Cell Fate Simulation Reveals Cancer Cell Features in the Tumor Microenvironment

Sachiko Sato[1], Ann Rancourt[1,2], and Masahiko S. Satoh[2]*

[1]Glycobiology and Bioimaging Laboratory of Research Center for Infectious Diseases and Axe of Infectious and Immunological Diseases, Research Centre of CHU de Quebec, Faculty of Medicine, Laval University and [2]Laboratory of DNA Damage Responses and Bioimaging, Research Centre of CHU de Quebec, Faculty of Medicine, Laval University, 2705 Boulevard Laurier, Quebec, G1V 4G2, Canada

*Correspondence to Masahiko S. Satoh

TEL: 1-418-525-4444 ext. 47340

FAX: 1-418-654-2159

e-mail: masahiko.sato@crchudequebec.ulaval.ca

Abstract

The characteristics of individual cancer cells within a population residing in the tumor microenvironment (TME) evolve over time, and these changes can significantly influence the fate of cancer. However, a research approach that provides insight into the spatiotemporal dynamics of these attributes and the TME status is currently lacking. To address this gap, we developed a pipeline by integrating a computerized single-cell tracking system, cell fate simulation, and a three-dimensional TME (3D TME) simulation, revealing distinct behaviors of cancer cells within the TME. Our study began by examining the spatiotemporal behaviors of individual cancer cells through single-cell tracking, focusing specifically on the α2-6 sialic acid (2-6Sia) modification of glycans, known for its association with cellular stemness and malignancy. Our findings indicate that pancreatic cancer-derived cells exhibit elevated 2-6Sia modification, correlating with an enhanced reproductive capacity. In contrast, only a minority of cervical cancer-derived cells display elevated 2-6Sia expression. We then generated inferred cell populations using tracking data and a cell fate simulation algorithm, varying levels of 2-6Sia expression, to further elucidate the impact of 2-6Sia expression on the fate of each cancer cell. Subsequently, we conducted a 3D TME simulation employing these cell populations and the immune cell landscape characteristic of 193 cervical and 172 pancreatic cancers. Our results suggest that pancreatic cancer cells are less influenced by the immune cell landscape within the TME compared to cervical cancer cells, highlighting the predominant role of intrinsic attributes, such as 2-6Sia expression, in determining the fate of cancer cells within the TME. Importantly, our approach has the potential to uncover novel and previously unexplored spatiotemporal dynamic behaviors of cancer cells within the TME.

Introduction

Cancer cell populations exhibit a range of phenotypical characteristics stemming from genetic sequence alterations and chromosomal variations (*1-8*). Notably, missense mutations in p53 are prevalent in over 50% of human cancer cases (*9-15*). Aneuploidy, characterized by abnormal chromosome numbers, can arise through various mechanisms, including DNA damage, impaired M phase checkpoints (*16, 17*), and multipolar cell division, frequently observed in cancer tissues (*18-20*). Furthermore, cancer cells often retain stemness-like properties, albeit to varying degrees within a cancer cell population (*21-24*). As a result, the phenotypical diversity within cancer cell populations accumulates during cancer development.

These heterogeneous cancer cell populations coexist within the tumor microenvironment (TME), which comprises various cell types such as infiltrated immune cells, tumor-associated fibroblasts, endothelial cells, and cancer cells (*25-28*). Together, they form a complex ecosystem characterized by conditions like hypoxia and abnormal levels of cytokines, growth factors, and metabolites (*25-28*). The Cancer Genome Atlas Program has provided insights into the TME's cellular composition in primary tumors through genomic, epigenetic, and protein-level molecular profiling (*29-33*). This landscape may represent a specific stage in the TME's evolution within the ongoing process of cancer development, although the landscape and the attributes of cancer cells in the TME could evolve, influencing the fate of cancer. However, a research approach that could shed light on the spatiotemporal changes of the attributes and the landscape is still missing.

Thus to gain insight into the spatiotemporal aspects of cancer cell evolution and its interlay with immune cells residing in TME, we developed a three-dimensional TME (3D TME) simulation utilizing Artificial Intelligence (AI). This simulation allows us to position cancer and immune cells, facilitating the investigation of cancer cell fate within the TME while considering the spatiotemporal aspects of their interactions. Our approach commenced with the acquisition of spatiotemporal data on individual cell behavior through empirical single-cell tracking of established cultured cancer cell lines (*20, 34*). To

generate single-cell tracking data suitable for bioinformatics analysis, we developed a computerized single-cell tracking system. This involved long-term live cell imaging, video generation, individual cell tracking, and the creation of cell lineage data. Our emphasis was particularly on aspects related to stemness, specifically those associated with oligosaccharides, notably the α2-6 sialic acid structure (2-6Sia) (*35-40*), and multipolar cell division, leading to the generation of cells with altered chromosome numbers (*18-20*). Our findings revealed that pancreatic cancer-derived cells exhibit elevated 2-6Sia expression, which correlates with enhanced reproductive capacity. In contrast, only a small subset of cervical cancer-derived cells displays elevated 2-6Sia expression, underscoring their distinct characteristics compared to pancreatic cancer cells. Subsequently, based on the single-cell data of cancer cell lines, we applied a cell fate simulation algorithm (*20*) to generate inferred cell populations. *In vivo*, cancer cells express varying levels of 2-6Sia (*41-44*). These inferred cell populations retain attributes of cancer cells composing cancer cell lines but signify specific 2-6Sia expression levels of cancer cells *in vivo*. These cells were then introduced into the 3D TME alongside immune cells, allowing us to simulate the fate of each cancer cell within the TME. We conducted a 3D TME simulation using immune cell landscapes characteristic of 193 cervical and 172 pancreatic cancers. Our results indicate that the fate of pancreatic cancer cells is less influenced by the immune cell landscape within the TME, whereas the fate of cervical cancer cells is notably impacted by the landscape. These findings suggest that the inherent attributes of cancer cells play a significant role in determining their fate within the TME, and the expression of 2-6Sia may be linked to the prognosis of these cancers. These results also suggest that our research approach can be used to gain insight into the spatiotemporal changes of the attributes and TME status, having the potential to simulate cancer cell fate in individual cancer cases.

Results

Overview of single-cell tracking and 3D TME simulation

The cancer cell population comprises cells with diverse phenotypic characteristics that evolve over time (Fig. 1A). The fate of these cells is influenced by cells residing in the TME, including immune cells. These immune cells can either suppress the growth of cancer cells or create a permissive microenvironment for the proliferation of cancer cells (*25-28*). To simulate the fate of cancer cells within the TME, we required detailed information about the spatiotemporal behavior of individual cancer cells.

To accomplish this, we conducted single-cell tracking analysis (*20, 34*) on established cancer cell lines (Fig. 1B). In order to generate data suitable for bioinformatic analysis, we devised a computerized single-cell tracking system (Supplementary Fig. 1). This procedure involved live cell imaging using near-infrared-illuminated Differential Interference Contrast (NER-DIC) microscopy (*20, 34*), which allowed us to create grayscale live cell imaging videos without introducing any phototoxicity. Then, we developed an algorithm to segment the grayscale DIC images with a wide range of cell densities, from sparse to densely populated populations (Supplementary Fig. 2A and B). This approach allowed us to track individual cell behaviors within these videos. Subsequently, automated cell tracking was executed using an algorithm that analyzes the surrounding areas of the tracking-target region. It identifies the tracking-target area in the subsequent image frame by considering information from the surrounding areas (Supplementary Fig. 3). Through this tracking process, various cellular events, including multipolar cell division, can be detected. This results in the establishment of accurate cell lineage data, including the spatiotemporal behaviors of a cell that commence with the tracking process and encompass all of its progeny (Fig. 1C). After live cell imaging, cells can be labeled with specific antibodies or lectins targeting biomolecules or post-translational modifications, such as oligosaccharides bound to membrane proteins, for fluorescent

imaging. The quantified fluorescent values associated with each cell were then integrated into the cell lineage data (Fig. 1C).

While cancer cell lines are derived from a single cell within the original cancer cell population, they may only partially represent the characteristics of the original population. To address this limitation, we employed a cell fate simulation algorithm (*20*) (Fig. 1B and D). This algorithm calculates the probabilistic likelihood of a specific cellular event occurring subsequent to a preceding event, considering the duration between these events (*20*). For instance, it can assess the likelihood of multipolar cell division occurring after bipolar cell division, along with the associated duration. Consequently, the algorithm generates cell lineage data based on these probabilistic values. The growth dynamics of each cell population can be defined by combining these probabilistic values. Therefore, the cell fate simulation algorithm has the capability to simulate the growth of various cell populations when the requisite information to define these probabilistic values is available. For example, if we possess knowledge that the expression of 2-6Sia is notably high in highly reproductive cells, the algorithm can generate an inferred cell population in which reproductive cells exhibit a high level of 2-6Sia expression. By leveraging empirical single-cell tracking data generated from established cell lines in combination with information regarding cells in their *in vivo* context, an inferred cell population that can reflect the attributes of growth patterns of cells *in vivo* can be generated. Our cell fate simulation algorithm generates diverse inferred cell populations, thereby encompassing the heterogeneity observed among cancer cells, specifically regarding 2-6 Sia expression in cancer tissue (*41-44*). Subsequently, this inferred cell population was introduced into a 3D TME (Fig. 1B and E) alongside immune cells, allowing us to simulate the fate of these cells within the TME.

Distinct Characteristics of Cell Lines Derived from Cervical and Pancreatic Cancer

We conducted single-cell tracking using HeLa (cervical cancer) and MiaPaCa2 (pancreatic cancer) cell lines. The cell population size at each time point was determined through single-cell

tracking data (Fig. 2A). Additionally, Supplementary Data 1, Supplementary Data 2, Supplementary Video 1, and Supplementary Video 2, provide cell lineage maps and videos illustrating the tracking processes. Our results reveal an increase in the population size of HeLa and MiaPaCa2 cells, with average cell doubling times of 26.6 and 22.7 hours, respectively.

Next, we assessed the occurrence of multipolar cell divisions, cell death, and cell fusion in HeLa and MiaPaCa2 cells (Fig. 2B). Notably, MiaPaCa2 cells exhibited a lower frequency of multipolar cell divisions, cell death, and cell fusion compared to HeLa cells. Subsequently, we delved into the cellular events following multipolar cell divisions and cell fusion. In Fig. 2C, we have listed the patterns we analyzed. For instance, 'CF (cell fusion) > BD (bipolar cell division)' illustrates that bipolar cell division occurs following cell fusion. In the case of HeLa cells, over 50% of the progeny arising from multipolar cell division (MD) and cells resulting from cell fusion ultimately underwent cell death (CD)(Fig. 2D and E, HeLa). In contrast, among MiaPaCa2 cells generated through multipolar cell division (MD) or cell fusion (CF), more than 50% retained the capability to undergo bipolar cell division (BD)(Fig. 2D and E, MiaPaCa2). This difference between HeLa and MiaPaCa2 cells suggests that while MiaPaCa2 cells experience multipolar cell division and cell fusion less frequently, the progeny resulting from these events have a higher probability of survival compared to HeLa cells (Fig. 2F, MD+CF > BD, HeLa vs. MiaPaCa2). Consequently, MiaPaCa2 cells have a greater chance of accumulating cells with altered chromosome numbers due to multipolar cell division and cell fusion, potentially comprising 3.2% of the total cell population of MiaPaCa2 cells after 170 hours of growth (Fig. 2G and H).

Expression of 2-6Sia in Cervical and Pancreatic Cancer Cell Lines

HeLa and MiaPaCa2 cells were subjected to immunostaining for the cancer marker TAG-72 (*45-47*), the stem cell marker CD133 (*48-50*), and the stem cell glycan marker fucose α1-2galactose α1-3-structure (*40*), recognized by the rBC2LCN bacterial lectin (Fig. 3A and B, respectively). The

majority of HeLa cells exhibited broad expression of these markers, with most MiaPaCa2 cells also displaying these markers, except for the rBC2LCN lectin staining.

The oligosaccharide structure attached to membrane proteins and terminated by α2-6 sialic acid structure, recognized by *Sambucus Nigra* lectin (Fig. 3C; SNA1), is suggested to be associated with the maintenance of stemness features and malignancy (*35-40, 43*). Interestingly, only a subset of HeLa cells expressed 2-6Sia (Fig. 3A, SNA1 staining), indicating that 2-6Sia-related stemness differs from other markers, such as CD133. Conversely, the majority of MiaPaCa2 cells displayed positive 2-6Sia expression (Fig. 3B, SNA1 staining). We employed single-cell tracking data linked to the SNA1 binding levels of individual cells to create a visual heatmap representation of these levels. In this visualization (Fig. 3D), individual cells were positioned on the surface of a sphere, illustrating that MiaPac2 cells exhibit elevated levels of 2-6Sia-related stemness compared to HeLa cells.

In Fig. 3E (HeLa) and F (MiaPaCa2), we plotted the total number of cells in a cell lineage on the horizontal axis, where a higher value represents a cell's ability to produce a larger number of progeny (indicative of reproductive ability), and the total SNA1 binding levels of progeny within the cell lineage on the vertical axis. Consequently, if a cell lineage comprises a large number of cells, all expressing high levels of 2-6Sia, it will be positioned towards the upper-right quadrant of the graph. Notably, only a subset of the HeLa cell population expressed 2-6Sia, and these cells exhibited a higher reproductive ability, while 2-6Sia-negative cells showed varied reproductive ability (Fig. 3E). In contrast, the majority of MiaPaCa2 cells expressed 2-6Sia, and the levels of expression displayed a linear relationship with their reproductive ability (Fig. 3F). We also conducted similar analyses by plotting the cell death count (Fig. 3G for HeLa and Fig. 3H for MiaPaCa2), multipolar cell division count (Fig. 3I for HeLa and Fig. 3J for MiaPaCa2), and cell fusion count (Fig. 3K for HeLa and Fig. 3L for MiaPaCa2). In all instances, it was observed that cell death, multipolar cell division, and cell fusion predominantly occurred in cells exhibiting lower 2-6Sia expression levels (below 2000 of total SNA1 binding). These findings not only underscore the connection between 2-6Sia expression and the

maintenance of cellular reproductive ability but also highlight its role in reducing the likelihood of cell death, multipolar cell division, and cell fusion. Taken together, these results suggest that only a subset of HeLa cells expressing 2-6Sia retain their stemness, contributing to a higher reproductive ability to maintain the cell population (Fig. 3E), consistent with our previous observation (*34*). The presence of 2-6Sia-negative cells increases the likelihood of cell death, multipolar cell division, and cell fusion (Fig. 3G, I, and K). In contrast, the majority of MiaPaCa2 cells maintain 2-6Sia-related stemness (Fig. 3F), resulting in fewer occurrences of cell death, multipolar cell division, and cell fusion (Fig. 3H, J, and L). This implies that 2-6Sia-related stemness plays a role in preserving the integrity of the cell population.

Generation of Inferred Cell Populations

Our findings thus far indicate distinct patterns of 2-6Sia expression in HeLa and MiaPaCa2 cells. This expression could be related to the reproductive capacity of these cells and plays a role in maintaining the integrity of the cell population. Specifically, it reduces the occurrence of cell fusion and multipolar cell division, events that can lead to the production of cells with altered chromosome numbers. In cancer tissues, it has been observed that the expression levels of 2-6Sia between cancer cells vary (*41-44*), implying that each cancer cell within a tissue may possess a unique reproductive ability and a chance to produce cells with altered chromosome numbers. This variation falls within the range of 10 to 200% of the average expression levels (*41-44*). To explore the effects of varying levels of 2-6Sia expression on cancer cell fate, we thus created inferred cell populations with 1.5, 1.0, 0.5, and 0.25-fold 2-6Sia expression relative to the levels found in HeLa and MiaPaCa2 cells. We accomplished this using the cell fate simulation algorithm (*20*) with some modifications. An overview of the inferred cell population generation is illustrated in Fig. 4A, with additional details provided in Supplementary Fig. 4 and Materials and Methods.

