Single Cell Ca^{2+} profiling for the discrimination of cancer cell lines.

Camille Caussette[1], Amélie Bura[1], Maxime Guéguinou[1], Dorine Bellanger[1], Alison Robert[1], Mathilde Cancel[1,2], Gaëlle Fromont-Hankard[1,3], Christophe Vandier[1], Karine Mahéo[1], Thierry Brouard[4]*, David Crottès[1]*

1. Inserm UMR1069 N2Cox « Niche Nutrition Cancer Oxydatif », Tours, France
2. CHRU Tours, Oncology, Tours, France
3. CHRU Tours, Pathological Anatomy and Cytology, Tours, France
4. LIFAT EA6300, Tours, France

* authors contributed equally to this work.
Corresponding author : david.crottes@univ-tours.fr

ABSTRACT:

Calcium (Ca^{2+}) is an essential and ubiquitous second messenger controlling numerous cellular functions. Ca^{2+} signaling relied on the finely tuned oscillations of the cytosolic Ca^{2+} concentrations induced by components of Ca^{2+} signaling toolkit (ion channels, pumps and ion exchangers). The regulation of these Ca^{2+} oscillations define a Ca^{2+} signature that is representative of the cellular identity and phenotype of a cell. In cancers, molecular actors of the Ca^{2+} signaling toolkit are aberrantly expressed. We hypothesized that Ca^{2+} signature of cancer cells are representative of their cellular identity, their tissue of origins (TOO) as well as their isolation site (IS). We defined the Ca^{2+} signature of prostate and colon cancer cell lines by collecting the profile of cytosolic Ca^{2+} responses evoked by a panel of agonists in 22904 individual cells. We then highlighted the heterogeneity of those Ca^{2+} profiles and successfully developed a classifier predicting the tissue of origins (TOO), the isolation site (IS) or the cellular identity of individual cancer cells using a supervised neural network. Unsupervised clustering revealed that Ca^{2+} profiles of single cancer cells derived from 3 main classes of Ca^{2+} responses sub-divided into 50 different clusters. Thus, we highlighted that supervised machine learning applied on top of single cell Ca^{2+} profiling is an effective method to discriminate cancer cells at single cell level and that the cancer cell Ca^{2+} signature can be summarized into 3 main profiles of Ca^{2+} responses.

INTRODUCTION:

Calcium (Ca^{2+}) is an essential and ubiquitous second messenger controlling numerous cellular functions (proliferation, apoptosis, migration, metabolism, etc.) and involved in multiple pathologies including cancer. Ca^{2+} signaling relied on the finely tuned oscillations of the cytosolic Ca^{2+} concentration induced by components of the Ca^{2+} signaling toolkit (ion channels, pumps and ion exchangers) in response to input signals. Ca^{2+} signaling transducers (NFAT, NF-kB, ..) decode oscillations of the cytosolic Ca^{2+} concentration and generate the appropriate cellular behavior (1, 2)

The spatio-temporal regulation of those cytosolic Ca^{2+} oscillations, so-called “Ca^{2+} signature”, is essential to define the appropriate cellular response (3). Thus, it has been reported that different cell types or tissues will present distinct characteristics of cytosolic Ca^{2+} oscillations (duration, frequency, amplitude, …) (4–8). Ca^{2+} signature can be associated to distinct phenotype of an isogenic cell type. In endothelial cells, depending of the extracellular concentration of VEGF, different Ca^{2+} responses are elicited, activating different signaling pathways and contributing to endothelium development either through proliferation or migration (9). In Dorsal Root Ganglion (DRG) neurons, 4 different profiles of Ca^{2+} responses corresponding to distinct subtypes of DRG neurons can be induced by Veratridine (inhibitor of voltage-gated sodium channels) (10). Moreover, cortical and cerebellar astrocytes exhibit particular profiles of Ca^{2+} responses when stimulated by different agonists (11). This heterogeneity of Ca^{2+} responses allow a complex pattern of signaling

responses contributing to the diversification of cellular responses and the establishment of multiples phenotypes (12, 13).

Calcium oscillations play a crucial role in encoding information to coordinate various transcriptional programs. Artificial manipulation of the profile of Ca^{2+} signal can triggers the activation of various signaling pathways (14–16). Transcription factors that are sensitive to Ca^{2+}, such as NFAT (Nuclear Factor of Activated T-cells) and NF-κB, are intricately regulated by cytosolic calcium waves. The frequent occurrence of Ca^{2+} oscillations has been observed to lead to the activation of NFAT. Those oscillations are tightly regulated by the specific action of Ca^{2+} channels (mainly voltage-gated Ca^{2+} channels in excitable cells and store-operated Ca^{2+} channels in non-excitable cells). In non-excitable cells, depending on the duration and intensity of responses generated by the specific recruitment of Ca^{2+} channels, namely Orai1, Orai2, and Orai3, it is possible to selectively influence the activation of NFAT C1 and C4 isoforms. This differential modulation allows for the induction of a unique cellular response, adding an additional layer of complexity to the regulatory pathways governed by Ca^{2+} signaling (16).

Cancer is a highly heterogeneous disease. In recent years, the genetic and phenotypic heterogeneity of cancer cells have been associated to tumor aggressiveness and treatment resistance (17–19). Multiple sub-populations of cancer cells can thus exist in a single tumor and present distinct oncogenic functions or synergistically enhance invasiveness (19). Ca^{2+} fluxes are modified in cancer cells and have been associated to hallmarks of cancers such as migration, proliferation or chemotherapy resistance (20–22). Consequently, it has been observed that ion channels involved in Ca^{2+} signaling are frequently over-expressed in multiple cancers (21, 23–25). Interestingly, aberrant expression of ion channels is heterogeneous and function of the cancer type. For instance, Orai1 is over-expressed in liver, esophageal, lung, breast and colorectal cancers while Orai3 are over-expressed in prostate, breast and lung cancers (22, 26). This suggest that the Ca^{2+} signature of cancer cells should be different according the profile of expression of Ca^{2+} signaling toolkit and may thus be representative of their origins or lineage.

In cancer cells, Ca^{2+} signaling has been largely investigated through the prism of the Store-Operated Ca^{2+} entry (SOCE) (22). In those cells, the activation of various plasma membrane (PM) receptors stimulates the production of the diffusible second messenger inositol 1,4,5-trisphosphate (IP3). IP3 then binds to IP3 receptor (IP3R) channels located in the endoplasmic reticulum (ER), leading to the release of ER luminal Ca^{2+} into the cytosol and resulting in ER store depletion. This depletion causes a conformational change in the Ca^{2+}-sensing protein stromal interacting molecule (STIM). STIM translocates to ER–PM junctions, where it traps and activates ORAI (primarily Orai1) and TRPC (mainly TRPC1) channels, initiating the SOCE. When ectopically co-expressed with STIM1, the Orai1 homologs Orai2 and Orai3 are capable of mediating SOCE (16).

In this study, we postulated that Ca^{2+} signature of cancer cells are representative of their cellular identity, their TOO as well as their IS. We developed a workflow allowing to define the Ca^{2+} signature of prostate and colon cancer cell lines by collecting the profile of cytosolic Ca^{2+} responses evoked by a panel of agonists in 22904 individual cells. We then highlighted the heterogeneity of those Ca^{2+} profiles and successfully developed a classifier predicting the TOO, the IS or the cellular identity of individual cancer cells using a supervised neural network. Our results suggested thus that the single cell Ca^{2+} profiling has the potential to discriminate various sub-populations of cancer cells with distinct genetic characteristics. Unsupervised clustering revealed that Ca^{2+} profiles of single cancer cells are all derivating of 3 main classes of Ca^{2+} responses sub-divided into 50 different clusters. Our study unveils the potential of Ca^{2+} profiling for the discrimination of cancer cells and the identification of their phenotypic characteristics at single cell level and provides a framework for researcher to investigate the remodeling of Ca^{2+} signature during cancer development.

RESULTS:

Store-Operated Ca^{2+} Entry is not associated to the tissue of origins (TOO) or the isolation site (IS) of cancer cell lines.

To reliably assess SOCE function, many studies have utilized protocols that trigger maximal SOCE activation. Thapsigargin, a sarcoplasmic ATPase/ER Ca^{2+} (SERCA) blocker, is commonly employed to completely deplete ER Ca^{2+} stores, inducing conditions of maximum reserve depletion. Under these conditions, the addition of Ca^{2+} to the extracellular medium typically results in sustained and stable Ca^{2+} entry into the cytoplasm, demonstrating maximum SOCE influx. However, it's important to note that this type of influx is rare, if not impossible, under real-life conditions.

We wondered if the measure of the Tg-induced SOCE could reflect the tissue of origins (TOO) or the isolation site (IS) of cancer cells. For that purpose, we selected a panel of cancer cell lines originated from colon or prostate cancers and isolated from non-cancerous tissues, primary or metastatic tumors. We observed a robust Tg-induced SOCE in all of the 16 human cell lines assayed (Figure 1A). While we distinguished minimal variation of the amplitude of the ER Ca^{2+} release and of the extracellular Ca^{2+} influx (Figure 1A and 1B), we did not observe any significant difference of the Tg-induced ER release and of the Tg-induced Ca^{2+} influx in function of the TOO or the IS of cell lines (Figure 1C and 1D). These results suggest that in our panel of cell lines, the SOCE is not different according their TOO or IS.

Single cell Ca^{2+} profiling analysis of cancer cell lines.

In most *in vivo* contexts, cells respond to agonists present in their microenvironment (whether physiological or pathological) with relatively low agonist concentrations (1 to 10 µM). Ca^{2+} signals in most cells typically manifest as transient regenerative oscillations or spikes. The frequency, amplitude, and duration of these Ca^{2+} oscillations can vary depending on the type and concentration of the agonist used, the expression of receptors targeted by the agonist, and, of course, the nature of the calcium channels or regulators expressed by the cell.

We then wondered if agonists-evoked Ca^{2+} responses could permit to discriminate cancer cell lines according their TOO or IS. To verify this hypothesis, we designed an assay recording agonists-elicited Ca^{2+} responses at single cell level. Agonists selected included Adenosine Tri-phosphate (ATP), Lysophosphatidic Acid (LPA), Acetyl-Choline (Ach.), Platelet-Aggregating Factor (PAF), Histamine (Hist.), Epidermal Growth Factor (EGF) and Prostaglandin E_2 (PGE_2). All agonists were assayed at physio-pathological concentrations previously reported to elicit Ca^{2+} responses. Thapsigargin (Tg) and the vehicle (Veh.) were used as a positive and negative control respectively. All 16 cell lines were stimulated by 13 different conditions. Altogether, we acquired 22904 single cell profile of Ca^{2+} responses (Figure 2). For each profiles, we computed the maximal amplitude, the AUC and the number of oscillations. Those values were averaged per cell lines and agonists tested.

