The University of Osaka
Institutional Knowledge Archive

Title	Single-cell RNA sequencing and machine learning provide candidate drugs against drug-tolerant persister cells in colorectal cancer
Author(s)	Nojima, Yosui; Yao, Ryoji; Suzuki, Takashi
Citation	Biochimica et Biophysica Acta - Molecular Basis of Disease. 2025, 1871(3), p. 167693
Version Type	VoR
URL	https://hdl.handle.net/11094/100580
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

BBA - Molecular Basis of Disease 1871 (2025) 167693

Contents lists available at ScienceDirect

BBA - Molecular Basis of Disease

journal homepage: www.elsevier.com/locate/bbadis

Single-cell RNA sequencing and machine learning provide candidate drugs against drug-tolerant persister cells in colorectal cancer

Yosui Nojima [a,*], Ryoji Yao [b], Takashi Suzuki [a,*]

[a] Center for Mathematical Modeling and Data Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
[b] Department of Cell Biology, Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto-ku, Tokyo 135-8550, Japan

ARTICLE INFO

Keywords:
Drug-tolerant persister cell
Patient-derived organoid
Familial adenomatous polyposis
Colorectal cancer
Single-cell RNA sequencing
Machine learning

ABSTRACT

Drug resistance often stems from drug-tolerant persister (DTP) cells in cancer. These cells arise from various lineages and exhibit complex dynamics. However, effectively targeting DTP cells remains challenging. We used single-cell RNA sequencing (scRNA-Seq) data and machine learning (ML) models to identify DTP cells in patient-derived organoids (PDOs) and computationally screened candidate drugs targeting these cells in familial adenomatous polyposis (FAP), associated with a high risk of colorectal cancer. Three PDOs (benign and malignant tumor organoids and a normal organoid) were evaluated using scRNA-Seq. ML models constructed based on public scRNA-Seq data classified DTP versus non-DTP cells. Candidate drugs for DTP cells in a malignant tumor organoid were identified from public drug sensitivity data. From FAP scRNA-Seq data, a specific TC1 cell cluster in tumor organoids was identified. The ML model identified up to 36 % of TC1 cells as DTP cells, a higher proportion than those for other clusters. A viability assay using a malignant tumor organoid demonstrated that YM-155 and THZ2 exert synergistic effects with trametinib. The constructed ML model is effective for DTP cell identification based on scRNA-Seq data for FAP and provides candidate treatments. This approach may improve DTP cell targeting in the treatment of colorectal and other cancers.

1. Introduction

In colorectal cancer (CRC), inactivation of the adenomatous polyposis coli (*APC*) gene often occurs as the initial event in the development of preneoplastic asymptomatic lesions, known as adenomatous polyps [1]. The step-wise accumulation of genetic alterations confers a malignant phenotype [1]. Familial adenomatous polyposis (FAP) is an inherited syndrome characterized by the development of multiple adenomas in the colorectum, a high risk of CRC, and extracolonic manifestations [2]. Germline *APC* mutations cause FAP, an autosomal dominant disorder [2–4]. A second hit causes adenomatous polyps, recapitulating the initial event in sporadic CRC development [1].

Organoid culture systems that partially mimic the 3D architecture of specific organs have been used extensively for cancer research [5]. Genomic, transcriptomic, and proteomic data for cancer organoids have contributed to the identification of novel driver genes and molecular-level characterization of various cancers [5–8]. We have previously established patient-derived organoids (PDOs) for FAP (i.e., HCT24-8 and HCT24-10, showing malignant and benign properties, respectively) [1]. HCT24-8 is resistant to MEK inhibitors and harbors an activating mutation in *KRAS* [1].

In many tumors, drug resistance is mediated by drug-tolerant persister (DTP) cell subpopulations [9]. The prevalence of DTP cells has implications for disease progression and management [9]. DTP cells do not appear to arise by a single event but rather involve a combination of changes; thus, a multimodal approach is needed for the characterization of these cells [9,10]. DTP cells have been reported in many cancers [9]. However, the identification of drugs targeting DTP cells is a still major challenge.

Single-cell RNA sequencing (scRNA-Seq) is an emerging technology used to measure transcript levels within each individual cell of a sample. It allows a representation of expression profiles in specific subpopulations of cells [11] and provides a new opportunity to probe cancer biology in a high-resolution and dynamic manner [12]. For example, Kalkavan et al. evaluated PC9 cells (a human lung adenocarcinoma cell line) before and after drug treatments using scRNA-Seq and identified DTP cells within the cell line [13]. Furthermore, scRNA-Seq data are used as inputs to feed linear or nonlinear models, contributing to the increasing use of machine learning (ML) methods in cancer diagnosis, prognosis, and treatment guidance [11,14].

* Corresponding authors.
E-mail addresses: nojima.yosui.mmds@osaka-u.ac.jp (Y. Nojima), ryao@jfcr.or.jp (R. Yao), suzuki@sigmath.es.osaka-u.ac.jp (T. Suzuki).

https://doi.org/10.1016/j.bbadis.2025.167693
Received 25 November 2024; Received in revised form 24 December 2024; Accepted 22 January 2025
Available online 25 January 2025
0925-4439/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

In this study, we used scRNA-Seq to characterize PDOs in FAP and built ML models using public data from Kalkavan et al. as training data to identify DTP cells within PDOs. This study provides a novel strategy for the identification of DTP cells using ML models. Furthermore, we identified candidate drugs against DTP cells in the PDOs of FAP using public drug sensitivity data.

2. Materials and methods

2.1. Establishment and culture of organoids

FAP tumor samples were obtained from consenting patients, and all procedures were approved by the Research Ethics Board at the JFCR Cancer Institute (Tokyo, Japan). Following surgical resection, tumors were immediately cut into small pieces and washed with PBS. Tumors were enzymatically dissociated in digestion buffer composed of DMEM (Thermo Fisher Scientific, Waltham, MA, USA), 0.0625 % collagenase (Sigma-Aldrich, St. Louis, MO, USA), 0.125 % dispase (Thermo Fisher Scientific), and 2.5 % FBS at 37 °C for 60 min. Tumors were vigorously agitated by pipetting in PBS, and the tumor fragments were collected and centrifuged at 1000 ×*g* for 5 min. The pellets were suspended in Matrigel (Corning Incorporated, Corning, NY, USA), and 25 μl of this mixture was dispensed in each well. Plates were incubated at 37 °C for 10 min. The basal medium consisted of Advanced DMEM/F12 (Thermo Fisher Scientific) supplemented with 2 mmol/l GlutaMAX, (Thermo Fisher Scientific), 10 mmol/l HEPES (Sigma-Aldrich), penicillin/streptomycin (Thermo Fisher Scientific), primocin (InvivoGen, San Diego, CA, USA), B27 supplement (Thermo Fisher Scientific), 1.25 mmol/l *N*-acetyl-L-cysteine (Sigma-Aldrich), and 10 nmol/l gastrin (Sigma-Aldrich). PDOs were cultured in ENR medium composed of 10 ng/ml recombinant mouse EGF (Thermo Fisher Scientific), 10 % Noggin conditioned medium, and 1 μg/ml recombinant mouse R-spondin-1 (R&D Systems, Minneapolis, MN, USA) and were maintained at 37 °C at 5 % O_2. Media were changed every 3 days.

2.2. Processing of single-cell data

Single-cell sequencing was performed following a previously reported method [15]. The fastq files were processed using Cell Ranger (version 3.1.0; 10× Genomics, Inc., Pleasanton, CA, USA) count pipelines with the GRCh38 reference human genome (version 3.0.0; built on November 19, 2018). The output files from Cell Ranger were input into velocyto (version 0.17.17) [16] to count the spliced and unspliced reads. Loom files were generated and input into the Seurat package (version 4.4.0) [17]. The "spliced" assay was utilized as the default assay. Cell quality was assessed by the percentage of mitochondrial gene expression. In the PDO datasets, cells with unique feature counts between 500 and 3000 and with <15 % mitochondrial genes were retained for downstream processing. In the GSE189638 dataset, unique feature counts and the percentage of mitochondrial genes were set to the same values used previously [13]. The filtered data were scaled and transformed using the SCTransform function [18], and linear regression was performed to remove variation due to cell quality (% mitochondrial reads). Then, the three scRNA-Seq data sets were integrated using the SelectIntegrationFeatures function with nfeatures = 3000 and the PrepSCTIntegration, FindIntegrationAnchors, and IntegrateData functions. Principal components (PCs) were selected by inspecting elbow plots (PDOs; Fig. S1B, GSE189638; Fig. S5B) and the percentage of variance explained was assessed. Then, non-linear dimensionality reduction was performed to visualize the results in uniform manifold approximation and projection (UMAP) plots using the first 30 PCs for each dataset. The cell clusters were detected by running the FindNeighbors function using the first 30 PCs for each dataset as the input followed by the FindClusters function. Each cell type was annotated according to maker gene expression using the RenameIdents function of the SeuratObject package (version 4.1.4).

Differentially expressed genes (DEGs) between tumor cluster 1 (TC1) of the malignant tumor organoid and clusters of other organoids were obtained using the FindMarkers function. DEGs were identified based on 5 % adjusted *P*-values and increased expression in TC1 cells of the malignant tumor organoid. A pathway analysis was performed using the enrichKEGG function of clusterProfiler package (version 4.6.2) [19]. The Wnt signaling pathway was visualized using the pathview package (version 1.38.0) [20]. These functions were run in R software (version 4.2.3).

2.3. Differential abundance analysis

Milo (version 1.6.0) [21] was used to evaluate the differential abundance (DA) of cells within defined neighborhoods between tumor and normal organoids. We first used the buildGraph function to construct a *k*-nearest neighbor (*k*-NN) graph with k = 20, using 30 PCs. Next, we used the makeNhoods function to assign cells to neighborhoods based on their connectivity over the *k*-NN graph. For computational efficiency, we subsampled 1000 cells for each cluster in the GSE189638 dataset. To test for DA, the testNhoods function was used. To control for multiple testing, we used a 10 % spatial false discovery rate (FDR) implemented in Milo as a threshold for significance. The spatial FDR and $\log_2$(fold change) of the number of cells between the tumor and normal organoids in each neighborhood were used for visualization. The Milo package was implemented in R software (version 4.2.3).

2.4. RNA velocity, single-cell trajectory, and pseudotime analyses

Spliced and unspliced counts were calculated using the velocyto package (see scRNA-Seq analysis). To estimate RNA velocities of single cells, scVelo (version 0.2.4) [22] was used, as implemented in Python (version 3.9.1). The top 2000 highly variable genes were selected among those that passed a minimum threshold of 20 for spliced and unspliced mRNA counts using the scv.pp.filter_and_normalize function. A nearest-neighbor graph (with 30 neighbors) was calculated based on Euclidean distances in PCA space (with 30 PCs) on logarithmized spliced counts using the scv.pp.moments function. Prior to the velocity estimation, first- and second-order moments (means and uncentered variances) were computed for each cell across its 30 nearest neighbors using the scv.pp.moments function. The velocities were estimated using the scv.tl.velocity function, and a velocity graph was generated using the scv.tl.velocity_graph function. A stream plot of velocities on the embedding was generated using the scv.pl.velocity_embedding_stream function.