The initial step involves creating a 2-6Sia expression list (Fig. 4A and Supplementary Fig. 4A), which encompasses the expression levels of all analyzed HeLa or MiaPaCa2 cells obtained through single-cell tracking. This list serves as a reference when assigning 2-6Sia expression levels to individual cells. Subsequently, the cell fate simulation algorithm was used to generate inferred cell populations (*20*). The first step within the generation is to determine the length of time progenitor cells spend before experiencing their first event (Supplementary Fig. 4B). Since each cell is in a different stage of the cell cycle, this length varies. The algorithm then assigns a cellular event to the progenitor cell. In the case of bipolar cell division, the algorithm generates two daughter cells. Next, 2-6Sia expression levels are assigned to each daughter cell using the 2-6Sia expression list (Supplementary Fig. 4C). This is achieved by generating a random value to select one of the expression values stored in the list. For HeLa cells, where only a subset expresses 2-6Sia, the majority of values in the list are expected to be zero, resulting in an inferred cell population that mirrors HeLa cells' 2-6Sia expression pattern. Similarly, an inferred cell population resembling the 2-6Sia expression pattern of MiaPaCa2 cells can be generated. By assigning 2-6Sia expression to inferred cells with similar cell doubling time as HeLa or MiaPaCa2 cells, the 2-6Sia expression can be assigned to cells in a manner that reflects their reproductive capacity. Finally, 2-6Sia expression levels were modified by factors of 1.5, 1.0, 0.5, and 0.25. Although the expression levels may exceed the highest 2-6Sia expression levels of HeLa or MiaPaCa2 cells, they were capped at the highest observed value. We started with 500 progenitor cells resulting in the creation of 500 cell lineages per inferred cell populations. These generated cell populations are designated as follows: Cervical 2-6Sia 1.5, 1.0, 0.5, and 0.25, and Pancreatic 2-6Sia 1.5, 1.0, 0.5, and 0.25. Cell lineage maps of Cervical 2-6Sia 1.5 and Pancreatic 2-6Sia 1.5 were provided in Supplementary Data 3 and Supplementary Data 4, respectively.

Fig. 4B (HeLa) and Fig. 4C (MiaPaCa2) depict the cell population expansion curves, which include empirical tracking results (light blue line) and inferred cell populations (blue lines). Inferred cell populations generated using single-cell tracking data of HeLa cells show a ~5% slower rate of

population expansion compared to actual HeLa cells, as the cell fate simulation algorithm tends to reflect the averaged proliferation pattern of the given cell population. In Fig. 4D, the levels of 2-6Sia in the inferred cell populations are presented. The 2-6Sia 1.0 cells correspond to cells generated without modifying the 2-6Sia 1.0 expression levels. By applying 1.5-fold, 0.5-fold, or 0.25-fold modifications, inferred cell populations with modified 2-6Sia expression levels were generated and the levels of 2-6Sia expression reflected the expression patterns of HeLa and MiaPaCa2 cells (Fig. 4D).

Designing the 3D TME Simulation

We developed a 3D TME simulation to investigate the spatiotemporal dynamics of cancer cells within the TME. In pursuit of this goal, we initially identified key parameters that define the status of cancer cells within the TME: [1] Attributes of each cancer cell and their spatiotemporal changes, [2] The TME cell landscape and its dynamic changes over time, [3] The mode of interaction between cancer cells and immune cells, and [4] The spatial location of cancer cells and immune cells within the TME. In our study, we focused on specific aspects and simplified certain elements due to the unavailability of complete information. As for [1], our focus was on 2-6Sia expression and the progeny resulting from multipolar cell division. Regarding [2], we recognized that available TME cell landscape information represents a specific moment in cancer development, but it is unclear how the landscape changes over time. Thus, we assumed that the immune cell landscape remains constant during the 3D TME simulation. Concerning [3], we considered the impact of the microenvironment on cancer cells as simulation variables, acknowledging that cancer cells can respond differently to various microenvironments. Finally, due to challenges in replicating the precise spatial distribution of cancer and immune cells within the TME using computer-generated 3D space, we randomly positioned these cells within a 3D space (sphere). In summary, we randomly placed inferred cancer cells and immune cells within a sphere, where cancer cells expressed various levels of 2-6Sia.

Categorization of TME-Resident Cells for 3D TME Simulation

To streamline the development of the 3D TME simulation algorithm, we categorized TME-resident cells into three groups: Suppressive, Permissive, and Lethal (Fig. 5A). Suppressive cells encompass $CD4^+$ T cells, $CD8^+$ T cells, and B cells. Permissive cells include macrophage M2, regulatory T cells, and endothelial cells, all of which contribute to creating a microenvironment conducive to cancer cell proliferation. Lethal cells consist of cytotoxic T cells and natural killer cells. Some cell types, such as M1 macrophages and neutrophils, can exhibit dual roles, either suppressing or promoting cancer cell growth depending on the context (*25-28*). Due to this complexity, we chose not to include these cells in the 3D TME simulation.

In the 3D TME, these cells coexisted with inferred cancer cells. Because the fate of each inferred cancer cell had already been predetermined during the population generation using the cell fate simulation algorithm, the proliferation of the inferred cancer cells followed the cell lineage data associated with each cell. Consequently, when an inferred cancer cell encounters a Suppressive, Permissive, or Lethal cell, the impact of these encounters on cancer cells is enacted by modifying the cell lineage data. In cases where a cancer cell encounters a Suppressive cell, both the cell's doubling time and that of its progeny were extended to reduce the proliferation rate of these cells (Fig. 5A). When a cancer cell encounters Permissive cells, the cell's doubling time was shortened, leading to an increased rate of cell proliferation, although the minimum cell doubling time was set to match the shortest cell doubling time within the given cell population. In instances where the cell encounters a Lethal cell, the cell is eliminated through cell death, and all its progeny are removed from the cell lineage data.

The 3D TME Simulation Processes

To perform the 3D TME simulation, several parameters needed to be defined, as outlined in Supplementary Pseudo Code 1. These parameters included: [1] Resistance levels of 2-6Sia-expressing

Cells: Determining the resistance levels of cells expressing 2-6Sia to suppressive and lethal effects, [2] Cell ratios: Establishing the relative numbers of Suppressive, Permissive, and Lethal cells in relation to cancer cells, and [3] Impact strength: The strength of the influence exerted by Suppressive, Permissive, and Lethal cells on cancer cells (see Materials and Methods for details).

As for [1], cells expressing 2-6Sia demonstrate approximately 2 to 4.5-fold resistance to TNF, cisplatin, and gemcitabine compared to non-expressing cells (*51-53*). To take into account this, we assigned a 5-fold resistance factor to cancer cells expressing the highest level of 2-6Sia when subjected to Suppressive or Lethal effects, relative to non-expressing cells. In this scenario, when a cancer cell encounters, for instance, a Lethal cell, there is a 20% chance of the cancer cell being eliminated (a 1 in 5 chance). Conversely, if a cancer cell lacks 2-6Sia expression and encounters a Lethal cell, it is eliminated with a 100% probability. The degree of resistance for other cells was calculated based on their 2-6Sia expression levels.

Regarding parameter [2], we conducted a 3D TME simulation involving cancer cells and either Suppressive, Permissive, or Lethal cells to determine the ratio that could effectively observe the impact of these cells on cancer cells (Supplementary Fig. 5A-H, using Cervical 2-6Sia 1.0 and Pancreatic 2-6Sia 1.0 cells). Within the ratios of inferred cancer cells to Suppressive cells, ranging from 1:0.1 to 1:0.9 (Supplementary Fig. 5A and E) and the ratios of cancer cells to Lethal cells, ranging from 1:0.1 to 1:0.46 (Supplementary Fig. 5C and G), we observed a decrease in the expansion of the cancer cell population, as expected. This reduction was primarily due to the prolongation of cell doubling time or cell elimination, with Lethal cells exerting a stronger effect compared to Suppressive cells. However, when we performed the simulation by varying the ratio of Permissive cells, it did not have a significant impact on the expansion of the cell population. This observation can be attributed to the fact that the minimal cell doubling time was restricted from falling below the shortest doubling time observed within the inferred cell population (Supplementary Fig. 5B and F). Inferred cancer cells, which had a doubling time close to the shortest observed, may have experienced a comparatively lower influence

from the presence of Permissive cells. Consequently, we conducted the 3D TME simulation with a fixed ratio of 1:0.2 for inferred cancer cells to Suppressive cells and a varying ratio of Permissive cells (cancer cells: Suppressive cells: Permissive cells; 1:0.2:0.2 to 1:0.2:0.5, Supplementary Fig. 5D and H). This approach allowed the effects of Permissive cells to counteract the prolongation of cell doubling time caused by Suppressive cells, resulting in the reduction of cell doubling time. As anticipated, we observed that Permissive cells effectively mitigated the reduction in population expansion induced by Suppressive cells. These results indicate that the effects of Suppressive, Permissive, and Lethal cells can be observed within the ratio range of 1:0.1 to 1:0.9 for Suppressive and Permissive cells and of 1:0.1 to 1:0.46 for Lethal cells.

Regarding parameter [3], we performed a 3D TME simulation to determine the range of strength for Suppressive, Permissive, and Lethal cells' effects on cancer cells (Supplementary Fig. 5I and P). For Suppressive cells, when the strength was set to 10, the cell doubling time of the cells encountering the Suppressive cell and its progeny was extended by 10% (for instance, if the doubling time was 25 hours, it was extended to 27.5 hours). Similarly, with a strength setting of 10 for Permissive cells, the cell doubling time of the encountering cell and its progeny was reduced by 10%. If the reduced doubling time fell below the shortest doubling time within the inferred cell population, it was capped at the shortest time. For Lethal cells, the chance of inducing cell death was set. For example, setting values of 20 implied that cell death occurred with a 20% chance when cancer cells encountered a Lethal cell. We conducted 3D TME simulations by varying these setting values and found a similar trend for Suppressive, Permissive, Lethal, and the combination of Suppressive and Permissive cells, as analyzed by cancer cell to immune cell ratio (Supplementary Fig. 5I-P). These results suggest that the effects of Suppressive, Permissive, and Lethal cells can be analyzed within the range of 10% to 100% for Suppressive and Permissive cells and a 10% to 50% chance for Lethal cells.

The 3D TME simulation process involved the identification of Suppressive, Permissive, and Lethal cells within a specified spatial range around a cancer cell (see Materials and Methods for

details). At each time point, Suppressive, Permissive, and/or Lethal cells within this spatial range were located, and the nearest cell to the cancer cell was selected. If the selected cell was a Lethal cell, first, the resistance levels related to the 2-6Sia expression on cancer cells were considered, followed by the Lethal strength. If cell death was determined to be induced, the process of eliminating the cancer cell and its progeny was executed. If the nearest cell was a Suppressive cell, the cell doubling time was extended, after considering 2-6Sia expression levels followed by the Suppressive strength. When the nearest cell was a Permissive cell, the cell doubling time was reduced, accounting for the Permissive strength. In cases where the nearest and second nearest cells were a combination of Suppressive and Permissive cells, or vice versa, the relative distance between both cells was factored into the calculation of strength. This process was applied to each inferred cancer cell, and the iterative simulation continued until reaching the final time point. In Fig. 5B, a graphical representation of the 3D TME is presented, depicting views at 0, 2000, 4000, and 6000 minutes. The simulation was performed with Cervical 2-6Sia 1.5 cells in the presence of Suppressive, Permissive, and Lethal cells, with the influence strength of each category set to 0 (Control) or Lethal strength 100.

Simulations with Pancreatic and Cervical Cancer TME.

We conducted 3D TME simulations using TME cell landscape data from 193 cervical and 172 pancreatic cancer patients sourced from the TIMEDB database (*54*). In this process, we leveraged the ConsensusTME dataset to classify immune cells, cancer-associated fibroblasts, and endothelial cells into three categories: Suppressive, Permissive, and Lethal cells. The percentage of each cell category was calculated, and the corresponding number of cells was introduced into the 3D TME based on these percentages that are unique for each cancer case (see Supplementary Data 5 (Cervical) and Supplementary Data 6 (Pancreatic) for the calculated percentages). For instance, if the Suppressive cell percentage was calculated to be 45%, 225 Suppressive cells alongside 500 inferred cancer cells were placed. In terms of Suppressive, Permissive, and Lethal strength, we selected 50%, 50%, and 100%,

respectively, which we confirmed had an effect on the population expansion of inferred cancer cells (Supplementary Fig. 5). We employed four inferred cell populations for both cervical (2-6Sia 1.5, 1.0, 0.5, and 0.25) and pancreatic (2-6Sia 1.5, 1.0, 0.5, and 0.25) cancers. Consequently, we performed 3D TME simulations for 193 cervical cancer cases with the aforementioned cell populations and 172 pancreatic cancer cases with the corresponding cell populations. Subsequently, we analyzed the simulation results in relation to an indicator for prognosis, Progression-Free Survival (PFS), available from TIMEDB (Fig. 6, Cervical cancer; A-C, and Pancreatic cancer; D-F). In the figures, each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case, and blue, brown, light blue, and green dots represent results obtained with 2-6Sia 1.5, 1.0, 0.5, and 0.25, respectively. We plotted various parameters against PFS, including cell population size (Fig. 6A and D), 2-6Sia levels (Fig. 6B and E), and the number of reproductive cells produced through multipolar cell division, which are likely to carry an altered number of chromosomes (Fig. 6C and F). However, in all cases, the coefficient of determination (R^2) was below 0.1, indicating a lack of significant correlation between the simulation results and PFS. This suggests that establishing an association between simulation outcomes and PFS would require additional information, particularly related to the responses of cancer cells and the TME to anti-cancer treatment. Detailed results of the simulation, PFS data, C1 to C6 classification (*55*), and categorized values are summarized in Supplementary Data 5 (Cervical) and Supplementary Data 6 (Pancreatic).

Relationship between simulated cell population size and 2-6Sia expression levels.

To examine the relationship between simulated cell population size and 2-6Sia expression levels, we conducted an analysis using simulation data from Cervical 2-6Sia 1.0 cells and Pancreatic 2-6Sia 1.0 cells within TME cell landscapes. We observed a positive correlation between the simulated cell population size and 2-6Sia expression levels (Fig. 6G, with $R^2 = 0.25$ for Cervical, and in Fig. 6H, with $R^2 = 0.41$ for Pancreatic). This positive connection between cell population size and 2-6Sia

expression levels indicates a significant role in the proliferation of 2-6Sia-expressing cells in sustaining the cell population. As for the associations between simulated population size and the number of reproductive cells generated through multipolar cell division, as well as between the number of reproductive cells and 2-6Sia expression levels, R^2 values remained below 0.1 (Supplementary Fig. 6). This suggests that the accumulation of such reproductive cells within the cervical and pancreatic cell population may vary depending on the unique characteristics of each individual cancer case.

Characteristics of cervical cancer cells within the TME.

We investigated the relationship between the simulated population size of Cervical 2-6Sia cells and the proportions of Suppressive (Fig. 7A), Permissive (Fig. 7B), or Lethal (Fig. 7C) cells, as determined based on the cell landscape data of the cervical and pancreatic cancer TME from the TIMEDB database (*54*). The results revealed correlations between the simulated cell population size of Cervical 2-6Sia cells and the proportions of Suppressive, Permissive, and Lethal cells within TME (Fig. 7A-C; $R^2 > 0.1$-0.55, each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case, see Supplementary Data 5 for the proportions). Notably, negative linear correlations were observed when examining cell size in relation to Suppressive (Fig. 7A) and Lethal (Fig. 7C) cell proportions within the 3D TME. In contrast, a positive correlation emerged when cell size was plotted against the proportion of Permissive cells (Fig. 7B). These findings suggest that the proliferation of Cervical 2-6Sia cells is indeed influenced by the composition of immune cells within the TME of each cancer case. Among the Cervical 2-6Sia cell populations, we observed a 10% increase in the average cell population size of Cervical 2-6Sia 0.25 compared to that of Cervical 2-6Sia 1.5 (Fig. 7D). Cervical 2-6Sia 1.5 cells exhibited the highest levels of 2-6Sia expression, which gradually decreased in the Cervical 2-6Sia 1.0, 0.5, and 0.25 populations, as expected (Fig. 7E). Regarding the reproductive progeny produced through multipolar cell division, their contribution was marginal (Fig.

7F), with the percentage of such progeny within the Cervical 2-6Sia cell population being less than 0.4% (Fig. 7G).

These findings underscore the distinctive characteristics of cervical cancer cells (Fig. 7H). The proliferation of cervical cancer cells is significantly influenced by the immune cell landscape in the TME.

Characteristics of pancreatic cancer cells within the TME.

We then explored the relationship between the simulated population size of Pancreatic 2-6Sia cells and the proportions of Suppressive (Fig. 8A), Permissive (Fig. 8B), or Lethal (Fig. 8C) cells. In contrast to Cervical 2-6Sia, we found almost no significant correlations ($R^2 < 0.15$, each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case, see Supplementary Data 6 for the proportions) in the case of Pancreatic 2-6Sia cells. This suggests that the proliferation of Pancreatic 2-6Sia cells was relatively unaffected by the immune cell composition within the TME. Among different Pancreatic 2-6Sia cell populations, we observed no significant differences in population size (Fig. 8D). Pancreatic 2-6Sia 1.5 cells exhibited the highest levels of 2-6Sia expression, which gradually decreased in the Pancreatic 2-6Sia 1.0, 0.5, and 0.25 populations, as expected (Fig. 8E). Concerning the number of reproductive cells produced through multipolar cell division, there was a reduction in Pancreatic 2-6Sia 0.5 cells but an increase in the Pancreatic 2-6Sia 0.25 cell population compared to Pancreatic 2-6Sia 1.5 cells (Fig. 8F and G, average 0.48 (Pancreatic 2-6Sia 1.5 cells) to 0.66 (Pancreatic 2-6Sia 0.25 cells)). This suggests that although varying levels of 2-6Sia did not substantially impact the simulated cell population size, cells with altered chromosome numbers expanded within the Pancreatic 2-6Sia cell population, particularly in those expressing 0.25% levels of 2-6Sia.