Depending of the agonist and cell line assayed, we observed heterogeneous profiles of Ca^{2+} responses (Figure 2). Taken individually, only the measure of the maximal amplitude of the Ca^{2+} responses elicited by EGF was able to discriminate cell lines according both TOO and IS (Figure 3A). We highlighted that the amplitude of Ca^{2+} responses elicited by Acetyl-Choline (1 µM) is significantly higher in cancer cell lines isolated from primary tumors compared to metastatic tumors (Figure 3B). Supra-physiologically concentrations of PAF produced a significantly higher AUC and number of oscillations in prostate cell lines compared to colon cell lines (Figure 3B and C). Other conditions assayed did not provide significant differences between cell lines according to their TOO or IS.

Thus, EGF is the only agonist of our panel that can potentially discriminate cell lines according both TOO and IS on the basis of the amplitude of the Ca^{2+} response elicited.

Artificial Neural Network trained with profile of Ca^{2+} responses discriminate single cancer cells.

With the increasing complexity and volume of data generated by modern high throughput technologies, the use of machine learning has become more spread. Machine learning algorithms take advantages of the complexity of data to either cluster samples into natural patterns based on their similarities (unsupervised machine learning) or classify samples according a specified labeling allowing for the prediction of the labeling of new samples (supervised machine learning) .

Here, we applied supervised machine learning to build a classifier for the prediction of the TOO, the IS and the cellular identity (cell line name) of each single cells based on their Ca^{2+} response. We developed an artificial neural network (ANN) composed of 3 layers. The first corresponds to input data (normalized and scaled fluorescence intensity recorded at each timepoints), the second is a hidden layer of half-sized the input layer and the final layer is the output composed of different possibles labels (Figure 4A). Each ANN was run over 500 epochs and 5 different runs were performed for each conditions tested. Varying the size of the subset of the dataset used for training from 90 to 70% did not influence much performance of the ANN. We started to observe reduced performance when using 50% of the dataset for training (Supplemental Figure 4). For subsequent experiments, we set a splitting factor of 0.2 between validation and training (80% of the dataset used for training and 20% for validation)

When trained with both the complete Ca^{2+} response and the name of the molecule tested as input, the ANN performs relatively well for discriminating TOO, IS or cellular identity of individual cells (Figure 4B). No strong difference were observed for the prediction of TOO by our ANN with 13.9 and 16.3% of errors of prediction for CRC and PCa respectively (Figure 4B). Regarding the IS, ANN has a better accuracy for the prediction of cells isolated from primary or metastatic tumors (respectively 83.6 and 82.8%) than for the prediction of cells isolated from non-cancerous tissues (63.9%). ANN trained to predict the cellular identity of individual cells is able to predict 12 cell lines with an accuracy of prediction over 50% (8x more than what obtained by random pickup). LNCap, C4-2, C4-2b and DU-145 are cell lines with the least prediction accuracy. Interestingly, by removing the name of the agonist tested for each single cell Ca^{2+} response used for the training of the ANN induced a significant drop of the accuracy of prediction (Figure 4C) suggesting that this information is critical for the classification built by the training of the ANN.

To go further, we explored the performance of ANN trained with single cell calcium responses elicited by a single agonist. Surprisingly, those ANNs had similar performance compared to the ANN trained with the entire dataset and better performance compared to the ANN trained with the full dataset omitting the labels of agonists as input. For the prediction of cell lineage, we observed that ANNs trained with Ca^{2+} responses elicited by Thapsigargin, Histamine or ATP (10 µM) are performing much better for the prediction of LNCap, C4-2, C4-2b and DU-145 cell lines.

This suggest that information contained in the measure of the Ca^{2+} responses elicited by a n agonist is sufficient to predict its TOO, IS and lineage by supervised machine learning.

Unsupervised clustering of heterogeneous Ca^{2+} responses identifies the Ca^{2+} signature of cancer cell lines as a mixture of 3 classes of Ca^{2+} responses.

We noticed an important heterogeneity of the profile of Ca^{2+} responses recorded (Figure 2) as illustrated by the density analysis of the maximal amplitude of Ca^{2+} responses (Figure 5A and Supplemental Figure 2). We observed that a large variability in the amplitude of Ca^{2+} responses elicited by a single agonist in a single cell line (Figure 5). This suggests an important intra-lineage heterogeneity of Ca^{2+} responses.

We then explored this heterogeneity to highlight natural patterns of Ca^{2+} responses. Dimensionality reduction has been proven to improve the performance of some unsupervised clustering algorithm while facilitating the visualization of complex data (27). Here, we applied uniform manifold approximation projection (UMAP) on the dataset to obtain a projection of each individual Ca^{2+} responses in a two-dimensional space (Figure 6A). We thus observed two massive structures, one “mass” central and a second describing an arc of circle on the outer border of the 2D space.

Overlapping values of amplitude, AUC and the number of oscillations of each individual Ca^{2+} responses on the two dimensional space defined by UMAP highlight that the Ca^{2+} responses with the highest values for each of these parameters do not localize into the same spatial location (Figure 6A, B, and C). Cells with the highest amplitude and number of oscillations tend to locate mainly in the central mass while cells with the highest AUC tend to be cluster at the interface of both structures on the UMAP representation. This suggest that independently of the cell line or the agonist assayed, the UMAP representation allows to discriminate cells with a very small or absent Ca^{2+} responses (located in the outer arc of circle) from cells displaying high Ca^{2+} responses (located mainly in the central mass).

By visualizing the Ca^{2+} responses evoked for a single cell line or a single agonist, we observed that a large overlapping between cells belonging to different cell lines and stimulated by different agonists suggesting that the clustering did not segregate cells based on their cellular identity or the agonist tested. However, we observed that LNCap and C4-2 tend to locate essentially in the left outer border of the UMAP representation while RWPE-1 and PC3 locate mainly in the lower right corner (Supplemental Figure 3A). Tg was the agonist inducing the most homogeneous Ca^{2+} responses are all cells tend to cluster on very restricted spatial location on the UMAP projection (Supplemental Figure 3B). Cells stimulated by EGF, PGE_2 and Vehicle (all inducing Ca^{2+} responses of small amplitude) were localized mainly in the outer border of the UMAP representation in the spatial location in agreement with our observation that this area gathered cells with the weakest Ca^{2+} responses. These observations suggested that the dimension reduction allow the discrimination of different clusters of cells based on the characteristic of Ca^{2+} responses and that Ca^{2+} responses evoked in individual cell lines by individual agonists will tend to clusters in different groups with an important overlapping.

To explore different patterns of Ca^{2+} responses present in our dataset, we classified calcium responses using an unsupervised approach. In order to limit bias in the choice of the number of clusters “k”, we selected algorithms that did not require a prior definition of k value such as affinity propagation, DBScan, Optics and shared-nearest neighbor (SNN). Each algorithm were assayed either on the raw scaled dataset or on dataset reduced to two or three components using UMAP (designed as “UMAP 2D” and “UMAP 3D” respectively). Algorithms were assayed on multiple fractions of various sizes of the dataset, each one encompassing a sample of all conditions assayed. We were expecting that as the fraction of the dataset increased, the number of clusters identified by the algorithm should converge to a maximum value indicating that a finite number of clusters has been identified. We observed that DBScan did not succeed to find a coherent number of clusters independently of the prior dimension reduction of the dataset. Optics and Affinity propagation identified a linearly increasing numbers of clusters depending on the size of the dataset used for clustering. Interestingly, we observed that SNN applied on the dimension reduced dataset tend to converge to a maximal number of clusters when 70% of the dataset was used (Supplemental Figure 4A). 50 different clusters were thus identified using SNN applied on the dataset reduced to two components (Figure 6D and Supplemental Figure 4B). To characterize each of these clusters, we defined their amplitude, AUC and the number of oscillations. Thus, we observed that clusters 48 to 50 displayed the high amplitude and AUC. Clusters 29, 33, 35 and 36 contained Ca^{2+} responses with the highest number of oscillations. Thus, the unsupervised classification provided by SNN discriminate groups of cells with particular characteristics of Ca^{2+} responses (Figure 6E).

We then investigated the distribution of cells into each clusters in function of their characteristics (TOO, IS, cell lines, agonist tested). Clusters were sorted by hierarchical clustering according their average values of the amplitude, AUC and number of oscillations of Ca^{2+} responses. Clustering cell lines according their relative distribution into each clusters highlighted 3 main clusters: one including only CRC cell lines (SW620, DLD1, NCM356, Lovo, SW48), a second one including only prostate cell lines (VCap, LNCap, C4-2, C4-2b) and one intermediate including a mix of CRC and Pca cell lines (PC3, RWPE-1, DU-145, 22RV1, HCT-116, HT-29, SW480). We also observed that LNCap and its two derivatives (C4-2, C4-2b) cluster together suggesting that Ca^{2+} profiling identified their lineage and shared cellular identity (Figure 6F). Similarly, clustering agonists according the relative distribution of the number of Ca^{2+} responses into each clusters highlighted

interesting patterns. First, we observed that Tg is really distinct of other agonists as expected based on the particular Ca^{2+} responses induced. We also noticed that agonists eliciting Ca^{2+} responses with the lowest amplitude (EGF, Vehicle, PGE2) were gather together.

Visual comparison of Ca^{2+} responses for each clusters defined by SNN and hierarchical clustering in function of characteristics of Ca^{2+} responses suggested that some clusters displayed a similar Ca^{2+} profile and potentially should be grouped together (Supplemental Figure 4B). To verify this assumption, we averaged Ca^{2+} responses included into each primary clusters, reduced their dimension by UMAP and apply SNN clustering. SNN identified 3 clusters with distinct spatial location on the two-dimensional representation (Figure 6G and H). For each cell lines exposed to each agonists, we defined the relative proportion of cells into each of three "secondary" clusters defined by SNN (Supplemental Figure 5). Then, we summarized for each cell line the average proportion of cells observed for each agonists into each of those three clusters (Figure 6I). We observed that each cell lines displayed a particular pattern of distribution. For example, PC3 is distributed into cluster 1 and 3 with a minor presence in cluster 2 while LNCap and C4-2 distributed mainly in cluster 1 and 2 with a total absence of cluster 3.

Altogether, unsupervised clustering highlights that 22904 Ca^{2+} responses collected in this study can be grouped into 50 clusters that belong to 3 main classes of Ca^{2+} responses. Each cell lines can be defined by the probability to elicit a mix of these 3 main profiles of Ca^{2+} responses.

DISCUSSION:

Implicated in virtually all physiological processes, Ca^{2+} signaling has been widely investigated over years. The profile of cytosolic Ca^{2+} responses to various stimuli has been associated to specific signaling pathways or to different cell types (4, 5). Interestingly, cortical or cerebellar astrocytes displayed different profiles of Ca^{2+} responses to a panel of restricted agonists (11) suggesting that different cell types can be discriminated through their Ca2+ signature.

In this study, we defined the Ca^{2+} signature of cancer cell lines by recording Ca^{2+} responses induced by a panel of agonists. Using supervised machine learning, we defined a classifier allowing to predict the TOO, the IS and the cellular identity of individual cells based on the profile of Ca^{2+} response induced by an agonist (Figure 4). This result highlights that Ca^{2+} profiling can encode a sufficient amount of information to permit the identification of single cancer cells.