Single-cell trajectory and pseudotime analyses were performed using the monocle3 package (version 1.3.1) [23] in R software (version 4.2.3). TC1 cell cluster, TC2 cell cluster, Stem cells of cluster 2, Stem cells of cluster 22, Enterocytes of cluster 5, and Enterocytes of cluster 9 were subsampled prior to running monocle3. The subsampled data were normalized using the preprocess_cds function. The align_cds function was then used to eliminate batch effects using the batchelor package (version 1.14.1) [24]. UMAP dimensionality reduction and cell clustering were applied to the data using the reduce_dimension and cluster_cells functions, respectively. A principal graph was learned from the reduced dimensions using the learn_graph function, and visualization through the UMAP was performed using the plot_cells function. The graph was used to order cells in pseudotime using the order_cells function.

2.5. Construction of ML models

Based on the DA results, the C2 cluster was defined as DTP cells in GSE189638. Since 2998 cells clustered in C2, we randomly selected 214 cells from each cluster except for the C2 cluster, resulting in 2994 cells classified as non-DTP cells. GSE189638 and PDO data were integrated, and genes that were expressed in >50 % of all cells were selected, resulting in 916 genes. Of these 916 genes, 897 overlapped with Cancer

Dependency Map (DepMap) gene expression data (version 23Q4) containing $\log_2(TPM + 1)$ values for protein-coding genes. Classification models were constructed using the train function of the caret package (version 6.0-94) [25] in R software (version 4.2.3). The training samples were split into training and validation sets at a ratio of 8:2 and trained using 5-fold cross-validation. Four algorithms, support vector machines (SVM) with radial basis function kernel (svmRadial), SVM with linear kernel (svmLinear), eXtreme gradient boosting (xgbTree), and random forest (RF), were applied. Optimal hyperparameters were searched during the learning process using a grid search. The hyperparameter settings for each algorithm in the grid search are shown in Table S1. Metrics of accuracy, area under the curve (AUC), and kappa values were calculated using the confusionMatrix function and the roc function of the pROC package (version 1.18.0). Using the classification models with the best-performing hyperparameter settings, each cell in the PDO scRNA-Seq data was classified as either a DTP or non-DTP cell using the predict function.

2.6. Identification of candidate drugs against DTP cells

A list of 195 cell lines with activating mutations in *KRAS* was obtained using the Cancer Cell Line Encyclopedia (CCLE) and the Catalogue of Somatic Mutations in Cancer (COSMIC) data downloaded from Harmonizome 3.0 (https://maayanlab.cloud/Harmonizome/). Of 195 cell lines, 172 with processed bulk RNA-Seq data from DepMap were classified as either DTP or non-DTP cells. Drug sensitivity data for 22 of 23 cell lines classified as DTP were obtained from Profiling Relative Inhibition Simultaneously in Mixtures (PRISM), Cancer Target Discovery and Development (CTD^2), and Genomics of Drug Sensitivity in Cancer (GDSC) data available from DepMap. For overlapping drugs among the PRISM, CTD^2, and GDSC datasets, the average AUC value was obtained. The top three drugs with the lowest AUC values for each of the 22 cell lines were selected, resulting in 27 drugs. A network composed of the 22 cell lines and 27 drugs was constructed using Cytoscape (version 3.10.0).

To rank the 27 drugs, scores were calculated using Eq. (1).

$$score = \sum_{i=1}^{n} (P_i - log_2 A_i + |log_2 C|) \quad (1)$$

Here, P denotes the probability that a cell line is classified as DTP cells. A denotes the AUC when a cell line was treated with the drug. C denotes the maximum AUC among all combinations of 27 drugs and 22 cell lines, and n denotes the number of edges bound to each drug, which also corresponds to the number of cell line nodes.

2.7. PDO viability assay

Trametinib and YM-155 were purchased from Selleck Biotechnology (Kanagawa, Japan). THZ2 and sangivamycin were purchased from MedChemExpress (Monmouth Junction, NJ, USA). Organoids were digested with TrypLE Express (Thermo Fisher Scientific) supplemented with 10 μmol/l Y27632 (Sigma-Aldrich) at 37 °C for 15 min with pipetting every 5 min. The cells were suspended in a basal medium, and clumps were removed by passing the suspension through a 40 μm cell strainer. The suspension was centrifuged at 1000 ×*g* for 5 min, and the cells were resuspended in the basal medium. Cells were counted and adjusted to 2.0×10^5 cells/ml in the Matrigel. Then, 10 ml of the cell suspension was dispensed into each well of U-bottom 96-well microplates (Thermo Fisher Scientific). PDOs were cultured for 3 days to allow the formation of organoids and then treated with drugs at the indicated doses for 7 days. The medium was changed every 3–4 days. We generated 8-step, 10-fold drug matrices in technical quadruplets. Cell viability was measured using a CellTiter-Glo 3D Cell Viability Assay Kit (Promega, Madison, WI, USA). Readings were obtained using a Mithras LB 940 luminometer (Berthold Technologies GmbH & Co., Bad Wildbad, Germany). In the 96-well plate, the average cell viability for each drug concentration was calculated using DMSO-treated cells as controls. Data were analyzed using GraphPad Prism 9 (GraphPad Software Inc., Boston, MA, USA).

Titrations to determine drug synergy were performed by plating the PDOs in 96-well plates as described above. After 3 days, PDOs were treated with the indicated combination. The drugs were titrated in a five-dose manner, ranging from four times the IC_{50} concentration to one-fourth of the IC_{50} concentration. Cell viability was determined using the CellTiter-Glo 3D Cell Viability Assay, as described above. Synergistic effects were determined using the median effect principle proposed by Chou-Talalay [26], defined by Eq. (2) and implemented in the R package medianeffect (version 0.9.0).

$$CI = \frac{(D)_1}{(D_x)_1} + \frac{(D)_2}{(D_x)_2} \quad (2)$$

Here, CI indicates the combination index. D denotes the respective combination doses of each drug resulting in 50 % growth inhibition. D_x denotes the dose of each drug resulting in 50 % growth inhibition when used alone.

3. Results

3.1. Identification of cell types and tumor clusters

Three organoids were evaluated using scRNA-Seq: HCT24-8 (malignant tumor organoid), HCT-10 (benign organoid), and HCT71-2 (derived from normal colorectal tissue of a patient with sporadic CRC). After applying a unified scRNA-Seq analysis pipeline (see Methods), we obtained 2435, 1975, and 2270 cells from the benign tumor, malignant tumor, and normal organoids, respectively. The cells from the three scRNA-Seq data sets were divided into 35 clusters, and each cluster was annotated according to their marker genes (Figs. 1A and S1A). Cells from clusters 18 and 19 did not show significant expression of any marker genes and were not abundant in the normal organoid; therefore, they were defined as TC1 and TC2, respectively. To evaluate differences in composition between tumor and normal organoids, a differential abundance (DA) analysis was performed using Milo [21], revealing 429 neighborhoods, including 31 with evidence for DA (Spatial FDR < 0.1, Fig. 1B). Of the 31 neighborhoods, 23 were from TC1 and TC2 (Fig. 1C), resulting in higher proportions of TC1 and TC2 in tumor organoids than in the normal organoid. Furthermore, there were more TC1 cells in the malignant tumor organoid than in the benign tumor organoid, whereas there were more TC2 cells in the benign tumor organoid than in the malignant tumor organoid (Fig. 1D).

3.2. Differentiation trajectories of TC1 and TC2 cells

To estimate the cell types from which TC1 and TC2 cells were derived, RNA velocity and single-cell trajectory analyses were performed. The RNA velocity analysis showed a strong directional flow from the stem cell cluster to TC1 and from the enterocyte cluster to TC2 (Fig. 2A). Using stem cells around TC1 (clusters 2 and 22 in Fig. S1A) and enterocytes around TC2 (clusters 5 and 9 in Fig. S1A), differentiation trajectories were evaluated (Fig. 2B). When both ends of the estimated differentiation trajectory were set to the early cells, TC2 cells exhibited the longest pseudotime, and these were contained mainly in the benign organoid and not in the malignant and normal organoids (Fig. 2C). The expression levels of the differentiation marker genes *CDKN1A* and *CDKN2A* were higher in TC2 cells of the benign tumor organoid than in other clusters and organoids (Fig. S2A–D), supporting the validity of the trajectory analysis. *PERP*, *TESC*, *AREG*, and *SOX4*, known marker genes of cancer stem cells (CSCs) in CRC [27–29], were more highly expressed in the TC1 cell cluster of the malignant tumor organoid than in other clusters and organoids (Table S2). A pathway