These results underscore the unique characteristics of pancreatic cancer cells (Fig. 8H). The majority of pancreatic cells expressing 2-6Sia seemed to be less influenced by the TME cell landscape.

In addition, cells with altered chromosome compositions accumulated in populations expressing lower levels of 2-6Sia, introducing genetic diversity within the pancreatic cell population.

Discussion

In this study, we empirically generated data related to the spatiotemporal behavior of individual cells using a computerized single-cell tracking system using two established cancer cell lines. We then created inferred cell populations, which signified a particular attribute of cancer cells in vivo, based on these datasets using the cell fate simulation algorithm we previously developed (*20*). Additionally, we simulated the impact of the immune cell landscape in the TME on these cells using a 3D TME simulation algorithm by focusing on the role of 2-6Sia-related stemness and the reproductive cells produced through multipolar cell division. To date, extensive data has been accumulated at the gene, protein, cell tissue, and clinical levels regarding cancer cells and the TME. However, often, these data represent cancer tissue or cells at a specific moment within the long processes of cancer development. For instance, the status of the TME at the time of cancer tissue analysis may differ from earlier stages and may continue to change as cancer progresses. While the information obtained from the TME status at a specific moment during the cancer progression can offer insights into its dynamic changes, we believe that incorporating a temporal context based on empirical data, coupled with AI-based simulations that account for the heterogeneous nature of cancer cell population, will provide a deeper understanding of the dynamics of cancer development. Fig. 9 presents a summarized conceptual flow of our approach, which combines the spatiotemporal behaviors of individual cancer cells with existing information on the TME. This approach is versatile and can utilize various attributes, not limited to 2-6Sia, to uncover the fate of cancer cells within the TME and shed light on the distinct impact of the TME landscape. Moreover, in the process of developing this approach, it can bring to the forefront the biological parameters that require exploration, thus revealing overlooked aspects of both cancer cells and TME cells. Thus, we believe that our approach has the potential to significantly contribute to the advancement of cancer research.

In our study on cervical cancer cells, we have demonstrated a significant influence of the immune cell landscape within the TME on the proliferation of cervical cancer cells. Given that the

majority of cervical cancer cells do not express 2-6Sia, Suppressive, Permissive, and Lethal cells exert their effects in a probabilistic manner. This observation allows us to predict that the majority of cervical cancer cell populations lacking stemness may exhibit sensitivity to cytotoxic treatments, which may be reflected in the general characteristics of cervical cancer response to cancer treatment. Conversely, a small subset of cells expressing 2-6Sia may selectively survive such treatments.

Unlike cervical cancer cells, our study reveals that the proliferation of pancreatic cancer cells is relatively less influenced by the immune cell landscape within the TME. This difference can be attributed to the fact that the majority of MiaPaCa2 cells express 2-6Sia, a factor known to confer resistance to cell killing (*51-53*). Consequently, cells with high 2-6Sia expression levels exhibit enhanced reproductive capabilities. Importantly, high 2-6Sia expression is closely associated with stemness (*35-40*). It has been reported that glycan modification status with 2-6Sia is subject to dynamic regulation. For example, blood-circulating α2-6Sialytransferase 1 can modify cell surface glycoproteins with 2-6Sia when CMP-sialic acid is released from inflammation-activated platelets. (*56*). This implies that cancer cells that do not express 2-6Sia can be converted into 2-6Sia-expressing cells, with some attributes related to a stemness, by this platelet-related mechanism, adding a layer of complexity that is not solely dictated by gene expression patterns in cancer development. Furthermore, the implication that cancer cells retain stemness may also contribute to the complexity of cancer cell development. Although it has been reported that malignant cancer cells tend to retain stemness (*24*), maintaining a high level of stemness could also preserve the integrity of cancer cells. We found that cell death, cell fusion, and multipolar cell division tend to occur in cells with lower levels of 2-6Sia. In addition, in our 3D TME simulation, we observed an increase in reproductive cells produced through multipolar cell division in cell populations with lower levels of 2-6Sia expression. As these reproductive cells could carry an altered number of chromosomes, it can potentially lead to increased phenotypic diversity within a cancer cell population. In connection with 2-6Sia expression, it has been reported that invasive cancer cells were generated among cells with reduced 2-6Sia expression levels

(*57*). This suggests that malignant cancer cells could arise among cells retaining a reduced level of stemness to introduce phenotypical diversity among the cancer cell population while maintaining a certain level of stemness to preserve their reproductive ability. For patients with pancreatic cancer, these complex parameters could affect their response to anti-cancer treatment and prognosis. Those could be factors related to a poor prognosis of pancreatic cancer and we believe that our simulation can be used to examine the possible impact of the fate of pancreatic cancer cells by creating various scenarios.

In this report, we have showcased the core framework of the 3D TME simulation. This framework can be further extended by increasing precision, such as adding more categories of immune cells, defining the rate of immune cell influx into the TME, incorporating the formation of certain tissue structures, and considering other cancer markers like the expression of mutant p53. Furthermore, the response of cells to anticancer treatment can be simulated by extending this framework. Concerning this, cancer treatments, which may alter the immune cell landscape and cancer attributes, could hinder attempts to predict the prognosis of cancer based on the immune cell landscape of the TME or characteristic alterations of cancer cells in primary tumors. Moreover, the lack of knowledge related to the alteration of the landscape or the attributes makes it challenging to validate any prognostic markers beyond establishing a statistical relationship between a marker and prognosis. This significant gap could be filled by our approach.

We thus believe that this integrated approach, which we have showcased here, utilizing 2-6Sia expression, known to be linked to stemness as an example, has the potential to facilitate more precise predictions of cancer prognosis in humans and enhance our understanding of the response of cancer cells to treatment.

Materials and Methods

Cell culture and cell plating.

HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) in a humidified atmosphere with 5% CO^2. MiaPaCa2 cells were maintained in DMEM supplemented with 10% FBS and 2% horse serum under a 5% CO_2-humidified atmosphere. For cell plating, approximately 3500 cells in 50 µl of cell suspension were added to the center of each well of a Lab-Tek™ II 8 Chamber Slide. Subsequently, 0.75 ml of culture medium was gently added to each well. The Lab-Tek™ II 8 Chamber Slide was placed on a microscope stage 24 hours after plating.

Culture conditions on the microscope stage.

Cells were maintained in an environmental chamber (Live Cell Instruments, Korea) set at 37°C with 80% relative humidity. To prevent an increase in medium pH within the Lab-TekII 8-well chamber, the CO_2 concentration was controlled at 7.5%. To minimize medium evaporation, a NIR-DIC optimal glass lid (Live Cell Instruments) was placed on the Lab-Tek™ II 8 Chamber Slide. The typical medium evaporation rate was 10 µl/24 hours, and as a result, 700–800 µl of medium was added per well, ensuring cells could be cultured for at least one week without the need for medium changes. For monitoring pH during long-term live cell imaging, phenol red was included in the medium. Importantly, this inclusion did not interfere with NIR-DIC imaging.

Live cell imaging microscopy.

A custom microscope was constructed using an Olympus IX81 microscope frame (Quorum Technologies, Ontario, Canada). NIR-DIC imaging was chosen to minimize phototoxicity and reduce sensitivity to light distortion caused by plastic partitions and surface irregularities in the culture medium. The microscope featured dual imaging capabilities, combining NIR-DIC and fluorescent

imaging for both live and fixed samples. Two distinct light paths were established within the microscope. For NIR-DIC imaging, NIR light was generated using light-emitting diodes, and passed through a polarizer, Nomarski prism, and condenser before illuminating the cells. The returning light passed through an objective lens and Nomarski prism, ultimately reaching the first charge-coupled device (CCD) camera (Camera 1, Hamamatsu Photonics, Image EM, 512×512 pixels). In the case of fluorescent imaging, laser light was directed onto cells stained with fluorescence-conjugated antibodies or proteins. This was accomplished by passing the laser light through a Nipkow disk (Yokokawa-Quorum Technologies) and an objective lens to excite fluorophores. The emitted light was collected by the Nipkow disk and captured using the second CCD camera (Camera 2, Hamamatsu, Image EM, 512×512 pixels). The system allowed the use of ×10–×40 objectives. In this study, an Olympus ×10 dry objective (UPlanSApo, 10×/0.40 NA, α/0.17/FN26.5) or ×20 dry objective (UPlanSApo, 20×/0.75 NA, α/0.17/FN2G.5) was employed. To achieve equivalent magnifications of ×15 and ×30, respectively, a ×1.5 coupler (Quorum Technologies) was inserted into the light path leading to the CCD cameras. A precision piezo XY stage with absolute measurement capabilities was utilized. An environmental chamber (Live Cell Instrument) was mounted on the piezo stage. The entire microscope system, along with the piezo XY stage, was controlled using Metamorph software (Quorum Technologies) on a Windows computer. Images were acquired using Metamorph's multi-dimensional acquisition mode (MDA). For NIR-DIC imaging, a 34 ms exposure time was used. The exposure time for fluorescence imaging was adjusted based on the fluorescence intensity. Typically, 20–80 z-planes at 1 μm intervals were acquired for NIR-DIC imaging. The resulting z-plane NIR-DIC images, captured by Camera 1, were saved as 512×512 pixel multi-layer TIFF files. For fluorescence imaging, z-plane NIR-DIC and fluorescence images generated by CCD cameras 1 and 2, respectively, were merged using a macro program in Metamorph. This produced 512×1024 pixel multi-layer TIFF files. To manage data, Computer 1 (Windows computer) was connected to Computer 2 (Macintosh) via an Ethernet cable. Computer 2 was responsible for creating live cell videos. Image files generated by

Computer 1 were promptly transferred to Computer 2 using an in-house *File Transfer* software program, with subsequent deletion of the image files from Computer 1.

Developed software for a single-cell tracking system.

Image files produced by Metamorph underwent processing through a series of in-house software tools (Supplementary Fig. 1). The computerized single-cell lineage tracking analysis system consisted of two computers: Computer 1 (Windows) controlled the microscope, while Computer 2 (Macintosh) handled image processing, single-cell tracking, and data analysis tasks. Computer 1 was equipped with commercially available image acquisition software (Metamorph) to oversee microscope control and image file creation. Computer 2 utilized several in-house software programs (italic names indicate the in-house software). *Image Processing Controller* managed other software programs. *File Transfer* and *Map* facilitated communication between Computers 1 and 2. *File Converter* imported images generated by other microscopes. *Name Assignment*, *Focal Image Selection*, and *Contrast Set* were responsible for creating live cell videos. *Data Backup* controlled file archiving and backup procedures. *Outline Drawing* performed image segmentation and established the object segmentation library. *Movie Viewer* played movies and fine-tuned image quality. *Object Tracking Controller*, in conjunction with *Automatic Object Tracking*, created the cell lineage database and facilitated data verification. *Data Analysis* offered various options for data analysis. The data generated through this process was subsequently used for cell fate simulation and 3D TME simulation.

Setting FOVs and adjustment of focus.

Cells cultured in each well of the Lab-Tek™ II 8 Chamber Slide were simultaneously monitored. To cover the area of interest with multiple fields of view (FOVs), 2D image acquisition arrays were established in each well (*20*). The configuration of FOVs for these arrays and the necessary focus adjustments were executed using *Map* software in conjunction with Metamorph. *Map* software

was specifically designed to coordinate and share information on the xyz positions between Metamorph on Computer 1 and software on Computer 2. To initiate this coordination, the objective lens was initially positioned at the left corner of Well 2 of the Lab-Tek™ II 8 Chamber Slide, and this position was registered in *Map* software. Subsequently, the outline of a Lab-Tek™ II 8 Chamber Slide was traced on the screen. The objective lens was then directed to a FOV of interest within a well using the outline as a guide. To view the cells in the surrounding area, 5x5 FOV dimensions were captured using Metamorph, which generated xyz position data for each FOV. A suitable area for live cell imaging was searched by repeating this process. When the suitable area was found, xyz position data saved on Computer 1 was transferred to Computer 2 using *Map* software. Then, the selection of the area to be imaged was finalized by setting the 2D image acquisition arrays that define each position of FOV using *Map* software. Next, the optimal objective z position for each FOV was determined, and the xyz position information created by Map software was compiled into a multi-dimensional acquisition (MDA) file format. This MDA file was then used by Metamorph to acquire an image of each FOV. The complied MDA file, created by *Map* software, was transmitted to Computer 1 and uploaded to Metamorph. Subsequently, Metamorph performed MDA image acquisition to generate images for the corresponding FOVs.

Starting image acquisition.

The first round of image acquisition was initiated by Metamorph through MDA image acquisition (Supplementary Fig. 1). Subsequently, *File Transfer* transmitted the multi-layer TIFF files (512×512 pixels) generated by Metamorph (Computer 1) to Computer 2. To facilitate file archiving, each multi-layer TIFF file was assigned a specific file name using *Name Assignment*. The image files underwent processing using *Focus Image Selection*, which involved selecting a focal image from among the images within a multi-layer TIFF file or generating an all-in-focus image. The all-in-focus image was constructed by selecting the optimal grayscale value from the z-axis data for each pixel.

Following this, the contrast of the resulting images was adjusted, and image backgrounds were corrected using *Contrast Set*. Since the background patterns varied for each FOV, custom background patterns were generated using data from single-layer TIFF files. Each of these background patterns was then applied to the corresponding FOVs. Ultimately, stitched images were created using *Contrast Set* after contrast adjustment, background correction, and fine-tuning of the position for each FOV.

Automated long-term live cell imaging video creation.

Following the initial setup, automated image acquisition commenced. Metamorph captured images of each FOV at 10-minute intervals. Subsequently, after each round of image acquisition, the process involving *File Transfer*, *Name Assignment*, *Focus Image Selection*, and *Contrast Set* was coordinated by *Image Processing Controller*. The status of the images could be monitored using *Image Viewer*, and *Data Backup* generated backup files as needed. This process continued until the completion of live cell imaging.

Staining of cells with SNAI-TRITC.

Fluorescent imaging was conducted at the end of the live cell imaging process. Before removing the Lab-Tek™ II 8 Chamber Slide from the microscope stage, a snapshot was captured using Map software to document the position of a unique object, such as a cell with a distinct shape. Subsequently, the chamber slide was removed from the stage, and the cells underwent three washes with phosphate-buffered saline (PBS). Following this, the cells were fixed with 3.7% paraformaldehyde for 15 minutes at room temperature and then washed three additional times with PBS. Next, the cells were treated with Carbo-Free Blocking Solution (×1) (Vector) for 1 hour at room temperature and subsequently incubated with TRITC-labelled SNAI (E-Y Laboratories) diluted with Carbo-Free Blocking Solution (×1) to a concentration of 50 μg/ml for 1 hour at 4°C. Following this incubation, the cells were washed with PBS three times and exposed to DAPI. This was performed by

diluting two drops of NucBlue Fixed Cell ReadyProbes Reagent with 1 ml Carbo-Free Blocking Solution (×1) for 15 minutes at room temperature, followed by three additional washes with PBS. The Lab-Tek™ II 8 Chamber Slide was then returned to the microscope stage, and its position was adjusted using the snapshot, which was taken before removing the chamber slide from the microscope stage. As returning it to the identical position as during live cell imaging proved to be challenging, the shift in the x and y position from the live cell imaging setup was calculated, and a new MDA file, accounting for this shift in position, was generated. Following the upload of this new MDA file, fluorescent image acquisition was conducted by Metamorph. Fluorescent imaging was conducted with a 250 ms exposure using lasers with wavelengths of 403 nm and 491 nm for DAPI and TRITC, respectively.

Segmentation of NIR-DIC images.