Our study is thus the first to propose supervised machine learning applied on top of Ca^{2+} profiling as an effective method to discriminate cancer cells at single cell level. The pipeline of analysis we developed has thus the potential to complete the arsenal of high-throughput technologies (such single cell RNASeq, spatial transcriptomics or multiplexed imaging methods) currently used for the investigation of intra-tumor heterogeneity and the identification of sub-populations of cancer cells. Advantage of Ca^{2+} profiling is to provide a rapid and inexpensive way to measure the functional identity of single cells while other methods rely on frozen / fixed samples at a defined time-point. However, further studies will be necessary to quantitatively assess the prediction accuracy of single cell Ca^{2+} profiling compared with single cell transcriptomics.

Inherent to all classifiers developed by supervised machine learning, our prediction model is limited by the finite list of existing output provided. To go further, it requires to compile Ca^{2+} signatures of additional cell lines of different TOO and IS and includes different phenotypes traits (such as chemotherapy resistance, aggressiveness, proliferative index, ...) in the list of outputs to be predicted. Recently, genetic algorithm has been used for optimizing neural network developed for the prediction of epileptic foci on iEEG recordings (28). In the study, the classifier developed is really simplistic containing only three layers of neurons and thus may benefit of such optimization strategies to improve its prediction accuracy while limiting over-fitting.

Cell lines may not recapitulate the full spectrum of characteristics of tumors observed in clinics despite preserving some key characteristics of tumors (29). In our assay, isogenic cell lines are cultured at small density and thus lack cell-cell interactions with surrounding cancer cells.

Compared to *ex-vivo* tissues, they also lack interactions with non-cancerous cells such as fibroblasts, stromal cells, adipocytes and immune cells. Thus, the definition of Ca^{2+} signatures for cancer cell lines provided in this study may not fully correspond to the Ca^{2+} signatures observed *in vivo*. Applying Ca^{2+} profiling for the investigation of sub-populations of cancer cells in heterogeneous tumor tissues will thus require to translate our workflow for the measure of Ca^{2+} responses *in vivo*. In this way, we previously developed a protocol allowing the measure of constitutive Ca^{2+} entry on organotypic cultures slice of prostate tumors (30). Prostate tumors being highly heterogeneous (31), the definition of Ca^{2+} signature at single cell level on this model may contribute to identify different clonal tumor site and being an interesting tool for the diagnosis and prognosis of those cancers. By being able to traceback the identity of cancer cells, single cell Ca^{2+} profiling could also be interesting for the identification of cancer with unknown origins (CUP). However, this will require to increment our database with example of profiles of Ca^{2+} responses obtained on cells isolated from CUP .

Ca^{2+} signaling has been widely investigated in cancers. Importantly, SOCE has been measured in different cancers and been associated to hallmarks of cancers (21, 22). In several cancer types, an up-regulation of SOCE in cancer cells compared to normal cells has been observed (22). Moreover, ion channels and proteins contributing to the SOCE have been observed as mis-expressed in cancers (22). Surprisingly, in our study, we noticed that SOCE is not significantly different in cell line according their TOO or their isolation site. Thus, this result suggesting that despite a different profile of expression of SOCE-related genes in our panel of cell lines, the mechanism by itself is moderately affected. This contrasting observation to previous literature can be explained by 1) our limited number of cell lines investigated 2) by our culture conditions (isolated cells on 2D culture plastic dishes) and 3) by the protocol and analysis workflow used to measure SOCE that may not permit the discrimination of subtle changes of SOCE in cancer cells according their isolation site. A more comprehensive analysis of SOCE in a larger number of cell lines or on isolated non-tumoral and tumoral tissues will be required in order to verify our assumptions.

Previous investigations have highlighted that profiles of cytosolic Ca^{2+} responses are highly heterogeneous and subsequently may constitute a marker of cellular heterogeneity. Hippocampal neurons stimulated with baclofen displayed 4 different Ca^{2+} profiles (32). Similarly in DRG neurons treated with an inhibitor of voltage-gated sodium channels present 4 distinct profiles of Ca^{2+} responses (10). Those investigations rely on a limited number of individual cells analyzed (up to 400 individual cells), of cell types (from 1 to 2) investigated and number of compounds tested (up to 4 compounds). In comparison, our study encompasses 16 cell lines from two different TOO and isolated from three different sites stimulated by 11 different compounds. Altogether, we have collected 22904 single cell Ca^{2+} responses. Using unsupervised clustering algorithm, we succeeded to cluster those profiles into 50 unique clusters belonging to 3 main classes of Ca^{2+} responses. Our database offers thus a deeper repertoire of Ca^{2+} responses to analyze cellular heterogeneity. Compared to other investigations using either manual annotation (10) or clustering based on a prior definition of the number of clusters to obtain (32), in our study we deliberately favored an approach using algorithms that did not require a pre-defined number of clusters. Thus, our analysis was less subjected to bias introduced by the definition of the number of k clusters. We have tested several algorithms including Affinity propagation, DBScan, OPTICS and Shared-nearest neighbors (SNN) on the dataset subjected or not to a dimensionality reduction. We found that SNN applied on two-dimensional reduced dataset provided the most promising clusterization of profiles of Ca^{2+} responses. Further investigations will be required to validate those findings and benchmark additional algorithms to evaluate their internal performance in a more unbiased manner.

It is interesting to note that our unsupervised clustering proposes 3 main classes of Ca^{2+} responses, a value really close to the 4 identified in other investigations (10, 32). This suggest that it exists a limited number of profiles of Ca^{2+} responses independently of the cell type investigated and that the declination of these main profiles generated the Ca^{2+} signature of each cell types. Further studies will be required to assay if this clustering can be observed in independent dataset

including other cell types and to determine molecular actors responsible for this declination into Ca^{2+} signatures for cell lines.

Altogether, in this study, we propose for the first time that single cell Ca^{2+} profiling as an effective tool to discriminate cancer cells according their genetic characteristics. Using unsupervised clustering, our study suggest that the heterogeneity of Ca^{2+} responses may be the result of the declination of 3 main Ca^{2+} responses. Further investigations will be required to extend this single cell Ca^{2+} profiling to additional cancers, to assess the discrimination of different cell types present in tumors *in vivo* and to improve our understanding of the definition of Ca^{2+} signatures specific of each cell types by identifying molecular actors involved.

ACKNOWLEDGEMENTS:

This work was supported by the Le Studium, Loire Valley Institute for Advanced Studies, Orléans & Tours, France (DC), Région Centre-Val de Loire APR-IR grant (CC, AB and KM), Inserm PCSI (CV) and ARFMAD association (MG). We acknowledge all lab members of Inserm U1069 and LIFAT for their help. DC designed the project. CC, AB, DB performed cell culture and experiments. TB and DC performed machine learning experiments. DC,TB, MG, CV and KM wrote the manuscript. Funding acquisition: DC, CV, KM, MG.

MATERIALS AND METHODS:

Cell culture

The CRC cell lines (HT29, DLD1, HCT116, Lovo, SW48, SW480, and SW620) were cultured in McCoy media + GlutaMax (Gibco, Thermo Fisher, Illkirch, France) supplemented with 10% fetal bovine serum (FBS; Eurobio, Les Ulis, France) and 1% penicillin and streptomycin (Eurobio). NCM356 cells (Incell Corporation, LLC, San Antonio, TX, USA) were cultured in high-glucose Dulbecco's Modifed Eagle Medium (DMEM) (Sigma-Aldrich, Missouri, USA) supplemented with 10% FBS and antibiotics. NCM356 cells (Incell Corporation, LLC, San Antonio, TX, USA) were cultured in high-glucose Dulbecco's Modifed Eagle Medium (DMEM) (Sigma-Aldrich, Missouri, USA) supplemented with 10% FBS and antibiotics. All cells were cultured in an incubator at 37°C with 5% CO_2. Prostate cancer cell lines 22RV1, DU-145, C4-2 and C4-2b were cultured in RPMI-1640 media + GlutaMax (Gibco, Thermo Fisher, Illkirch, France) supplemented with 10% fetal bovine serum (FBS; Eurobio, Les Ulis, France) and 1% penicillin and streptomycin (Eurobio). PC3 were cultured in RPMI-1640 media + GlutaMax (Gibco, Thermo Fisher, Illkirch, France) supplemented with 5% fetal bovine serum (FBS; Eurobio, Les Ulis, France) and 1% penicillin and streptomycin (Eurobio). VCap cell line was cultured in DMEM supplemented with 2 mM Glutamine (Eurobio), 10% fetal bovine serum (FBS; Eurobio, Les Ulis, France) and 1% penicillin and streptomycin (Eurobio).

SOCE measurement by Fura-10 AM

Intracellular Ca^{2+} imaging was performed as described previously (33). Cells were plated in 96-well plates at 1×10^5 cells per well 24 h before the experiment. Adherent cells were for loaded with the ratiometric dye Fura10-acetoxymethyl ester (AM; 4.5 µM) at 37°C for 45 min and then washed with PBS supplemented with Ca^{2+}. During the experiment, the cells were incubated with Ca^{2+}-free physiological saline solution (PSS) solution and treated with 2 µM Thapsigargin (TG) (Reference: T7458, Life Technologies, Thermo Fisher) to deplete intracellular store of Ca^{2+}. Ca^{2+} entry was stimulated by injecting 2 mM of $CaCl_2$. Fluorescence emission was measured at 510 nm using the FlexStation-3 (Molecular Devices) with excitation at 354 and 415 nm.

Single cell calcium Imaging:

Single cell calcium imaging was performed as described previously (34). Briefly, 24 h before the experiment, 5×10^4 cells were plated on 35-mm disches. On the day of the experiment, cells were loaded with Cal520-AM (4.5 µM) at 37°C for 30 min in physiological saline solution (PSS) containing: 150 mM NaCl, 4 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM Hepes, and 10 mM D-glucose. Cells were then washed twice in PSS and incubate in PSS solution for 3 min before acquisition. For stimulation of cells, an equal volume of a 2× concentrated compound of interest

(prepared in PSS) was added to the dish after the first minute of acquisition. Fluorescence was acquired using a Nikon TiEclipse inverted scope equipped a DS-Qi2 camera (Nikon), a combination of dichroic cube, excitation and emission filters at 488 and 515 nm (SemRock), and a SOLA SE-5-LCR-SA Lamp (Photometrics). Images were acquired every 3 s for 15 min. Fluorescence intensity of individual cells was obtained by defining a region-of-interest for each individual cell, substracting the background fluorescence and normalizing to the fluorescent intensity measured in the first image (F/F_0). All image processing was performed using ImageJ Software (NIH).

Processing of calcium profiling:
All profiles of calcium measurement were processed using R and RStudio. For each single cells, amplitude and AUC were defined after the first minute of acquisition. Cells are considered as positively responsive for the tested compounds if the fluorescence is superior to 110% of the initial fluorescence (F_0). Oscillations were determined using *find_peaks* functions from *GCalcium* R package (n.points = 3). Only oscillations with a fluorescence above 110 % of the initial fluorescence (F_0) are counted.

