Microfluidics guided by deep learning for cancer immunotherapy screening

Zheng Ao1,2, Hongwei Cai1, Zhuhao Wu3, Liya Hu3, Asael Nunez3, Zhuolong Zhou4,5, Hongcheng Liu3, Maria Bondesson3, Xiongbin Lu3,6, Xin Lu7,8,9, Ming Dao10,11, and Feng Guo12,6

Immunocyte infiltration and cytotoxicity play critical roles in both inflammation and immunotherapy. However, current cancer immunotherapy screening methods overlook the capacity of the T cells to penetrate the tumor stroma, thereby significantly limiting the development of effective treatments for solid tumors. Here, we present an automated high-throughput microfluidic platform for simultaneous tracking of the dynamics of T cell infiltration and cytotoxicity within the 3D tumor cultures with a tunable stromal makeup. By recourse to a clinical tumor-infiltrating lymphocyte (TIL) score analyzer, which is based on a clinical data-driven deep learning method, our platform can evaluate the efficacy of each treatment based on the scoring of T cell infiltration patterns. By screening a drug library using this technology, we identified an epigenetic drug (lysine-specific histone demethylase 1 inhibitor, LSD1i) that effectively promoted T cell tumor infiltration and enhanced treatment efficacy in combination with an immune checkpoint inhibitor (anti-PD1) in vivo. We demonstrated an automated system and strategy for screening immunocyte-solid tumor interactions, enabling the discovery of immune- and combination therapies.


Competing interest statement: Z.A., H.C., and F.G. have filed a patent based on the work presented in this paper. This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS.

Published November 7, 2022.

This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence may be addressed. Email: mingdao@mit.edu, or fengguo@iu.edu.

2Authors contributed equally to this work.

Significance

Immune-cell infiltration and cytotoxicity to pathogens and diseased cells are ubiquitous in health and disease. To better understand immune-cell behavior in a 3D environment, we developed an automated high-throughput microfluidic platform that enables real-time imaging of immune-cell infiltration dynamics and killing of the target cancer cells. We trained a deep learning algorithm using clinical data and integrated the algorithm with our microfluidic platform to effectively identify epigenetic drugs that promote T cell tumor infiltration and enhance cancer immunotherapy efficacy both in vitro and in vivo. Our platform provides a unique method to investigate immune-tissue interactions, which can be widely applied to oncology, immunology, neurology, microbiology, tissue engineering, regenerative medicine, translational medicine, and so on.

An effective immune system is critically important for fighting against cancer. However, immune evasion is one of the key hallmarks of cancer. During tumor development, lymphocytes infiltrate the tumor and inhibit its progression. Meanwhile, tumors evolve to evade such immune surveillance by downregulating antigen presentation, secreting extracellular matrix (ECM) to physically limit lymphocyte infiltration as well as chemokine/cytokines to repel lymphocytes. (1) High densities of tumor-infiltrating lymphocytes (TILs) correlate with improved prognosis in many cancer types, including breast, colorectal, ovarian, skin, and pancreatic cancers (2–4). In addition to TIL density, clinical evidence shows that the type, infiltration depth, clustering index, and activation status of TILs inside solid tumors are closely associated with disease progression and treatment outcomes (5). Recent advances in cancer immunotherapy, such as immune checkpoint inhibitors (ICIs) that remove the “brakes” on T cell–mediated antitumor immunity (6), point to potentially better outcomes in solid tumor treatment. However, the responses of cancer patients to treatment protocols that entail ICI drugs vary greatly, with overall response rates ranging from <10% to ~40% for solid tumors (7). The ineffectiveness of ICI therapy can be largely attributed to limited TIL infiltration (8). However, TIL infiltration behavior has not been thoroughly investigated and exploited for cancer immunotherapy discovery. Thus, methods based on TIL behavior may become important tools for the discovery of effective immunotherapies and/or combination therapies that not only improve T cell cytotoxicity but also simultaneously promote T cell tumor infiltration.

To date, efforts to discover cancer immunotherapies have been performed mainly using in vitro two-dimensional (2D) cell cultures and/or in vivo animal models. However, these models have some distinct limitations. The 2D cultures fail to recapitulate key physiological aspects of tumors, such as the tumors’ stromal physical barrier and hypoxic cores (9). Animal models are limited by their scalability and are time-consuming and labor intensive (10). The 2D culture-based screening assays often rely on easily obtainable readouts, such as cytotoxicity, activation status, and cytokine profiles of T cells, while ignoring the more complex dynamics of T cell tumor infiltration behavior. In contrast, 3D tumor cultures, such as organoids/spheroids, can be constructed to include key physiological aspects of an in vivo tumor, including tumor genotypes and phenotypes, cell–cell contacts, ECM barriers, and hypoxic cores (11, 12), thereby benefitting from their potential for elucidating dynamic immunocyte–tumor interactions (13). For example, a large number of colon tumor organoids cultured in a thick basement membrane matrix such as Matrigel were used to isolate tumor-reactive...
T cells based on the cytotoxicity and proliferation of T cells (14). However, challenges in 3D tumor culturing, such as difficulties in generating scalable, uniformly sized, and standardized organoids/spheroids, as well as obstruction of free immunocyte infiltration and interaction by thick ECM embedding, have limited utility for high-throughput T cell behavior–based phenotypic screening (15). Furthermore, manual acquisition and analysis of the vast amount of data generated from 3D cultures pose barriers to establishing efficient and scalable high-throughput screens. Recently, automated data analysis approaches based on deep learning methods have improved the consistency and accuracy of analyzing TIL scores in pathological slides (16–18). Despite their promise, deep learning methods have not yet been used for TIL pattern analysis in in vitro models. Thus, the development of high-throughput, standardized, and automated technologies incorporating deep learning for automated TIL behavior analysis would pave the way for further advances in immunotherapy screening.

Here, we report a high-throughput, automated system for T cell behavior (e.g., tumor infiltration)–based immunotherapy discovery. This constitutes a unique demonstration of cancer immunotherapy screening that can simultaneously interrogate both T cell infiltration behavior and cytotoxicity. Through a relatively simple yet effective pillar-lattice-array design, our platform provides scalable, standardized, and tunable “core/tumor-shell/stroma” spheroids for robust drug discovery applications. In addition, we demonstrate a strategy using a deep learning algorithm to instruct drug prioritization from real-world features and clinical and pathological datasets. As a proof of concept, we used our intelligent microfluidic system to screen a library of 141 small-molecule epigenetic modulators as potential drug candidates. Among these epigenetic modulators, we have identified a lead compound, GSK-LSD1, which effectively promotes T cell infiltration, and consequently promotes deep tumor cytotoxicity once combined with anti-programmed cell death protein 1 (PD1) treatment. Furthermore, we validated the efficacy of this compound in a B16F10 mouse model in vivo.