Grayscale live cell imaging videos were then subjected to image segmentation. To this end, we developed a method referred to as Stepwise Area Expansion, which was performed by *Outline Drawing* software (Supplementary Fig. 2). Four images were created from the original grayscale image by applying four different thresholds (TH) values (TH1–4, Supplementary Fig. 2A). First, pixels above the threshold value (e.g. 200, 170, 140, and 110 of 0–255 grayscale for TH1, 2, 3, and 4, respectively) were extracted from the grayscale image. Segmentation was then performed step by step, starting from the TH1 image (Supplementary Fig. 2B). Connectivity analysis was performed on the TH1 image to identify groups of pixels that were attached and an edge circle that surrounded the group of pixels was then determined (Supplementary Fig. 2B, Panel a, pink circles). The TH1 edge circles were overlaid on the corresponding grayscale image (Supplementary Fig. 2B, Panel b). The pixels composing the circles were referred to as the original pixels. Using the overlaid grayscale image as a reference, the software examined the pixel values for each original pixel. The values of pixels located in a circle from 12 o'clock from the original pixels were examined (Supplementary Fig. 2B, Panel b and Line extension (magnified)). The edges of the pixels were extended until the values became either 100 (representing

the grayscale background) or 50 above or below the original pixel, as long as those extensions were towards the outside of the edge circles. The final pixel positions were determined by this technique. Similar examinations were performed for the other three directions (Supplementary Fig. 2B, Panel b: 3, 6, and 9 o'clock directions; yellow, light blue, and blue lines, respectively) and the final pixel positions were then linked to create four directional lines corresponding to the 12, 3, 6, and 9 o'clock directions (Supplementary Fig. 2B, Panel c and Edge linking (magnified), L 12', L3', L6', and L9'). Finally, the four lines were linked to make an edge circle (Supplementary Fig. 2B, Panel d, green circles), which completed the initial segmentation for the first TH1 image. The TH1 edge circles were then overlaid on the TH2 image and another connectivity analysis was performed, except for the TH1 edge circle areas (Supplementary Fig. 2B, Panel e, green circles). Two types of edge circles (Supplementary Fig. 2B, Panel e, green and pink circles) were thus created at this stage. The pink circles that emerged in the TH2 image were overlaid on the corresponding grayscale image, the pixel value of each original pixel in the pink circles was examined, the final pixel positions were determined (Supplementary Fig. 2B, Panel f), and the positions were linked to create four lines (Supplementary Fig. 2B, Panel g) and a new edge circle was created (Supplementary Fig. 2B, Panel h), as described for TH1. The software applied a different expansion approach for the green edge circles that had been determined during TH1 processing (Supplementary Fig. 2B, Panel e, green circles): the location of each pixel on the edge of the circles (Supplementary Fig. 2B, Panel k) was moved by 1–4 pixels outside the original (Supplementary Fig. 2B, Panel l, Edge expansion (magnified)) to create the red circles (Supplementary Fig. 2B, Panel l). Both the green TH2 (Supplementary Fig. 2B, Panel h) and red TH2 circles (Supplementary Fig. 2B, Panel l) were then overlaid onto the TH2 image (Supplementary Fig. 2B, Panel i), and the circle overlaps were removed to create new edge circles (Supplementary Fig. 2B, Panel j). Those new edge circles were then applied to TH3, and the above-mentioned processes were repeated for TH3 and TH4 images. Areas surrounded by edge circles were numbered.

This approach proves valuable in scenarios where cell segmentation encounters challenges. For instance, when a flat cell coexists with larger, bright cells that are easily distinguishable in the NIR-DIC image, the flat cell may become obscured by the brightness of the larger cells, making segmentation difficult. In the Stepwise Area Expansion method, the initial step involves defining the boundaries of the brighter objects within the TH1 image. Since the TH1 image typically does not encompass the darker flat cell, there is a possibility that the flat cell lies outside this boundary. In subsequent steps utilizing images created with lower threshold levels, the darker flat cell can be included in the image. By excluding the area within the previously determined boundary during segmentation, the region corresponding to the flat cell can be segmented. This approach facilitates accurate segmentation of grayscale images that contain cells with varying brightness levels and shapes. Supplementary Fig. 2A (depicted by blue lines) provides examples of segmentation outcomes using both high and low-density HeLa cell cultures.

Assignment of cell lineage number and cell number.

To assign cell lineage numbers and cell identifiers to the cells within the Time 1 image, a verification and correction process was carried out for the segmentation results using the *Object Tracking Controller* software. In cases where multiple segments were incorrectly associated with a single cell, one segment was chosen to represent the cell, and segments were merged as needed. Conversely, if multiple cells were grouped within a single segment, the segment was divided to accurately represent each individual cell. The cells identified within the Time 1 image were classified as progenitor cells and a unique cell lineage number was assigned to each progenitor cell. Progenitor cell numbers were designated as 0. Data was then recorded in the Cell Lineage Database.

Automated object tracking (single-cell tracking).

After assigning cell lineage numbers to segmented areas, the automatic cell tracking process was initiated using the *Automatic Cell Tracking* software. Supplementary Fig. 3 provides an overview of the automatic tracking process for the segmented area, outlined by green lines. In Supplementary Fig. 3A (Time A), orange characters represent segmented areas identified as cell representations. The blue area with a white asterisk denotes the cell being tracked, while the yellow and magenta areas indicate segmented areas representing neighboring cells. In the subsequent time point (Supplementary Fig. 3B, Time A+1), the positions of cells and segmentation patterns changed. To perform single-cell tracking, it was essential to determine the segmented area corresponding to the cell being tracked. To achieve this, the blue area from Time A was overlaid onto the Time A+1 image (Supplementary Fig. 3C). However, the blue segmented area from Time A did not exist in Time A+1; instead, it overlapped with the light blue and white segmented areas. The *Automatic Cell Tracking* software couldn't ascertain which area corresponded to the tracked cell. To resolve this, the yellow and magenta areas were overlaid on the Time A+1 image (Supplementary Fig. 3D). This process revealed that the light blue area was related to the yellow area. Furthermore, the magenta area overlapped with the white area, but the larger portion of the white area overlapped with the blue area. Consequently, the *Automatic Cell Tracking* software determined that the white area represented the cell being tracked (Supplementary Fig. 3E). This process was systematically repeated for all cells recorded in the database.

Tracking data verification.

Since achieving 100% accuracy in automatic single-cell tracking is not always feasible in practice, the tracking data underwent verification using the interactive features of the *Object Tracking Controller* software. If errors were identified, the single-cell tracking data were corrected, ultimately enabling the generation of nearly 100% accurate single-cell tracking data.

Data analysis software.

Bioinformatics analysis was conducted using the *Data Analysis* software, which encompasses a range of functions such as generating cell lineage maps, calculating the individual cell doubling time, and determining cell population expansion curves. Notably, cell fate simulation is one of the features offered by this software.

Staining of cells with antibodies against a stem cell marker.

HeLa and MiaPaCa2 cells were fixed with 3.7% paraformaldehyde for 15 minutes at room temperature, followed by three additional washes with PBS. Subsequently, the cells underwent treatment with Carbo-Free Blocking Solution (×1) (Vector) for 1 hour at room temperature. After this blocking step, the cells were incubated with the following primary antibodies or lectin: Anti-CD133 antibody (Developmental Studies Hybridoma Bank) at a dilution of 1:50, which corresponds to 5 μg/ml, in Carbo-Free Blocking Solution (×1) for 30 minutes at room temperature, Anti-TAG-72 antibody (B72.3) at a 1:50 dilution in Carbo-Free Blocking Solution (×1) for 1 hour at room temperature, or FITC-labelled rBC2LCN (Wako) at a 1:100 dilution in Carbo-Free Blocking Solution (×1) for 30 minutes at room temperature. Following the primary antibody or FITC-labelled rBC2LCN incubation, the cells were washed three times with PBS. For cells incubated with primary antibodies (anti-CD133 and anti-TAG-72), a secondary antibody (Invitrogen, goat anti-mouse IgG Alexa Fluor 488) diluted 1:1000 was applied for 1 hour at room temperature. After the secondary antibody incubation, the cells were subjected to three additional washes with PBS. Subsequently, the cells were stained with DAPI and underwent three final washes with PBS. Fluorescent imaging was conducted with a 250 ms exposure using lasers with wavelengths of 403 nm for DAPI and 491 nm for Alexa Fluor 488 and FITC.

Generation of inferred cell populations with varied 2-6Sia expression levels.

The initial step in generating the inferred cell population involved creating a 2-6Sia expression list (Supplementary Fig. 4A). In this list, we considered examples from Lineage 1 and Lineage 2. Indirect immunofluorescence using TRITC-tagged SNA1 was performed at the end of live cell imaging, allowing us to determine the SNA1 binding levels of cells present at the end of the imaging period. Lineage 1 exhibited no expression of 2-6Sia in any of the cells present at the end of live cell imaging, while Lineage 2 showed various levels of 2-6Sia expression in its cells. Based on the 2-6Sia expression levels of cells at the end of live cell imaging, we traced back the 2-6Sia expression levels along the cell lineage map to determine the expression levels of their parent cells. This was done by calculating the average expression of daughter cells, thus estimating the evolution of 2-6Sia expression levels within a cell lineage. Subsequently, we compiled a 2-6Sia expression list that included the 2-6Sia expression levels of all tracked HeLa or MiaPaCa2 cells. For the Lineages 1 and 2 examples, the list contained 17 data items with a value of 0 and 17 data items with varying levels of expression values.

Next, we generated the inferred cell population using the cell fate simulation algorithm (*20*), as illustrated in Supplementary Fig. 4B. This algorithm analyzed cell lineage data and calculated the probability of certain types of cell events occurring following a given event, along with the time intervals between events. The combinations of events that the algorithm analyzed were previously described (*20*). The cell fate simulation algorithm employed these probabilistic values to generate an inferred cell population. Initially, the algorithm assigned the length of time to the First event (the event that occurred in the progenitor cell) to the progenitors. In the actual process, 500 progenitors were generated, each at a different stage of the cell cycle, resulting in varying time intervals until the First event.

Subsequently, a cellular event was assigned to a progenitor cell based on the probabilistic values. If a bipolar cell division was assigned, two daughter cells were created. In Example 1, a longer time interval until the next event was assigned compared to Example 2. After the assignment, the algorithm referred to the cell doubling time (the time between bipolar cell divisions) of 2-6Sia-

expressing cells. As 80% of the cell doubling times for 2-6Sia-expressing cells fell within ±10% of the average cell doubling time for these cells, the algorithm checked whether any of the 2-6Sia-expressing cells had cell doubling times within ±10% of the time assigned to daughter cells. If no such cells were found, 2-6Sia expression levels were not assigned to the daughter cells (Example 1). However, if some cells met this criterion, 2-6Sia expression levels were assigned to the daughter cells (Example 2). This approach allowed us to generate inferred cells with reproductive abilities similar to those of 2-6Sia-expressing cells.

Subsequently, the algorithm assigned 2-6Sia expression levels to the daughter cells by referring to the 2-6Sia expression level list (Supplementary Fig. 4C). A random value was generated to select one value from the list, which was then assigned to one of the daughter cells. In the case of the Lineage 1 and 2 examples (Supplementary Fig. 4A), there was a 50% chance of selecting value 0. As the variation in 2-6Sia expression levels typically fell within ±15%, the algorithm determined the 2-6Sia expression level of the second daughter cell based on the first daughter cell's level, accounting for the ±15% variation. Finally, 2-6Sia expression levels were modified to create inferred cell populations with various levels of 2-6Sia expression. This was achieved by multiplying the value of the 2-6Sia expression by factors such as 1.5, 0.5, and 0.25. In cases where the calculated value exceeded the highest 2-6Sia expression value observed in the HeLa or MiaPaCa2 cell population, the value was capped at the highest observed level.

The 3D TME simulation methodology.

The outline of the 3D TME simulation was presented in Supplementary Persuade Code 1. The codes for the 3D TME were developed using C, C++, and Objective-C with the utilization of SceneKit (macOS 12). To execute the 3D TME simulation, various parameters pertaining to graphic display and simulation execution were configured. Graphic display parameters for both cancer cells and immune cells included options such as cell shape (sphere, box, or capsule), displayed cell size, color, alpha

value, motility of cells, nucleus size, and nucleus color. These parameters could be individually adjusted for cancer cells, Suppressive cells, Permissive cells, and Lethal cells. Additionally, for fluorescence display reflecting the quantitated fluorescent value of each cancer cell, the level of expression could be visualized using either a heatmap scale (ranging from blue to red, indicating low to high expression) or manually set colors.

Concerning Suppressive, Permissive, and Lethal cells, the ratio of each cell type relative to cancer cells can be specified. In a typical simulation, 500 progenitors of the inferred cell population were employed, and the numbers of Suppressive, Permissive, and Lethal cells were determined based on the specified ratio relative to the total of 500 cells. For instance, if the ratio was set to 1:0.5 (cancer cells to either Suppressive, Permissive, or Lethal cells), then 250 of the chosen cell types were placed accordingly. These ratios could be customized independently for each category of cells, and the range of the ratio as well as the increment could be set. For instance, the 3D TME simulation could be automatically conducted by accounting for the defined range and increment. As an example, if the range was set from 1:0 to 1:1 for each category with an increment of 0.1, this would result in 1331 simulations being carried out automatically.

The strength of the impact could also be configured. In the case of Suppressive and Permissive cells, the setting values, such as 10, represented the percentage for prolonging or shortening the cell doubling time. For example, if an inferred cell had a doubling time of 25 hours and encountered a Suppressive cell, the doubling time would be extended by 10%, making it 27.5 hours. In the case of encountering Permissive cells, the cell doubling time for inferred cells and their progeny would be reduced by 10%. However, this reduced cell doubling time could not fall below the shortest cell doubling time observed within the inferred cell population, and in such cases, the shortest time was assigned.

As for Lethal cells, the setting value represented the probability of inducing cell death. For instance, a setting of 100 indicated a 100% chance of inducing cell death when cancer cells

encountered a Lethal cell. When cell death was triggered for a cancer cell, all its progeny were removed from the cell lineage database. The range of these setting values, along with the increment, could be customized, and the 3D TME simulation could be automatically executed, considering the specified increment value. The resistance of cancer cells to Suppressive and Lethal effects could also be set. For example, if the maximum resistance level was defined as 5-fold, then cancer cells subjected to Suppressive or Lethal effects would have a 20% chance (1 in 5 chance) of experiencing these effects.

These cells were positioned within spheres, each with a radius of 50 pixels. Cells were randomly distributed within the space, typically with a radius of 6 pixels from the cancer cell to search for Suppressive, Permissive, and Lethal cells. Under this setting, neighboring cancer cells were found within an 80% chance within the 6-pixel radius. Multiple cell lineage data of the inferred cell population can be loaded and the primary simulation process commenced with cancer cells identified at Time 1 and involved searching within a specified radius. If Suppressive, Permissive, and/or Lethal cells were discovered, the two nearest cells were selected. If the nearest cell was a Lethal cell, the resistance level of a cancer cell was initially considered, followed by the strength factor. If a cancer cell was determined to be killed, the cell death process was executed. If the nearest cell was a Suppressive cell, and the second nearest cell was not a Permissive cell, the resistance level of a cancer cell was initially taken into account, followed by the strength factor. If the cancer cell doubling time was prolonged, the process for prolongation was executed. If the nearest cell was a Permissive cell, and the second nearest cell was not a Suppressive cell, the strength factor was taken into account. If the cancer cell doubling time was determined to be reduced, the process for reduction was executed. If the nearest cell was a Suppressive cell, and the second nearest cell was a Permissive cell, the distances of both cells from a cancer cell were calculated, and the Suppressive strength was adjusted, considering the presence of Permissive cells. For instance, if the initial strength of both Suppressive and Permissive cells was set to 50 and the distance of Suppressive and Permissive cells from a cancer cell was 3 and 6 pixels, respectively, the effect of Permissive strength was considered as half of the Suppressive strength,

resulting in an adjusted Suppressive strength of 25 (50 (Suppressive) - 25 (Permissive)). If the nearest cell was a Permissive cell, and the second nearest cell was a Suppressive cell, a similar adjustment process was applied. These processes were iteratively carried out for each cancer cell identified at each time point and continued until reaching the final time point.

Parameters for the Methodology of 3D TME Simulation.

In the 3D TME simulation, the motility of cancer cells, Suppressive cells, Permissive cells, and Lethal cells was restricted to a maximum displacement of 2 pixels from their initial positions within the 3D TME sphere. This constraint was crucial to confine cancer cells and immune cells within the simulation sphere. However, this constraint also fixed the relative positions of cancer cells concerning Suppressive, Permissive, and Lethal cells. For example, once a cancer cell was positioned within the 3D TME, the spatial relationship between that cancer cell and nearby Suppressive, Permissive, or Lethal cells remained constant throughout the cell's simulation. As a result, if Suppressive, Permissive, and/or Lethal cells located within 6 pixels of a cancer cell were monitored at each time point, the same combination of Suppressive, Permissive, and/or Lethal cells would persist. This implied that if a Suppressive cell was initially the closest to a cancer cell, it would always remain the closest cell throughout the cancer cell's lifetime.

During the simulation, when a cancer cell experienced a Suppressive or Permissive effect, it did not receive additional Suppressive or Permissive effects. In other words, if a cell encountered a Suppressive cell and had the effect induced, it would not undergo another Suppressive effect. This rule applied uniformly to all progeny cells. However, these progeny cells could subsequently encounter additional Suppressive, Permissive, or Lethal effects.

When cancer cells produced daughter cells, the locations of these daughter cells within the 3D space were randomly determined within 6 pixels from the parent cancer cell. This randomization altered the relative positions of Suppressive, Permissive, and Lethal cells concerning the daughter cells

compared to their parent cell. Consequently, if a Suppressive effect was applied to a daughter cell due to its parent cell's previous interaction with a Suppressive cell, and the daughter cell encountered a Permissive cell, the daughter cell would be subject to the Permissive effect. Similarly, daughter cells could receive a Suppressive or Lethal effect if either of these cells became the closest to a daughter cell.