Supervised Machine learning and Neural Networks:
Artificial Neural network (ANN) was built using *tensorflow, keras, scikit-learn* and *numpy* Python packages. ANN models contains three layers of neurons: 1) *input* layer containing normalized and scaled fluorescence intensity recorded at each timepoints, 2) *C1 "hidden"* layer of half-sized the input layer and 3) *output* layer composed of different possibles labels. Unless otherwise stated, default training:validation ratio of 0.8:0.2 was applied on the dataset. Mean squared error was used as a loss function. Learning rate and batch size were set at 9e−6 and 32 respectively. ANN were trained for 500 epochs and for each models, 5 different runs were performed.

Dimensionality reduction (UMAP):
Profiles of Ca^{2+} responses are normalized and subjected to Uniform Manifold Approximation and Projection (UMAP) using *umap* R package. n_neighbors and min_dist parameters were modulated to assay for the best visualization. Unless otherwise is stated, min_dist = 0,1 and n_neighbors = 15 were used for UMAP representation.

Unsupervised machine learning:
Each signal acquisition was normalized and scaled using *scikit-learn* python library. Unsupervised clustering algorithms such as Affinity Propagation, DBScan and Optics were applied using *scikit-learn* and *scipy* Python libraries. Affinity Propagation is used with following parameters: damping=0.95, maximal iteration=500, convergence iteration=20. Shared-Nearest Neighbor (SNN) was performed using *bluster* and Louvain clustering was applied using *igraph* R packages.

REFERENCES:

1. P. A. Brodskiy, J. J. Zartman, Calcium as a signal integrator in developing epithelial tissues. *Phys Biol* **15**, 051001 (2018).

2. M. J. Berridge, M. D. Bootman, H. L. Roderick, Calcium signalling: dynamics, homeostasis and remodelling. *Nat Rev Mol Cell Biol* **4**, 517–529 (2003).

3. M. J. Berridge, P. Lipp, M. D. Bootman, The versatility and universality of calcium signalling. *Nat Rev Mol Cell Biol* **1**, 11–21 (2000).

4. E. Smedler, P. Uhlén, Frequency decoding of calcium oscillations. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1840**, 964–969 (2014).

5. M. J. Boulware, J. S. Marchant, Timing in cellular Ca2+ signaling. *Curr Biol* **18**, R769–R776 (2008).

6. R. C. Evans, K. T. Blackwell, Calcium: Amplitude, Duration, or Location? *The Biological Bulletin* **228**, 75–83 (2015).

7. O. Gryshchenko, J. V. Gerasimenko, S. Peng, O. V. Gerasimenko, O. H. Petersen, Calcium signalling in the acinar environment of the exocrine pancreas: physiology and pathophysiology. *J. Physiol. (Lond.)* **596**, 2663–2678 (2018).

8. M. J. Berridge, Calcium oscillations. *Journal of Biological Chemistry* **265**, 9583–9586 (1990).

9. D. P. Noren, *et al.*, Endothelial cells decode VEGF-mediated Ca^{2+} signaling patterns to produce distinct functional responses. *Sci. Signal.* **9** (2016).

10. Z. A. Mohammed, C. Doran, D. Grundy, M. A. Nassar, Veratridine produces distinct calcium response profiles in mouse Dorsal Root Ganglia neurons. *Sci Rep* **7**, 45221 (2017).

11. L. R. James, *et al.*, High-throughput analysis of calcium signalling kinetics in astrocytes stimulated with different neurotransmitters. *PLoS One* **6**, e26889 (2011).

12. Y. Yokota, *et al.*, Endothelial Ca 2+ oscillations reflect VEGFR signaling-regulated angiogenic capacity in vivo. *Elife* **4**, e08817 (2015).

13. V. Scheuss, R. Yasuda, A. Sobczyk, K. Svoboda, Nonlinear [Ca2+] signaling in dendrites and spines caused by activity-dependent depression of Ca2+ extrusion. *J Neurosci* **26**, 8183–8194 (2006).

14. R. E. Dolmetsch, K. Xu, R. S. Lewis, Calcium oscillations increase the efficiency and specificity of gene expression. **392** (1998).

15. A. B. Parekh, Decoding cytosolic Ca2+ oscillations. *Trends in Biochemical Sciences* **36**, 78–87 (2011).

16. R. E. Yoast, *et al.*, The native ORAI channel trio underlies the diversity of Ca2+ signaling events. *Nat Commun* **11**, 2444 (2020).

17. G. Gambardella, *et al.*, A single-cell analysis of breast cancer cell lines to study tumour heterogeneity and drug response. *Nat Commun* **13**, 1714 (2022).

18. R. A. Burrell, N. McGranahan, J. Bartek, C. Swanton, The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–345 (2013).

19. Asadullah, *et al.*, Combined heterogeneity in cell size and deformability promotes cancer invasiveness. *Journal of Cell Science* **134**, jcs250225 (2021).

20. N. Prevarskaya, R. Skryma, Y. Shuba, Ion channels and the hallmarks of cancer. *Trends in Molecular Medicine* **16**, 107–121 (2010).

21. S. Marchi, C. Giorgi, L. Galluzzi, P. Pinton, Ca2+ Fluxes and Cancer. *Mol Cell* **78**, 1055–1069 (2020).

22. Y.-F. Chen, P.-C. Lin, Y.-M. Yeh, L.-H. Chen, M.-R. Shen, Store-Operated Ca2+ Entry in Tumor Progression: From Molecular Mechanisms to Clinical Implications. *Cancers (Basel)* **11**, 899 (2019).

23. E. Terrié, V. Coronas, B. Constantin, Role of the calcium toolkit in cancer stem cells. *Cell Calcium* **80**, 141–151 (2019).

24. M. Guéguinou, *et al.*, KCa and Ca2+ channels: The complex thought. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 2322–2333 (2014).

25. O. Romito, *et al.*, Calcium signaling: A therapeutic target to overcome resistance to therapies in cancer. *Cell Calcium* **108**, 102673 (2022).

26. J. I. E. Bruce, A. D. James, Targeting the Calcium Signalling Machinery in Cancer. *Cancers (Basel)* **12**, 2351 (2020).

27. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction (2020) (September 25, 2023).

28. K. Pijackova, *et al.*, Genetic algorithm designed for optimization of neural network architectures for intracranial EEG recordings analysis. *J Neural Eng* **20** (2023).

29. A. Warren, *et al.*, Global computational alignment of tumor and cell line transcriptional profiles. *Nat Commun* **12**, 22 (2021).

30. S. Figiel, *et al.*, Functional Organotypic Cultures of Prostate Tissues: A Relevant Preclinical Model that Preserves Hypoxia Sensitivity and Calcium Signaling. *Am. J. Pathol.* **189**, 1268–1275 (2019).

31. M. C. Haffner, *et al.*, Genomic and phenotypic heterogeneity in prostate cancer. *Nat Rev Urol* **18**, 79–92 (2021).

32. S. Swain, *et al.*, Confocal Imaging and *k* -Means Clustering of $GABA_B$ and mGluR Mediated Modulation of Ca^{2+} Spiking in Hippocampal Neurons. *ACS Chem. Neurosci.* **9**, 3094–3107 (2018).

33. M. Guéguinou, *et al.*, Curcumin and NCLX inhibitors share anti-tumoral mechanisms in microsatellite-instability-driven colorectal cancer. *Cell Mol Life Sci* **79**, 284 (2022).

34. D. Crottès, *et al.*, TMEM16A controls EGF-induced calcium signaling implicated in pancreatic cancer prognosis. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13026–13035 (2019).

FIGURE 1.

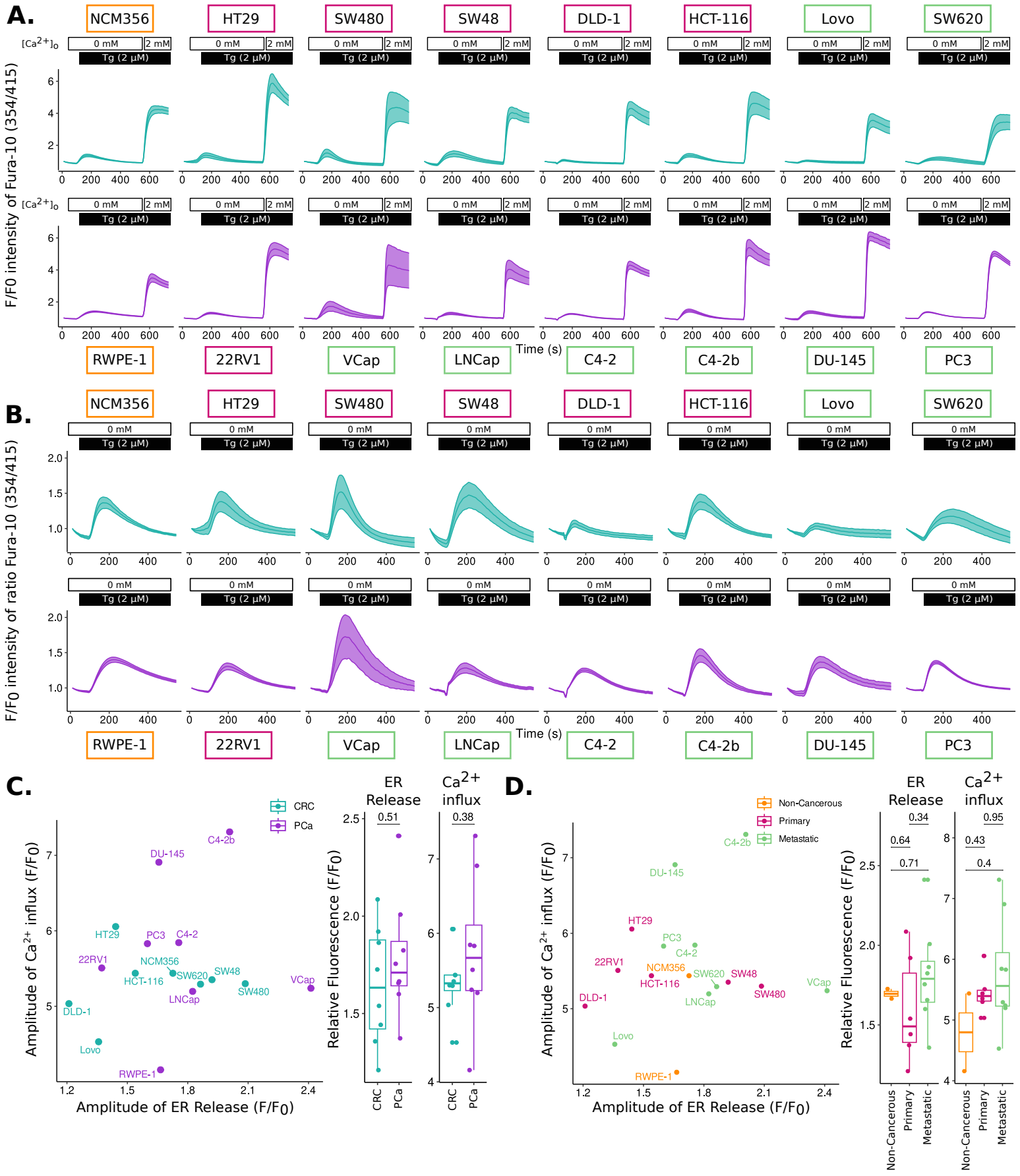

A.
NCM356
HT29
SW480
SW48
DLD-1
HCT-116
Lovo
SW620
[Ca2+]o
0 mM
2 mM
Tg (2 μM)
F/F0 intensity of Fura-10 (354/415)
Time (s)
RWPE-1
22RV1
VCap
LNCap
C4-2
C4-2b
DU-145
PC3
B.
F/F0 intensity of ratio Fura-10 (354/415)
C.
CRC
PCa
Amplitude of Ca2+ influx (F/F0)
Amplitude of ER Release (F/F0)
ER Release
Ca2+ influx
Relative Fluorescence (F/F0)
0.51
0.38
D.
Non-Cancerous
Primary
Metastatic
0.34
0.64
0.71
0.95
0.43
0.4

FIGURE 2.