Results

Integration of Automated Screening System. Our intelligent microfluidic screening system incorporates two components (Fig. 1A). (1) A microfluidic immune-tumor interaction platform: Through the introduction of pillar-lattice-arrays into well plates, our platform comprises 7,680 uniformly sized heterotypic tumor spheroids per well plate. Each spheroid consists of a uniform tumor core and stromal outer layer, representing the essential features of a primary tumor. Both the size of the core and the thickness of the outer layer are tunable. With the Matrigel-free culture and pillar immobilization of spheroids, our platform enables perfusion, free interaction, and time-lapse tracking of immunocyte dynamic behaviors within a large number of spheroids under different treatment conditions (e.g., 96, 384, or 1,536 conditions per run, depending on the choice of well plates), mitigating technical inconsistencies of long-term coculturing and imaging, such as fusion, size variation, and motility of spheroids. (2) A deep learning–based TIL score analyzer: To process the vast amount of T cell behavior data from our microfluidic platform, we incorporated automated image acquisition and image-based screening. The deep learning TIL scoring algorithm is trained using clinical data, including digital pathology images and survival data from 411 tumors from 397 melanoma patients in The Cancer Genome Atlas (TCGA) database, and our TIL score analyzer then automatically generates the T cell infiltration scores that correspond to TIL distribution patterns learned from the images of the high-survival (>3 y) or low-survival patient group, respectively. We integrated our microfluidic platform with the deep learning–based TIL score analyzer to produce an automated screening system to determine immunocyte behaviors and enable screening for immunotherapeutic drugs in a high-throughput and automated manner.

Formation of Scalable Tumor-Stroma Spheroid Array. We fabricated scalable uniformly sized 3D core/tumor-shell/stroma spheroids (Fig. 1B and SI Appendix, Fig. S1A) in the pillar-lattice-arrays.
within a well plate using a two-step cell-seeding procedure. Initially, tumor cell suspensions (labeled with membrane DiO green dye, pseudocolored as blue) were seeded into the pillar-lattice-arrays to self-assemble into tumor spheroid “cores” within 12 h. The spheroid formation was driven by the geometry of the hydrophobic pillar lattices, which repel the cells. After 1 d of culture, cancer-associated fibroblasts (CAFs, labeled with membrane Dil red dye) isolated from orthotopic primary tumors were seeded on top of the tumor “cores” to form heterotypic spheroids. A large number of spheroids (7,680) per plate can be formed within pillar-lattice-arrays within one well; they reached their maximum diameter on day 3 and were arrested at uniform sizes (160 ± 9.3 μm) by the pillar-lattice-array design (Fig. 1C). These immobilized uniform-sized spheroids maintained high viability over prolonged culture (SI Appendix, Fig. S1B). This method avoids the risks of spheroids merging or moving in the well during medium changes, which is a common problem with other spheroid fabrication and culture methods, thus highlighting the advantage of our design for high-throughput screening applications. Notably, the 3D core/tumor-shell/stroma spheroids were fabricated to recapitulate tumor immunosuppressive architectures and microenvironments in that CAFs secrete ECM to physically restrict T cell infiltration and secrete chemokines to repel T cells (19). We also demonstrated the fine-tuning of dimension and composition of heterotypic spheroids by controlling the ratio and initial seeding density of tumor cells and CAFs (SI Appendix, Fig. S1C). In addition, we confirmed that tumor-stroma spheroids proliferate faster compared to regular tumor spheroids (SI Appendix, Fig. S1D), indicating that the CAFs retained their in vivo function to promote tumor cell proliferation.

Tracking Dynamic T Cell Behavior within 3D Cultures. Through time-lapse imaging, our microfluidic platform tracks the dynamic interactions of T cells with tumor-stroma spheroids at a single-cell resolution while facilitating high throughput. To visualize tumor antigen-specific T cell interactions with the tumors, we subjected the tumor core consisting of ovalbumin (OVA257–264) presenting tumor cells with a shell of CAFs to OVA antigen-specific T cells (OT-I cells). Using our microfluidic platform and fluorescence-labeled cells (T cells labeled with blue CMAC cell tracker dye, tumor cells labeled with DiO green dye, CAFs labeled with Dil red dye, and dead cells detected via the SYTOX deep-red dye), we observed and quantified the dynamic migration and killing behaviors of T cells within the tumor spheroids. We defined the infiltration depth (or killing depth) of a T cell as the radial distance between the T cell and the tumor spheroid surface (Fig. 2A). We investigated the impact of antigen (OVA) presentation on tumor cores consisting of UN-KC6141 cells, as well as the stromal layer of shells/CAFs, on the dynamics of immune-tumor interactions (Fig. 2 B and C and SI Appendix, Discussion S1). We found that the CAF shells inhibited both T cell infiltration and cytotoxicity (details in SI Appendix, Fig. S2), and that the presentation of OVA antigen of tumor cores enhanced T cell cytotoxicity (details in SI Appendix, Fig. S3). By using this engineering platform and cell system, we simultaneously tracked the dynamic processes of the tumor infiltration and cytotoxicity of an individual T cell. We further observed that a T cell swarmed toward a tumor spheroid at high speed, infiltrated the spheroid with a slower speed, performed killing of a tumor cell, and then continued to move at elevated speed to locate the next target tumor cell (Fig. 2 D and E and Movie S1). Moreover, by comparing the behaviors of cytotoxic and noncytotoxic T cells (n = 30), we found that the cytotoxic T cells had lower median speeds as well as fewer straight tracks compared to noncytotoxic T cells (SI Appendix, Fig. S1E), which are consistent with other reports in in vivo models (20).