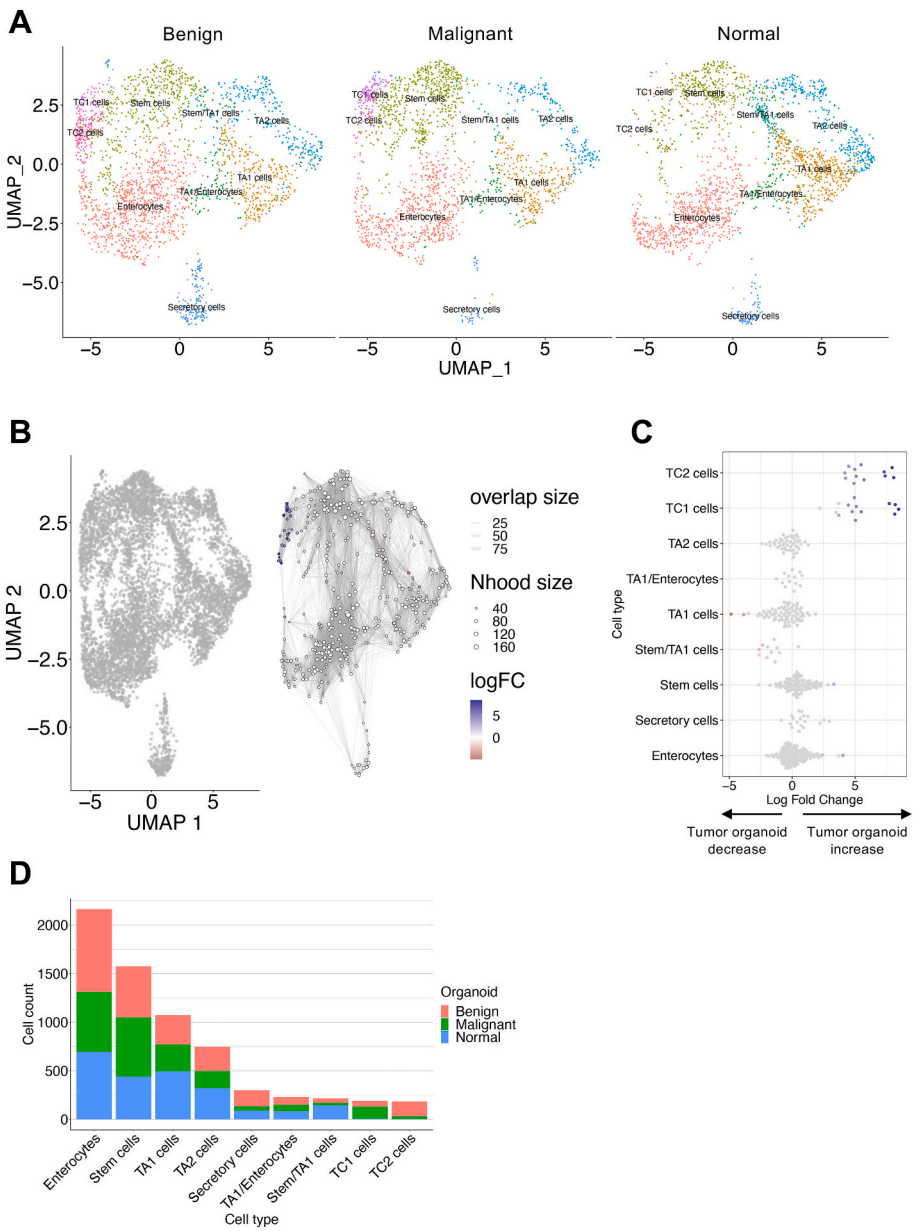

Fig. 1. Single-cell RNA sequencing (scRNA-Seq) data for patient-derived organoids (PDOs). (**A**) Uniform manifold approximation and projection (UMAP) embedding of scRNA-Seq data. Nine cell types were defined based on marker gene expression (Figure S1C). (**B**) Graph representation of neighborhoods (Nhoods) identified using a differential abundance analysis [21]. Nodes are Nhoods, colored by log2(FC) values between the tumor organoids and normal organoid. Nondifferential abundance Nhoods (Spatial FDR ≥ 0.1) are shown in white, and sizes correspond to the number of cells in a Nhood. Graph edges depict the number of cells shared between adjacent Nhoods. (**C**) Beeswarm plot showing the distribution of adjusted log2(FC) values for abundance between the tumor and normal organoids in Nhoods according to nine cell types. Colors are defined as in (B). (**D**) Bar plot showing the cell count for each cell type in each organoid.

analysis based on DEGs between TC1 cells of the malignant tumor organoid and other clusters and organoids showed enrichment for the Wnt signaling pathway and Epstein-Barr virus infection (Fig. S3A). The Wnt signaling pathway contributes to the maintenance of stemness in CSCs of CRC [30]. The Wnt signaling pathway-related genes *WNT6*, *BAMBI*, *NKD1*, *NOTUM*, and *APCDD1* were detected among the DEGs (Fig. S3B). They were highly expressed in TC1 cells (Fig. S3C), especially in the malignant tumor organoid (Fig. S3D), and were slightly expressed in TC2 cells (Fig. S3C). A pathway analysis based on DEGs between TC2 cells of the benign tumor organoid and other clusters and organoids showed enrichment for focal adhesion, *Salmonella* infection, and the p53 signaling pathway (Fig. S4A). Integrins (ITGs) are heterodimeric transmembrane receptors consisting of α and β subunits, and they play a key role in focal adhesion [31]. Notably, heterodimeric integrin αvβ6 (ITGAV/B6) promotes tumor formation in colorectal cancer [31], heterodimeric integrin α3β4 (ITGA3/A4) is highly expressed in ovarian cancer [32], and ITGA2 is a candidate tumor marker for colorectal cancer [33]. These five integrin genes, *ITGA2*, *ITGA3*, *ITGAV*, *ITGB4*, and *ITGB6*, were more highly expressed in the TC2 cell cluster of the benign tumor organoid than in other clusters and organoids (Fig. S4B–D).

3.3. Construction of ML models for the classification of DTP and non-DTP cells

Drug resistance is frequently facilitated by a subpopulation of DTP cells in the tumor [9]. The malignant tumor organoid of FAP is resistant to trametinib treatment [1], suggesting that it contains DTP cells.

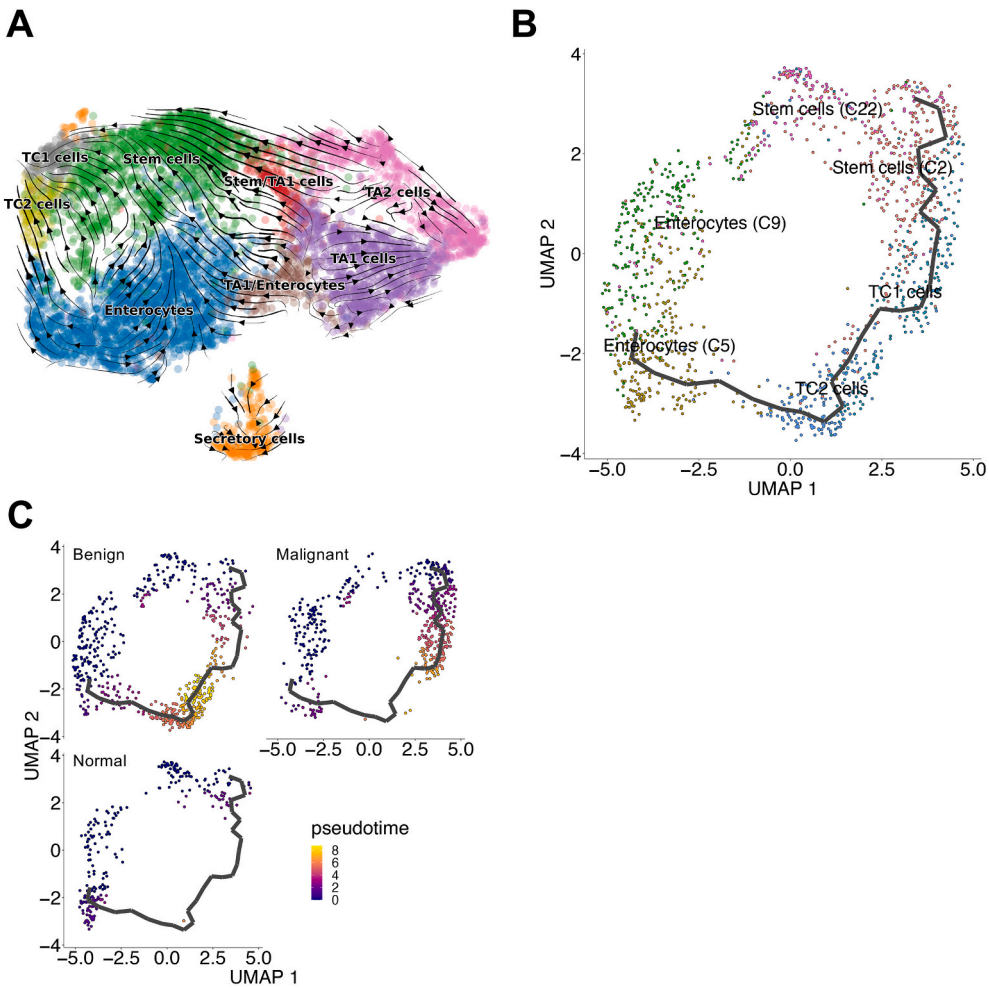

Fig. 2. RNA velocity, single-cell trajectory, and pseudotime inference from PDO scRNA-Seq data. (**A**) RNA velocities derived from the dynamical model for nine clusters from PDOs are projected onto a UMAP embedding. (**B**) Single-cell differentiation trajectory inferred using monocle3 overlaid on a UMAP embedding. The origin positions were set to both branch ends. TC1 cell cluster, TC2 cell cluster, Stem cells of cluster 2, Stem cells of cluster 22, Enterocytes of cluster 5, and Enterocytes of cluster 9 were subsampled prior to running monocle3. (**C**) Pseudotime inferred using monocle3 overlaid on a UMAP embedding.

Kalkavan et al. performed scRNA-Seq (GSE189638) using PC9, a human cell line derived from a patient with lung adenocarcinoma, parental (PT) cells, and corresponding persister cells on days 1, 3, and 7 (PS1D, PS3D, and PS7D, respectively) after BH3-mimetic treatment [13]. They utilized two BH3-mimetics, ABT737 (a BCL-2, BCL-X_L, and BCL-W inhibitor) [34] and S63845 (a selective MCL-1 inhibitor) [35]. BCL-2, BCL-X_L, BCL-W, and MCL-1 are associated with apoptosis and pathways related to MEK [36–40]. Thus, we hypothesized that the gene expression profiles in DTP cells enriched by treatment with trametinib, a MEK inhibitor, and those in DTP cells enriched by BH3-mimetic treatment are highly similar. To distinguish the DTP cells based on scRNA-Seq data for the malignant tumor organoid, we built ML models using the GSE189638 dataset. After applying the unified scRNA-Seq analysis pipeline (see Methods), we obtained 3620, 13,584, 13,411, and 6280 cells from PT, PS1D, PS3D, and PS7D, respectively. The cells evaluated using scRNA-Seq were divided into 15 clusters (Fig. S5A), defined as C0 to C14 (Fig. 3A). PS1D is enriched for DTP cells, suggesting a gradual return to the PT phenotype after BH3-mimetic treatment [13]. A DA analysis between PS1D and other samples identified 963 neighborhoods, including seven with evidence for DA (Spatial FDR < 0.1, Fig. 3B). Six of the seven neighborhoods were from the C2 cluster (Fig. 3C); the proportion of C2 was higher in the PS1D sample than in other samples. Furthermore, more than half of the cells in the C2 cluster were from PS1D (Fig. 3D). Thus, we defined the cells clustered in C2 as DTP cells and utilized them as training data to build ML models for classification.

Four algorithms, svmRadial, svmLinear, xgbTree, and RF, were utilized for the construction of ML models. Optimal hyperparameters were searched during the learning process using a grid search (Fig. S6A–E) and were evaluated via 5-fold cross-validation. The AUC, accuracy, and kappa values for each fold in four algorithms are shown in Fig. S7A–D. The kappa values for the classification models constructed using the svmRadial, xgbTree, and RF algorithms were generally over 0.6, whereas that for svmLinear was approximately 0.5. These values indicate good model performance [41].