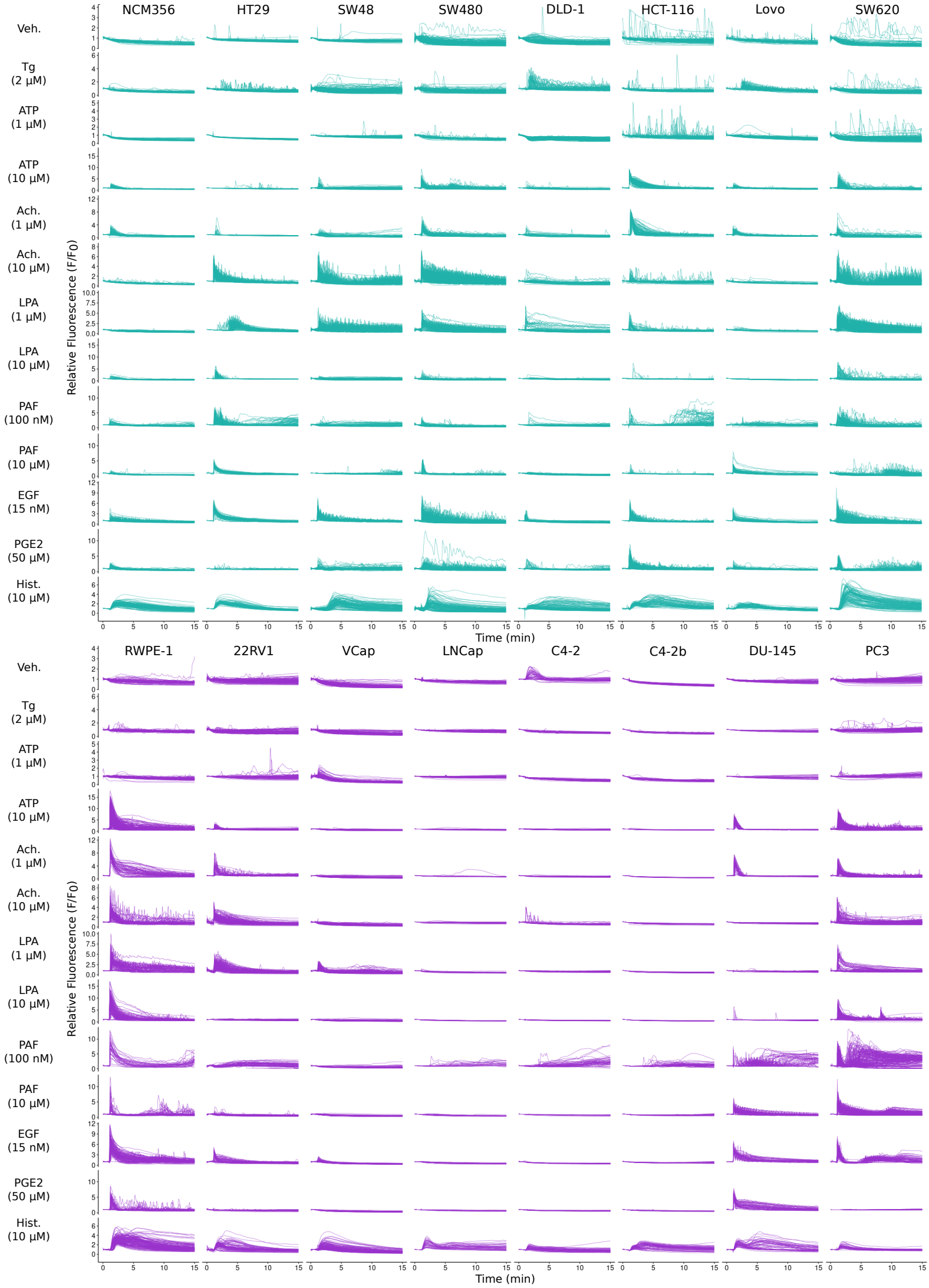

NCM356
HT29
SW48
SW480
DLD-1
HCT-116
Lovo
SW620
Veh.
Tg (2 μM)
ATP (1 μM)
ATP (10 μM)
Ach. (1 μM)
Ach. (10 μM)
LPA (1 μM)
LPA (10 μM)
PAF (100 nM)
PAF (10 μM)
EGF (15 nM)
PGE2 (50 μM)
Hist. (10 μM)
Relative Fluorescence (F/F0)
Time (min)
RWPE-1
22RV1
VCap
LNCap
C4-2
C4-2b
DU-145
PC3
Veh.
Tg (2 μM)
ATP (1 μM)
ATP (10 μM)
Ach. (1 μM)
Ach. (10 μM)
LPA (1 μM)
LPA (10 μM)
PAF (100 nM)
PAF (10 μM)
EGF (15 nM)
PGE2 (50 μM)
Hist. (10 μM)
Relative Fluorescence (F/F0)
Time (min)

FIGURE 3.