Other Parameters for 3D TME Simulation Methodology.

Upon placing Suppressive, Permissive, and Lethal cells in the 3D TME, individual cell lineages were established, each characterized by its specific attribute (e.g., suppressing, permissive, or lethal). The 3D TME simulation algorithm executed simulation processes by referencing these attributes. This flexibility allowed the creation of comprehensive simulation scenarios by modifying attributes, including alterations at specific stages of the simulation to reflect changes in cell behavior. \

Each cancer cell was associated with attributes, such as 2-6 Sia expression levels, and by incorporating multiple attributes, the simulation could encompass various characteristics of cancer cells. Thus, the 3D TME simulation, based on lineage data and associated attributes for each cell, facilitated the execution of complex simulations, provided that the biological implications of each cell's attributes were defined.

Statistical analysis.

Statistical analyses were conducted using Prism 10 software.

ACKNOWLEDGMENTS

We thank the Bioimaging Platform of the Research Centre for Infectious Diseases, Axe of Infectious and Inflammatory Diseases at the CHU de Quebec Research Centre for their invaluable technical support with the microscopes. We also acknowledge the Developmental Studies Hybridoma Bank for providing

the monoclonal antibody against CD133, which is created by the U.S. National Institutes of Health, Institute of Child Health and Human Development, and maintained at the Department of Biology, University of Iowa (Iowa City, IA, USA). This research received financial support from the Canadian Foundation for Innovation and the Canadian Institutes for Health Research.

AUTHOR CONTRIBUTIONS

M.S.S. and S.S. designed the study, conducted experiments, and contributed to the manuscript's writing. M.S.S. developed the software for the computerized single-cell tracking system, cell fate simulation, and 3D TME simulation. A.R. conducted experiments, analyzed single-cell tracking data, and contributed to the manuscript's editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in the paper and supporting files.

References

1. U. Ben-David, A. Amon, Context is everything: aneuploidy in cancer. *Nat Rev Genet* **21**, 44-62 (2020).
2. I. Dagogo-Jack, A. T. Shaw, Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol* **15**, 81-94 (2018).
3. A. J. Holland, D. W. Cleveland, Losing balance: the origin and impact of aneuploidy in cancer. *EMBO Rep* **13**, 501-514 (2012).
4. S. F. Bakhoum, L. C. Cantley, The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment. *Cell* **174**, 1347-1360 (2018).
5. C. Z. Zhang *et al*., Chromothripsis from DNA damage in micronuclei. *Nature* **522**, 179-184 (2015).
6. B. Vogelstein *et al*., Cancer genome landscapes. *Science* **339**, 1546-1558 (2013).
7. I. Martincorena, P. J. Campbell, Somatic mutation in cancer and normal cells. *Science* **349**, 1483-1489 (2015).
8. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719-724 (2009).
9. W. A. Freed-Pastor, C. Prives, Mutant p53: one name, many proteins. *Genes Dev* **26**, 1268-1286 (2012).
10. C. Kandoth *et al*., Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-339 (2013).
11. F. Mantovani, L. Collavin, G. Del Sal, Mutant p53 as a guardian of the cancer cell. *Cell Death Differ* **26**, 199-212 (2019).
12. P. A. Muller, K. H. Vousden, p53 mutations in cancer. *Nat Cell Biol* **15**, 2-8 (2013).

13. A. Petitjean *et al*., Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* **28**, 622-629 (2007).

14. C. Prives, S. W. Lowe, Cancer: Mutant p53 and chromatin regulation. *Nature* **525**, 199-200 (2015).

15. X. Yue *et al*., Mutant p53 in Cancer: Accumulation, Gain-of-Function, and Therapy. *J Mol Biol* **429**, 1595-1606 (2017).

16. N. Shindo, M. Otsuki, K. S. K. Uchida, T. Hirota, Prolonged mitosis causes separase deregulation and chromosome nondisjunction. *Cell Rep* **34**, 108652 (2021).

17. K. S. K. Uchida *et al*., Kinetochore stretching-mediated rapid silencing of the spindle-assembly checkpoint required for failsafe chromosome segregation. *Curr Biol* **31**, 1581-1591 e1583 (2021).

18. D. Gisselsson *et al*., When the genome plays dice: circumvention of the spindle assembly checkpoint and near-random chromosome segregation in multipolar cancer cell mitoses. *PLoS One* **3**, e1871 (2008).

19. A. J. Holland, D. W. Cleveland, Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nat Rev Mol Cell Biol* **10**, 478-487 (2009).

20. A. Rancourt, S. Sato, M. S. Satoh, Empirical single-cell tracking and cell-fate simulation reveal dual roles of p53 in tumor suppression. *eLife* **https://doi.org/10.7554/eLife.72498**, (2022).

21. I. Ben-Porath *et al*., An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**, 499-507 (2008).

22. K. Eppert *et al*., Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med* **17**, 1086-1093 (2011).

23. J. Kim *et al*., A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* **143**, 313-324 (2010).

24. T. M. Malta *et al*., Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **173**, 338-354 e315 (2018).

25. G. Anderson, Tumor Microenvironment and Metabolism: Role of the Mitochondrial Melatonergic Pathway in Determining Intercellular Interactions in a New Dynamic Homeostasis. *Int J Mol Sci* **24**, (2022).

26. S. M. Ansell, R. H. Vonderheide, Cellular composition of the tumor microenvironment. *Am Soc Clin Oncol Educ Book*, (2013).

27. A. Bozyk, K. Wojas-Krawczyk, P. Krawczyk, J. Milanowski, Tumor Microenvironment-A Short Review of Cellular and Interaction Diversity. *Biology (Basel)* **11**, (2022).

28. E. Karamitopoulou, Tumour microenvironment of pancreatic cancer: immune landscape is dictated by molecular and histopathological features. *Br J Cancer* **121**, 5-14 (2019).

29. A. Blum, P. Wang, J. C. Zenklusen, SnapShot: TCGA-Analyzed Tumors. *Cell* **173**, 530 (2018).

30. L. Fishbein, M. D. Wilkerson, Chromaffin cell biology: inferences from The Cancer Genome Atlas. *Cell Tissue Res* **372**, 339-346 (2018).

31. J. E. Rood, A. Maartens, A. Hupalowska, S. A. Teichmann, A. Regev, Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486-2496 (2022).

32. O. Rozenblatt-Rosen *et al*., The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell* **181**, 236-249 (2020).

33. A. T. Ruffin *et al*., Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment. *Nat Rev Cancer* **23**, 173-188 (2023).

34. S. Sato, A. Rancourt, Y. Sato, M. S. Satoh, Single-cell lineage tracking analysis reveals that an established cell line comprises putative cancer stem cells and their heterogeneous progeny. *Sci Rep* **6**, 23328 (2016).

35. D. R. Christie, F. M. Shaikh, J. A. t. Lucas, J. A. Lucas, 3rd, S. L. Bellis, ST6Gal-I expression in ovarian cancer cells promotes an invasive phenotype by altering integrin glycosylation and function. *J Ovarian Res* **1**, 3 (2008).

36. K. Hasehira *et al*., Structural and quantitative evidence for dynamic glycome shift on production of induced pluripotent stem cells. *Mol Cell Proteomics* **11**, 1913-1923 (2012).

37. M. J. Schultz *et al*., The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. *Cancer Res* **76**, 3978-3988 (2016).

38. A. F. Swindall *et al*., ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. *Cancer Res* **73**, 2368-2378 (2013).

39. H. Tateno *et al*., alpha2-6 sialylation is a marker of the differentiation potential of human mesenchymal stem cells. *Glycobiology* **26**, 1328-1337 (2016).

40. H. Tateno *et al*., Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem* **286**, 20345-20353 (2011).

41. P. B. dos-Santos *et al*., Lectin histochemistry reveals SNA as a prognostic carbohydrate-dependent probe for invasive ductal carcinoma of the breast: a clinicopathological and immunohistochemical auxiliary tool. *Int J Clin Exp Pathol* **7**, 2337-2349 (2014).

42. S. P. Sunny *et al*., CD44-SNA1 integrated cytopathology for delineation of high grade dysplastic and neoplastic oral lesions. *PLoS One* **18**, e0291972 (2023).

43. E. Kurz *et al*., Integrated Systems Analysis of the Murine and Human Pancreatic Cancer Glycomes Reveals a Tumor-Promoting Role for ST6GAL1. *Mol Cell Proteomics* **20**, 100160 (2021).

44. T. Wagatsuma *et al*., Discovery of Pancreatic Ductal Adenocarcinoma-Related Aberrant Glycosylations: A Multilateral Approach of Lectin Microarray-Based Tissue Glycomic Profiling With Public Transcriptomic Datasets. *Front Oncol* **10**, 338 (2020).

45. S. P. Povoski, I. S. Hatzaras, C. M. Mojzisik, E. W. Martin, Jr., Oncologic theranostics: recognition of this concept in antigen-directed cancer therapy for colorectal cancer with anti-TAG-72 monoclonal antibodies. *Expert Rev Mol Diagn* **11**, 667-670 (2011).

46. B. E. Rogers *et al*., Intraperitoneal radioimmunotherapy with a humanized anti-TAG-72 (CC49) antibody with a deleted CH2 region. *Cancer Biother Radiopharm* **20**, 502-513 (2005).

47. J. Schmolling, J. Reinsberg, U. Wagner, D. Krebs, Anti-TAG-72 antibody B72.3--immunological and clinical effects in ovarian carcinoma. *Hybridoma* **16**, 53-58 (1997).

48. S. Miraglia *et al*., A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood* **90**, 5013-5021 (1997).

49. A. Weigmann, D. Corbeil, A. Hellwig, W. B. Huttner, Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc Natl Acad Sci U S A* **94**, 12425-12430 (1997).

50. A. H. Yin *et al*., AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* **90**, 5002-5012 (1997).

51. A. T. Holdbrooks, C. M. Britain, S. L. Bellis, ST6Gal-I sialyltransferase promotes tumor necrosis factor (TNF)-mediated cancer cell survival via sialylation of the TNF receptor 1 (TNFR1) death receptor. *J Biol Chem* **293**, 1610-1622 (2018).

52. A. Chakraborty *et al*., ST6Gal-I sialyltransferase promotes chemoresistance in pancreatic ductal adenocarcinoma by abrogating gemcitabine-mediated DNA damage. *J Biol Chem* **293**, 984-994 (2018).

53. M. J. Schultz *et al*., ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J Ovarian Res* **6**, 25 (2013).

54. X. Wang *et al*., TIMEDB: tumor immune micro-environment cell composition database with automatic analysis and interactive visualization. *Nucleic Acids Res* **51**, D1417-D1424 (2023).

55. V. Thorsson *et al*., The Immune Landscape of Cancer. *Immunity* **51**, 411-412 (2019).

56. M. M. Lee *et al*., Platelets support extracellular sialylation by supplying the sugar donor substrate. *J Biol Chem* **289**, 8742-8748 (2014).

57. D. Pally *et al*., Heterogeneity in 2,6-Linked Sialic Acids Potentiates Invasion of Breast Cancer Epithelia. *ACS Cent Sci* **7**, 110-125 (2021).

Figure legends

Fig. 1. Conceptual Workflow for Performing a 3D TME Simulation

A. The cancer cell population within the TME exhibits inherent phenotypic diversity and dynamically evolves over time. **B.** To uncover the temporal dynamics of cancer cells within the TME, a conceptual workflow for performing a 3D TME simulation was developed. It started with the acquisition of spatiotemporal information about individual cancer cells using a computerized single-cell tracking system. Subsequently, inferred cell populations that signify particular attributes of cells in vivo were generated using a cell fate simulation algorithm, followed by the execution of a 3D TME simulation. **C.** To capture the spatiotemporal behaviors of cancer cells, single-cell tracking was performed using established cancer cell lines. This involved creating live cell imaging videos of cultured cancer cells and tracking individual cells. The resulting data resources included cell lineage information. Quantitative analysis of 2-6Sia expression levels was conducted at the end of imaging and integrated into the cell lineage data. **D.** Utilizing the cell lineage data, inferred cell populations were generated by focusing on specific characteristics of cancer cells using a cell fate simulation algorithm. **E.** The inferred cancer cell population, along with immune cells, was introduced into a 3D TME generated by artificial intelligence for the simulation. In this simulation, inferred cancer cells (depicted in blue) coexisted with immune cells (depicted in yellow, green, and red) within a sphere.

Fig. 2. Characteristics of cervical cancer cell line (HeLa) and pancreatic cancer cell line (MiaPaCa2).

A. Single-cell tracking was conducted with HeLa and MiaPaCa2 cells, and cell counts at each time point were plotted to determine cell population expansion curves. **B**. The counts of multipolar cell divisions (MD), cell deaths (CD), and cell fusions (CF) occurring within cell lineages were recorded. Notably, HeLa cells exhibited a higher count than MiaPaCa2 cells. Statistical analysis was performed using ordinary one-way ANOVA (Tukey's), and statistical significance was indicated as ****P <

0.0001. A DIC image illustrates a cell undergoing tripolar cell division. **C**. The fate of cells resulting from cell fusion and the progeny produced by multipolar cell division were explored. Possible patterns are illustrated. For instance, 'CF > BD' illustrates that bipolar cell division (BD) occurs following cell fusion (CF). **D**. The fate of cells resulting from cell fusion is shown. **E**. The fate of progeny produced through multipolar cell division is presented. **F**. The total numbers of bipolar cell divisions (BDs), multipolar cell divisions (MDs), cell deaths (CDs), and cell fusions (CFs) that occurred following cell fusion (CF) or multipolar cell division (MD) were determined. Statistical analysis was conducted using ordinary one-way ANOVA (Kruskal-Wallis test), and statistical significance was indicated as **** P < 0.0001. **G**. Cell fate simulation was performed using single-cell tracking data of HeLa and MiaPaCa2 cells, and the number of reproductive progeny derived from cells produced through multipolar cell division was determined. **H**. A magnified result is shown. MD-Progeny: Reproductive progeny produced through multipolar cell division. The percentage indicated in the parentheses is the percentage of reproductive progeny within the total HeLa or MiaPaCa2 cell population.

Fig. 3. Analysis of 2-6Sia expression using SNA1.

A. Fluorescence-tagged rBC2LCN, SNA1, anti-CD133, and anti-TAG-72 staining was performed on HeLa cells (**A**) and MiaPaCa2 cells (**B**), followed by visualization using Alexa 488-labeled secondary antibodies and fluorescence imaging. Cell nuclei were counterstained with DAPI (blue). DIC images are provided for reference. **C**. Illustration depicting membrane proteins with attached oligosaccharides. These oligosaccharides feature a sialic acid α2-6 galactose structure (2-6Sia) at the termini, recognized by Sambucus Nigra lectin (SNA1). **D**. Visualization of individual cell 2-6Sia expression levels by arranging cells on a spherical surface and color-coding them using a heatmap scale. **E**. Count of cells within a HeLa cell lineage and determination of the total SNA1 binding value associated with the lineage. A dotted circle indicates the cell lineage with higher SNA1 binding values. **F**. Count of cells within a MiaPaCa2 cell lineage and determination of the total SNA1 binding value linked to the

lineage. Statistical analysis was performed using simple linear regression, resulting in an R^2 value of 0.51. **G** (HeLa) and **H** (MiaPaCa2). Count of cell deaths within a cell lineage plotted against the total SNA1 binding value associated with the lineage. Cell death counts were categorized into groups: below SNA1 binding value 2000, 2000 to 3999, and more than 4000, with percentages shown. **I** (HeLa) and **J** (MiaPaCa2). Count of multipolar cell divisions within a cell lineage plotted against the total SNA1 binding value associated with the lineage. Multipolar cell division counts were categorized into groups: below SNA1 binding value 2000, 2000 to 3999, and more than 4000, with percentages shown. **K** (HeLa) and **L** (MiaPaCa2). Count of cell fusions within a cell lineage plotted against the total SNA1 binding value associated with the lineage. Cell fusion counts were categorized into groups: below SNA1 binding value 2000, 2000 to 3999, and more than 4000, with percentages shown.

Fig. 4. Generation of inferred cell populations.