3.4. Identification of drug-tolerant persister cells in the malignant tumor organoid of FAP

To identify DTP cells based on scRNA-Seq data for FAP, all cells derived from the three organoids were classified as either DTP or non-DTP using the four ML models. Using the svmRadial and svmLinear models, the number of cells classified as DTP was higher in the malignant tumor organoid than in the benign tumor organoid and normal organoid (Figs. 4A and S8A). In contrast, using the xgbTree model, the number of cells classified as DTP was higher in the benign tumor organoid and normal organoid than in the malignant tumor organoid (Fig. S8B). Using the RF model, more cells were classified as DTP than as non-DTP in all organoids (Fig. S8C). Among all TC1 cells, the proportions classified as DTP were 35.6 %, 29.3 %, 13.6 %, and 63.9 % using svmRadial, svmLinear, xgbTree, and RF models, respectively

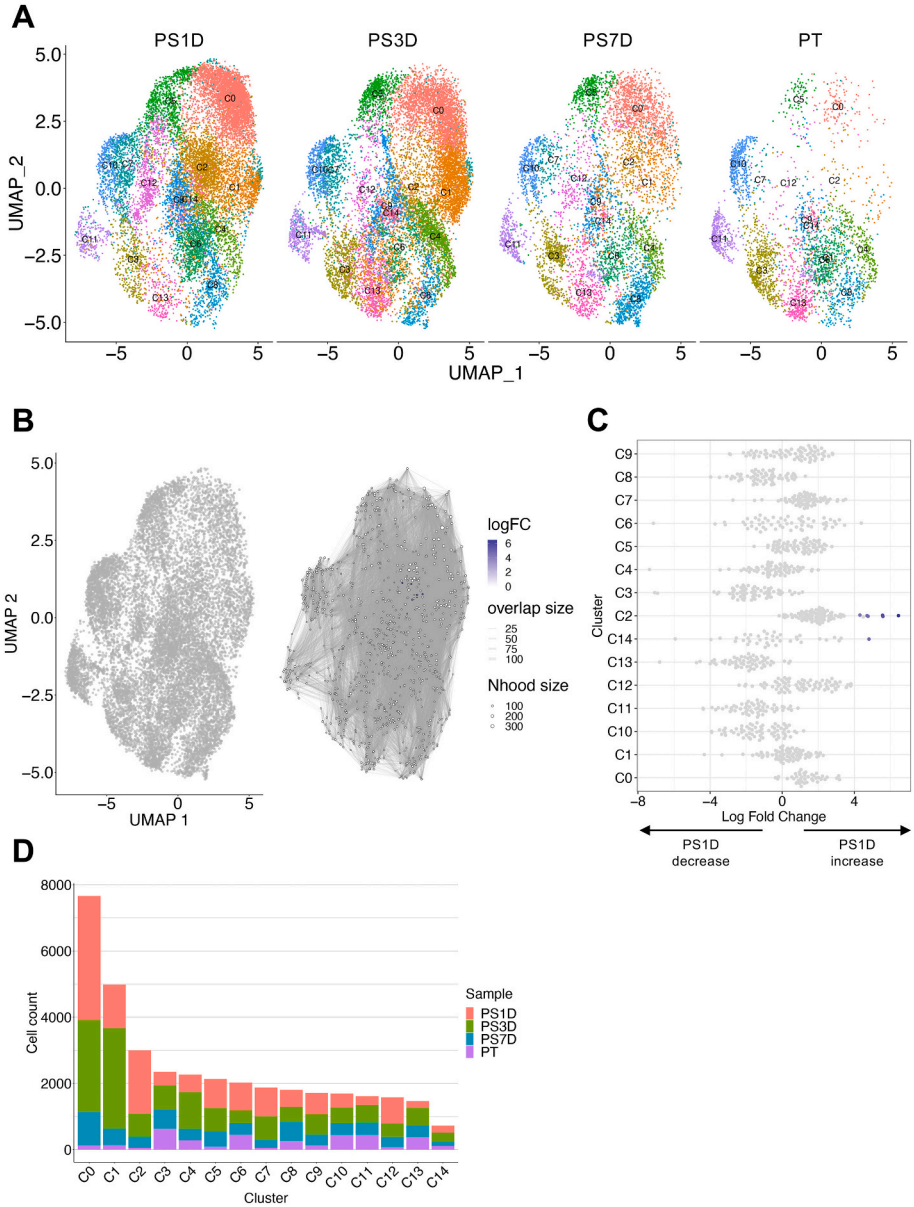

Fig. 3. Identification of drug-tolerant persister cells using public scRNA-Seq data. (**A**) UMAP embedding of public scRNA-Seq data. (**B**) Graph representation of Nhoods identified using a differential abundance analysis [21]. Nodes are Nhoods, colored based on $\log_2$(FC) values between persister cells in the 1st day (PS1D) and other samples. Nondifferential abundance Nhoods (Spatial FDR ≥ 0.1) are shown in white, and sizes correspond to the number of cells in a Nhood. Graph edges depict the number of cells shared between adjacent Nhoods. (**C**) Beeswarm plot showing the distribution of adjusted $\log_2$(FC) in abundance between PS1D and other samples in Nhoods according to 15 clusters. Colors are defined in the same way as in (B). (**D**) Bar plot showing the cell count for each cluster.

(Figs. 4B, C and S8D–I). Using svmRadial and svmLinear, the proportion of cells classified as DTP was highest for TC1 cells, followed by stem cells, whereas it was <10 % in other cell clusters (Figs. 4B, C and S8D, E). Using xgbTree, there was a slight difference between TC1 cells and other cell types (Fig. S8F and G). Using RF, TA1 cells had the highest proportion of cells classified as DTP, followed by Stem/TA1 cells, whereas the proportion of cells classified as DTP in TC1 cells was low (Fig. S8H and I).

3.5. Identification of drug combinations involving trametinib and validation of synergistic effects

The ML models mainly classified TC1 cells as DTP cells, suggesting that the drug resistance of the malignant tumor organoid is due to TC1 cells. Thus, we identified candidate drugs against TC1 cells using public cell line data, including mutation, gene expression, and drug sensitivity data. A flowchart for drug identification is shown in Fig. 5A. As the malignant tumor organoid harbors activating mutations in *KRAS* [1], we identified 195 cell lines with *KRAS* mutations from the Cancer Cell Line Encyclopedia (CCLE) and Catalogue of Somatic Mutations in Cancer (COSMIC). Bulk RNA-Seq data for 172 of 195 cell lines were available from DepMap. The ML models built using svmRadial classified 23 out of 172 cell lines as DTP cells. Of these 23 cell lines, drug sensitivity data for 22 cell lines were available from DepMap. We selected the top three drugs with strong effects on the 22 cell lines, resulting in a total of 27 drugs (Fig. 5A). The probabilities of DTP cells were highest (exceeding 0.8) for DAUDI (derived from a patient with Burkitt lymphoma), followed by PL21 (derived from a patient with acute myeloid leukemia) and HCC44 (derived from a patient with lung adenocarcinoma) (Fig. 5B and Table S3). The candidate drugs with the top three scores were ranked as follows: YM-155, THZ2, and sangivamycin (Fig. 5C).

We examined the synergistic effects of YM-155, THZ2, or

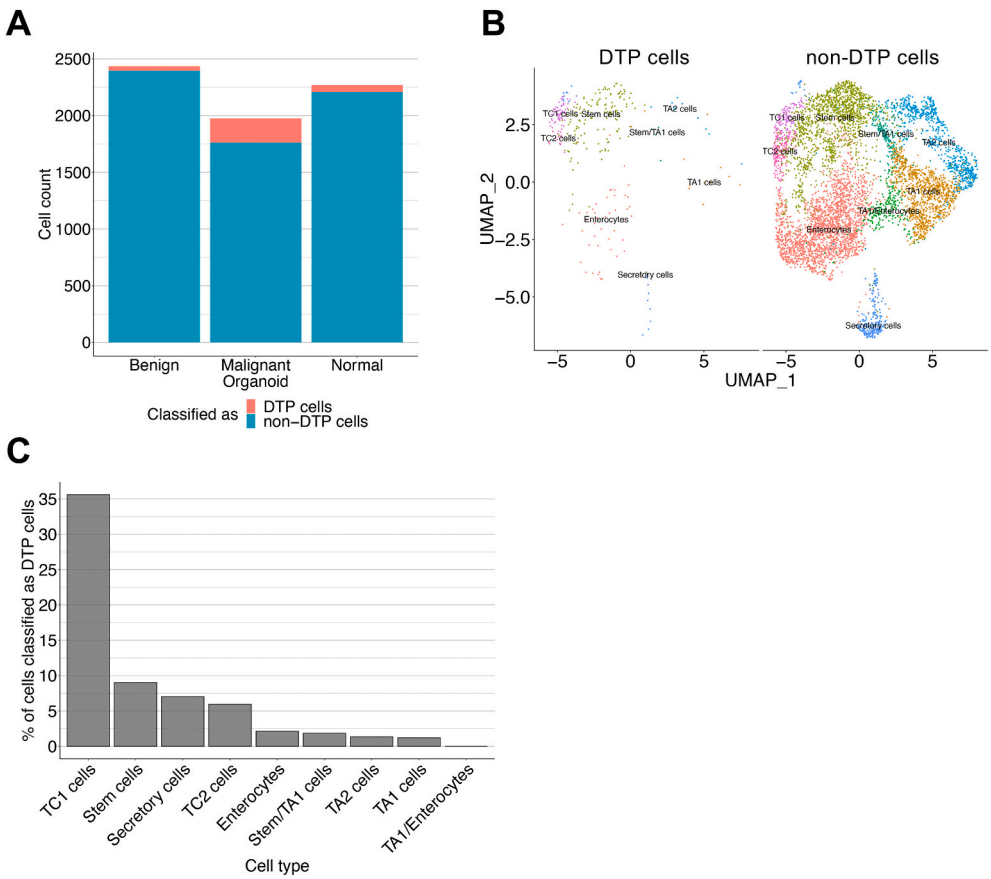

Fig. 4. Classification of single cells in PDOs using the svmRadial model. (**A**) Bar plot showing the counts of cells classified as DTP or non-DTP cells in each organoid. (**B**) Cells classified as either drug-tolerant persister (DTP) or non-DTP cells are projected into a UMAP embedding. (**C**) Percentage of cells classified as DTP cells in each cell type of PDOs.

sangivamycin with trametinib against the malignant tumor organoid using the Chou-Talalay method for drug combinations [26]. The combination indexes for the combination of YM-155 or THZ2 with trametinib were <1.0 in each fraction affected (Fa) (Fig. 6A and B), whereas those for sangivamycin with trametinib were >1.0 at a lower Fa and <1.0 at a higher Fa (Fig. 6C).

4. Discussion

Tumor development in patients with FAP recapitulates that of sporadic CRC [1]. PDOs have many clinical features of the original patient tumors, including similar genetic mutations and responses to chemotherapeutic agents [1,5,7,42]. Therefore, FAP organoids in this study were used as models for the progression of sporadic CRC.