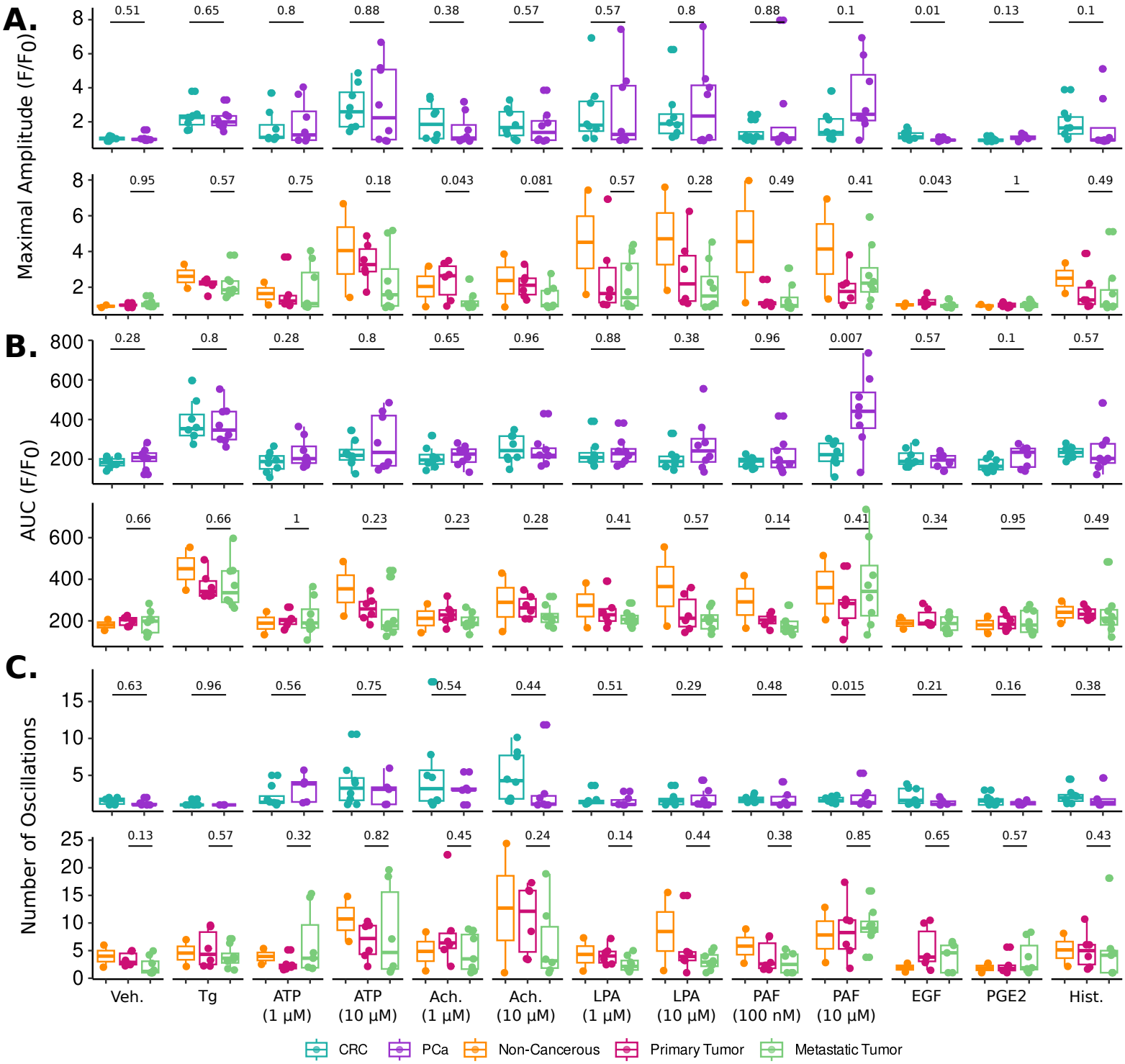
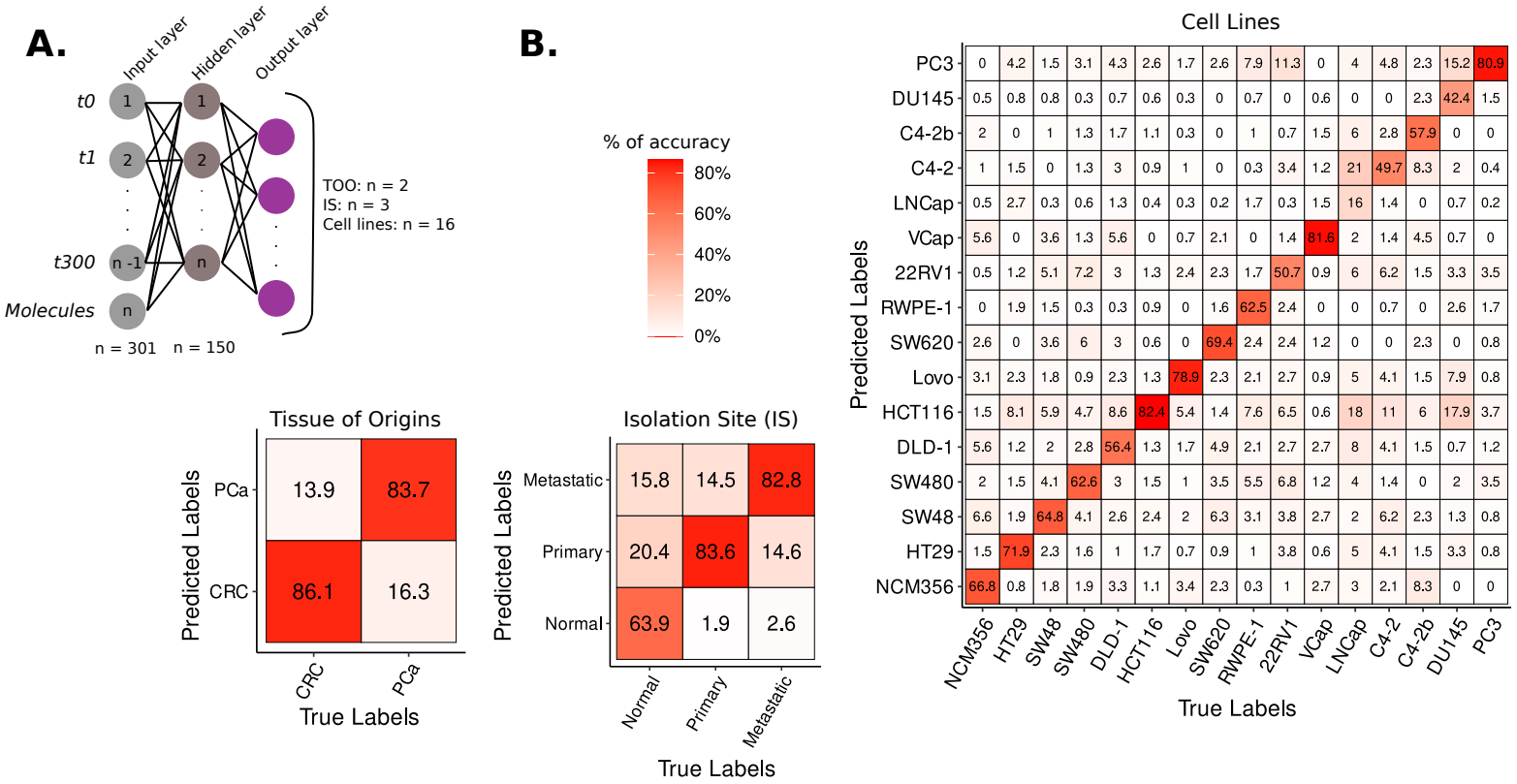
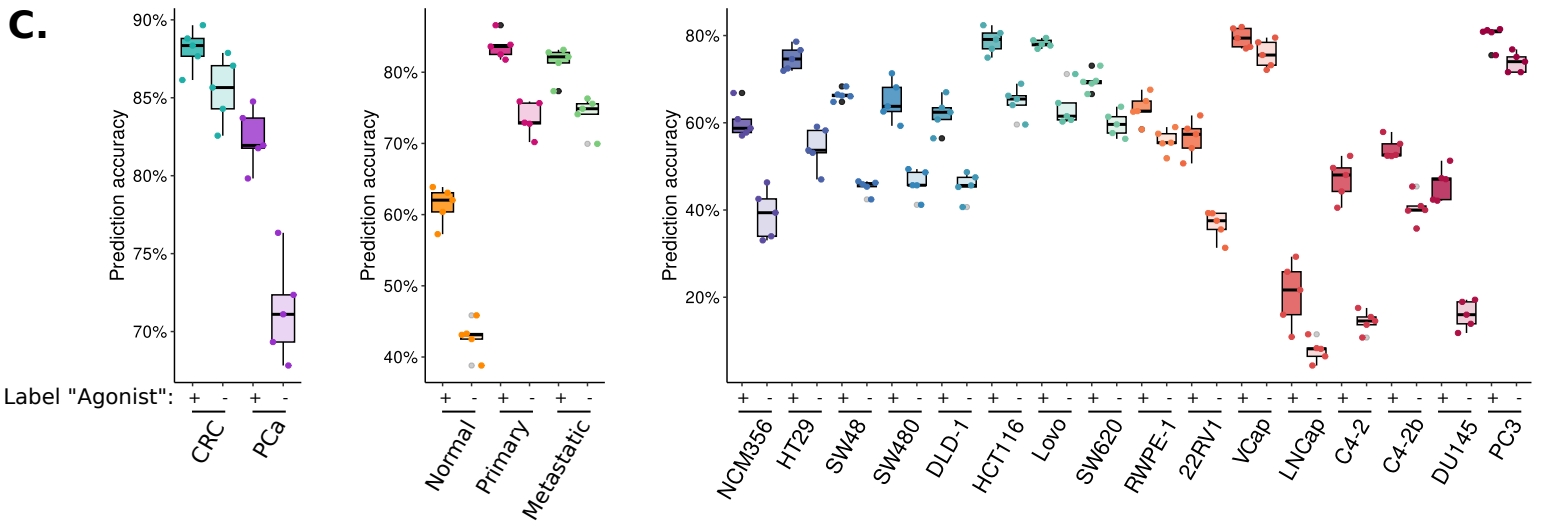

A.
Maximal Amplitude (F/F0)
0.51 0.65 0.8 0.88 0.38 0.57 0.57 0.8 0.88 0.1 0.01 0.13 0.1
0.95 0.57 0.75 0.18 0.043 0.081 0.57 0.28 0.49 0.41 0.043 1 0.49
B.
AUC (F/F0)
0.28 0.8 0.28 0.8 0.65 0.96 0.88 0.38 0.96 0.007 0.57 0.1 0.57
0.66 0.66 1 0.23 0.23 0.28 0.41 0.57 0.14 0.41 0.34 0.95 0.49
C.
Number of Oscillations
0.63 0.96 0.56 0.75 0.54 0.44 0.51 0.29 0.48 0.015 0.21 0.16 0.38
0.13 0.57 0.32 0.82 0.45 0.24 0.14 0.44 0.38 0.85 0.65 0.57 0.43
Veh.
Tg
ATP (1 μM)
ATP (10 μM)
Ach. (1 μM)
Ach. (10 μM)
LPA (1 μM)
LPA (10 μM)
PAF (100 nM)
PAF (10 μM)
EGF
PGE2
Hist.
CRC
PCa
Non-Cancerous
Primary Tumor
Metastatic Tumor

FIGURE 4.


A.
Input layer
Hidden layer
Output layer
t0
t1
t300
Molecules
n = 301
n = 150
TOO: n = 2
IS: n = 3
Cell lines: n = 16
B.
% of accuracy
80%
60%
40%
20%
0%
Tissue of Origins
Predicted Labels
True Labels
PCa
CRC
13.9
83.7
86.1
16.3
Isolation Site (IS)
Metastatic
Primary
Normal
15.8
14.5
82.8
20.4
83.6
14.6
63.9
1.9
2.6
Cell Lines
PC3 0 4.2 1.5 3.1 4.3 2.6 1.7 2.6 7.9 11.3 0 4 4.8 2.3 15.2 80.9
DU145 0.5 0.8 0.8 0.3 0.7 0.6 0.3 0 0.7 0 0.6 0 0 2.3 42.4 1.5
C4-2b 2 0 1 1.3 1.7 1.1 0.3 0 1 0.7 1.5 6 2.8 57.9 0 0
C4-2 1 1.5 0 1.3 3 0.9 1 0 0.3 3.4 1.2 21 49.7 8.3 2 0.4
LNCap 0.5 2.7 0.3 0.6 1.3 0.4 0.3 0.2 1.7 0.3 1.5 16 1.4 0 0.7 0.2
VCap 5.6 0 3.6 1.3 5.6 0 0.7 2.1 0 1.4 81.6 2 1.4 4.5 0.7 0
22RV1 0.5 1.2 5.1 7.2 3 1.3 2.4 2.3 1.7 50.7 0.9 6 6.2 1.5 3.3 3.5
RWPE-1 0 1.9 1.5 0.3 0.3 0.9 0 1.6 62.5 2.4 0 0 0.7 0 2.6 1.7
SW620 2.6 0 3.6 6 3 0.6 0 69.4 2.4 2.4 1.2 0 0 2.3 0 0.8
Lovo 3.1 2.3 1.8 0.9 2.3 1.3 78.9 2.3 2.1 2.7 0.9 5 4.1 1.5 7.9 0.8
HCT116 1.5 8.1 5.9 4.7 8.6 82.4 5.4 1.4 7.6 6.5 0.6 18 11 6 17.9 3.7
DLD-1 5.6 1.2 2 2.8 56.4 1.3 1.7 4.9 2.1 2.7 2.7 8 4.1 1.5 0.7 1.2
SW480 2 1.5 4.1 62.6 3 1.5 1 3.5 5.5 6.8 1.2 4 1.4 0 2 3.5
SW48 6.6 1.9 64.8 4.1 2.6 2.4 2 6.3 3.1 3.8 2.7 2 6.2 2.3 1.3 0.8
HT29 1.5 71.9 2.3 1.6 1 1.7 0.7 0.9 1 3.8 0.6 5 4.1 1.5 3.3 0.8
NCM356 66.8 0.8 1.8 1.9 3.3 1.1 3.4 2.3 0.3 1 2.7 3 2.1 8.3 0 0
NCM356 HT29 SW48 SW480 DLD-1 HCT116 Lovo SW620 RWPE-1 22RV1 VCap LNCap C4-2 C4-2b DU145 PC3


C.
Prediction accuracy
90%
85%
80%
75%
70%
Label "Agonist":
CRC
PCa
Normal
Primary
Metastatic
NCM356
HT29
SW48
SW480
DLD-1
HCT116
Lovo
SW620
RWPE-1
22RV1
VCap
LNCap
C4-2
C4-2b
DU145
PC3