A. A list containing the 2-6Sia expression levels of all analyzed HeLa or MiaPaCa2 cells obtained through single-cell tracking was compiled. This list served as a reference when assigning 2-6Sia expression levels to inferred cells. Subsequently, a progenitor cell was created by assigning the length of time until a cellular event occurred and the type of that event using the cell fate simulation algorithm. In the case of bipolar cell division, two daughter cells were generated, and each was assigned a 2-6Sia expression level using the 2-6Sia expression level list. These expression levels were later modified to introduce variations in 2-6Sia expression levels. Additional details are available in the Materials and Methods and Supplementary Fig. 4. **B** (HeLa; light blue line and Cervical 2-6Sia; blue lines) and **C** (MiaPaCa2; light blue line and Pancreatic 2-6Sia; blue lines). Inferred cell populations were generated using a cell fate simulation algorithm by varying 2-6Sia expression levels. The cell population expansion curves for these cells are shown. **D.** The expression levels of 2-6Sia in the inferred cell population are presented. As the cell fate simulation algorithm incorporated the recorded 2-6Sia expression patterns from cell lineage data, the inferred cell population's 2-6Sia expression levels

reflect those of HeLa and MiaPaCa2 cells. These cell populations were designated as Cervical 2-6Sia 1.5, 1.0, 0.5, and 0.25, and Pancreatic 2-6Sia 1.5, 1.0, 0.5, and 0.25.

Fig. 5. Categorization of immune cells.

A. Immune cells within the TME were classified into Suppressive, Permissive, and Lethal cells. When a cancer cell is located within a defined spatial distance of a Suppressive cell, it experiences an extended doubling time. The extension occurs once during the lifetime of the cell. The doubling time of its progeny is also extended, but the progeny could be subject to another Suppressive, Permissive, or Lethal effect. When a cancer cell is within a specified spatial distance of a Permissive cell, it experiences a shortened doubling time. The shortening occurs once during the lifetime of the cell. The doubling time of its progeny is also shortened, but the progeny could be subject to another Suppressive, Permissive, or Lethal effect. When both Suppressive and Permissive cells are within a specified distance, their impact on cancer cells depends on their respective distances from the cancer cell. When a cancer cell is within a specified spatial distance of a Lethal cell, cell death is marked to the cell, and its progeny are removed from the cell lineage database. **B.** A 3D TME simulation was conducted using Cervical 2-6Sia 1.5 cells in the presence of Suppressive, Permissive, and Lethal cells, represented by red, light blue, and yellow spheres, respectively (denoted as smaller spheres). The expression levels of 2-6Sia were visualized using a heatmap scale, ranging from blue to red (denoted as a larger size of the spheres). For Control, the Suppressive, Permissive, and Lethal strengths were set to 0. For Lethal strength 100, the Suppressive, Permissive, and Lethal strengths were set to 0, 0, and 100, respectively.

Fig. 6. Relationships of simulation results with Progress-free survival and 2-6Sia expression.

Four different populations of Cervical 2-6Sia cells (**A-C**): 2-6Sia 1.5, 2-6Sia 1.0, 2-6Sia 0.5, and 2-6Sia 0.25, along with four populations of Pancreatic 2-6Sia cells (**D-F**): 2-6Sia 1.5, 2-6Sia 1.0, 2-6Sia 0.5, and 2-6Sia 0.25, were employed in 3D TME simulations using TME cell landscape data from 193

cervical and 172 pancreatic cancer cases. Assessed parameters include the cell population size at the end of the simulation (**A** and **D**), 2-6Sia expression levels (**B** and **E**), and the number of reproductive cells generated through multipolar cell division (**C** and **F**). Simulation results were plotted against Progression-Free Survival (PFS). Variation of each data relative to the linear regression line was calculated. **G** and **H**. Simulation results are presented for Cervical 2-6Sia 1.0 (**G**) and Pancreatic 2-6Sia 1.0 (**H**). The simulated cell population size is plotted against 2-6Sia expression levels in (**G** and **H**). The resulting R^2 values were shown. Each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case.

Fig. 7 Characteristics of cervical cancer cells within the TME.

A-C. The simulated population size of Cervical 2-6Sia cell populations (2-6Sia 1.5, 2-6Sia 1.0, 2-6Sia 0.5, and 2-6Sia 0.25) is plotted against the initial proportion of Suppressive cells (A), Permissive cells (**B**), and Lethal cells (**C**). Variation of each data relative to the linear regression line was calculated. The resulting R2 values were shown. **D-G**. Cell population sizes (**D**), 2-6Sia expression levels (**E**), and the number of reproductive cells produced through multipolar cell division (**F**) of Cervical 2-6Sia (1.5, 1.0, 0.5, and 0.25) were analyzed. The percentage of the number of reproductive cells produced through multipolar cell division relative to the total cell population is also shown **(G).** Statistical analysis was conducted using ordinary one-way ANOVA (Tukey's), and significance is denoted as ****$P < 0.0001$, while "ns" indicates no significance. The results of the analysis are shown between Cervical 2-6Sia 1.5 and the other populations. **H**: The summarized characteristics of cervical cancer cells within the TME context are presented. Each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case.

Fig. 8 Characteristics of pancreatic cancer cells within the TME.

A-C. The simulated population size of Pancreatic 2-6Sia cell populations (2-6Sia 1.5, 2-6Sia 1.0, 2-6Sia 0.5, and 2-6Sia 0.25) is plotted against the initial proportion of Suppressive cells (**A**), Permissive cells (**B**), and Lethal cells (**C**). Variation of each data relative to the linear regression line was calculated. The resulting R^2 values were shown. **D-G.** Cell population sizes (**D**), 2-6Sia expression levels (**E**), and the number of reproductive cells produced through multipolar cell division (**F**) of Pancreatic 2-6Sia (1.5, 1.0, 0.5, and 0.25) were analyzed. The percentage of the number of reproductive cells produced through multipolar cell division relative to the total cell population is also shown (**G**). Statistical analysis was conducted using ordinary one-way ANOVA (Tukey's), and significance is denoted as **P < 0.01 and ****P < 0.0001, while "ns" indicates no significance. The results of the analysis are shown between Pancreatic 2-6Sia 1.5 and the other populations. **H**: The summarized characteristics of pancreatic cancer cells within the TME context are presented. Each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case.

Fig. 9: Conceptual overview of 3D TME simulation with spatiotemporal information of individual cancer cells

The spatiotemporal behaviors of cancer cells are unveiled through single-cell tracking of established cancer cell lines. Furthermore, insights into the characteristics of cancer cells are obtained from existing published data. Merging this information, inferred cell populations are created using a cell fate simulation algorithm. These resulting cell populations are then utilized to simulate the fate of cancer cells within the TME, which comprises immune cells and other resident cells.

Supplementary Video 1. Single-cell tracking of HeLa cell.

A video for single-cell tracking of HeLa cells is shown (time points 1–693). A unique color was assigned to each cell lineage. The total number of tracked cells was 11,640.

Supplementary Video 2. Single-ell tracking of MiaPaCa2 cell.

A video for single-cell tracking of MiaPaCa2 cells is shown (time points 1–597). A unique color was assigned to each cell lineage. The total number of tracked cells was 6,987.

Supplementary Fig. 1. Computerized single-cell lineage tracking analysis system.

The computerized single-cell lineage tracking analysis system included Computer 1 to control the microscope and Computer 2 to perform image processing, single-cell tracking, and data analysis. Commercially available image acquisition software (Metamorph) was installed on Computer 1 to control the microscope and create image files. Computer 2 employed various custom software programs; *Image Processing Controller* controlled other software programs; *File Transfer* and *Map* communicated between Computers 1 and 2; *File Converter* imported images created by other microscopes; *Name Assignment*, *Focal Image Selection*, and *Contrast Set* produced live cell videos; *Data Backup* controlled file archiving and backup; *Outline Drawing* performed image segmentation and created the object segmentation library; *Movie Viewer* played movies and fine-tuned image quality; *Object Tracking Controller* created the cell lineage database by controlling *Automatic Object Tracking*, and facilitated data verification; and *Data Analysis* provided various options for data analysis. DIC/FL; differential interference contrast/fluorescence. FOV; field of view.

Supplementary Fig. 2. Image segmentation.

A. Image segmentation was performed using the Stepwise Area Expansion method. Pixels with values exceeding predefined thresholds (THs) ranging from TH1 to TH4 were extracted from the NIR-DIC image and assigned a grayscale value of 150 (ranging from 0 to 255 grayscale, such as 200, 170, 140, and 110 for TH1, 2, 3, and 4, respectively). The segmented results are displayed, with areas enclosed by blue circles representing the segmented regions. **B.** Connectivity analysis was conducted using TH1 to identify connected pixel edges, which were marked with circles (referred to as 'edge circles,' as

shown in **a**: Edge circles, indicated by pink lines. These circles were overlaid onto the NIR-DIC image, and lines were extended from pixels within these regions toward the 12, 3, 6, and 9 o'clock directions (**b**: Line extension, depicted in red, yellow, light blue, and blue regions, respectively, along with a magnified view of Line extension). The endpoints of each line were linked (**c**: Edge linking, and a magnified view of Edge linking), creating new edge circles by connecting the linked lines (**d**: Creation of new areas, marked by green lines). These newly formed edge circles were then overlaid onto the TH2 image (**e**: represented by green circles), and pixels that did not overlap with green circles were identified (**f**: indicated by pink circles). The processing of pink circles followed a similar procedure as for TH1 (**f-h**). The green circles were expanded by 1–4 pixels outside the edge circles (**k** and **l**, along with an enlarged view of Edge expansion). Finally, the green and red edge circles were overlaid to detect overlaps (**i**: Overlay), which were subsequently eliminated (**j**: Removal of overlaps). These processes were replicated for the TH3 and TH4 images.

Supplementary Fig. 3. Object tracking (single-cell tracking).

A. At Time A, the segmented areas identified as cell representations are denoted by orange characters (cell lineage numbers). The cell being tracked is highlighted by a blue area with a white asterisk. Additionally, the yellow and magenta areas represent segmented regions corresponding to neighboring cells. **B.** At Time A+1, there were changes in the positions of cells and segmentation patterns (indicated by black characters representing segmented area numbers). **C.** The blue area from Time A was overlaid onto the Time A+1 image. However, the blue segmented area observed at Time A was not present in Time A+1; instead, it was partially overlaid with the light blue and white segmented areas. **D.** To identify which segmented area corresponds to the tracked cell, the yellow and magenta areas were overlaid on the Time A+1 image. This process revealed that the light blue area was associated with the yellow area. Furthermore, the magenta area overlapped with the white area, with the larger portion of the white area overlapping with the blue area. **E.** Consequently, it was determined that the white area

represents the segmented region corresponding to the tracked cell. This systematic process was repeated for all cells recorded in the database.

Supplementary Fig. 4. Outline of the Generation of Inferred Cell Populations.

A. The initial step in generating the inferred cell population involved creating a list of 2-6Sia expression levels. SNA1 binding levels of individual cells, measured at the end of the live cell, are displayed on the cell lineage map using a heatmap scale, where blue to red corresponds to low to high expression. For example, Lineage 1 exhibited no expression of 2-6Sia in any of the cells, whereas Lineage 2 showed various levels of 2-6Sia expression in its cells. Based on the 2-6Sia expression levels of cells, we traced back along the cell lineage map. Subsequently, a list of 2-6Sia expression levels for all tracked HeLa or MiaPaCa2 cells was compiled. B. The inferred cell population was generated using a cell fate simulation algorithm. The algorithm assigned the length of time for the First event (the event that occurred in the progenitor cell) to the progenitors. Then, a cellular event was assigned to a progenitor cell. If a bipolar cell division was assigned, two daughter cells were created. In Example 1, a longer time interval until the Next event was assigned compared to Example 2. After the assignment, the algorithm referred to the cell doubling time (the time between bipolar cell divisions) of 2-6Sia-expressing cells. It checked whether any of the 2-6Sia-expressing cells had cell doubling times within $\pm$10% of the time assigned to daughter cells. If no such cells were found, 2-6Sia expression levels were not assigned to the daughter cells (Example 1). If some cells met this criterion, 2-6Sia expression levels were assigned to the daughter cells (Example 2). **C.** The algorithm assigned 2-6Sia expression levels to the daughter cells by referencing the 2-6Sia expression level list. A random value was generated to select one value from the list, which was then assigned to one of the daughter cells, followed by assignment to the second daughter cell. Finally, 2-6Sia expression levels were modified to create inferred cell populations with various levels of 2-6Sia expression. For additional details, see the Materials and Methods.

Supplementary Fig. 5. Validation 3D TME simulations.

A-H: Validation simulations of the 3D TME were conducted by varying the ratio of cancer cells to Suppressive, Permissive, or Lethal cells. **A-C**: Suppressive, Permissive, or Lethal cells were introduced alongside Cervical 2-6Sia 1.0 cells in the 3D TME. (**A**) Suppressive cells, (**B**) Permissive cells, and (**C**) Lethal cells. Simulations were performed by changing the ratio from 1:0 to 1:0.9 for Suppressive and Permissive cells and from 1:0 to 1:0.46 for Lethal cells. **D**: Both Suppressive and Permissive cells were placed in the 3D TME, with a cancer cell to Suppressive cell ratio of 1:0.2, and the ratio of Permissive cells changed from 1:0 to 1:0.5. **E-H**: Suppressive, Permissive, or Lethal cells were introduced alongside Pancreatic 2-6Sia 1.0 cells in the 3D TME. (**E**) Suppressive cells, (**F**) Permissive cells, and (**G**) Lethal cells. Simulations were performed by changing the ratio from 1:0 to 1:0.9 for Suppressive and Permissive cells and from 1:0 to 1:0.46 for Lethal cells. **H**: Both Suppressive and Permissive cells were placed in the 3D TME, with a cancer cell to Suppressive cell ratio of 1:0.2, and the ratio of Permissive cells changed from 1:0 to 1:0.5. **I-L**: Control simulations of the 3D TME were conducted by varying the Suppressive, Permissive, and Lethal strength. Suppressive, Permissive, or Lethal cells were placed alongside Cervical 2-6Sia 1.0 cells in the 3D TME. (**I**) Suppressive cells, (**J**) Permissive cells, and (**K**) Lethal cells. Simulations involved changing the Suppressive strength (doubling time increase) from 0% to 90% and the Permissive strength (doubling time decrease) from 0% to 90%. For Lethal cells, the simulation varied the chance of inducing cell death from 0% to 44.5%. **L**: Both Suppressive and Permissive cells were placed in the 3D TME, with a 32.5% increase in the doubling time for Suppressive cells, and the ratio of Permissive cell doubling time changed from 0% to 100%. **M-P**: Suppressive, Permissive, or Lethal cells were introduced alongside Pancreatic 2-6Sia 1.0 cells in the 3D TME. (**M**) Suppressive cells, (**N**) Permissive cells, and (**O**) Lethal cells. **P**: Both Suppressive and Permissive cells were placed in the 3D TME.

Supplementary Fig. 6. Characteristics related to 2-6Sia expression.

Simulation results are presented for Cervical 2-6Sia 1.0 (**A** and **B**) and Pancreatic 2-6Sia 1.0 (**C** and **D**). The simulated cell population size is plotted against the number of reproductive cells produced through multipolar cell division in (**A** and **C**), and the number of reproductive cells is plotted against 2-6Sia expression levels in (**B** and **D**). XY plots were performed and variation of each data relative to the linear regression line was calculated. The resulting R^2 values were shown. Each dot signifies the outcome of simulations using the TME cell landscape data of each cancer case.

Supplementary Data 1. Cell lineage maps of HeLa cells.

Cell-lineage maps of HeLa cells were generated using data from the cell-lineage database. In these maps, the following symbols are used to represent specific events: Light blue circle, Mitosis; Pink square, Cell death; Orange circle and orange vertical line, Cell fusion; Black vertical line, Bipolar cell division; and Red vertical line, Multipolar cell division. Levels of SNA1 binding to each cell were shown at the end of the cell lineage map using a heat map scale from blue to red, indicating low to high binding.

Supplementary Data 2. Cell lineage maps of MiaPaCa2 cells.

Cell-lineage maps of MiaPaCa2 cells were generated using data from the cell-lineage database. In these maps, the following symbols are used to represent specific events: Light blue circle, Mitosis; Pink square, Cell death; Orange circle and orange vertical line, Cell fusion; Black vertical line, Bipolar cell division; and Red vertical line, Multipolar cell division. Levels of SNA1 binding to each cell were shown at the end of the cell lineage map using a heat map scale from blue to red, indicating low to high binding.

Supplementary Data 3. Cell lineage maps of Cervical 2-6Sia 1.5 cells.

Cell-lineage maps of Cervical 2-6Sia 1.5 cells were generated using data from the cell-lineage database. In these maps, the following symbols are used to represent specific events: Light blue circle, Mitosis; Pink square, Cell death; Orange circle and orange vertical line, Cell fusion; Black vertical line, Bipolar cell division; and Red vertical line, Multipolar cell division. Levels of 2-6Sia expression of individual cells were shown using a heat map scale from blue to red, indicating low to high binding.

Supplementary Data 4. Cell lineage maps of Pancreatic 2-6Sia 1.5 cells.