Fig. 2. Tracking dynamics of T cell tumor infiltration and cytotoxicity. (A) Schematics of dynamic interactions between T cells and a tumor-stroma spheroid. T cell infiltration depth is defined as the radial distance from the cell to the spheroid surface, and infiltration (or killing) depth is positive (or negative) once the T cell has infiltrated (or is outside) the spheroid. (B) Dependence of T cell infiltration on the antigen presentation of the tumor core (Antigen (+)) and the presence of a stroma/CAF shell (CAF(+)). All of the lines are plotted with a 68% confidence interval (CI). (C) Dependence of T cell killing depth and capability (n: killed cell number) on the antigen presentation of tumor core (Antigen (+) and the stroma/CAF shell (CAF(+)). (D) Time-lapse images and their extractions of the simultaneous infiltration and killing dynamics of an individual T cell within a tumor-stroma spheroid (from Movie S1). The white (or magenta) arrow and circle indicate the T cell (or dead tumor cells), negative values of infiltration depth (Id) indicate T cells outside of spheroids, and positive values indicate T cells inside of spheroids. (E) Quantification of infiltration depth and speed of the same T cell (in D) over time. The green and red dashed lines indicate the infiltration and killing events of this T cell. (Scale bar in D: 50 μm.)
Training the Deep Learning-Based TIL Score Analyzer Using a Clinical Database. To process the data from very large numbers of images tracking T cell behaviors on many spheroids immobilized in our microfluidic platform, we developed a deep learning-based TIL score analyzer (Fig. 3A and SI Appendix, Fig. S4 and Discussion S2) that consists of an image processor, a deep learning algorithm, and a classifier. We integrated an image processor that extracted TIL maps by digitalizing hematoxylin and eosin (H&E) images of solid tumors (411 images from TCGA database) into a bicolor map in red (lymphocytes) and blue (tumors) (21). Next, we trained the deep learning-based TIL score analyzer to score TIL infiltration based on clinical patient survival data (Fig. 3A). The analyzer used a deep convolutional neural network to extract the features of TILs from training bicolor data (Fig. 3B). The analyzer used a deep convolutional neural network to extract the features of TILs from training bicolor maps and associated them with discrete patient survival time. Then, the analyzer could assign a TIL score of “0” or “1” by using a 3-y patient survival classifier with an area under the receiver operating characteristic curve of 0.8051 (SI Appendix, Fig. S4A). The TIL score correlates with the TIL number, infiltration depth, and clustering index such as the Ball-Hall or Banfield-Raftery indexes (21), and shows a better correlation with patient survival as compared with single TIL metrics (SI Appendix, Fig. S5). Using the well-trained analyzer, we evaluated the TIL scores of a particular drug X as TIL score (drug X). We defined the TIL score (drug X) by dividing TIL images into high- and low-infiltration groups by binning the cases based on predicted TIL score quartiles. Each screening result of “drug X” was given a TIL score of “0” or “1.” To generate the final TIL score (drug X), 10 T cell tumor infiltration images were randomly selected, and the final TIL score (drug X) was calculated by averaging these scores, (Fig. 3B). This automated strategy enables us to evaluate the efficacy of each drug in a large library in an automated, objective, and scalable manner.

Screening of Epigenetic Drugs. After the establishment of the method, we test-ran our intelligent microfluidic screening system in a proof-of-concept application by screening a library of drugs with epigenetic activities. It is known that epigenetic drugs can promote antigen presentation, reverse T cell exhaustion, or augment inflammation-related genes through the activation of endogenous retrovirus–mediated pathways (22). We chose a commercially available library of 141 epigenetic modulators against common targets, such as histone deacetylase, DNA methyltransferase, histone methyltransferase, and bromodomain and extra-terminal repeat. Since these small molecules may have cytotoxic activities (e.g., anti-proliferation, direct killing) and/or immunomodulating effects (e.g., altering T cell cytotoxicity, T cell infiltration), we performed the screening using our screening system in two steps: (1) drug cytotoxicity screening: Using our automated system, a total of 135,360 tumor-stroma spheroids (from melanoma B16F10 or pancreatic UN-KC6141 cell lines) were formed in our microfluidic platform in 48 h (480 spheroids/6 replicate wells per condition). After treating each epigenetic modulator at a screening concentration of 5 μM, we excluded any single agent that causes >20% cytotoxicity/anti-proliferative activity, and finally chose a total number of 50 noncytotoxic agents for B16F10 spheroids and 48 noncytotoxic agents for UN-KC6141 spheroids for further immunotherapeutic screening (SI Appendix, Fig. S6). (2) Immunotherapy screening: Initially, using our engineering platform, 23,520 tumor-stroma spheroids (melanoma: B16F10-CAF; pancreatic UN-KC6141-CAF) were formed in 48 h and treated with the noncytotoxic epigenetic agents that at a screening concentration of 5 μM for 24 h. After removal of the epigenetic agents, the tumor spheroids were subjected to fluorescence-labeled T cells (OT-I cells), following tracking of T cell infiltration and cytotoxicity within the tumor spheroids for 12 h. A total of 980 sets of z stack time-lapse images were collected after screening the full epigenetic drug library. Our screening system generated TIL scores for all of the drugs (Fig. 4A and SI Appendix, Fig. S7). The drug GSK-LSD1, a small-molecule inhibitor of LSD1, reached a TIL score of 0.6, which is significantly higher than that of positive control anti-PD1 treatment (0.3), untreated blank control (0.1), and antigen negative control (0.0) (Fig. 4A and SI Appendix, Fig. S4D). This
finding echoes previous studies that used animal models with genetic knockout of LSD1, in which the recruitment and activation of T cells are promoted (23).

In Vitro Validation of the Top Drug Candidate and Development of Combination Therapy. To further investigate the efficacy of GSK-LSD1 alone and in combination therapy, we simultaneously tracked the infiltration and cytotoxicity of T cells. We analyzed a total of 19,032 T cell tracks from 40 tumor spheroids treated with GSK-LSD1, anti-PD1, or a combination of GSK-LSD1 and anti-PD1 (Fig. 4B and Movie S2). Analysis of a total of 1,261 infiltrated T cells showed that (1) GSK-LSD1 treatment was applied in combination with anti-PD1 treatment, a greater tumor growth inhibition was observed, with two of eight mice having tumors completely regressed (Fig. 5B). Upon closer examination, a significantly higher number of CD8+ T cell infiltrated into GSK-LSD1-treated tumors compared to tumors of control or anti-PD1-treated mice, as analyzed by immunofluorescence staining (Fig. 5C) and flow cytometry (Fig. 5D and E and SI Appendix, Fig. S8). Our results are consistent with previous reports using LSD1 knockout mice or LSD1 inhibitor treatment in melanoma and breast cancer models (23, 24). Our in vivo test further validated our in vitro screening results. Thus, we demonstrated that our automated screening system can identify valid T cell tumor infiltration–promoting agents that translate into in vivo results.

In Vivo Validation of the Top Drug Candidates. To validate our in vitro results, we tested our top candidate drug in vivo in the B16F10 syngeneic tumor model (Fig. 5A). We examined the antitumor effect of GSK-LSD1 treatment alone and in combination with anti-PD1. A moderate tumor growth inhibition was seen in the GSK-LSD1 treatment group. Moreover, when GSK-LSD1 treatment was applied in combination with anti-PD1 treatment, a greater tumor growth inhibition was observed, with two of eight mice having tumors completely regressed (Fig. 5B). Upon closer examination, a significantly higher number of CD8+ T cell infiltrated into GSK-LSD1-treated tumors compared to tumors of control or anti-PD1-treated mice, as analyzed by immunofluorescence staining (Fig. 5C) and flow cytometry (Fig. 5D and E and SI Appendix, Fig. S8). Our results are consistent with previous reports using LSD1 knockout mice or LSD1 inhibitor treatment in melanoma and breast cancer models (23, 24). Our in vivo test further validated our in vitro screening results. Thus, we demonstrated that our automated screening system can identify valid T cell tumor infiltration–promoting agents that translate into in vivo results.