The scRNA-Seq data for FAP highlighted the high frequencies of TC1 and TC2 cell clusters in the tumor organoids. Cancer cells are derived from terminally differentiated cells by genetic mutations [43]. TC2 cells showed high pseudotime values and high differentiation marker gene expression; in addition, their differentiation trajectory predicted that they originate from enterocytes. Thus, TC2 cells might arise from differentiated cells by genetic mutations or other factors. TC2 in the benign tumor organoids expressed high levels of *ITGA2*, *ITGA3*, *ITGAV*, *ITGB4*, and *ITGB6*. Although we have previously confirmed the lack of tumor formation in the benign tumor organoid using mouse xenograft experiments [1], these five integrin genes are involved in carcinogenesis [31–33]. Thus, the TC2 cells in the benign tumor organoid may be in a precancerous condition, consistent with the characteristics of FAP [44]. CSCs arise from normal stem cells via genetic mutations and other mechanisms [45]. In this study, the differentiation trajectory predicted that TC1 cells originate from stem cells; in addition, levels of Wnt signaling pathway-related genes were increased in TC1 cells of the malignant tumor organoid. The Wnt signaling pathway contributes to the maintenance of stemness in CSCs of CRC [30]. In addition, the *PERP*, *TESC*, and *AREG* genes were upregulated in TC1 cells. These genes were among the 50 marker genes identified by Lin et al. for CSCs in CRC [27]. Furthermore, *SOX4* is expressed in intestinal stem cells, maintains the stemness of cancer cells in CRC, and is regulated by the Wnt signaling pathway in intestinal stem cells [28,29]. We also detected *SOX4* upregulation in TC1 cells. In addition, mutant *KRAS* promotes CSC properties via the Wnt signaling pathway in colorectal cancer [46]. The malignant tumor organoid is *KRAS*-positive [1], and TC1 was enriched in the malignant tumor organoid. These results suggest that TC1 cells possess cancer stem-like properties. CSCs typically exist as single cells or small clusters of up to 20 cells [47]. Therefore, TC1 cells might be enriched during the establishment and/or culture of the organoid.

CSCs are characterized by their ability to maintain an asymmetric self-renewal cellular state, expressing exclusive "stemness" markers and thus occupying the apex of the differentiation hierarchy [47]. In addition, CSCs have the potential for immune evasion and treatment resistance [47–49]. The selective pressures imposed by systemic therapies might eliminate non-CSCs while sparing a population of intrinsically resistant CSCs, leading to CSC enrichment [47,50]. These surviving CSCs can be effectively re-populated, resulting in chemotherapy failure and cancer recurrence [51]. This hypothesis is called the CSC theory [52]. During chemotherapy, the cell cycle of CSCs slows down and the cells fall into a "quiescent" state [53], resulting a slow rate of proliferation [54]. However, conflicting results have been obtained regarding the proliferation rates of DTPs, with some reports suggesting slow

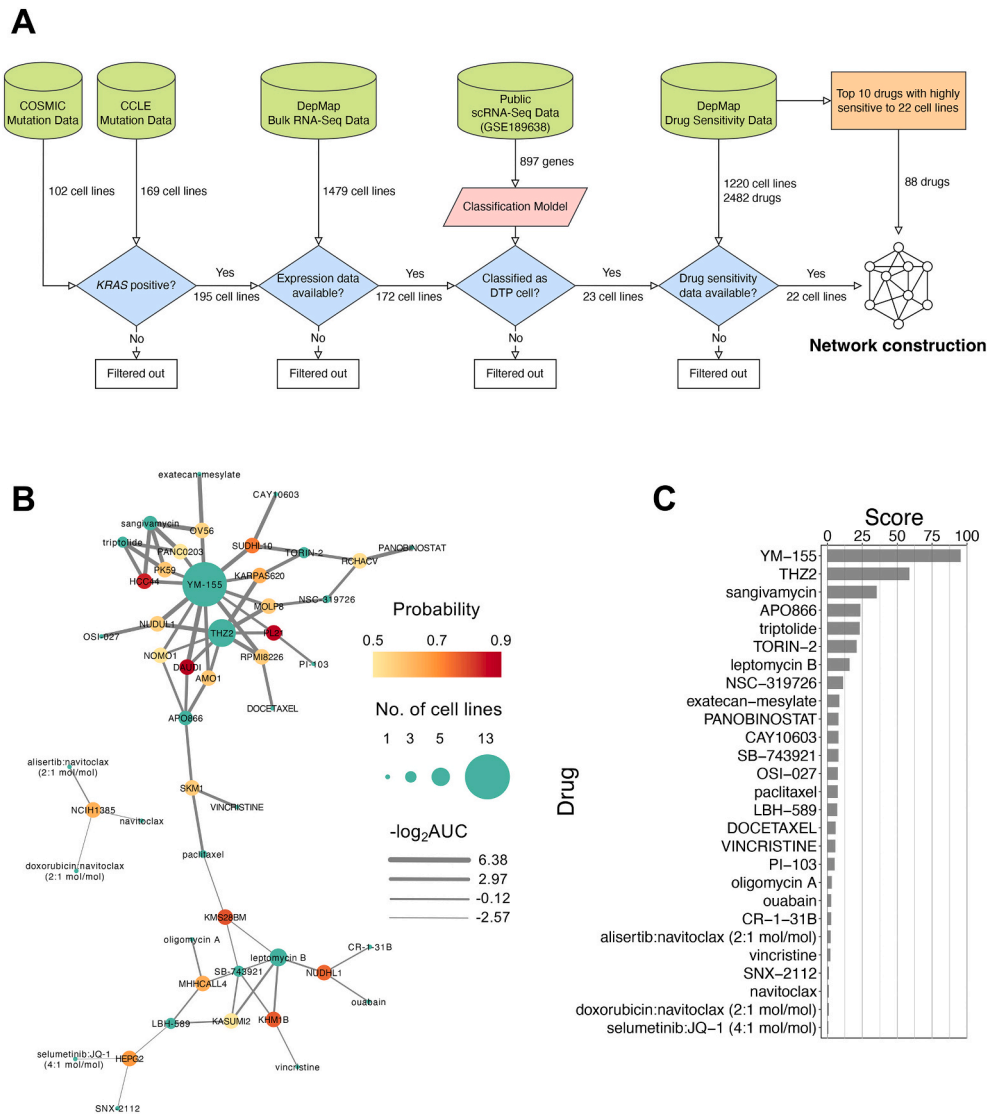

Fig. 5. Identification of candidate drugs against DTP cells in the malignant tumor organoid. (**A**) Flowchart for the identification of candidate drugs against DTP cells. COSMIC, Catalogue of Somatic Mutations in Cancer; CCLE, Cancer Cell Line Encyclopedia; DepMap, Cancer Dependency Map. (**B**) Network composed of 22 cell lines classified as DTP cells and 27 drugs predicted to be effective against these cells. Green nodes indicate drugs. Nodes with probability estimates indicate cell lines. The edges indicate -$\log_2$AUC. AUC; area under the curve. (**C**) Scores for each of the 27 drugs (see Methods).

proliferation [9,55] and others suggesting that they proliferate rapidly or grow at rates similar to those of non-DTP cells [9,55,56]. Although persistence and cancer stemness are two sides of the same coin, little is known about their relationship, in part owing to a lack of experimental systems that enable the investigation of both persistent cancer cells and CSCs [50,52]. Therefore, CSCs cannot necessarily be classified as DTP cells. Additionally, there is currently no consensus on DTP markers for CRC.

To investigate whether TC1 cells are DTP cells, we built ML models for classification using a public dataset, GSE189638. The frequency of C2 was elevated in PS1D, enriched for DTP cells [13], indicating that C2 is composed of DTP cells. Typically, DTP cells are a rare population found in malignant tumor tissues, existing at a very low frequency in clinical samples [9]. Therefore, the svmRadial and svmLinear results were closer to the actual pathology than the xgbTree and RF results. Additionally, svmRadial demonstrated higher accuracy than that of svmLinear in cross-validation. Thus, we concluded that the svmRadial results are the most reasonable in terms of model accuracy and pathology. The svmRadial model classified approximately 36 % of the TC1 cluster cells as DTP cells. This proportion was significantly higher than those for other cell types, suggesting that TC1 cells include DTP cells. ABT737 and S63845 are BH3-mimetic, and trametinib is a MEK inhibitor [34,35,57]. The target proteins of these inhibitors each function in the same apoptosis-inducing pathway [36–40]. Therefore, it is assumed that DTP cells in both the malignant tumor organoid and in PS1D result from treatment with inhibitors that have similar mechanisms of action. As the result, the classification model could detect the DTP cells in the malignant tumor organoid, identified as TC1 cells.

Based on these results, drug resistance in the malignant tumor organoid results from TC1 cells; thus, we searched for drugs expected to be effective against TC1 cells. We used DepMap, CCLE, and COSMIC, which include data for various cancer cell lines, including genetic mutation, gene expression, and drug sensitivity data [58,59], with detailed genetic and pharmacological annotations [60]. *KRAS* mutations drive primary resistance to anti-epidermal growth factor receptor (EGFR) treatment [61]. Activating mutations in *KRAS* confer resistance to EGFR inhibitors in FAP organoids [1]. Thus, we focused on cell lines with activating mutations in *KRAS*, and classified the cell lines into DTP or non-DTP cells using the ML model. Among the cell lines classified as DTP cells, we investigated 22 for which gene expression and drug sensitivity

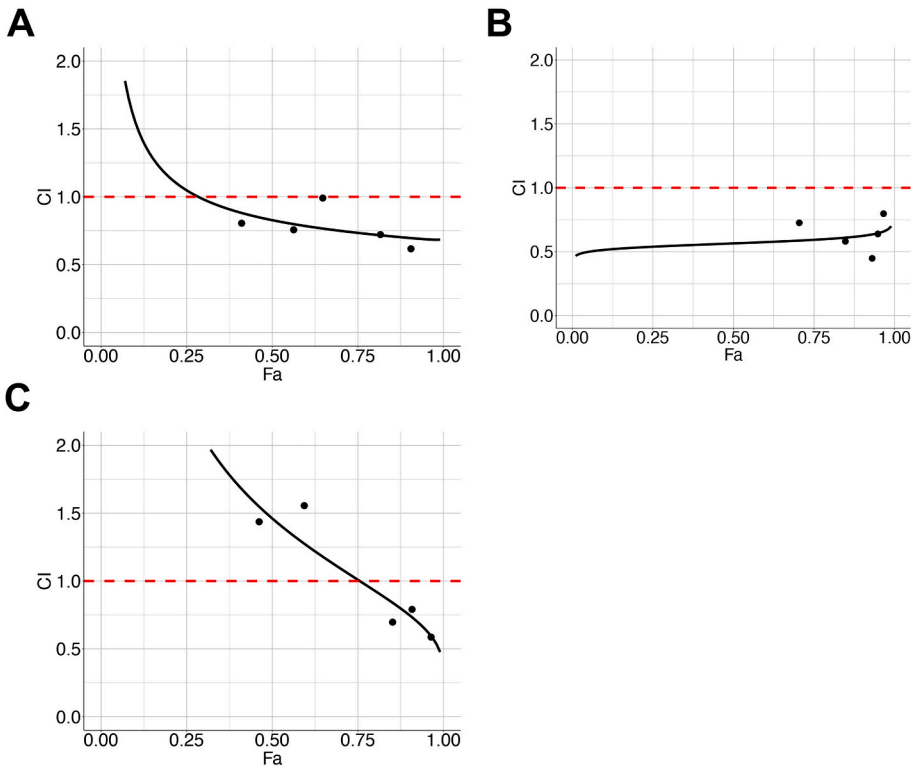

Fig. 6. Validation of the synergistic effects of drug combinations against the malignant tumor organoid. Fraction affected (Fa)-Combination index (CI) plot reveals whether drug combinations exert synergistic, additive, or antagonistic effects. CIs of less than, equal to, and greater than 1 indicate synergy, additivity, and antagonism, respectively [26]. Each CI-Fa plot shows assay results when YM-155 (**A**), THZ2 (**B**), or sangivamycin (**C**) is combined with trametinib. Each dot and each curve indicate the average of the CI and the curve fitted from the CI, respectively. The red dashed line indicates the threshold of CI < 1.

data were available. We validated the effects of YM-155, THZ2, and sangivamycin (with the highest scores) on the malignant tumor organoid. The target proteins of YM-155, THZ2, and sangivamycin are survivin, cyclin-dependent kinase 7 (CDK7), and protein kinase C, respectively [62–64].

