the feature channels and steadily decreasing the spatial resolution, enforcing the network to learn even more abstract representations of the data (i.e., features) in the deeper layers, before mapping the information relevant to cancer cells back to the original resolution in the decoding upward path. This happens by up-sampling the abstract, low-resolution information from lower layers and concatenating it with the less abstract, but higher-resolution information from the encoding path via skip connections (some exemplary visualizations of the computational stages are presented in Figure S4A–C).

To assess the reliability of our automated deep learning architecture, we applied it to a fresh test set of a full-body scan, which was neither used for training the CNNs nor to optimize hyperparameters. The data sets were manually annotated by human experts and any disagreements between experts were jointly reviewed and discussed in order to derive a refined, commonly agreed reference annotation (see STAR Methods for details). We then systematically compared the performance of our deep learning approach to that of established detection methods as well as the performance of a single human annotator, calculating F1-score (also known as Dice score), a common performance measure based on both the metastasis detection rate (recall) and false positive rate (precision).

As shown in Figure 3D, we found that DeepMACT reached an F1-score of 80%, outperforming existing filter-based detectors such as the ImageJ Object Detector (18%) or a custom-made filter-based detector (36%) by a large margin. The similar performance of 3D CNNs such as a customized 3D U-net (38%) highlights the benefit of the specialized DeepMACT approach for the tumor models we tested. Indeed, the detection performance of DeepMACT comes very close to the level of a single human expert annotator with an F1-score of 83%. The slightly higher F1-score of the human annotator is mainly driven by the high precision. However, the human annotator missed around 29% of all micrometastases (examples are shown in Figure S4D–F) and detecting those false negatives would require a repetitive and very laborious re-analysis of the entire animal scans, requiring up to several months of human work time. On the other hand, the F1-score of DeepMACT is a result of a balance between precision and recall, which can be freely adjusted via the model's threshold. For DeepMACT, we can increase detection rate (recall) over 95%. While this also increases the false-positive rate, correcting the false positive data requires only a review of detected signals by a human annotator, which we completed within 1 hour per mouse in this study (a typical example for a false positive detection is shown in Figure S4G–I). Combining the DeepMACT prediction with this quick review yielded an F1-score of 89%, exceeding the performance of a single human annotator. A more detailed analysis on the trade-off between precision and recall is shown in Figure S4J. Notably, DeepMACT could detect micrometastases about 30 times faster than filter-based detectors and over 300 times faster than a human annotator (Figure 3E). Even taking the time for a manual review of the DeepMACT prediction into account, the total processing speed was still 8 times faster than filter-based detectors and over 60 times faster than a human annotator, who was already supported by a dedicated and interactive software, custom-built for this task and these data; without annotation software, the human manual annotation would be estimated to take several months for a single mouse. Thus, DeepMACT can complete months to years of human labor within hours without compromising on segmentation quality.

DeepMACT reliably detects micrometastases in different tumor models

After establishing the DeepMACT pipeline, we used it to analyze full mouse bodies. Apart from the primary tumor and the macrometastasis in the axillary lymph node, we could detect hundreds of micrometastases of varying sizes throughout the body, especially in the lungs (Figure 4A–B). Overall, DeepMACT identified 520 micrometastases throughout the entire body in this particular mouse, of which there were 306 in the lungs and 214 throughout other organs of the body (Figure 4C). We found that micrometastases are mostly located in the inner tissue layers (about 1 cm depth from the surface), as shown by color-coding in Figure 4D, making them particularly difficult to detect by other methods. To analyze the spatial distribution with regard to the lung anatomy, we registered all 306 lung micrometastases to the mouse lung lobes. We found that micrometastases were evenly distributed in all lobes (Figure 4E). Interestingly, the micrometastases were randomly distributed throughout the lungs regardless of their size, suggesting independent colonization at multiple sites. Furthermore, we quantified the size and relative location of all micrometastases in the entire body (Figure 4F–L). While 79% of micrometastases were within 1 mm to the nearest neighboring micrometastasis, we also found highly isolated micrometastases as distant as 9.3 mm apart from their nearest neighbor (Figure 4G). Importantly, we found a large number of micrometastases with estimated cell counts of a few hundred cells or less (Figure 4F,H) and diameters less than 50-100 μm (Figure 4I), which would be very difficult to detect in mice by other methods. Comparing the micrometastases in the lungs with those in the torso, we found that the tumor burden in the lungs was more than a hundred times higher in this tumor model (Figure 4J). Also, micrometastases in the lungs were, on average, 30% larger in diameter (Figure 4K), with a more than two-fold higher estimated cell number per metastasis, compared to micrometastases in the rest of the torso (Figure 4L).

To verify the robustness and applicability of the DeepMACT pipeline for a wider range of experimental settings, we conducted additional studies. First, we implanted a solid tumor (MDA-MB-231 breast cancer grown in another mouse for 10 weeks) into a healthy mouse and analyzed it right away, leaving no time for metastases to form. As expected, no metastases could be found in this control, indicating that tumor cells do not detach from a solid tumor during the tissue clearing procedure and that no artefacts (such as potential unspecific nanobody accumulations during the staining procedure) would be mistaken for metastases (Figure 5A).

Second, we applied the pipeline to 3 different tumor models with distinct metastatic propensity and organotropism. A nude mouse intracardially injected with human MCF-7 estrogen receptor (ER)-positive breast cancer cells developed metastases throughout the body, with a substantial burden in the lungs (49 metastases), liver (18) and kidneys (11), but also in the bones (2; indicated by yellow arrows) and the brain (1) (Figure 5B). A C57BL/6 mouse transplanted with murine syngeneic R254 pancreatic cancer cells, however, did not develop any metastases in the brain, kidneys or bones, but rather in the lungs (8), the liver (6) and also in distinct tissues such as the peritoneum (Figure 5C; metastasis in peritoneum indicated with a magenta arrow). A further model using the human brain metastatic lung cancer cell line H2030-BrM3 transplanted in nude mice only showed few metastases in the liver (2) or kidneys (1) but many in the brain (31) (Figure 5D). These experiments

demonstrate that the DeepMACT pipeline can reliably detect micrometastases in a variety of tumor models with distinct organotropisms, including different immunodeficient or immunocompetent mouse strains, syngeneic tumors and xenotransplants. Furthermore, metastases can be quantified and assessed by DeepMACT in organs in which this is difficult to achieve by other methods, such as bones and the brain.

In a third experiment, we tested the potential of DeepMACT to study the progression of the metastatic process over time. We injected MDA-MB-231 cancer cells intracardially and analyzed the distribution of metastases 2 days, 6 days, and 14 days after injection (Figure 5E–G). We found metastases in the brain, lungs, liver, kidneys, bones, and other organs at all time points. Moreover, our results showed a substantial increase in the total metastatic burden in the mouse bodies as well as in the lungs as the primary metastatic organ (Figure 5H).

Importantly, neither the increase in overall tumor burden for the time-course study nor the differential distribution of metastases across organs for any of the cancer models tested were clearly revealed by bioluminescence images (Figure S5). Thus, our pipeline is the first to enable quantitative analyses of micrometastases in full-body scans, greatly enhancing our ability to assess the metastatic process in a comprehensive manner.

DeepMACT reveals therapeutic antibody targeting at the level of single metastases

A number of tumor-targeting monoclonal antibodies have become part of the standard treatment for various solid and hematological malignancies and many more are in early or late stages of clinical development (Barker and Clevers, 2006; Pandey and Mahadevan, 2014). However, so far there has been no methodology to determine the distribution of therapeutic antibodies across the entire body, down to the level of single micrometastases. Here, we used DeepMACT to assess the biodistribution of the therapeutic monoclonal antibody 6A10 directed against human carbonic anhydrase XII (CA12) (Battke et al., 2011; Gondi et al., 2013; von Neubeck et al., 2018). CA12 is overexpressed in various types of cancers and blocking its activity with the antibody 6A10 reduces tumor growth (Gondi et al., 2013) and increases the sensitivity of tumors to chemotherapy (von Neubeck et al., 2018). We intravenously injected 20 μ g of 6A10 conjugated to Alexa-568 nine weeks after transplantation of MDA-MB-231 cells and perfused the mouse two days after the antibody injection for full-body-scale analysis, enhancing the tumor signal with Atto-647N. Because Alexa-568 excitation/emission spectra overlap with the endogenous mCherry signal of the transplanted cancer cells, we confirmed that the vDISCO pipeline completely eliminates the signal from endogenously expressed mCherry (Cai et al., 2018) (Figure S6).

We first acquired 2D images with epifluorescence microscopy and observed an accumulation of the 6A10 antibody at the primary tumor (Figure 6A,E; tumor shown in magenta, therapeutic antibody in cyan) and the metastases at the axillary lymph node (Figure 6A,B). Focusing on the lungs, we detected micrometastases that were targeted by the 6A10 antibody (Figure 6C, white arrow) and others that were not (Figure 6D, yellow arrow). Acquiring 3D scans with light-sheet microscopy, we assessed the complete biodistribution of the therapeutic antibody and micrometastases throughout the mouse body (Figure 6F–H, Video S3). The axillary lymph node metastases and the micrometastases in the lungs are

shown in Figure 6F. Analyzing the signal of individual micrometastases and the 6A10 antibody by light-sheet microscopy in 3D, we could evaluate the efficiency of antibody drug targeting for all the micrometastases (Figure 6G, white arrows). We also verified the targeting of micrometastases by the 6A10 antibody in different organs such as lungs and kidney, using confocal microscopy (Figure S7).

Next, we used DeepMACT to systematically assess and quantify the efficiency of antibody drug targeting in full body scans at the level of single micrometastases (Figure 6I). While overall 77% of metastases were targeted by the antibody, we found that significantly more micrometastases were targeted in the lungs (85%) as compared to the rest of the body (66%) (Figure 6J, Video S3 and S4). To further assess the efficiency of drug targeting for micrometastases in the lung versus the rest of the body, we evaluated the antibody concentration by quantifying the antibody signal contrast (relative signal strength versus local surrounding; see STAR Methods for details) (Figure 6K). Metastases in the lungs generally tended to have a higher antibody signal ratio, in line with the higher share of targeted metastases. In addition, the antibody signal ratio was much more narrowly distributed compared with micrometastases outside the lungs. The lower average and wider distribution of antibody signal ratio in the micrometastases in the rest of the body indicate that there is a substantially higher variance in the antibody targeting to the cells of those micrometastases. While some are very strongly targeted, many others are not targeted at all. The largest quartile of micrometastases was significantly more likely targeted (88%) than the smallest quartile (67%) (Figure 6L). We also identified various off-target binding sites throughout the body, i.e. binding of the therapeutic antibody to mouse tissues, which is presumably due to unspecific interactions since 6A10 does not bind to murine CA12 (cyan inset in Figure 6H). Overall, these data demonstrate that DeepMACT provides a powerful platform to track the biodistribution of therapeutic antibodies along with micrometastases in mouse bodies. Thus, it represents the first method that allows quantitative analysis of the efficiency of antibody-based drug targeting at the full body scale, with a resolution down to the level of individual micrometastases.

Exploring potential mechanisms of antibody drug targeting

The above results demonstrated that antibody-based drugs, which are the basis of many targeted/personalized treatments, may miss as many as 23% of the micrometastases. Next, we aimed to explore potential mechanisms that might explain this failure. We first hypothesized that the efficiency of targeting of micrometastases might depend on the availability of nearby blood supply transporting the therapeutic antibody. To explore if the vascularization of defined tissue regions can have an effect on antibody drug targeting, we performed lectin labeling of vessels in the lungs, where most of the micrometastases are located. Analyzing diverse micrometastases of different sizes, we found that each of them had blood vessels within a distance of 1-6 μm (Figure 7A,B). This distance is smaller than even a single cell diameter ($\sim 10 \mu\text{m}$) suggesting that absence of nearby blood vessels could not be the major reason for the lack of antibody drug targeting (Tabrizi et al., 2010).

Next, we hypothesized that the tumor microenvironment at the sites of metastases could be related to the efficiency of targeting. If so, we would expect a non-random spatial

distribution of targeted and untargeted metastasis on a local scale. To address this, we turned to DeepMACT and assessed the local clustering of micrometastases targeted by the antibody. We quantified the distances between micrometastases and their nearest neighbor for all micrometastases in the entire body, differentiating between targeted and untargeted nearest neighbors. The distance between two neighboring metastases is smaller for two targeted metastases (about 0.8 mm) than for two untargeted or a mixed pair of an untargeted and a targeted metastasis (consistently at about 1.7-2.0 mm) (Figure 7C,D). Importantly, the average distance from an untargeted to the nearest targeted metastasis is significantly larger than from a targeted one. This would not be expected in a random distribution and indicates a clustering on a local scale. Thus, these analyses suggest the existence of factors in tumor microenvironments influencing the efficiency of antibody drug targeting.

DISCUSSION

Unbiased, comprehensive detection of cancer metastases and the biodistribution of tumor-targeting therapeutics at the level of single micrometastases would substantially accelerate pre-clinical cancer research. Towards this goal, we capitalized on a powerful tissue clearing and imaging method combined with deep learning-based analysis, enabling us to visualize and analyze cancer metastasis in transparent mouse bodies. The resulting DeepMACT workflow is a straightforward method for systemic analysis of micrometastases and therapeutic antibody drug distribution at the full body scale and with a resolution down to individual micrometastases within days, a task that would otherwise take several months to years of human labor. Thus, DeepMACT-based evaluation of entire transparent mouse bodies instead of selected tissues/organs can foster the development and translation of new therapies from pre-clinical research much more efficiently than traditional methods.

To further facilitate easy adoption of our technology by diverse labs, we provide 1) a handbook (Methods S1) with detailed step-by-step instructions for carrying out the DeepMACT pipeline; 2) various resource videos and troubleshooting tips; 3) a package including the trained DeepMACT algorithm and annotated data; and 4) an online version of the DeepMACT algorithm that can be executed via any web browser (hosted by the Code Ocean initiative) without downloading any code nor installing any software (links to these resources are provided in STAR Methods).