Discussion

Immune infiltration and cytotoxicity in solid tissues are essential for immune surveillance, inflammation, autoimmune disease, and immunotherapy. However, current in vitro models and methods are largely lacking in high-throughput tracking and analysis of dynamic behaviors of immunocytes within 3D tissues. To address these shortcomings, we developed an automated screening system that generates scalable uniform-sized core/tumor-shell/stroma spheroid arrays, enables free perfusion of immune cells, and achieves automated tracking and analysis of immune cell–tissue interaction dynamics. As a proof-of-concept application for cancer immunotherapy drug screening, our system was used to screen a library of epigenetic modulators in a high-throughput and automated manner. We identified and in vivo validated an epigenetic drug (GSK-LSD1) that effectively promotes T cell tumor infiltration and enhances T cell infiltration and cytotoxicity in combination with anti-PD1 treatment.
Scalable, Standardized 3D Cultures That Represent Key Physiological Features. Compared with 2D in vitro cultures, 3D cultures can recapitulate key physiological and pathological aspects of primary tumors, including tumor architecture, microenvironment, and drug response vulnerabilities (25). Recently, patient-derived organotypic cultures and tumor spheroids/organoids/clusters were established to model cancer immunity and test treatments, highlighting the uniqueness of these 3D models in preserving the genomic and/or phenotypic features of clinical tumors (26–28). However, current patient-derived 3D cultures suffer from a high heterogeneity in cell composition, cell architecture, and size and shape of the cell mass, limiting their applications for high-throughput screening and robust testing. Attempts have been made to develop methods for the standardization of organoid fabrication and minimization of interorganoid heterogeneity (29). Thus, it is critical to generate scalable standardized 3D cultures that represent key physiological features. In this work, considering tumor stroma as one of the main barriers to immune invasion, we developed a pillar-lattice-array–based microfluidic platform to fabricate uniformly sized core/tumor-shell/stroma spheroids on the scale of 10⁵ spheroids with tunable size, structure, and composition. We believe that our microfluidic platform and core-shell spheroid fabrication strategy can also be used for scalable, standardized 3D cultures with other key physiological features, such as endothelium and immune components.

Automated Culturing and Imaging of 3D Cultures. High-throughput screening platforms based on 2D culture with automated robotic handling for culturing cells and with imaging or readout systems for analysis have been developed. An automated system has also been published for culturing and maintaining 3D organoid cultures (29). However, challenges remain for tracking the dynamics of drug response or immune interactions of 3D cultures, partly caused by changes that occur during long-term cultures, such as uneven growth, random fusion, movement, rotation, and necrotic cores. To overcome these problems, we integrated a transparent thin polydimethylsiloxane (PDMS) layer of engineered pillar lattices into commonly used well plates for the generation and immobilization of a large number of uniform-sized spheroids. After the incorporation of automated in situ time-lapse imaging and image processing approaches, our platform allows the real-time tracking of 100 spheroids as well as dynamic infiltration and cytotoxic behaviors of T cells from over 7,000 immobilized tumor spheroids under 30 treatment conditions (SI Appendix, Discussion S3), resulting in 1,040 sets of T cell–tumor interaction images. Further efforts will be needed to integrate robotic medium/drug-handling approaches with our automated microfluidic platform for a fully automated high-throughput screening system.

Machine Learning–Based Analysis of the Vast Amount of Screening Data. Recently, TIL infiltration patterns in patient tumor slides were shown to predict patient survival (5). However, the current evaluation of TIL infiltration mainly relies on a visual inspection and manual qualitative or quantitative analysis by experienced pathologists. Machine learning methods have been used to address the challenges in imaging processing and feature extraction of TIL maps from a vast amount of clinical data, avoiding the inconsistency and bias stemming from human interpretations (16–18). Moreover, simple metrics such as TIL number, infiltration depth, and clustering have been explored individually as predictors of patient survival or immunotherapy outcome, but have not been unified as a comprehensive single “TIL score.” Here, using such clinical TIL map and survival data from the TCGA database, we demonstrated that a machine learning method generated an integrated TIL score that better correlates with patient survival time based on the feature pattern of the TIL map, achieving an excellent median c index of 0.674 and a higher correlation (a coefficient of 0.231) than current methods based on any selected individual parameter, such as TIL clustering index or density (SI Appendix, Fig. S8). Since the distribution of CD8⁺ T cells highly correlated to that of TILs (21, 30), we used this algorithm and an effect threshold to score for the effect of drugs based on the T cell tumor infiltration images. We believe that the incorporation of machine learning methods is crucial for developing automated, nonbiased, high-throughput screening of TIL infiltration. In addition, TIL infiltration could vary based on disease indication, metastatic sites, genetic background, and prior treatments. Thus, we envision this strategy and algorithm could be adapted,
Materials and Methods

Device Fabrication. The pillar-array-device arrays were fabricated by using the standard soft lithography and PDMS fabrication procedure as described in the SI Appendix, SI Methods and Materials.

Cell Cultures. The mouse melanoma cell line B16F10 was purchased from the American Type Culture Collection. The mouse pancreatic tumor cell line UN-16141 was a kind gift from Dr. Surinder K. Batra (University of Nebraska). CAFs were isolated from orthotopic tumors. The OT-I CD8+ T cells were isolated from OT-I mouse spleen. All of the procedures were approved by the Indiana University Institutional Animal Care and Use Committee. The cell isolation and culture followed standard procedures as detailed in the SI Appendix, SI Materials and Methods.

On-Chip Investigation. The hybrid tumor-CAF spheroids were formed using the fabricated devices. The on-chip T cell infiltration of tumor spheroids was tracked and recorded by a Leica SP8 confocal microscope or Olympus OSR spinning disk confocal microscope, and raw images captured were exported as TIFF image stacks in ImageJ. The detailed on-chip investigation of T cell tumor infiltration as described in the SI Appendix, SI Materials and Methods.

TIL Score Analyzer. The TIL scoring deep learning system was developed to score TIL infiltration based on clinical pathology data and patient survival data (https://gdc.cancer.gov/). The network architecture, training, prediction, and validation of this system were detailed in the SI Appendix, SI Materials and Methods.

Screening of Epigenetic Drugs. An epigenetic drug library (Cayman Chemicals, Epigenetics Screening Library, cat. no. 11076) was used to test the intelligent screening system. The TIL score for each drug was determined by using the system. The details of the drug screening process were described in the SI Appendix, SI Materials and Methods.

Validation of Candidate Drug. The antitumor activity of screened top drug candidates was tested in vivo using a syngeneic mouse tumor model. The details of in vivo validation were described in the SI Appendix, SI Materials and Methods.

Data, Materials, and Software Availability. All of the study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We gratefully acknowledge the insightful guidance and detailed input from Prof. Subra Suresh at Nanyang Technological University. F.G. acknowledges the departmental startup funds of Indiana University Bloomington, and NIH awards (R03EB030331, DP2AI160242, R01DK133864, and U01DA056242). M.D. acknowledges support from NIH (R01HL154150). Xiongbin Lu, Xin Lu, and F.G. also acknowledge support from the Indiana Clinical and Translational Sciences Institute (funded, in part by grant no. UL1TR002529 from the NIH National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors thank the Indiana University Imaging Center (NIH1ST00024988) and Indiana University Flow Cytometry core facility and the Indiana University Nanoscale Characterization Facility for use of their instruments.