YM-155 and THZ2 (but not sangivamycin) showed synergistic effects with trametinib on the malignant tumor organoid. Although the synergistic effect of YM-155 or THZ2 with trametinib was not strong, this finding has important implications considering the drug resistance exhibited by DTP cells. This study demonstrates that DTP cells are not resistant to all drugs. Survivin, the target protein of YM-155, is the smallest member of the inhibitor of apoptosis family and is overexpressed in various malignant tumors and CSCs [65,66]. The elevated survivin levels indicate poor responses to chemotherapy and drug resistance [66]. Furthermore, other groups have demonstrated that YM-155 inhibits the stemness of CSCs in breast, lung, and gastric cancer cells [67–69]. CDK7, a target protein of THZ2, contributes to the maintenance of CSC activity through NOTCH1-cMYC signaling and its inhibition attenuates CSC activity in anaplastic thyroid cancer [70]. These previous findings support the validity of the strategy in this study.

As the carcinogenic mechanisms in FAP and sporadic CRC are identical [1], the strategy utilized in this study is expected to be useful for proposing candidate drugs against DTP cells in sporadic CRC. Although PC9 cells used in GSE189638 are derived from non-FAP tissues, the ML models classified up to 36 % of TC1 cells as DTP cells. This suggests that the classification model may also be able to identify DTP cells resulting from treatment with MEK inhibitors or BH3 mimetics in other types of cancers. The drug types could affect model performance. However, the ability of the model to identify DTP cells treated with drugs targeting pathways other than those related to MEK inhibitors and BH3 mimetics was not evaluated in this study. This study demonstrated that DTP cells within organoids before drug treatment can be identified using ML models constructed using cell line data. In addition to the combination of BH3-mimetic and PC9 cells, numerous combinations of drugs and drug-resistant cell lines have been reported [71]. By adding these scRNA-Seq data, it may be possible to construct a DTP cell identification model able to handle various drugs and cancer types in the future. Challenges remain in translating single-cell findings into clinical practice and developing personalized treatment strategies. Thus, it is difficult to immediately apply the approach used in this study to clinical practice. However, the clinical application of scRNA-Seq is being discussed in rheumatic and allergic diseases [72], and with these advancements, the feasibility of this approach in clinical practice is expected to increase.

Ethical approval

FAP tumor samples were obtained from consenting patients, and all procedures were approved by the Research Ethics Board at the JFCR Cancer Institute.

Funding

This work was supported by the Japan Agency for Medical Research and Development (AMED) under grant number JP23ck0106885 to Y.N. and the Japan Science and Technology Agency CREST under grant number JPMJCR2022 to T.S.

CRediT authorship contribution statement

Yosui Nojima: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Ryoji Yao:** Data curation, Investigation,

Resources, Validation, Writing – review & editing. **Takashi Suzuki:** Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbadis.2025.167693.

Data availability

All data supporting the findings of this study are available within the paper and its supplemental information. The scRNA-Seq data were deposited into the NBDC Human Database with accession code JGAS000139 (https://humandbs.dbcls.jp/en/hum0106-v1). The raw data for the GSE189638 dataset are available from the National Center for Biotechnology Information (NCBI) SRA (https://www.ncbi.nlm.nih.gov/sra). Other relevant data are available from the corresponding authors upon reasonable request. All analyses were performed using the cited software, packages, and pipelines, whose codes are publicly available. Further information on the code used in this study can be obtained from the corresponding author on reasonable request.

References

[1] M. Sakahara, T. Okamoto, J. Oyanagi, H. Takano, Y. Natsume, H. Yamanaka, D. Kusama, M. Fusejima, N. Tanaka, S. Mori, H. Kawachi, M. Ueno, Y. Sakai, T. Noda, S. Nagayama, R. Yao, IFN/STAT signaling controls tumorigenesis and the drug response in colorectal cancer, Cancer Sci. 110 (2019) 1293–1305, https://doi.org/10.1111/cas.13964.

[2] M.L. Leoz, S. Carballal, L. Moreira, T. Ocaña, F. Balaguer, The genetic basis of familial adenomatous polyposis and its implications for clinical practice and risk management, TACG 8 (2015) 95–107, https://doi.org/10.2147/TACG.S51484.

[3] I. Nishisho, Y. Nakamura, Y. Miyoshi, Y. Miki, H. Ando, A. Horii, K. Koyama, J. Utsunomiya, S. Baba, P. Hedge, A. Markham, A.J. Krush, G. Petersen, S. R. Hamilton, M.C. Nilbert, D.B. Levy, T.M. Bryan, A.C. Preisinger, K.J. Smith, L.-K. Su, K.W. Kinzler, B. Vogelstein, Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients, Science 253 (1991) 665–669, https://doi.org/10.1126/science.1651563.

[4] J. Groden, A. Thliveris, W. Samowitz, M. Carlson, L. Gelbert, H. Albertsen, G. Joslyn, J. Stevens, L. Spirio, M. Robertson, L. Sargeant, K. Krapcho, E. Wolff, R. Burt, J.P. Hughes, J. Warrington, J. McPherson, J. Wasmuth, D. Le Paslier, H. Abderrahim, D. Cohen, M. Leppert, R. White, Identification and characterization of the familial adenomatous polyposis coli gene, Cell 66 (1991) 589–600, https://doi.org/10.1016/0092-8674(81)90021-0.

[5] Y. Abe, A. Tada, J. Isoyama, S. Nagayama, R. Yao, J. Adachi, T. Tomonaga, Improved phosphoproteomic analysis for phosphosignaling and active-kinome profiling in Matrigel-embedded spheroids and patient-derived organoids, Sci. Rep. 8 (2018) 11401, https://doi.org/10.1038/s41598-018-29837-1.

[6] D. Gao, I. Vela, A. Sboner, P.J. Iaquinta, W.R. Karthaus, A. Gopalan, C. Dowling, J. N. Wanjala, E.A. Undvall, V.K. Arora, J. Wongvipat, M. Kossai, S. Ramazanoglu, L. P. Barboza, W. Di, Z. Cao, Q.F. Zhang, I. Sirota, L. Ran, T.Y. MacDonald, H. Beltran, J.-M. Mosquera, K.A. Touijer, P.T. Scardino, V.P. Laudone, K.R. Curtis, D. E. Rathkopf, M.J. Morris, D.C. Danila, S.F. Slovin, S.B. Solomon, J.A. Eastham, P. Chi, B. Carver, M.A. Rubin, H.I. Scher, H. Clevers, C.L. Sawyers, Y. Chen, Organoid cultures derived from patients with advanced prostate cancer, Cell 159 (2014) 176–187, https://doi.org/10.1016/j.cell.2014.08.016.

[7] M. van de Wetering, H.E. Francies, J.M. Francis, G. Bounova, F. Iorio, A. Pronk, W. van Houdt, J. van Gorp, A. Taylor-Weiner, L. Kester, A. McLaren-Douglas, J. Blokker, S. Jaksani, S. Bartfeld, R. Volckman, P. van Sluis, V.S.W. Li, S. Seepo, C. Sekhar Pedamallu, K. Cibulskis, S.L. Carter, A. McKenna, M.S. Lawrence, L. Lichtenstein, C. Stewart, J. Koster, R. Versteeg, A. van Oudenaarden, J. Saez-Rodriguez, R.G.J. Vries, G. Getz, L. Wessels, M.R. Stratton, U. McDermott, M. Meyerson, M.J. Garnett, H. Clevers, Prospective derivation of a living organoid biobank of colorectal cancer patients, Cell 161 (2015) 933–945, https://doi.org/10.1016/j.cell.2015.03.053.

[8] S.F. Boj, C.-I. Hwang, L.A. Baker, I.I.C. Chio, D.D. Engle, V. Corbo, M. Jager, M. Ponz-Sarvise, H. Tiriac, M.S. Spector, A. Gracanin, T. Oni, K.H. Yu, R. van Boxtel, M. Huch, K.D. Rivera, J.P. Wilson, M.E. Feigin, D. Öhlund, A. Handly-Santana, C.M. Ardito-Abraham, M. Ludwig, E. Elyada, B. Alagesan, G. Biffi, G. N. Yordanov, B. Delcuze, B. Creighton, K. Wright, Y. Park, F.H.M. Morsink, I. Q. Molenaar, I.H. Borel Rinkes, E. Cuppen, Y. Hao, Y. Jin, I.J. Nijman, C. Iacobuzio-Donahue, S.D. Leach, D.J. Pappin, M. Hammell, D.S. Klimstra, O. Basturk, R. H. Hruban, G.J. Offerhaus, R.G.J. Vries, H. Clevers, D.A. Tuveson, Organoid models of human and mouse ductal pancreatic cancer, Cell 160 (2015) 324–338, https://doi.org/10.1016/j.cell.2014.12.021.

[9] P.K. Dhanyamraju, T.D. Schell, S. Amin, G.P. Robertson, Drug-tolerant persister cells in cancer therapy resistance, Cancer Res. 82 (2022) 2503–2514, https://doi.org/10.1158/0008-5472.CAN-21-3844.

[10] M.L. De Angelis, F. Francescangeli, F. La Torre, A. Zeuner, Stem cell plasticity and dormancy in the development of cancer therapy resistance, Front. Oncol. 9 (2019), https://doi.org/10.3389/fonc.2019.00626.

[11] M. Kuksin, D. Morel, M. Aglave, F.-X. Danlos, A. Marabelle, A. Zinovyev, D. Gautheret, L. Verlingue, Applications of single-cell and bulk RNA sequencing in onco-immunology, Eur. J. Cancer 149 (2021) 193–210, https://doi.org/10.1016/j.ejca.2021.03.005.

[12] K. Sehgal, A. Portell, E.V. Ivanova, P.H. Lizotte, N.R. Mahadevan, J.R. Greene, A. Vajdi, C. Gurjao, T. Teceno, L.J. Taus, T.C. Thai, S. Kitajima, D. Liu, T. Tani, M. Noureddine, C.J. Lau, P.T. Kirschmeier, D. Liu, M. Giannakis, R.W. Jenkins, P. C. Gokhale, S. Goldoni, M. Pinzon-Ortiz, W.D. Hastings, P.S. Hammerman, J. J. Miret, C.P. Paweletz, D.A. Barbie, Dynamic single-cell RNA sequencing identifies immunotherapy persister cells following PD-1 blockade, J. Clin. Invest. 131 (2021), https://doi.org/10.1172/JCI135038.