DeepMACT technology

Here we set out to make use of recent technologies that can provide scalable and unbiased histological assessment of entire biological specimens. Most full body scale clearing and imaging studies have so far relied on visualization of endogenous fluorescent signal, which is not sufficiently strong to allow imaging and quantification of metastases in transparent mice (Kubota et al., 2017; Pan et al., 2016). To overcome this, we adopted the vDISCO mouse clearing and staining technology, as it can enhance the fluorescent signal in fixed and cleared tissues by more than 100 times (Cai et al., 2018), ensuring reliable detection of micrometastases. Because vDISCO employs nanobody enhancement of the endogenous fluorescent signal, currently up to 21 types of fluorescent proteins can be labelled with available nanobodies. In addition, conjugation of existing nanobodies with fluorescent dyes

at diverse spectra, including those in the near infrared range would help to generate more options for multiplex experiments including imaging of more than one type of fluorescently labeled cell along with conjugated therapeutic antibodies.

Secondly, we developed a highly efficient deep learning architecture based on U-net like CNNs exploiting 2D maximum-intensity projections with high SNR to reliably detect metastases in 3D. Deep learning-based detection not only serves the purpose of automation but also provides a very effective tool in finding metastases that would be easily overlooked by humans. In our data, an expert human annotator missed around 29% of all metastases. This is in line with previous studies where human experts missed 1 in 4 breast cancer metastases in histopathology (Vestjens et al., 2012), a problem that is further exacerbated if humans work under time pressure (Ehteshami Bejnordi et al., 2017). Motivated by this, deep-learning-based approaches for cancer and metastasis detection have recently started gaining substantial momentum for various imaging modalities, also beyond microscopy (Litjens et al., 2016; Liu et al., 2018; Steiner et al., 2018; Wang et al., 2017).

Here we used an MDA-MB-231 cancer cell-based tumor model to train the algorithms. While training deep networks in general may require large training datasets to diversify their applications, the U-net-like architecture at the core of DeepMACT can be easily adopted to other cancer models (Bhatia et al., 2019; Falk et al., 2019; Wang N., 2018). Indeed, after learning to detect the characteristic shape and appearance of micrometastases against the background signal, DeepMACT successfully analyzed 3 additional tumor models we used here without further training: MCF-7 estrogen receptor positive breast cancer model, H2030-BrM3 lung cancer model, and R254 syngeneic pancreatic cancer model. Therefore, it would require little effort to apply our algorithms to different types of tumor models. Also, adapting the algorithm to applications in which, for instance, shape and size differ substantially from MDA-MB-231 metastases, would not require training from scratch. Adjusting design parameters such as the size of subvolumes (see the STAR Methods and the detailed handbook (Methods S1) for DeepMACT that we provide) allows the straightforward adaptation of the algorithm to new data with different SNR, metastasis sizes, or spatial resolution of the scan. Furthermore, building upon our pre-trained algorithms, which are freely available online, allows retraining the algorithm with substantially less training data.

To ensure high computational efficiency, our approach solves the three-dimensional task of detecting and segmenting the metastases by exploiting two-dimensional representations of the data. This is important because 2D maximum-intensity-projections increase SNR when there is little background noise owing to the high specificity of the labels in vDISCO clearing. 3D convolutions are exponentially more expensive in terms of model complexity (number of parameters) as well as computational load than 2D convolutions, thus requiring more powerful computing resources and more data annotated in 3D to train the algorithm. Importantly, the increased number of parameters is detrimental to model performance, unless the amount of training data is further increased. In this study, the 3D CNNs we tested failed to reach a high level of detection performance due to limited availability of training data, a common constraint in practice given the cost associated with annotating data (especially in 3D). In addition, the more efficient nature of our approach allows training the

entire algorithm on a standard workstation with an ordinary GPU within a few hours; applying the trained algorithm to a new dataset takes in the order of 15 minutes, highlighting the scalability and cost-efficiency of our pipeline. Thus, the DeepMACT architecture is designed to enable widespread adoption of our approach by minimizing data annotation and computing requirements while allowing for easy adaptation to other experimental setups (such as different tumor models).

DeepMACT detection of micrometastases and tumor-targeting drugs

Methods such as magnetic resonance imaging (MRI), computed tomography (CT), and bioluminescence imaging have been widely used to visualize cancer growth at the primary site and distant body regions (Condeelis and Weissleder, 2010; Massoud and Gambhir, 2003, 2007; Ntziachristos, 2010; Pichler et al., 2008; Timpson et al., 2011). While these methods provide crucial longitudinal information on the size of the primary tumor and large metastases, they typically can only resolve structures larger than 75 μm , hence they do not have the resolution to detect smaller micrometastases consisting of fewer cells.

Unbiased high-throughput mapping of tumor micrometastases in full body scans of rodents can be a valuable tool to uncover the biology behind the dissemination of tumor cells. We show here that DeepMACT is a powerful pipeline for detecting and mapping cancer metastases in mouse bodies, allowing identification of the precise locations of even the smallest disseminated tumors. Complex analysis, e.g., of the size, location, and density of micrometastases could be performed in a short time throughout the body, without dissecting any pre-defined region. In addition to detecting the micrometastases in the selected organs such as the lungs and liver, we also identified numerous micrometastases throughout the torso. For the MDA-MB-231 breast cancer line we could show that metastases are present in deep tissues such as the brain or distant locations such as hindlimb bones as early as 2 days after cardiac injection. DeepMACT also allowed us to assess important differences between distinct cancer models in terms of overall metastatic propensity and organotropism. For example, as expected from previous reports (Nguyen et al., 2009) the H2030-BrM3 lung cancer line developed the highest fraction of brain metastases among all cancer models tested and also metastasized to bones. However, DeepMACT allowed us to comprehensively characterize the distribution of metastases throughout the body, revealing for instance that this model has a high propensity to metastasize to lungs, but produces much fewer liver metastases than any other model tested. The pancreatic tumor, on the other hand, metastasized neither to the brain nor to the kidneys but disseminated for instance into the peritoneum – a pattern observed commonly in human patients as well as in several different mouse models of the disease (Hingorani et al., 2003; Lenk et al., 2017; Ryan et al., 2014; Schonhuber et al., 2014). Overall, we find that our results agree with the existing literature, but while previous studies were structurally limited to selective analysis of micrometastases in small tissue samples, the results shown here represent the first systematic, unbiased, and comprehensive full-body scale screening for micrometastases for these cancer models.

While precise assessment of therapeutic antibody biodistribution is critical for evaluating its specificity and utility for tumor treatment, there has been no method so far that can provide such information down to the level of individual micrometastases on full body scale scans.

Here, we applied DeepMACT to study not only the distribution of single metastases but also of a therapeutic monoclonal antibody. We demonstrated that the “on” and “off” targeting of antibody drugs throughout the body can readily be assessed by DeepMACT. For example, we observed that not all micrometastases in the lungs were targeted by the anti-CA12 therapeutic antibody 6A10. Understanding why antibody-based therapeutics do not target all metastases would be important for developing more effective treatments. Towards this goal, we studied the potential mechanisms that could contribute to the lack of targeting. Vascular staining demonstrated that blood vessels were present in the immediate vicinity of all examined metastases in the lungs, suggesting that insufficient vascularization is unlikely to be a common cause for the failure of antibody drug targeting. Interestingly, DeepMACT analysis found that micrometastases located in close proximity are more likely to be targeted. This suggests that the local microenvironment within metastatic niches plays an important role in determining the efficiency of antibody targeting, e.g. by altering antibody penetration, binding affinity and clearance. Furthermore, heterogeneity of antigen expression on the surface of tumor cells and internalization and degradation of antigen/antibody complexes might also affect therapeutic antibody targeting efficiency. While our findings are based on one therapeutic antibody, they nevertheless highlight a potential use case for applying the DeepMACT pipeline in pre-clinical studies aimed at understanding and improving the specificity and efficacy of tumor treatments.

In conclusion, DeepMACT is a powerful technology combining unbiased full body scale imaging with automated analysis. It enables visualization, quantification, and analysis of tumor micrometastases and antibody-based therapies in mice with high resolution and an accuracy equivalent to that of human experts but speeding up the workflow by orders of magnitude compared to traditional methods. Because this technology is time- and cost-efficient, scalable, and easily adoptable, it can be used to study metastasis and optimize antibody-based drug targeting in diverse tumor models.

STAR METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Ali Ertürk (erturk@helmholtz-muenchen.de). The lab protocol as well as the algorithms and data for the DeepMACT pipeline are freely available and have been deposited to <http://discotechnologies.org/DeepMACT/>.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Spontaneous breast cancer metastasis model—Female NSG (NOD/SCID/IL2 receptor *gamma* chain knockout) mice were obtained from Jackson Laboratory and housed at the animal facility of the Helmholtz Center Munich and the Institute of Stroke and Dementia research Munich. All animal experiments were conducted according to institutional guidelines of the Ludwig Maximilian University of Munich and Helmholtz Center Munich after approval of the Ethical Review Board of the Government of Upper Bavaria (Regierung von Oberbayern, Munich, Germany). MDA-MB-231 breast cancer cells transduced with a lentivirus expressing mCherry and enhanced firefly luciferase (Vick et al.,

2015) were counted, filtered through a 100 μm filter and resuspended in RPMI 1640 medium (Gibco, 11875093). 2×10^6 cells per mouse were injected transdermally in a volume of 50 μl into the 4th left mammary fat pad of 3-4 months old female NSG mice. For the intracardial injection model used for the time-course study, 1×10^5 cells per mouse were injected in a volume of 100 μl PBS into the left ventricle of female NSG mice as described before (Campbell et al., 2012). In brief, the mice were anesthetized using an isoflurane vaporizer and placed ventral side up on a heating pad to keep the body temperature around 37 °C. Then the chest area was shaved and cleaned by 70% ethanol. The midway point between top of xiphoid process and the sternal notch was marked and the injection point slightly on the left (anatomical) side of sternum was defined. Finally, the injection was conducted by a 0.5 ml insulin syringe (B.Braun, Omnican 50, U100 Insulin 0.5 ml / 50 I.U., 30G x 1/2", 9151125) with a bright red blood pulse back in the syringe as a successful sign. After injection, gentle pressure around the injection site was applied to prevent inner bleeding and the mice were kept in a recovery chamber (Mediheat, 34-0516) at 30 °C until they fully recovered from the anesthesia.

Tumor growth was monitored by bioluminescence measurement (photons/second) of the full body using an IVIS Lumina II Imaging System (Caliper Life Sciences) as described (Vick et al., 2015). Briefly, mice were anesthetized with isoflurane, fixed in the imaging chamber and imaged 15 minutes after Luciferin injection (150 mg/kg; i.p.). Bioluminescence signal was quantified using the Living Image software 4.2 (Caliper Life Sciences).

Estrogen positive breast cancer model and brain metastatic lung cancer model

—Animal experiments were approved by the veterinary department of the regional council in Darmstadt, Hesse, Germany. Xenograft transplantations were performed in athymic 5-6 week old female NMRI nu/nu mice (Janvier Labs) that were kept in a specific pathogen-free animal facility according to the institutional guidelines of the University of Giessen and University of Frankfurt. Intracardial injections were performed as described before (Sevenich et al., 2014). In brief, prior to the tumor cell injection, subconfluent cells, lentivirally transduced with a construct expressing mCherry and enhanced firefly luciferase, were harvested and kept on ice in sterile PBS until the inoculation. Mice were anesthetized with 100 mg/kg ketamine and 10 mg/kg xylazine and the depth of anesthesia was confirmed by the absence of toe reflexes. The chest was sterilized using 70% ethanol and 1×10^5 MCF-7 cells, or 5×10^4 H2030-BrM3 in a total volume of 100 μl PBS were injected stepwise into the left cardiac ventricle using a 26G needle. Success of the injections was monitored by pulsating reflux of arterial blood into the syringe. Metastatic growth was monitored by in vivo bioluminescence using an IVIS Lumina II Imaging System 5 minutes after an intraperitoneal injection of 150 mg/kg luciferin.

Pancreatic cancer model—Immunocompetent (wildtype C57BL/6) mice were housed at the animal facility of the Klinikum rechts der Isar of TUM. All animal studies were conducted in compliance with European guidelines for the care and use of laboratory animals and were approved by the Institutional Animal Care and Use Committees (IACUC) of Technische Universität München, Regierung von Oberbayern and UK Home Office. The low passaged primary pancreatic cancer cell line R254, derived from a genetically

engineered KPC mouse (LSL-KrasG12D/+;LSL-Trp53R172H/+;Ptf1aCre/+) on a C57BL/6 background as described previously (Eser et al., 2013; von Burstin et al., 2009; von Burstin et al., 2008), was transduced with lentiviral particles expressing EGFP and Firefly Luciferase. 2.5×10^3 cells per mouse in 20 μ l Dulbecco's modified Eagle medium were implanted orthotopically into the pancreas of 2-3 months old male mice. Tumor growth was monitored by bioluminescence measurement of the entire body. In brief, mice were anesthetized with midazolam/medetomidine/fentanyl, injected with D-luciferin (Synchem, Kassel, Germany) at 150 mg/kg intraperitoneally (IP) and imaged after 10 minutes using a cooled back-thinned, charge-coupled device camera (OrcaII ER, Hamamatsu, Herrsching, Germany) equipped with an image intensifier for 10-120 seconds; bin size, 2; gain, 700. A photographic grayscale image was taken, and the bioluminescent signals were displayed in pseudocolors and projected on the grayscale image using SimplePCI software (Hamamatsu).

Injection of therapeutic antibody—9 weeks after tumor cell injections, one mouse was randomly chosen for different experimental procedures including injection of a human carbonic anhydrase (CA) XII-specific antibody (6A10) (Battke et al., 2011). In brief, 20 μ g of 6A10 antibody conjugated with Alexa-568 was injected into the tail vein of the mouse. 48 hours later, the mouse was perfused for vDISCO pipeline including enhancing endogenous mCherry fluorescence and clearing as described in the METHOD DETAILS section.

METHOD DETAILS

Perfusion and tissue preparation—The mice were deeply anesthetized using a combination of midazolam, medetomidine and fentanyl (MMF) (1ml/100g of body mass for mice; i.p.). Then, the chest cavity of the animals were opened for the standard intracardial perfusion with heparinized 0.01 M PBS (10-25 U/ml of Heparin as final concentration, Ratiopharm, N68542.03; 100-125 mmHg pressure using a Leica Perfusion One system) for 5-10 minutes at room temperature until the blood was washed out, followed by 4% paraformaldehyde (PFA) in 0.01 M PBS (pH 7.4) (Morphisto, 11762.01000) for 10-20 minutes. The skin was carefully removed and the mouse bodies were post-fixed in 4% PFA for 1 day at 4 °C and transferred to 0.01 M PBS.