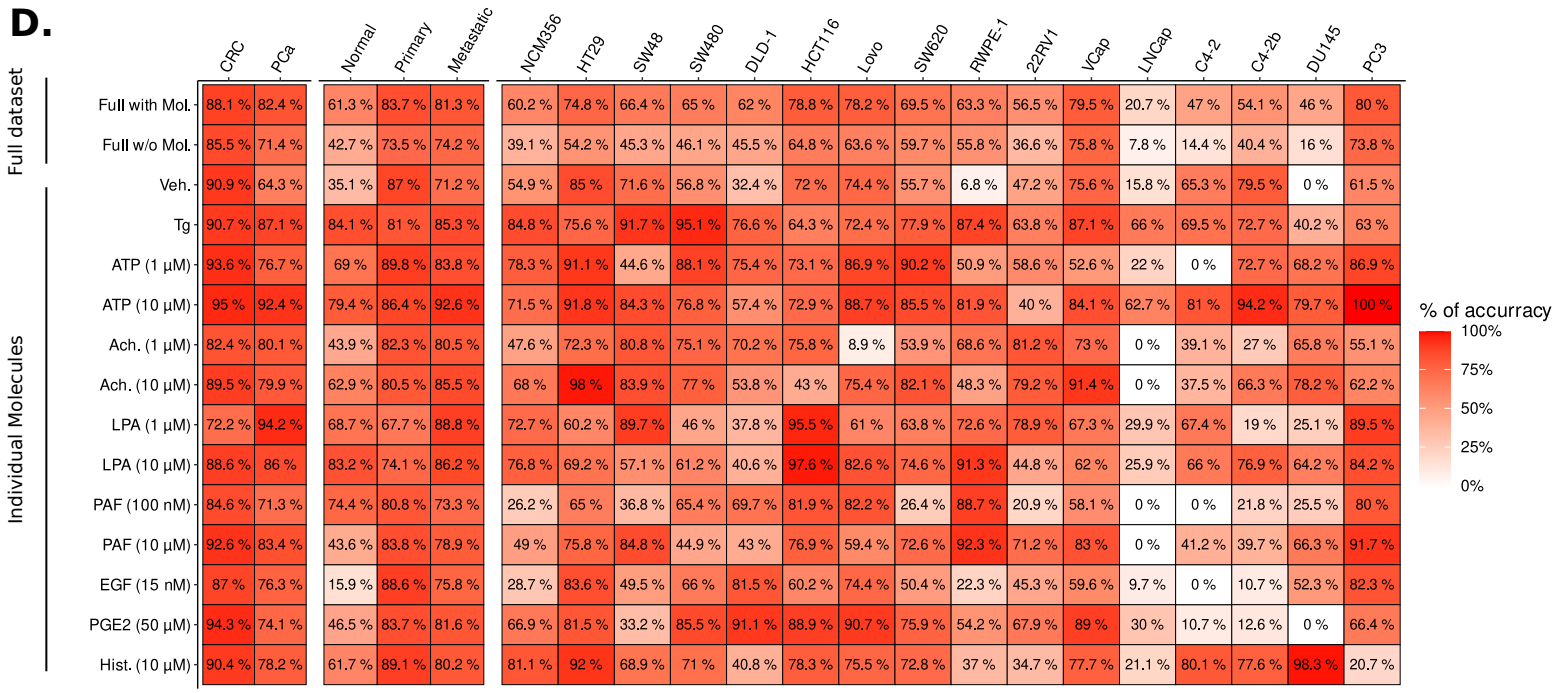

D.
Full dataset
Individual Molecules
CRC PCa Normal Primary Metastatic NCM356 HT29 SW48 SW480 DLD-1 HCT116 Lovo SW620 RWPE-1 22RV1 VCap LNCap C4-2 C4-2b DU145 PC3
Full with Mol. 88.1 % 82.4 % 61.3 % 83.7 % 81.3 % 60.2 % 74.8 % 66.4 % 65 % 62 % 78.8 % 78.2 % 69.5 % 63.3 % 56.5 % 79.5 % 20.7 % 47 % 54.1 % 46 % 80 %
Full w/o Mol. 85.5 % 71.4 % 42.7 % 73.5 % 74.2 % 39.1 % 54.2 % 45.3 % 46.1 % 45.5 % 64.8 % 63.6 % 59.7 % 55.8 % 36.6 % 75.8 % 7.8 % 14.4 % 40.4 % 16 % 73.8 %
Veh. 90.9 % 64.3 % 35.1 % 87 % 71.2 % 54.9 % 85 % 71.6 % 56.8 % 32.4 % 72 % 74.4 % 55.7 % 6.8 % 47.2 % 75.6 % 15.8 % 65.3 % 79.5 % 0 % 61.5 %
Tg 90.7 % 87.1 % 84.1 % 81 % 85.3 % 84.8 % 75.6 % 91.7 % 95.1 % 76.6 % 64.3 % 72.4 % 77.9 % 87.4 % 63.8 % 87.1 % 66 % 69.5 % 72.7 % 40.2 % 63 %
ATP (1 µM) 93.6 % 76.7 % 69 % 89.8 % 83.8 % 78.3 % 91.1 % 44.6 % 88.1 % 75.4 % 73.1 % 86.9 % 90.2 % 50.9 % 58.6 % 52.6 % 22 % 0 % 72.7 % 68.2 % 86.9 %
ATP (10 µM) 95 % 92.4 % 79.4 % 86.4 % 92.6 % 71.5 % 91.8 % 84.3 % 76.8 % 57.4 % 72.9 % 88.7 % 85.5 % 81.9 % 40 % 84.1 % 62.7 % 81 % 94.2 % 79.7 % 100 %
Ach. (1 µM) 82.4 % 80.1 % 43.9 % 82.3 % 80.5 % 47.6 % 72.3 % 80.8 % 75.1 % 70.2 % 75.8 % 8.9 % 53.9 % 68.6 % 81.2 % 73 % 0 % 39.1 % 27 % 65.8 % 55.1 %
Ach. (10 µM) 89.5 % 79.9 % 62.9 % 80.5 % 85.5 % 68 % 98 % 83.9 % 77 % 53.8 % 43 % 75.4 % 82.1 % 48.3 % 79.2 % 91.4 % 0 % 37.5 % 66.3 % 78.2 % 62.2 %
LPA (1 µM) 72.2 % 94.2 % 68.7 % 67.7 % 88.8 % 72.7 % 60.2 % 89.7 % 46 % 37.8 % 95.5 % 61 % 63.8 % 72.6 % 78.9 % 67.3 % 29.9 % 67.4 % 19 % 25.1 % 89.5 %
LPA (10 µM) 88.6 % 86 % 83.2 % 74.1 % 86.2 % 76.8 % 69.2 % 57.1 % 61.2 % 40.6 % 97.6 % 82.6 % 74.6 % 91.3 % 44.8 % 62 % 25.9 % 66 % 76.9 % 64.2 % 84.2 %
PAF (100 nM) 84.6 % 71.3 % 74.4 % 80.8 % 73.3 % 26.2 % 65 % 36.8 % 65.4 % 69.7 % 81.9 % 82.2 % 26.4 % 88.7 % 20.9 % 58.1 % 0 % 0 % 21.8 % 25.5 % 80 %
PAF (10 µM) 92.6 % 83.4 % 43.6 % 83.8 % 78.9 % 49 % 75.8 % 84.8 % 44.9 % 43 % 76.9 % 59.4 % 72.6 % 92.3 % 71.2 % 83 % 0 % 41.2 % 39.7 % 66.3 % 91.7 %
EGF (15 nM) 87 % 76.3 % 15.9 % 88.6 % 75.8 % 28.7 % 83.6 % 49.5 % 66 % 81.5 % 60.2 % 74.4 % 50.4 % 22.3 % 45.3 % 59.6 % 9.7 % 0 % 10.7 % 52.3 % 82.3 %
PGE2 (50 µM) 94.3 % 74.1 % 46.5 % 83.7 % 81.6 % 66.9 % 81.5 % 33.2 % 85.5 % 91.1 % 88.9 % 90.7 % 75.9 % 54.2 % 67.9 % 89 % 30 % 10.7 % 12.6 % 0 % 66.4 %
Hist. (10 µM) 90.4 % 78.2 % 61.7 % 89.1 % 80.2 % 81.1 % 92 % 68.9 % 71 % 40.8 % 78.3 % 75.5 % 72.8 % 37 % 34.7 % 77.7 % 21.1 % 80.1 % 77.6 % 98.3 % 20.7 %
% of accurracy
100%
75%
50%
25%
0%

FIGURE 5.

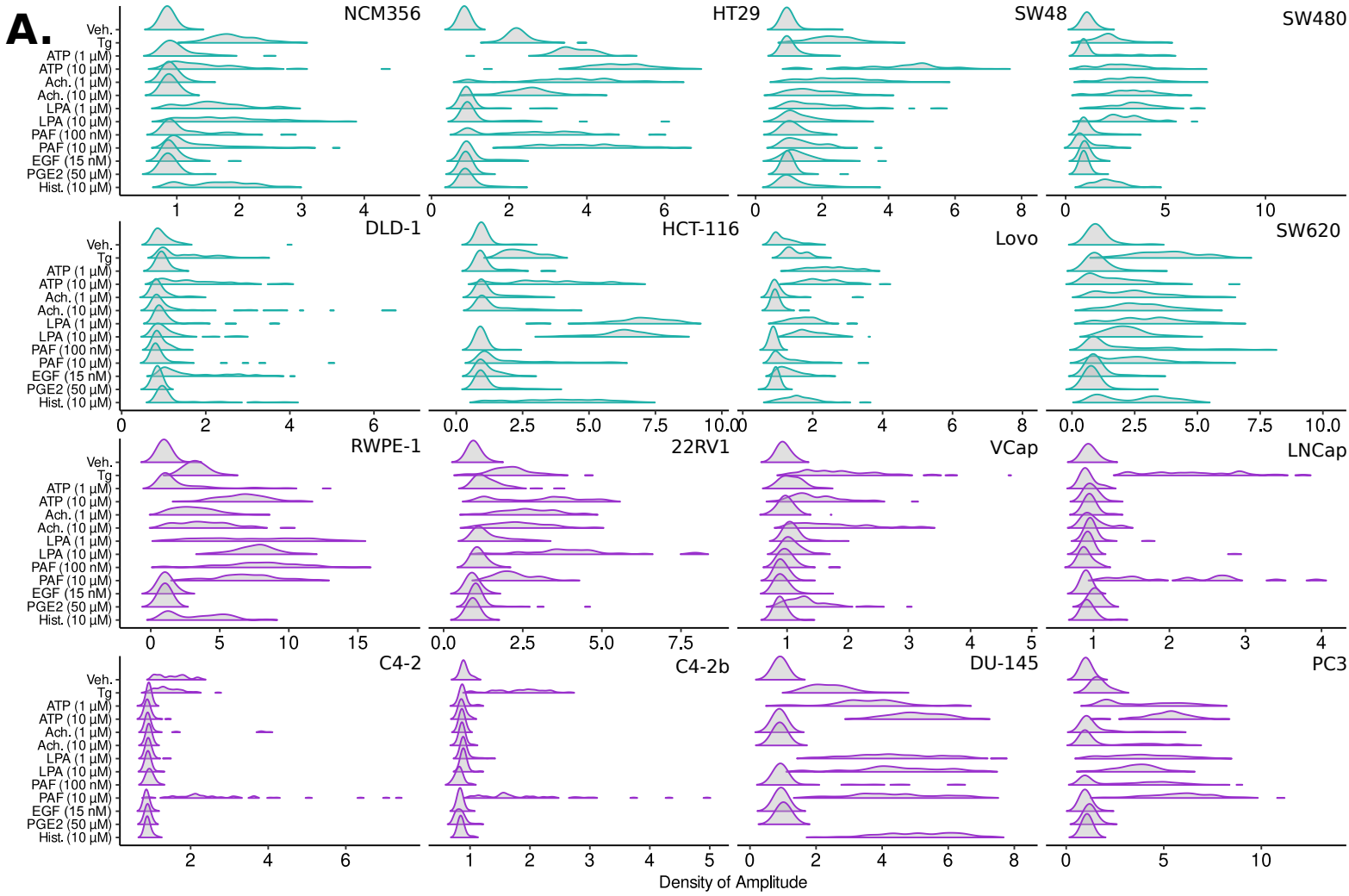

A.
NCM356
HT29
SW48
SW480
DLD-1
HCT-116
Lovo
SW620
RWPE-1
22RV1
VCap
LNCap
C4-2
C4-2b
DU-145
PC3
Veh.
Tg
ATP (1 μM)
ATP (10 μM)
Ach. (1 μM)
Ach. (10 μM)
LPA (1 μM)
LPA (10 μM)
PAF (100 nM)
PAF (10 μM)
EGF (15 nM)
PGE2 (50 μM)
Hist. (10 μM)
Density of Amplitude

FIGURE 6.