[13] H. Kalkavan, M.J. Chen, J.C. Crawford, G. Quarato, P. Fitzgerald, S.W.G. Tait, C. R. Goding, D.R. Green, Sublethal cytochrome c release generates drug-tolerant persister cells, Cell 185 (2022) 3356–3374.e22, https://doi.org/10.1016/j.cell.2022.07.025.

[14] R. Petegrosso, Z. Li, R. Kuang, Machine learning and statistical methods for clustering single-cell RNA-sequencing data, Brief. Bioinform. 21 (2020) 1209–1223, https://doi.org/10.1093/bib/bbz063.

[15] T. Okamoto, D. duVerle, K. Yaginuma, Y. Natsume, H. Yamanaka, D. Kusama, M. Fukuda, M. Yamamoto, F. Perraudeau, U. Srivastava, Y. Kashima, A. Suzuki, Y. Kuze, Y. Takahashi, M. Ueno, Y. Sakai, T. Noda, K. Tsuda, Y. Suzuki, S. Nagayama, R. Yao, Comparative analysis of patient-matched PDOs revealed a reduction in OLFM4-associated clusters in metastatic lesions in colorectal cancer, Stem Cell Rep. 16 (2021) 954–967, https://doi.org/10.1016/j.stemcr.2021.02.012.

[16] G. La Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M.E. Kastriti, P. Lönnerberg, A. Furlan, J. Fan, L.E. Borm, Z. Liu, D. van Bruggen, J. Guo, X. He, R. Barker, E. Sundström, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, P.V. Kharchenko, RNA velocity of single cells, Nature 560 (2018) 494–498, https://doi.org/10.1038/s41586-018-0414-6.

[17] Y. Hao, S. Hao, E. Andersen-Nissen, W.M. Mauck, S. Zheng, A. Butler, M.J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E.P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L.M. Fleming, B. Yeung, A.J. Rogers, J. M. McElrath, C.A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data, Cell 184 (2021) 3573–3587.e29, https://doi.org/10.1016/j.cell.2021.04.048.

[18] C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression, Genome Biol. 20 (2019) 296, https://doi.org/10.1186/s13059-019-1874-1.

[19] G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.

[20] W. Luo, C. Brouwer, Pathview: an R/Bioconductor package for pathway-based data integration and visualization, Bioinformatics 29 (2013) 1830–1831, https://doi.org/10.1093/bioinformatics/btt285.

[21] E. Dann, N.C. Henderson, S.A. Teichmann, M.D. Morgan, J.C. Marioni, Differential abundance testing on single-cell data using k-nearest neighbor graphs, Nat. Biotechnol. 40 (2022) 245–253, https://doi.org/10.1038/s41587-021-01033-z.

[22] V. Bergen, M. Lange, S. Peidli, F.A. Wolf, F.J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling, Nat. Biotechnol. 38 (2020) 1408–1414, https://doi.org/10.1038/s41587-020-0591-3.

[23] J. Cao, M. Spielmann, X. Qiu, X. Huang, D.M. Ibrahim, A.J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F.J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis, Nature 566 (2019) 496–502, https://doi.org/10.1038/s41586-019-0969-x.

[24] L. Haghverdi, A.T.L. Lun, M.D. Morgan, J.C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors, Nat. Biotechnol. 36 (2018) 421–427, https://doi.org/10.1038/nbt.4091.

[25] M. Kuhn, Building predictive models in R using the caret package, J. Stat. Softw. 28 (2008) 1–26, https://doi.org/10.18637/jss.v028.i05.

[26] T.-C. Chou, Drug combination studies and their synergy quantification using the Chou-Talalay method, Cancer Res. 70 (2010) 440–446, https://doi.org/10.1158/0008-5472.CAN-09-1947.

[27] K. Lin, S. Chowdhury, M.A. Zeineddine, F.A. Zeineddine, N.J. Hornstein, O. E. Villarreal, D.M. Maru, C.L. Haymaker, J.-N. Vauthey, G.J. Chang, E. Bogatenkova, D. Menter, S. Kopetz, J.P. Shen, Identification of colorectal cancer cell stemness from single-cell RNA sequencing, Mol. Cancer Res. 22 (2024) 337–346, https://doi.org/10.1158/1541-7786.MCR-23-0468.

[28] J. Liu, J. Qiu, Z. Zhang, L. Zhou, Y. Li, D. Ding, Y. Zhang, D. Zou, D. Wang, Q. Zhou, T. Lang, SOX4 maintains the stemness of cancer cells via transcriptionally enhancing HDAC1 revealed by comparative proteomics study, Cell Biosci. 11 (2021) 23, https://doi.org/10.1186/s13578-021-00539-y.

[29] T. Fevr, S. Robine, D. Louvard, J. Huelsken, Wnt/β-Catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells, Mol. Cell. Biol. (2007), https://doi.org/10.1128/MCB.01034-07.

[30] P. Ordóñez-Morán, C. Dafflon, M. Imajo, E. Nishida, J. Huelsken, HOXA5 counteracts stem cell traits by inhibiting Wnt signaling in colorectal cancer, Cancer Cell 28 (2015) 815–829, https://doi.org/10.1016/j.ccell.2015.11.001.

[31] P. Busenhart, A. Montalban-Arques, E. Katkeviciute, Y. Morsy, C.V. Passen, L. Hering, K. Atrott, S. Lang, J.F.G. Garzon, E. Naschberger, A. Hartmann, G. Rogler, M. Stürzl, M.R. Spalinger, M. Scharl, Inhibition of integrin αvβ6 sparks T-cell antitumor response and enhances immune checkpoint blockade therapy in colorectal cancer, J. Immunother. Cancer 10 (2022) e003465, https://doi.org/10.1136/jitc-2021-003465.

[32] A. Wu, S. Zhang, J. Liu, Y. Huang, W. Deng, G. Shu, G. Yin, Integrated analysis of prognostic and immune associated integrin family in ovarian cancer, Front. Genet. 11 (2020), https://doi.org/10.3389/fgene.2020.00705.

[33] T. Liu, Y. Gu, Y. Zhang, Y. Li, Integrin α2 in the microenvironment and the tumor compartment of digestive (gastrointestinal) cancers: emerging regulators and therapeutic opportunities, Front. Oncol. 14 (2024), https://doi.org/10.3389/fonc.2024.1439709.

[34] M. Konopleva, R. Contractor, T. Tsao, I. Samudio, P.P. Ruvolo, S. Kitada, X. Deng, D. Zhai, Y.-X. Shi, T. Sneed, M. Verhaegen, M. Soengas, V.R. Ruvolo, T. McQueen, W.D. Schober, J.C. Watt, T. Jiffar, X. Ling, F.C. Marini, D. Harris, M. Dietrich, Z. Estrov, J. McCubrey, W.S. May, J.C. Reed, M. Andreeff, Mechanisms of apoptosis sensitivity and resistance to the BH3 mimetic ABT-737 in acute myeloid leukemia, Cancer Cell 10 (2006) 375–388, https://doi.org/10.1016/j.ccr.2006.10.006.

[35] A. Kotschy, Z. Szlavik, J. Murray, J. Davidson, A.L. Maragno, G. Le Toumelin-Braizat, M. Chanrion, G.L. Kelly, J.-N. Gong, D.M. Moujalled, A. Bruno, M. Csekei, A. Paczal, Z.B. Szabo, S. Sipos, G. Radics, A. Proszenyak, B. Balint, L. Ondi, G. Blasko, A. Robertson, A. Surgenor, P. Dokurno, I. Chen, N. Matassova, J. Smith, C. Pedder, C. Graham, A. Studeny, G. Lysiak-Auvity, A.-M. Girard, F. Gravé, D. Segal, C.D. Riffkin, G. Pomilio, L.C.A. Galbraith, B.J. Aubrey, M.S. Brennan, M. J. Herold, C. Chang, G. Guasconi, N. Cauquil, F. Melchiore, N. Guigal-Stephan, B. Lockhart, F. Colland, J.A. Hickman, A.W. Roberts, D.C.S. Huang, A.H. Wei, A. Strasser, G. Lessene, O. Geneste, The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models, Nature 538 (2016) 477–482, https://doi.org/10.1038/nature19830.

[36] A. Najem, M. Krayem, F. Salès, N. Hussein, B. Badran, C. Robert, A. Awada, F. Journe, G.E. Ghanem, P53 and MITF/Bcl-2 identified as key pathways in the acquired resistance of NRAS-mutant melanoma to MEK inhibition, Eur. J. Cancer 83 (2017) 154–165, https://doi.org/10.1016/j.ejca.2017.06.033.

[37] R.B. Corcoran, K.A. Cheng, A.N. Hata, A.C. Faber, H. Ebi, E.M. Coffee, P. Greninger, R.D. Brown, J.T. Godfrey, T.J. Cohoon, Y. Song, E. Lifshits, K.E. Hung, T. Shioda, D. Dias-Santagata, A. Singh, J. Settleman, C.H. Benes, M. Mino-Kenudson, K.-K. Wong, J.A. Engelman, Synthetic lethal interaction of combined BCL-XL and MEK inhibition promotes tumor regressions in KRAS mutant cancer models, Cancer Cell 23 (2013) 121–128, https://doi.org/10.1016/j.ccr.2012.11.007.

[38] A.A. Rambal, Z.L.G. Panaguiton, L. Kramer, S. Grant, H. Harada, MEK inhibitors potentiate dexamethasone lethality in acute lymphoblastic leukemia cells through the pro-apoptotic molecule BIM, Leukemia 23 (2009) 1744–1754, https://doi.org/10.1038/leu.2009.80.

[39] C. Alcon, F. Martín, E. Prada, J. Mora, A. Soriano, G. Guillén, S. Gallego, J. Roma, J. Samitier, A. Villanueva, J. Montero, MEK and MCL-1 sequential inhibition synergize to enhance rhabdomyosarcoma treatment, Cell Death Dis. 8 (2022) 1–12, https://doi.org/10.1038/s41420-022-00959-w.

[40] M. Winkler, J. Friedrich, C. Boedicker, N. Dolgikh, Co-targeting MCL-1 and ERK1/2 kinase induces mitochondrial apoptosis in rhabdomyosarcoma cells, Transl. Oncol. 16 (2022) 101313, https://doi.org/10.1016/j.tranon.2021.101313.

[41] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1977) 159–174, https://doi.org/10.2307/2529310.