Tissue clearing and staining

uDISCO mouse body clearing: The uDISCO protocol to clear bodies of mice was described previously (Pan et al., 2016). In brief, a transcardial-circulatory system was established involving a peristaltic pump (ISMATEC, REGLO Digital MS-4/8 ISM 834; reference tubing, SC0266). Two channels from the pump were set for the circulation through the heart into the vasculature: the first channel pumped the clearing solution into the mouse body and the second channel collected the solution exiting the mouse body and recirculated the solution back to the original bottle. For the outflow tubing of the first channel, which injected the solution into the heart, the tip of a syringe (cut from a 1 ml syringe-Braun, 9166017V) was used to connect the perfusion needle (Leica, 39471024) to the tubing. Meanwhile, the inflow tubing of the second channel, which recirculated the clearing solutions, was fixed to the glass chamber containing the mouse body. The amount of solutions for circulation depended on the capacity of the clearing glass chamber. For

example, if the maximum volume of glass chamber is 400 ml, 300 ml of volume of solution was used for circulation.

All clearing steps were performed in a fume hood. Firstly, the mouse body was put in a glass chamber and the perfusion needle was inserted into the heart through the same hole that was used for PFA perfusion. Then, after covering the chamber with aluminum foil the transcardial circulation was started with a pressure of 230 mmHg (60 rpm on the ISMATEC pump). The mouse bodies were perfused for 6 hours with the following gradient of *tert*-butanol (Carl Roth, AE16.3): 30 Vol%, 50 Vol%, 70 Vol%, 90 Vol% (in distilled water), 100 Vol% twice, and finally with the refractive index matching solution BABB-D4 containing 4 parts BABB (benzyl alcohol + benzyl benzoate 1:2, Sigma, 24122 and W213802), 1 part diphenyl ether (DPE) (Alfa Aesar, A15791) and 0.4% Vol vitamin E (DL-alpha-tocopherol, Alfa Aesar, A17039), for at least 6 hours until achieving transparency of the bodies. As the melting point of *tert*-butanol is between 23 to 26 °C, a heating mat set at 35-40 °C was used for the two rounds of 100% *tert*-butanol circulation to prevent the solution from solidifying.

vDISCO mouse body immunostaining and clearing: The detailed protocol of vDISCO was described previously (Cai et al., 2018). The following nanobodies and dyes were used for mouse body immunostaining: Atto647N conjugated anti-RFP/mCherry signal-enhancing nanobodies (Chromotek, rba647n-100), Atto594 conjugated anti-RFP/mCherry signal-enhancing nanobodies (Chromotek, rba594-100), Atto647N conjugated anti-GFP signal-enhancing nanobodies (Chromotek, gba647n-100), Hoechst 33342 (Thermo Fisher Scientific, 21492H), Propidium iodide (PI, Sigma, P4864). Please note that different batch of nanoboosters coming from different companies can have different penetration and stability performances. Please check <http://www.discotechnologies.org/vDISCO/> for updates on which nanoboosters to use.

To perform the mouse body immunolabeling, a simplified transcardial-circulatory system using the same type of peristaltic pump was established (ISMATEC, REGLO Digital MS-4/8 ISM 834; reference tubing, SC0266). In short, one reference tubing was connected by two connectors (Omnilab, 5434482) from both ends and extended by additional PVC tubing (Omnilab, 5437920). The head part from a 1 ml syringe (Braun, 9166017V) was cut and inserted into the outflow PVC tubing as a connector for the perfusion needle (Leica, 39471024). Next, a PBS perfused and PFA fixed mouse body was placed into a 250-ml glass chamber (Omnilab, 5163279) and 200 ml of 0.01 M PBS was filled immediately into the chamber. Note that the sample will be kept in the same chamber through the entire immunolabeling and clearing process, it should be always embedded in the respective solutions till the moment of imaging. Then, the inflow tubing of the transcardial-circulatory system was fixed underneath the surface of PBS in the glass chamber using adhesive tape and the pumping circulation was started until the air bubbles were completely removed from the tubing system. The mouse body decolorization, decalcification and immunolabeling steps were conducted subsequently after inserting and fixing the perfusion needle into the heart of the sample through the same pinhole made during sample preparation.

In general, the animals were firstly perfused with decolorization solution for 2-3 days at room temperature to remove remaining heme and blood before immunostaining. The

decolorization solution which is a 1:4 dilution of CUBIC reagent 1 (Susaki et al., 2014) in 0.01 M PBS was refreshed twice during the decolorization step. CUBIC reagent 1 was prepared as a mixture of 25 wt% N,N,N,N'-tetrakis (2-hydroxypropyl) ethylenediamine (Sigma-Aldrich, 122262), 25 wt% urea (Carl Roth, 3941.3) and 15 wt% Triton X-100 in 0.01 M PBS, as described in the original publication. Next, after washing with 0.01 M PBS for 3 hours 3 times, the samples were perfused with the decalcification solution (10 wt/vol% EDTA in 0.01 M PBS, pH to 8–9, Carl Roth, 1702922685) for 2 days and for 1 more day with permeabilization solution containing 0.5% Triton X-100, 1.5% goat serum (Gibco, 16210072), 0.5 mM of Methyl-beta-cyclodextrin (Sigma, 332615), 0.2% trans-1-Acetyl-4-hydroxy-L-proline (Sigma, 441562), 0.05% sodium azide (Sigma, 71290) in 0.01 M PBS. Before the immunostaining step, additional 0.22 µm syringe filters (Sartorius 16532) were attached to the inflow tubing to prevent the potential accumulation of nanobody aggregates and high pressure pumping at 160–230 mmHg (45–60 r.p.m.) was maintained through the entire labeling process. Then the immunostaining solution was prepared as a mixture of permeabilization solution and 35 µl of nanobody (stock concentration 0.5 – 1 mg/ml), 10 µg/ml Hoechst or 300 µl of propidium iodide (stock concentration 1mg/ml) and filtered by the same 0.22 µm syringe filter before use. Subsequently the animals were perfused for 5-6 days with 200 ml of immunostaining solution at room temperature and further passively labeled in the same staining solution with extra 5 µl of signal-enhancing nanobody with gentle shaking for 2 days at 37 °C or at room temperature. Then the mice were connected back to the circulation system and perfused with washing solution (1.5% goat serum, 0.5% Triton X-100, 0.05% of sodium azide in 0.01 M PBS) for 12 hours twice at room temperature and at the end with 0.01 M PBS for 3 hours 3 times at room temperature.

After completing the mouse body immunolabeling step, the mouse bodies were passively cleared using 3DISCO (Ertürk et al., 2012) at room temperature with gentle shaking (IKA, 2D digital) under a fume hood. For dehydration, the mouse bodies were incubated in 200 ml of the gradient tetrahydrofuran (THF, Sigma, 186562) in distilled water (6-12 hours for each step): 50 Vol% THF, 70 Vol% THF, 80 Vol% THF, 100 Vol% THF and again 100 Vol% THF; then the mouse bodies were incubated for 3 hours in dichloromethane (Sigma, 270997), and finally in BABB until the tissue were rendered completely transparent. During all clearing steps, the glass chamber was sealed with parafilm and covered by aluminum foil to prevent extra solution evaporation and fluorescence quenching. For details, see also the step-by-step handbook (Methods S1).

Rehydration and immunostaining of cleared tissue: Anti-Firefly Luciferase (dilution 1:2000, Abcam, ab21176) and AlexaFluor 488 goat antirabbit IgG (H+L) (dilution 1:400, Life Technologies, A11034) were used to verify the specificity of anti-RFP/mCherry signal-enhancing nanobody labeling. After identification of metastases in the lungs of vDISCO-processed mice, lung tissue was dissected and rehydrated by applying the reverse gradient of tert-butanol used for uDISCO clearing, as follows (6 hours each at 37°C with gentle shaking): 100 Vol% twice, 90 Vol%, 70 Vol%, 50 Vol%, 30 Vol% and 0.01 M PBS twice at room temperature. Rehydrated samples were cut into 1 mm sections using a vibratome (Leica, VT1200S) and were incubated in 0.01 M PBS containing 0.2% gelatin (Sigma, G2500), 0.5% Triton X-100, 0.05% sodium azide and 5% normal goat serum for 1 day at 37

°C. The sections were then incubated with the primary antibodies diluted in the same solution overnight at 37 °C, washed twice in PBS at room temperature, incubated with secondary antibodies diluted in the same solution for 4 hours at 37 °C and at the end washed in PBS three times at room temperature (related to Figure S2B).

Lectin vasculature labeling in lung tissue: The bodies of mice were perfused and collected as described above. After checking with epifluorescence stereomicroscopy (Zeiss AxioZoom EMS3/SyCoP3), the lung lobes with multiple metastases were dissected and sliced into 20 µm thick tissue sections by using a cryostat (Leica, CM3050S). The lung sections were washed 2 times with 0.01 M PBS and then incubated in Alexa 488 conjugated Lectin (4 µg/ml, invitrogen, W11261) at 4 °C overnight. The sections were then stained with Hoechst 33342 (10 µg/ml, Thermo Fisher Scientific, 21492H) for 5 minutes at room temperature to visualize the nucleus. After washing 2 times with PBS, the slides were mounted with fluorescent mounting medium (Dako, 10097416) and were ready for confocal microscopy (related to Figure 7A).

mCherry signal enhancement in lung tissue: 20 µm thick lung tissue sections were washed with 0.01 M PBS 2 times before starting the enhancement process. One-hour incubation in blocking solution containing 1% Bovine Serum Albumin (Sigma, A7906), 2% goat serum (Gibco, 16210-072), 0.1% Triton X-100 and 0.05% Tween 20 (Bio-Rad, 161-0781) in PBS, was performed at room temperature. Then the staining solution was prepared in 1% Bovine Serum Albumin and 0.5% Triton X-100 in PBS. Atto647N conjugated anti-RFP/mCherry signal-enhancing nanobodies was diluted 1:500 in the staining solution and the lung sections were incubated overnight at 4°C. After the treatment with nanobodies, the lung sections were washed 3 times with PBS for 5 minutes with gentle shaking. After nuclear staining by Hoechst 33342 (10 µg/ml) and post wash with PBS as described before, the slides were mounted with fluorescence mounting medium and were ready for confocal microscopy (related to Figure S2C).

Image acquisition

Epifluorescence stereomicroscopy imaging: Cleared mouse bodies were fixed in the original clearing chamber and were imaged with Zeiss AxioZoom EMS3/SyCoP3 fluorescence stereomicroscope using a 1x long working distance air objective lens (Plan Z 1x, 0.25 NA, Working distance (WD) = 56 mm). The magnification was set as 7x and imaging areas were selected manually to cover the entire mouse bodies. The images were taken with GFP, RFP and Cy5 filters and files were exported as RGB images in JPEG format. For high resolution imaging of individual metastasis, higher zoom factor can be applied up to 112x.

Light-sheet microscopy imaging: Single plane illumination (light-sheet) image stacks were acquired using an Ultramicroscope II (LaVision BioTec), allowing an axial resolution of 4 µm. For low magnification full-body screening of tumor and antibody signals we used a 1x Olympus air objective (Olympus MV PLAPO 1x/0.25 NA [WD = 65 mm]) coupled to an Olympus MVX10 zoom body, which provides zoom-out and -in ranging from 0.63x up to 6.3x. Using 1x objective, we imaged a field of view of 2 x 2.5 cm, covering the entire width

of the mouse body. Tile scans with 60% overlap along the longitudinal y-axis of the mouse body were obtained from ventral and dorsal surfaces up to 13 mm in depth, covering the entire volume of the body using a z-step of 10 μm . Exposure time was 150 ms, laser power was 3 to 4 mW (70% to 95% of the power level) and the light-sheet width was kept at maximum. Alternatively, the mouse bodies were scanned with a dipping 1.1x objective (LaVision BioTec MI PLAN 1.1x/0.1 NA [WD = 17 mm]) coupled with an Olympus revolving zoom body unit (U-TVCAC). In brief, 3x8 tile scans with 25% overlap were obtained from both sides to 11 mm in depth, covering the entire volume of the body using a z-step of 6 μm . Light-sheet width was set at 80% and exposure time was 80 ms. The laser power was adjusted depending on the intensity of the fluorescent signal to avoid reaching saturation. For the mouse displayed in Figure 6H, the lower part of the jaw was removed to fit the mouse head into imaging chamber; this was not the case for any other mouse presented in this study.

After low magnification imaging of the full scale mouse body, organs (including lungs, liver, kidneys, brain, spleen, intestines and bones) were imaged individually using high magnification objectives (Olympus XLFLUOR 4x corrected/0.28 NA [WD = 10 mm], LaVision BioTec MI PLAN 12x/0.53 NA [WD = 10 mm] and Zeiss 20x Cfr Plan-Neofluar/0.1 NA [WD = 4 mm]) coupled to an Olympus revolving zoom body unit (U-TVCAC) kept at 1x. High magnification tile scans were acquired using 20% overlap and the light-sheet width was reduced to obtain maximum illumination in the field of view. For the data used for the comparison of signal profile plots of lung metastases taken in red and far-red channels and for the analysis of endogenous fluorescence signal depletion after the uDISCO protocol, we used the same MVX10 zoom body, coupled this time with a 2x objective (Olympus MVPLAPO2XC/0.5 NA [WD = 6 mm]) at zoom body magnification 6.3x and 2.5x respectively.

Confocal microscopy imaging: For imaging the thick cleared specimens such as dissected tissues, pieces of organs or whole organs were placed on 35 mm glass bottom petri dishes (MatTek, P35G-0-14-C), then the samples were covered with one or two drops of the refractive index matching solution such as BABB or BABB-D4. Sealing of this mounting chamber was not necessary. The samples were imaged with an inverted laser-scanning confocal microscopy system (Zeiss, LSM 880) using a 40x oil immersion lens (Zeiss, EC Plan-Neofluar 40x/1.30 Oil DIC M27) and a 25x water immersion long-working distance objective lens (Leica, NA 0.95, WD = 2.5 mm), the latter one was mounted on a custom mounting thread. The z-step size was 1-2.5 μm . For imaging the lung tissue sections with Lectin staining, with nanobodies or with Anti-Firefly Luciferase staining, the slides were imaged with the same inverted laser-scanning confocal microscopy system with a z-step size of 2 μm .