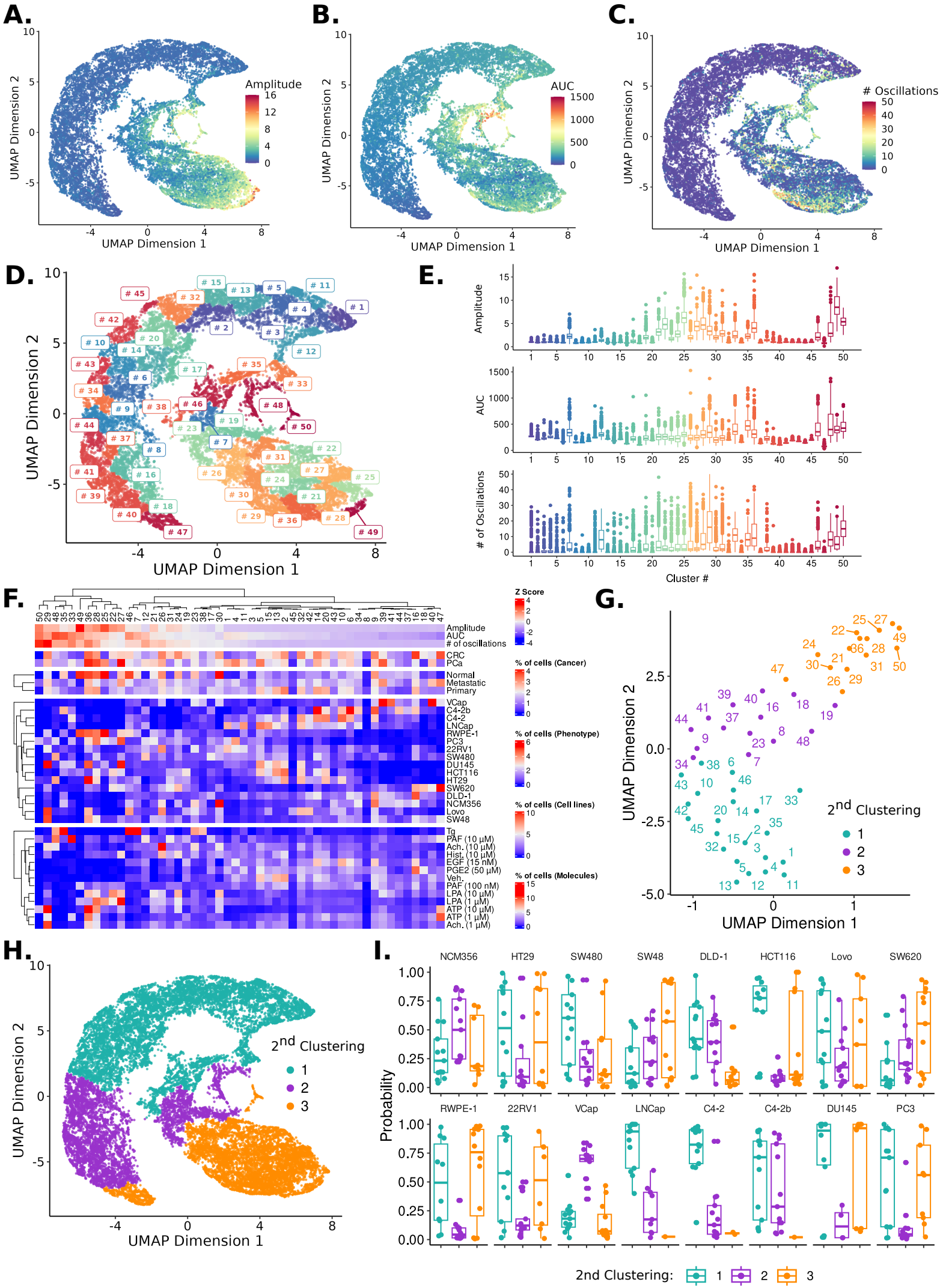

A.
B.
C.
UMAP Dimension 1
UMAP Dimension 2
Amplitude
AUC
Oscillations
D.
E.
Amplitude
AUC
# of Oscillations
Cluster #
F.
Z Score
Amplitude
AUC
# of oscillations
CRC
PCa
Normal
Metastatic
Primary
VCap
C4-2b
C4-2
LNCap
RWPE-1
PC3
22RV1
SW480
DU145
HCT116
HT29
SW620
DLD-1
NCM356
Lovo
SW48
Tg
PAF (10 μM)
Ach. (10 μM)
Hist. (10 μM)
EGF (15 nM)
PGE2 (50 μM)
Veh.
PAF (100 nM)
LPA (10 μM)
LPA (1 μM)
ATP (10 μM)
ATP (1 μM)
Ach. (1 μM)
% of cells (Cancer)
% of cells (Phenotype)
% of cells (Cell lines)
% of cells (Molecules)
G.
2nd Clustering
H.
I.
NCM356
HT29
SW480
SW48
DLD-1
HCT116
Lovo
SW620
RWPE-1
22RV1
VCap
LNCap
C4-2
C4-2b
DU145
PC3
Probability
2nd Clustering:

FIGURE LEGENDS:

Figure 1 : Store-operated Ca^{2+} entry induced by Thapsigargin in a panel of cell lines.
A. Ratiometric fluorescent measures of the Thapsigargin (Tg)-induced SOCE in a panel of CRC and PCa cell lines (Mean ± SEM, n = 4 to 12).
B. Ratiometric fluorescent measures of the Tg-induced ER released in a panel of CRC and PCa cell lines (Mean ± SEM, n = 4 to 12).
C. Characteristics of the SOCE in function of the cancer type. Left panel: Scatter plot representing the amplitude of the Tg-induced Ca^{2+} influx in function of the Tg-induced ER-release in each cell line. Right panel: boxplots representing the amplitude of the Tg-induced ER-release and Ca^{2+} influx per cell lines in function of their original tissue (Colon or Prostate).
D. Characteristics of the SOCE in function of the cancer cell type. Left panel: Scatter plot representing the amplitude of the Tg-induced Ca^{2+} influx in function of the Tg-induced ER-release in each cell line. Right panel: Boxplots representing the amplitude of the Tg-induced ER-release and Ca^{2+} influx per cell lines in function of their original cancer phenotype (Non-cancerous, Primary tumors or Metastasis).

Figure 2 : Single cell Ca^{2+} imaging of colon and prostate cell lines.
Single cell Ca^{2+} responses of a panel of CRC (in green) and PCa cell lines (in purple) has been recorded during 15 minutes. After one minute, differents agonists were added extracellularly. (ATP: Adenosine Tri-phosphate, Ach: AcetylCholine, EGF: Epidermal Growth Factor, LPA: Lysophosphatidic Acid, PAF: Platelet Aggregating Factor, PGE2: Prostaglandin E2, Tg: Thapsigaring, Veh.: Vehicle)

Figure 3 : Characteristics of single cell Ca^{2+} responses induced by agonists in function of the TOO or the IS. (A. to C.) Boxplots comparing the average of the amplitude (**A.**), the AUC (**B.**) and the number of oscillations (**C.**) of Ca^{2+} responses elicited by a panel of molecules between CRC and PCa cell lines (*Upper panel*) or between cell lines originated from non-cancerous tissues, primary or metastatic tumors (*Lower panel*). (Veh. : Vehicle, Tg : Thapsigargin, ATP : Adenosine tri-phosphate, Ach : Acetyl-choline, LPA : Lysophosphatidic acid, PAF : Platelet-Aggregating Factor, EGF : Epidermal Growth Factor, PGE2 : Prostaglandin E2, Hist : Histamine).

Figure 4 : Supervised clustering of single cell Ca^{2+} responses.
A. Single cell Ca^{2+} profiles were submitted to supervised machine learning. An artificial neural network (ANN) of 3 layers was generated. The first layer is composed of 301 nodes and take as input the entire vector of fluorescence intensity + the name of the molecule tested (one value per nodes). The second layer is composed of the half number of nodes of the first layer. Third layer is the output of the ANN and is composed of the number of nodes corresponding to the number of output desired (2 for tissue origins, 3 for cancer phenotype, 16 for lineage).
B. Representative confusion matrix of the prediction accuracy (%) obtained after training the ANN with 80% of the dataset and testing the model with 20% of the dataset.
C. Boxplot representing the prediction accuracy obtained for different output using the full dataset including or not the label of molecules used (left panel: tissue origin, middle panel: cancer phenotype, right panel: lineage).
D. Heatmap representing the prediction accuracy of different output (tissue origins, cancer phenotype, lineage) obtained by ANN built using as input the full dataset with or without the labeling of molecules or datasets corresponding of calcium responses to individual molecules.

Figure 5: Heterogeneity of single cell Ca^{2+} responses.
A. Density of the amplitude, the AUC and number of oscillations of Ca^{2+} responses measured in in each cell lines stimulated by molecules.

Figure 6: Unsupervised classification of single cell Ca^{2+} responses.

A. Visualization of Ca^{2+} responses measured in individual cells in two-dimensional space using UMAP dimensionality reduction and colored for each clusters defined by SNN.

B. Boxplot representing the amplitude, the AUC and the number of oscillations of each clusters defined by SNN.

(C-E). Visualization in two-dimensional space of the amplitude (**C.**), the AUC (**D.**) and the number of oscillations of Ca^{2+} responses (**E.**)

F. Heatmap representing the distribution of cells in individual clusters defined by SNN according their characteristics (tissue origins,. Clusters are sorted according the average maximal amplitude, AUC and number of oscillations.

G. Visualization in two-dimensional space of Ca^{2+} responses averaged for each clusters defined by SNN in A. and subjected to SNN classification.

H. Visualization in two-dimensional space of individual cells colored according the second SNN classification.

I. Probability of the distribution of each cell lines into different second clusters defined SNN.