[42] G. Vlachogiannis, S. Hedayat, A. Vatsiou, Y. Jamin, J. Fernández-Mateos, K. Khan, A. Lampis, K. Eason, I. Huntingford, R. Burke, M. Rata, D.-M. Koh, N. Tunariu, D. Collins, S. Hulkki-Wilson, C. Ragulan, I. Spiteri, S.Y. Moorcraft, I. Chau, S. Rao, D. Watkins, N. Fotiadis, M. Bali, M. Darvish-Damavandi, H. Lote, Z. Eltahir, E. C. Smyth, R. Begum, P.A. Clarke, J.C. Hahne, M. Dowsett, J. de Bono, P. Workman, A. Sadanandam, M. Fassan, O.J. Sansom, S. Eccles, N. Starling, C. Braconi, A. Sottoriva, S.P. Robinson, D. Cunningham, N. Valeri, Patient-derived organoids model treatment response of metastatic gastrointestinal cancers, Science 359 (2018) 920–926, https://doi.org/10.1126/science.aao2774.

[43] M.P. Ponnusamy, S.K. Batra, Ovarian cancer: emerging concept on cancer stem cells, J. Ovarian Res. 1 (2008) 4, https://doi.org/10.1186/1757-2215-1-4.

[44] W. Friedl, S. Aretz, Familial adenomatous polyposis: experience from a study of 1164 unrelated German polyposis patients, Hered. Cancer Clin. Pract. 3 (2005) 95, https://doi.org/10.1186/1897-4287-3-3-95.

[45] R. Bjerkvig, B.B. Tysnes, K.S. Aboody, J. Najbauer, A.J.A. Terzis, The origin of the cancer stem cell: current controversies and new insights, Nat. Rev. Cancer 5 (2005) 899–904, https://doi.org/10.1038/nrc1740.

[46] J.-H. Hwang, J. Yoon, Y.-H. Cho, P.-H. Cha, J.-C. Park, K.-Y. Choi, A mutant KRAS-induced factor REG4 promotes cancer stem cell properties via Wnt/β-catenin signaling, Int. J. Cancer 146 (2020) 2877–2890, https://doi.org/10.1002/ijc.32728.

[47] Y. Pu, L. Li, H. Peng, L. Liu, D. Heymann, C. Robert, F. Vallette, S. Shen, Drug-tolerant persister cells in cancer: the cutting edges and future directions, Nat. Rev. Clin. Oncol. 20 (2023) 799–813, https://doi.org/10.1038/s41571-023-00815-5.

[48] P. Valent, D. Bonnet, R. De Maria, T. Lapidot, M. Copland, J.V. Melo, C. Chomienne, F. Ishikawa, J.J. Schuringa, G. Stassi, B. Huntly, H. Herrmann, J. Soulier, A. Roesch, G.J. Schuurhuis, S. Wöhrer, M. Arock, J. Zuber, S. Cerny-Reiterer, H.E. Johnsen, M. Andreeff, C. Eaves, Cancer stem cell definitions and terminology: the devil is in the details, Nat. Rev. Cancer 12 (2012) 767–775, https://doi.org/10.1038/nrc3368.

[49] B. Beck, C. Blanpain, Unravelling cancer stem cell potential, Nat. Rev. Cancer 13 (2013) 727–738, https://doi.org/10.1038/nrc3597.

[50] E. Batlle, H. Clevers, Cancer stem cells revisited, Nat. Med. 23 (2017) 1124–1134, https://doi.org/10.1038/nm.4409.

[51] P.K. Das, F. Islam, A.K. Lam, The roles of cancer stem cells and therapy resistance in colorectal carcinoma, Cells 9 (2020) 1392, https://doi.org/10.3390/cells9061392.

[52] Y. Ohta, M. Fujii, S. Takahashi, A. Takano, K. Nanki, M. Matano, H. Hanyu, M. Saito, M. Shimokawa, S. Nishikori, Y. Hatano, R. Ishii, K. Sawada, A. Machinaga, W. Ikeda, T. Imamura, T. Sato, Cell–matrix interface regulates dormancy in human colon cancer stem cells, Nature 608 (2022) 784–794, https://doi.org/10.1038/s41586-022-05043-y.

[53] H.-M. Zhou, J.-G. Zhang, X. Zhang, Q. Li, Targeting cancer stem cells for reversing therapy resistance: mechanism, signaling, and prospective agents, Sig. Transduct. Target Ther. 6 (2021) 1–17, https://doi.org/10.1038/s41392-020-00430-1.

[54] T. Shibue, R.A. Weinberg, EMT, CSCs, and drug resistance: the mechanistic link and clinical implications, Nat. Rev. Clin. Oncol. 14 (2017) 611–629, https://doi.org/10.1038/nrclinonc.2017.44.

[55] H.F. Cabanos, A.N. Hata, Emerging insights into targeted therapy-tolerant persister cells in cancer, Cancers 13 (2021) 2666, https://doi.org/10.3390/cancers13112666.

[56] G. De Conti, M.H. Dias, R. Bernards, Fighting drug resistance through the targeting of drug-tolerant persister cells, Cancers 13 (2021) 1118, https://doi.org/10.3390/cancers13051118.

[57] T. Yamaguchi, R. Kakefuda, N. Tajima, Y. Sowa, T. Sakai, Antitumor activities of JTP-74057 (GSK1120212), a novel MEK1/2 inhibitor, on colorectal cancer cell lines in vitro and in vivo, Int. J. Oncol. 39 (2011) 23–31, https://doi.org/10.3892/ijo.2011.1015.

[58] A. Tsherniak, F. Vazquez, P.G. Montgomery, B.A. Weir, G. Kryukov, G.S. Cowley, S. Gill, W.F. Harrington, S. Pantel, J.M. Krill-Burger, R.M. Meyers, L. Ali, A. Goodale, Y. Lee, G. Jiang, J. Hsiao, W.F.J. Gerath, S. Howell, E. Merkel, M. Ghandi, L.A. Garraway, D.E. Root, T.R. Golub, J.S. Boehm, W.C. Hahn, Defining a cancer dependency map, Cell 170 (2017) 564–576.e16, https://doi.org/10.1016/j.cell.2017.06.010.

[59] S.A. Forbes, G. Tang, N. Bindal, S. Bamford, E. Dawson, C. Cole, C.Y. Kok, M. Jia, R. Ewing, A. Menzies, J.W. Teague, M.R. Stratton, P.A. Futreal, COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer, Nucleic Acids Res. 38 (2010) D652–D657, https://doi.org/10.1093/nar/gkp995.

[60] Y. Nojima, M. Aoki, S. Re, H. Hirano, Y. Abe, R. Narumi, S. Muraoka, H. Shoji, K. Honda, T. Tomonaga, K. Mizuguchi, N. Boku, J. Adachi, Integration of pharmacoproteomic and computational approaches reveals the cellular signal transduction pathways affected by apatinib in gastric cancer cell lines, Comput. Struct. Biotechnol. J. 21 (2023) 2172–2187, https://doi.org/10.1016/j.csbj.2023.03.006.

[61] S. Misale, F. Di Nicolantonio, A. Sartore-Bianchi, S. Siena, A. Bardelli, Resistance to anti-EGFR therapy in colorectal cancer: from heterogeneity to convergent evolution, Cancer Discov. 4 (2014) 1269–1280, https://doi.org/10.1158/2159-8290.CD-14-0462.

[62] T. Nakahara, M. Takeuchi, I. Kinoyama, T. Minematsu, K. Shirasuna, A. Matsuhisa, A. Kita, F. Tominaga, K. Yamanaka, M. Kudoh, M. Sasamata, YM155, a novel small-molecule survivin suppressant, induces regression of established human hormone-refractory prostate tumor xenografts, Cancer Res. 67 (2007) 8014–8021, https://doi.org/10.1158/0008-5472.CAN-07-1343.

[63] Y. Wang, T. Zhang, N. Kwiatkowski, B.J. Abraham, T.I. Lee, S. Xie, H. Yuzugullu, T. Von, H. Li, Z. Lin, D.G. Stover, E. Lim, Z.C. Wang, J.D. Iglehart, R.A. Young, N. S. Gray, J.J. Zhao, CDK7-dependent transcriptional addiction in triple-negative breast cancer, Cell 163 (2015) 174–186, https://doi.org/10.1016/j.cell.2015.08.063.

[64] C.R. Loomis, R.M. Bell, Sangivamycin, a nucleoside analogue, is a potent inhibitor of protein kinase C, J. Biol. Chem. 263 (1988) 1682–1692, https://doi.org/10.1016/S0021-9258(19)77930-7.

[65] Y. Cao, H. Tang, G. Wang, P. Li, Z. Song, W. Li, X. Sun, X. Zhong, Q. Yu, S. Zhu, L. Zhu, Targeting survivin with Tanshinone IIA inhibits tumor growth and overcomes chemoresistance in colorectal cancer, Cell Death Dis. 9 (2023) 1–12, https://doi.org/10.1038/s41420-023-01622-8.

[66] A. Rauch, A. Carlstedt, C. Emmerich, A.-H.M. Mustafa, A. Göder, S.K. Knauer, M. Linnebacher, T. Heinzel, O.H. Krämer, Survivin antagonizes chemotherapy-induced cell death of colorectal cancer cells, Oncotarget 9 (2018) 27835–27850, https://doi.org/10.18632/oncotarget.25600.

[67] X.J. Cheng, J.C. Lin, Y.F. Ding, L. Zhu, J. Ye, S.P. Tu, Survivin inhibitor YM155 suppresses gastric cancer xenograft growth in mice without affecting normal tissues, Oncotarget 7 (2016) 7096–7109, https://doi.org/10.18632/oncotarget.6898.

[68] C.-C. Cheng, J. Chang, S.C.-C. Huang, H.-C. Lin, A.-S. Ho, K.-H. Lim, C.-C. Chang, L. Huang, Y.-C. Chang, Y.-F. Chang, C.-W. Wu, YM155 as an inhibitor of cancer stemness simultaneously inhibits autophosphorylation of epidermal growth factor receptor and G9a-mediated stemness in lung cancer cells, PLoS ONE 12 (2017) e0182149, https://doi.org/10.1371/journal.pone.0182149.

[69] R.A. Syahrani, E. Yunita, S.I. Wanandi, Suppression of rotenone-treated human breast cancer stem cell survival using survivin inhibitor YM155 is associated to oxidative stress modulation, Asian Pac. J. Cancer Prev. 21 (2020) 2631–2637, https://doi.org/10.31557/APJCP.2020.21.9.2631.

[70] W.K.L. Doolittle, L. Zhao, S.-Y. Cheng, Blocking CDK7-mediated NOTCH1-cMYC signaling attenuates cancer stem cell activity in anaplastic thyroid cancer, Thyroid 32 (2022) 937–948, https://doi.org/10.1089/thy.2022.0087.
[71] Q. Wu, Z. Yang, Y. Nie, Y. Shi, D. Fan, Multi-drug resistance in cancer chemotherapeutics: mechanisms and lab approaches, Cancer Lett. 347 (2014) 159–166, https://doi.org/10.1016/j.canlet.2014.03.013.
[72] M. Nishide, H. Shimagami, A. Kumanogoh, Single-cell analysis in rheumatic and allergic diseases: insights for clinical practice, Nat. Rev. Immunol. 24 (2024) 1–17, https://doi.org/10.1038/s41577-024-01043-3.