Reconstructions of full-body scans: For epifluorescence microscopy reconstructions (2D montage of entire mouse), the collected images were stitched semi-automatically using Adobe Photoshop photomerge function. After saving the stitched images as TIFF or JPEG files, the signal from individual imaging channels can be extracted by using Split Channels function in ImageJ.

For light-sheet microscopy reconstructions (3D montage of entire mouse), the image stacks were acquired and saved by ImSpector (LaVision BioTec GmbH) as 16-bit grayscale TIFF images for each channel separately. The stacks were first stitched with Fiji (ImageJ2) and fused together with Vision4D (Arivis AG). For details, see also the step-by-step handbook (Methods S1). Further image processing was done mostly in Fiji: first, the autofluorescence channel (imaged in 488 excitation) was equalized for a general outline of the mouse body. The organs were segmented manually by defining the regions of interests (ROIs). Data visualization was done with Amira (FEI Visualization Sciences Group), Imaris (Bitplane AG), Vision4D in both volumetric and maximum intensity projection color mapping.

QUANTIFICATION AND STATISTICAL ANALYSIS

General data processing—All data processing after image volume reconstruction was performed in Python using custom scripts based on publicly available standard packages comprising SciPy (Jones et al., 2001), Seaborn (Waskom, 2012), and Pandas (McKinney, 2008). Deep Learning models were built using the PyTorch framework (Paszke, 2016). Since a single full body scan is in the order of several terabytes due to its high resolution (the data used for training had a voxel size of $(10\mu\text{m})^3$), the volume was divided into 1176 subvolumes of $(350\text{px})^3$ (or $(3.5\text{mm})^3$) to enable efficient processing. Subvolumes were overlapping by 50px to ensure any given metastasis is fully captured by at least one subvolume to avoid artefacts of divided metastases at subvolume interfaces. Please note that the size and overlap of subvolumes are design choices that allow easy adaptation to different data sets, e.g. with different SNR, metastasis sizes, or spatial resolution of the scan. Final analyses were conducted on the re-assembled full volume whereby reconcatenation ruled out any double-counting at previously overlapping subvolumes.

Data annotation by human experts—To provide ground truth in the form of a commonly agreed upon reference annotation for training, as well as for evaluation of the algorithms developed, full body scans of two mice (with MDA-MB-231 tumor cells transplanted in the mammary fat pad) were manually annotated by a group of human experts. This manual process was augmented with a set of tools to reduce the total workload from an estimated total duration of several months down to 150 person-hours net annotation time.

Automatic annotation with fixed filter kernel: To avoid starting from scratch to annotate two volumes of several thousand z-slices, an automatic detection and segmentation method was applied to provide a basis for manual correction. Due to the insufficient performance of established methods (in this case: the 3D Object Detector for ImageJ (Bolte and Cordelières, 2006)), we developed a custom-made filter based detector tailored to the specifics of this dataset. In brief, we handcrafted a spatial filter kernel optimized to detect the most common metastases and applied it with 3D convolutions to the dataset; subsequent binarization and connected-component analysis yielded *seed points* collocated with metastases. This allowed for further analyses of the immediate local neighborhood of these candidate regions; a local 3D segmentation was derived by selective region growing around these seed points based on the local signal intensity distribution up to a mean *foreground signal* limited to 4 standard deviations above the mean signal in the local surrounding. Finally, obvious false positives

were filtered out. Together, this approach generated a first proposal for the data annotation that at least captured the most obvious metastases while producing an acceptable rate of false positives. As shown in the results section, the quality of this proposal was about twice as good as compared to the 3D Object Detector in ImageJ (35% instead of 18% in F1-score). Importantly, further fine-tuning of filters and parameters and any additional automated pre- or post-processing did not improve the results, indicating that a F1-score of 35% may be close to the performance limit of such approaches with fixed filter kernels and fixed decision rules for such kind of data.

Manual annotation correction by human experts: This first proposal served as a basis for human annotation. In general, three kinds of manual correction were needed to derive a good annotation: removal of false positives, addition of false negatives (previously missed metastases) and adjustment of the 3D segmentation of each metastasis. To avoid the need to perform this task individually for each of the 350 layers of a (350px)³ data subvolume, a custom tool with an interactive graphical user interface was developed. Based on maximum intensity projections along each dimension, the tool allowed to review, adjust, add, and remove each potential metastasis in the subvolume with a few mouse clicks, drastically speeding up the annotation process from hours to minutes per subvolume. Different perspectives (X, Y, Z) and viewing modes (e.g., projections, orthogonal slices, adjusted contrasts, 3D renderings) for each individual metastasis allowed the annotator to take maximally informed decisions even in less obvious cases.

Refinement of annotation to ground truth: A small fraction (3%) of the entire data set was labeled several times by the annotators without their awareness to assess human labeling consistency. Since the difference in annotation for a given subvolume between two trials of a single annotator was about as big as between two independent annotators and quite substantial (the agreement between two trials of the same annotator or between annotators only reached an F1-score of 80-85%) we decided to invest additional time to refine the entire data set. First, all experts (3 graduate students with extensive experience in the field of imaging and tumor biology) jointly discussed examples of annotation differences to build a common understanding. Annotations of subvolumes with the biggest discrepancies were again reviewed and refined. Furthermore, this analysis revealed that the most prevalent source of annotation error was overlooked metastases (false negatives). Here, around 29% of metastases were missed in the human annotation, in line with previous studies (Ehteshami Bejnordi et al., 2017; Vestjens et al., 2012). To effectively identify all missed metastases in the entire data set, our deep learning algorithm (see next section) was trained on the status quo of the annotations and applied to the data set with high sensitivity. This yielded a long list of potential candidates. With the help of another custom-built, interactive graphical user interface, all potential candidates were manually reviewed by the annotators and either discarded or manually adjusted and added to the segmentation (this is the manual refinement step referred to in Figure 3D–E). A small set of potential metastases, for which human annotators could not take a conclusive decision even after joint discussion, was recorded separately, but not added to the segmentation. These laborious steps ensured the generation of a high-quality *ground truth* for training the algorithm and, importantly, for evaluating its performance in comparison to a single human annotator. Here, this selectively iterative

approach of refining annotations based on the input of several human experts was chosen due to the substantial amount of manual work involved with reviewing our high-resolution scans. Since a full review of one person takes about a month of full-time work, repeating this process several times would be desirable but too costly. In applications where several, independent full annotations are available, advanced mathematical frameworks for refining decisions from different experts to a single decision can be applied in order to avoid a bias towards individual decisions (Mavandadi et al., 2012a; Mavandadi et al., 2012b).

Deep learning for metastasis detection

Implementation of DeepMACT model architecture: Inspired by the established U-net architecture (Ronneberger et al., 2015), we designed a deep learning approach that is depicted in Figure 3A and briefly described in the results section. The architecture of the CNN at its core (Figure 3C) is characterized by an encoding downward path and a decoding upward path comprising a total of 7 levels, in which each level also has a lateral skip-connection that bypasses the deeper levels and feeds the output of the encoding unit directly to the corresponding decoding unit. Each encoding unit increases the number of feature channels per pixel with the help of two kernel-based convolutions (kernel size: 3; padding: 1; dilation: 1; stride: 1) followed by batch normalization and a rectifying linear unit (ReLU). While the first convolutional step increases the number of feature channels, this number stays constant for the second convolutional step. Before being passed on to the next encoding stage, the spatial resolution is halved using maxpooling (kernel size: 2, stride: 2). Decoding units take two inputs: the output from the previous layer is spatially upsampled by a factor of two (bilinearly) and concatenated along the feature dimension with the output of the corresponding encoding unit, bypassing the deeper levels. A first convolutional step (same parameters as before) decreases the number of feature channels, which is again kept constant in the two subsequent convolutions. The 24-feature channel output of the last decoder is mapped to logits in the 2D space with a convolutional step without padding, batch normalization, or a rectifying linear unit.

Implementations of customized 3D U-net: To compare the DeepMACT approach with CNNs that operate on volumetric data with 3D convolutions, we implemented several customized versions of 3D U-nets (Çiçek et al., 2016). In this alternative approach, the volumetric data is directly fed to the network (instead of projections) to predict a 3D probability volume (without reconstruction from 2D predictions). While following the overall architectural approach of the DeepMACT implementation, we replaced 2D with 3D convolutional operators. We implemented a variety of derivations by changing the total number of levels of en- and decoding units and thus, the maximum number of feature channels, which both drive model complexity in terms of number of parameters. The best-performing implementation consisted of 3 levels of en- and decoding units with a maximum number of feature channels of 48; the corresponding performance values are reported in Figure 3D–E (leaner or more complex models yielded comparable or lower performance). All other parameters and procedures (e.g., for training and testing) are identical to the DeepMACT implementation.

Training, validation, and test sets: Following established standards, model training and evaluation was based on k-fold cross-validation (k=5). Thus, the annotated data set was split into mutually exclusive sets for training and validation (80%) and for testing (20%). This process was repeated k times, yielding a total of 5 mutually exclusive test sets that are collectively exhaustive. The network weights and all design choices and hyperparameters (such as batch size, learning rate, etc) were optimized solely with the training and validation set to avoid overfitting on the specifics of the test set. The data set was confined to subvolumes within the torso of the mouse body as subvolumes containing near-zero values outside the body contain no useful information to train or test on. In contrast to all metastases in the entire body, the tumor tissues of the primary tumor and the auxiliary lymph node are several orders of magnitude larger (i.e., they follow very different statistics than all micrometastases) and were thus excluded. The signal from one subvolume was corrupted by a dirt particle and thus also excluded. In total, these exclusions made up less than 1% of the total scan volume. The split between the three subsets (training, validation, testing) of the data was done on a subvolume level (from which the three projections are created afterwards) to avoid information leak between different projections from the same subvolume.

Training procedure: The model training was conducted in two steps. First, a large number of models spanning a broad set of different hyperparameters were trained for 10 epochs using another (nested) k-fold cross-validation (k=5) within the training and validation set. Second, the model with the best-performing set of hyperparameters (presented here) was trained for the remaining epochs. Thus, any hyperparameter choice was made without looking at the performance on the test set. The model was trained for 40 epochs of the entire training data set, using random vertical and horizontal flips of the data to augment its variance (further training epochs did not improve the predictive power). We used a batch size B of 4 but found that other batch sizes work similarly well. Each input was normalized by its local subvolume peak value, which was found to work better than normalization to the global volume peak value or non-linear normalizations. To calculate the gradients for network weight optimization (i.e., to train the model), we used weighted binary cross entropy as a loss function for a given prediction \hat{Y} compared to the ground truth Y , giving more weight w for foreground (FG) pixels p versus background pixels (BG) to account for the *class imbalance* (i.e., that metastases are very sparsely distributed in space):

$$w_{BCE\hat{Y}, Y} = (-1) \sum_b^B \sum_p^P (w_{FG} y_{b,p} \log(y_{b,p} + \epsilon) + w_{BG} (1 - y_{b,p}) \log(1 - y_{b,p} + \epsilon))$$

A small numerical offset $\epsilon = 10^{-4}$ was applied for numerical stability. We found equal weights or a slightly stronger bias to foreground to work almost equally well (here, we used $w_{FG}=2$ and $w_{BG}=0.5$), larger biases had negative effects. Additionally, we allowed the network to optimize the share of training data that contains at least some foreground by ignoring parts of training data in which no foreground is present. A share of 90% training data with at least some foreground optimized the performance on the validation set and was thus chosen. The network was trained using the Adam optimizer (Kingma and Ba, 2014); the initial learning rate was set to 10^{-4} and was gradually decreased by a factor of 10 to a

minimum of 10^{-7} every time the loss function reached a plateau for more than 2 epochs. A single training run over 40 epochs takes only around 20-30 minutes on a normal workstation equipped with a NVIDIA Titan XP GPU.

Testing procedure and inference mode: As mentioned before, we applied k-fold cross-validation. Thus, in each of the $k=5$ folds the model was tested on fresh data that was not seen by the model during training and validation. Together, all 5 sets span the entire annotated data set. As depicted in Figure 3A, the trained algorithm in inference mode was used to generate probability masks for each of the three projection perspectives (P_{XY} , P_{YZ} , P_{ZX}), in which the pixel value indicates the network's confidence that this pixel is part of a metastasis in the given sub-volume s . Building the outer product of the three probability masks allows to recombine the three judgements of the network in 3D space:

$$P_{V,s} \in [0, 1]^3 = \left(P_{XY,s} \in [0, 1]^2 \otimes P_{YZ,s} \in [0, 1]^2 \otimes P_{ZX,s} \in [0, 1]^2 \right)^{1/3}$$

This 3D recombination P_v of the 2D probability maps yields a final predicted segmentation mask after binarization. By default, the confidence threshold was set to 50%; however, changing this parameter allows to manually adjust the trade-off between sensitivity and specificity, if desired (also see Figure S4). Please note that the F1-score for evaluation is not affected by this trade-off (i.e., a better detection rate at the cost of a higher false positive rate would not artificially increase the F1-score and vice versa). Subsequent connected-component analysis converts the output to an explicit list and segmentation of predicted metastases in 3D space.

Importantly, DeepMACT did not have to be re-trained to be applied to the full body scans ($n=7$) obtained for other tumor models (intracardially injected MDA-MB-231 breast cancer cells, MCF-7 breast cancer cells, R254 pancreatic cancer cells, H2030 lung cancer cells). Thus, the metastasis distribution for these scans could be readily inferred within minutes without further training data annotation.