Author affiliations: “Department of Intelligent Systems Engineering, Indiana University, Bloomington, IN 47405; “Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202; “Melvin and Bren Simon Cancer Center, Indiana University School of Medicine, Indianapolis, IN 46202; “Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611; “Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556; and “Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
Supplementary information

Supplementary Methods and Materials

Microfluidic device fabrication. The PDMS-based pillar-lattice-arrays were designed by using software (AutoCAD) and cast from SU8 mold. The distance of two adjacent cylindrical pillars is spaced identically to 35 μm, which would allow for the unrestricted motility of T cells. Cuboid pillars, which are about 220 μm in diameter, make it easier to peel off the solidified PDMS compared to all the same cylindrical pillars. Standard photolithography techniques were used to fabricate the SU8 mold of the pillar lattices. Briefly, the photoresist (SU8-2100) was evenly coated onto a silicon wafer at a 150-μm thickness (10s at 100 rpm, and 30s at 1800 rpm), followed by processes of the soft bake, UV-light exposure, post-exposure bake, develop and hard bake. The mold was exposed to chlorotrimethylsilane (Sigma, USA) for 5min to facilitate the peeling of the PDMS pillars. Subsequently, the PDMS mixture was poured onto the wafer (Sylgard 184, USA, weight ratio A/B = 10:1) and coated upon SU8 mold by using a spinning coating machine (10s at 100 rpm, 30s at 500 rpm). A thin PDMS layer was obtained and degassed in a vacuum desiccator for 15 minutes. The wafer with the PDMS layer was baked in a conventional oven at 80 °C for 1 hour. After curing, the thin PDMS layer was carefully peeled off and tailored to fit the standard 96-well plate. Finally, PDMS-based pillar-lattice-arrays were autoclaved and placed in an imaging 96-well plate (MatTek). The pillars were then soaked in 1% (w/v) Pluronic 108 solution overnight, sterilized with 70% ethanol for 15 minutes, and washed twice with DMEM culture medium.

Cell isolation and culture.

Cancer cell culture: The mouse melanoma cell line B16F10 was purchased from ATCC. The mouse pancreatic tumor cell line UN-KC6141 was a kind gift from Dr. Surinder K. Batra (University of Nebraska). Both tumor cell lines were cultured in DMEM supplemented with 10% fetal bovine serum (FBS) (Gibco, USA) and penicillin-streptomycin (100 U/mL) (Gibco, USA). Cells were maintained in a 5% CO2 supplemented, 37°C humidified incubator, and passaged at 70% confluency by trypsin-EDTA (Gibco, USA).

Primary CAF culture. Cancer-associated fibroblasts (CAFs) were isolated from orthotopic tumors. Briefly, 200,000 UN-KC6141 or B16F10 tumor cells in 30% (v/v) Matrigel (BD, USA) were injected into the pancreas and subcutaneously (s.c.), respectively, into 4-6 weeks old C57BL6 mice. All procedures were approved by the Indiana University Institutional Animal Care and Use Committee (IACUC). Tumors were allowed to grow for 3 weeks in vivo. After 3 weeks, mice were euthanized to collect xenograft tumor tissue. Tumor tissue was digested by a mouse tumor dissociation kit (Miltenyi, Germany) using a gentleMACS dissociator. Dissociated tumor single-cell suspensions were filtered using a 70 μm cell strainer, resuspended in fibroblast growth medium-2 (FGM-2) (Lonza, USA), and the immediately cultured adherently in a 5% CO2 supplemented, 37°C humidified incubator for 3 weeks to allow for fibroblasts outgrowth. The CAF cells were validated using FAP/alpha-SMA immunofluorescence staining to confirm their identity.
**CD8+ T cell isolation and culture.** OT-I mice [C57BL/6-Tg(TcraTcrb)1100Mjb/J] were purchased from the Jackson Laboratory (JAX) and bred in house. All procedures were approved by the Indiana University Institutional Animal Care and Use Committee (IACUC). The OT-I mice contain Tcra-V2 and Tcrb-V5 transgenes and produce MHC class I-restricted, ovalbumin-specific, CD8+ T cells (OT-I cells). Newborn offspring were genotyped for OT-I T cell receptor (TCR) expression according to protocols provided by JAX. Genotyped newborns of 4-10 weeks old were euthanized to collect spleen. Spleen was grinded using a 30 µm cell strainer. Cells passed through the strainer were then treated with ACK lysis buffer (Gibco) to lyse red blood cells. Then the CD8+ T cells were isolated using a naive CD8a T Cell isolation kit (Miltenyi, Germany). The isolated CD8+ T cells were cultured in DMEM medium supplemented with 10% FBS, 30U /mL rIL-2 (Peprotech, USA) and penicillin-streptomycin (100 U/mL). Dynabeads CD3/CD28 were added to T cell culture at a bead-to-cell ratio of 1:1 to stimulate T cell expansion in vitro.

**On-chip investigation of T cell-tumor infiltration.**

**Hybrid tumor-stroma spheroid formation.** To form hybrid tumor-CAF spheroids, 0.4 million tumor cells (UN-KC6141-OVA or B16F10-OVA) were labeled with membrane DiO dye (Green) (Invitrogen, USA) and seeded on the pillar-lattice-arrays in each well of 96 well imaging plate (MatTek, USA). After 24 h, tumor cells spontaneously aggregated into tumor spheroids. Next, 0.6 million CAF cells were labeled with Dil dye (red) (Invitrogen, USA) and seeded on the pillar-lattice-arrays. The newly seeded CAF cells spontaneously assembled on top of the tumor spheroids to form a tumor core surrounded by a shell of CAF cells, enclosing the tumor core in 12 h (Fig.S1).

**T cell track analysis.** To analyze T cell infiltration tracks, raw images captured on the confocal microscopes were exported as TIFF image stacks in ImageJ. The image stacks were imported into Imaris 9.0 software (Bitplane, Switzerland) for T cell track detection and analysis. T cells were detected by analyzing the blue CMAC tracker channel with a spot radius of 10 μm with background subtraction. Tracking was performed using the Brownian motion model with a maximum displacement of 50 μm and a maximum gap of 1 frame. Tumor spheroids were detected by a surface with a radius of 300 μm. The detected tumor spheroid was analyzed by distance transformation (Matlab plugin) to calculate the distance inside and outside the spheroid surface. The final infiltration depth (Id) was calculated by deducting the “outside distance” from the “inside distance”. We filtered out T cells that interacted with the spheroid by thresholding the maximum T cell track distance of the Id at 0.1 μm. T cell cytotoxicity was analyzed as described previously,(3, 4) briefly dead tumor cells were detected by the staining of SYTOX deep red dye at 0.2 µM, detected in the deep red channel with a spot radius of 10 μm. T cell and dead cell colocalization were detected using a distance threshold of 10 μm.