Cell-lineage maps of Pancreatic 2-6Sia 1.5 cells were generated using data from the cell-lineage database. In these maps, the following symbols are used to represent specific events: Light blue circle, Mitosis; Pink square, Cell death; Orange circle and orange vertical line, Cell fusion; Black vertical line, Bipolar cell division; and Red vertical line, Multipolar cell division. Levels of 2-6Sia expression of individual cells were shown using a heat map scale from blue to red, indicating low to high binding.

Supplementary Data 5. Summary of results (Pancreatic 2-6Sia 1.5, 1.0, 0.5 and 0.25).

The summary table comprises the following information: patient ID, the Best call of immune cell landscape (C1 to C6), Progression-free survival (PFS), cell population size, 2-6Sia levels, the number of reproductive cells generated through multipolar cell division (MD), and the relative composition of Suppressive, Permissive and Lethal cells.

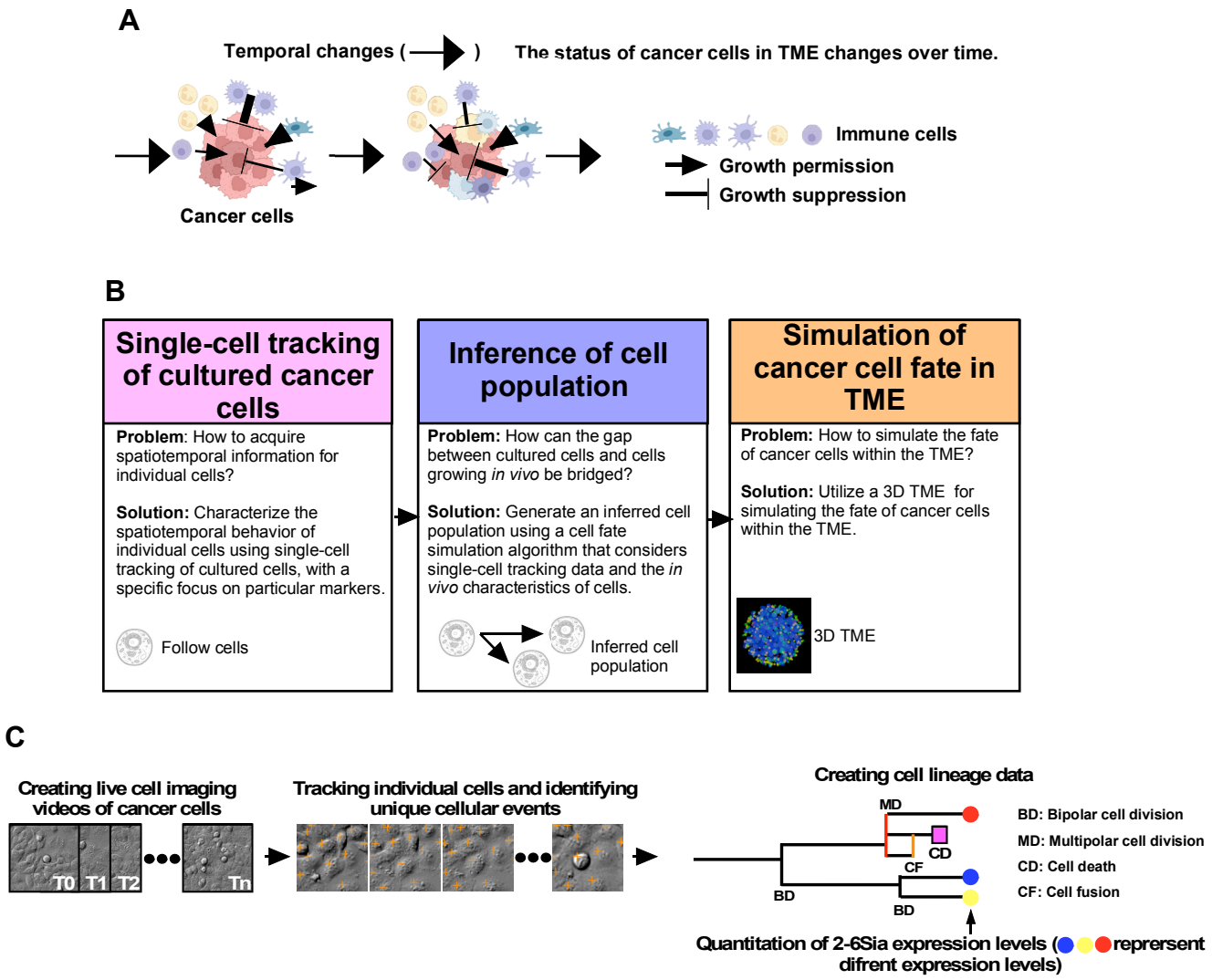
Supplementary Data 6. Summary of results (Cervical 2-6Sia 1.5, 1.0, 0.5 and 0.25).

The summary table comprises the following information: patient ID, the Best call of immune cell landscape (C1 to C6), Progression-free survival (PFS), cell population size, 2-6Sia levels, the number of reproductive cells generated through multipolar cell division (MD), and the relative composition of Suppressive, Permissive and Lethal cells.

Supplementary pseudocode 1.

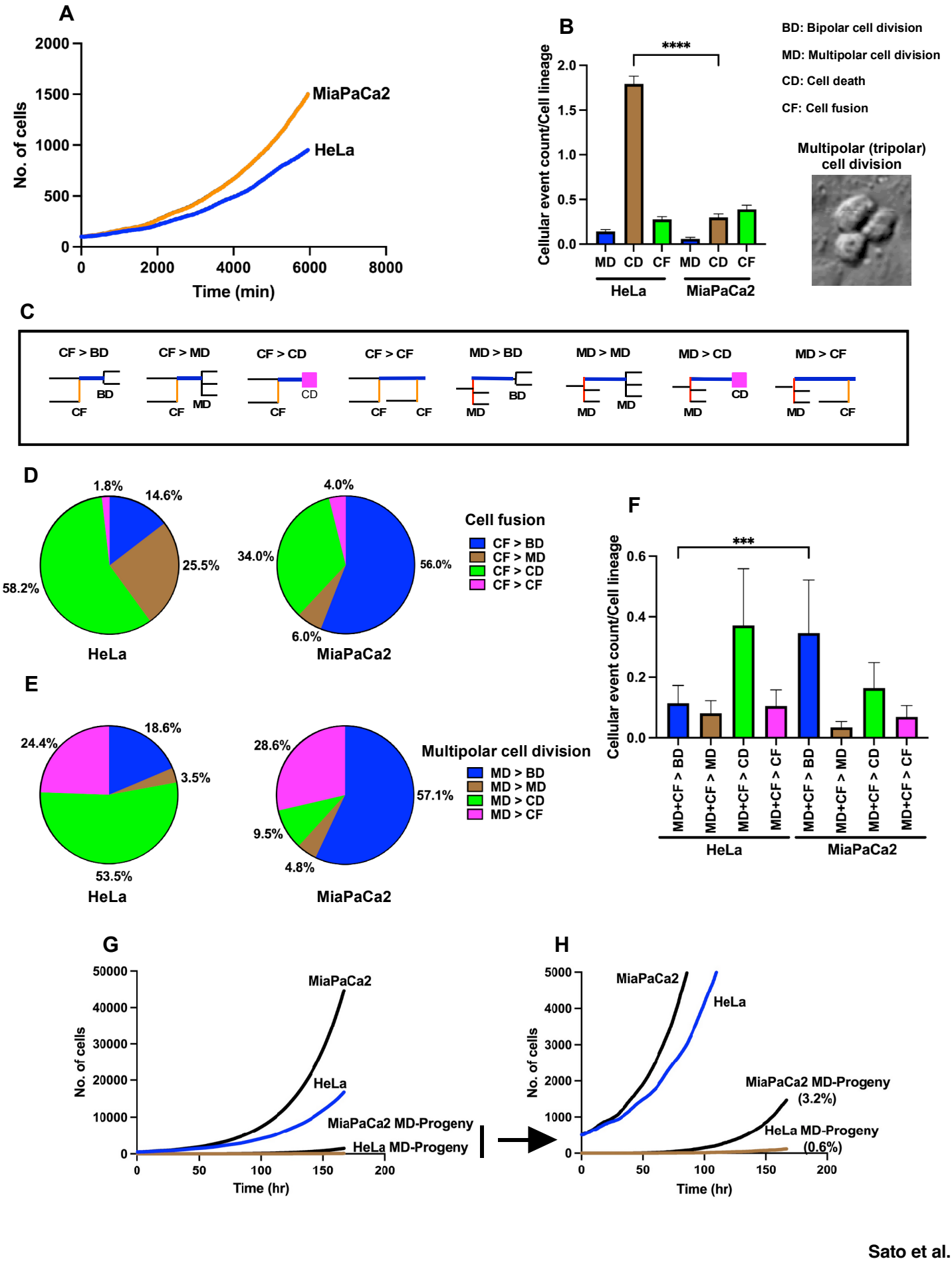
A pseudocode for the 3D TME simulation was provided. See details for Materials and Methods.

Fig. 1



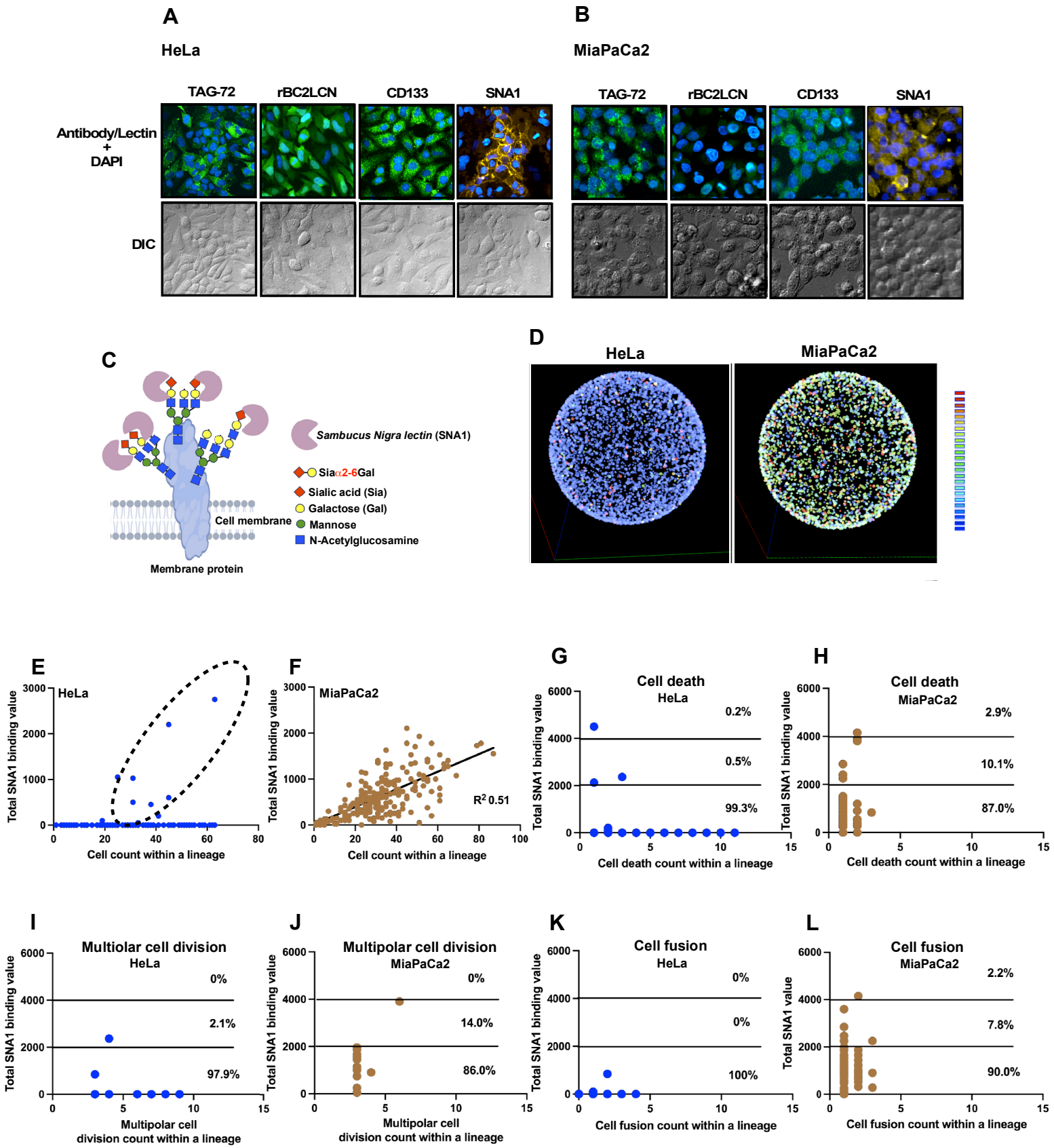
Sato et al.

Fig. 2



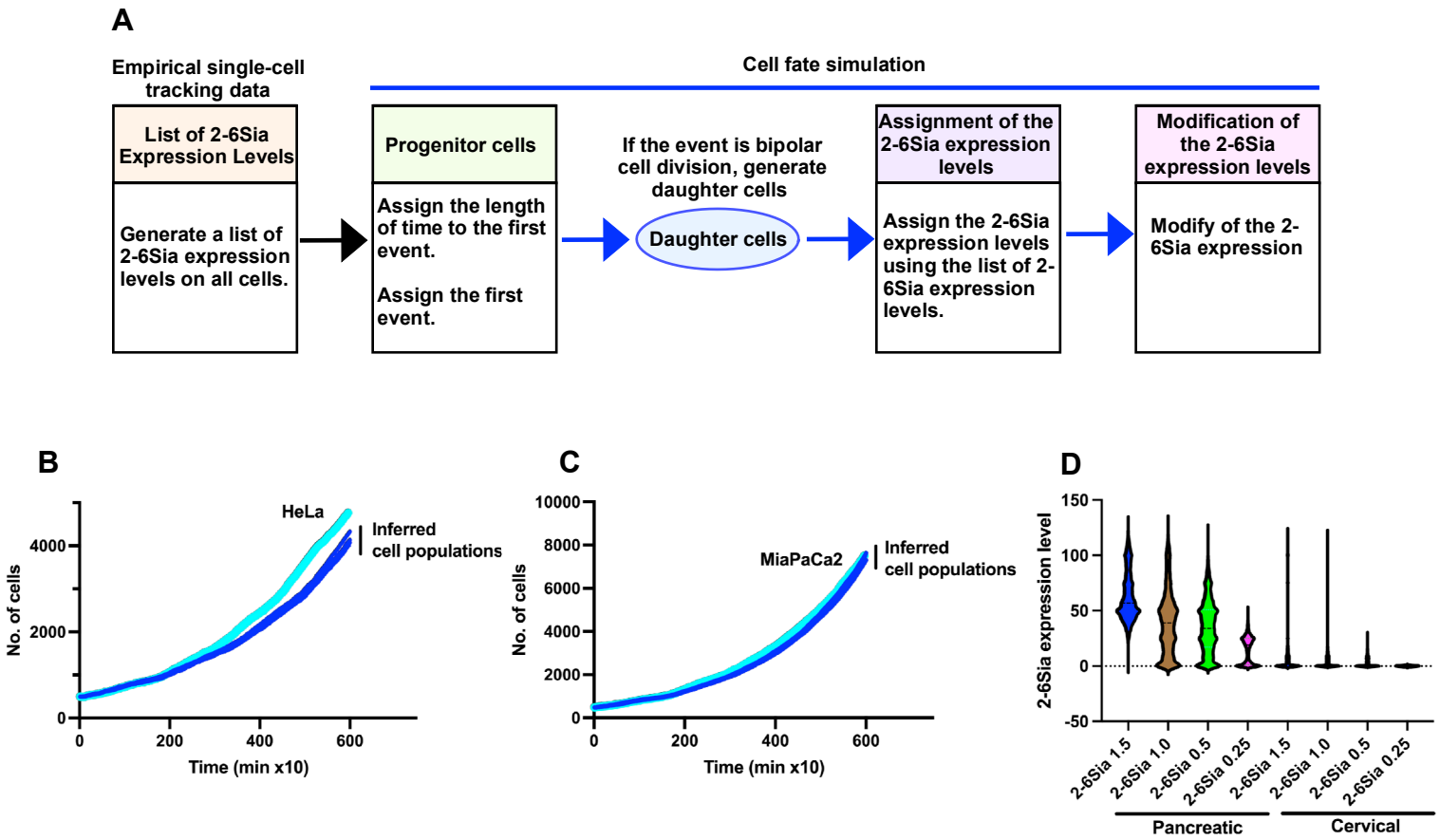
Sato et al.