Performance evaluation: The same performance evaluation procedure was used for the comparison shown in Figure 3D, including the performance of a single human annotator. A standard test for detection tasks, the F1 score quantifies the accuracy of a model by combining precision (share of true positives among all positive predictions, including false positives) and recall (share of predicted positives among the sum of the true positive and false negative predictions). It is mathematically equivalent to the Sørensen–Dice coefficient (“Dice score”), which is the commonly used name for pixel-wise image segmentation. The F1 score is defined as:

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

For all comparisons of detection and segmentation performance, the ground truth (refined by several human experts as described above) is used as a reference. We quantified the performance of the proposed deep learning algorithm based on its prediction of the test set.

For a comparison, the segmentations as provided by established tools like the 3D Object Detector in ImageJ, our custom-made detector as described above, as well as the annotation as provided by a single human annotator (before joint refinement) were quantified in the same manner. Overlapping segmentations for metastases were counted as true positive predictions, non-overlapping predictions as false positives and metastases not detected by the prediction as false negatives. Predictions corresponding to the small set of cases unclear to the group of human experts (see above) were neither counted as true positive predictions nor as false negatives, i.e. they neither increased nor decreased the performance evaluation. All performance evaluations were conducted on the entirety of the test set as a whole. To quantify the inherent variance, the distribution of performance results was estimated with $n=1000$ resampled test sets (of same size) using the bootstrapping approach. The correctness of the exact three-dimensional outline of each metastasis (segmentation) was verified by assessing the volumetric overlap with the three-dimensional outline drawn by human annotators. This confirmed an overlap accuracy in 3D of 90% for the worst segmentation; importantly, 90% of all detected metastases were segmented with an accuracy of 97.5% or higher.

Analysis of individual metastases

Organ registration: For the full body scale light-sheet scans (e.g., Figure 4C,F, Figure 5) the outlines of selected organs of interest (all lung lobes, brain, both kidneys, heart, liver) were manually segmented as multi-point polygons in a stack of slices in 3D using Fiji. For each metastasis detected by our deep learning architecture we assessed whether its center of mass falls into the 3D segmentation of one of those organs using a custom Python script. Any metastasis not registered to one of these organs is referred to as located in “the rest of the torso” in this manuscript. The 3D segmentation of the lungs was also used to compute the overall lung volume to assess the tumor density in Figure 4J, which we quantified as the share of the sum of the volume of all metastases registered to an organ of the entire organ volume.

Metastasis characterization: The output of our deep learning architecture is a binary segmentation volume for all metastases. We applied connected component analysis to derive an explicit list of metastases fully characterized in 3D. Based on each metastasis 3D shape and voxel-based volume V , we computed its average diameter as

$$d_{avg} = 2\sqrt[3]{\frac{3V}{4\pi}}$$

In order to put the metastasis size into context, we computed an estimate of the number of cells per metastasis. To this end, we measured the diameter of single cells and estimated the number of cells for a given metastasis volume based on volumetric extrapolation. We confirmed the accuracy of the estimations by the number of nuclei (PI or Hoechst labeled) in $n=26$ samples. On average, the estimates were off by 3.5 cells (10.3%). For 73% of all samples checked, the estimates were off by less than 20%; the worst estimate was 35% off the actual count (39 instead of 60 cells). The estimated numbers may include other cell types present in the metastases apart from tumor cells, e.g. immune cells, vascular cells and

fibroblasts. Given that metastasis sizes in full body scans can vary by orders of magnitude (see Figure 4H) this estimation accuracy was deemed sufficient to derive insightful conclusions from cell count estimates.

The distance of each metastasis i to its nearest neighboring metastasis was measured in 3D space as the Euclidian distance between their center of masses CoM :

$$distNN_i = \min_j \sqrt{(CoM_{i,x} - CoM_{j,x})^2 + (CoM_{i,y} - CoM_{j,y})^2 + (CoM_{i,z} - CoM_{j,z})^2}$$

Drug targeting analysis: We assessed the 6A10 antibody targeting of a given metastasis by analyzing the distribution of the fluorescent signal strength within the 3D segmentation of each metastasis (ξ_m) versus the distribution in its local surrounding (250 μm around the metastasis) ξ_s . For each signal distribution, the number of voxels within the metastasis segmentation n_m or in its local surrounding n_s can be seen as the number of observations of the underlying true (but unknown) distributions. The degree of targeting was estimated by quantifying the ratio of mean signal strength within the segmentation to the mean signal strength in its surrounding (e.g., in Figure 6K). We refer to this as antibody signal ratio. A ratio larger than 1 means that the antibody signal strength within the 3D segmentation of the metastasis is higher than around it (see dashed line in Figure 6K). Whether or not a metastasis was deemed “targeted” was assessed with a version of the t -test to determine whether mean of the observed signal distribution in the metastasis ξ_m was significantly at least $\approx 50\%$ (ratio of 1.5) above the mean of the observed signal distribution in the local surrounding ξ_s . Importantly, a t -test is valid for the signals despite their highly non-normal underlying distribution as the number of observations far exceeds the requirements of the central limit theorem (i.e., while the signals are not normally distributed, the estimation of their means is normally distributed due the high number of observations). This was verified manually. However, due to a typically much larger number of observations in the local surrounding ξ_s than for the metastasis itself ξ_m , the statistical test was not performed with a *Student’s t-test* but with the *Welch’s t-test* that corrects the *degrees of freedom* for an unequal number of observations for both distributions:

$$t = \frac{\text{mean}(\xi_m) - (1 + \Delta)\text{mean}(\xi_s)}{\sqrt{\frac{\text{std}(\xi_m)^2}{n_m} + \frac{\text{std}(\xi_s)^2}{n_s}}}$$

$$DF_{adjusted} = \left(\frac{\text{std}(\xi_m)^2}{n_m} + \frac{\text{std}(\xi_s)^2}{n_s} \right)^2 / \left(\frac{\left(\frac{\text{std}(\xi_m)^2}{n_m} \right)^2}{n_m - 1} + \frac{\left(\frac{\text{std}(\xi_s)^2}{n_s} \right)^2}{n_s - 1} \right)$$

Analysis of fluorescence signal profiles: We considered the fluorescence signal profiles from each channel: excitation 470 nm, 561 nm and 647 nm. These profiles were plotted in the same z-stack and normalized as percentage over the maximum peak. To compare the

reduction of the background and the improvement of the signal over background ratio (SBR) in far-red and near far-red channels, we analyzed lung metastases expressing mCherry imaged with excitation 545/561 nm after uDISCO clearing, lung metastases labeled with anti-mCherry nanobody conjugated with Atto594 imaged with excitation 590 nm and lung metastases labeled with anti-mCherry nanobody conjugated with Atto647N imaged with excitation 640 nm after vDISCO clearing (n=3 tumors per each experimental group which consisted of 3 animals per each imaging modality). The signal profile was measured from a defined straight line covering the tumors and surrounding tissue background and all the values of the plot from a representative animal per each experimental group were shown in a representative line chart (Figure S1D). Finally, the normalized plots represented in Figure S1E were calculated by normalizing the plots of lung metastases obtained as described above over the average signal intensity of the respective surrounding background.

To compare the signal-to-background ratio (SBR) in Figure S6C,D, the samples were labeled with anti-mCherry nanobody conjugated with Atto647N and primary tumors were imaged with excitation 470 nm, 561 nm and 640 nm respectively after vDISCO clearing. Fluorescence signal intensity profiles and background normalized profiles for each channel were plotted with the same strategy as described above.

Metastasis diameter and vessel distance: Metastasis diameters were verified manually. For quantifying the distance between metastases and vessels, ten points on the border of each metastasis were randomly selected and the shortest distance from these points to the closest vessel wall were measured. The presented distance between each metastasis and nearest vessels was quantified by averaging these ten measurements. In Figure 7B, 50 metastases were quantified to generate the scatter plot; each scatter point represents one single metastasis.

DATA AND CODE AVAILABILITY

The lab protocol as well as the algorithms and data for the DeepMACT pipeline are freely available and have been deposited to <http://discotechnologies.org/DeepMACT/> For convenience, this includes a fully functional demonstration script (including data).

ADDITIONAL RESOURCES

Fully functional online demo of DeepMACT: <https://codeocean.com/capsule/8c13691f-7f9a-4af4-8522-c26f581c9e83/tree?ID=a8ba18d2bf5046b08fafa2d6a42bfd7a>

Further details on the vDISCO protocol: <http://discotechnologies.org/vDISCO/>

Registration for in-person workshops: <http://discotechnologies.org/workshop/>

Videos on tissue clearing: https://youtube.com/channel/UCAVXKhQ_ZjEdkAdFR5HwjRQ

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the Vascular Dementia Research Foundation, Synergy Excellence Cluster Munich (SyNergy), the Helmholtz-Center for Environment Health (grants to A.E. and R.Z.), Fritz Thyssen Stiftung and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (A.E.), the German Federal Ministry of Education and Research via the Software Campus initiative (O.S.), DFG (SFB 1321 P06 and S01 to D.S.), the European Research Council (ERC CoG No. 648521, to D.S.), and the Excellence Cluster Cardio-Pulmonary Institute (to T.A.). I.J. was funded by ERC Consolidator Grant 681524, Mildred Scheel Professorship by DFG, Collaborative Research Center 1243, project A05. Furthermore, NVIDIA supported this work with a Titan XP via the GPU Grant Program. We thank the Monoclonal Antibody Core Facility at the Helmholtz Center Munich for providing the 6A10 antibody, Lisa Sevenich at the Georg Speyer House for help with the i.c. metastasis model, Annette Frank and Stephanie Hoffmann for performing bioluminescence in vivo imaging for MDA-MB-231 cancer model, and Dr. W. Ouyang for constructive comments on the manuscript. M.I.T is member of Graduate School of Systemic Neurosciences (GSN), Ludwig Maximilian University of Munich.

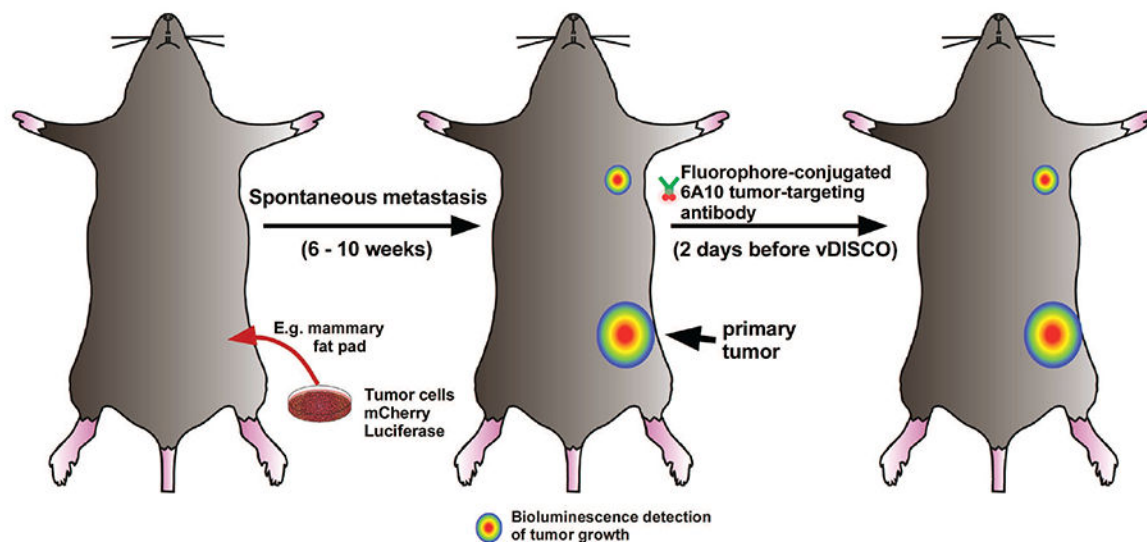
REFERENCES

- Barker N, and Clevers H (2006). Mining the Wnt pathway for cancer therapeutics. *Nat Rev Drug Discov* 5, 997–1014. [PubMed: 17139285]
- Batke C, Kremmer E, Mysliwicz J, Gondi G, Dumitru C, Brandau S, Lang S, Vullo D, Supuran C, and Zeidler R (2011). Generation and characterization of the first inhibitory antibody targeting tumour-associated carbonic anhydrase XII. *Cancer Immunol Immunother* 60, 649–658. [PubMed: 21298264]
- Bhatia S, Sinha Y, and Goel L (2019). *Lung Cancer Detection: A Deep Learning Approach* (Singapore: Springer Singapore).
- Bolte S, and Cordelieres F (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy* 224, 213–232. [PubMed: 17210054]
- Cai R, Pan C, Ghasemigharagoz A, Todorov MI, Forstera B, Zhao S, Bhatia HS, Parra-Damas A, Mrowka L, Theodorou D, et al. (2018). Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull-meninges connections. *Nat Neurosci*.
- Camacho DM, Collins KM, Powers RK, Costello JC, and Collins JJ (2018). Next-Generation Machine Learning for Biological Networks. *Cell* 173, 1581–1592. [PubMed: 29887378]
- Campbell JP, Merkel AR, Masood-Campbell SK, Elefteriou F, and Sterling JA (2012). Models of bone metastasis. *Journal of visualized experiments : JoVE*, e4260. [PubMed: 22972196]
- Christiansen EM, Yang SJ, Ando DM, Javaherian A, Skibinski G, Lipnick S, Mount E, O'Neil A, Shah K, Lee AK, et al. (2018). In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell* 173, 792–803 e719. [PubMed: 29656897]
- Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, and Ronneberger O (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation* (Cham: Springer International Publishing).
- Condeelis J, and Weissleder R (2010). In vivo imaging in cancer. *Cold Spring Harb Perspect Biol* 2, a003848. [PubMed: 20861158]
- de Jong M, Essers J, and van Weerden WM (2014). Imaging preclinical tumour models: improving translational power. *Nat Rev Cancer* 14, 481–493. [PubMed: 24943811]
- Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak J, the CC, Hermsen M, Manson QF, et al. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* 318, 2199–2210. [PubMed: 29234806]
- Ertürk A, Becker K, Jährling N, Mauch CP, Hojer CD, Egen JG, Hellal F, Bradke F, Sheng M, and Dodt H-U (2012). Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols* 7, 1983. [PubMed: 23060243]
- Eser S, Reiff N, Messer M, Seidler B, Gottschalk K, Dobler M, Hieber M, Arbeiter A, Klein S, Kong B, et al. (2013). Selective requirement of PI3K/PDK1 signaling for Kras oncogene-driven pancreatic cell plasticity and cancer. *Cancer cell* 23, 406–420. [PubMed: 23453624]
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, and Thrun S (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115–118. [PubMed: 28117445]