**Deep-learning-based TIL score analyzer.**

**Data and TIL maps.** Whole-slide images and clinical follow-up data were obtained from TCGA (https://gdc.cancer.gov/) Skin Cutaneous Melanoma (SKCM) projects. The TIL localization inside tumors were extracted from the images of diagnostic H&E-stained sections from the Skin Cutaneous Melanoma as previously report.(5) The TIL map were colored in red and blue representing TILs (or T cells) and tumor. A total of 411 whole-slide images from 397 unique patients were processed to generate
TIL maps for analysis. The TIL maps were sampled with 128 x 128 pixels and resized to 256 x 256 pixels.

**Network architecture and training procedures.** The TIL scoring deep learning system was developed to score TIL infiltration based on clinical pathology data and patient survival data. The TIL scoring algorithm combines elements of the 18-layer residual neural network (Resnet18)(6) with a discrete-time hazards model to predict time-to-event data from images.(7, 8) Image feature extraction is achieved by five groups of convolutional layers arranged in residual blocks. The extracted features are output to a sequence of two fully connected layers containing 1,000, 20 nodes, respectively. The terminal fully connected layer outputs a prediction of the hazard function \( h(\tau_j | x) = \frac{1}{1+\exp[-\phi_j(x)]} \) at the discrete time-point \( \tau_j \) associated with the image, where \( \phi(x) \in R^m \) represents the output of the network to \( m \) (20 in our case) discrete-time point, \( x \) is the input of network corresponding to each image sample. The loss function for backpropagation is the modified cross-entropy function to account for the censored data. The full loss function can be written as follows:(9, 10)

\[
loss = -\frac{1}{n} \sum_{i=1}^{n} \left( d_i \log[\sigma_k(\tau_i)(\phi(x_i))] + (1 - d_i) \log \left[ \sum_{k=\sigma_k(\tau_i)+1}^{m+1} \sigma_k(\phi(x)) \right] \right)
\]

where \( \sigma_k(\tau_i) \) denotes the duration index of individual \( i \)'s event time. As the hazard function was trained, a 3-year survival classification was used to give TIL score.

**Training sampling.** To establish the scoring algorithm that scores TIL pattern based on patient survival time, our deep learning algorithm was fine-tuned on pre-trained Resnet18 for 30 epochs. The stochastic gradient descent (SGD) optimizer was used to minimize the negative log-likelihood via backpropagation to optimize model weights of fully connected layers.(11) Model weights of modified fully connected layers were initialized using the variance scaling method, and a weight decay was applied to the fully connected layers during training (decay rate = 0.1). Models were trained for 30 epochs (1 epoch is one complete cycle through all training samples) using mini-batches consisting of 16 image samples each. Minibatch samples were random assigned at the beginning of each epoch for robustness. During training, a single area was sampled from each slide, and these sampled areas were treated as semi-independent training samples. Each sample was labeled with the corresponding patient survival time for training, duplicating survival time for patients that preserved multiple slides. The areas were sampled randomly at the beginning of each training epoch and regarded as an entirely new set of samples. These sample areas were randomly transformed to improve the robustness to account for tissue orientation, color variations, and anisotropy. The contrast[0.5], brightness[0.5], saturation[0.5] and hue[0.5] of the samples were also randomly transformed using the “colorJitter” PyTorch operations. These sampling and transformation procedures have the effect of augmenting the effective size of the limited clinical labeled training data. Similar augmentation approaches have been widely adapted and shown considerable improvements in neural network performance in a variety of imaging applications.(12)

**Prediction sampling and averaging.** Sampling was also performed to increase the robustness and stability of predictions. (i) Nine areas are first sampled from slides corresponding to each patient. (ii) The survival time expectation (Se) of each sample for each patient is then predicted using the trained
deep learning system. (iii) The median Se of nine predictions from samples is calculated as the final survival time prediction for the patient.

**Validation procedures.** To train and evaluate model performance, patients were randomly assigned to non-overlapping training (80%) and test (20%) sets. If a patient was assigned to one set, then all H&E slides corresponding to that patient were assigned to the corresponding set. This ensured that no slices from a single patient were assigned to both training and testing sets to avoid overfitting and ensure validation accuracy. The randomized assignment of patient datasets was performed 15 times and each of these training/testing sets was employed to train and evaluate a model. Prediction accuracy was measured using a time-dependent concordance index (c index) to measure the concordance between predicted survival and actual survival for testing samples. A c index of 1 indicates perfect concordance between predicted risk and overall survival, and a c index of 0.5 corresponds to random concordance. The following 3-year classification outcomes were also validated by cross validation, and roc curves from the cross validation were drawn to evaluate the classification performance.

**Screening of epigenetic drugs.**

**Epigenetic drug library screening.** To screen effective epigenetic drugs that promote T cell infiltration, tumor-stroma spheroids at day 3 (24 hours post CAF cell addition) were treated with 5 μM of chemicals of an epigenetic drug library (Cayman Chemicals, USA, Epigenetics Screening Library, Catalog# 11076) for 24 hours. Post epigenetic drug treatment, culture medium was refreshed, and OT-I T cells, labeled with blue CMAC tracker dye (Invitrogen, USA), were added at 50,000 cells per well. Additionally, anti-mouse-PD-1 antibody (clone RMP1-14, Bioxcell, USA) was added at a concentration of 3.5 μg/mL together with OT-I T cells. SYTOX deep red nucleic acid stain was loaded at a final concentration of 0.2 μM to visualize cell death as it stains newly killed cells’ nuclei once the cell membrane integrity is compromised. The plate was then set up for time-lapse imaging on a Leica SP8 confocal microscope or Olympus OSR spinning disk confocal microscope for 12 h in a 5% CO2 supplemented, 37°C humidified incubation chamber. Confocal images were taken every 15 min with 4 channels.

**Drug screening result TIL scoring.** The final images of T cell distributions in tumor-stroma spheroids after 24 hours of screening were analyzed for TIL scores by our deep learning algorithm. For each treatment, ten spheroids’ images were collected and used for TIL scoring. Each screening result images was given a TIL Score of “0” when the TIL infiltration pattern was classified as “hazard”; or a TIL Score of “1” when it was classified as “non-hazard”. The final infiltration score for each drug, calculated by averaging the Score from 10 screening results from the same drug treatment condition.