Fig. 3



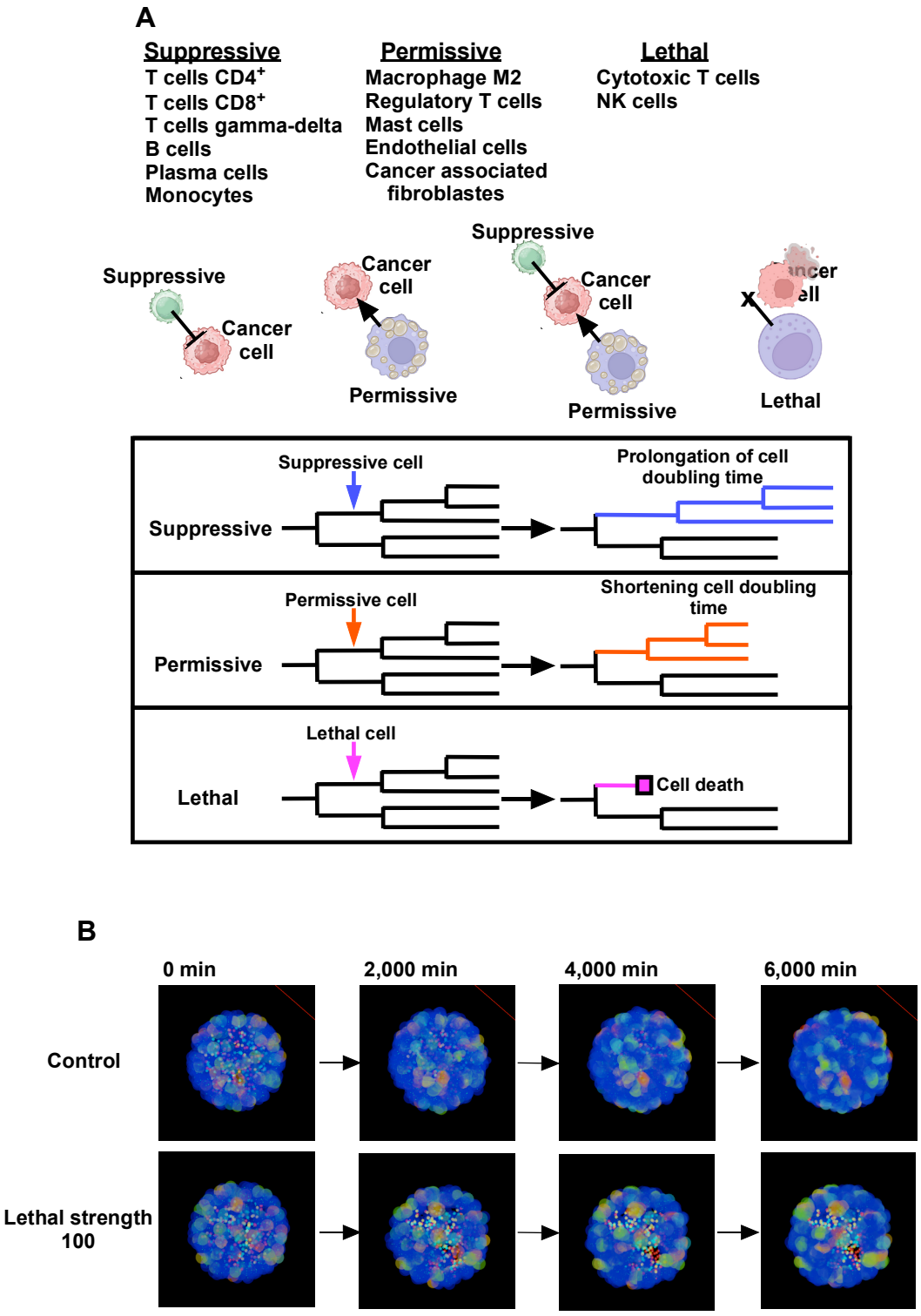
Sato et al.

Fig. 4



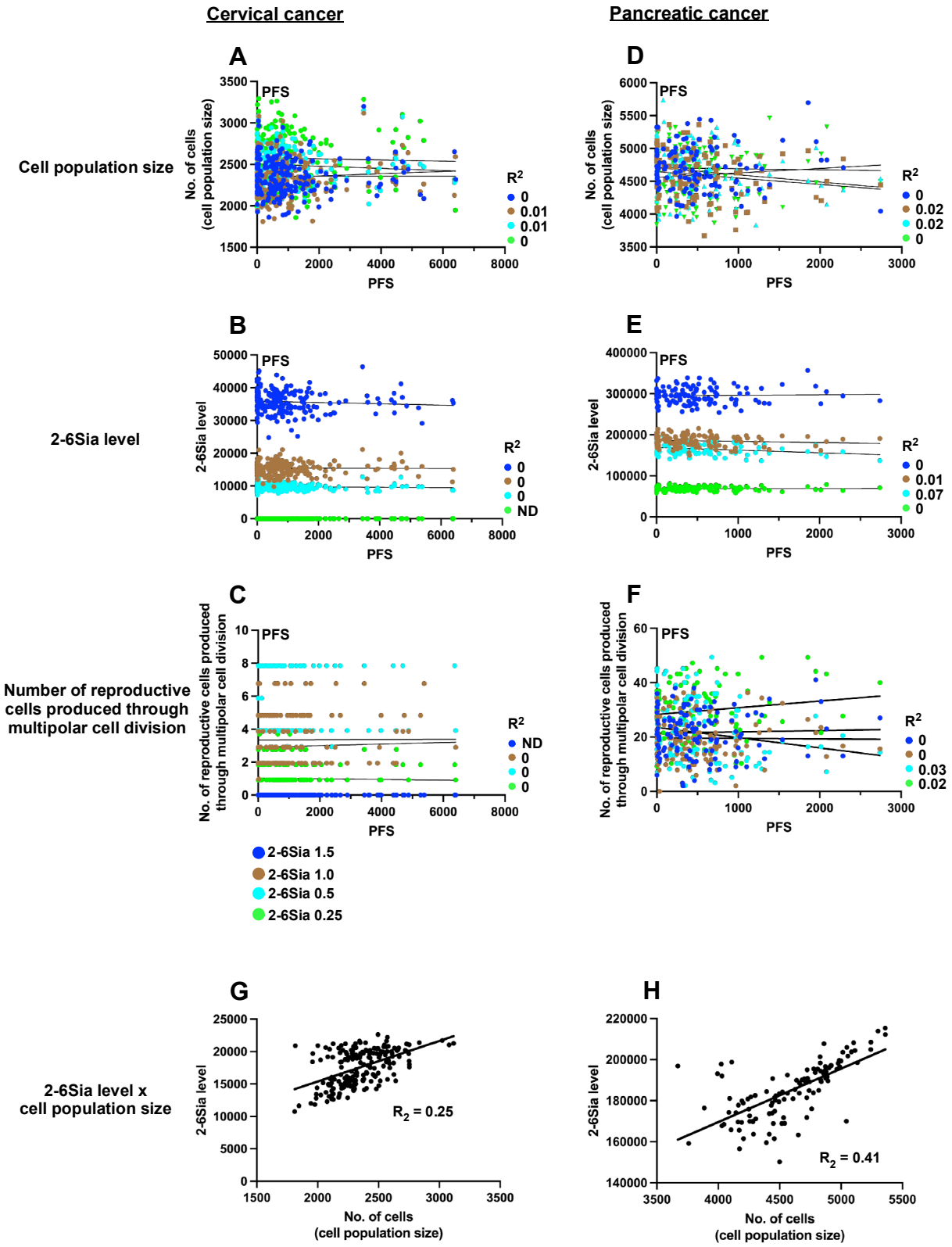
Sato et al.

Fig. 5



Sato et al.

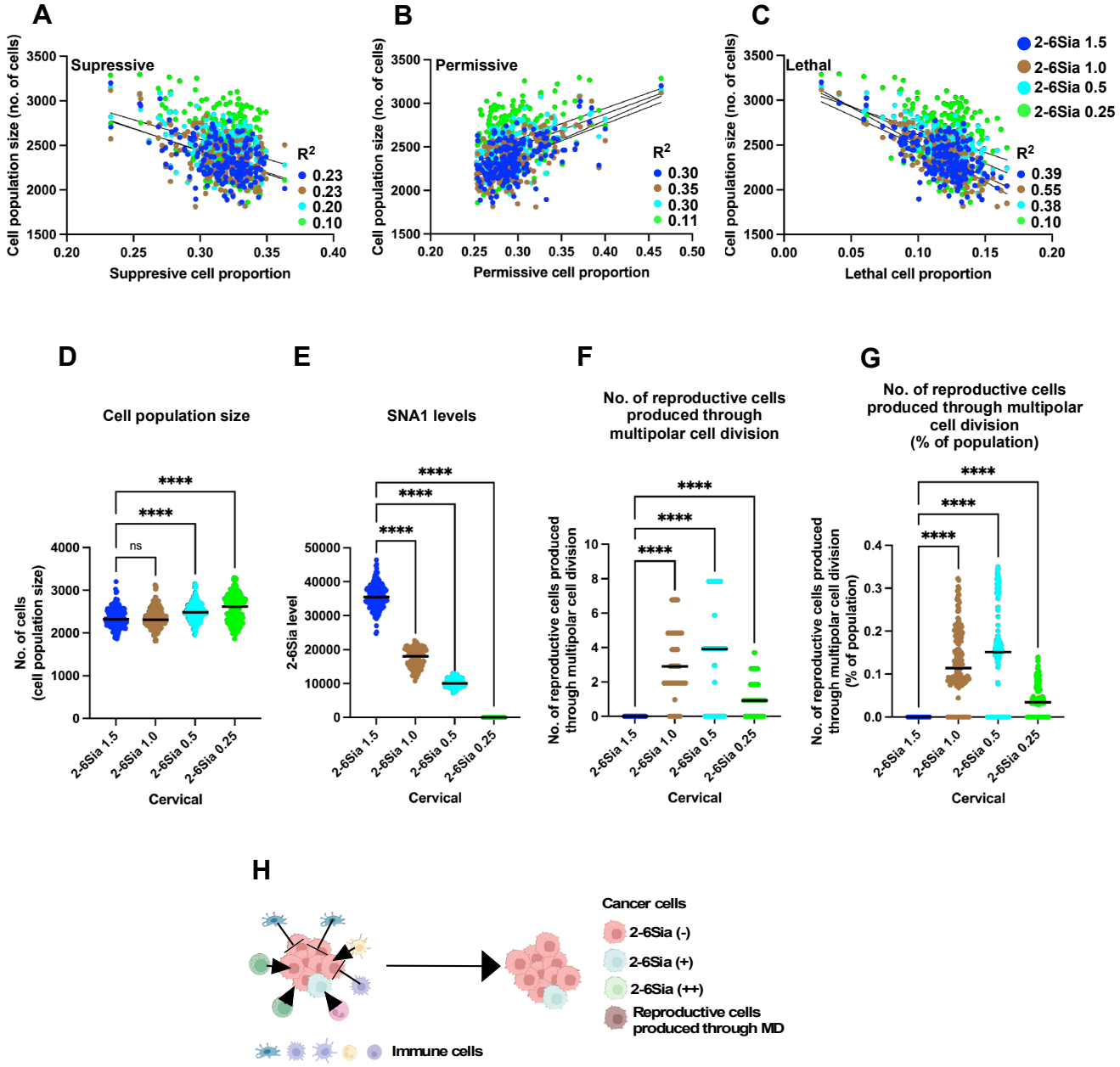
Fig. 6



Sato et al.

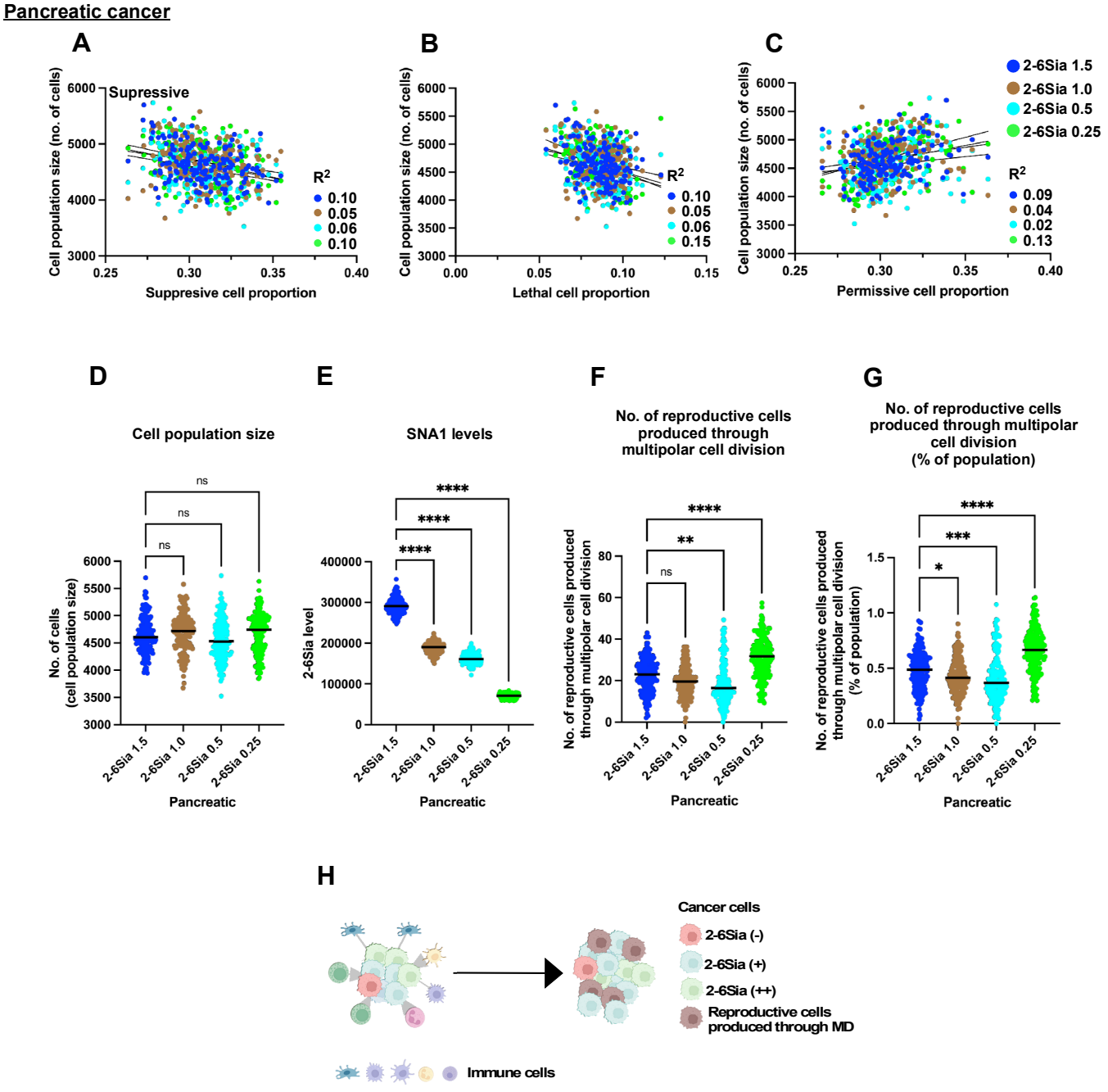
Fig. 7

Cervical cancer



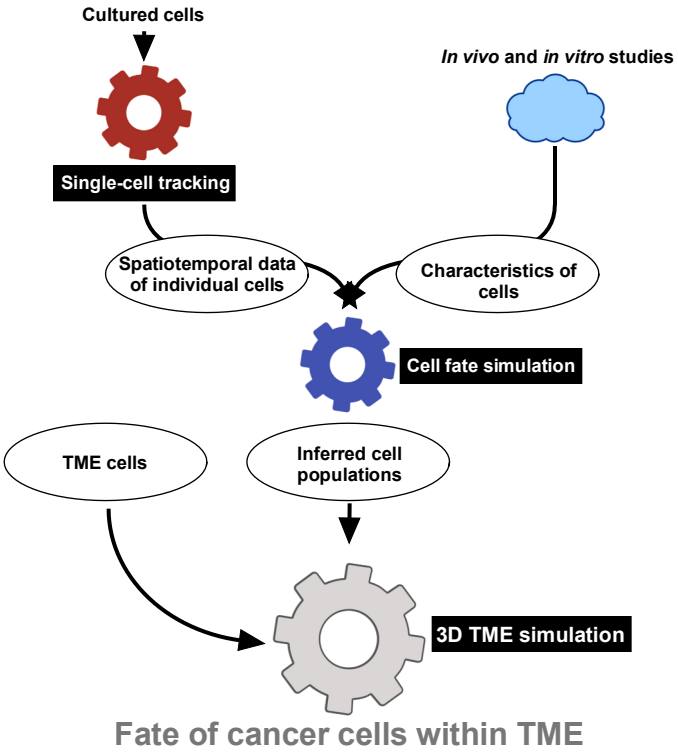
Sato et al.

Fig. 8



Sato et al.

Fig. 9


Cultured cells
In vivo and *in vitro* studies
Single-cell tracking
Spatiotemporal data of individual cells
Characteristics of cells
Cell fate simulation
TME cells
Inferred cell populations
3D TME simulation
Fate of cancer cells within TME