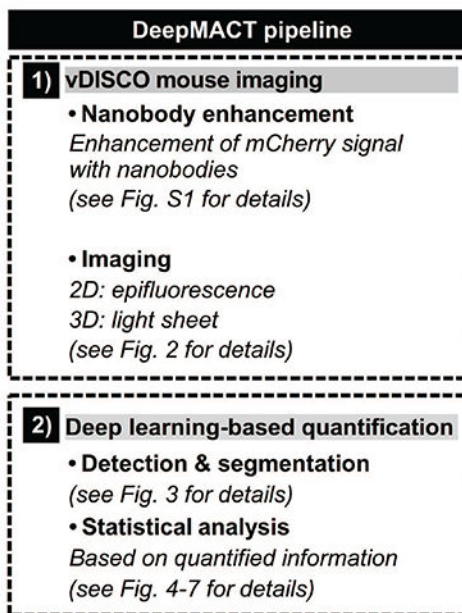
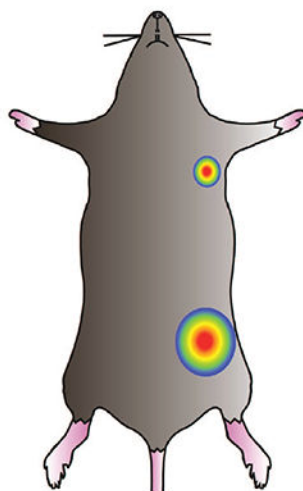
- Falk T, Mai D, Bensch R, Cicek O, Abdulkadir A, Marrakchi Y, Bohm A, Deubner J, Jackel Z, Seiwald K, et al. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods* 16, 67–70. [PubMed: 30559429]
- Gondi G, Mysliwicz J, Hulikova A, Jen JP, Swietach P, Kremmer E, and Zeidler R (2013). Antitumor efficacy of a monoclonal antibody that inhibits the activity of cancer-associated carbonic anhydrase XII. *Cancer research* 73, 6494–6503. [PubMed: 24030978]
- Hanahan D, and Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. [PubMed: 21376230]
- Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA, Ross S, Conrads TP, Veenstra TD, Hitt BA, et al. (2003). Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer cell* 4, 437–450. [PubMed: 14706336]
- Iorns E, Drews-Elger K, Ward TM, Dean S, Clarke J, Berry D, El Ashry D, and Lippman M (2012). A new mouse model for the study of human breast cancer metastasis. *PLoS One* 7, e47995. [PubMed: 23118918]
- Jones E, Oliphant T, and Peterson P (2001). SciPy: Open Source Scientific Tools for Python. URL <http://www.scipy.org/>
- Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, et al. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 172, 1122–1131 e1129. [PubMed: 29474911]
- Kingma DP, and Ba J (2014). Adam: A Method for Stochastic Optimization. In arXiv e-prints.
- Kubota SI, Takahashi K, Nishida J, Morishita Y, Ehata S, Tainaka K, Miyazono K, and Ueda HR (2017). Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution. *Cell reports* 20, 236–250. [PubMed: 28683317]
- Lambert AW, Pattabiraman DR, and Weinberg RA (2017). Emerging Biological Principles of Metastasis. *Cell* 168, 670–691. [PubMed: 28187288]
- Lenk L, Pein M, Will O, Gomez B, Viol F, Hauser C, Egberts JH, Gundlach JP, Helm O, Tiwari S, et al. (2017). The hepatic microenvironment essentially determines tumor cell dormancy and metastatic outgrowth of pancreatic ductal adenocarcinoma. *Oncoimmunology* 7, e1368603. [PubMed: 29296518]
- Litjens G, Sanchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, Bult P, van Ginneken B, and van der Laak J (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific reports* 6, 26286. [PubMed: 27212078]
- Liu Y, Kohlberger T, Norouzi M, Dahl GE, Smith JL, Mohtashamian A, Olson N, Peng LH, Hipp JD, and Stumpe MC (2018). Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection. *Arch Pathol Lab Med*.
- Massague J, and Obenauf AC (2016). Metastatic colonization by circulating tumour cells. *Nature* 529, 298–306. [PubMed: 26791720]
- Massoud TF, and Gambhir SS (2003). Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev* 17, 545–580. [PubMed: 12629038]
- Massoud TF, and Gambhir SS (2007). Integrating noninvasive molecular imaging into molecular medicine: an evolving paradigm. *Trends Mol Med* 13, 183–191. [PubMed: 17403616]
- Mavandadi S, Dimitrov S, Feng S, Yu F, Sikora U, Yaglidere O, Padmanabhan S, Nielsen K, and Ozcan A (2012a). Distributed medical image analysis and diagnosis through crowd-sourced games: a malaria case study. *PLoS One* 7, e37245. [PubMed: 22606353]
- Mavandadi S, Feng S, Yu F, Dimitrov S, Nielsen-Saines K, Prescott WR, and Ozcan A (2012b). A mathematical framework for combining decisions of multiple experts toward accurate and remote diagnosis of malaria using tele-microscopy. *PLoS One* 7, e46192. [PubMed: 23071544]
- McKinney W (2008). Pandas. URL <https://pandas.pydata.org/>
- Nguyen DX, Chiang AC, Zhang XH, Kim JY, Kris MG, Ladanyi M, Gerald WL, and Massague J (2009). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 138, 51–62. [PubMed: 19576624]
- Ntziachristos V (2010). Going deeper than microscopy: the optical imaging frontier in biology. *Nat Methods* 7, 603–614. [PubMed: 20676081]

- Pan C, Cai R, Quacquarelli FP, Ghasemigharagoz A, Lourbopoulos A, Matryba P, Plesnila N, Dichgans M, Hellal F, and Erturk A (2016). Shrinkage-mediated imaging of entire organs and organisms using uDISCO. *Nat Methods*.
- Pandey M, and Mahadevan D (2014). Monoclonal antibodies as therapeutics in human malignancies. *Future Oncol* 10, 609–636. [PubMed: 24754592]
- Paszke A (2016). PyTorch. URL <https://pytorch.org/>
- Pichler BJ, Wehrl HF, and Judenhofer MS (2008). Latest advances in molecular imaging instrumentation. *J Nucl Med* 49 Suppl 2, 5S–23S. [PubMed: 18523063]
- Ronneberger O, Fischer P, and Brox T (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (Cham: Springer International Publishing).
- Ryan DP, Hong TS, and Bardeesy N (2014). Pancreatic adenocarcinoma. *The New England journal of medicine* 371, 1039–1049. [PubMed: 25207767]
- Schonhuber N, Seidler B, Schuck K, Veltkamp C, Schachtler C, Zukowska M, Eser S, Feyerabend TB, Paul MC, Eser P, et al. (2014). A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med* 20, 1340–1347. [PubMed: 25326799]
- Sevenich L, Bowman RL, Mason SD, Quail DF, Rapaport F, Elie BT, Brogi E, Brastianos PK, Hahn WC, Holsinger LJ, et al. (2014). Analysis of tumour- and stroma-supplied proteolytic networks reveals a brain-metastasis-promoting role for cathepsin S. *Nature cell biology* 16, 876–888. [PubMed: 25086747]
- Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, Thng F, Peng L, and Stumpe MC (2018). Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am J Surg Pathol* 42, 1636–1646. [PubMed: 30312179]
- Sullivan DP, Winsnes CF, Akesson L, Hjelmare M, Wiking M, Schutten R, Campbell L, Leifsson H, Rhodes S, Nordgren A, et al. (2018). Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat Biotechnol* 36, 820–828. [PubMed: 30125267]
- Susaki EA, Tainaka K, Perrin D, Kishino F, Tawara T, Watanabe TM, Yokoyama C, Onoe H, Eguchi M, Yamaguchi S, et al. (2014). Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* 157, 726–739. [PubMed: 24746791]
- Tabrizi M, Bornstein GG, and Suria H (2010). Biodistribution mechanisms of therapeutic monoclonal antibodies in health and disease. *The AAPS journal* 12, 33–43. [PubMed: 19924542]
- Tainaka K, Kubota SI, Suyama TQ, Susaki EA, Perrin D, Ukai-Tadenuma M, Ukai H, and Ueda HR (2014). Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* 159, 911–924. [PubMed: 25417165]
- Timpson P, McGhee EJ, and Anderson KI (2011). Imaging molecular dynamics in vivo—from cell biology to animal models. *J Cell Sci* 124, 2877–2890. [PubMed: 21878495]
- Topol EJ (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25, 44–56. [PubMed: 30617339]
- Tuchin VV (2016). Editor's Introduction: Optical Methods for Biomedical Diagnosis.
- Vestjens JH, Pepels MJ, de Boer M, Borm GF, van Deurzen CH, van Diest PJ, van Dijk JA, Adang EM, Nortier JW, Rutgers EJ, et al. (2012). Relevant impact of central pathology review on nodal classification in individual breast cancer patients. *Ann Oncol* 23, 2561–2566. [PubMed: 22495317]
- Vick B, Rothenberg M, Sandhofer N, Carlet M, Finkenzeller C, Krupka C, Grunert M, Trumpp A, Corbacioglu S, Ebinger M, et al. (2015). An advanced preclinical mouse model for acute myeloid leukemia using patients' cells of various genetic subgroups and in vivo bioluminescence imaging. *PLoS One* 10, e0120925. [PubMed: 25793878]
- von Burstin J, Eser S, Paul MC, Seidler B, Brandl M, Messer M, von Werder A, Schmidt A, Mages J, Pagel P, et al. (2009). E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex. *Gastroenterology* 137, 361–371, 371 e361–365. [PubMed: 19362090]
- von Burstin J, Eser S, Seidler B, Meining A, Bajbouj M, Mages J, Lang R, Kind AJ, Schnieke AE, Schmid RM, et al. (2008). Highly sensitive detection of early-stage pancreatic cancer by multimodal near-infrared molecular imaging in living mice. *Int J Cancer* 123, 2138–2147. [PubMed: 18709639]

- von Neubeck B, Gondi G, Riganti C, Pan C, Parra Damas A, Scherb H, Erturk A, and Zeidler R (2018). An inhibitory antibody targeting carbonic anhydrase XII abrogates chemoresistance and significantly reduces lung metastases in an orthotopic breast cancer model in vivo. *Int J Cancer* 143, 2065–2075. [PubMed: 29786141]
- Wang H, Rivenson Y, Jin Y, Wei Z, Gao R, Gunaydin H, Bentolila LA, Kural C, and Ozcan A (2019). Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat Methods* 16, 103–110. [PubMed: 30559434]
- Wang J, Fang Z, Lang N, Yuan H, Su MY, and Baldi P (2017). A multi-resolution approach for spinal metastasis detection using deep Siamese neural networks. *Comput Biol Med* 84, 137–146. [PubMed: 28364643]
- Wang N, Yi Wang CB, Min Xu, Chenchen Qin, Xin Yang, Tianfu Wang, Anhua Li, inggang Shen, Dong Ni (2018). Densely Deep Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 – 21st International Conference, 2018, Proceedings*, 641–648.
- Waskom M (2012). seaborn: statistical data visualization. URL <https://seaborn.pydata.org/>
- Yang B, Treweek JB, Kulkarni RP, Deverman BE, Chen CK, Lubeck E, Shah S, Cai L, and Gradinaru V (2014). Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* 158, 945–958. [PubMed: 25088144]
- Zipfel WR, Williams RM, Christie R, Nikitin AY, Hyman BT, and Webb WW (2003). Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proc Natl Acad Sci U S A* 100, 7075–7080. [PubMed: 12756303]

A Tumor cells transplanted in mice**B**

Limited detection of cancer metastases by bioluminescence imaging



Detection of all cancer metastases and analysis of antibody distribution



Figure 1. Experimental design and schematic of the DeepMACT pipeline for analysis of cancer metastases and antibody drug targeting

(A) Illustration of the experimental workflow for tumor transplantation and antibody application. (B) Steps of the DeepMACT pipeline on full-body mouse scans. First, the mice are fixed and processed with the vDISCO protocol to enhance the fluorescent signal of cancer cells. Transparent mice are subsequently imaged from head to toe using light-sheet microscopy, revealing all metastases. Light-sheet images are assembled into a complete 3D image of the mouse. Next, convolutional neural networks are trained to identify and segment

all micrometastases in the fluorescence signal. The trained algorithms are then applied to 3D images to detect cancer metastases and an antibody-based drug targeting in full-body mouse scans.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