**Validation of selected drug candidates.**

**In vivo validation of GSK-LSD1.** To validate the effect of the anti-tumor activity of GSK-LSD1 in vivo, 200,000 B16F10 cells were injected subcutaneously (s.c.) into 40 C57BL6 mice of 5 weeks age (Envigo). On day 7, mice were randomized into 4 groups to have similar starting average tumor size (p>0.05), with 10 mice per group. Each animal was injected intraperitoneal (i.p.) with 200 μg anti-PD-1 (clone RMP1-14, Bioxcell, USA) and/or GSK-LSD1 (20mg/kg) every other day from day 8 to day 17. Tumor size was measured every day and calculated as tumor volume = (length*width²)/2.
Spheroid and tumor immunofluorescence staining. To visualize infiltrated T cells, the spheroid and tumor immunofluorescence staining were conducted using our previously developed protocols.(13-16) Tumor spheroids or tumor tissues were fixed with 4% PBS buffered formalin (Sigma, USA) overnight then transferred to 70% ethanol for dehydration. The dehydrated spheroid/tissues were then embedded in paraffin and sectioned at 30 μm thickness and mounted onto slides. The mounted tissue sections were then treated for antigen retrieval by heat-mediated antigen retrieval and stained with Alexa 594 labelled anti-CD8 antibody (Cat# 100758, Biolegend, USA) to identify infiltrating T cells. Stained slides were visualized using an Olympus epi-fluorescence microscope.

Statistical analyses. The c indices generated by Monte Carlo cross-validation were performed using the Wilcoxon signed-rank test. This paired test was chosen because each method was evaluated using identical training/testing sets. Statistical analysis of Kaplan–Meier plots was performed using the Mantel-Cox log-rank test. Other statistical analysis comparing 2 groups was performed using the student's t-tests, comparing 3 or more groups was performed using one-way ANOVA. P-value was denoted as following: *p<0.05, **p<0.01, ***p<0.005, ****p<0.001.

Supplementary Discussion

Discussion. S1 Impact of CAF layer and tumor antigen on T cell infiltration and killing. Tumor stroma was reported to prohibit T cell infiltration via physical ECM barrier and secretory factors (e.g., chemokines, cytokines, and microRNA), and tumor with thicker stroma such as pancreatic cancer tends to permit less immune infiltration.(17) However, it has been not systematically studied how tumor stromal components affect T cell penetration quantitatively. Thus, using our microfluidic platform, we fabricated tumor spheroids with stromal “shell” with tunable thickness by tuning tumor cell to CAF ratio. We found that CAF can effectively prohibit T cell infiltration and killing (Fig. 2b), Moreover, higher CAF composition resulted in less T cell infiltration, as well as reduced T cell killing. Interestingly, although fewer T cells can enter the spheroid in the higher CAF ratio conditions, the ones that did successfully enter the spheroid did not halt their infiltration until they reached the tumor core/CAF shell boundary, which still could perform killing (Fig. S2), suggesting the physical barrier created by CAF has the limited effect to prohibit T cell infiltration, which is consistent to that observed in patient tumor sections(18). Another key factor to impact T cell infiltration is the tumor antigen presentation, as acquired ICI resistance can be partially attributed to antigen loss/downregulation of antigen presentation molecules as seen from patients’ pathology slides.(19) To study this phenomenon, we also tuned antigen-positive tumor cell ratio in our model to observe its impact on T cell infiltration. We mixed OVA+ UN-KC6141 cells and OVA- wild type UN-KC6141 cells at a ratio of 0, 25, 50, and 100%. We observed that T cell has minimum infiltration inside the 0 and 25% low OVA+ cell ratio spheroids, whereas T cell infiltration number and depth significantly increased in the 50% and 100% OVA+ group, indicating tumor antigen presentation is another key factor in T cell infiltration (Fig. S3). Surprisingly, the 50% OVA+ spheroids permit a slightly higher number and deeper T cell infiltration as compared with 100% OVA+ spheroids, despite more killing are seen with 100% OVA+ spheroids. This is likely due to the frequent pauses of T cells at the surface of 100% OVA+ spheroids to perform killing, limiting
its penetration (Fig. S3). This observation also highlights the importance of studying T cell infiltration as an independent event, as it does not always correlate positively with antigen presentation or T cells’ capacity to kill.

**Discussion.** S2 Development of the deep-learning-based TIL score analyzer. We trained the deep-learning-based TIL score analyzer using clinical data of pathology H&E images and survival data from 411 slides, 397 cases from The Cancer Genome Atlas (TCGA) Skin Cutaneous Melanoma (SKCM) projects. The TIL score is to best assign higher TIL scores to patients with better survival. Although cancer patient survival can be attributed to many factors such as treatment history and molecular status of the tumors in addition to TIL infiltration, abundant evidence have shown that patients’ survival correlate well with TIL scores. Thus, we attempted to develop the analyzer that projects TIL maps (e.g., TIL number, the relative depth, distribution, and clusters of TILs within tumors) to the patient survival. The core element of the deep learning algorithm consisted of convolutional neural network (CNN) modules, image sampling, and risk filtering to improve prediction accuracy and stability (Fig.S4). The TILs and tumor were first extracted from H&E-stained tissue sections to digitized images colored in red (TILs) and blue (tumor). These digitized images of TIL/tumor distribution were then used to train a deep convolutional network integrated with a discrete-time hazards model for predicting outcomes. The network consisted of interconnected residual blocks and nonlinear functions to transfer the images to highly predictive prognostic features. Fully connected layers perform additional transformations on these image-derived features, and a discrete-time hazard layer generates a prediction of the survival likelihood distribution, and then a SoftMax layer outputs the classification of 3-year survival potential. To improve the deep learning performance, a sampling and risk filtering technique was adapted to address intra-tumoral heterogeneity and limited clinical samples. In the training, new samples were randomly sampled from each slide image at the start of each training iteration, providing the CNN with a fresh look at each patient’s TIL distribution and capturing heterogeneity within the slides. Each sample is preprocessed using data augmentation techniques that randomly transform the images to reinforce network robustness to tumor orientation and variations. For prospective prediction, we first took multiple samples within each slide to generate a representative batch of fields for each patient, and then, risk output from these samples were sorted and filtered to predict a more robust patient-level risk that reflects the aggressiveness of their disease. These sampling and filtering procedures were described in detail in Methods. The prognostic accuracy of our deep learning algorithm was assessed using Monte Carlo cross-validation. We randomly split our cohort into paired training (80%) and testing (20%) sets to generate 15 training/testing set pairs. We trained the model using each training set and then, evaluated the prognostic accuracy of these models on the paired testing sets, generating a total of 15 accuracy measurements. Using the sampling and filtering techniques, our deep learning system reached a median c index of 0.674 to predict the survival probability distribution. To access the performance of the classifier, we employed 10 times cross-validation to test the sensitivity and specificity. And the 3-year survival classification obtained from the survival probability to score the TIL patterns reached a median AUC of 0.8051 (Fig.4Sc), where AUC evaluates the performance of the classifier globally, an AUC of 0.8-0.9 means a good classification. Compared with current methods, there are two main improvements we’ve made: (1) Whereas the previous reference calculated “Banfield Raftery” index (“count of TIL clusters”) or
“Ball Hall” index (“cluster extend”) from the extracted TIL map and analyzed the correlation of these 2 indexes with patient survival, we further developed this by directly inputting these TIL maps into a deep convolutional neural network and trained it to associate the TIL patterns with discrete patient survival time to generate a TIL score. This TIL score correlated better with patient survival as compared to any single parameter in the previous study including the “Banfield Raftery” index or “Ball Hall” index alone (Fig.S5). (2) This scoring algorithm in combination with our microfluidic platform allowed us to apply this to evaluate our in vitro screening TIL pattern images to find epigenetic drugs that resulted in better TIL infiltration patterns. Thus, with minimal pre-processing, we could directly input our TIL pattern images into the trained convolutional neural network and obtain the corresponding TIL scores.

**Discussion. S3 Throughput of our system.** Our platform can fabricate 80 spheroids per well in a 96-well plate format. In a typical drug screening experiment using 20 plates, we could achieve a throughput of 153,600 spheroids per screen in combination with a high-content imager. For T cell infiltration analysis, the throughput is limited by an additional factor: the limitation of high-speed imaging. In this study, our strategy was to perform an initial T-cell infiltration screening using a fast-scanning Olympus OSR spinning disk microscope allowing for 2-color (T cells: Green and tumor-CAF spheroids: Yellow) imaging of 100 positions with a 15-minutes imaging interval for 12 hours. This enabled high-dimensional time-series data recording of single T cell infiltration dynamics. Top candidates were then selected and validated with detailed interrogation of T cell infiltration and killing dynamics in detail using the 4-color (T-cell, CAF shell, tumor core, cell death indicator dye) Leica SP8 confocal microscope. Due to the limited scanning speed of our Leica SP8 confocal microscope set-up, we were only able to perform a detailed validation of 30 spheroids at a time with a 15-minutes imaging interval. Moreover, we only used the 12-hour end-point T cell infiltration position map for drug scoring in our experiments. In the future, the deep-learning-guided system can be developed to eliminate the prolonged time-lapse recording, and scan all 7,680 spheroids at the end, allowing for screening of 768 agents with 10 spheroids per drug or 1,536 agents with 5 spheroids per drug, greatly expand the screening throughput.
Supplementary Figures

Fig. S1 Microfluidic fabrication of “core/tumor-shell/stroma” spheroid array. a. Formation of "core/tumor-shell/stroma" spheroids by sequential addition of tumor and CAF cells. b. Tumor spheroids’ viability over the prolonged culture period. b. Tunable CAF layer by changing tumor cell UN-KC6141 and CAF ratio. c. Tunable tumor/CAF ratio resulted in variable stromal thickness in "core/tumor-shell/stroma" spheroids. d. Controlled initial seeding of ~9,000 cells per spheroid with pure 9,000 UN-KC6141 cells or 3,000 UN-KC6141 cells + 6,000 CAF cells resulted in differential spheroid growth. e. Distinct speed and track straightness was observed in killing and non-killing T cells. Speed and track straightness was calculated in Imaris software based on the Euclidean distance between T cell position displacement in consecutive imaging frames. Scale bar: 200 µm.
Fig. S2 Impact of the tunable stromal layer of tumor spheroids on T cell tumor infiltration and killing behavior. **a.** The "core/tumor-shell/stroma" spheroids with tunable CAF layer thickness were subjected to T cell infiltration (tumor cells: green, CAF cells: yellow). **b.** Tunable stromal layer thickness in the various tumor:CAF ratio conditions. **c.** T cell infiltration depth over time in heterotypic spheroids with 1:1.5, 1:3, and 1:5 tumor to CAF ratio. **d.** Total infiltrated T cell number and infiltration depth distribution in all 3 types of spheroids with various CAF ratios. **e.** Total killing events and killing depth distribution in all 3 types of tumor spheroids with various CAF ratios. Swarm plot width is scaled by event numbers. Scale bar: 50 µm.
Fig. S3 Impact of tumor antigen presentation on T cell tumor infiltration and killing behavior. a. Tuning OVA+ versus OVA- cell ratio in "core/tumor-shell/stroma" spheroids to makeup OVA+ cell ratio as 0%, 25%, 50%, and 100%. b. T cell infiltration depth over time inside all four different types of spheroids with various OVA+ cell ratio. c. Total infiltrated T cell number and infiltration depth distribution in all 4 types of spheroids. 100% OVA+ spheroids have less infiltrating T cells likely due to the frequent pausing and killing of T cells during the 12-hour imaging period. d. Total killing events and killing depth distribution in all 4 types of spheroids. 100% OVA+ spheroids have less infiltrated T and infiltration depth yet more killing events as compared with 50% OVA+ group. Swarm plot width is scaled by event numbers. Scale bar: 30 µm.
Fig. S4 The deep-learning-based TIL scoring system. a. Detailed diagram of the architecture. The architecture is a variation of the Resnet18 network with modified output layers. b. Representative training curves of training loss and validation loss. c. The cross-validated time-dependent ROC curve was generated by the classifier for 3-year survival predictions with an AUC of 0.8051. d. Predicted TIL score of different drugs (n=10).
Fig. S5 Deep learning TIL scoring method is better than other methods based on single TIL parameter. a. Spearman’s rank correlation coefficient of deep-learning-based TIL scores and other single T cell infiltration parameters with patients’ overall survival (OS). b. Comparison of deep learning TIL scores spearman’s rank correlation coefficient with other T cell single-parameter based readouts from TCGA skin cancer pathology slides to predict patients’ overall survival.
Fig. S6 Cytotoxicity screening of epigenetic drugs. UN-KC6141 spheroids and B16F10 spheroids were treated with epigenetic drug at 5 µM for 24 hours and cytotoxicity was measured by CCK-8 assay. Drugs without cytotoxicity (cell viability>80% and p-value>0.05) highlighted by green were chosen for T cell infiltration screening.
Fig. S7 Drug cores of epigenetic drug library in "core/tumor-shell/stroma" spheroids (B16F10 and UN-KC6141). Negative control refers to a score of OVA negative cell spheroids against T cells (OVA-specific CD8+ T cells). No treatment control refers to OVA positive cell spheroids against T cells (OVA-specific CD8+ T cells).
Fig. S8 Flow cytometry gating strategy for quantification of tumor infiltrating CD8+ T cells. The total tumor cell population was first gated for FSC-H and FSC-A for single cells, then CD45+ cells were selected and further gated into CD4+ and CD8+ T cells.
Supplementary Movies

Movie. S1 Tracking of simultaneous T cell infiltration and killing. An individual T cell’s killing of two tumor cells is recorded as the T cell (green) infiltrates into "core/tumor-shell/stroma" spheroids (yellow surface) and kills (red). Track color scale: timestamped.

Movie. S2 Effect of treatments on T cell infiltration. T cell infiltration and killing of tumor cells inside the "core/tumor(cyan)-shell/stroma(yellow)" spheroids. Videos show different T cell behaviors under control (no treatment), anti-PD1 treatment, GSK-LSD1 treatment, and combo (anti-PD1+ GSK-LSD1) treatments. Track color scale: timestamped.
References

1. Y. Wang et al., In situ differentiation and generation of functional liver organoids from human iPSCs in a 3D perfusable chip system. Lab on a Chip 18, 3606-3616 (2018).