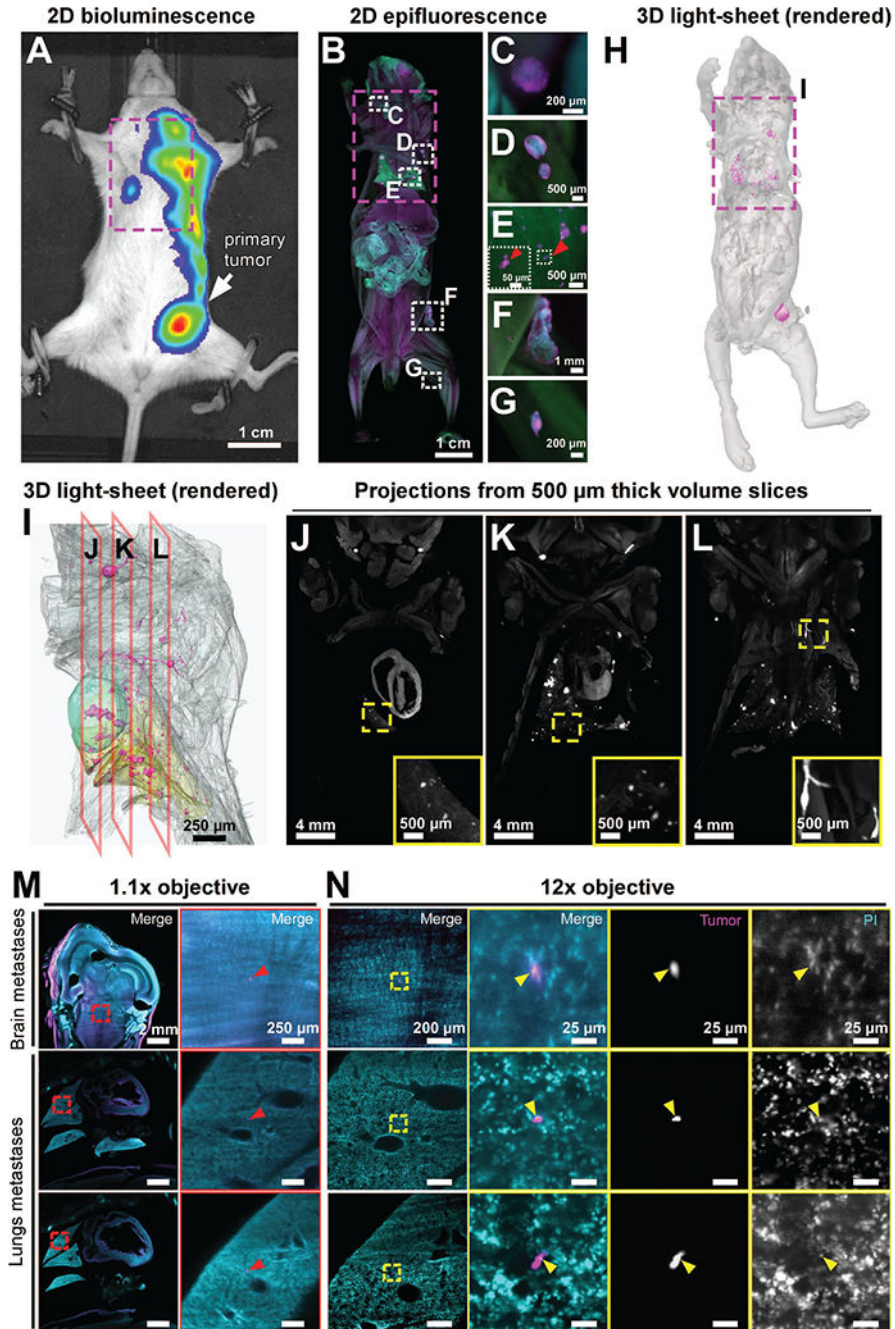


Figure 2. DeepMACT step 1: vDISCO visualization of metastases in a full-body scan of a mouse
(A) Bioluminescence image of a NSG female mouse before vDISCO which was taken 2 months after MDA-MB-231 cancer cell implantation into the mammary fat pad. **(B-G)** Epifluorescence images of the same mouse after vDISCO show metastases (magenta) in greater detail compared to bioluminescence, including small micrometastases that can be readily detected in the lungs (**E**, red arrowhead) and in the leg (**G**), in addition to the primary tumor (**F**) and major metastases (**C** and **D**) that are also visible in bioluminescence as bulk signal (**A**). **(H)** 3D visualization of the transparent mouse body imaged by light-sheet

microscopy. **(I)** Lateral views of the 3D segmentation obtained from the light-sheet imaging data corresponding to the magenta-boxed region indicated in (A, B, and **H**). For simplicity, only a few organs are segmented: the heart (cyan) and the lungs (yellow); the mouse body is shown in transparent gray and the metastases are in magenta. **(J-L)** Original light-sheet microscopy data (500 μm projections) showing metastases from the sagittal planes indicated in (J). **(M-N)** Single cell metastases identified in the brain and in the lungs by full-body light-sheet microscopy scans using a 1.1x objective with 6 μm lateral resolution (tumor cells in magenta and nucleus labeled with propidium iodide, PI, in cyan) (red arrowheads in M). The same metastases were further imaged by light-sheet microscopy with a 12x objective. Single plane images showed the colocalization of each micrometastasis with a single nucleus (yellow arrowheads in N). See Figure S1–S3 and Videos S1, S2.

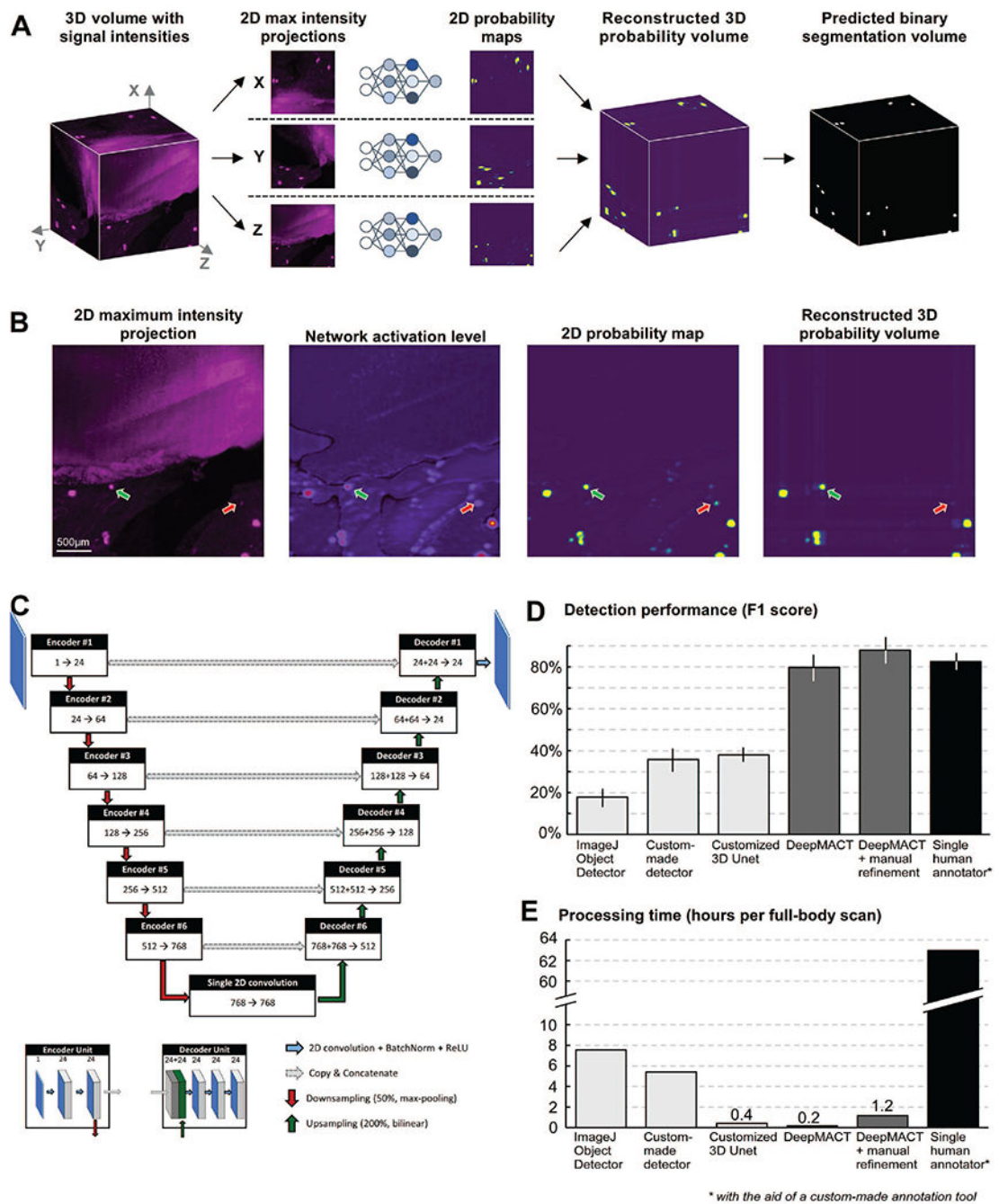


Figure 3. DeepMACT step 2: Schematic and performance of the deep learning algorithm

(A) Representation of the deep learning inference workflow to efficiently derive 3D detection and segmentation exploiting three 2D computational operations. (B) Visualization of the computational stages; the green arrow shows successful detection of a metastasis, the red arrow shows elimination of a false positive detection in the 3D reconstruction stage. (C) High-level representation of the network architecture with an encoding and a decoding path. (D-E) Comparison of our deep learning pipeline, DeepMACT, to alternative automated

methods and manual segmentation by a human expert in terms of detection performance (**D**) and processing time (**E**).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

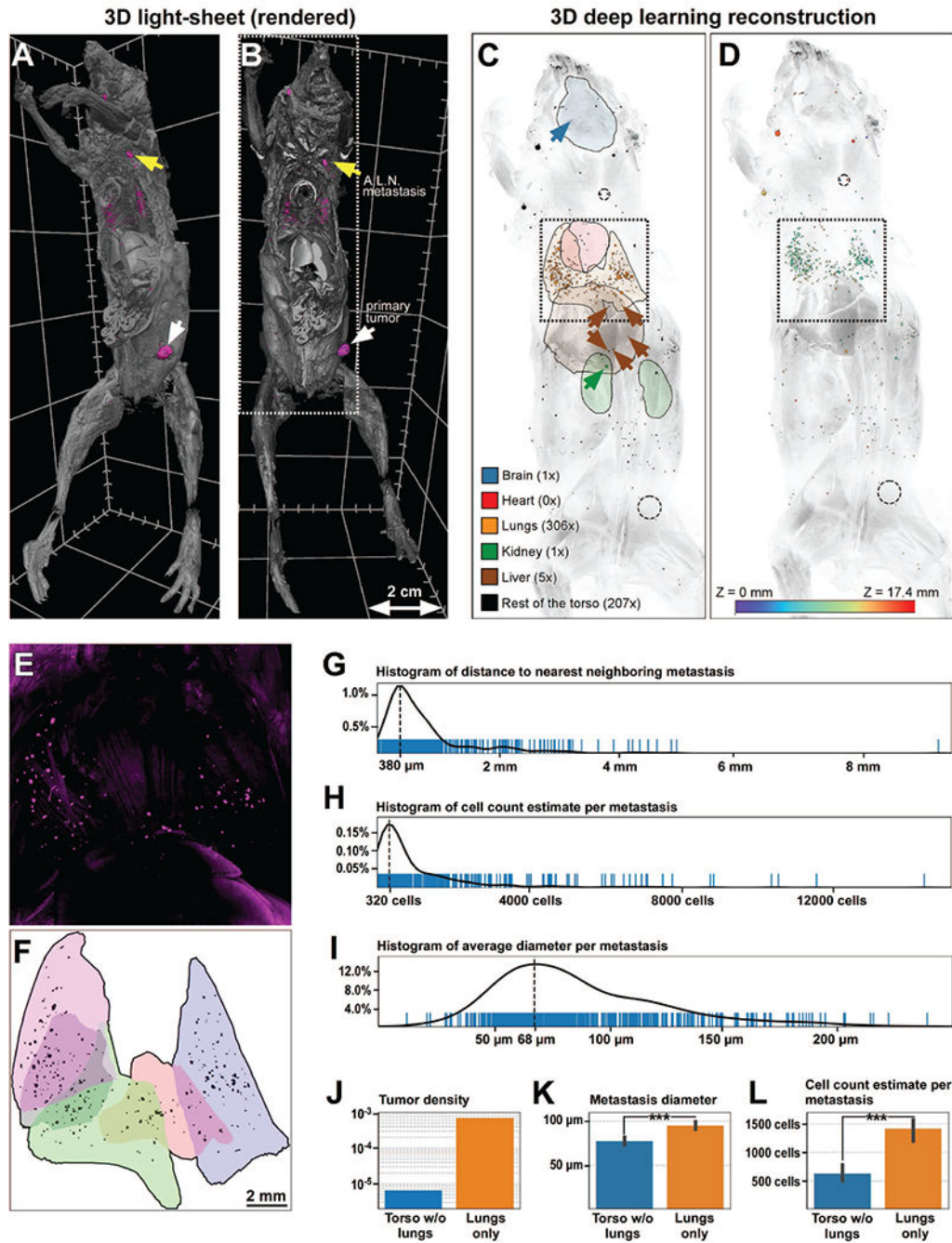


Figure 4. Deep learning-based detection and segmentation enables quantitative analysis at the level of individual metastases

(A,B) 3D rendering of a mouse transplanted with MDA-MB-231 cells in the mammary fat pad after light-sheet microscopy imaging in lateral and ventral views, respectively. Metastases in the mouse body are shown in magenta. The white arrow indicates the primary tumor and the yellow arrow indicates metastases in the axillary lymph node (A.L.N.). (C,D) Deep learning reconstructions of all detected metastases (A.L.N. and primary tumor indicated with dashed circles) color-coded by organ (C) and depth along the z-axis (D),

cropped to the white box in (B) to show higher level of detail. **(E,F)** Detailed view of metastases in the lung region (corresponding to the black box in C) in a maximum intensity projection of a 3D light-sheet scan (E) and projection of 3D deep learning-based detection, with metastases registered to individual lung lobes (shown in different colors) (F). **(G-L)** Deep learning-based distributions; blue bars show individual metastases, the black line shows the Gaussian kernel density estimation. **(G)** 3D distance to nearest neighboring metastasis. **(H)** Estimates of cell counts per metastasis. **(I)** Metastasis diameter averaged in 3D space. **(J-L)** Quantitative comparison between metastases in the lungs and the rest of the torso; bars indicate 95%-confidence intervals. **(J)** Tumor density as share of metastatic tissue of the entire volume is two orders of magnitude higher in lungs versus rest of torso. **(K)** Metastasis diameter (averaged in 3D space) is significantly higher in lungs ($p < 0.001$; two-sided t-test). **(L)** Cell count estimate per metastasis is significantly higher in lungs ($p < 0.001$; two-sided t-test).

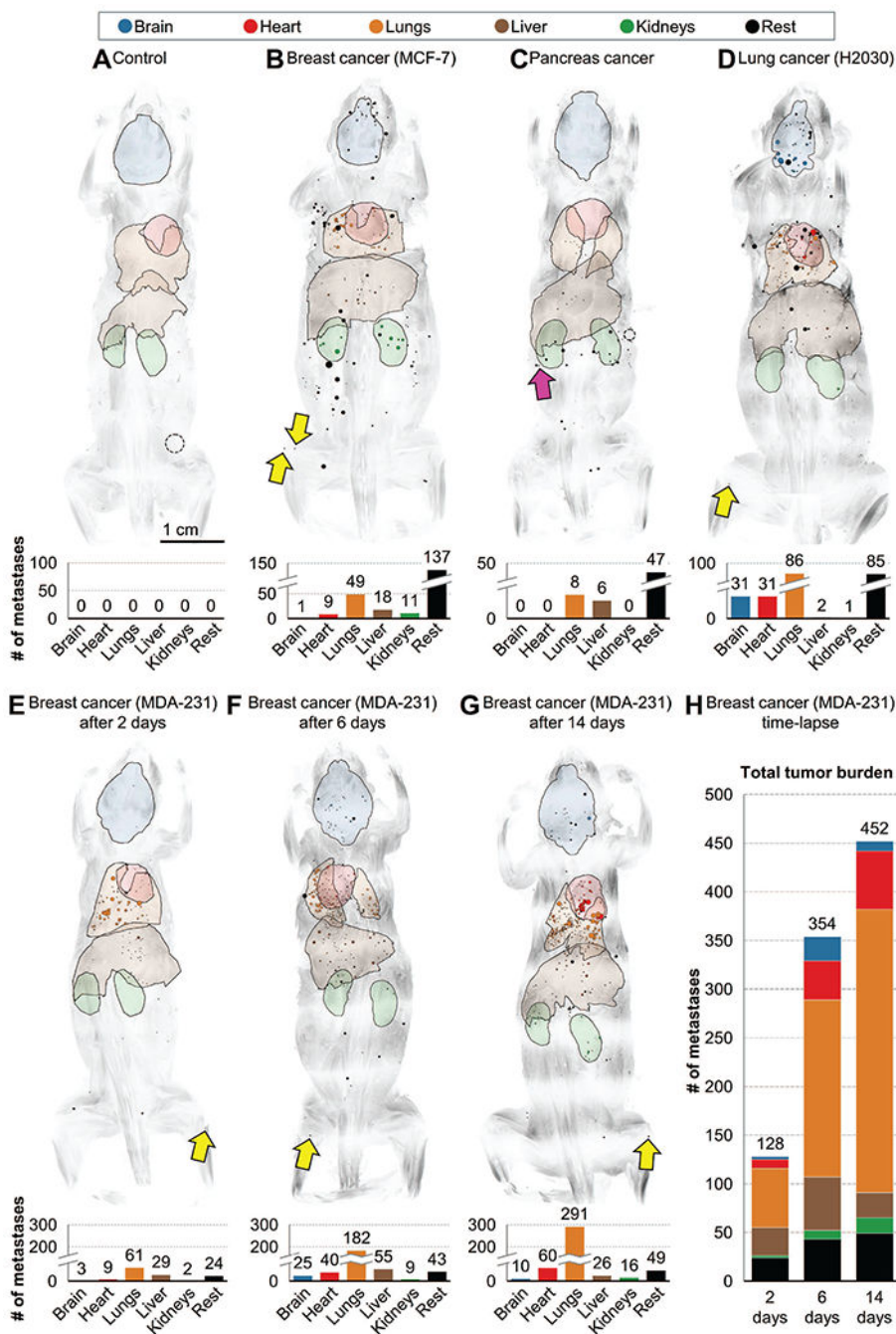


Figure 5. DeepMACT reliably detects metastases in all organs for a variety of tumor models
 Metastasis detections in full-body 3D light-sheet microscopy scans; each dot represents a metastasis, color-coded by organ; black metastasis within organ outlines are not inside that organ but rather above or below it. (A) A control mouse was perfused immediately after implant of a solid tumor (MDA-MB-231; dashed circle), leaving no time for metastases to form. (B) MCF-7 breast cancer cells were intracardially injected in a nude mouse. (C) Pancreatic cancer cells (R254) were transplanted into the pancreas (dashed circle) of a C57BL/6 mouse. (D) H2030-BrM3 lung cancer cells were intracardially injected in a nude

mouse. **(E-G)** Three NSG mice were intracardially injected of MDA-MB-231 breast cancer cells and sacrificed after 2 days **(E)**, 6 days **(F)**, and 14 days **(G)**. **(H)** DeepMACT analysis shows increase in tumor burden over time. Yellow arrows indicate metastases in bones; the magenta arrow indicates a metastasis in the peritoneum.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

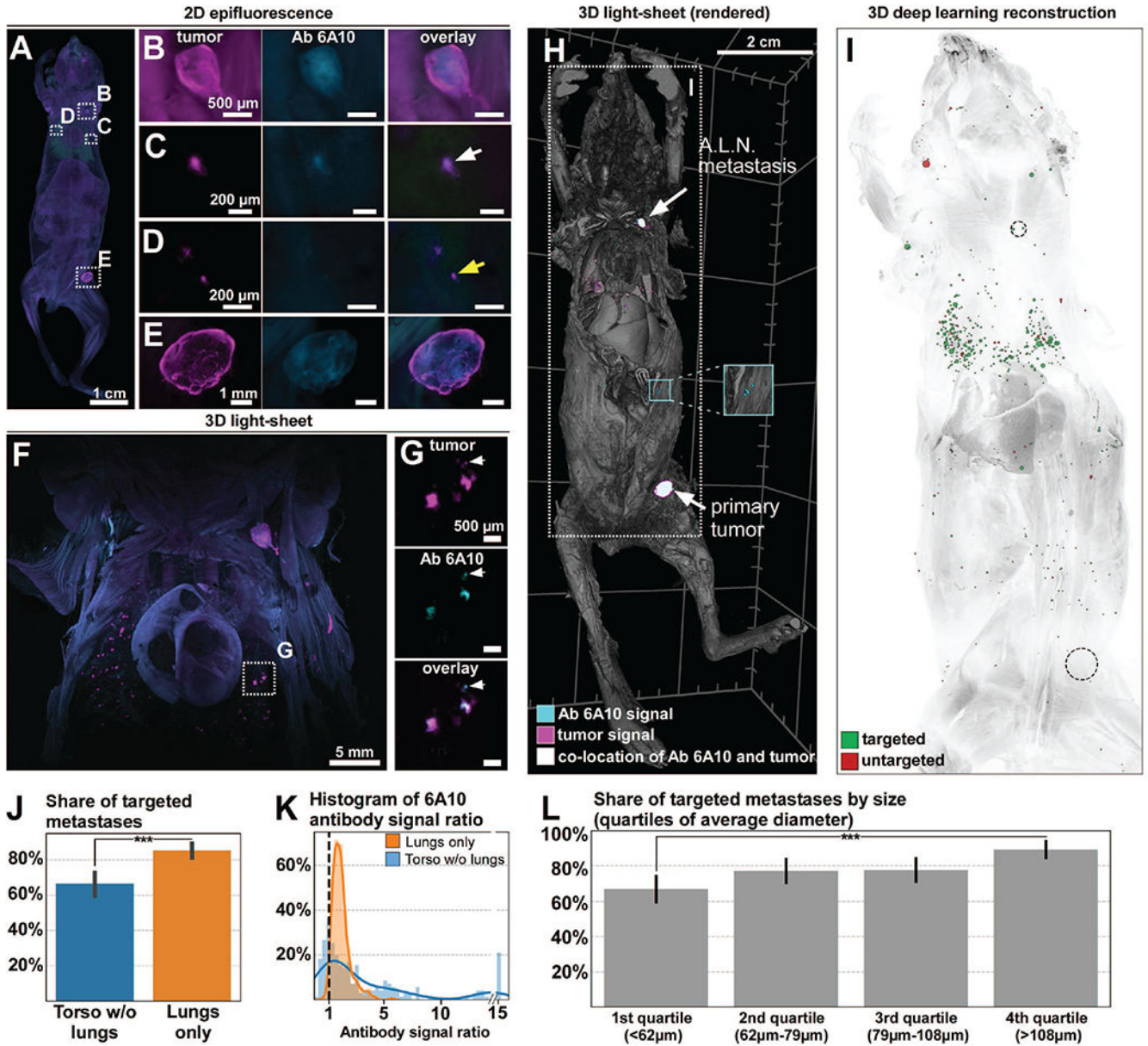


Figure 6. The DeepMACT pipeline enables quantitative analysis of drug delivery efficacy at the level of single metastases

A mouse transplanted in the mammary fat pad with MDA-MB-231 cells was intravenously injected with 6A10 anti-CA12 antibody 9 weeks later. (A) Epifluorescence images of a processed mouse show details (B-E) of both tumor metastases (boosted with Alexa647N nanobody, shown in magenta) and 6A10 antibody (conjugated with Alexa568, shown in cyan) distributions and their overlay. While most of the micrometastases are targeted by the antibody (C, white arrow), there are some that are not (D, yellow arrow). (F) Full-body 3D light-sheet scan, cropped to the chest region, shows the distributions of metastases (magenta) and antibody (cyan). (G) Detailed view of the boxed region in (F) showing very small micrometastases targeted by the therapeutic antibody (white arrows). (H) 3D rendering of a mouse body light-sheet scan showing the tumor signal in magenta and the 6A10 antibody

signal in cyan (co-localization of the signals is shown in white). The cyan inset shows an example of off-target accumulation of the 6A10 antibody. **(I)** Deep learning-based reconstruction of the animal in **(H)** showing targeted metastases in green and untargeted metastases in red; the dashed circles represent the primary tumor A.L.N metastases. **(J)** A significantly higher share of metastases are targeted in the lungs versus the rest of torso ($p < 0.001$, two-sided t-test). **(K)** Comparison of the distributions of 6A10 antibody signal ratio (signal strength in metastasis versus local surrounding; see the methods for further details) per metastasis in the lungs versus the rest of torso. The dashed line indicates a ratio of 1 (equal signal strengths). **(L)** Share of targeted metastases as a function of their size (split into quartiles of average metastasis diameter) ($p < 0.001$, two-sided t-test).

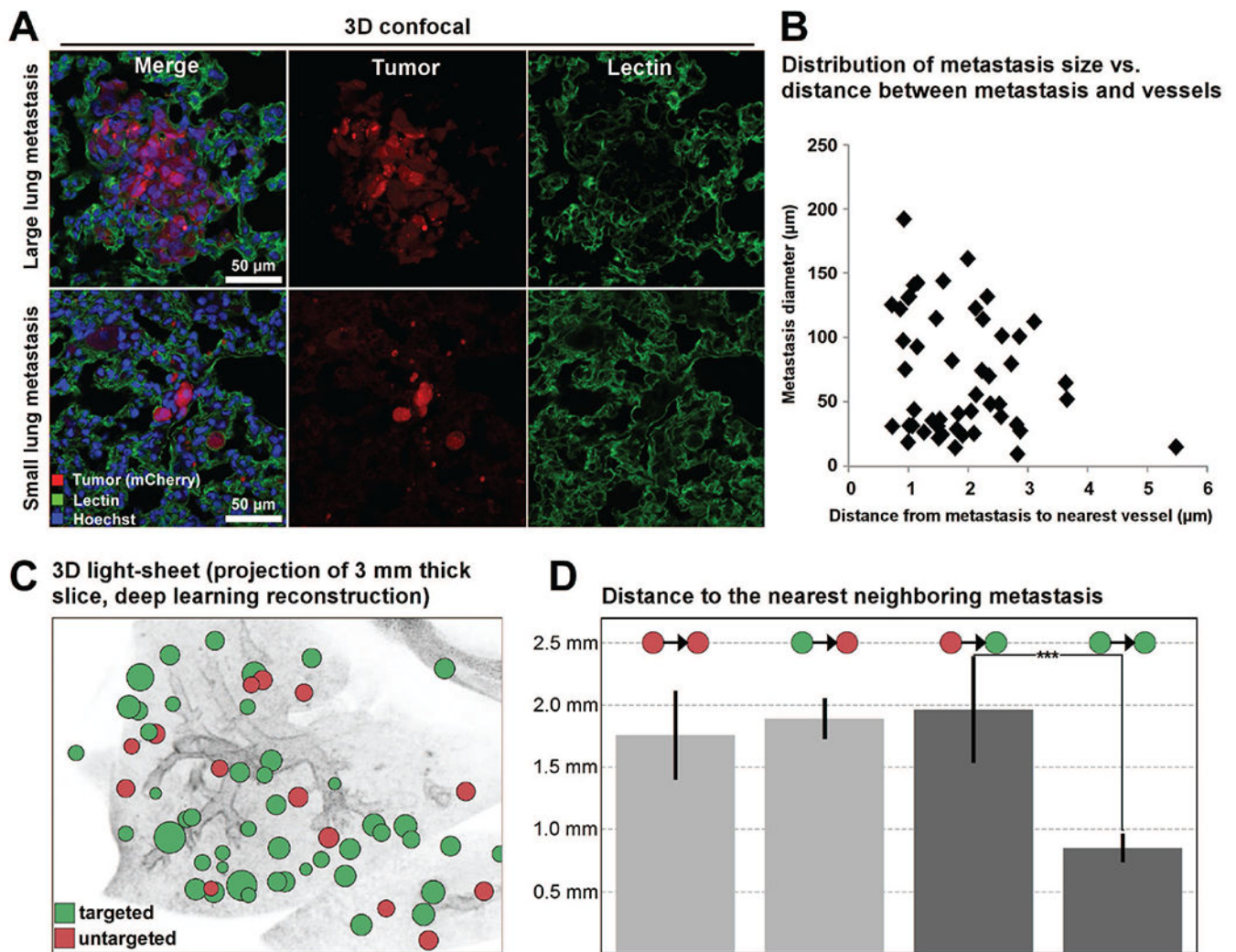


Figure 7. Potential mechanisms of metastasis targeting by therapeutic antibody

(A) Confocal images of a large and a small metastasis (less than 5 cancer cells) in lungs of a mouse transplanted with MDA-MB-231 cells and intravenously injected with 6A10 anti-CA12 antibody, labeled with lectin (green) and Hoechst (blue). (B) Distribution of metastasis size and distance to the nearest vessel, showing that most of the metastases are close to vessels (distance less than 6 μm) (n=50). (C) Deep learning-based reconstruction of lung metastases with and without 6A10 antibody targeting. (D) Deep learning-based quantification of distance between metastases and their nearest neighbor. The average distance from an untargeted to the nearest targeted metastasis is significantly ($p < 0.001$; two-sided t-test) larger than from a targeted one; this shows local clustering of targeted and untargeted metastases (see the methods for further details).